



Project no. 004089
AMMA
African Monsoon Multidisciplinary Analysis
Instrument: IP
Thematic priority: 1.1.6.3 Global change and Ecosystems

WP 3.2: Human process, adaptation and environmental interactions

Du 3.2 b Stratification of Ecosystems

Spatial and temporal indicators for identifying the cultivated domain in the Sudano-Sahelian environment

Start of the project: 1 January 2006

Duration: 60 months

In charge of the deliverable: CIRAD

Marion Houles¹
Christian Feau, Gérard Lainé, Christian Baron, Dany Lo Seen²

March 2007

¹ Master 2 professionnel Géomatique, Université de Toulouse France

² CIRAD, Centre de coopération Internationale en Recherche et Développement, France

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

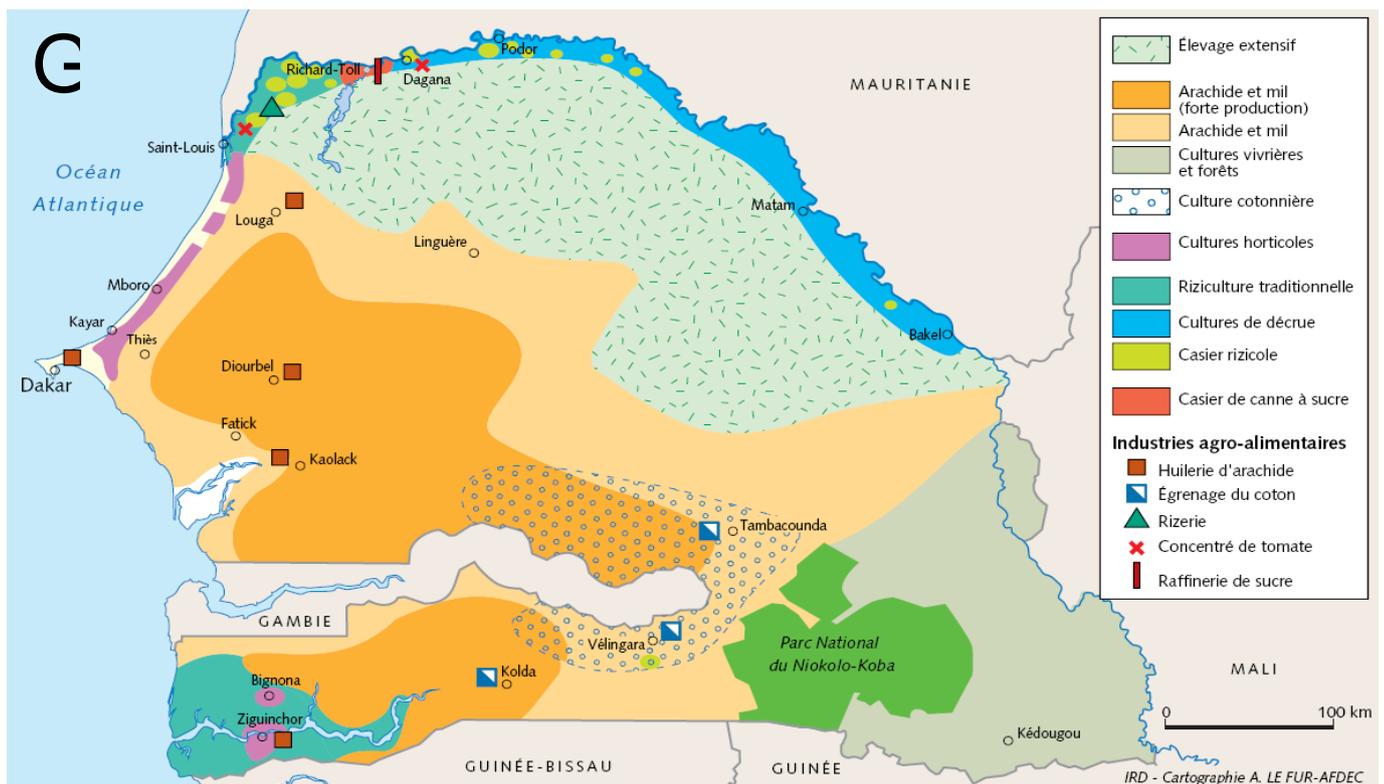
I INTRODUCTION

I.1 Aim of the deliverable

One of the preliminary stages necessary for analysing the impact of climate change on agriculture and food security is the spatialisation of the cultivated domain at a national scale (scale used for studying climate change) with a precision approaching that of regional-scale spatialisation. This spatialisation can be based on remote sensing by using satellite data such as MODIS, Spot Vegetation or Landsat, whose low acquisition costs allow for the development of tools and of a methodology that can be used by our African partners. The present deliverable has been prepared based on the work carried out by M. Houllès during her Master Degree (Houllès, 2007).

I.2 Selection of the study area

Senegal was selected as the study area in view of the project participants' knowledge of the terrain, derived from studies they were conducting there, as well as the partnerships they had established there.



Map 1: Agricultural land base of Senegal

Sources: IRD – cartography A. LE FUR-AFDEC

I.3 Issues and scientific context

Changes in the farming landscape are caused, among other reasons, by the interaction between climate change and human activity. These landscapes and their changes are

detectable by remote sensing at different spatial levels (from the farm to the region) and temporal levels.

Since the 1980s, the use of remote sensing at low spatial resolutions (kilometric resolutions) has become increasingly widespread because it allows the study and monitoring of vast ecosystems at national or continental scales. Low-resolution satellite images allow regular monitoring of renewable resources at very low costs. The loss in precision due to the low spatial resolution is compensated for by a large coverage area and by an excellent temporal resolution. Daily repetitivity is, in fact, essential for regular monitoring of resources, especially vegetation resources. Nevertheless, the low (kilometric) spatial resolution of this data does not, by itself, provide a satisfactory representation of the landscape. The pixel size of low-resolution images is too coarse, for example, to identify crops at the plot scale.

But recent medium-resolution sensors, such as MODIS/TERRA, with spatial resolutions as low as 250 m, offers new possibilities in the study of the cultivated domain. With an increase in spatial resolution from one kilometre to 250 m, the detection of groups of plots can now be considered. However, this remains to be evaluated and confirmed for agricultural landscapes that are fragmented. This brings us to the question whether high- and low-resolution satellite data, in combination with external data (thematic maps, statistics, etc.), can provide a correct representation of the distribution of the cultivated domain at the country scale.

I.4 General methodology

The aim of this work is to establish a methodology that can identify as precisely as possible the cultivated domain at a national scale, based on satellite images. The cultivated domain will be considered to include all crops and recently fallowed land.

Therefore spatial, spectral and temporal indicators capable of separating crops from other land uses need to be proposed and tested. More precisely, among temporal indicators, we need to know whether it is possible to identify the cultivated domain on the basis of the phenological behaviour of the vegetation, as derived from the temporal variation of a vegetation index, such as the NDVI (Normalized Difference Vegetation Index).

