Analysis of GAC NDVI Data for Cropland Identification and Yield Forecasting in Mediterranean African Countries

Fabio Maselli and Felix Rembold

Abstract

The utilization of NOAA-AVHRR NDVI data for crop yield forecasting is of particular importance in semiarid regions where there are strong inter-year yield fluctuations due to meteorological vagaries. The present work deals with the use of monthly GAC NDVI data for the early estimation of cereal crop yield in Mediterranean African countries. A preliminary analysis showed that relatively high correlations were present between crop yield and mean NDVI values of specific months computed at national levels. The stratification of the countries according to the USGS global land-cover map brought only marginal correlation increases. Greater improvements were instead reached by a statistical method which allows the estimation of the per-pixel fractions of agricultural and non-agricultural vegetation. When compared to available independent maps, the areas identified in this way were confirmed to be mainly covered by crop and forest land, respectively. The methodology for cropland identification and yield forecasting was finally evaluated for operational applications.

Introduction

The necessity for timely and accurate crop monitoring systems working at a regional scale is particularly felt in arid and semi-arid countries, where temporal and geographical rainfall variability leads to high inter-year fluctuations in primary production and increases the risk of severe famines (Hutchinson, 1991). These environmental situations, along with the wide extent of the areas to monitor and the generally poor availability of efficient communications, enhance the utility of remote sensing techniques, which can view large land surfaces synoptically and with high temporal frequency (Prince, 1990).

Among the available remote sensing systems, the Advanced Very High Resolution Radiometer (AVHRR) mounted aboard the National and Oceanic Atmospheric Administration (NOAA) satellites is presently the most efficient for operational large-scale crop monitoring (Prince, 1991). The NOAA-AVHRR Normalized Difference Vegetation Index (NDVI) is in fact linked to several vegetation properties (percentage cover, leaf area index, active green biomass) and is an indirect indicator of primary productivity through its quasi-linear relation with the percentage of absorbed photosynthetically active radiation (%APAR) (Prince, 1990; Los, 1998). Such a relationship makes possible the early estimation of crop yield, because this parameter is mainly determined by the photosynthetic activity of agricultural plants in certain periods prior to harvest (Baret and Olioso, 1989; Rossini and Benedetti, 1993). Previous studies built on this property found useful statistical relationships between NDVI values at the peak of the growing season and final crop yield once the perturbing effects of the geographical variability in environmental features (natural vegetation and soil types and conditions, topography, etc.) had been removed (Maselli *et al.*, 1993; Hayes and Decker, 1996). The same and other investigations showed that yield forecasting can be obtained by the use of NDVI data of specific periods which depend on the eco-climatic conditions of the areas and the types of crop grown (Hayes and Decker, 1996; Lewis *et al.*, 1998; Maselli *et al.*, 2000). Most recent research also indicated that a decisive improvement in crop yield forecasting capability is linked to the selective consideration of NDVI values from cropped areas, because other vegetation types, having different seasonal developments, may introduce noise in the relationships NDVI/yield (Genovese *et al.*, 1999; Maselli *et al.*, 2000). This should be particularly the case in Mediterranean countries, where areas cultivated with cereal crops are generally intermingled with other vegetation types, such as forests and *maquis*, which have a different phenology and response to meteorological factors (Lacaze *et al.*, 1996).

In the current work, carried out within the EU (European Union) project CAMELEO (Changes in Mediterranean Semi-arid Ecosystems on the Long Term through Earth Observation), the use of multitemporal NDVI data was tested for the estimation and forecasting of cereal crop yield in Mediterranean African countries. In particular, the first efforts were devoted to finding which was the most suitable period for yield forecasting by NDVI data in each country. The research was then directed to the identification of cropped areas which could be selectively considered for improving the forecasting capability. This was attempted first by focusing on different land-cover types derived from an existing supervised classification. Next, a new statistical method based on correlation analysis was applied to identify single pixels specifically containing agricultural and non-agricultural vegetation. The results obtained using the two methods were evaluated in terms of their correlation with final crop yield. Also, the areas identified as crop and forest land were compared to existing cartography. The utility of the optimal methodology found was finally evaluated for operational crop yield forecasting.

Study Area

Environmental Features

The study area corresponds to the four North African countries involved in the CAMELEO project: Morocco, Algeria, Tunisia,

IATA-CNR, P.le delle Cascine 18, 50144 Firenze, Italy (maselli@iata.fi.cnr.it).

Photogrammetric Engineering & Remote Sensing
Vol. 67, No. 5, May 2001, pp. 593–602.

0099-1112/01/6705–593$3.00/0

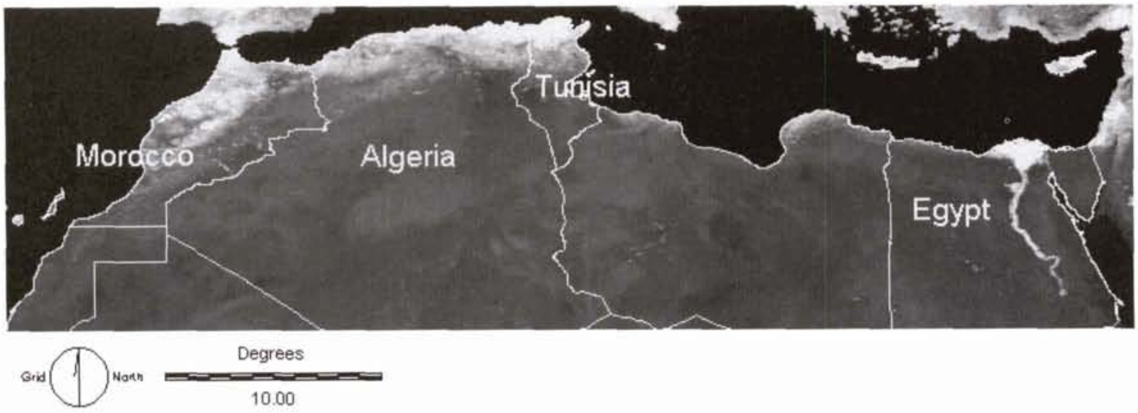

Figure 1. Example of an NDVI MVC image (April 1992) showing the four North African countries involved in CAMELEO (Morocco, Algeria, Tunisia, and Egypt).

and Egypt (Figure 1). With the exclusion of the completely desert zones without seasonal changes, nearly the whole area is part of the arid and hyper arid ecosystems of the Mediterranean bioclimatic region, according to the classification of Emberger (Emberger, 1955). In all four countries, the Mediterranean climate is typical for a large east-west belt bounded by the Mediterranean Sea on the north and by the 125-mm (annual) isohyete on the south. The north-south extension of this belt varies between 30 and 300 km. The Mediterranean area includes the Atlas Mountains in Morocco and Tunisia and the coast massifs in Algeria, which form an effective barrier against the increase in width of the Sahara Desert with its extreme aridity and high temperatures. Where mountains are absent, the Sahara Desert almost reaches the Mediterranean coast. Annual rainfall ranges from 125 mm near the Sahara to 1000 mm in some upper zones of the Atlas Mountains and the High Tell Mountains in Tunisia. Maximum annual rainfall is concentrated in the coldest months of the year (from November-December to March-April), and is the main limiting factor for plant production in the whole area with the exception of the Nile valley, where water is provided by the river floods.

