

MIVIS Hyperspectral Sensors for the Detection and GIS Supported Interpretation of Subsoil Archaeological Sites

Arianna Traviglia

Department of Antiquity and Near Eastern Sciences
University Ca' Foscari of Venice
Venice, Italy
arianna.traviglia@unive.it

Abstract

In recent years, Italian researchers have paid particular attention to the use of airborne hyperspectral sensors applied to archaeology, where the MIVIS sensor has found large application proving to be a very important complementary source of information for archaeological goals. In this paper, the results of research based in Aquileia (northeast Italy) are presented. The spectral content of the MIVIS images has been used to give prominence to the archaeological traces on the base of the different spectral characteristics of the terrain and the vegetation, allowing the detection of an elevated number of surface features. The paper will describe the various processes that have been applied to the images according to the different environmental situations: Vegetation Indices have allowed detection of traces over several types of vegetated surfaces; Soil Index has supported discoveries over bare soil; and Principal Component Analysis provided help in terms of better discrimination of contiguous surfaces.

1 Introduction

Archaeological remote sensing using satellite and airborne sensors to collect multi- and hyper-spectral data is a well-defined domain with huge opportunities for the efficient detection, mapping and management of cultural heritage sites. The possibility of locating unidentified archaeological sites by spectral recognition is, in fact, a valuable addition to traditional survey methods as already demonstrated by a decade of different projects from all over the world (Sever and Parry 2006:439-446, 477-502).

In recent years, there has been particular attention focused on the use of airborne hyperspectral sensors applied to archaeological research in Italy where, for concurrent reasons, the MIVIS sensor has found large application. MIVIS (Multispectral Infrared and Visible Imaging Spectrometer) is a simultaneous multispectral imaging system that operates in the wide range of wavelengths from visible to Thermal-IR regions of the spectrum, with a high spectral resolution and elevated number of channels (102). The runs are taken from a distance of around 5000-6000 feet from the Earth surface: based on flight altitude their pixel resolution is usually around 10x10 ft.

The main advantages of MIVIS reside in the relative ease of transportation of the sensor and in the fact that the shots can be scheduled with the aircraft company in charge of the shooting activity. This, unlike multispectral satellites images, where the purchasing politics limit the possibilities to define with good approximation the moment of acquisition without extra charges, make easier the access to data, the possibility of their collection in specific moment of the day and period of the year, and the opportunity to check in advance the weather conditions to make sure they fit with the research needs. The positive results of all the different projects involving its use in Italy (see Traviglia 2005:139 and Traviglia In Press for a summary) have shown that MIVIS imagery can be considered an interesting additional

data resource for archaeological goals.

In the current research, the spectral content of the MIVIS images is used to reveal the presence of ancient buried sites and structures on the basis of the different spectral characteristics of the terrain and of the vegetation in the area surrounding the ancient Roman town of Aquileia (northeast Italy). Various processes have been applied to the images and their results compared in order to identify the ones that better match the different research targets. The goal of the enhancement techniques is to increase and improve the optic distinction between traces recorded in the scene and the surrounding soil by generating a new image where the useful information is more easily identifiable.

Among them, vegetation indexes have found prominence. These indices in fact can enable the identification of underground archaeological deposits that enhance or, in opposition, limit the growth of the vegetation. As well known from the long tradition of studies applying aerial photographs, heterogeneity of the texture of the subsoil has a strong reflection on the growth of the vegetation, determining the appearance of the so called “crop-, grass-, or weed- marks” (Wilson 2000:67-80). The mechanism of formation of these traces on the vegetation relies on the fact that, when some kinds of solid deposit—possibly archaeological—are present in the subsoil, the vegetation over them will have a slower growth rate compared to the surrounding plants. From the other side, when in the subsoil there are ditches of some sort, the growth of the plants over them will be faster than that of surrounding ones and the vigor more accentuated. Despite the large and often unjustified use of the NDVI (Normalized Difference Vegetation Index) in analogous studies, it can be demonstrated that different vegetation indices can work better for archaeological research, in accordance with certain environmental situations: for this reason, in this work experience, various indices have

been calculated, tested, and compared to determine the best method for evaluating vegetation.

With the same criteria, problems related to situations in which the landscape appears deprived of vegetation have also been taken into consideration. The possibilities offered by studying the tonal variations of the bare soil (damp-mark) are another traditional trend in archaeology: their existence, in fact, can be an indication of the presence in the subsoil of archaeological structures or ditches that inhibit or facilitate the absorption of rainwater and the increase of humidity (Wilson 2000:53-67). The study of the different degrees of water absorption of a soil becomes, therefore, particularly useful for the identification of buried archaeological sites. A specific soil vegetation index needed, consequently, to be developed to support the investigation of the numerous traces over non-vegetated soil; emphasizing the wetness or the dryness of a portion of the ground, it allows the possible identification of underground structures.

In the course of this research, different processes typical of the multispectral analysis have also been successfully applied to hyperspectral shots, like Principal Component Analysis (PCA). As with many other multispectral images, MIVIS images have a strong interband correlation. Contiguous bands of the images convey basically the same information (see below, Figure 7). This produces adjacent bands very similar to each other where the dimensionality of the dataset is increased and the redundancy in the data is high. The PCA is a procedure commonly used for reducing this redundancy (Lillesand et al. 2004:536-538). Useful in common multispectral analysis, PCA finds even better employment in hyperspectral datasets, where the increased number of bands magnifies the need for reducing the dimensionality of the data. This type of transformation has been employed in this research mainly in the enhancement processes preceding visual interpretation of the data, and as a pre-processing procedure prior to further data processing such as the classifications. In addition to the traditional PCA performance, a number of principal component images have been created through spectral subsets in order to convey only the information of spectral regions of interest. A Selective Principal Component Analysis (SPCA) has been computed for groups of homogeneous bands belonging to different spectral regions or to different sensors' spectrometers. In archaeological terms, PCA and SPCA allow for a better discrimination of contiguous surfaces, which become more distinguishable in visual analysis allowing for easier recognition of variations in the texture of surfaces.

Finally, geographic information systems (GIS) technology was used to provide and manage all the archaeological and topographical data necessary for the interpretation of the processed images, and for the eventual recognition of the surface anomalies on the images as ancient origin traces. In the GIS environment, the treated images are subjected to optic analysis and the identified terrain traces or anomalies are graphically traced on a separate layer in order to be compared to the information extracted from different sources, such as archaeological, thematic and cartographic data, and in order to select the ones that with good approximation are linked to underground archaeological structures.

1.1 Methodological Approach

The methodological approach of the research consisted basically of two main phases that are not necessary always in a clear chronological sequence, but that can interact in all the different sub-phases in which they can be subdivided. The first part of the process resides in the observation and treatment of the hyperspectral images; the second part lies in the digital outlining of the identified anomalies and traces in the GIS, and in their selection and interpretation.

The first phase invariably starts after the needed pre-processing operations for noise removal, with a performed-at-screen visual analysis of the images; the goal of this optical survey is not only to identify surface anomalies and traces, that without a previous treatment of the pictures could not be so easily recognized, but also to form a general idea of the landscape and of its characteristics. In this step, the shots are usually analyzed both as single bands and as composites, in true and false colors, using dedicated software for image processing (ENVI 4.0). In the next step, the visual analysis is supported by simple emphasizing techniques like linear, Gaussian, equalization stretches, filters, density slicing, and de-correlation stretch. Finally, a string of advanced processes are applied to the images: depending on the target area and the searched results, arithmetic and statistic operations are applied to the whole image or to part of it, using all or part of the bands composing a run. In this paper, only some of the tested processes, the ones that provided better results, will be presented.

