

## Combining TanDEM-X with multi-temporal, multi-source satellite data for the reconstruction of the Bronze Age landscapes of the Indus Civilisation.

H.A. Orengo<sup>1</sup>, A. Garcia<sup>2</sup>, F.C. Conesa<sup>2</sup>, A. Green<sup>2</sup>, R.N. Singh<sup>3</sup>, & C.A. Petrie<sup>2,4</sup>

<sup>1</sup> Ramón y Cajal researcher, Catalan Institute of Classical Archaeology

<sup>2</sup> McDonald Institute for Archaeological Research, University of Cambridge

<sup>3</sup> Department of AIHC and Archaeology, Banaras Hindu University

<sup>4</sup> Department of Archaeology, University of Cambridge

### **Introduction**

The ancient Indus was, along with Mesopotamia and Egypt, one of the three so-called cradles of Old World civilisation. Many Indus sites, some of them large urban nuclei such as Harappa and Mohenjo Daro, lie along major river systems that stretch across modern Pakistan and northwest India. The extremely flat physiography of the Indus River floodplain in conjunction with its proximity to the Himalayas and the distribution of winter westerly rain and rain from the Indian Summer Monsoon combine to produce a distinctive environmental zone where water availability is extremely seasonal, has significant variability, and can produce very active fluvial processes. This variability had enormous consequences for early human habitation of the Indus zone. The relict Ghaggar-Hakra and its tributaries, for example, once extended much of the eastern part of the Indus River basin and appear to have supported

concentrations of settlements across greater Punjab, Haryana and Rajasthan in India and Punjab in Pakistan. This area was a locus for clusters of settlements related to the Indus Civilisation (fig. 1). The Ghaggar-Hakra's middle and lower courses are both flanked and obscured by the advancing dune systems of the Thar Desert of Cholistan.

There has been considerable historical interest in the hydrology of the Indus plain, and since the late 1970s, satellite images have been employed to trace dry riverbeds and reconstruct the ancient hydrological system in relation to the distribution of known sites in the area. However, these analyses have typically relied on visual interpretation of images with visibility being determined by changing environmental conditions, and no systematic and quantifiable attempts have been made to investigate the relationship between landscape features and settlements. Furthermore, the location and distribution of known sites are in many cases unreliable as they were surveyed using 1:63,000 scale maps and before the systematic adoption of handheld GPS receivers.

This paper aims to re-evaluate the relationship between settlement and the changing hydrological systems of the Indus River Basin and explore the mechanisms involved in the human adaptation to and management of changing water availability conditions. Water management and availability are relevant to a range of current debates, archaeological and otherwise, particularly those related to food security, sustainability and resilience. Beyond archaeology, water management has significance for European and global policy, particularly in the discussion on shifts in climate towards more arid conditions. This research thus provides important information on the causes and consequences of water deprivation

at a landscape level, and on the strategies adopted by past human groups to deal with changing conditions on water availability. The insights gained from this analysis have particular importance for understanding the sustainability in the Mediterranean basin, for example, where lack of water is increasingly becoming a problem.

## Methods and results

The team set out to develop and employ a new combined methodological approach for the detection and analysis of palaeorivers and features of archaeological interest in the Sutlej-Yamuna