Agricultural features

In general, North African agriculture has been intensifying for most of the present century. Over the past three decades, this process has accelerated significantly (Swearingen, 1998). Intensification has occurred in two key directions, horizontally and, above all, vertically. Horizontally, North African farmers have progressively expanded cultivation to lower rainfall areas. In addition, they have expanded livestock raising to more marginal areas that previously were little used. Vertically, farmers have intensified agriculture through irrigation development, reduction of fallow, and intensification of stock raising. In the near term, the intensification processes have fueled economic development and enabled these countries to feed their growing populations. However, North African agriculture is now approaching its environmental limits due to diminishing land and water resources.

As regards the crops grown, rainfed spring cereals are by far the most dominant in the three western study countries (Morocco, Algeria, and Tunisia), with wheat representing 50 to 70 percent of total agricultural production. In these areas wheat is sown in November and December and is harvested from April to June. The sowing period is generally longer in Algeria and Tunisia than in Morocco, where the growing season usually begins earlier (MARS, 1997). A particular case is represented by Egypt, where the peculiar cultivation systems used are fed by the regular Nile's floods and relevant irrigation systems. Two growing seasons are therefore present, one with the peak in winter and the other with the peak in summer. During the former, wheat is mainly grown while, in the latter, there is a prevalence of maize and rice, which together represent more than 50 percent of total agricultural production (FAO, 1999).

Data Used

Cartographic Data

Digitized administrative boundaries of all North African countries were obtained from the Environmental Systems Research Institute (ESRI) ArcWorld™ 1:3M Continental Coverage CD ROM. These are very accurate national, regional, and provincial boundaries reported as vector files in a geographic projection.

A one-km digital land-cover map produced by the U.S. Geological Survey (USGS), the University of Nebraska-Lincoln (UNL), and the European Commission's Joint Research Centre (JRC) was acquired through the Internet (Brown *et al.*, 1993; Eidenshink and Faundeen, 1994). This is one portion of the global land-cover characteristics database that was developed on a continent-by-continent basis. All continents in the global database share the same map projections (Interrupted Goode Homolosine), have one-km nominal spatial resolution, and are based on one-km AVHRR data spanning from April 1992 through March 1993. Each continental database has unique elements that are based on the salient geographic aspects of the specific continent. The legend for the four study countries includes 11 land-cover classes which, from a preliminary analysis, were found to be rather mixed (Maselli *et al.*, 1998a).

Regarding analog cartography, an agro-pedo-climatic map was available only for Algeria. This map, produced by a national agency (BNEDER, 1994), reported the distribution of ten main land units at one-million scale. For Morocco, some cartographic data were available regarding cropland distribution in limited areas.

Crop Yield Data

Agricultural statistics for the main crops grown in the study countries were taken from the Food and Agriculture Organization (FAO) statistical database through the Internet. In particular, time series of annual harvested area, total production, and productivity for wheat, maize, and total cereals over the period 1982 through 1994 were taken from the database. Unfortunately, these data were available only on a country level, i.e., total values for each country and each year.

Out of these data, total cereal yield was considered for further processing as the most interesting agricultural parameter to estimate at the national level. According to what was stated above about the intensification of agricultural practices, cereal yield showed a general increasing trend in all study countries during the 13 years considered. However, a high inter-annual variability existed in Morocco, Algeria, and Tunisia, presumably due to meteorological (rainfall) fluctuations. Again, Egypt represented an exception because mean cereal yield was far higher than in the other countries (5 versus 1 T/ha) and there was a lower inter-annual variability thanks to the more regular water availability provided by Nile's floods.

Satellite Data

Following previous investigations, NOAA-AVHRR NDVI Maximum Value Composites (MVC) images were considered as the most appropriate for the estimation of crop yield (Groten, 1993; Maselli *et al.*, 1993; Hayes and Decker, 1996). In particular, attention was directed to the collection of imagery covering the whole belt of North African countries (Morocco, Algeria, Tunisia, Egypt) with the highest possible spatial and temporal resolutions and for a relatively long time period (at least 10 to 12 years).

Only two data sets with these requisite data were found, both derived from Global Area Coverage (GAC) images. The first were ten-day NDVI MVC images for the period 1981 through 1999 produced by the Global Inventory Monitoring and Modeling System (GIMMS) (Los *et al.*, 1994), and the second were ten-day NDVI MVC images for the period 1981 through 1994 produced by the National Aeronautic and Space Administration (NASA) Mission to Planet Earth program (James and Kallury, 1994).[1] Both of these data sets were freely distributed and were recovered through the Internet. After a preliminary evaluation, the latter set was chosen, even if shorter, for its higher geometric and radiometric consistency in the North African region (Maselli *et al.*, 1998a).

The original NOAA AVHRR images used to generate this data set came from Goddard Distributed Active Archive Center (DAAC) Pathfinder Land (PAL) data set and from NOAA's Satellite Active Archive (SAA). The final ten-day NDVI MVC imagery was produced through

- Unpacking and storing of GAC orbital data; this includes also retrieval of ancillary data needed for subsequent processing;
- Geopositioning of each scan using an orbital model;
- Application of calibration (Rao et al., 1993a; Rao et al., 1993b) and atmospheric corrections (Gordon *et al.*, 1988);
- Calculation and appending of cloud flags;
- Resampling of data to an 8- by 8-km resolution and conversion into the Goode Homolosine projection (Goode, 1925; Steinwand *et al.*, 1995); and
- Derivation of NDVI data and computation of ten-day composites (Holben, 1986); only pixels within 42 degrees of nadir were used in the composites.

Data Processing and Results

The entire image processing chain was carried out using both commercial software packages (IDRISI and ENVI) and specific programs written in-house in Fortran and C codes. All commercial packages and some specific programs were used on PC Pentium platforms, while the other programs ran on a Digital Microvax 3500 computer system.

Data Pre-Processing

All georeferenced digital maps and NDVI images were first converted to IDRISI and ENVI formats. Then, if in a different projection, they were re-projected into the geographic reference system maintaining approximately the original resolutions. This operation could not be performed by the commercial packages for the Pathfinder data, which were originally in a projection (Interrupted Goode Homolosine) unrecognized by IDRISI and ENVI, and a specific C program freely provided by NASA was used in this case. The agro-pedo-climatic map of Algeria was digitized by a scanner, and georeferenced directly to the geographic projection.

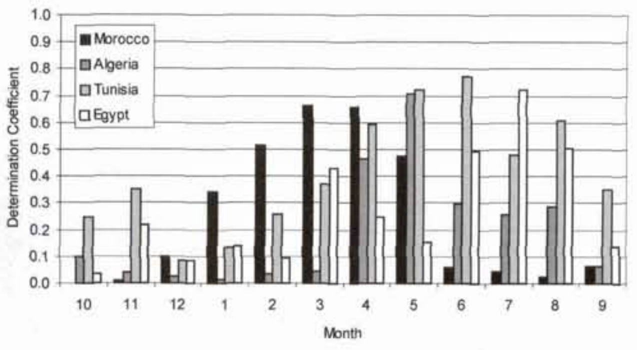

Figure 2. Determination coefficients found between mean monthly NDVI values of each country and final crop yield in the 13 study years (1982–1994).

Because the ten-day MVCs were still affected by cloud perturbations and other atmospheric effects, all of these images were further composed over one-month periods, obtaining in this way 12 images for each of the 13 study years. These monthly MVCs were visually checked for both residual atmospheric disturbances and geometric distortions, and were found to be of acceptable quality under both conditions.