In the second phase of the research, the treated MIVIS images are imported in the GIS environment where they can be georeferenced, read, and interpreted following the criteria of archaeological photo-interpretation. The first step in the process of recognition of archaeological features consists of the observation and reading of the images, basically a visual analysis aiming to simply identify categories of objects on the landscape. Either during or after the completion of the photo-observation and interpretation, the traces recognized on the images are recorded by graphically transferring them to distinct vector layers of the GIS module. At this point the interpretation of the features can start. The nature and characteristics of each element are determined and its causes investigated, also in relationship with the surrounding elements of the landscape and with published information and archaeological thematic cartography, in order to separate the traces due to modern landscape alterations from the ones related to the ancient landscape exploitation (Figure 1).

1.2 The Case Study Area

Located in northeast Italy, the case study area includes the Roman foundation town of Aquileia and its surrounding countryside, comprising the neighboring Communes of Terzo d'Aquileia and Fiumicello, all located in the Province of Udine (Figure 2).

Probably an important center even before the coming of the Romans, the town was established by them as a colony in 181 BC and a frontier fortress on the northeast. Because of its strategic position, it soon became a leading center of

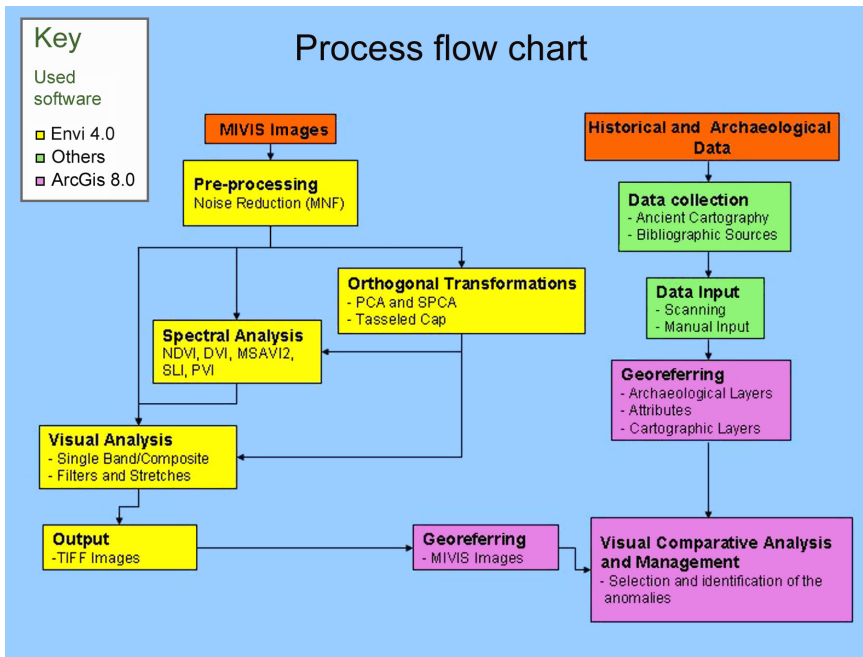


Figure 1. Process flow chart: the GIS manages the processed remote sensing data and the archaeological research data.

trade, especially in agricultural products, and was elevated to the rank of *municipium*, probably in 90 BC. From that moment its importance grew ceaselessly, especially since Aquileia was the starting-point of several important roads leading to the northeastern portion of the Roman domain, and it reached its acme in the 4th century when it became a naval station and along with its already established status as a fortress of the Empire against the barbarians of the North and East. Its destruction at the hand of Attila's Huns was the beginning of its demise, since, although rebuilt and repopulated, it was much diminished and continued to exist until was once more destroyed (AD 590) by the Lombards; with this destruction followed a slow decline in the subsequent decades.

At the current state of the archaeological research, that has more than a century-long tradition, the issues related to the Roman urban area have been favored, with approaches

peculiarities of this territory, an alluvial flat plain mostly devoted to agricultural exploitation, with its strong archaeological potential and a long tradition of studies through extensive archaeological excavations, field-walking survey, aerial photography interpretation, and historical sources, make it ideal for the use of remotely sensed data such as MIVIS images to detect unknown suburban settlements or infrastructures, and as support in solving some of the problems concerning the area. These are in part connected with the landscape morphology, shaped by a long and complex process linked with the rising sea level, and the consequent migration of the local rivers that played, in the past, a fundamental role as part of the communication and trade network. Especially with these kinds of issues, the remote sensing data can take on an important role since the full comprehension of the changes in settlement distribution over time and the forms of spatial organization of the suburban area can

support the comprehension of the planimetric distribution of the town itself.

2 Image Processing

2.1 Pre-processing Operations

The MIVIS images purchased in raw format are strongly affected from panorama distortions due to scanner geometry and from effects introduced by perturbations in position and attitude of the airborne platform. The rectification and georeferencing of the data is consequently a primary difficulty. A geometric correction that rectifies the images is made

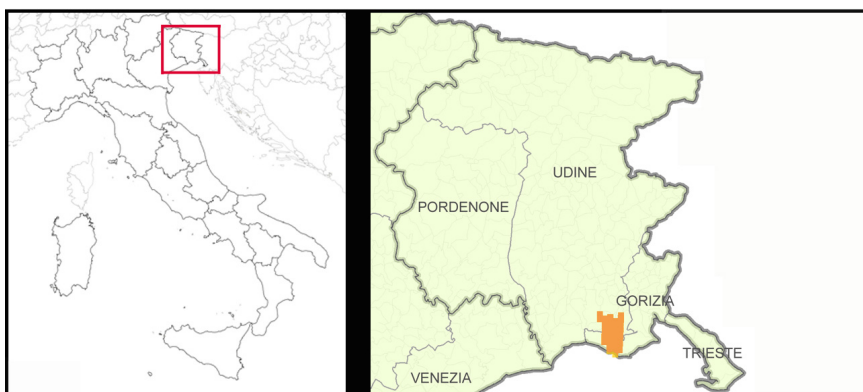


Figure 2. The case study area. In the image on the right, the study area is indicated in darker color.

necessary not only to make the features on the landscape more understandable and recognizable but also to integrate the MIVIS data in the GIS environment. However, all of the most common procedures followed to rectify hyperspectral images risk compromising the radiometric information of the dataset. Thus, the procedure followed uses only raw data in the analysis stage and rectifies and georeferences the images for their integration into the GIS.

Another aspect that has to be taken in consideration before processing the images is the atmospheric distortion and the consequent need for a correction. At the current state of the research, the Italian National Research Council (CNR) is still working on the development of appropriate algorithms to apply to MIVIS images, so it is not currently possible to perform a specific atmospheric correction to the images. However, it can be said that in the kind of analysis and processes that have been applied, the distortion introduced from the atmosphere has been very small and effectively irrelevant.

In order to overcome the shortcomings of the available datasets and improve the accuracy of the obtainable results, a noise removal process was tested and applied, obtaining a drastic reduction of the unwanted disturbance due to limitations of the signal digitization and data recording process. In some cases, in fact, noise was not only degrading the true radiometric information content of the images but also masking it, to the point that several of them were useless. The noise removal process consists of two steps: the first operation, the MNF (Minimum Noise Fraction), extracts the noise through the inherent dimensionality of image data, segregates it, and reduces the computational requirements for subsequent processing (Boardman and Kruse 1994), the second step is the inverse transformation of the previous step (Inverse MNF) that allows for the elimination of the system noise from the original bands by performing the calculation excluding the noise components.

The images submitted to this process have shown a clear improvement of the information in terms of signal noise and have been used in lieu of original data for most of the analysis processes (Figure 3).

2.2 Vegetation Indices

Arithmetic operations on Red and NIR bands have been widely used in the study of vegetation monitoring since they were first demonstrated to be sensitive indicators of the presence and condition of green vegetation. Typically, they provide a black-and-white image that can show variations in the state of health of the plants, where vegetated areas appear in bright color (the higher the value, the healthier the vegetation) and remaining objects have darker colors.

