A semi-empirical index for estimating soil moisture from MIVIS data to identify sub-surface archaeological sites.

Arianna TRAVIGLIA (*)
(*) GeoNetLab, Università degli Studi di Trieste, Piazzale Europa 1, Trieste, travigli@units.it

RIASSUNTO
Sono note in archeologia le possibilità offerte dallo studio dei passaggi tonali del suolo nudo ("damp-mark"): il loro verificarsi infatti può essere indice di presenza nel primo sottosuolo di strutture archeologiche o fossati che rispettivamente inibiscono o agevolano l'assorbimento dell'acqua piovana e la risalita dell'umidità. Lo studio del diverso grado di assorbimento idrico di un suolo diviene quindi di primaria utilità per l'individuazione di siti archeologici non emergenti in contesti extraurbani.
Nel presente lavoro verrà presentato un indice del suolo (S.L.I.) per dati iperspettrali MIVIS che mira a costituire un supporto nell'identificazione di tracce sul suolo non vegetato, permettendo l'enfatizzazione dell'umidità o della aridità di una porzione di terreno.

ABSTRACT
The possibilities offered by studying the tonal variations of the bare soil ("damp-mark") are well known 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 of the rising of humidity. The studying of the different degrees of water absorption of a soil becomes therefore particularly useful for the identification of buried archaeological sites in extra-urban context.
In this work will be presented a Soil Index for hyperspectral MIVIS data that aims to constitute a support for the identification of traces over not-vegetated soil, emphasizing the wetness or the dryness of a portion of the ground.

INTRODUCTION
The study of tone variations of bare soil ("damp marks") has a long tradition in archaeological research: their occurrence in fact can be an indication of the presence in the subsoil of archaeological structures or ditches that respectively inhibit or facilitate the absorption of rain water and the rising of humidity. The soil in fact assumes different tonality when it lies over ruins of constructions, roads etc., that is to say in all the situations in which in the past there was a presence or accumulation of materials having a different nature with respect to that of the surrounding soil.
Consequently, the possibility to emphasize the wetness or the dryness of a portion of the ground is potentially very useful for archaeological goals. In order to pursue this goal, the concept of Soil Line (S.L.), related to site-specific soil conditions within a field, has been exploited and a Soil Line Index (S.L.I.) has been produced, providing as results new images where the dry-wet discrimination is accentuated.
The S.L.I. has been tested and applied on hyperspectral MIVIS data, provided from Regione Friuli Venezia Giulia. The Regione Friuli Venezia Giulia holds a consistent collection of hyperspectral remotely sensed images that constitute an important source of data suitable for the archaeological research and the investigation of the development of regional settlement patterns in ancient periods[1]. The coverage of MIVIS hyperspectral images is almost complete from the point of view of

[1] I would 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 for allowing the use the MIVIS images.
For MIVIS images: received authorization from Regione Autonoma Friuli Venezia Giulia, P.M.T./1295/2100, Jan. 25th 2005.

geographic extension, covering a large part of the extension of the region, with shots taken in a period going from 1998 to 2003.

The surface used as a case study includes the urban and suburban areas of Aquileia and the neighboring Communes of Terzo di Aquileia and Fiumicello (UD, Italy). Until few years ago, the archaeological researches have favored the issues related to the urban area. Only recently topographic researches aiming at the reconstruction of the settlement system, of the resource exploitation and of the distributive arrangement of the suburban spaces[2] have been undertaken.

METHODOLOGY

The Soil Line is a well-known linear relationship between the near-infrared and red reflectance[3], a hypothetical line in spectral space that describes the variation in the spectrum of bare soil in the image. The concept that a definable region in a N.I.R-Red scatter plot of the 2-dimensional space (based on the multispectral bands) is occupied by agricultural crops and that a region occupied by pixels recognizable as soil is a thin, lengthened ellipsoid in this 2-dimensional space has been introduced in the 1970s for identifying agricultural crops[4]. In the axis of this ellipsoid, pixels representing soils range from soils of low reflectance to those of high reflectance; the locations occupied by vegetation, soil and water can be seen in three distinct areas of the scatter plot forming a triangle. Later studies have shown some of the limits of the equation defining the Soil Line theorized in the P.V.I. (*Perpendicular Vegetation Index*)[5] and the need of the definition of the Soil Line using empirical data. Starting from this concept and verified that the assumption of the location of the three groups of pixels identified for Landsat data was true also for MIVIS data, the Soil Line Index was determined by identifying a Soil Line directly in the scatter plot of the Red and N.I.R. values and then by measuring the distance of the remaining pixels from a set point along the detected Soil Line, after excluding the pixels representing the vegetation.

