

# Palaeo-landscape feature identification: a FOSS cloud-based Python approach through Google Earth Engine (GEE)

Guillem Domingo-Ribas <sup>1,\*</sup> and Filippo Brandolini <sup>1,\*</sup>

<sup>1</sup>McCord Centre for Landscape - School of History, Classics and Archaeology, Newcastle University, UK, Newcastle upon Tyne, NE4 5HP, UK; \*

\*Correspondence: g.domingo-ribas2@newcastle.ac.uk; filippo.brandolini@newcastle.ac.uk  
Presented at the ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

## Abstract:

Landscapes are the result of complex long-term interactions between natural and cultural forces, hence requiring techniques that enhance past landscape mapping and sustainable development perspectives. This paper presents a FOSS cloud-based protocol that employs Sentinel-2 imagery in GEE to identify palaeo-landscape features. Developed in Google Colab, the procedure is tested in the Po Plain riverscape (Italy), and it can be easily adapted and replicated for any area of the world.

**Keywords:** multispectral analysis, Sentinel-2, Google Earth Engine, riverscape, FOSS

**FOSS software used and licence:** , Google Earth Engine (*Google Earth Engine Licence Agreement*), Python Programming Language (*PSF Licence Agreement*), QGIS, (GNU Public Licence (GPL) Version), Google Colab (Modified BSD Licence).

**Open dataset and licence:** Copernicus Sentinel-2 satellite data (*Open Access compliant Creative Commons CC BY-SA 3.0 IGO licence*).

## Introduction

Contemporary landscapes materialise from complex long-term interactions between natural and cultural forces. Human populations reorganise the land, adapt its use and its spatial structure to meet their needs. Their use of the land has altered the natural environment at least since the late Pleistocene and, over recent millennia, these processes have accelerated. Therefore, given the threat of human activities as active geomorphological agents that modify the physical landscape, it becomes paramount to develop techniques that enhance past landscape reconstruction while aiding future landscape planning and sustainable development perspectives.

In this line, interdisciplinary GIS and remote sensing techniques have provided significant enhancements in landscape research. These tools have improved the analysis of landscape dynamics over different spatial and temporal scales, enabling the identification of land use change, inferring population dynamics, economic development and, in turn, providing examples of human resilience through past sustainable development practices. Thus, GIS and remote sensing tools also permit the assessment of human impact on natural environments,

hence requiring an approach in which the natural and cultural heritage domains are perceived together (Harrison 2015: 36). At the same time, other scholars highlight that the customizability and the potential for streamlining and automatising feature identification processes offered by remote sensing tools significantly reduce time-consuming manual mapping tasks, notwithstanding the need for analysis on the accuracy of the results (Traviglia and Torsello 2017).

However, considering that archaeological research or commercial archaeological companies do not normally receive a high investment, software companies' licences can hinder the usage of GIS. In this line, remote sensing software licences, powerful working stations with large storage space and certain satellite images can also require higher funding efforts (Parcak 2009). Instead, free and open source software (FOSS) constitute an alternative that does not only imply saving money, but also benefits from a vast active online community as well as installing, modifying and sharing the software when necessary (Ducke 2012). At the same time, in order to move towards a good scientific practice, it is necessary to publish the data, the source code and the software environment that generated the results, which is a growing tendency amongst open source projects and enables other users to reproduce the same process (Donoho et al. 2009).