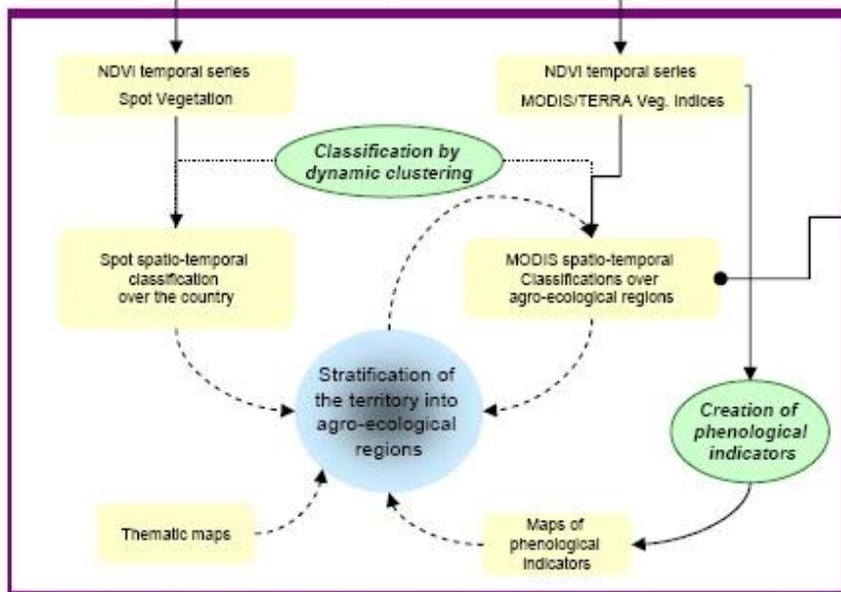
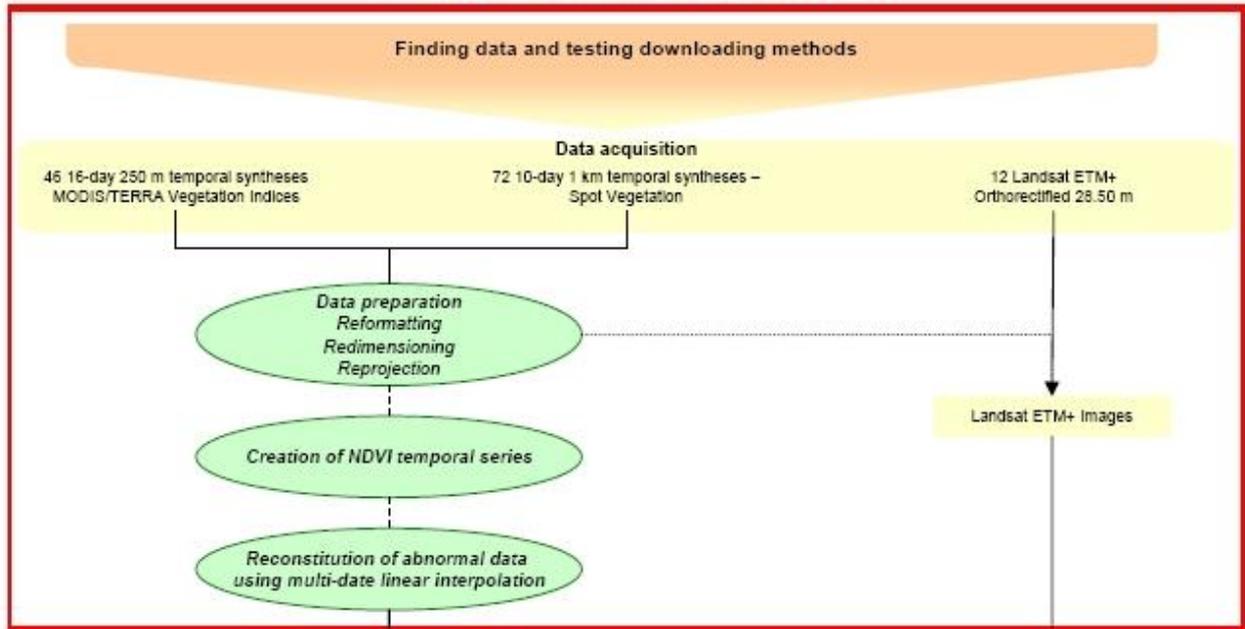
The proposed methodology should also be easily transposable to other areas and available at low cost because it should allow regular monitoring of agricultural resources of countries concerned by drought and food insecurity.

The approach adopted in this study ranges from the general to the specific (Bertrand et al., 1985). At each working scale, from the country to the large agro-ecological regions to the micro-regions, the satellite images that are the most appropriate in terms of spatial resolution have been used.

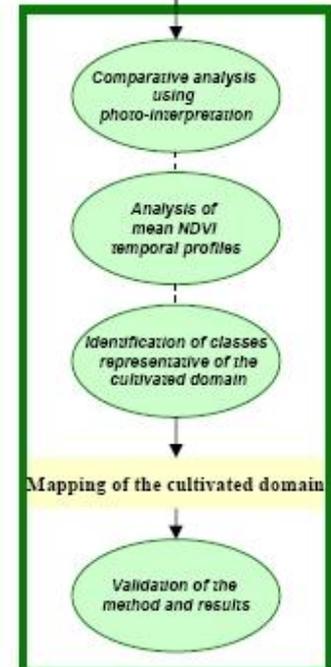
The approach proposed to achieve the aims listed above can be broken down into several stages:

- Data acquisition and preparation
- Stratification of the territory into homogeneous agro-ecological zones
- Identification of the cultivated domain

1 . Data acquisition and preparation



2 . Stratification of the territory into homogeneous agro-ecological zones



3. Identification of the cultivated domain

Figure 1: Flow-chart of processes

II ACQUISITION, ORGANIZATION AND PRE-PROCESSING OF DATA

II.1 Thematic maps: organization into a GIS

Within the context of this study, thematic maps will help, on the one hand, in the stratification of Senegalese territory into homogeneous agro-ecological zones, and on the other, in the analysis and comprehension of results obtained after the spatio-temporal classifications. In the context of information sharing, the data will be organized in a GIS and metadata will be created and managed within a catalogue.

We relied primarily on two sites that allowed the downloading of free GIS data on Senegal:

<http://www.roselt-oss.org/accueil.php>

<http://www.ibimet.cnr.it/Case/ap3a/index.htm>

A total of 35 layers of information on agriculture, climate, demography, environment, geology, hydrography, relief as well as on the infrastructure were downloaded. Only reliable data with clearly explained structure and originating from dependable sources were retained.

Thematic data was then sorted and organized in ArcCatalog©. Metadata was created with the Export MDWeb© tool installed in ArcCatalog© (cf. appendix 2). MDWeb is a tool used to create data catalogues. It is an ArcGIS© 8.x extension which allows automatic generation of metadata in the XML ISO 19115 format, conforming to MDWeb's template standards. Export MDWeb© is available to the community under the CeCill open license.

II.2 Satellite images

Choice of satellite images

We have chosen satellite images based on their potential for identifying the cultivated domain at the national scale and depending on a few practical considerations such as the cost or the ease of acquisition.

To start with, it seemed necessary to us to acquire images at low spatial resolutions and high repetitivity for exploring the possibility of distinguishing crops on the basis of the phenological evolution.

We chose the images from the MODIS/TERRA sensor which produces images at a spatial resolution of 250 m, as opposed to 1 km for its rivals, with a daily temporal resolution.

We also chose to acquire the Spot Vegetation temporal series because their coarser spatial resolution (1 km) will allow us to analyse vegetative activity on a larger scale. One of the advantages of Spot Vegetation is the availability of ten-day maximum value composites (MVC) virtually free of cloud cover. These two types of images are available for free on the Internet.

As far as high-resolution images are concerned, we have chosen to acquire Landsat ETM+ orthorectified images because some of them are available for free on the Internet.

MODIS TERRA Vegetation Indices 250 m data

From amongst the products offered by the MODIS sensor, we have selected the ‘Vegetation Indices 16-Day L3 Global 250m SIN Grid’ temporal syntheses for our study because they offer the best spatial resolution.

Spot Vegetation 1 km data

The VGT-D10 temporal syntheses were chosen for the study because the method used minimizes the noise on the temporal profiles thanks to better correction of directional effects. Moreover, the method facilitates comparisons between different dates because the data is in the same angular configuration (at the nadir).

Landsat ETM+ orthogeocover data

The Landsat ETM+ Orthorectified data is geometrically corrected taking into account the relief. This technique considerably improves the georeferencing quality.