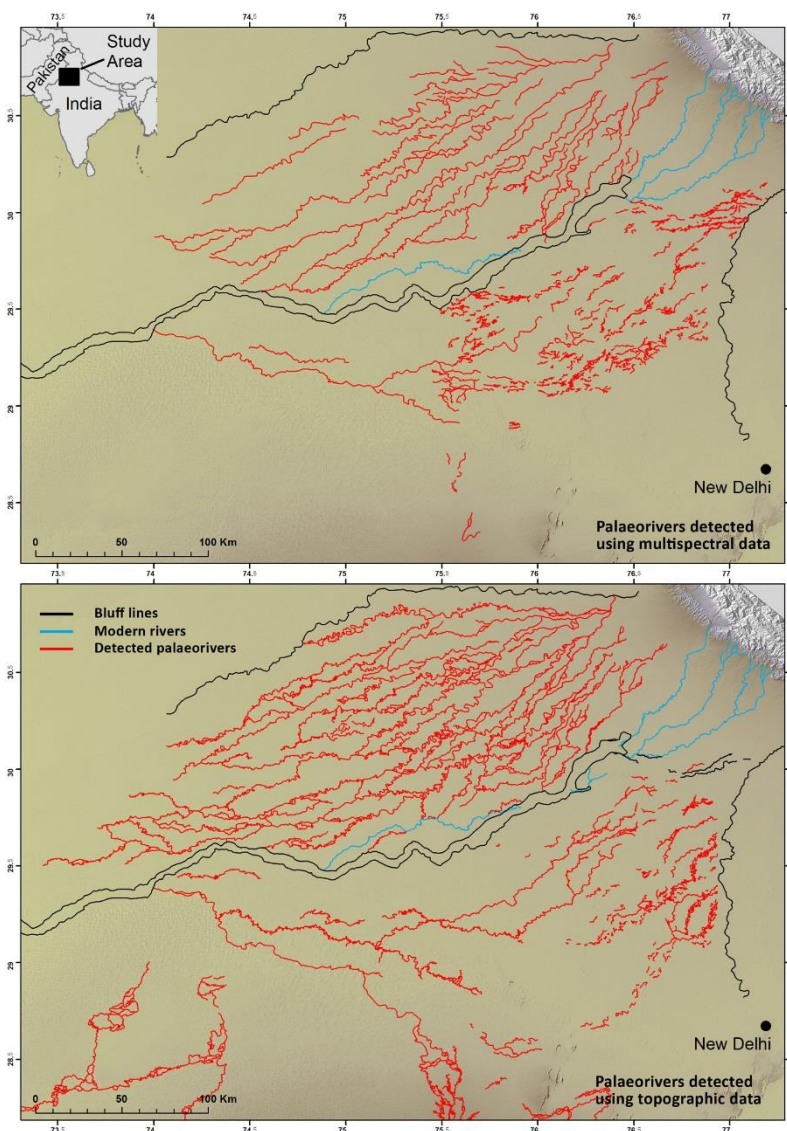
interfluvium in northwest India in order to make a step-change advance on the results of previous remote sensing-based studies. The scope of research has now been extended to encompass areas in Pakistani Punjab. Our research methods can be summarised as follows:

### i. Multispectral-based reconstruction of Indus palaeohydrology.

We applied multi-temporal multi-spectral imagery using both long-term seasonal vegetation indices and spectral decomposition techniques such as Tasseled Cap Transformation and Principal Component Analysis

(Orenco and Petrie 2017). The algorithms were written in Google Earth Engine (GEE)'s implementation of JavaScript. The use of GEE served a triple purpose: (a) it allowed us to tap Google's parallel computing, which was a necessary resource given the compute-intensive character of the analyses the very large size (80,000 km<sup>2</sup>) of our study area; (b) GEE offers access to a large repository of geospatial datasets including freely available multispectral satellite imagery and many global and national DSMs, saving the user the time to download and mosaic the different datasets employed; and (c) GEE is freely accessible (upon registration) and the code employed could be easily distributed for other researchers to test it and employ it. These analyses allowed us to reconstruct around 8,000 km of relict watercourses (fig. 1, top).

Our method dealing with the use of multi-temporal multi-spectral passive imagery for the detection of palaeorivers



**Figure 1.** Comparison between the use of Seasonal multi-temporal multispectral data analysis (above) and MSRM (below) for the detection of palaeorivers in the study area.

has been the object of an open access publication in the journal *Remote Sensing* (freely accessible [here](#)).

**ii. Topographic reconstruction of Indus palaeohydrology.** We have carried out a micro-topography-based geomorphological analysis of northwest India using TanDEM-X data. Different micro-topography filtering methods developed for the analysis of LiDAR data were tested at different spatial resolutions and kernel sizes. It soon became evident that none of these could produce satisfactory results for the detection of large micro-topographic features, in this case palaeorivers (which can be very large in size but presenting an extremely subtle and variable topographic imprint). All these methods were constrained by their focus on a specific kernel size, which resulted in very a partial detection of features. Consequently, we developed an algorithm focused on the micro-topographic analysis of multi-scale features in medium and small-scale DSMs. This algorithm, which we have named Multi-Scale Relief Model (MSRM) (Orengo and Petrie 2018) was also implemented in GEE for the reasons stated above. MSRM using TanDEM-X was successful in detecting more than 15,000 km (fig. 1, bottom) of relict watercourses. The results of the analysis using the MSRM algorithm also allowed us to map these watercourses in detail and, in doing so, describe their morphologic traits and interpret their sedimentary history, which is an important step in the analysis of past human-environment relations. TanDEM-X data was essential in obtaining the high-resolution results in the study area, however, SRTM and other 30 m/px DSMs were also employed outside the study area in northwest India (for which we did not have