Therefore, the advent of free-ware cloud computing services such as Google Earth Engine (GEE) represents a step forward towards sustainable and cost-effective landscape monitoring through the processing of more than 40 years of free and open satellite imagery and earth observation tools. Moreover, the possibility of employing GEE through its Python API in Google Colaboratory (commonly known as Colab) allows the use of a Python development

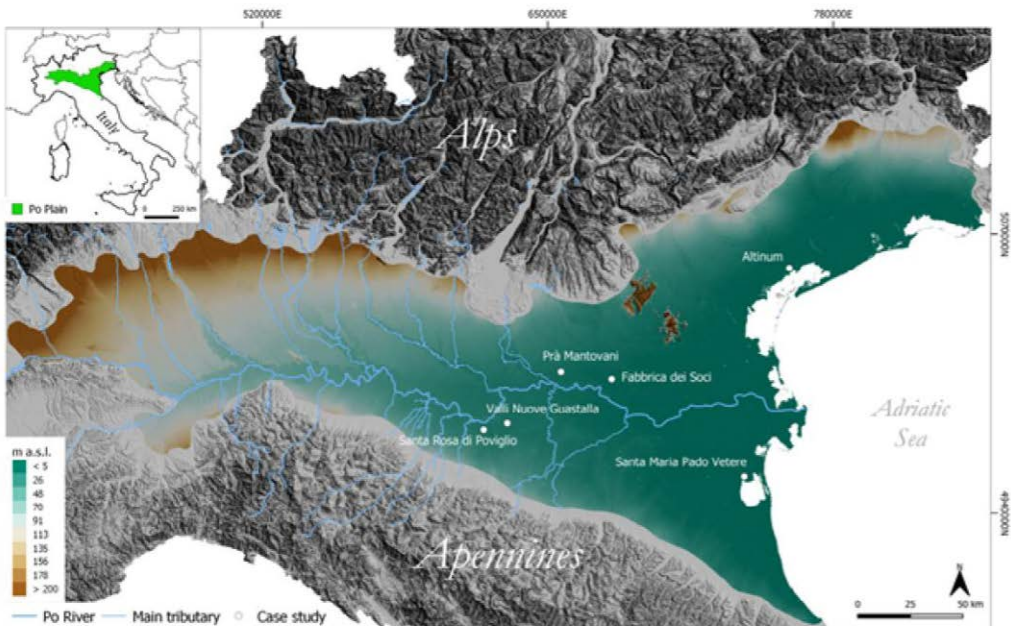


Figure 1. Test case area and locations of the selected case study locations (Brandolini et al. 2021)

environment that runs in the browser using Google Cloud, facilitating the creation of extensions, packages and scripts that broaden the applications of FOSS, particularly for processing more than 40 years of free and open satellite data, whose integration within archaeological projects is expected to continue increasing (Orengo 2015). In this line, while the capabilities of Python in modelling landscape dynamics are widely known, few research revolves around the potential of the GEE Python API (Vos et al. 2019), with a few exceptions that present the potential of the tool and its data catalogue in archaeological research (e.g. Liss, Howland, and Levy 2017; Fattore et al. 2021). Thus, we propose a complete FOSS-cloud approach to identify palaeo-landscape features through the use of GEE Python API in Google Colab. The present paper summarises our presentation at the ArcheoFOSS 2021 Conference, and further information about this methodological approach is available at (Brandolini et al. 2021).

### **Test case area**

In order to test the potential of our FOSS approach, riverine landscapes have been selected for assessing the applicability and effectiveness of the methodology that we propose. Fluvial/alluvial environments have played a key role since prehistory, and archaeological research has shown that human agency has altered deeply the spatial configuration and rate of fluvial processes. Therefore, riverine landscapes are the result of complex relationships between human activities and environmental factors (Fryirs and Brierley 2012). Moreover, the large scale of palaeo-riverscape features has triggered their identification through remote sensing and satellite imagery, hence becoming a solid example for testing this type of approach.

In this line, the Po Plain (Northern Italy) (fig. 1), which is the largest floodplain in Italy, is rich in field and remotely sensed geomorphological data. Thanks to its complex settlement and land-management history, this test case area represents an ideal space in which to test the possibilities of our FOSS-cloud approach to detect riverscapes' palaeo-features. Thus, in this contribution we will introduce the first Python application of Sentinel-2 data for heritage research in a European riverscape, while illustrating the possibility of detecting and interpreting buried anthropogenic landscape features originating in different periods.

