Potential of hyperspectral remote sensing for field scale soil mapping and precision agriculture applications

Raffaele Casa,[1] Fabio Castaldi,[1] Simone Pascucci,[2] Stefano Pignatti[2]

[1]Dipartimento di Scienze e Tecnologie per l'Agricoltura, le Foreste, la Natura e l'Energia (DAFNE), Università della Tuscia, Viterbo;

[2]Istituto di Metodologie per l'Analisi Ambientale, Consiglio Nazionale delle Ricerche (CNR IMAA), Tito Scalo (Potenza), Italy

Abstract

Mapping within-field variation in soil properties opens up the possibility of employing variable agronomic management and precision farming technologies with potential environmental and economic benefits. However, the excessive cost of systematic direct soil sampling severely constrains the practical feasibility of site specific management based on soil variability information. Remote sensing offers a cost effective and efficient means for gathering a great deal of information on soil properties. The aim of the present work was to assess the potential of satellite hyperspectral imagery for the mapping of soil properties in the tilled layer of agricultural fields, in the context of precision agriculture applications. CHRIS-PROBA satellite images were acquired over two bare soil fields and their capability to provide estimates of soil texture and soil organic matter (SOM) at the field scale was assessed. Partial least squares regression (PLSR) models were developed on datasets spatially independent from those used for validation. Clay and sand could be estimated with intermediate accuracy, with values of RPD (ratio of performance to deviation) higher than 1.4. Root mean squared error (RMSE) values of 3.7 and 5.2 were obtained for clay in the two fields respectively. SOM estimates were not satisfactory, probably because of the limited range of spatial variation in the studied fields.

Maps of uniform soil zones were obtained from measured and estimates soil texture data by means of fuzzy c-means classification. The resulting maps were then used for the parameterization of a simple water balance model, i.e. CropWat8.0, in order to simulate and compare uniform and variable-rate irrigation strategies. Simulation results suggest that site-specific irrigation allows to reduce significantly water losses by deep percolation, which occur when irrigation scheduling and volumes are calculated on the basis of average field soil properties. The present paper demonstrates the usefulness of satellite hyperspectral data for mapping soil spatial variability at the field scale, providing useful information for precision agriculture applications.

Introduction

Information on the variability of soil properties within agricultural fields is of paramount importance for the purpose of implementing site-specific agronomic management strategies. Site-specific farming practices take into account such variability and allow a more efficient use of resources such as water and fertilizers (Gebbers and Adamchuk, 2010). However, practical solutions for rapid and low-cost mapping of soil properties are still lacking. The availability of detailed information on soil properties at the field scale is generally insufficient for the application of site-specific techniques, being currently a limiting factor for their implementation, given the excessive cost of spatially dense soil sampling and analysis. Consequently, there is great interest in the development of low cost and rapid soil mapping methods, and in this context proximal soil sensing, based on the measurement of geo-electrical or spectral soil properties, seems to offer good opportunities (Viscarra Rossel et al., 2011). Remote sensing from airborne and especially from space-borne platforms, would be in theory even more appealing than proximal soil sensing, due to its lower cost and the possibility of avoiding the use of specific field equipment. Hyperspectral remote sensing in particular, has potential in this sense (Ben-Dor et al., 2009), though there are several problems limiting this kind of application: small range of variation of within-field soil properties, high spatial resolution required, difficulty to relate sensed soil surface properties to variable of agronomic interest throughout the tilled layer.

There are very few examples of studies for this kind of applications using satellites with hyperspectral capabilities, in particular employing

CHRIS-PROBA (European Space Agency's pushbroom Compact High Resolution Imaging Spectrometer). In the next few years, the launch of satellites carrying high spatial resolution hyperspectral sensors, such as the German EnMAP (Kaufman *et al.*, 2006) or the Italian PRISMA (Labate *et al.*, 2009), could open up new possibilities for the remote estimation of soil properties.

The objectives of this study was that of assessing the potential of CHRIS, and in perspective of forthcoming sensors, for soil properties estimation and of evaluating their potential usefulness for precision agriculture applications though an example concerning variable irrigation strategies.