For the current research, different vegetation indices were performed and tested. The results gained were compared among themselves to determine the best method for evaluating vegetation health in the target area and with the original MIVIS image to verify the improvement in visibility that they offer. The comparison of the process results using a multi-criteria analysis has shown that the type of surface under examination must be taken into consideration, and that a single type of vegetation index cannot be applied to all the situations. This means that in studying a vast area presenting variations of vegetation cover, the vegetation indices must be singularly applied based on the type of canopy of the target fields. In the investigated area, very different types of ground cover conditions were encountered, ranging from differential vegetation grow (even inside the same field) to the heterogeneous density of the canopy.

In case of low vegetation cover, the better discrimination of the quality of the vegetation (and consequently in visibility of surface traces) were obtained using the DVI (Difference Vegetation Index). The DVI consists of a subtraction operation involving the use of the Red and NIR bands and performed on a pixel-by-pixel basis to assess the degree of change in the images used (Tucker 1979:128; Lillesand et al. 2004:468).

The formula to obtain the index is well known:

$$\text{DVI} = \text{NIR} - \text{Red} \quad (1)$$

As result of this arithmetic operation, healthy-vegetated areas in the picture show high values because of the relatively high Near Infrared reflectance and low visible reflectance (Figure 4b); less-healthy-vegetated areas show lower values, while rock and bare soil areas have resulted in vegetation indices near 0.

In those areas of inhomogeneous crop canopy, the NDVI (Normalized Difference Vegetation Index) has demonstrated to have very good sensitivity to changes in vegetation cover and to provide better results than any other vegetation index, but only in cases

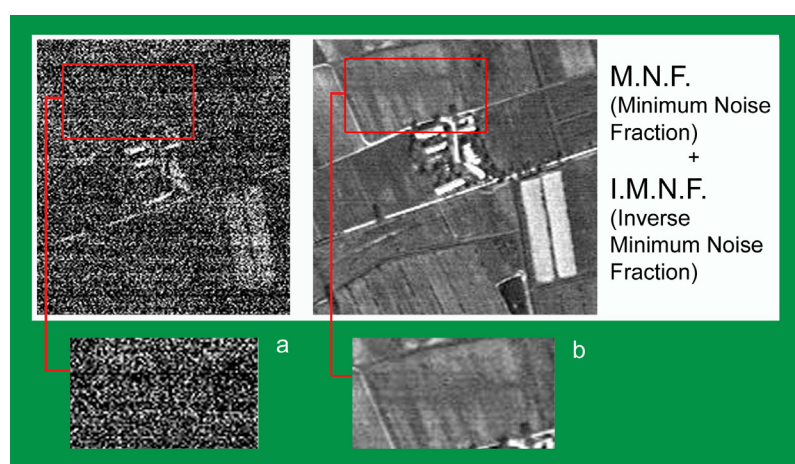


Figure 3. Comparison of an original MIVIS band: a) to one subjected to MNF/Inverse MNF transformation; b) note the improvement of the data quality in terms of signal noise, especially in the detail panels.

of high vegetation cover. It is common knowledge that the NDVI (Rouse et al. 1973) is the difference of the Red and Near Infrared band combination divided by the sum of the Red and Near Infrared band combination or:

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red}) \quad (2)$$

As in DVI, after its application, areas with vegetation yield high values and appear brighter in the picture, with their brightness varying based of their health and maturity. The strength of the NDVI is in its ratio-based concept, which reduces some of the types of noise (illumination differences, cloud shadows, topographic variations) present in multiple bands (Figure 4c); the main disadvantage, as stated, is its sensitivity to canopy background variations in the case of low or sparse vegetation.

Archaeological research has used NDVI almost exclusively, but often more as a matter of routine rather than for a valid reason, which is demonstrated by the fact that it has been utilized indifferently over every kind of canopy condition. NDVI, instead, should be preferred to simple DVI index for vegetation monitoring only in certain types of situations where normalization is needed to help compensate for changes in illumination conditions, surface slope, aspect, and other extraneous factors. In fact, the application of the NDVI over the study area demonstrated that it is less effective with respect to the desired goals—such as the discrimination of the growth of vegetation and of its health—compared to the DVI in those areas where the situations (e.g., very homogeneous) did not require normalization.

In order to bypass the problems of the NDVI in areas of low plant cover, MSAVI2 (2nd Modified Soil Adjusted Vegetation Index), a recursion of MSAVI, was successfully applied. This vegetation index is computed through the formula (Qi et al. 1994):

$$\text{MSAVI2} = (1/2) * [2(\text{NIR} + 1) - \sqrt{(2(\text{NIR} + 1)^2 - 8(\text{NIR} - \text{Red}))}] \quad (3)$$

As with the MSAVI, it was created to minimize the reflectance effects of the soil background on the Vegetation Indices for situations in which the background soil introduces significant variations in the spectral response of the

vegetation. These effects are particularly evident in areas where the coverage of the turf or of the cultivations is lean and/or scattered. This being the case in some portions of the target area, since the shots were taken during the autumn when most of the crops were harvested, the index has shown encouraging results. It is important to stress, however, that, despite the improvement in the attainable results in terms of accuracy, a difficulty that has to be kept in mind in using a vegetation index is an increase in the sensitivity to variations in the atmosphere, which alters the light seen by the instruments. This can cause variations in the calculated values of vegetation indices (Qi et al. 1994; Leprieur et al. 1996). Since no atmospheric correction has been performed on MIVIS images, the atmospheric distortion could be the cause of some of the dissatisfying results obtained with a few of the runs. Like the NDVI, the MSAVI2 annuls the effects of shadows, allowing better discrimination of the traces previously hidden by the shadow itself (Figure 4d).

In conclusion, it is important to stress the need for a preliminary analysis of the target area to identify the type of vegetation cover and the morphological situation, and to evaluate the presence of objects that can cast large shadows over the considered areas before applying any kind of index. Based on the positive experiences with some vegetation indexes, the aim of future research is to test a larger number of them that might better fit specific situations not investigated before, for example, partially submerged canopies or crops at an advanced level of maturation. The ultimate goal here should be to apply this comparative method across a large variety of environmental situations that can be encountered, especially when dealing with large surface areas.

2.3 Soil Line Index

The Soil Line Index (SLI) created in this research for MIVIS data aims to provide an aid in the identification of anomalies on bare soil by increasing the optical distinction between the wetness or the dryness of a portion of the ground.

The soil line is a well-known linear relationship between the near-infrared and red reflectance used originally in studies applied to Landsat MSS (Kauth and Thomas 1976;

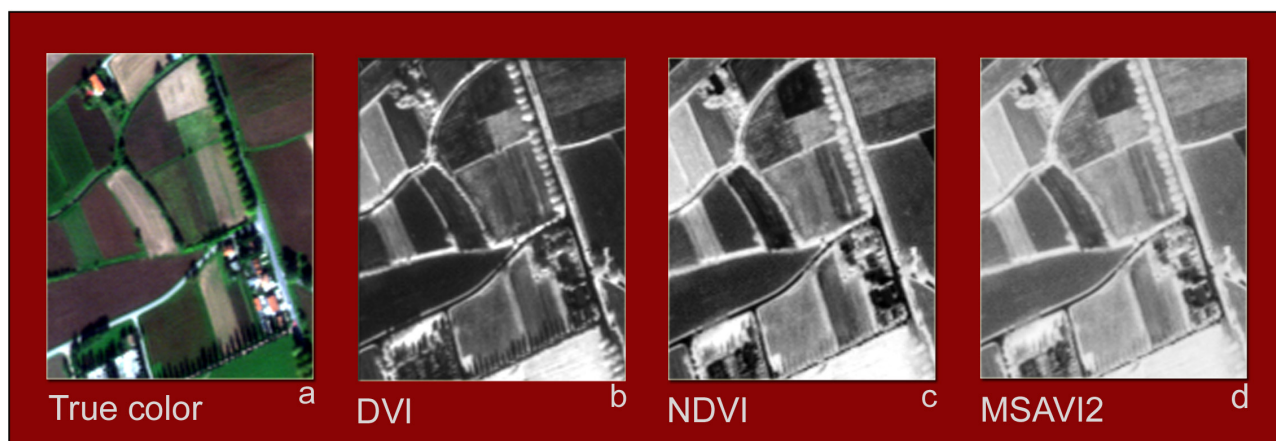


Figure 4. Comparison among vegetation indexes in a detail of one of the MIVIS runs: a) true color image; b) DVI; c) NDVI; d) MSAVI2.