Empirical Determination of a MIVIS Soil Line

An empirical method of identification of a S.L., based on the interaction on the scatter plot of the images' values, have been elaborated and tested, in order to investigate the validity of the premises for MIVIS images. The MIVIS images used in this research[6] have been provided at a low level of pre-processing, in only the radiometrically corrected format, as a cost-benefit compromise, and they have been subjected to a noise removal process, obtaining a drastic reduction of the unwanted disturbance due to limitations of the signal digitization and data recording process[7].

The first step of the empirical identification of the Soil Line is the creation of a scatter plot of the radiance measured in the visible Red band against radiance in the N.I.R. in the used image, showing Red band on the x-axis and N.I.R. band on the y-axis[8] (see Figure 1). In this two-dimensional feature space, the distribution of the pixels is determined by the absorption and reflective characteristics of the different physical materials in the image represented by the individual pixels.

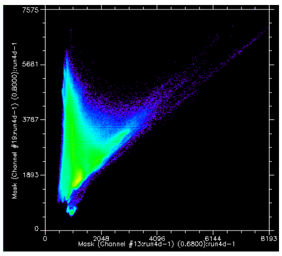

Figure 1 - 2D feature space Red (band 13, x) vs. N.I.R. (band 19, y) scatter plot for a MIVIS run.

---

[2] See Maggi, Oriolo, 1999; Oriolo, 1998.

[3] The first studies treating the topic were Kauth, Thomas, 1976, for Tasseled Cap Transformation and Richardson, Wiengand, 1977 for the P.V.I.; more recently see also Wiegand, Richardson, 1982 and Baret, 1993 and the related, updated, bibliography.

[4] These studies were primarily applied to Landsat MSS; for a short resume about the principles of the various theories of Soil Line and related, updated bibliography see Mather, 2004, pp. 145-148.

[5] See Mather, 2004, p. 146 for a short report.

[6] The MIVIS images used in this research have been captured during October 1998.

[7] See Traviglia, In Press, par. 6.1.

[8] The bands chosen are: band 13 (R.) and band 19 (N.I.R.).

As well known, in a scatter plot of a Landsat MSS image, the pixel locations occupied by vegetation, soil and water can be seen as occupying three pronounced locations forming a triangle (see Figure 2). Bare soil pixels will lie along the line S1-S2 with the level of wetness of the soil being more elevated at the S1 end of the Soil Line than at the S2 end. Pixels representing vegetation lie above and to the left of the Soil Line and the perpendicular distance to the Soil Line is, in the P.V.I. theory, the measure that is correlated with the biomass.

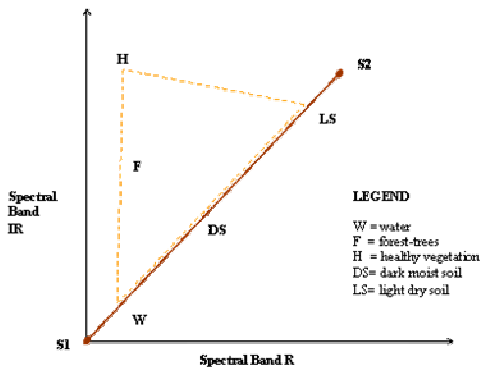

Figure 2 – N.I.R. vs R. 2-D feature space showing the soil, vegetation and water triangle.