Correlation Analysis with Global NDVI Data

A first evaluation of the available NDVI data for crop yield estimation was made by computing the monthly NDVI MVC averages of each country and regressing them with the annual cereal yield values (Anderson, 1984). Each regression analysis considered 13 points and was repeated for each study country for all months of the growing season. As previously explained, in the three western countries cereal crops are seeded in the fall and harvested in the spring of the following year, while in Egypt there are two growing seasons with the second ending in late summer. On the basis of this information and to maintain uniformity, all correlations were computed between crop yield of a season and monthly NDVI values from October of the previous year to September of the same year. In this way, it was possible to identify the most suitable month for final yield forecasting in each country as well as the efficiency of the process.

The results of all regression analyses are summarized in Figure 2 in the form of determination coefficients (r^2). As can be seen, the maximum correlations between monthy NDVI values and final crop yield were found in spring-summer months and were rather high even though the NDVI values were computed over entire countries. In practice, all maximum correlations were significant at the 99 percent confidence level, which, with 13 data points, corresponds to a determination coefficient of 0.47. The maximum correlations were found in different months for the four countries. The peak of correlation was in fact in March–April for Morocco, May for Algeria, May–June for Tunisia, and July for Egypt, where also a secondary peak in winter was present. These correlation patterns can be explained on the basis of the above-mentioned differences in the phenology of cereals, which is anticipated in the western zones and mainly linked to the summer growing season in Egypt.

---

[1] The lengths reported are referred to the two data series available when the current work was completed (summer 1999).

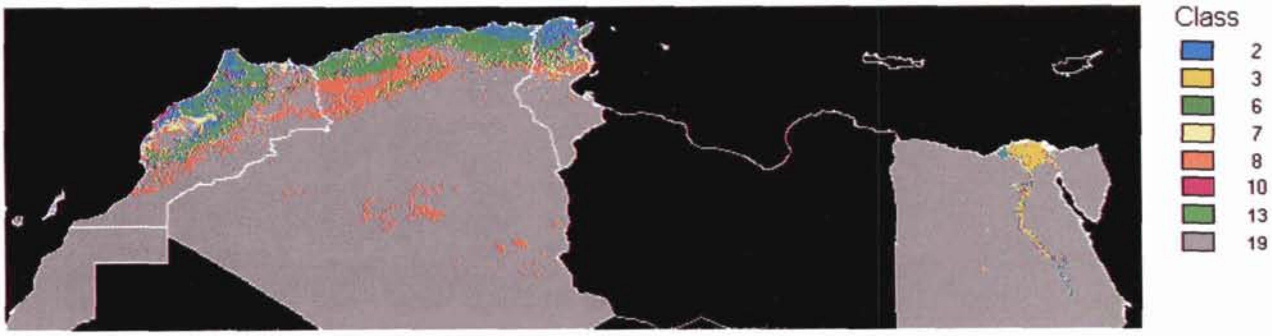

Plate 1. USGS supervised classification windowed to the study area and degraded to the GAC pixel size (8 by 8 km).

TABLE 1. THE USGS CLASSES WHICH ARE PRESENT IN THE STUDY AREA

USGS Class	Description
2	Dryland cropland and pasture
3	Irrigated cropland and pasture
6	Cropland/Woodland mosaic
7	Grassland
8	Shrubland
10	Savanna
13	Evergreen broadleaf forest
19	Barren or sparsely vegetated land

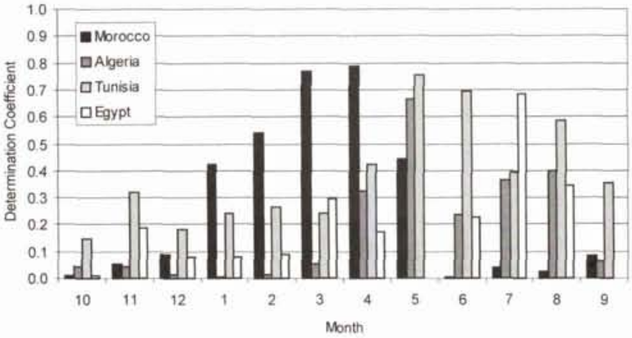

Figure 3. Determination coefficients found between mean monthly NDVI values of the most correlated USGS land-cover class of each country (class 6 for Morocco, Tunisia, and Algeria; class 3 for Egypt) and final crop yield in the 13 study years (1982–1994).

Correlation Analysis with NDVI Data of Single Land-Cover Classes

After the preliminary analysis at the national level, attention was focused on the identification of areas where NDVI values were most correlated with crop yield. For this scope, the land-cover classification produced by the USGS was used as indicator of homogeneous zones (Brown *et al.*, 1993). The global land-cover map was therefore windowed to an area corresponding to that of the GAC images. Plate 1 shows a version of the classification degraded to the GAC pixel size (8 by 8 km). As previously mentioned, this reported 11 land-cover classes for the study area, which were successively reduced to eight for Egypt and seven for the other countries by retaining only classes with more than 30 GAC pixels for each country and including a desert class. The legend of the modified USGS classification is summarized in Table 1.

Because the USGS classification was available at a 1-km resolution, a methodology was applied to extract the NDVI profiles of the pure classes at GAC resolution. This was a method recently proposed by our research group for the integration of data with different spatial and temporal resolutions (Maselli *et al.*, 1998b). The method is based on an assumption of linearity in the composition of NDVI values from different cover classes which, even though strictly valid only for the original bands, was found to hold with a good approximation also for the already composed index (Kerdiles and Grondona, 1995). First, a Boolean mask for each class of the 1-km USGS classification image was created. Next, the mask images of the eight classes were degraded to GAC resolution by simple averaging, obtaining abundance images which reported the proportion of each land-cover class in each GAC pixel (Kerdiles and Grondona, 1995). These abundance images were then regressed against the NDVI images by a multivariate linear procedure applied independently to each country. The extrapolation of the regression models found to complete class cover (fraction equal to 1 for the class considered and to 0 for all others) allowed the computation of NDVI profiles analogous to end-member spectra for each USGS class and each country (Maselli *et al.*, 1998b). Finally, the monthly NDVI values of the classes were regressed against the annual yield data in the same way as for the previous global analysis.

The results of the NDVI/yield regressions were evaluated again by considering determination coefficients. To summarize the results obtained, the agricultural classes with highest correlations with cereal crop yield were selected, which corresponded to class 6 (cropland/woodland mosaic) for Morocco, Algeria, and Tunisia and to class 3 (irrigated cropland and pasture) for Egypt. The determination coefficients of the NDVI/yield regressions for these classes are shown in Figure 3. As can be seen, the seasonal correlation patterns were very similar to those obtained for entire countries, and also the absolute values were almost the same for Algeria, Tunisia, and Egypt, while a certain improvement was found for Morocco. This indicated that the use of the supervised classification brought only marginal improvements to the yield forecasting capability, probably due to the presence of mixed vegetation types within each land-cover class (Maselli *et al.*, 1998a).