II.3 Acquisition and preparation of satellite images

- 46 MODIS temporal syntheses of vegetation indices were acquired for free over the Internet, corresponding to a period of 2 years, from 21/03/2004 to 06/03/2006 (cf. table 5). To be able to read the native format of MODIS data and to reproject it into a standard projection such as UTM WGS84 Zone 28N (the projection used for Senegal), the MODIS Reprojection Tool (MRT) software, freely available on the [lpdaac](http://lpdaac.crk.usgs.gov/) site, was used. This tool also allowed us to spatially clip the MODIS images.
- 72 10-day Spot Vegetation images were acquired, equivalent to a 2-year period, from 01/03/2004 to 21/02/2006 (cf. table 5). They were downloaded for free from the <http://free.vgt.vito.be/> site.
- 12 Landsat ETM+ Orthogeocover images were downloaded to cover Senegal fully. They were directly downloaded in GeoTIFF format in the UTM WGS84 Zone 28N projection from the <http://glovis.usgs.gov/> site. No conversion of format or projection was thus necessary. Their acquisition dates are September, October, November 1999, 2000 and 2001.

Creation of temporal series

We thus had at our disposal 46 MODIS images with 7 bands (B, R, NIR, MIR, NDVI, EVI, NDVI quality) and 72 Spot VGT-D10 images with 6 bands (B, R, NIR, MIR, NDVI, BDC Status). For our study, NDVI as well as quality (NDVI quality and BDC Status Map) channels were needed. To facilitate the use of this data, it was necessary to reorganize the latter into temporal series.

For the MODIS/TERRA ‘Vegetation Indices 16-Day L3 Global 250m SIN Grid’ we have created 4 temporal series:

- one NDVI temporal series spanning 2 years
- one NDVI temporal series for the year 2004 (from 21/03/2004 to 06/03/2005)
- one NDVI temporal series for the year 2005 (from 22/03/2005 to 06/03/2006)
- one NDVI quality temporal series spanning 2 years

For the ‘Spot VGT-D10’ images, two temporal series were created:

- one NDVI temporal series spanning 2 years
- one MIR temporal series spanning 2 years

These temporal series then allow the display of temporal profiles.

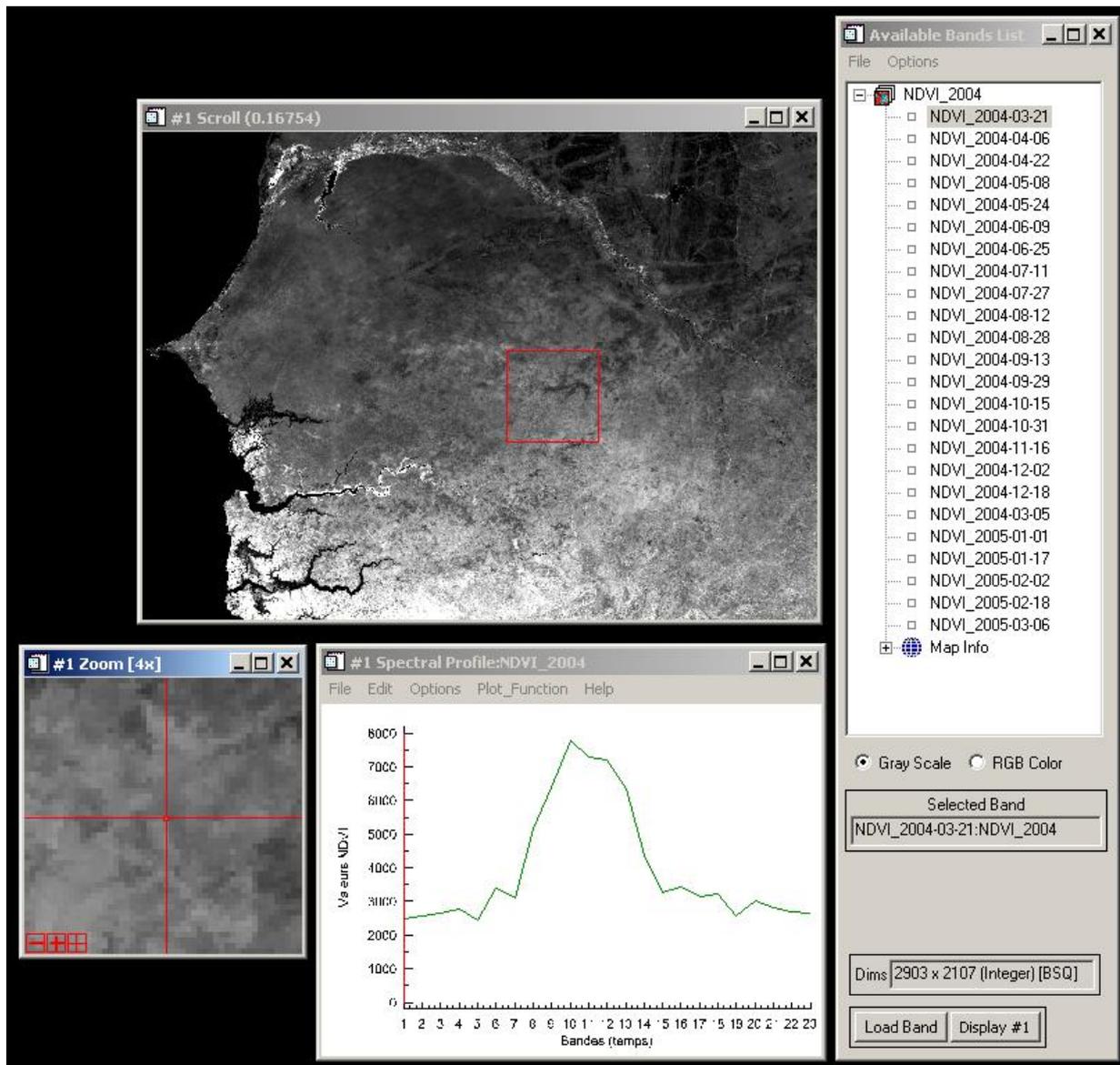


Figure 2: Display of the NDVI temporal profile of pixel 1765/1145 (centre cross) - MODIS NDVI 2004 temporal series

Evaluation of the quality of temporal series

On the NDVI temporal profiles, we observe abnormal drops in NDVI values. Several different reasons can contribute to this noise: atmospheric conditions, viewing geometry, etc.

It has therefore been found necessary to assess and quantify the quality of MODIS and Spot Vegetation images. To do this, the ‘NDVI quality’ channel of MODIS and the ‘BDC Status Map’ channel of Spot Vegetation have been used.

The quality assessment of MODIS images has revealed:

- a significant cloud-cover level between 25 June and 29 September 2004 and between 9 May and 30 September 2006. This level can reach 40% in the middle of the rainy season (July, August). The clouds are concentrated mainly in the south and west of the country.
- a significant correlation between the cloud cover and the mean, doubtful and low values of usage quality indices. This confirms the fact that atmospheric conditions form a significant parameter in the evaluation of the usage quality (cf. appendix 4).

A more in-depth analysis of the MODIS images allowed us to observe that the presence of cloudy zones translates into abnormal NDVI drops in the NDVI temporal profiles, imparting to the curve a saw-tooth aspect (e.g., figure 12).