In order to verify if the assumption was valid for MIVIS images and further to identify the position of a Soil Line on those, the clusters of pixels and their relation to the physical material were empirically investigated[9]. A number of pixels have been manually selected on the scatter plot and their relative corresponding pixels on the original True Color composite image have been color masked, visually marking their apparent nature. As result, the previously observed linear relationship between Red and Near-infrared reflectance for Landsat data was demonstrated to be true also for MIVIS images. The cluster in the lower right was confirmed to be the pixels for the entire sea water portion of the scene (see Figure 3) and also the other pixels were established to represent the expected landscape features.
The next step was the identification in the plot of the hypothetical line that, in this spectral space, describes the variation in the spectrum of bare soil: the Soil Line.

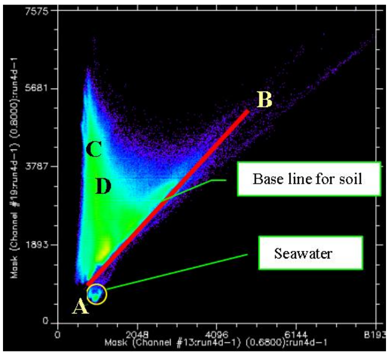

Figure 3 - SoilLine placement and water position in the scatter plot of a MIVIS run. The Soil Line extends from darker soils (A) to an upper region of bright soils (point B). Point C represents a pure vegetation pixel and Point D represents a partially vegetated pixel.

The Soil Line AB (see Figure 3) was determined by drawing a boundary parallel to the trend of the scatter on the lower right side and manually moving this boundary up on the scatter plot in a parallel shift until pixels representing soil began to be masked and colored on the corresponding True Color

---

[9] The software that was used for the processing of MIVIS images is R.S.I. Envi 4.0 ©, that has been provided in temporary license from R.S.I Italy (thanks to Dr. P. Filippi).

image. In order to describe mathematically the line AB, two points that determine the inclination of the line have to be identified. The slope “m” was established mathematically through the application of the Slope Equation formula:

$$m = y2-y1 / x2-x1 \qquad [1]$$

and subsequently by applying the Line Equation formula:

$$y = mx + b \qquad [2]$$

A point (Z) has to be now identified in the scatter plot: it is chosen arbitrarily, identifying the very first and lowest pixel[10] among the ones lying between the Soil Line and the vegetation threshold, and represents the most left pixel of the ones representing wet soil, that is to say the highest humidity of the scatter plot area: it can be taken as the lower left most point on the scatter plot. The distance from the projection of point Z to the projection of any pixel in the scatter plot onto the Soil Line (see Figure 4) can be consider as an indication to the moisture content of the soil, ignoring the differences due to soil type and texture; this distance is referred as Soil Line Index.
One problem that has to be stressed here is that the vegetation results to interfere with the Index, since what is obtained through the application of the Soil Line Index is an image where in the same light color are represented both the dry soil and vigorous vegetation. Consequently, it is necessary to mask it. In order to do that, a vegetation threshold has to be to identify (see Figure 4) to mask out from the image the reflectance due to the vegetation.

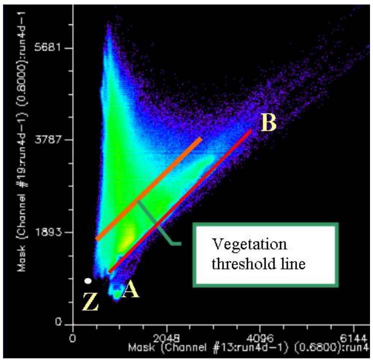

Figure 4 - Vegetation threshold

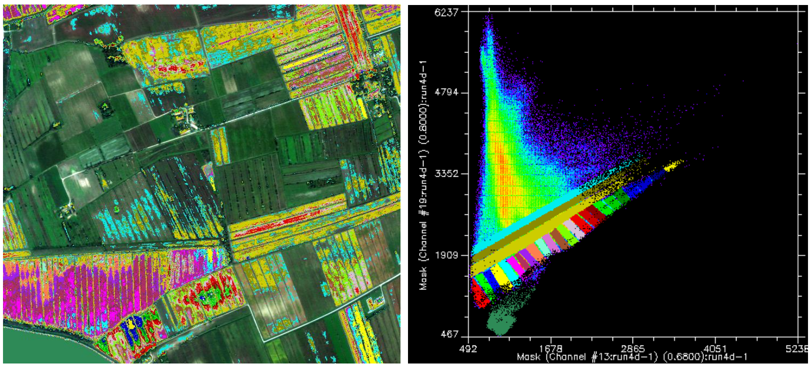

Figure 5 - Verification of the soil moisture variation.