Richardson and Wiengand 1977; Wiegand and Richardson 1982; Baret et al. 1993). The relationship was introduced in the 1970s for identifying agricultural crops and was based on the concept that in a Red-NIR scatter plot of the two-dimensional (2D) space, a definable region is occupied by agricultural crops and another one is occupied by pixels recognizable as water or soil; this last can be identify as a thin, lengthened ellipsoid. In the axis of this ellipsoid, pixels representing soils range from soils of low reflectance to those of high reflectance. The locations occupied in the scatter plot by vegetation, soil, and water can be seen in three distinct areas forming a triangle, which Kauth and Thomas (1976) described as “a triangular, cap shaped region with a tassel.” Later studies have shown some of the limits of the equation defining mathematically the soil line (Mather 2004:146) theorized in the PVI (Perpendicular Vegetation Index) and the need of the definition of the soil line using empirical data. Starting from this concept, with verification that the assumption of the location of the three types of pixels created for Landsat data was true also for MIVIS data, the SLI was determined by identifying a soil line directly in the scatter plot of the radiance measured in the visible Red band against radiance in the NIR (Figure 5a), then determining mathematically the slope “m” through the application of the slope equation formula:

$$m = (y_2 - y_1) / (x_2 - x_1) \quad (4)$$

and subsequently by applying the line equation formula:

$$y = mx + b \quad (5)$$

In the scatter plot, a point (Z) has to be identified that is chosen arbitrarily, identifying the very first and lowest pixel among the ones lying between the soil line and the

vegetation threshold, and representing the left-most pixel of the ones representing wet soil—that is to say, the highest humidity of the scatter plot area—and which can be taken as the lower-left-most point on the scatter plot. The distance from point Z to the projection of any pixel in the scatter plot onto the soil line can be consider as an indication of the moisture content of the soil, ignoring the differences due to soil type and texture (Figure 5b). The distance formula is a simple application of the Pythagorean Theorem. However, this provides the distance from Z to a pixel and not the value projected along the soil line, so it is necessary to look for the soil index component of that distance measure by projecting the distance onto the soil line. This is done trigonometrically through the function Cos of the angle (γ) formed between the segment connecting the point Z to each point P and the soil line:

$$SLI \text{ (Soil Line Index)} = D_{ZP} \text{ Cos } (\gamma) \quad (6)$$

When applying the formula, the obtained image represents a soil humidity index wherein light colors represent the dryness of the soil and dark colors the humidity.

For most of the sample areas, the application of the SLI has shown a clear improvement in the differentiation of the typologies of the soils, accentuating the dry-wet discrimination and thereby making clear the distinction of lines or zones of different condition from the surrounding ground. A particularly interesting group of traces that were detected using this index is located in the west side of the urban settlement, close to the eastern wall of the Roman *circus* (Figure 6). The features, previously noted also through other processing, have been strongly emphasized by the SLI and they have become more clearly distinguishable. Perfectly aligned in parallel and perpendicular to the town orientation, their position and thickness allows one to suppose

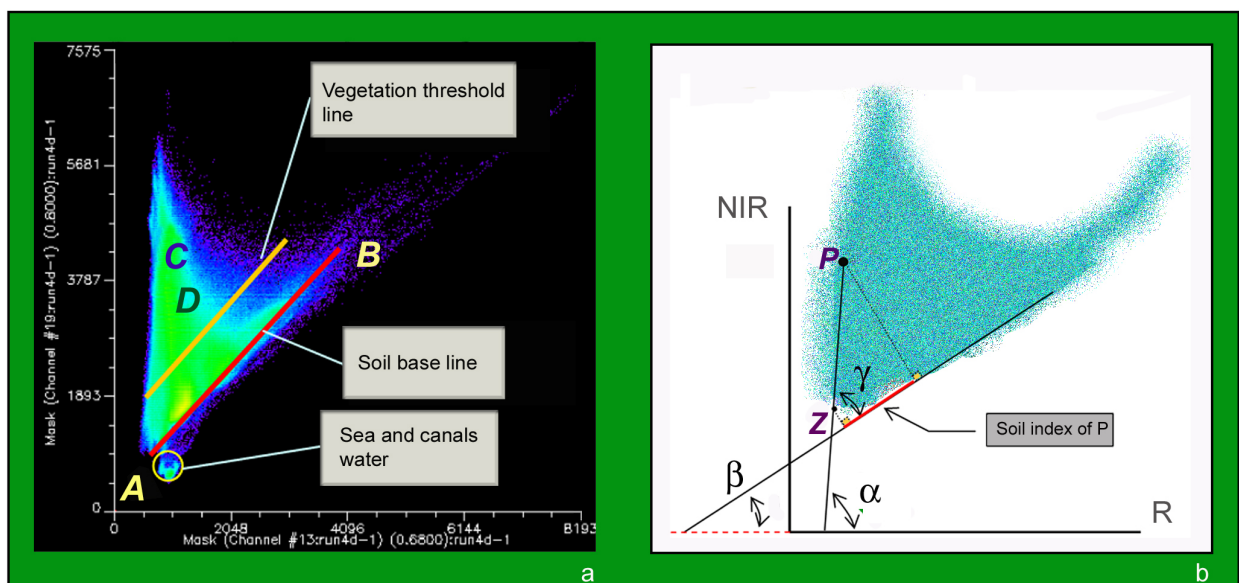


Figure 5. a) Soil line position in the R-NIR scatter plot of a MIVIS run (bands 13-19). The soil line extends from darker soils with low R and NIR image intensity (point A) to an upper region of bright soils with high R and NIR image intensity (point B). Point C represents a pure vegetation pixel and Point D represents a partially vegetated pixel; b) representation of the segment ZP's projection onto the soil line in the R-NIR scatter plot.

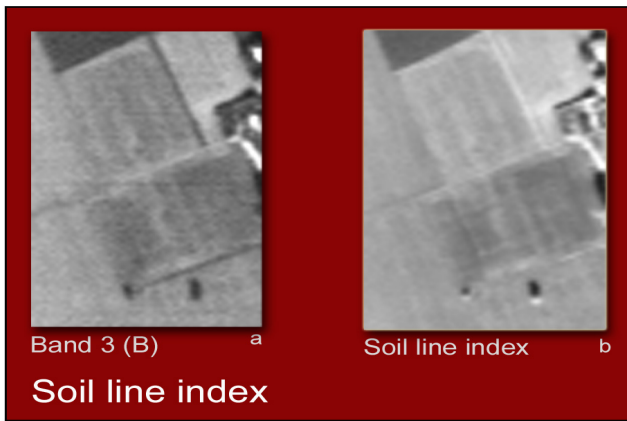


Figure 6. Result of the application of a Soil Line Index (SLI): a better discrimination of the soils can be seen when applying the SLI (6 b) compared to an original band (6 a), and it reveals a larger number of details. The detail shows a group of traces that have been interpreted as the alimentary market of Roman Aquileia.

that they could refer to the alimentary market of the Roman *municipium*, built after the demolition of the previous one (Tiusi 2004:282).

2.4 Principal Component Analysis and Selective Principal Component Analysis

Principal Component Analysis (PCA). The application of the PCA to archaeological study has great potential utility since it provides supplementary information compared to the original bands and avoids useless loss of time in case of a preliminary surveying of the images. Through the PCA the redundancy in data is reduced by transforming a set of correlated variables into a new set of uncorrelated ones. This transformation entails a rotation of the original axes to new orientations that are orthogonal to each other, and therefore, there is no correlation between variables. The goal of the

process is to reduce the information previously contained in the original n-band dataset into a smaller number of new bands that can be used in place of the original ones (Lillesand et al. 2004:536-542).