### **Materials and methods**

Several applications of remote sensing techniques and GIS to record past landscape settings have been applied in the Po Plain over the last decades. Nevertheless, the current state-of-the-art in FOSS software and the availability of open source satellite data enable the development of protocols that require less computational power and are more user-friendly for non-experienced researchers.

Our approach tests the potentiality of Sentinel-2 (S2) satellite constellation, which was developed by the European Space Agency (ESA) and provides data in 13 separate bands with a spatial resolution up to 10 metres.

In this line, it is worth mentioning that buried features can sometimes be visible through crop marks and soil marks on the surface, given the difference in moisture that they retain in comparison to their surrounding area. However, even with high-resolution satellite sensors,

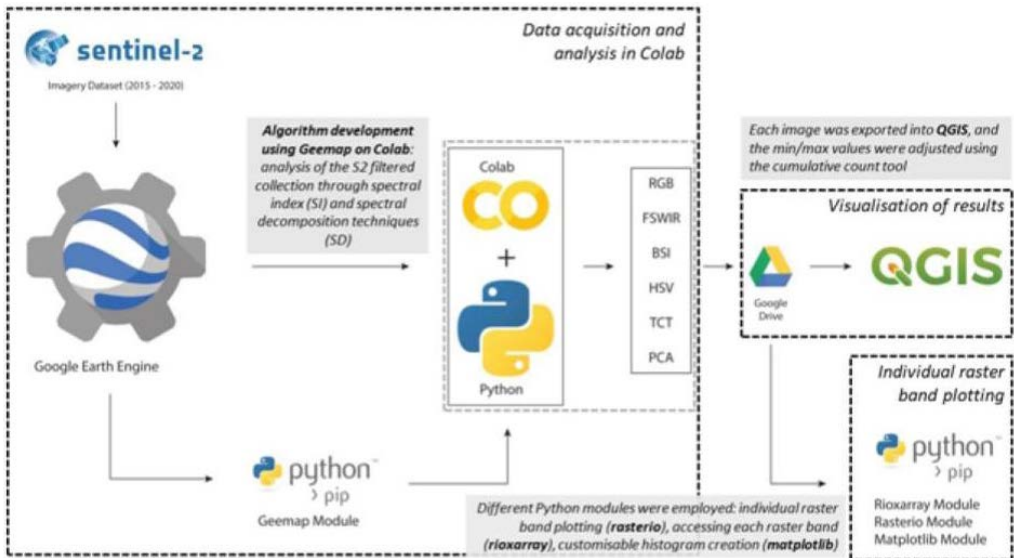


Figure 2. Methodological workflow with the FOSS tools employed. All the steps from data acquisition, processing, analysis and plotting are RGB: Red-Green-Blue colour composite; FSWIR: False colour composite; BSI: Bare Soil Index; HSV: Hue, Saturation and Value; TCT: Tasseled Cap Transformation; PCA: Principal Component Analysis. Based on Brandolini et al. 2021

crop marks are often dependent on various aspects, amongst which the phenological stage of the crops stands out. Therefore, in order to overcome the problematic nature of crop mark identification, a multi-temporal approach has been taken by calculating the mean values of bands in the most significant periods, depending on drought and ploughing episodes, for the identification of crop/soil marks between the years 2015 – 2020.