Assuming the theory that the vegetation influence is perpendicular to the Soil Line and that all the iso-vegetation lines remain parallel to the Soil Line[11], so that perpendicular variation from this line indicates increased vegetation cover, the vegetation threshold can be found by manually moving a line parallel to the baseline in an increased N.I.R. reflectance direction, until no more soil pixels in the scatter plot are included in the group of those highlighted in the display of the image comprising vegetation[12]. Every point over the vegetation threshold can now be excluded.
An empirical verification that every pixel between the baseline and the vegetation threshold line refers only to soil or urban structures can be easily performed by classifying in the scatter plot the entire supposed soil zone pixels between the two lines (i.e. Soil Line and vegetation threshold) using small perpendicular rectangular classes to verify the soil moisture variation on the True Color image (see Figure 5).
This empirical classification through visual inspection makes it possible to investigate on the nature of the soils represented by the pixels, revealing that the lower end pixels in the soil zone reference canals and shadows and that the upper end pixels reference dry soil in the target area.

---

[10] Some researches suggest to use the y-intercept of the Soil Line and some even suggest the x-intercept, which could be done mathematically. The process used here gave better distribution of DN values and avoided the need for further image enhancement (i.e. stretching).
[11] See Richardson, Wiegand, 1977, pp. 1543-1547:
P.I.V.= sin (α)N.I.R.-cos (α)Red
where α is the angle between the Soil Line and the N.I.R. axis.
[12] Empirically, the Soil Line must be moved up along the N.I.R. axis and this is easily done by adjusting the y-intercept value.

Soil Line Index

The selection of the point with highest humidity of the scatter plot area (Z), having excluded the seawater lying below the Soil Line, is performed identifying the very first and lowest pixel[13] among the ones lying between the Soil Line and the vegetation threshold. The point happens to correspond to both shadows and water in the canals on the MIVIS image, the later probably because of the interference of the soil spectrum in shallow waters.

The next step consists in measuring the distance of every pixel P along the axis of the Soil Line starting from Z (D_{ZP}). 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 on 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:

$$\gamma = \alpha - \beta \qquad [3]$$

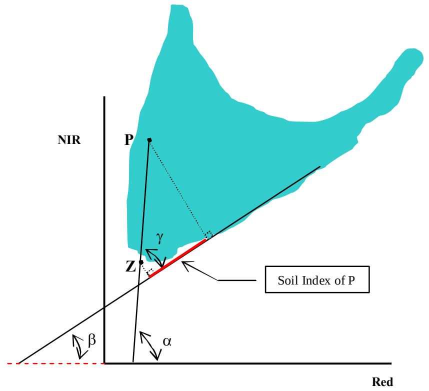

Figure 6 – Graphic interpretation of segment ZP's projection onto the Soil Line.

Since the inclination (m) of the Soil Line is known, the angle β between the Soil Line and the Red axis is determined by the Arctan (tan -1) of this inclination:

$$\beta = \tan^{-1}(m) \qquad [4]$$

Since α is equal to the Arctan of the slope of the segment ZP:

$$\alpha = \tan^{-1}(m\ ZP) \quad \text{where } m_{ZP} = (Py-Zy)/(Px-Zx) \qquad [5]$$

from the equation above:

$$\gamma = \tan^{-1}(Py-Zy)/(Px-Zx) - \tan^{-1}(m) \qquad [6]$$

The Cos of γ is the coefficient of the projection of every point P along the soil axe.

$$\text{Cos } \gamma = \cos[\tan^{-1}(Py-Zy)/(Px-Zx) - \tan^{-1}(m)] \qquad [7]$$

At this point is possible to calculate the Soil Line Index:

$$S.L.I. = D_{ZP} \text{ Cos } (\gamma) \qquad [8]$$

The formula can be applied through band math modules to the MIVIS data and an image representing the soil humidity index is obtained, where light colors represent the dryness of the soil and dark colors the humidity.