Materials and Methods

Satellite hyperspectral images were acquired by CHRIS-PROBA, over two bare soil fields in Maccarese (Roma, Italy) on 27 April 2007 and 25 May 2010. CHRIS was configured in mode 5, including 37 bands in the visible and near infrared range (442-1019 nm) with a pixel size of 17 m. CHRIS data were pre-processed using the CHRIS-Box software (VISAT 4.7). Vertical striping was reduced using the algorithm of Gómez-Chova *et al.* (2008). Atmospheric correction was performed using the FLAASH module, implemented in ENVI 4.7 (ITT VIS, Boulder, CO, USA). The reflectance values obtained were further geocoded in UTM ED 1950 mean projection and corrected using 30 Ground Control Points for each field, extracted from Latium Regional Technical Map at a scale of 1:10,000. Further refinement included the application of a local warping procedure (nearest neighbor resampling) using an ortho-photo (0.5 m resolution) of the selected area as reference (Dai and Khorram, 1998). A mean RMS of 0.61 pixels was achieved with respect to the ortho-photo.

In the days immediately following the satellite data acquisition, a systematic soil sampling along an approximate grid of 40 m was carried out. Samples were collected in the 0-30 cm soil layer in a 11 ha section of a field coded as B064 and in two sections, respectively of 15 and 12 ha of a field coded as B071. Samples were dried and sieved at 2 mm. Particle size analysis and soil organic matter (SOM) analysis were carried out for all the 97 samples of field B064 and the 117 samples of field B071. The data were then interpolated by block kriging, using stable or spherical variogram models, to the same spatial resolution and grid of the CHRIS pixels. Therefore, the reflectance spectrum of each pixel in the CHRIS image was associated to the corresponding soil properties.

The spectral behaviour of a set of 146 soil samples at a range of soil moisture contents was studied in the lab using the Analytical Spectral Devices (ASD) Field Spec Fr Pro 104 spectroradiometer, which measures radiance in the 350-2500 nm spectral range, converted to reflectance using a calibrated Spectralon panel. The soil samples were placed in Petri dishes over a black surface and the spectra were acquired in the dark using the ASD contact (plant) probe with an internal light source at a distance of a few mm from the soil. Spectroscopic measurements were carried out at different moisture contents by first leaving samples to equilibrate on wet sand until saturation, then spectra were collected, the samples were weighted and subsequently placed in an oven at 80°C for about 30 min before acquiring again other spectra. This sequence was repeated until the samples were completely dry.

The objective of such tests was to verify wether the variation of soil moisture was a confounding factor for the determination of soil texture and organic matter. For that reason, preliminary tests were performed by carrying out estimates separately for three groups of dry, wet and intermediate moisture samples as well for mixed moisture sets, *i.e.* including both dry and moist samples. Since soil moisture measurements had not been carried out at the times of satellite acquisitions, it was not possible to carry out tests using the same soil moisture level

corresponding to CHRIS-PROBA acquired spectra. In all trials, better results were always obtained when using mixed moisture sets. Therefore, in all subsequent tests, we used data sets containing spectra at different soil moistures. Lab spectra were split in two randomly extracted sets of two thirds for calibration and one third for validation, 51 and 26 for field B064 and 46 and 23 for field B071, respectively.

All spectra were converted into absorbance units as log(1/R), where R is the reflectance value, since preliminary trials showed that using absorbance enabled to reduce the number of components in the PLSR (see following). Smoothing was applied to lab spectra using a Savitzky-Golay filter (span=15, degree=2). In addition, the latter spectra were also resampled to the same wavebands of the CHRIS sensor, in order to investigate the effect of reducing spectral range and resolution.

Partial least squared regression (PLSR) was employed in order to develop statistical models for the estimation of soil properties for lab and satellite data sets. This multivariate statistical method allows to reduce the number of predictors, providing the optimal linear model through a restricted number of factors (Viscarra Rossel, 2008; Wold *et al.*, 1983). Leave one-out cross validation was used to determine the number of factors to retain, using calibration data sets, by computing the relative root mean squared error of prediction (RRMSEp). In this procedure, the variation of RRMSEp values in response to an increasing number of factors, allows to identify the lowest number of factors for which the model performs well, avoiding over-fitting, since choosing too many factors will provide models performing well only on the data on which they were calibrated (Viscarra Rossel, 2008). Once determined the optimal number of factors from the leave one-out cross validation, the calibration models were then applied to independent validation sets. CHRIS image calibration data were obtained from a central strip of the field B064 (147 pixels, *i.e.* spectra) and validation data from lateral parts of the sampled area (232 pixels), whereas for field B071 the 15 ha NW section of the field (549 pixels) was used for calibration and the 12 ha SE section (389 pixels) for validation. The performance of the estimations was assessed by computing the root mean squared error (RMSE) and the ratio of performance to deviation (RPD) (Chang and Laird, 2002). It is assumed that models having a RPD>2 can accurately predict, models having an RPD between 2 and 1.4 are intermediate and models having RPD<1.4 have no prediction ability (Chang and Laird, 2002; Gomez *et al.*, 2008).