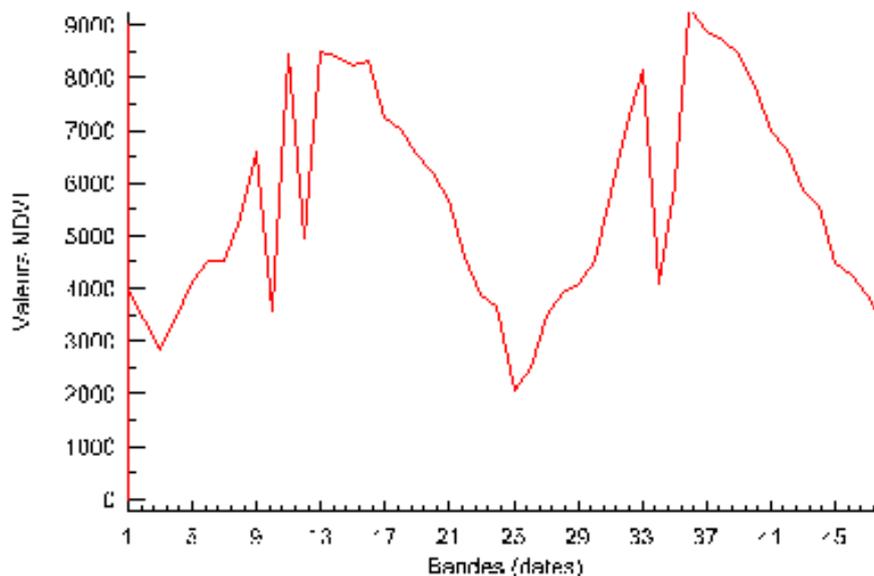


Figure 3: NDVI temporal profile of pixel 1730/1784

The cloud cover in the Spot Vegetation syntheses, on the other hand, is very low when compared to the MODIS syntheses (cf. figure 13). It attains a maximum of 10% for the ten-day period dated 01/08/2005. Just like for the MODIS images, this level increases during the rainy season, between the beginning of July and the beginning of September. The analysis of NDVI temporal profiles of these different pixel types has shown that only those pixels contaminated by clouds presented abnormal drops in NDVI.

Abnormal data thus highlighted should be identified for all the pixels of the temporal series and masked in such a way that it can be considered missing and thus calculated by estimation.

Reconstitution of missing data by interpolation

The first stage is to identify 'abnormal' pixels. To do this, we use the 'NDVI quality' channel for the MODIS images and the 'BDC Status Map' channel for the Spot Vegetation images.

The evaluation of the quality of MODIS and Spot Vegetation images has allowed us to recognize abnormal pixel as those with information code:

- 'mean' usage quality (MODIS)

- 'doubtful' usage quality (MODIS)
- 'low' usage quality (MODIS)
- 'clouds' production quality (MODIS)
- 'not produced because of poor quality' production quality (MODIS)
- cloud presence (Spot VGT)

The abnormal pixels are assigned a new arbitrary unique value which corresponds to missing data.

The second stage consists of creating new values for the missing data. This is done using a multi-date linear interpolation.

The principle is as follows:

- For the first date, for each pixel identified as abnormal, we take the first 2 values considered as correct for the same pixel on later dates, and we do a linear interpolation thus:

$$y = (y_2 \times (x_3 - x_2) + (y_3 - y_2) \times (x - x_2)) / (x_3 - x_2)$$

where

y is the value of the abnormal pixel

y₂ is the correct value of the same pixel on the first later date

y₃ is the correct value of the same pixel on the second later date

x is the date (or no. of the band) of the abnormal pixel

x₂ is the date (or no. of the band) of the correct pixel on the first later date

y₃ is the date (or no. of the band) of the correct pixel on the second later date

- For the second to the second penultimate date, for each pixel identified as abnormal, we take the value of the same pixel on the just preceding date and the first correct value of a later date and we do a linear interpolation using the same formula as above.
- For the last date, for each pixel identified as abnormal, we take the values of the pixel on the two preceding dates and we interpolate as above.

Tools

To reconstitute missing data, several tools have been created with IDL (Interactive Data Language), which is a scientific programming language specialized for data analysis. This is a language on which ENVI was developed. For this study, with so many images (72 Spot Vegetation and 66 MODIS), IDL proved very useful because it permitted the simultaneous analysis and processing of a large data set and thus the automation of data processing.

The different tools created for the interpolation of missing data have been included in the ENVI menu.

To conduct the first stage of the interpolation, i.e., the masking of 'abnormal' pixels, we have to use tools that we call 'modis_multimask' and 'Spot_multimask', using the one that corresponds to the data type we want to correct. This tool, whose use need not be restricted to this application, masks those pixels, for each band (or date), having values in the quality channel equal to the values specified by the user.

Once the NDVI temporal series is masked, we use the 'multi_interpolation' tool that applies on each masked pixel the multi-temporal linear interpolation method described above.

Interpolation results

As seen in the example for pixel 1335/1691 of the MODIS temporal series (cf. figure 16), the reconstitution of missing data using interpolation considerably improves the temporal profiles. The anomalies are removed and the profiles are smoothed.

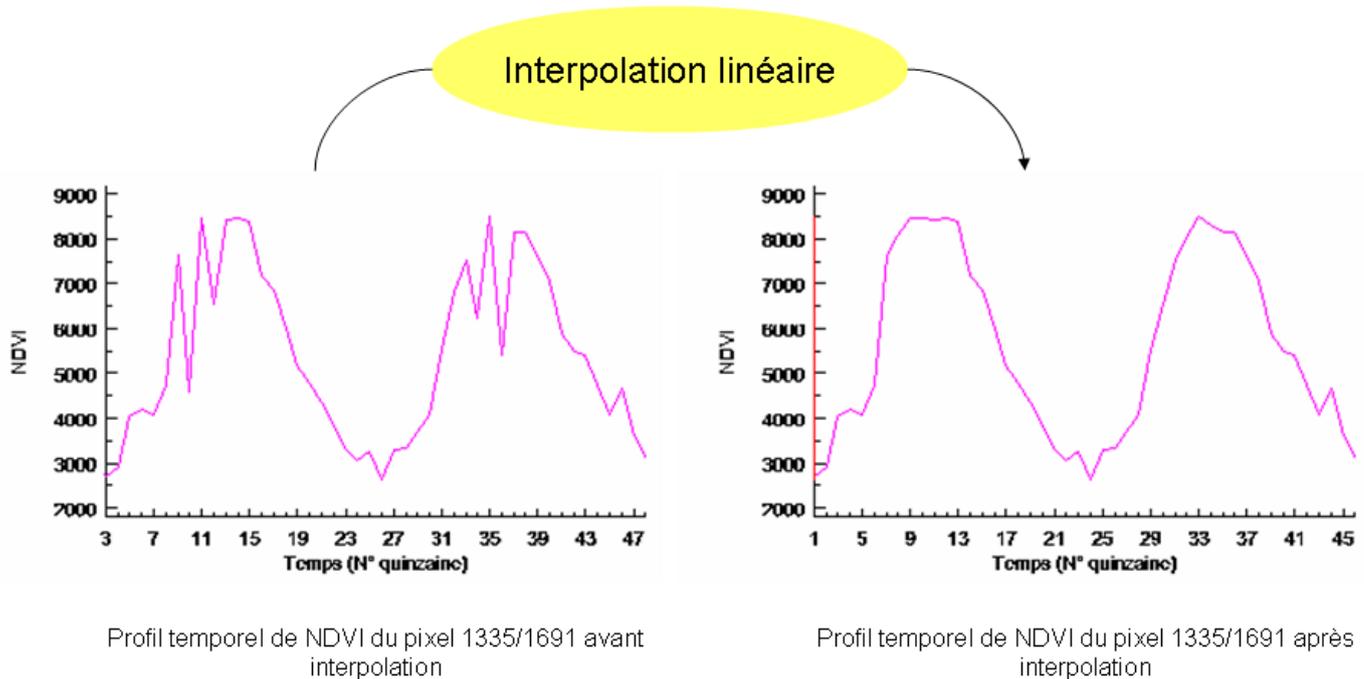


Figure 4: Temporal linear interpolation of pixel 1335/1691

Data thus prepared and corrected can now be processed and analysed. We have to now extract a maximum amount of information from these corrected temporal profiles and try to understand their significance. From the spatialised phenological information, derived from the temporal profiles, we will have to get spatialised ‘thematic’ information, i.e., associate to the MODIS and Spot VGT temporal classes a type of land use, by trying to identify one or several classes corresponding to crops.

III IDENTIFICATION OF THE CULTIVATED DOMAIN

Precise and consistent identification of the cultivated domain at a regional scale is especially important in arid and semi-arid countries, where the spatial and temporal variability of precipitation leads to large fluctuations in the primary production and thus to a significant famine risk (Maselli F., Rembold F., 2001).

The new remote-sensing techniques that allow observation of large territories with an improved temporal frequency are ideally suited for this environmental situation and to these countries’ cartographic needs.