Correlation Analysis with NDVI Data of Selected Pixels

On the basis of the previous considerations, a different approach was tested for the identification of cropland pixels and the selective extraction of their NDVI profiles. Because no reliable information was available on the spatial distribution of the cultivated areas at the GAC pixel resolution in all study countries, an image-based approach was applied for the identification of cropland pixels. In practice, it was hypothesized that the pixels of each country most intensively cultivated had NDVI

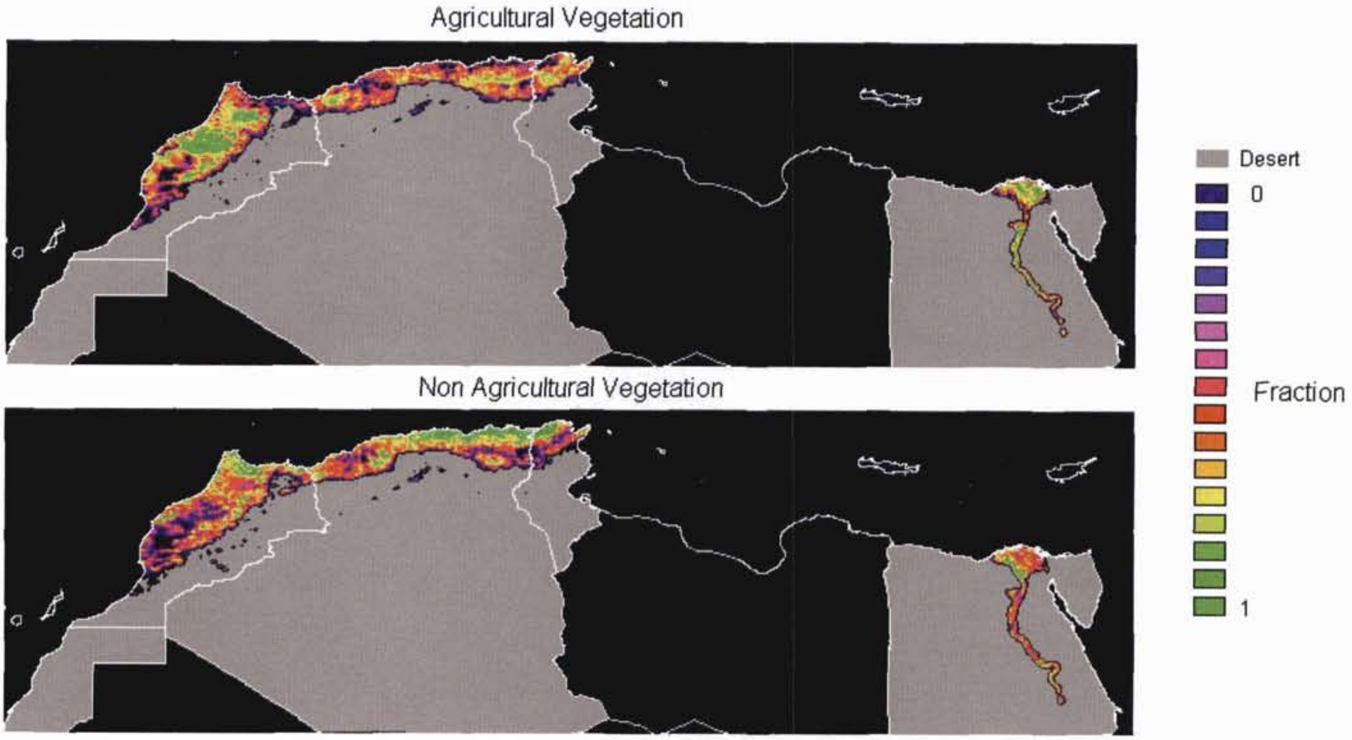

Plate 2. Fraction images of agricultural and non-agricultural vegetation obtained with the procedure described in the text.

values in the best yield predicting period most correlated with the relevant final yields. This hypothesis relied on the assumption that the spatial distribution of cropland was approximately stable during the study years, at least at a GAC pixel scale. The best monthly NDVI MVCs to predict crop yield were therefore selected for each country, which corresponded to images of April for Morocco, May for Algeria and Tunisia, and July for Egypt. A synthetic image was then generated for every country with each pixel expressing the correlation coefficient between these NDVI values in the 13 study years and the relevant values of final crop yield. This correlation image was considered as expressing information on the fractional cropland cover within each pixel. Because, however, it was noted that areas with very low NDVI values also could show high correlation with crop yield, probably due to the presence of grasses with phenology similar to that of cereals, a vegetation mask was additionally applied. In practice, three images expressing vegetation quantity in the selected months (April, May, and July) were computed by averaging the relevant NDVI MVCs over the 13 years. These mean vegetation images were stretched between 0.2 and 0.4 NDVI, which approximately corresponded to the limits of desert and quite homogeneous vegetation cover (Maselli *et al.*, 2000). Each resulting "mean vegetation" image, re-scaled between 0 and 1, was finally multiplied with two images, the first expressing the correlations with crop yield (values between 0 and 1) and the second the complement to one of these values. A third image was obtained by subtracting the first two from 1. In this way, three synthetic images were obtained for each country, all with values between 0 and 1, indicating, respectively,

- The fraction of vegetation correlated with crop yield (agricultural vegetation),
- The fraction of vegetation not correlated with crop yield (non-agricultural vegetation), and
- The fraction of barren land.

All three images were subjected to a Gaussian filtering with a 0.5-pixel standard deviation to reduce random noise. The three fraction images of each country were then re-composed to produce unique image for each land-use type (agricultural and non-agricultural vegetation and barren land) covering the entire study area. The first two of these images are shown in Plate 2 (the third is simply the complement to one of these two).

Based on the above considerations, these images were considered similar to abundance images, which could be used to extract the NDVI profiles of the three "land-use types" as done previously with the abundance images derived from the USGS classification. This again assumed a linearity in the composition of the NDVI values from each fraction image, which, even though not strictly demonstrated, was considered reasonable because the synthetic images were derived from linear correlation analysis with crop yield. The same methodology used before (multivariate regression and NDVI extrapolation to full class fraction for each country) was therefore applied to compute NDVI profiles of the three land-use types which again could be considered analogous to end-member spectra. The average annual profiles of the four countries over the study period are shown in Figure 4. It is worth noting that, in the three western countries, these profiles were typical of

- Areas cultivated with cereal crops, with a clear NDVI maximum in the spring months and low NDVI values in summer;
- Forest areas with a slight NDVI maximum in winter-spring months but with high NDVI during the whole year; and
- Desert areas, with NDVI always close to 0.