In order to assess the potential usefulness of the information on within-field soil texture variability, simulations of different irrigation scenarios were carried out by applying a daily soil moisture balance and irrigation scheduling model to 28 years of weather data from Maccarese. The model used was CropWat 8.0 (FAO, 2009). The soil parametrization was based on the use of uniform soil texture zones from either block kriging or PLSR estimates. For this purpose, soil data were clustered into classes by means of an unsupervised fuzzy c-means classification, implemented through the Management Zone Analyst (MZA) software (Fridgen *et al.*, 2004), allowing the definition of an optimal number of soil classes on the basis of minima of the fuzzyness performance index and the normalized classification entropy. These classes were assumed to correspond to uniform soil zones. The classification was performed using soil data both from block kriging and from PLSR estimates obtaining 6 classes for each field for the two data sources. The average soil properties of each soil class, plus the field averages, were used to calculate the soil parameter requested by CropWat, by means of the pedotransfer functions (Saxton and Rawls, 2006), assuming their homogeneity along the profile. Simulations using the CropWat model were used to assess i) the potential usefulness of a spatially variable irrigation strategy, and ii) the consequences of the use of imperfect knowledge of soil properties, such as that derived from CHRIS PLSR estimates, on the management of spatially variable irrigation.

Three irrigation scenarios were compared: i) uniform irrigation

(UNI), in which an optimal irrigation schedule, corresponding to a decision to irrigate when the readily available soil moisture was depleted, refilling the soil to field capacity at each irrigation, was calculated on the basis of field average soil properties, but applied to the different soils of the block kriging classes; ii) variable rate irrigation based on measured soil properties (VRIs), in which the optimal irrigation schedule was calculated on the basis of the soil properties of the block kriging classes and applied these same classes; iii) variable rate irrigation based on soil properties estimated from PLSR (VRI-PLSR), in which the optimal irrigation schedule was calculated on the basis of the soil properties of the PLSR classes and applied the classes obtained from block kriging of soil measurements. It should be noted that the UNI scenario represents the practice adopted when only average field soil properties are known, VRIs represents the practice adopted when soil properties variability across the fields is accurately known and the VRI-PLSR scenario simulates the situation in which imperfect data on soil properties are available, such as from from CHRIS hyperspectral data estimates. For scenario VRIs, in which variable rate irrigation (VRI) is applied using daily soil water balance on the basis of soil properties assumed to be correct, the potential yield is always attained and no irrigation loss by deep percolation occurs. Conversely, when the irrigation schedule is defined on the basis of an *average soil* for the whole field, but applied to the different soil classes identified from MZA clustering of block kriging data (*i.e.* the UNI scenario), both irrigation losses and reductions from the potential yield can occur.

The seasonal totals for the whole fields of irrigation amounts, deep percolation irrigation losses and potential yield reductions were calculated, for each scenario for the 28 years of data, by taking into account the area covered by each soil class. CropWat calculates yield reduction due to soil moisture stress as a percentage of the maximum production achievable in the area under optimal conditions using the FAO33 approach (Doorenbos and Kassam, 1979).

Results

The PLSR carried out using full range ASD spectra (400 to 2500 nm) allowed to develop models with a good prediction ability for field B064, for the estimation of the percentage of clay, sand (RMSE=2.78, RPD=3.02, with 9 components) and silt (RMSE= 1.07, RPD=3.97, with 16 components). The percentage of variance explained by the PLSR components minimizing the RRMSEp in the leave one-out cross validation was 93% for clay, 94% for sand and 99% for silt. Conversely, SOM was estimated with a RMSE of 0.13 and a RPD of 1.07 by the PLSR with 14 components. The poor performance for organic matter estimation is probably due to the fact that low values of SOM (<3.5%) have been found to have only a minimal effect on spectral response (Gomez *et al.,* 2008). In addition, the range of measured SOM values was rather narrow, spanning from 1.76% to 2.24%, wich was probably insufficient for the development of an adequately calibrated model.