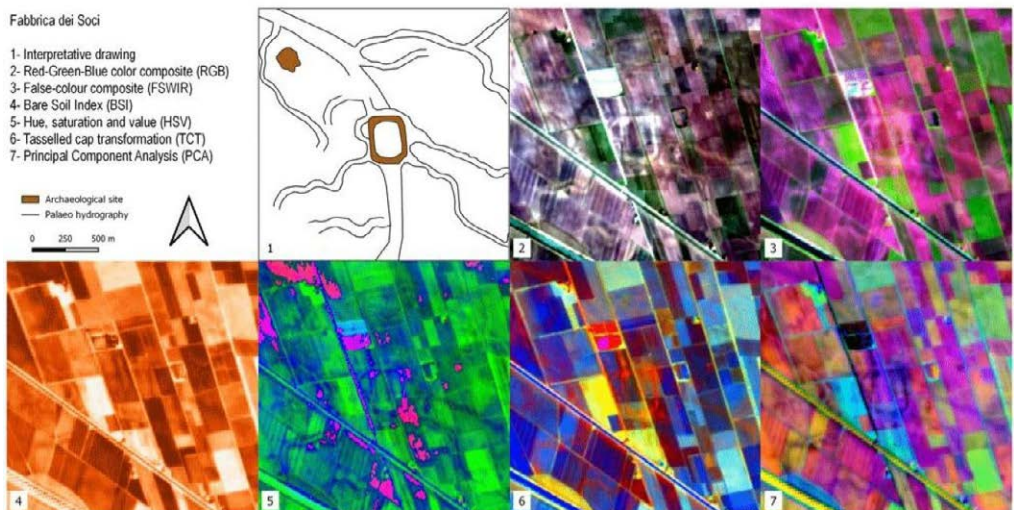


Figure 3. Outputs of the Fabbri dei Soci site (Brandolini et al. 2021)

The S2 satellite data were accessed through the Python module `geemap` in Colab, a serverless Jupyter notebook computational environment for interactive development. The native GEE Python API has relatively limited functionality for visualising the results, hence the `geemap` Python module was created specifically to fill this gap. The Python code developed enables the analysis of the S2 filtered image collection through spectral index (compositions RGB (bands 4-3-2), false short wave infrared colour (FSWIR, bands 12-8-4) and Bare Soil Index (BSI)), as well as spectral decomposition techniques (hue, saturation and value (HSV), tasselled cap transformation (TCT) and principal component analysis (PCA)). The Python module `rasterio` was also used to create individual plots for each raster band. Additionally, the Python packages `rioxarray` and `matplotlib` were employed, respectively, to access each raster band and create customisable histograms of their values (fig. 2).

## Results

As presented in fig. 1, different sites in the Po plain were employed to test the presented methodological approach and the output of the algorithm after the processing and analysis of the satellite data. Fig. 3 shows the results for Fabbria dei Soci, a relevant and well-documented settlement for the Terramara culture in the selected area, which consisted of a square-shaped village and a hydraulic system that administered water coming from a palaeo-channel river to the surrounding fields.

RGB and FSWIR outputs permit an easy identification of palaeo-channels, moats and canals, as well as the square-shaped habitat area in the middle, while HSV highlights the buried earthworks. At the same time, the PCA and, to a certain extent, the TCT outputs also display the shape of the buried features.

## Discussion

Sensing tool and its applications in archaeological research. First, `geemap`'s inspector and plotting tools enable the user to assess the spectral signature value of different points, such as the buried features and their surrounding area, showing how detecting different values can indicate the location and shape of the earthworks.

Secondly, the second half of the protocol is dedicated to plotting each output band separately. By accessing the data of each band and plotting them against their histogram, it is possible to assess the performance of single bands. As depicted in fig. 4, some bands seem to offer a more relevant contrast between the features and their adjacent areas, performing better than the combined outputs. The most suitable bands for the identification of features can vary according to the characteristics of each study area.

However, it is also worth mentioning the limitations of this approach. Beyond the resolution that GEE outputs can reach, which may not allow to identify buried palaeo-features of smaller size, our approach requires the user to choose the right timespan for the study case. Crop rotation and meteorological conditions must be considered when selecting the best period of visibility, as soil moisture is generally correlated with crop/soil marks.