In the analyzed MIVIS images subjected to PCA, more than 74% of the variability in the data is carried in the first “principal component” (PC) and around 24% in the second PC: together, the PC1 and PC2 in most cases account for 98.9% of the total variability in the original bands. Usually, about 1% is found in 3rd PC. PC1, PC2, and PC3 show virtually all of the variance in the scene (on average 99.6 %) and, consequently, of the total information. PCs from 4 and higher, together, usually contain only 0.60% of the variation in the data. However, some of these higher-order components have been demonstrated to contain useful information, recognizable and identifiable only through a visual check of the image itself. The presence of significant additional information that is not present in some of the lower orders has been noticed even up to the 30-40th PCs, the value varying from run to run. Technically, since in the PCA process the information is accumulated in the lower-order components and in the dataset the noise is evenly distributed, the lower-order principal components might be expected to have a higher signal-to-noise ratio than the higher-order PCs, which could lead one to think that higher-order PCs are not worthy of consideration. This was not always the case with MIVIS images and necessitated visually checking of the PC images, singularly, using the knowledge of the study area, rather than relying solely upon the magnitudes of the eigenvalues as an indicator of information content. All of the PC images for each run have been taken in consideration, at least in the first stage of the work, in order to define which of them contained information that could be consider useful for the research. PC images are analyzed subjecting them to visual interpretation of the data both as separate black-and-white images and as any three-component images combined to form a color composite Their enhancements are generated by displaying contrast-stretched images of the transformed

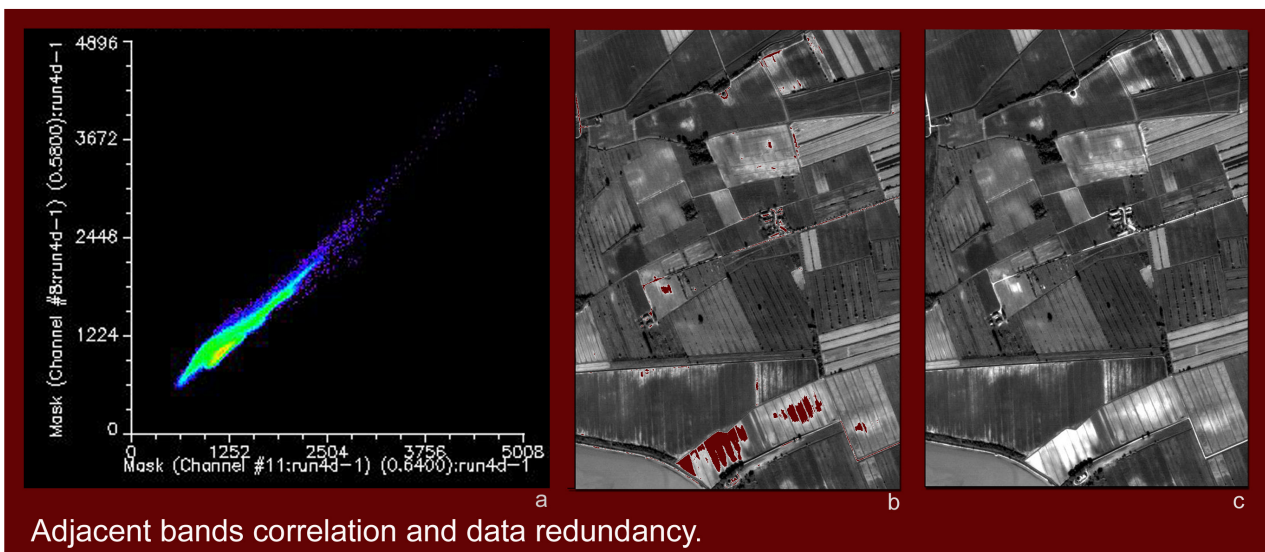


Figure 7. a) 2D scatter plot of Green band 8 (vertical axis) vs. Red band 11 (horizontal axis) showing high correlation between the two bands; b) Band 8 of a MIVIS run; c) Band 11 of the same MIVIS run.

pixel values. PC bands produce more colorful composite images than spectral color composite images because the data, once run through the process, are uncorrelated. The most common procedure consists of generating a RGB false color composite with the first three PCs of a scene (called PC1, PC 2, PC3): this creates six different combinations, but, in the case of MIVIS data, this means excluding from the composite many of the higher-order PCs that previously were shown to carry interesting information. Therefore, tests were made increasing the number of PCs used, bringing it to four. Using any three of these four PC images in a color composite with various assignments of blue, green, and red produces a total of 24 different combinations. Of those, the image composed of PC4 shown in blue, PC1 shown in green, and PC3 shown in red was the most interesting in terms of visibility of the soil irregularities. A small percentage of the images that passed the first selection were chosen to be subjected to a deeper visual analysis and to be integrated into the GIS for the subsequent archaeological interpretation.

The PC images (single component or composite) have proven to be a good starting point for a general overview of the target landscape, and they enabled the detection of several new features in the target area. However, many details that are visible analyzing the original bands singularly are not always recognizable in the PC because they are covered by the overlaying of information from other bands.

Selective Principal Component Analysis (SPCA). In order to overcome this problem and convey only the information of spectral regions of interest, a number of PC images can be created through spectral subsetting, producing a Selective Principal Component Analysis (SPCA); that is to say, a PCA computed for a group of bands (Table 1), for example for the group of channels of the 1st spectrometer or for the channels of a spectral region (e.g., Blue or Red), instead of using all the bands of the run. This selection preserves the separation of the spectra and the possibility for combination choices.

The visual analysis of the newly created bands of each spectral subset (SPCs) as single images reveal that the lower-order SPCs of each grouping (like SPC1, SPC2, and so on

of each subset) could provide a large amount of information that cannot be recognized in the original bands of each relative spectral group. The first three PCs of the subset of the visible bands (i.e., the SPCA computed on the 1st spectrometer), for example, show what could be considered the best results of the groups of the subsets: several of the traces and anomalies on the investigated surface can be identified and better discriminated from the surrounding environment. The lower-order components of the subset of the 2nd spectrometer, summarizing the NIR bands, prove to be a valid instrument in studying the vigor of the vegetation. From the other side, they show little reliability in the discrimination of different soils. These can be better analyzed through the 1st SPC image of the thermal subset of bands (i.e., spectrometer 4) where possible variations due to the presence of particular kind of sediments, rocks, or lateritic material in the subsoil can result in variation of the temperature. The useful information that can be provided by the single thermal bands is magnified and more clearly observable here. The least useful appears to be the MIR subset of bands (i.e., spectrometer 3), where the differentiation between soils or quality of vegetation is only slightly recognizable.

The best outcomes in terms of recognizable features can be reached through the composite of different spectral subsets, using a dedicated correlation matrix in order to identify the minimum set of SPCs able to provide complete information without loss of spectral coverage. This process is made necessary by the large number of combinations possible for the many SPCs. The selection of the SPCs to use has been based on the Intrinsic Dimensionality (ID) of the data for each subset; that is, the amount of SPCs of each subset able to carry interesting and original information. The ID of the subset of the 1st spectrometer, for example, is 3, since the first three SPCs contain useful information, while for the subset of all the Red bands the ID is equal to 1, since SPCs from Red channels after the 1st do not contain unique information (Table 2). Consequently, for each of the listed subsets of bands, the Eigenvalues have been evaluated in order to define which of the SPCs were going to be selected and used, keeping only SPCs holding high values and consequently showing the maximum information for a given subset.

Table 1. Table of the spectral subset for MIVIS runs: specific bands have been grouped together and subjected to SPCA.