Similar good estimation results for soil texture using full spectra were obtained for field B071. Clay and sand were estimated from 5 factors PLSR, obtaining good RPD values, respectively of 2.1 and 2.2 and RMSE values of 4.9 and 5.5. Silt showed intermediate performances, with an RPD of 1.5 and RMSE of 2.8, whereas models developed for soil organic matter had poor performances, with RPD values always below 1.4. The use of lab data with a reduced spectral range and resolution, corresponding to CHRIS wavebands, had an impact on the performance of the PLSR models developed. Nevertheless prediction ability was still good, with a 12 components PLSR which explained 92% of the variance for clay (RMSE=2.26, RPD=2.12), sand (RMSE=2.90, RPD=2.89) with 8 components explaining 95% of the variance and silt (RMSE=0.94, RPD=4.50) with 8 components explaining 93% of the variance. SOM

estimation was slightly worse than with full range spectra (RMSE= 1.10, RPD= 0.41 with 8 components which account for 60% of the variance). For field B071 lab data resampled to CHRIS spectral range and resolution allowed to obtain models with 6 factors with intermediate prediction ability for clay (RPD=1.6 and RMSE=6.1) and sand (RPD=1.6 and RMSE=7.6), whereas unsatisfactory results were obtained for silt (RPD=1.2 and RMSE=3.5) and soil organic matter (RPD<1.4). The use of real satellite data, employing spectra extracted from the CHRIS image pixels, still allowed the estimation of clay for field B064, though with a much poorer performance as compared to lab spectra (Figure 1). The cumulative variance explained by the 6 PLSR components was 67% and the RPD value indicated that the PLSR model was just above the threshold for being considered to have any prediction ability, whereas other soil properties were below this value. On the other hand, for all other properties, the RPD values were below 1.4 (*data not shown*). For field B071, models with intermediate prediction ability were obtained for clay and sand (Figure 2), though neither for silt nor for SOM. The maps of properties estimated from PLSR having RPD>1.4 showed a good match with the maps of kriged clay measurements (Figure 3). On the basis of these results, clay and sand estimates were retained, whereas silt was estimated as the complement to 100 and the mean soil organic matter was assumed to be constant for the whole fields. These estimates, as well as the block kriging data obtained from soil measurements, were employed for the delineation of uniform soil texture zones using the MZA software (Fridgen *et al.,* 2004). For each case, 6 classes were obtained (Figure 4). It should be noted that classes obtained from block kriging are assumed to represent the true soil types, whereas those from PLSR include uncertainties and errors derived from the estimation procedure.

CropWat simulation revealed that when using the uniform irrigation strategy (UNI), losses due to percolation were on the average 79 mm for field B071 and 15.4 mm for B071. For both fields, B064 and B071, the total irrigation amounts applied throughout the season were not significantly different among the three scenarios (Figures 5a and 6a). However, differences in the irrigation lost by percolation where statistically significant among irrigation strategies, though largely differing between fields. In the field B064, when the UNI strategy is adopted, only 4 mm more irrigation water is lost on average as compared to VRI-

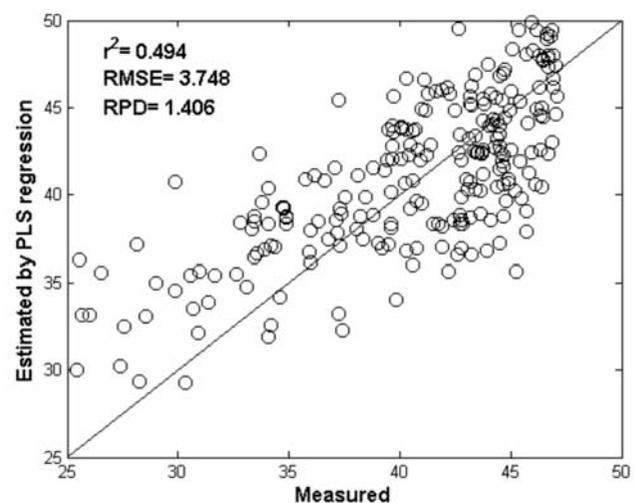r^2= 0.494
RMSE= 3.748
RPD= 1.406

Figure 1. Measured vs estimated clay values resulting from the application of PLSR to the validation set for field B064 using spectral reflectance data extracted from the CHRIS- PROBA image.

PLSRB (Figure 6b), though this difference is statistically significant for a paired means *t*-test (P<0.01). In field B071 the adoption of the UNI strategy leads on the average to the loss of 62 mm more than the VRI-PLSR strategy (P<0.001) (Figure 5b). Yield reductions were on the average 8.1% and 6.5%, respectively, when UNI and VRI-PLSR were applied to field B071 (Figure 5c). On field B064 yield reduction differences between UNI and VRI-PLSR were larger, *i.e.* they amounted to 7.6% and 1.1%, respectively (Figure 6c).