# PALAEO-LANDSCAPE FEATURE IDENTIFICATION

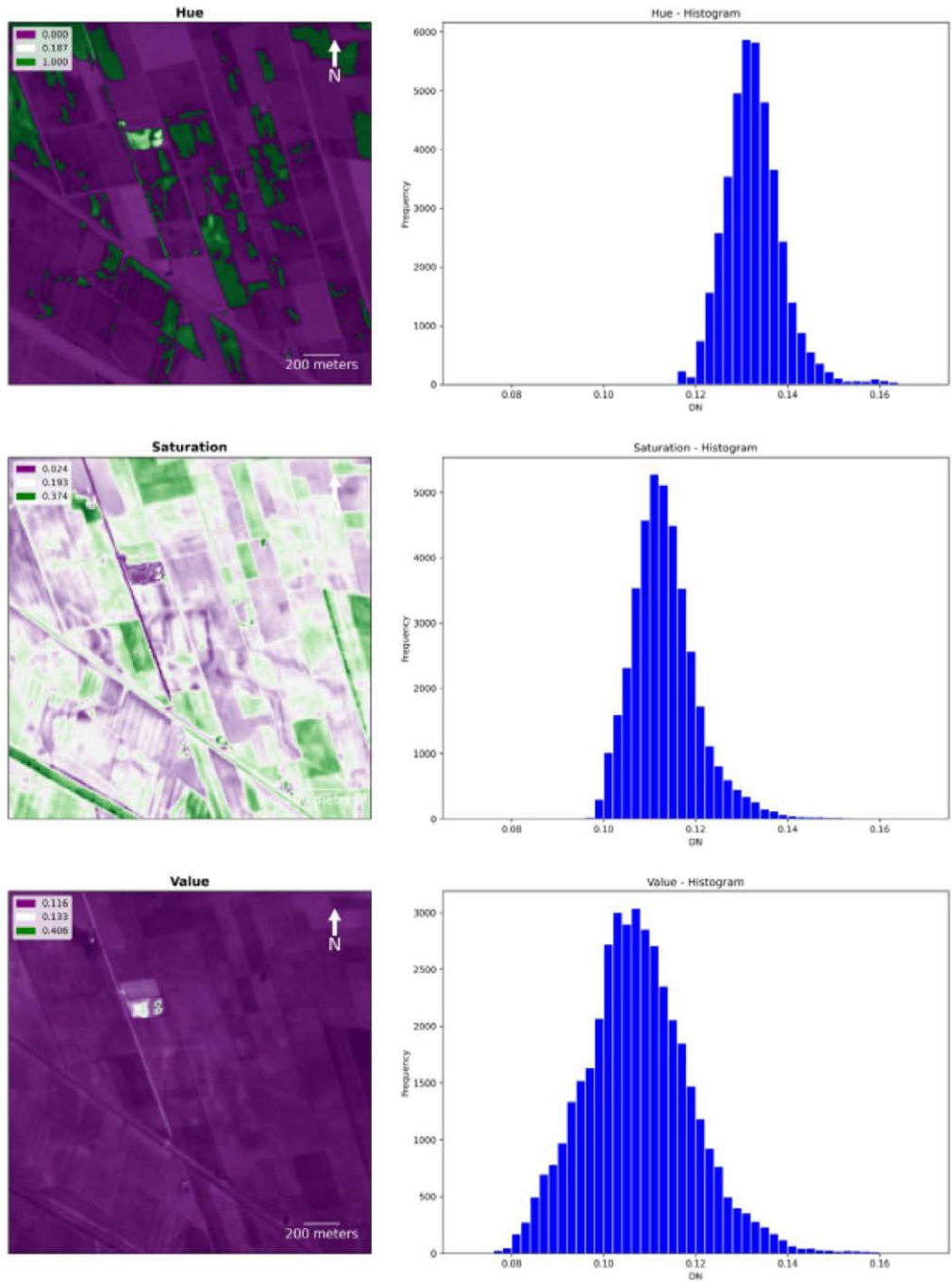


Figure 4. Hue, Saturation and Value plots and histograms of the Fabbri dei Socci site (Brandolini et al. 2021)

## Final remarks

While the choice of timespan is the only part of the protocol that needs to be customised by the final user according to the environmental conditions of each study area, the FOSS-cloud protocol presented here offers significant advantages regarding the mitigation of specialist software, data licensing and computational power. Overall, the outputs generated for the test locations of the Po Plain show some of the potentialities and limits of the GEE Python API in Colab as an alternative remote sensing tool to identify buried natural and anthropogenic palaeo-riverscape features.