But the identification of the cultivated domain over a large territory on the basis of low-resolution satellite data is complex and delicate. Two main problems are encountered in identifying the cultivated domain. The first arises from the fact that Senegalese territory is vast and very heterogeneous. It will, therefore, have to be stratified into homogeneous agro-ecological zones so that classifications can be adapted to it. The second problem is caused by the spatial resolution of MODIS data and the variability of natural environments. In reality, the cultivated areas are often intermixed with other vegetation types with different phenologies and which thus perturb the NDVI temporal series.

Therefore, in this part, we try to overcome these difficulties as best we can so that we can identify the cultivated domain in Senegal as precisely as possible.

III.1 Post-classification stratification

To identify the cultivated domain at the national scale, we will rely on spatio-temporal classifications. Several studies have shown that classifications created for a very heterogeneous environment lead to much confusion between different natural environments. A post-classification stratification of territory thus seems to be necessary to reduce spectral and temporal confusions between different types of land use. A new classification could then be applied within each of these stratified regions and the results then combined for the entirety of the region under study. This method, which we call ‘post-classification stratification’, improves the classification precision.

So in this part, we will identify homogeneous agro-ecological zones. To do this, a methodology which uses feedback in its processes and iterative refinement of the stratification will be applied (cf. figure 5). First, we use the thematic maps (pedology, vegetation, climatology, etc.) to perform an initial delimitation based on criteria to be explained in the following section. This delimitation will then be refined with the help of NDVI temporal profiles obtained from MODIS images and high-resolution images from Landsat ETM+ and Google Earth©. Then, the spatio-temporal classifications will be applied to each of the agro-ecological regions and, depending on the results and possible confusion between the classes, the stratification will be adjusted.

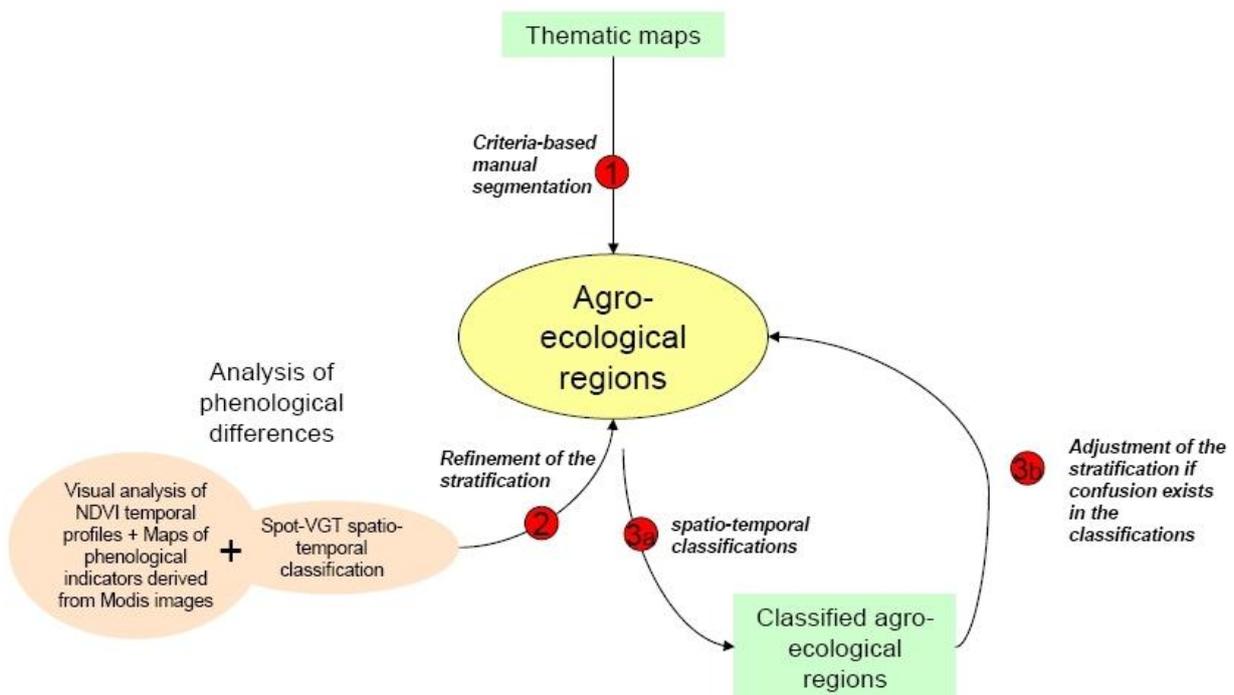


Figure 5: Model of processes for the delimitation into agro-ecological zones

Stratification on the basis of thematic maps

The first stage in the stratification process therefore consists of delimiting the entire Senegalese territory into several zones of similar agro-ecological criteria. To determine these zones, the methodology adopted is the one of superimposing information layers and cross

analysis. Two methods can be used to perform this cross analysis. The first consists of defining a certain number of criteria, of attributing a certain weight to them in the stratification and to perform calculations using SQL requests in a GIS. This rigorous method is of interest but is difficult to apply in this study because of the broad heterogeneity of the thematic maps as far as their sources, scales and production dates are concerned. Moreover, some maps that would be useful for stratification, such as the ethnic map, are available only in paper form and cannot be included in the calculations. Thus, it is the second solution, which consists of performing a visual cross analysis, that has been retained. As for the previous method, agro-ecological criteria have been defined and hierarchized as a function of their importance. Then the paper or digital thematic maps have been visually analysed to determine homogeneous zones based on the prescribed criteria.

This method is subject to criticism due to its subjective nature and thus uncertain repeatability but it allows inclusion of sources and complex criteria which cannot be calculated by GIS functions.

The criteria selected to conduct the visual stratification of the territory are as follows:

- the relief
- the geology and the soil
- the vegetation
- the duration of the seasons, the duration of the vegetative season and rainfall
- the ethnic groups and population density
- agricultural and animal production

The combination of these criteria has allowed us to define 7 agro-ecological zones. However, the radiometry and the temporal evolution of NDVI, criteria that will serve as the basis of classifications, have not yet been taken into account for the stratification. To minimize the confusion errors during classification, the agro-ecological zones have to be radiometrically and spatio-temporally homogeneous. What we have to do is to refine this first stratification by analysing the NDVI temporal profiles derived from the MODIS images, the spatio-temporal classification of the Spot VGT image and the high-resolution Landsat ETM+ and Google Earth© images.

Improving the stratification on the basis of phenological characteristics of crops

The low spatial resolution of the MODIS images does not allow us to visually distinguish crop zones because plots are not discernible. After all, to analyse the NDVI temporal profiles of crops, we have to first locate them. The high-resolution images from Landsat ETM+ and Quickbird of Google Earth© allowed us to identify crop zones and, by a simple geographical linking, it became possible to display the NDVI temporal profiles of the corresponding zones on the MODIS images.

The observation of NDVI temporal profiles derived from the MODIS images allowed us to highlight the differences in phenological behaviour of crops according to the zones. In the north-west, they seem to have a very short vegetative cycle (about a hundred days) with NDVI values that are on the low side (around 0.5). More to the south, the crops have a longer agricultural season, with the beginning of growth between end-May and beginning-June and the end of senescence around December. During the period of vegetative maturity, the vegetation indices reach much higher values: between 0.6 and 0.8.

It thus seemed indispensable to analyse more precisely these differences in the vegetative cycle and to take into consideration these phenological parameters in the stratification of Senegal.

But the zone-by-zone visual analysis of the crops' vegetative cycle is time consuming and exacting, and the delimitation of homogeneous zones based on phenology proves difficult. To facilitate this task, it seemed necessary to decompose and spatialise the phenological information by creating maps of phenological indicators and by establishing a spatio-temporal classification of the Spot VGT image for the entire country.