Discussion

This work represents a first step towards the development of a quantitative and comparative assessment of the potential of hypespectral satellite data in the context of soil mapping for precision agriculture applications. Several efforts have been made in the last decades to use satellite remote sensing for the estimation of soil properties (Ben-Dor *et al.,* 2009; Ge *et al*., 2011; Mulder *et al*., 2011), but until recently, most attempts were performed using low spectral and/or spatial resolution

Figure 3. Comparison of maps obtained from the application of PLSR to CHRIS Proba validation data sets (left) and from application of block kriging to soil sampling measurements (right). A) field B071 clay. B) field B071 sand. C) field B064 clay.
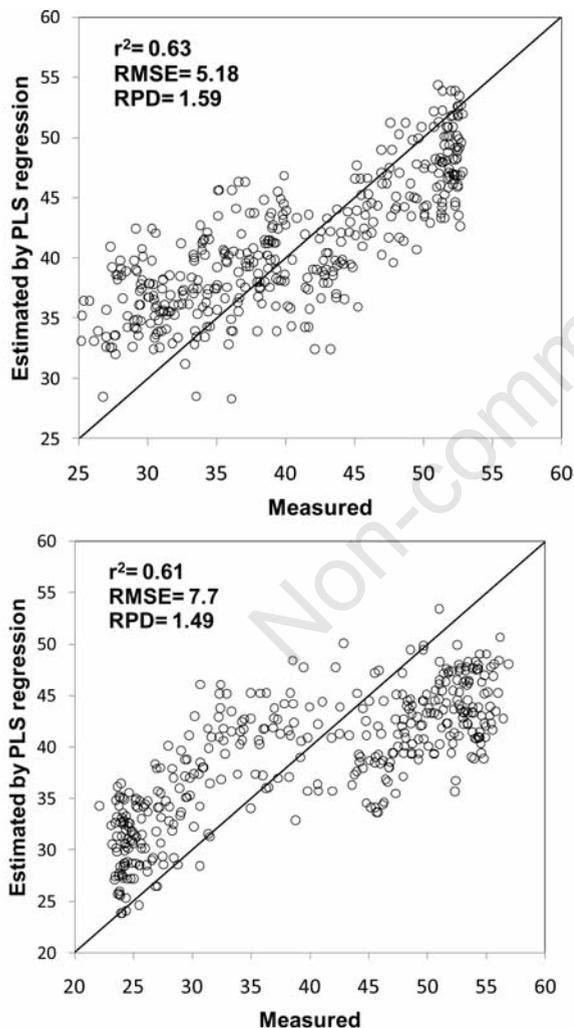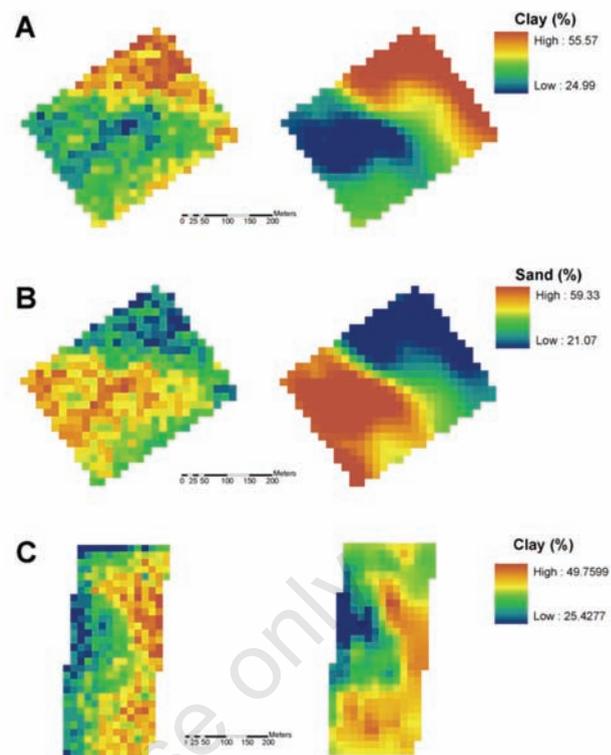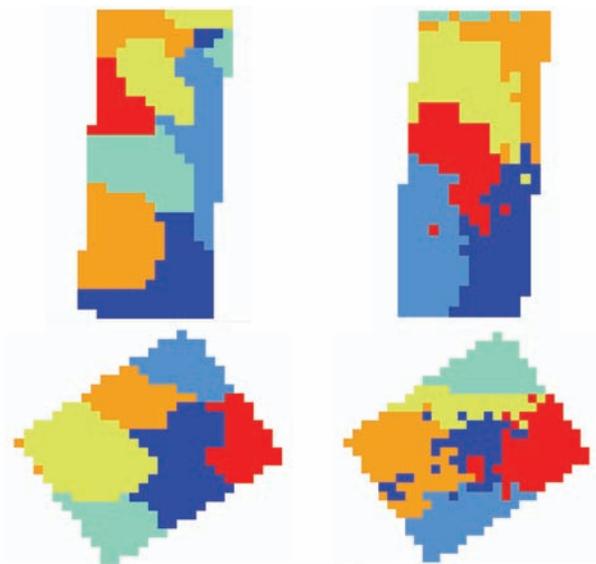