Egypt was again peculiar for the bimodal profiles of both agricultural and non-agricultural vegetation. In this case, the profile of the former land-use type differed from that of the latter for its higher winter and summer NDVI peaks and lower spring and fall minima. These features, as expected, were indicative of a more intense vegetation activity of cropped areas dur-

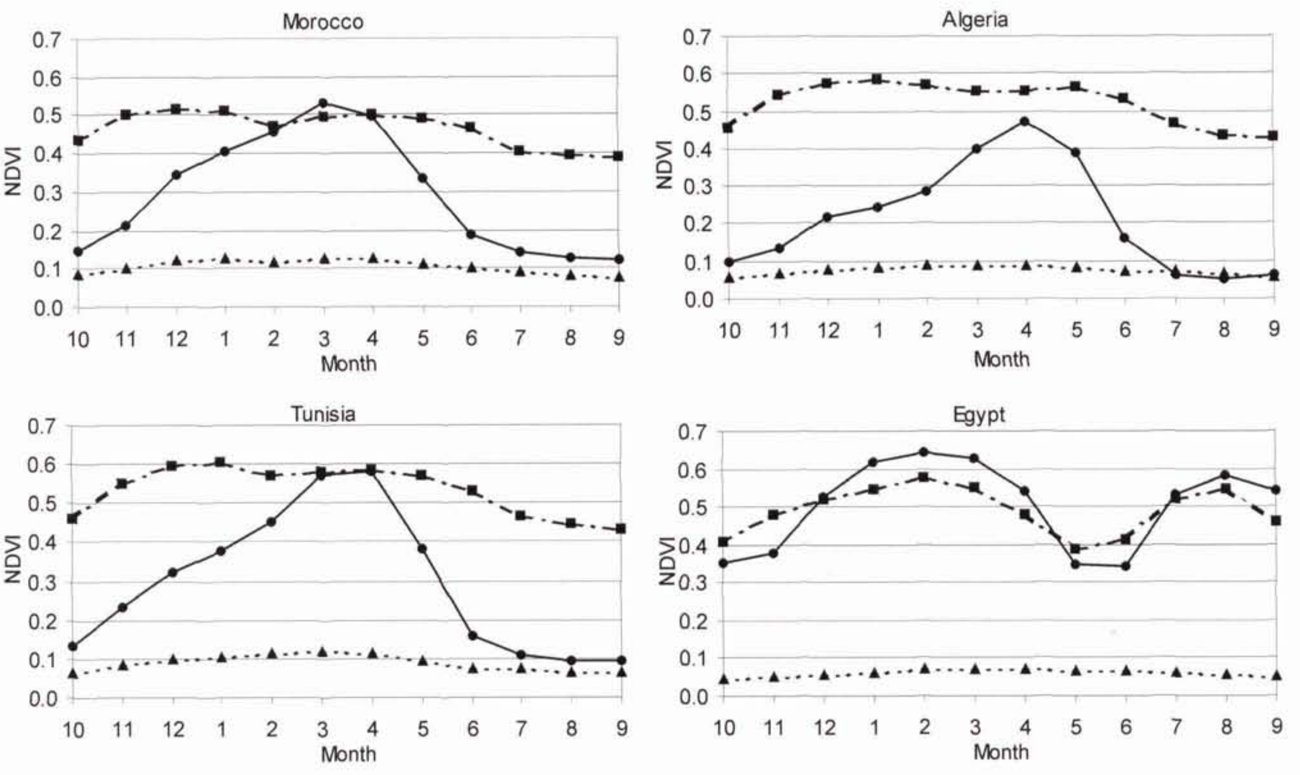

Figure 4. Average NDVI profiles for the three land-use classes (circle = agricultural vegetation, square = non-agricultural vegetation, triangle = barren land) identified using the described procedure.

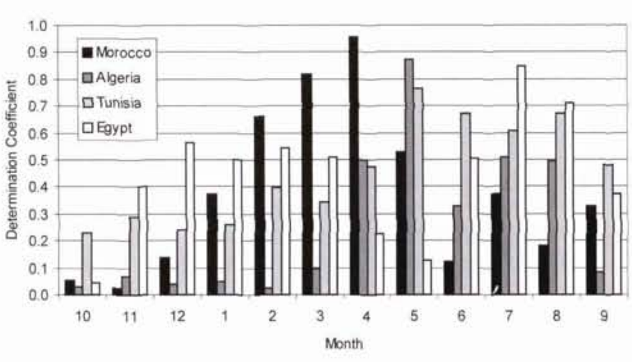

Figure 5. Determination coefficients found between monthly NDVI values of agricultural vegetation for each country and final crop yield in the 13 study years (1982–1994).

ing the two growing seasons. Globally, the methodology applied seemed to decompose the NDVI multitemporal data according to the main land use types of each country.

Regression analysis was again applied between the estimated NDVI profiles and crop yield data for the 13 years. The determination coefficients found for the first land-cover type in the four countries are shown in Figure 5. As can be seen, a decisive improvement in correlation was obtained for Morocco, Algeria, and Egypt. It is also worth noting that the other two land-use types, and particularly non-agricultural vegetation, showed very low and non-significant determination coefficients with final crop yield (0 to 0.3). This was interpreted as a further indication that the methodology applied separated well the main land-use components in the study area.

From the comparison of the annual NDVI profiles of Figure 4 with the relevant correlation trends of Figure 5, another interesting observation was derived about the temporal relationships NDVI/yield. Where in fact wheat was the main spring crop (Morocco, Algeria, and Tunisia), the NDVI peak preceded the period of maximum correlation with final yield of about one month. This is in accordance with previous investigations which found that the most influential period for wheat yield determination follows the flowering phase (which approximately coincides with the NDVI peak) of 20 to 30 days (Rossini and Benedetti, 1993). On the contrary, this was not the case where maize and rice were the prevalent summer crops (Egypt). In this country, the NDVI peak of August in fact followed the period of maximum correlation of July, presumably due to the different physiological behaviors of these crops.

Evaluation of the Maps Obtained

An evaluation of the three fraction images obtained in the previous section was first carried out by comparing them to the USGS land-cover classification. In Table 2 the fractional covers of the three land-use types are reported for each USGS class in the study area. It can be noted that, apart from USGS classes 8 (shrubland) and 19 (barren or sparsely vegetated land), which correspond to our non-vegetated land-use type, no complete correspondence exists. USGS class 7 (grassland) mostly corresponds again to barren land, while classes 2 (dryland cropland and pasture), 3 (irrigated cropland and pasture), 6 (cropland/woodland mosaic), 10 (savanna), and 13 (evergreen broadleaf forest) are differently split among our three land-use types. In practice, the best correspondence with our agricultural vegeta-

Table 2. Fractions of the Three Land-Use Types Identified by the Current Methodology for Each USGS Land-Cover Class

USGS Class	Agricultural Vegetation	Non-Agricultural Vegetation	Barren Land
2	0.33	0.43	0.24
3	0.57	0.40	0.03
6	0.44	0.40	0.16
7	0.20	0.15	0.65
8	0.04	0.03	0.93
10	0.29	0.31	0.40
13	0.27	0.31	0.42
19	0.00	0.00	1.00