RESULTS

The application of the S.L.I. to some sample images has shown a clear improvement in the differentiation of the typologies of the soils, accentuating the dry-wet discrimination. The S.L.I. have been tested in some sample areas of the images where previously has been verified through other processes the presence of line-shaped traces coherent with the settlement planning of the Aquileia area, in particular in the west suburban area. The use of the process provided very good results in terms of increasing of the visibility both of dry-soil lines and of humid areas. Dry-soil

---

[13] Some researches suggest to use the y-intercept of the Soil Line and some even suggest the x-intercept, which could be done mathematically. The process used here gave better distribution of DN values and avoided the need for further image enhancement (i.e. stretching).

lines after the application of the S.L.I. were more visible because of the increased contrast with the surrounding soils; wet areas were more easily identifiable in their shape and the application of S.L.I. permitted also, in some cases, to define the presence of wet-soil lines that were not previously visible, confused among highly moist spots, by determining sharper contours. The possibility to mask the vegetation furthermore allows to not confusing it with relevant anomalies that could be erroneously interpreted as having an archaeological nature, this due to the not elevated resolution of the MIVIS sensor (3m by 3m).
From an archaeological point of view, this image process results in the possibility to identify subsoil structures and ditches having a filling different from the contiguous soils, represented in the images as dry/wet soil traces. Accordingly with the presence of subsoil structures or underground ditches or canalizations in fact the observable traces end respectively in light color traces or darker color traces. The S.L.I. application was proved to emphasize the phenomena. In one of the test areas for example (see Figure 7), the S.L.I. allowed to identify the presence of a net of dry/wet soil lines, orthogonal among them and organic with the Aquileia road system, whose interpretation is still in progress.

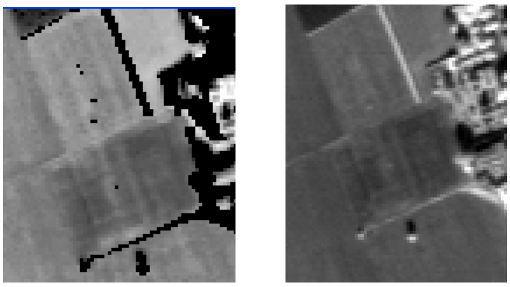

Figure 72 – A detail showing the result of the application of a Soil Line Index (S.L.I.): a better discrimination of the soils can be noticed when applying the S.L.I. (left) compared to the original Red band (right). (S.L.I. with vegetation masking filter –black areas-).

In conclusion, the encouraging results reached in the realized multiple tests invite to precede in this direction, demonstrating that the application of a S.L.I. align with the goals of the archaeological research.

REFERENCES

Baret F. et al. (1993), "The Soil Line concept in remote sensing", *Remote Sensing Reviews*, 7: 65-82

Maggi P., Oriolo F. (1999), "Dati d'archivio e prospezioni di superficie: nuove prospettive di ricerca per il territorio suburbano di Aquileia", *AAAd*, 45: 99-123

Kauth R.J., Thomas G.S. (1976), "The Tasseled Cap. A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat", *Proceeding of the Symposium on Machine Processing of Remotely Sensed Data*, West Lafayette – Indiana, 41-51

Mather P.M. (2004), *Computer Processing of Remotely-Sensed Images: an introduction*, Chichester

Oriolo F. (1998), "Nuovi dati sulle vie sepolcrali di Aquileia", *Xantener Berichte*, 7: 177-180

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

Traviglia A. In Press, "Integration of MIVIS hyperspectral remotely sensed data and Geographical Information Systems to study ancient landscapes: the Aquileia case study", *Agri Centuriati*, 2

Wiegand C.L., Richardson A.J. 1982, "Comparisons among a new soil index and other two and four dimensional vegetation indices", *Proceedings of the 48th annual meeting of the American Society of Photogrammetry*, 210-227.