Figure 2. Measured *vs* estimated clay (top) and sand (bottom) values resulting from the application of PLSR to the validation set for field B071 using spectral reflectance data extracted from the CHRIS-PROBA image.

Figure 4. Soil classes resulting from fuzzy c-means classification carried out from block kriging soil data (left) and form PLSR results (right). Top: field B064, bottom: field B071.

data, allowing prevailingly qualitative determinations at the regional scale. For more quantitative applications at the field scale, high spectral and spatial resolution data are necessary (Ben-Dor *et al.*, 2009; Mulder *et al.*, 2011). The availability of fuller spectral signatures from hyperspectral sensors and the inclusion of short wave infrared (SWIR) bands and the improved radiometric data quality, are expected to boost estimation capabilities. Multivariate calibration and chemometric methods, widely used for laboratory and field applications can be employed with hyperspectral data for the estimation of several soil properties (Viscarra Rossel, 2008; Ge *et al.*, 2011).

Our results - obtained with data acquired in the lab - seem to suggest that despite the reduced spectral range of the CHRIS instrument, which cuts off features known to be significantly related to soil properties such as clay and SOM (Ben-Dor *et al.*, 2009), it is still possible to estimate soil texture variables quantitatively. The use of CHRIS satellite data involves the inclusion of errors related to radiometric, geometric and atmospheric effects as well as scaling issues related to pixel size and kriging of ground truth soil properties. It is therefore not surprising that PLSR models developed using real CHRIS data have a much poorer performance as compared to lab data with CHRIS spectral range and resolution. Soil properties estimation is also possibly affected by the confounding effect of soil moisture. Nevertheless, despite these shortcomings, models with intermediate prediction ability could be obtained in field B064 for clay and in field B071 for clay and sand.

The potential value of this information for precision agriculture application was assessed through a very simple example, running simulations of different irrigation strategies. The model employed for such an assessment, CropWat8.0, was chosen because of its extreme semplicity, thus requiring a minimum number of input parameters. Much more complex models could have been used allowing more sophisticat-
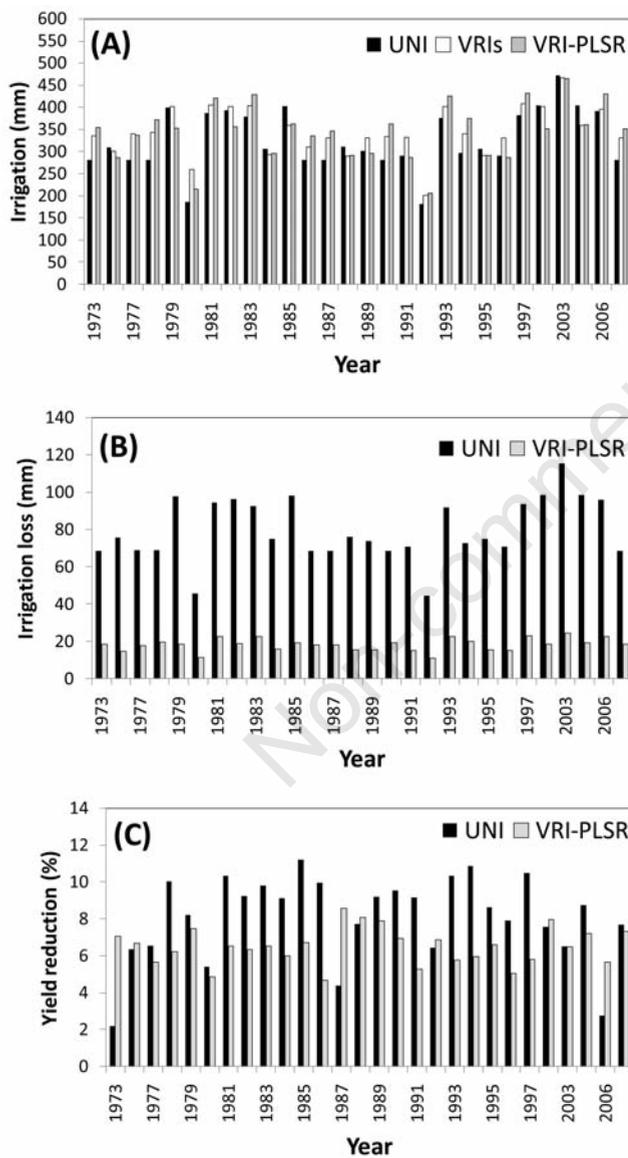