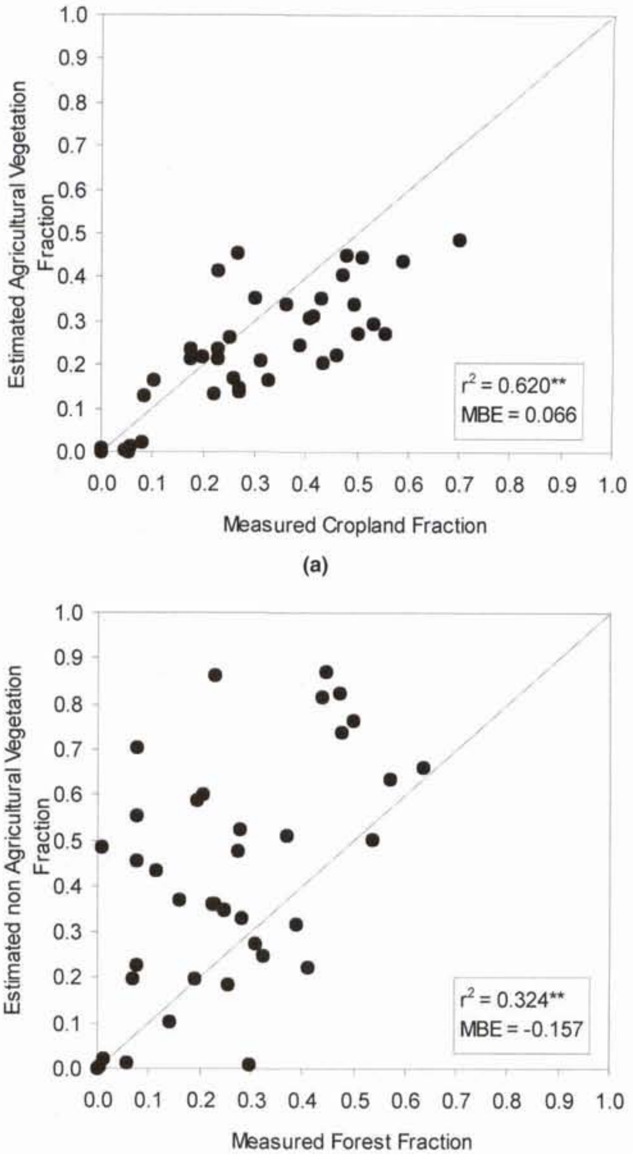

Figure 6. Linear regressions between the fractions of agricultural (a) and forest (b) vegetation derived from the agro-pedo-climatic map (measured fraction) and our land-use fraction images (estimated fraction) for the main 39 northern provinces of Algeria (** = correlation significant at the 99 percent confidence level).

tion type was found for the USGS classes 3 (present almost exclusively in Egypt) and 6.

For a clearer interpretation of these findings, another comparison was made against the independent agro-pedo-climatic map of Algeria produced by BNEDER (1994). Unfortunately, the geometric accuracy of the digitized map was found to be rather poor, having a mean error of about two GAC pixels. Thus, only comparisons over very large areas were deemed feasible. The fractions of the seven USGS classes and our three land-use types were, therefore, computed for the 39 largest northern districts of Algeria and were compared to the relevant fractions derived from the agro-pedo-climatic map. The results, expressed as determination coefficients, indicated only a clear correspondence between the USGS class 6 (cropland/woodland mosaic) and the agricultural agro-pedo-climatic class ($r^2 = 0.56$). All other correlations with the USGS class fractions were very low (r^2 below 0.2). On the contrary, the fractions of agricultural and non-agricultural vegetation from our images were correlated to the two corresponding classes of the agro-pedo-climatic map. As can be seen in Figure 6, the fractions of agricultural vegetation from the two sources had $r^2 = 0.62$, while a lower, but still significant, correlation was found for forest areas ($r^2 = 0.32$). The area correspondence of non-agricultural vegetation from the two sources, measured as mean bias error (MBE), was rather poor, but this was expected considering that the agro-pedo-climatic map reported other classes with mixed agricultural, pastoral, and forest uses.

Because incomplete cartographic information was available for Morocco, only a visual evaluation of the USGS classes and our images was possible for this country. This subjective evaluation confirmed that only USGS class 6 had a certain correspondence with agricultural areas, while our fraction images were highly informative on crop and forest distribution. In the central and northern part of Morocco, for example, areas known to have agricultural and forest cover corresponded well to pixels identified as crop and forest lands, respectively.

Evaluation of the Procedure for Operational Yield Forecasting

The identification of cropland areas is especially significant in the context of an operational yield forecasting system (Hutchinson, 1991). This implies that, in addition to being correctly identified, cropland pixels should remain approximately stable in different years. Having NDVI images and ground data for only 13 years, this stability was checked by a leave-one-out approach. The same procedure for cropland identification (creation of correlation images, masking with mean vegetation images, and Gaussian filtering) was therefore repeated 13 times using each time data for 12 years, with the exclusion of a rotating year. The three fraction images created for each time (with agricultural and non-agricultural vegetation and barren land) were then used to derive the three pure NDVI profiles for the remaining year by applying the multivariate regression method described previously to the relevant NDVI data. After 13 repetitions of this procedure, the NDVI values found for the best month of each country (April, May, and July) were again regressed against the relevant crop yield. In practice, this procedure simulated a situation in which the fraction images identified in a certain period are applied to crop yield forecasting in a following year.

The results for the four countries are shown in Figure 7. Even if correlations were lower than previously obtained, the prediction capability of the procedure remained high. Determination coefficients higher than 0.64 and root-mean-square errors (RMSE) lower than 18 percent of the mean yield values were in fact achieved for all countries using NDVI images taken 1 to 2 months earlier than crop harvesting.

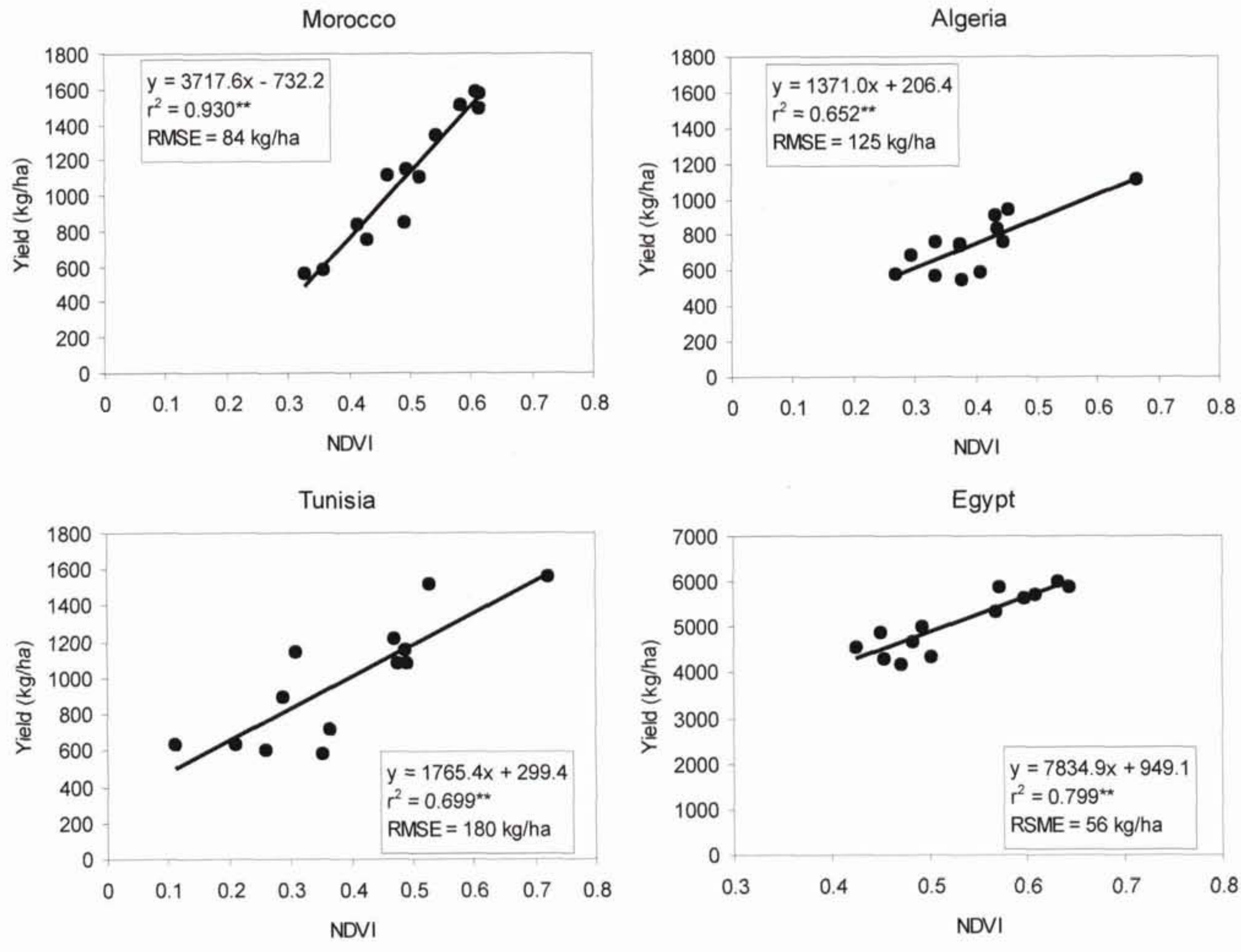

Figure 7. Linear regressions between final crop yield and NDVI values obtained using the leave-one-out procedure for the four study countries (** = correlation significant at the 99 percent confidence level).