GROUPING	USED BANDS	BAND EDGES	SPECTRAL REGIONS INVOLVED	TOTAL # OUTPUT PC IMAGES
Optical port 1	1-20	0.431-0.833	RGB+NIR (3 bands)	20
Optical port 2	21-28	1.15-1.2	NIR (8 bands)	8
Optical port 3	29-92	1.985-2.479	MIR	64
Optical port 4	93-102	8.21-12.7	TIR	10
Blue spectral region	1-4	0.481-0.512	B	4
Green spectral region	5-9	0.52-0.611	G	5
Red spectral region	10-17	0.611-0.773	R	8
NIR spectral region	18-28	0.772-0.833/1.15-1.55	NIR (11 bands)	11
NIR Optical port 1	18-20	0.772-0.833	NIR (3 bands)	3
Visible spectral region	1-17	0.431-0.773	RGB	17

The 12 selected SPCs were merged into one multi-bands image and used to calculate a correlation matrix. As a result, the SPCs mostly correlated (both negative and positive) to every other SPCs, were identified, which allowed the discard of those components most correlated and consequently carrying redundant information (Table 3).

Using the matrix to identify the correlations between components, a minimal image subset containing all the information was selected yielding a combination of bands corresponding to PC1 of the 1st spectrometer, PC2 of the 1st spectrometer, PC1 of the 4th spectrometer, and PC1 of the Blue region subset bands.

A series of 24 combinations of the four selected PC images were used to create composites in the major part of the executed tests; the most successful visualizations appear to be the ones where the PC1 of the 1st spectrometer, carrying information on the visible bands and the NIR, and the PC2 of the 1st spectrometer, which is correlated with NIR, MIR, and Blue, were present (Table 4).

Table 2. PC images selected on the base of Intrinsic Dimensionality: the selected images are the ones containing the maximum amount of information for each groups subjected to SPCA. In the 3rd column the "image #" identifying the band number assigned to the image for the correlation matrix (see next table).

GROUPING	I.D.	IMAGE #
Optical port 1	PC1	1
	PC2	2
	PC3	3
Optical port 2	PC1	4
	PC2	5
	PC3	6
Optical port 3	PC1	7
Optical port 4	PC1	8
Blue spectral region	PC1	9
Green spectral region	PC1	10
Red spectral region	PC1	11
NIR spectral region	PC1	12

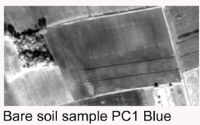
Again, in cases like this, one is faced with an elevated number of images to analyze visually: therefore, some kind of selection of the best results needs to be applied. An evaluation of the quality of the composite can be done by selecting sample areas representative of anomalies already identified through other processes (in this example, four sample areas have been used: a sample area of traces on dry bare soil, one on wet bare soil, one on vegetation, one on sea water) and attributing to them a quality score from 1 (low visibility) to 5 (high visibility). The sum of the four grades given for each combination provides the scale index of the quality of the composite and the first three or four images (with the highest index) of the scale are the ones that should be selected for visual analysis.

Table 4. Combinations of possible RGB color composites of SPCs (see Table 2 for the key of the numbers identifying the SP components "image #") listed in order of significance: their position in the list is based on the sum of the quality scores attributed to sample traces.

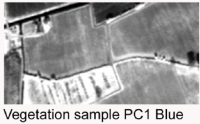
R	G	B	Water	D.b.s.	Veg.	W.b.s.	TOT
1	2	9	3	4	3	4	14
1	2	8	4	3	3	3	13
2	1	9	3	3	2	2	10
2	8	1	3	3	1	2	9
2	9	1	2	2	1	3	8
8	2	1	3	2	2	1	8
1	9	2	1	3	1	2	7
1	9	8	1	3	1	2	7
2	1	8	2	2	1	2	7
8	2	9	3	1	2	1	7
9	1	2	1	3	1	2	7
9	2	8	2	1	3	1	7
1	8	2	2	2	1	1	6
2	8	9	2	2	1	1	6
8	9	2	1	2	2	1	6
9	8	2	2	2	1	1	6
9	2	1	1	1	2	2	6
1	8	9	1	2	1	1	5
2	9	8	1	1	1	2	5
8	1	2	1	2	1	1	5
8	1	9	1	2	1	1	5
8	9	1	1	2	1	1	5
9	1	8	1	2	1	1	5
9	8	1	1	1	1	1	4

Key

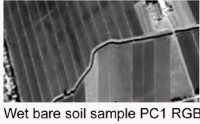
Water = sea water
D.b.s. = dry bare soil
Veg. = vegetation
W.b.s. = wet bare soil



Bare soil sample PC1 Blue



Vegetation sample PC1 Blue



Wet bare soil sample PC1 RGB

Table 3. Correlation matrix of the 12 SPCA images (see Table 2 for the key of the numbers identifying the SPCs "image #"). The dark highlighted values identify the highest positive or negative correlations of the other SPCs in the same column to the value of the examined SPC.

	1	2	3	4	5	6	7	8	9	10	11	12
	PC1 I.O.P.	PC2 I.O.P.	PC3 I.O.P.	PC1 II O.P.	PC2 II O.P.	PC3 II O.P.	PC1 III O.P.	PC1 IV O.P.	PC1 B.R.	PC1 G.R.	PC1 R.R.	PC1 N.I.R.
1	1,000	0,000	0,000	0,958	0,073	-0,083	0,370	0,438	0,531	0,651	0,986	0,971
2	0,000	1,000	0,000	-0,012	-0,912	-0,050	0,785	0,452	0,805	0,748	0,161	-0,235
3	0,000	0,000	1,000	0,179	-0,242	0,269	0,315	0,180	-0,242	-0,122	0,047	0,017
4	0,958	-0,012	0,179	1,000	0,000	0,000	0,445	0,521	0,459	0,591	0,949	0,942
5	0,073	-0,912	-0,242	0,000	1,000	0,000	-0,871	-0,538	-0,649	-0,599	-0,085	0,277
6	-0,083	-0,050	0,269	0,000	0,000	1,000	0,060	-0,273	-0,185	-0,111	-0,075	-0,064
7	0,370	0,785	0,315	0,445	-0,871	0,060	1,000	0,690	0,762	0,784	0,504	0,185
8	0,438	0,452	0,180	0,521	-0,538	-0,273	0,690	1,000	0,584	0,589	0,509	0,333
9	0,531	0,805	-0,242	0,459	-0,649	-0,185	0,762	0,584	1,000	0,972	0,641	0,323
10	0,651	0,748	-0,122	0,591	-0,599	-0,111	0,784	0,589	0,972	1,000	0,756	0,454
11	0,986	0,161	0,047	0,949	-0,085	-0,075	0,504	0,509	0,641	0,756	1,000	0,920
12	0,971	-0,235	0,017	0,942	0,277	-0,064	0,185	0,333	0,323	0,454	0,920	1,000

Correlation matrix of Principal Components

Results of the PCA Application. The application of the PCA to the MIVIS images has shown satisfactory results in the recognition of anomalous features on the landscape. Condensing in a few bands all the information previously scattered in hundreds of bands permits faster and more accurate research on traces over vegetation, bare soil, and water. Visual interpretation of the data has been strongly aided by the combination of PC images used to form a color composite: this was particularly valid when applied to PCs obtained from subsets of bands. As a method for using this kind of process, it is advisable to use primary PCs calculated using all the bands to gain a general knowledge of the studied landscape: this can provide the preliminary information that can be further probed using PCs from the subsets. However, the PCA calculated using all the bands or a portion of them (meaning the exclusion of the thermal bands) has been shown to be the least effective process, from the numerical point of view, of landscape feature identifications even if it is the best process to apply in terms of visibility of the traces. Consequently, the composite realized through the use of these types of PCs provide poorer results than the ones obtained by separating the bands.