Figure 5. Simulation results for the comparison of irrigation strategies for field B071. A) total seasonal irrigation. B) irrigation losses due to percolation. C) reduction from the potential yield.
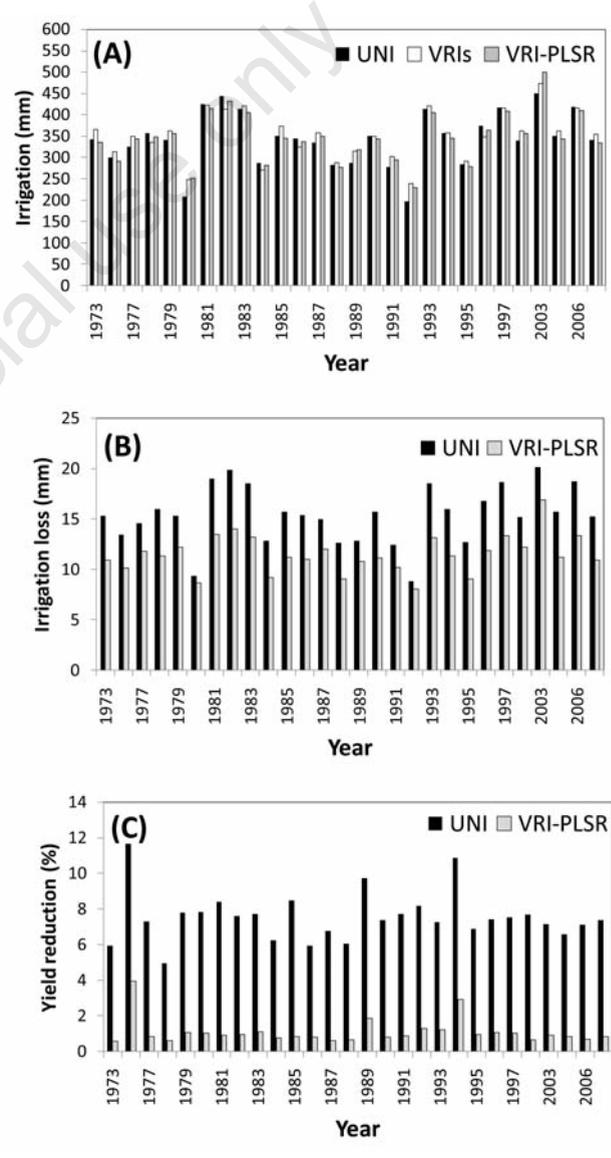
Figure 6. Simulation results for the comparison of irrigation strategies for field B064. A) total seasonal irrigation. B) irrigation losses due to percolation. C) reduction from the potential yield.

ed analyses, though their parametrisation could have posed more problems. In addition, it should be noted that some rather unrealistic assumptions were adopted for this assessement. These included: i) the uniformity of soil properties troughout the whole soil profile; ii) the goodness of the pedotransfer functions used to derive soil parameters from texture and SOM; iii) the goodness of the block kringing estimates; iv) the goodness of the MZA clustering.