It can also be noted that, similar to the previous analysis with data for all years, different NDVI/yield slopes were obtained for the four countries, which indicates an increasing efficiency of NDVI conversion to grain products going from Algeria to Tunisia, Morocco, and, finally, Egypt (Prince, 1991). This is again in accordance with existing information, because Algeria is known to have the poorest agricultural technologies, while Egypt has the highest productivity thanks to its peculiar water provision systems (FAO, 1999).

Summary and Conclusions

The importance of yield forecasting in semiarid developing countries is clearly apparent and does not need to be further highlighted. Recent research demonstrated that NOAA-AVHRR NDVI MVCs are particularly suitable for this purpose, because they are sensitive indicators of vegetation conditions during crop growth and can therefore be related to final productivity. In the current work, we performed investigations on this subject in four North African countries involved in the EU project CAMELEO. For the investigation, we had to rely on GAC NDVI data, which, even though not optimal for spatial and radiometric properties, were the most consistent set available over the entire study area for a relatively long period (13 years). Moreover, because only annual crop yield estimates for each country were available, the comparisons had to be made at the national level, which obviously prevented a higher spatial detail of the analysis.

In spite of these limitations, the first testing demonstrated that interesting relationships between final crop yield and NDVI values of suitable periods can be found, even at a country level. In particular, different optimal linear relations were found for each country when using NDVI data for the spring (April-May) and summer (July) months.

Next, it was verified that the USGS land-cover classification was not very effective for cropland identification, probably due to the definition of broad classes which contain mixed vegetation types. A specific method was therefore developed based on the assumption that cropped areas had NDVI values in the optimal yield forecasting period more correlated with final yield than did non-cropped areas. Correlation images were thus created and composed with mean NDVI images in order to estimate the distribution of agricultural and non-agricultural vegetation. The consideration of these images as abundance estimates allowed the extraction of NDVI multitemporal profiles of three main land-use types which were analogous to end-member spectra and particularly correlated and uncorrelated with final crop yield. Also, the areas identified as agricultural and non-agricultural vegetation were found to correspond to crop and forest land derived from independent sources. The application of a leave-one-out strategy finally confirmed the utility of the approach for operational crop yield forecasting.

The methodology proposed obviously relies on statistical considerations which may not be valid in other environmental situations. The whole approach is, in fact, based on the assumptions that the phenological development of the main cereal crops is approximately synchronous and constant within each country and that the distribution of these crops is stable in different years, at least at the GAC pixel scale. These assumptions do not generally hold in temperate countries, where crops with different phenologies are usually intermingled, plant development is variable in time due to several factors, and inter-year crop rotations are often applied. The existence of suitable conditions for the application of the method must therefore be ascertained in each single case, considering both local environmental features and agricultural practices. Also, further modifications are possible in order to improve the performance of the procedure and to adapt it to different ecological situations. In general, however, the operational utility of the method can be notable in semiarid areas where scarce reference data are available, because it only requires low-resolution NDVI data and large-scale ground yield assessments.

The current findings also have important implications for the evaluation of the GAC data set utilized. Even though these data are known to suffer from relevant technical limitations, they proved to contain information which, if suitably extracted, can be very useful for land analysis, not only at a regional level, but also at per-pixel and sub-pixel scales. In this sense, the availability of long-period image series represents a counterbalance for the decreased spatial resolution of the data. The use of multitemporal NDVI profiles, in fact, allows the extraction of sub-pixel information, at least in these semiarid environments with few principal land-use components.

A final remark is worth noting about the future extension of the research work within the framework of the CAMELEO project. The ultimate objective of this project is the identification of land-cover changes which can be indicative of possible degradation and/or desertification phenomena. Because vegetation variations can be generally ascribed almost exclusively to meteorological vagaries and human interventions, the identification of areas where these two factors are contemporaneously or separately active can be the basis for the analysis of relevant long-term vegetation responses. As a first step, the identification of cropped and non-cropped areas could serve to concentrate desertification assessment on the latter, because the former are mainly controlled by human activities. Within a more advanced strategy, the NDVI signal correlated to crop yield could be removed on a per-pixel basis, and the remaining information could be related to meteorological driving factors (mainly rainfall) to detect possible variations in water use efficiency not related to agricultural intensification and therefore indicative of environmental changes (Davenport and Nicholson, 1993). These are the main research lines towards which our activity is presently directed within the CAMELEO project and other collaborations with the Agrhymet Centre of Niamey (Niger).

Acknowledgments

The authors wish to thank the U.S. Geological Survey (USGS), the University of Nebraska-Lincoln (UNL), and the European Commission's Joint Research Centre (JRC) for generating the 1-km resolution global land-cover characteristics database and distributing it through the Earth Resources Observation System Data Center (EDC) Distributed Active Archive Center (DAAC). The authors also wish to thank the Distributed Active Archive Center (Code 902.2) at the Goddard Space Flight Center, Greenbelt, Maryland, for producing the data in their present form and distributing them. The original data products were produced under the NOAA/NASA Pathfinder program by a processing team headed by Ms. Mary James of the Goddard Global Change Data Center; and the science algorithms were established by the AVHRR Land Science Working Group, chaired by Dr. John Townshend of the University of Maryland. Goddard's contributions to these activities were sponsored by NASA's Mission to Planet Earth program. The agricultural statistics were taken from the FAO interactive archive (Rome). Copies of the agro-pedo-climatic map of Algeria and the agricultural maps of Morocco were kindly provided by the Istituto Agronomico per l'Oltremare (IAO), Florence. These contributions are gratefully acknowledged. The work of Felix Rembold was financed by a CNR grant supported by the EU Project CAMELEO (Contract No. IC18-CT97-0155). Finally, the authors want to thank Paolo Rossini and three anonymous PE&RS referees for their useful comments and suggestions on the first draft of the paper.

References

Anderson, T.W., 1984. *An Introduction to Multivariate Statistical Analysis, Second Edition*, Wiley and Sons, New York, N.Y., 102 p.

BNEDER, 1994. Carte de Synthèse des grands ensembles agro-pedo-climatiques 1:1.000.000, Bureau national d'Etudes de développement rurale, Alger, Algerie.

Baret, F., and A. Olioso, 1989. Estimation à partir de mesures de réflectance spectrale du rayonnement photosynthétiquement actif absorbé par une colture de blé, *Agronomie*, 9:885–895.