TanDEM-X data) to complete the hydrological network. Although the application of MSRM to SRTM data produced noticeably less clear results than in its application to TanDEM-X data, our algorithm was still able to extract meaningful hydrological information from SRTM and improve its readability, reducing stripping and other artefacts cause by the acquisition process. The use of this microtopographic approach complemented the multitemporal imagery analysis (section ii. above), allowing us to complete the reconstruction of the palaeohydrological network in northwest India and obtaining matching information on how these rivers behaved in the past.

Our MSRM algorithm dealing with the microtopographic analysis of multi-scale features has been the object of an open access publication in the journal *Earth Surf. Process. Landforms* (freely available [here](#)).

Geoarchaeological sampling of palaeoriver sediments has been conducted by Prof. C.A.I. French, Dr Julie Durcan and Ms J. Walker. These sediments are being dated and analysed at our geoarchaeological lab at Cambridge and in collaboration with the Department of Geography in Oxford, and will provide important information related to the date and nature of fluvial sedimentation in the study area.

**iii. Machine learning based reconstruction of Indus settlement distribution.** All existing historical maps of the Survey of India for the study area in northwest India have been obtained in digital form (scanned at 700 dpi) and georeferenced (Petrie *et al.* 2019). Features of archaeological interest (mostly mounds that could correspond to ancient habitation sites) have been extracted in vector form and associated to a

geodatabase containing information about their morphology. Also, all Corona declassified satellite images available for our study area have been acquired and georeferenced to obtain high-resolution images of these possible archaeological sites. Fieldwork in India was conducted to test the identifications made using old maps and obtain detailed information about the date and character of the sites (Singh *et al.* in press a, in press b; Green *et al.* in prep). This database is now being employed to develop training data for large-scale machine learning classification of archaeological sites in the study area. The character of both environmental variables and archaeological sites in the study area is highly variable. The former ranges from desert to alluvial plains with large irrigated sectors. Archaeological sites can be preserved in the form of mounds but many have been flattened, are under cultivation or are only preserved under modern population centres. The machine-learning approach required a complex design, based on the use of custom-built image composites created from multi-source data. These include multi-spectral imagery, SAR and topography. Custom-built data include multi-temporal vegetation indices, as underground archaeological sites under cultivated areas will slightly reduce agricultural production, and kernel-specific MSRM, as even flattened mounds will preserve some microtopographic elevation with respect to its immediate surroundings. Preliminary results of automated site-detection in the Cholistan desert provide success rates of 99.9% using known archaeological sites as test data, and our approach has been successful in identifying more than 400 archaeological sites in this area, most of which were previously unknown or inaccurately located.

## Conclusions

TanDEM-X has been transformative to the work that has been carried out in northwest India. The higher spatial resolution and accuracy of TanDEM-X has allowed us to detect multiple palaeorivers in northwest India, which were not visible through conventional optical sensors-based methodologies. In combination with multi-temporal multispectral imagery analysis, we have been able to reconstruct almost completely the palaeohydrology of an area of 80,000 km<sup>2</sup>. Tracing more than 20,000 km of palaeorivers. We hope to be able to expand the scope of our analysis to include the whole of Pakistani Punjab, and ultimately the entirety of the Indus River Basin. TanDEM-X has also been an essential part of our machine learning algorithms for the detection of archaeological sites allowing the detection of sites by means of their subtle topographical imprint. MSRM with its capacity to investigate multi-scale micro-topographies, in combination with the accuracy and resolution of TanDEM-X data present unique potential to make an important contribution to disciplines with interests in surface processes and landforms, such as geomorphology, geology, archaeology, geography, and hydrology.