Definition of phenological parameters

To create maps of phenological indicators from the MODIS NDVI temporal series, we have to first of all define the different phenological parameters which will be mapped and determine a way of measuring them. In fact, indicators have been appropriately defined as certain parameter values of the NDVI curves, in such a way that a comparative analysis can be carried out on the basis of common references:

- The start of the cycle with photosynthetic activity: It is often easy to identify visually. It corresponds to a sudden increase of the NDVI followed by a period of steady and significant growth until it attains its maximum. Occasionally, however, significant variations in NDVI take place before the start of the cycle, during the dry season, but these do not correspond to the start of the cycle of photosynthetic activity.

Therefore, to the indicator of growth, we have to add an indicator of the continuity of growth.

The beginning of the growth cycle shall therefore be determined as follows: it is the first fortnight whose NDVI increases by more than $x\%$ between it and the next fortnight and whose mean growth rate for the following 3 fortnights is greater than $x\%$. x will be determined by the user after having analysed his data and the mean growth rate which determines the start. For the MODIS data, x is between 10 and 15%. The most satisfactory results are obtained with a level of 13%. If the program identifies no date that fulfils these conditions, the growth rate specified by the user, i.e. x , will be progressively decreased by 1 count, until the start of the cycle is identified.

- The end of the photosynthetic activity cycle: The end of the cycle is preceded by a steady and significant decrease in the NDVI from its maximum and is followed by a period of very slow decline, or even stability. It will therefore be calculated in the same way as the start of the cycle but starting with the last fortnight.

It is thus the first fortnight, starting from the last, whose NDVI has decreased by more than $x\%$ since the preceding fortnight and whose mean decreasing level over the previous 3 fortnights is greater than $x\%$. We take the same x as used to determine the start of the growth cycle, i.e., 13% for the MODIS data.

- Date and maximum NDVI value: These two parameters are easy to calculate. The maximum value of NDVI corresponds to the highest NDVI value recorded for the pixel over the year. And it is also straightforward to simply look up the date on which this maximum value was reached.

- Length of the growth period and of the increase in green biomass: This is the difference between the date of the start of the cycle of photosynthetic activity and the date the maximum NDVI was reached.
- Length of the period of decline/senescence of the biomass: This is the difference between the date of the maximum NDVI and the end of the cycle of photosynthetic activity.
- Growth/decline ratio: This is equal to the length of the growth period divided by the length of the decline period.
- Total duration of the cycle (growth, maturity, senescence): This is the difference between the dates of the end and the beginning of the cycle of photosynthetic activity.

Maps of phenological indicators

To measure and map these different phenological indicators, IDL tools for ENVI have been specially created.

Finally, eight maps per year were created. They allowed us to highlight the significant spatial differences and to readjust the stratification accordingly. After this second stratification stage, 13 agro-ecological regions have been identified.

Analysis of phenological differences

Observation of NDVI temporal profiles and the analysis of maps revealed information that is significant for the rest of the study:

- In the north-west of the country, vegetative activity, as represented by NDVI values, was much weaker in 2004 than in 2005. The NDVI temporal profiles in 2004 in some zones were 'flattened'. Mbaye Diop, scientist at the Laboratory for Geomatic Study and Research (LERG) at the Senegal Institute of Agronomic Research (ISRA), during a visit to CIRAD, advanced a possible reason for this: the locust outbreak during the 2004 farming season. This information is significant for the rest of the data processing because data for 2004 cannot be used for some zones.

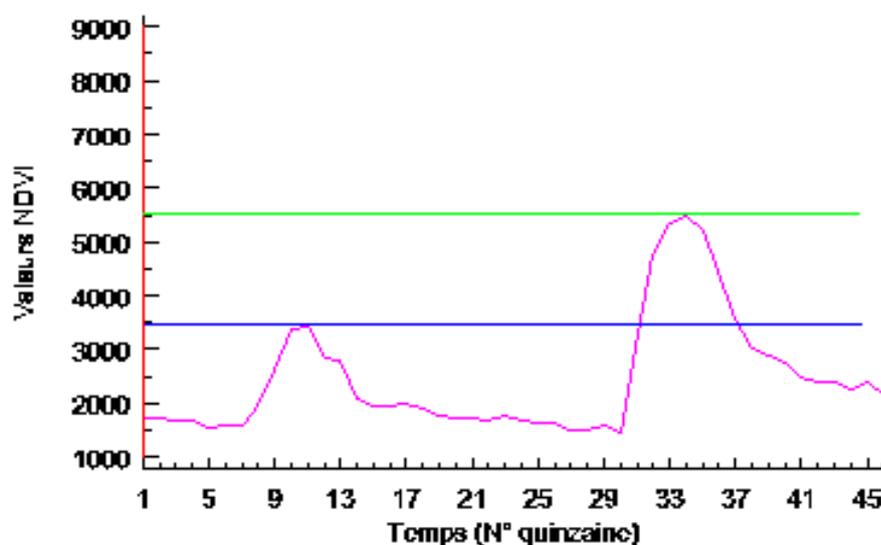


Figure 6: NDVI temporal profile for a pixel affected by the locust outbreak in 2004

- The photosynthetic activity, as represented by the maximum NDVI value and the length of the cycle of activity, is more significant in the south. This north-south gradient seems to be perfectly correlated to isohyets.
- The vegetative maturity, as represented by the date when the maximum NDVI was reached, is earlier in the east than in the west.
- The cultivated zones, notably the groundnut-grown area in the East and the bottom of small valleys, present growth periods that are longer than their periods of senescence, as expressed by the growth/decline ratios being greater than 1, and often greater than 1.5.

Classifications and fine tuning of the stratification

For each of the agro-ecological zones obtained by stratification, a spatio-temporal classification on the MODIS images was carried out. The spatio-temporal classification allows the grouping, within the same class, of all the pixels with similar temporal behaviour.

Each classification was then analysed briefly to detect possible confusions between different vegetation types caused by the heterogeneity of the environment. Depending on the level of confusion, the delimitation of the agro-ecological zone was modified to improve its homogenisation. It is this permanent iteration between stratification and classifications that allowed us to establish a definitive stratification of the Senegalese territory into 16 homogeneous agro-ecological zones.

III.2 Spatialisation, interpretation and significance of MODIS classes

During the final stratification stage, the spatio-temporal classifications were carried out to allow for adjustments of the delimitations based on the level of confusion. The method used to classify the temporal pixels is the classification by dynamic clustering. This technique groups pixels having the same temporal behaviour, characteristic of photosynthetic activity. From this spatialised phenological information, we have to derive spatialised ‘thematic’ information, i.e., associate to the each temporal class a plant formation and try to identify the classes that are representative of the cultivated domain.

Classification by dynamic clustering

The MODIS data set corresponds to a collection of ‘temporal pixels’ characterised by their own temporal profile. These pixels have to be classified and given a spatial representation, expressing the phenological evolution of different types of vegetation cover. The procedure chosen to classify ‘temporal pixels’ is the classification by dynamic clustering perfected by Viovy N., (1990). This is a fully automatic classification which does not require any external information. This method can delimit pixel groups having homogeneous temporal behaviours.

The number of classes for each agro-ecological region varied depending on the degree of the environment’s heterogeneity. For a relatively homogeneous agro-ecological region where the agricultural plots were clustered together, some twenty classes were sufficient. On the other hand, if the agro-ecological region presented several natural environments, or if the agricultural compartments were sparse or scattered and/or mixed with the natural vegetation, the number of classes doubled to some forty classes. This allows discrimination of a larger number of items.

A mean NDVI temporal profile corresponds to each class. It is from these mean profiles and high-resolution images that the classes corresponding to the cultivated domain will be identified.