**Supplementary Materials:** Link to the Python script (DOI: 10.5281/zenodo.5235030): <https://doi.org/10.5281/zenodo.5235030#.YUJu13AIsIk.link>

**Funding:** This research was funded by the NUAcT PhD fellowship, Newcastle University (UK) (G. Domingo-Ribas), and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No [890561], Historic Landscape and Soil Sustainability (HiLSS) (F. Brandolini).

## References

- Brandolini, F., G. Domingo-Ribas, A. Zerboni, and S. Turner. 2021. "Engine-enabled Python approach for the identification of anthropogenic palaeo-landscape features [version 2; peer review: 2 approved, 1 approved with reservations]." in *Open Research Europe*, 1: 22. DOI: <https://doi.org/10.12688/openreseurope.13135.2>
- Donoho, D. L., A. Maleki, I. Ur Rahman, M. Shahram, and V. Stodden. 2009. "Reproducible research in computational harmonic analysis." in *Computing in Science & Engineering*, 11 (1): 8–18. DOI: <https://doi.org/10.1109/MCSE.2009.15>
- Ducke, B. 2012. "Natives of a connected world: free and open source software in archaeology." in *World Archaeology*, 44 (4): 571–579. DOI: <https://doi.org/10.1080/00438243.2012.743259>
- Fattore, C., N. Abate, F. Faridani, N. Masini, and R. Lasaponara. 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." in *Sensors*, 21 (5): 1791. <https://doi.org/10.3390/s21051791>
- Fryirs, K. A., and G. J. Brierley. 2012. "Geomorphic Analysis of River Systems: An Approach to Reading the Landscape". In *Geomorphic Analysis of River Systems: An Approach to Reading the Landscape*, edited by Kristie A. Fryirs and Gary J. Brierley, 1–8. Chichester: Wiley-Blackwell. DOI: 10.1002/9781118305454
- Harrison, R. 2015. "Beyond "Natural" and "Cultural" Heritage: Toward an Ontological Politics of Heritage in the Age of Anthropocene." in *Heritage & Society*, 8 (1): 24–42. DOI: <https://doi.org/10.1179/2159032X15Z.00000000036>
- Liss, B., M. D. Howland, and T. E. Levy. 2017. "Testing Google Earth Engine for the automatic identification and vectorization of archaeological features: A case study from Faynan, Jordan." in *Journal of Archaeological Science: Reports*, 15: 299–304. <http://dx.doi.org/10.1016/j.jasrep.2017.08.013>

- Orengo, H. 2015. "Open source GIS and geospatial software: towards their integration into everyday archaeological practice". In *Open source archaeology - Ethics and Practice*, edited by Andrew T. Wilson and Ben Edwards, 64-82. Warsaw: De Gruyter. DOI: <https://doi.org/10.1515/9783110440171-006>
- Parcak, S. H. 2009. *Satellite Remote Sensing for Archaeology*. London: Routledge. <https://doi.org/10.4324/9780203881460>
- Traviglia, A., and A. Torsello. 2017. "Landscape Pattern Detection in Archaeological Remote Sensing." in *Geosciences*, 7: 128.
- Vos, Kilian, K. D. Splinter, M. D. Harley, J. A. Simmons, and I. L. Turner. 2019. "CoastSat: A Google Earth Engine-enabled Python toolkit to extract shorelines from publicly available satellite imagery." in *Environmental Modelling and Software*, 122: 104528. DOI: <https://doi.org/10.1016/j.envsoft.2019.104528>