Despite these limitations, the results of this analysis indicate a potential advantage in the management of irrigation by taking into account soil spatial variability. Indeed, the tests carried out using CropWat highlight the potential advantages resulting from the adoption of a VRI strategy, even when the irrigation scheduling is based on imperfect information such as that deriving from the CHRIS Proba PLSR results. The use of a variable irrigation strategy did not lead to the use of less water, but rather to a more efficient use, decreasing losses due to excessive irrigation rates and reducing water stress leading to yield reduction. Quite interesting differences were found between the two fields of the experiment. Differences in irrigation losses between the UNI and the VRI-PLSR strategies were much higher for field B071 than for field B064. This can be explained by the larger spatial variability in soil texture found in this field as compared to B064. The coefficients of variation for clay, silt and sand were 12%, 11% and 20% for field B064 and 21%, 29% and 21% for B071, respectively. This highlights the fact that the potential advantage of adopting a site-specific irrigation strategy depends on the soil spatial variability of the field. In conclusion, the value of the knowledge of field scale soil texture variability for precision agriculture applications suggests to consider satellite hyperspectral data, which will be increasingly available in the future, as an additional information source leading to a possible synergic use with other information, such as that coming from geo-electric and proximal soil sensing tools (Viscarra Rossell *et al.,* 2011).

References

Ben Dor E, Chabrillat S, Dematté JAM, Taylor GR, Hill J, Whiting ML, Sommer S, 2009. Using imaging spectroscopy to study soil properties. Remote Sens. Environ. 113:S38-S55.

Berk A, Bernstein LS, Anderson GP, Acharya PK, Robertson DC, Chetwynd JH, Adler-Golden SM, 1998. MODTRAN Cloud and Multiple Scattering Upgrades with Application to AVIRIS. Remote Sens. Environ. 65:367-375.

Chang CW, Laird DA, 2002. Near-infrared reflectance spectroscopic analysis of soil C and N. Soil Sci. 167:110-116.

Dai X, Khorram S, 1998. The effects of image mis-registration on the accuracy of remotely sensed change detection. IEEE Trans. Geosci. Remote Sens. 36:1566-1577.

Doorenbos J, Kassam AH, 1979. Yield response to water. FAO Irrigation and Drainage Paper No. 33. FAO Publ., Roma, Italy.

FAO, 2009. CROPWAT 8.0 for Windows. Water Resources Development and Management Service. Land and Water Development Division. Available from: http://www.fao.org/nr/water/infores_databases_cropwat.html

Fridgen JJ, Kitchen NR, Sudduth KA, Drummond ST, Wiebold WJ, Fraisse CW, 2004. Management Zone Analyst (MZA): Software for Subfield Management Zone Delineation. Agron. J. 96:100-108.

Ge Y, Thomasson A, Sui R, 2011. Remote sensing of soil properties in precision agriculture: a review. Front. Earth Sci. 5:229-238.

Gebbers R, Adamchuk VI, 2010. Precision agriculture and food security. Science 327:828-831.

Gómez-Chova L, Alonso L, Guanter L, Camps-Valls G, Calpe J, Moreno J, 2008. Correction of systematic spatial noise in push-broom hyperspectral sensors: application to CHRIS/PROBA images. Appl. Opt. 47: F46-F60.

Gomez C, Viscarra Rossel RA, McBratney AB, 2008. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: an Australian case study. Geoderma 146:403-411.

Kaufmann H, Segl K, Chabrillat S, Hofer S, Stuffier T, Mueller A, Richter R, Schreier G, Haydn R, Bach H, 2006. EnMAP - A hyperspectral sensor for environmental mapping and analysis. pp 1617-1619 in Proc. Int. Geoscience and Remote Sensing Symp. (IGARSS), Denver, CO, USA.

Labate, D., Ceccherini, M., Cisbani, A., Decosmo, V., Galeazzi, C., Giunti, L., Melozzi, M., Pieracini, S., Stagi, M., 2009. The PRISMA payload optomechanical design, a high performance instrument for a new hyperspectral mission. Acta Astronaut. 65:1429-1436.

Mulder VL, de Bruin S, Schaepman ME, Mayr TR, 2011. The use of remote sensing in soil and terrain mapping - A review. Geoderma 162:1-19.

Saxton KE, Rawls WJ, 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Sci. Am. J. 70:1569-1578.

Viscarra Rossel RA. 2008. ParLeS: software for chemometric analysis of spectroscopic data. Chemometr. Intell. Lab. 90:72-83.

Viscarra Rossel RA, Adamchuk VI, Sudduth KA, McKenzie NJ, Lobsey C, 2011. Proximal Soil Sensing. An effective approach for soil measurements in space and time. Adv. Agron. 113:237-282.

Wold S, Martens H, Wold H, 1983. The multivariate calibration problem in chemistry solved by the PLS method. In: A. Ruhe and B. Kågström (eds.), Proc. Conf. on Matrix Pencils. Lect. Notes Math. 973:286-293.