Interpretation of classifications

At this stage of the study, 16 spatio-temporal classifications corresponding to 16 agro-ecological regions were identified. Now the classes have to be interpreted thematically and those amongst them that represent the cultivated domain identified. But this interpretation is extremely delicate because of the nature of the MODIS data. Since the spatial resolution of the MODIS images is 250 m, i.e., objects with a surface area of less than $250\text{ m} \times 250\text{ m}$, or $62,500\text{ m}^2$, cannot be identified. But agricultural plots in Senegal are normally smaller than 0.5 ha (5000 m^2). So some twelve plots have to be grouped together to make up a 'pure' crop pixel. Most often, however, the plots are grouped in islands around villages or at the bottom of small valleys. And, in reality, the plots are often mixed with natural vegetation or even fallow lands. The resulting pixels are mixed pixels, or 'mixels'.

It was thus decided to sort the classes into three categories: 'crop' classes, 'crops mixed with natural vegetation' classes and 'other' classes. We will consider as 'crops' all the classes consisting of more than 80% of crops. Based on this level, we assume that the natural vegetation has very little influence on the mean NDVI temporal profile. The 'crops mixed with natural vegetation' class represents the areas consisting of between 50% and 80% of crops. The remainder is represented by the natural vegetation. The 'other' classes consist of all other types of land use.

The allocation of each of the classes to one of these three categories required a visual comparison of classifications with the high-resolution images and an analysis of each class's temporal curves.

Using photo-interpretation to identify the classes representative of the cultivated domain

The method initially used to identify MODIS classes that are representative of the cultivated domain was the comparative analysis of spatio-temporal classifications and high-resolution Landsat ETM+ and Quickbird images.

This comparative analysis was conducted visually, by photo-interpretation. Visual photo-interpretation, as conducted in this study, allowed for logical deductions based on the interpreter's experience. For example, in the Sahelian regions such as in Senegal, it is possible to locate crop islands from the position of villages since the crops generally form circular islands around the villages.

The procedure used to allocate a category (crops, crops + vegetation, others) to each spatio-temporal class is therefore based on a simple geographical link between the MODIS classifications and the high-resolution Landsat ETM+ images. Using photo-interpretation, we could therefore determine the dominant natural environment in each class.

Criteria used to identify crops

Whether on the plains, on dunal formations, or on slopes and glacis, the agricultural plots are usually clearly seen on the high-resolution images. It is therefore possible to identify crops from their shapes.

The spatial environment criterion is also very important for identifying crops. As mentioned above, the crops are usually found around villages and it is often possible to

discern a star-shaped road network originating from a village. In Casamance and in the Kaffrine and Tambacounda regions, the crops are concentrated on slopes and glacis, along hydrographical networks.

The high-resolution images from Landsat ETM+ and Quickbird of Google Earth© were taken at the end of the rainy season and the beginning of the dry season. The crops were therefore at the end of senescence and the soil practically bare of vegetation. The soil, generally very reflective, therefore influences substantially the crops' spectral signatures, with these latter presenting a mainly light colour on the Quickbird high-resolution images. On the composite used for the photo-interpretation, a Landsat RGB NIR/R/G colour composite, they present a light cyan shade tending towards pale pink. In regions with more water, where the farming season is longer, crops are still in vegetative activity and present a mainly pinkish colour.

Most often crops present a smooth texture, but occasionally we observe striations and stripations which correspond to harvest residues (cf. figure 23).

The size of agricultural plots is highly variable and depends on the crop type and the mode of farming (family farm, marabout-owned farm, commercial production).

Finally, another method is to eliminate other types of plant formations that are easier to separate.

It is thus this set of criteria, added to the knowledge of how agriculture is organized in Senegal, that allows us to identify the crop zones and to allocate to the MODIS classes one of the three categories: 'crops', 'crops mixed with natural vegetation' and 'others'.

Analysis of temporal-behaviour curves

In parallel with the comparative analysis by photo-interpretation, an analysis of NDVI temporal profiles of each class was conducted. This analysis allowed us to overcome doubts in some cases and even to revise a decision made using photo-interpretation.

The first stage of the analysis was to understand the crops' phenological behaviour to try to identify the temporal profiles that could correspond to this behaviour or to eliminate those that did not.

In Senegal, as in all the countries in the Sahel, the crop development is highly dependent on precipitation. A rainfall event of more than 20 mm is sufficient to allow germination of seeds of rain-fed crops (millet, sorghum, groundnuts, cowpea, etc.). If the period immediately after sowing, is favourable, i.e., with good soil moisture, the crop can start growing (Soumaré P.B., Martini M.). Because crops require a greater amount of energy to germinate, the start of their vegetative development is normally slower than that of the surrounding natural vegetation. The subsequent period, that of growth, is generally steady until the maturity stage when plant production saturates. From that point on, the harvests start and plant activity drops off dramatically. The vegetation remaining after the harvest then declines slowly.

Based on this information, we can assume that rain-fed crops present an NDVI temporal profile resembling the profile type below (cf. figure 6): a first ascending phase mainly concave due to the initial slow growth, then rising more rapidly, then a second steady phase, followed by a third phase lightly convex due to the slowing of growth in the maturity period. After that comes the first rather steep descending phase after the harvest followed by, finally, the last stage of slow and steady decline.