The best results in terms of amount of recognizable features were reached though the composites of spectral subset components, selected from the correlation matrix. The features that were best emphasized when applying this

operations were lineation of dry or wet soil in contrast with the surrounding soil, alterations in the health of the vegetation and its growth, filled riverbeds, and unexplainable alterations in the surface texture. A remarkable result achieved through this analysis is the increased quantity of details in relation to the route of the ancient *Via Annia* (Annian Way), an important artery that, starting from the Roman town of Adria, ran across the Adriatic arch up to Aquileia. The track of the ancient consular way exiting on the east side of the town, known in many of its stretches, appears in the images obtained through the SPCA, acquiring definition and visibility, especially where the trace is visible on bare soil.

3 Archaeological Interpretation of Remotely Sensed Data through GIS

The processed MIVIS images, imported in the GIS environment and georeferenced, are interpreted according to the principles of archaeological photo-interpretation. The features present in these images are in fact identified through the recognition of the traditional factors of photo-interpretation: proximity, dimension, alignment, orientation, shape, texture, pattern, tone, and size of the features. Obviously, not all of the traces visible on the landscape surface can be ascribed to factors of archaeological origin, since the

mechanisms that produce them could have happened in any historical period. For this reason, the nature and attributes of each feature must be investigated and their causes determined in relationship with the contiguous aspects of the landscape and with all the other data extracted from various sources that can lead to a valid interpretation. This collateral documentation, organized as an archaeological thematic map in vector format, is created by a methodic collection of published information and thematic cartography. Traces and anomalies identified through MIVIS and aerial pictures can be compared at this point with modern and archaeological elements of the landscape. Based on the comparison, a selection can now be done to eliminate the features that are not ascribable to archaeological presences but must be due, more realistically, to modern interventions on the territory.

To assist in the analysis of the identified anomalies, which are recorded by drawing them in layers of the GIS, several attributes are also recorded, such as the mechanism of their formation

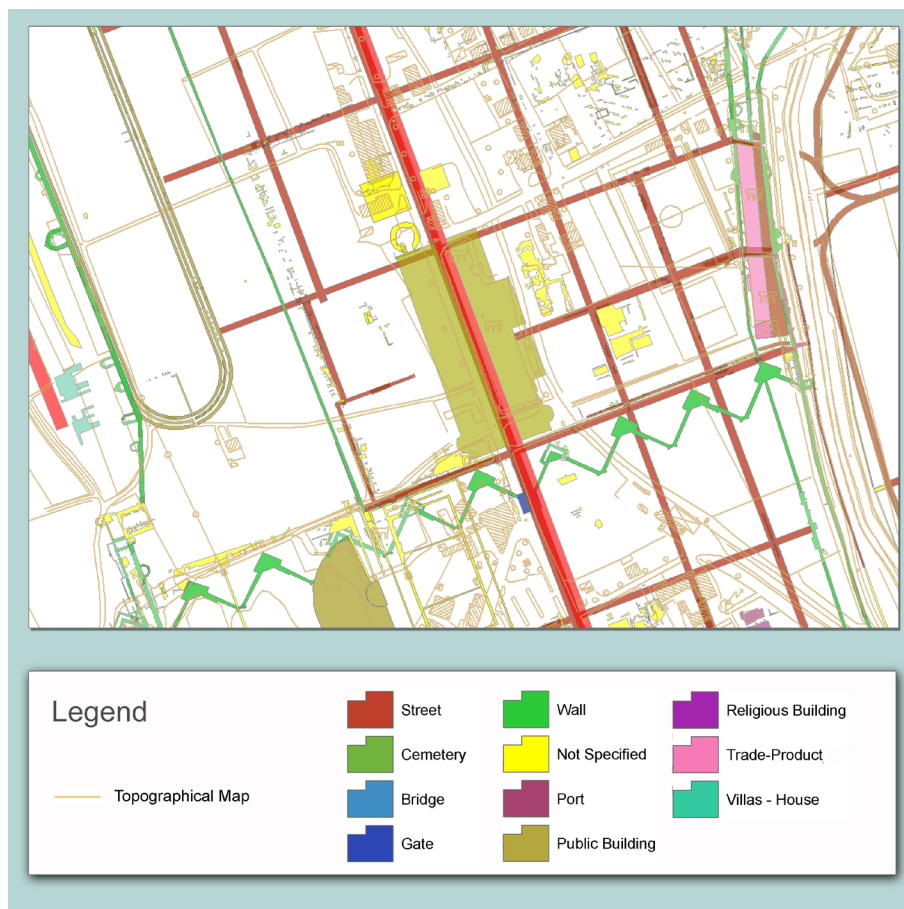


Figure 8. Different archaeological objects on the archeological thematic map layer (hypothetical tracks or locations are semi-transparent). The detail shows the central urban area of Aquileia.

(positive, negative, and unknown) and the type of image processing from which they have been extracted (Figure 9).

In order to overcome the lack of ground truth information, still irreplaceable in order to give trustworthiness to the analysis, interpretation, and confirmation of the data, two indices have been created: the “archaeological reliability value” and the “visibility index.” These indices express an evaluation of the quality of the observed feature. Although unable to substitute the value of field-walking survey, they can at least emphasize those locations on the landscape in which there is a higher probability to find archaeological deposits.

The index of “archaeological reliability” derives from the combination of data about the area surrounding a specific, identified trace. The index is expressed by a percentage from 10 to 100%, which indicates the level of reliability that a feature is of archaeological origin. When multiple data concur to confirm the hypothesis that the feature could be archaeological, the value of the percentage is elevated; on the contrary, when there are one or more hints showing that the feature could be from modern times, the percentage will be low (Figure 9). The “Index of Visibility” is an indicator of the quality of visual appearance of the traces; it is evaluated through a scale of values from one to five, where one indicates slight possibility to identify the trace and five means that the trace is extremely visible.

4 Conclusions

At the onset of this research, the main goal was to ascertain the usefulness of the MIVIS images in the identification of archaeological sites in the target area, simply meaning to identify the quality and quantity of new information produced. While still working toward new developments, it is possible to assert that MIVIS shots and the processing of them have given very positive verification and demonstrate MIVIS to be a valid instrument for archaeological research if integrated in a global information architecture managing various data. Due to its intrinsic characteristics, the level of detail provided by MIVIS images has proven to be useful for detection of potentially archaeological landscape features and that it allows the identification of a large range of objects, showing to be efficient in recognizing also relatively small traces. It must be

clearly stated, however, that in the case of the identification of small objects, they were, most of the times, detected in the GIS module on the basis of indications offered from other remotely sensed imagery (like Orthophoto), and from the archaeological layers, emphasizing the need for an approach integrating multiple tools and data. An efficient research strategy must in fact be oriented to the combined use of different remote sensing products managed in GIS environments, and to merge them with other information layers in order to obtain results that are otherwise unreachable. This also clarifies that the values of the MIVIS images are more in their hyperspectral content than in their spatial resolution, which cannot be considered completely adequate for the search of small archaeological remains.

It is easy to see, also, that the application of different image processing techniques and analyses tends to a combinatorial explosion of the number of newly produced images. Consequently, since many of the different processes produce similar results that provide the same traces, the number of repeatedly identified and recorded features turns out to be very high. It becomes, then, necessary to create correct methodologies that reduce the redundancy of the data and that define a selection process to identify the smallest set of processes that give complete coverage of all the