Brown, J., T.R. Loveland, J.W. Merchant, B.C. Reed, and D.O. Ohlen, 1993. Using multisource data in global land-cover characterization: Concepts, requirements, and methods, *Photogrammetric Engineering & Remote Sensing*, 59:977–987.

Davenport, M.L., and S.E. Nicholson, 1993. On the relation between rainfall and the Normalized Vegetation Index for diverse vegetation types in East Africa, *International Journal of Remote Sensing*, 14:2369–2389.

Eidenshink, J.C., and J.L. Faundeen, 1994. The 1 km AVHRR global land data set—First stages in implementation, *International Journal of Remote Sensing*, 15:3443–3462.

Emberger, L., 1955. Une classification biogéographique des climats, *Natur. Monspel., série Botanique*, 7:3–43.

FAO, 1999. *FAOSTAT 98 CD-ROM. FAO Statistical Databases*, ISBN 9250042302, Food and Agricultural Organization, Rome, Italy.

Genovese, G., C. Vignolles, T. Nègre, and G. Passera, 1999. The use of CORINE land cover to improve vegetation monitoring through NOAA-AVHRR/NDVI profiles, *Proceedings of the International Symposium "Modelling Cropping Systems,"* 21–23 June 1999, Lleida, Spain, pp. 83–84.

Goode, J.P., 1925. The Homolosine projection: A new device for portraying the Earth's surface entire, *Association of American Geographers, Annals*, 15:119–125.

Gordon, H.R., J.W. Brown, and R.H. Evans, 1988. Exact Rayleigh scattering calculations for use with the Nimbus-7 coastal zone colour scanner, *Applied Optics*, 27:2111–2122.

Groten, S.M.E., 1993. NDVI-crop monitoring and early yield assessment of Burkina Faso, *International Journal of Remote Sensing*, 8:1,495–1,515.

Hayes, M.J., and W.L. Decker, 1996. Using NOAA AVHRR data to estimate maize production in the United States Corn Belt, *International Journal of Remote Sensing*, 17:3189–3200.

Holben, B.N., 1986. Characteristics of maximum-value composite images from temporal AVHRR data, *International Journal of Remote Sensing*, 7:1,417–1,434.

Hutchinson, C.F., 1991. Uses of satellite data for famine early warning in sub-Saharan Africa, *International Journal of Remote Sensing*, 12:1405–1421.

James, M.E., and S.N.V. Kalluri, 1994. The Pathfinder AVHRR land dataset: An improved coarse resolution dataset for terrestrial monitoring, *International Journal of Remote Sensing*, 15:3347–3363.

Kerdiles, H., and M.O. Grondona, 1995. NOAA-AVHRR NDVI decomposition and subpixel classification using linear mixing in the Argentinean Pampa, *International Journal of Remote Sensing*, 16:1,303–1,325.

Lacaze, B., V. Caselles, C. Coll, H. Hill, C. Hoff, S. De Jong, W. Mehl, J.F.W. Negendank, H. Riesebos, E. Rubio, S. Sommer, J. Teixeira Filho, and E. Valor, 1996. *DeMon—Integrated Approaches to Desertification Mapping and Monitoring in the Mediterranean Basin*, Final Report of DeMon-1 Project, Joint Research Centre of European Commission, Ispra (VA), Italy, 165 p.

Lewis, J.E., J. Rowland, and A. Nadeau, 1998. Estimating maize production in Kenya using NDVI: Some statistical considerations, *International Journal of Remote Sensing*, 19:2609–2617.

Los, S.O., 1998. *Linkages between Global Vegetation and Climate. An Analysis Based on NOAA Advanced Very High Resolution Radiometer Data*, PhD. Dissertation, Vrije Universiteit, Goddard Space Flight Center, Greenbelt, Maryland, 179 p.

Los, S.O., C.O. Justice, and C.J. Tucker, 1994. A global 1° by 1° NDVI dataset for climate studies derived from the GIMMS continental NDVI, *International Journal of Remote Sensing*, 15:3493–3518.

MARS, 1997. *MARS Bulletin, 1996–1997 Season. Situation from the 1 September to 31 December 1996*, Vol.5, 1. J.R.C. Ispra, Italy, 32 p.

Maselli, F., C. Conese, L. Petkov, and M.A. Gilabert, 1993. Environmental monitoring and crop forecasting in the Sahel through the use of NOAA NDVI data. A case study: Niger 1986–89, *International Journal of Remote Sensing*, 14:3471–3487.

Maselli, F., F. Rembold, and P. Rossini, 1998a. Preliminary evaluation of different NOAA-AVHRR data sets for the assessment of land surface changes in North African countries, *CAMELEO (Changes in Arid Mediterranean Ecosystems on the Long Term through Earth Observation), Annual Progress Report Year 1998* (R. Escadafal, H. Bohbot, and J. Mégier, editors), Space Applications Institute, JRC, Ispra, Italy, pp. 67–83.

Maselli, F., M.A. Gilabert, and C. Conese, 1998b. Integration of high and low resolution NDVI data for monitoring vegetation in Mediterranean environments, *Remote Sensing of Environment*, 63:208–218.

Maselli, F., S. Romanelli, L. Bottai, and G. Maracchi, 2000. Processing of GAC NDVI MVCs for yield forecasting in the Sahelian region, *International Journal of Remote Sensing*, 21(18):3509–3523.

Prince, S.D., 1990. High temporal frequency remote sensing of primary production using NOAA AVHRR, *Applications of Remote Sensing in Agriculture* (M.D. Steven and J.A. Clark, editors), Butterworths, London, U.K., 4:169–183.

———, 1991. Satellite remote sensing of primary production: Comparison of results for Sahelian grasslands 1981–1988, *International Journal of Remote Sensing*, 12:1,301–1,312.

Rao, C.R.N., J.T. Sullivan, C.C. Walton, J.W. Brown, and R.H. Evans, 1993a. *Nonlinearity Corrections for the Thermal Infrared Channels of the Advanced Very High Resolution Radiometer: Assessment and Recommendations*, NOAA Technical Report NESDIS-69, NOAA/NESDIS, Washington, D.C., 31 p.

———, J. Chen, F.W. Taylor, P. Abel, Y.J. Kaufman, E. Vermote, W.R. Ross, and C. Brest, 1993b. *Degradation of the Visible and Near Infrared Channels of the Advanced Very High Resolution Radiometer on the NOAAP9 Spacecraft: Assessment and Recommendations for Corrections*, NOAA Technical Report NESDIS-70, NOAA/NESDIS, Washington, D.C., 25 p.

Rossini, P., and R. Benedetti, 1993. On the use of NDVI profiles as a tool for agricultural statistics: The case study of wheat yield estimate and forecast in Emilia Romagna, *Remote Sensing of Environment*, 45:311–326.

Steinwand, D.R., J.A. Hutchinson, and J.P. Snyder, 1995. Map projections for global and continental data sets and an analysis of pixel distortion caused by reprojection, *Photogrammetric Engineering & Remote Sensing*, 61:1,487–1,497.

Swearingen, W., 1998. Agriculture and environmental constraints in North Africa, *Proceedings, Conference on Agricultural Development in Central Asia, between Russia and the Middle East*, 20–22 November, University of Washington, Seattle (in press).

(Received 16 September 1999; accepted 09 May 2000; revised 11 August 2000)