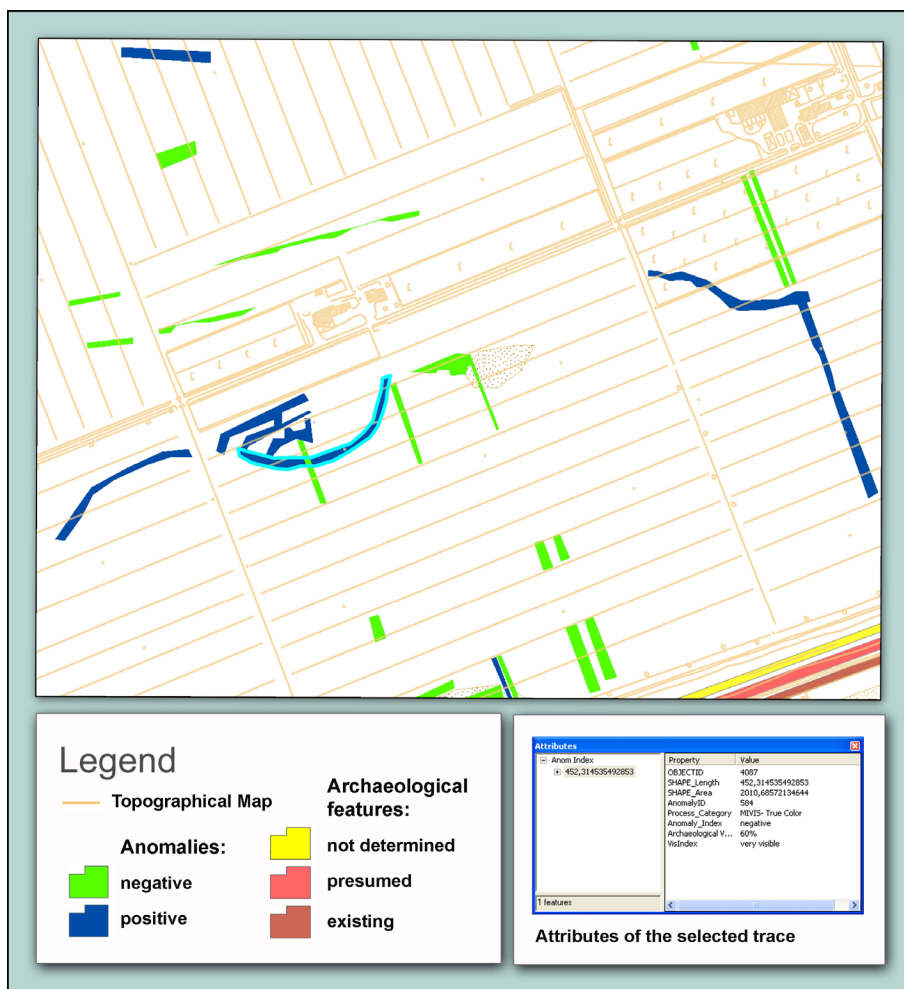


Figure 9. Traces recognized on a processed MIVIS image and reported in a drawn layer of the GIS; they are differentiated on the base of the nature of their formation (positive, negative, unknown) by different colors.

detected traces. In doing this, we must rely increasingly on external statistical tools and on the GIS. Until now the GIS platforms have been sparsely employed in archaeological remote sensing research or used only to provide and manage all the archaeological and topographical data, combining them through a simple overlay. The increased number of processed images and consequent data amount will impose the use of a more complex GIS architecture, the new needs pushing in the direction of more advanced structures that allow for more sophisticated storing and visualization techniques. In this way the GIS will assume an indispensable role in the interpretation of remotely sensed data leading to a multi-tool approach which in a multi-scale and diachronic context is the only reliable way to deliver effective contributions to the understanding of settlement patterns.

Acknowledgements

I take here the opportunity to thank the “Servizio sistema informativo territoriale e cartografia,” Department of the “Direzione centrale pianificazione, mobilità e infrastrutture di trasporto” of Regione Autonoma Friuli Venezia Giulia in the person of the Arch. Mario Ghidini for allowing the use of tools and data (MIVIS images, Orthophotos, CTR) indispensable for this work. MIVIS images, Orthophotos, Regional Technical Map (CTR): received authorization from Regione Autonoma Friuli Venezia Giulia, P.M.T./1295/2100, Jan. 25th 2005. Orthophotos: property of Compagnia Generale Ripresearee S.p.A. Parma. I am also grateful to Eng. Stephen A. White for his valuable suggestions and technical support.

References Cited

Baret, F., Jacquemond, S., and Hanocq, J. F. 1993. The soil line concept in remote sensing. *Remote Sensing Reviews* 7:65-82.

Bertacchi, Luisa. 2003. *Nuova pianta archeologica di Aquileia*. Udine: Edizioni del Confine.

Boardman, Joseph W. and Kruse, Fred A. 1994. Automated spectral analysis: a geological example using AVIRIS data, northern Grapevine Mountains, Nevada. In, *Proceedings of the Tenth Thematic Conference on Geological Remote Sensing, 9-12 May 1994, San Antonio, Texas, Environmental Research Institute of Michigan, Ann Arbor*, pp. 407-418. Ann Arbor, MI: ERIM.

Bottazzi, Gianluca and Buora, Maurizio. 1999. Nuovi dati sul territorio di Aquileia romana. *Antichità Altoadriatiche* 45:61-77.

Carre, Marie-Brigitte and Maselli Scotti, Franca. 2001. Il porto di Aquileia: dati antichi e ritrovamenti recenti. *Antichità Altoadriatiche* 46:211-243.

Kauth, Richard J. and Thomas, G. S. 1976. The tasseled cap. A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In, *Proceeding of the Symposium on Machine Processing of Remotely Sensed Data, June 29-July 1, 1976, the Laboratory for Applications of Remote Sensing, Purdue University of West Lafayette, Indiana*. P. H. Swain and D.B. Morrison, eds., vol. 76 CH, pp. 41-51. New York, NY: Institute of Electrical and Electronics Engineers.

Leprieur, C., Kerr, Y.H., and Pichon, Jean-Marc. 1996. Critical assessment of vegetation indices from AVHRR in a semi-arid environment. *International Journal of Remote Sensing* 17:2549-2564.

Lillesand, Thomas M., Kiefer, Ralph W., and Chipman, Jonathan W. 2004. *Remote Sensing and Image Interpretation*. New York: Wiley.

Maggi, Paola and Oriolo, Flaviana. 1999. Dati d'archivio e prospezioni di superficie: nuove prospettive di ricerca per il territorio suburbano di Aquileia. *Antichità Altoadriatiche* 45:99-123.

Mather, Paul M. 2004. *Computer Processing of Remotely-Sensed Images: an Introduction*. Chichester: Wiley.

Qi, J., Chehbouni, A., Huete, A. R., and Kerr, Y. H. 1994. Modified Soil Adjusted Vegetation Index (MSAVI). *Remote Sensing of Environment* 48:119-126.

Richardson, A. J., and Wiengand, C. L. 1977. Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing* 43:1541-1552.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W. 1973. Monitoring vegetation systems in the Great Plains with ERTS. In, *Proceedings Third ERTS Symposium, NASA SP-351, Goddard Space Flight Center, Washington D.C.*, vol. 1, pp. 309-317. Washington, DC: NASA.

Sever, Tom L. and Parry J. T. 2006. Archaeological Remote Sensing of early human settlements. In, *Remote Sensing of Human Settlements. Manual of Remote Sensing Volume 5*, M. K. Ridd and James D. Hipple, eds., pp. 431-502. Bethesda, MD: ASPRS.

Tiussi, Cristiano. 2004. Il sistema di distribuzione di Aquileia: mercati e magazzini. *Antichità Altoadriatiche* 59:257-316.

Traviglia, Arianna. 2005. Integration of MIVIS hyperspectral remotely sensed data and geographical information systems to study ancient landscapes: the Aquileia case study. *Agri Centuriati. An International Journal of Landscape Archaeology* 2:139-170.

Traviglia, Arianna. In press. Identificazione di tracce archeologiche sul territorio tramite immagini da sensore aviotrasportato MIVIS: l'esempio di Aquileia. *Archeologia aerea* 3.

Tucker, Compton J. 1979. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sensing of Environment* 8:127-150.

Wiegand, C. L. and Richardson, A. J. 1982. Comparisons among a new soil index and other two- and four-dimensional vegetation indices. In, American Congress on Surveying and Mapping, American Society of Photogrammetry Convention; APS Annual Meeting, 48th, Denver, CO; United States; 14-20 Mar. 1982, pp. 210-227. Falls Church, VA: American Society of Photogrammetry.

Wilson, David R. 2000. *Air Photo Interpretation for Archaeologists*. London: Tempus.