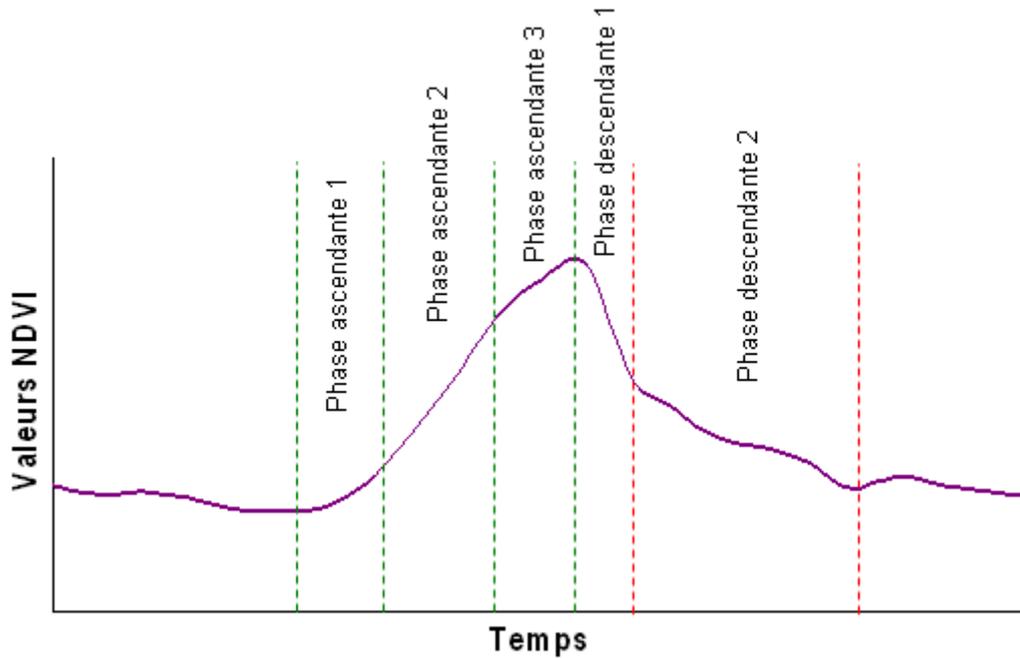


Figure 7: NDVI temporal profile type for a rain-fed crop

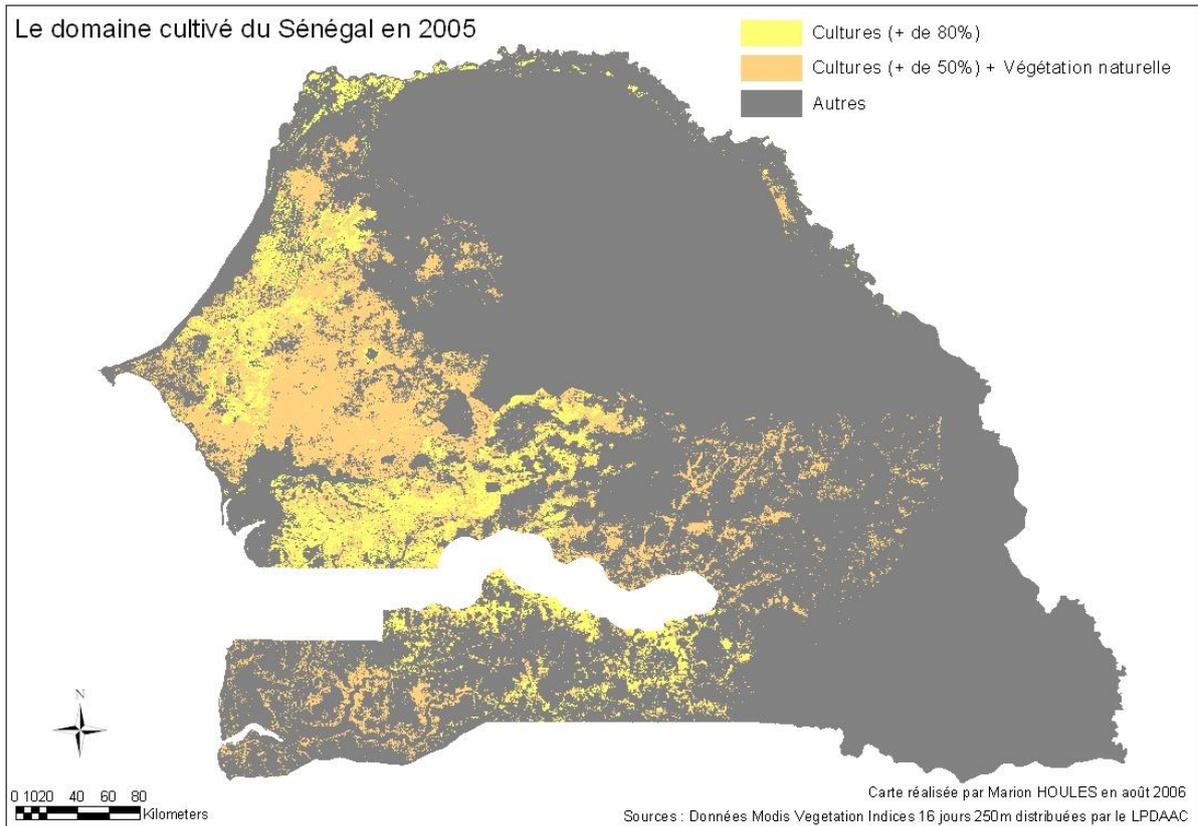
Based on this profile type, it was possible to eliminate all the classes represented by temporal curves very different from this one. However, we took care not to eliminate regions with irrigated crops or floodplain crops, which can present different profiles.

IV ANALYSIS AND VALIDATION OF RESULTS

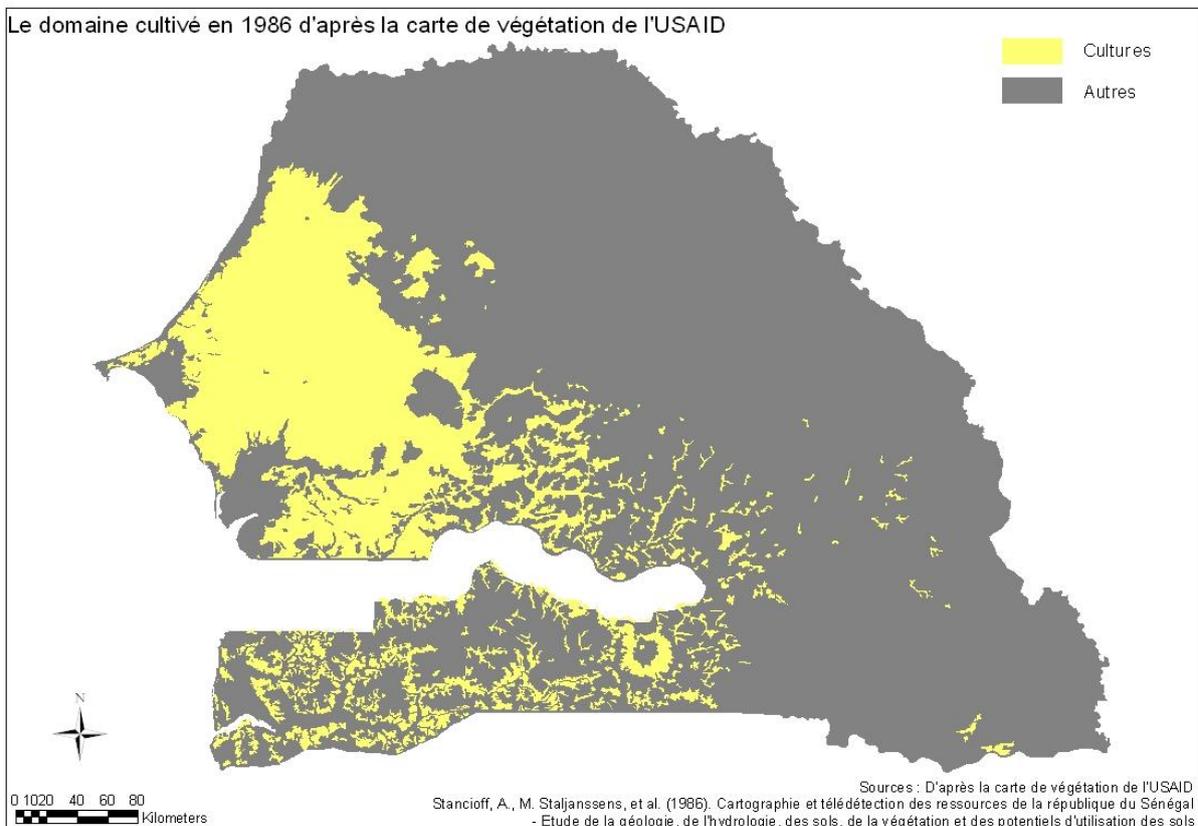
The identification of ‘crops’ and ‘crops mixed with natural vegetation’ classes allowed us to map the cultivated domain in 2005 (cf. map 2).

The photo-interpretation stage allowed us to get an idea of the possibilities of using low-resolution images for identifying the cultivated domain. On the whole, the results were very satisfactory at the national scale. However, we could observe that the results were more or less good depending on the zones. The method seems to be most suitable for regions where crops present large phenological differences with the surrounding natural environment. For example, in Casamance, where crops are concentrated at the bottom of small valleys and are surrounded by woodland and tree savannas, there is a marked difference between the phenological behaviour of crops and the savannas, and the crop identification is particularly good. On the other hand, in pastoral zones, where the crop islands are few in number and scattered within steppes and herbaceous prairies with phenological behaviour similar to that of crops, the results are less convincing. But given the relatively low area under crop cultivation in these zones, our inability to identify them definitely is not significant at the national scale.

Unfortunately, the only map we have of the cultivated domain at the national scale dates back to 1986; we do not have a more recent map. This map therefore is not a good basis to compare the results we have obtained. But we can nevertheless observe that, at the national scale, the similarities are many as far as the distribution of crops is concerned. However, if we analyse the two maps in some detail at the regional scale, and compare them to the high-resolution images, we can observe a much greater accuracy in the map generated by this present study.



Map 2: The cultivated domain in Senegal in 2005 based on spatio-temporal classifications derived from MODIS data



Map 3: The cultivated domain in 1986, from a USAID vegetation map

V ASSESSMENT AND PERSPECTIVES

First results have been very encouraging. The cultivated domain can be separated from other land-use types on the basis of its temporal behaviour. This constitutes a satisfactory improvement in the mapping of the cultivated domain; an annual update, at minimal cost, will then be possible. Moreover, tools developed in this study can be reused for other regions.

In perspective, it would be of interest to test other indices, such as the EVI, to see whether the results obtained with the NDVI can be improved. NDWI could also be considered; it allows the plant's water stress to be evaluated. Temporal series can be made with with all types of channels or indices and it would therefore be of interest to test as many as possible to evaluate their potential.

Since the crops' phenological behaviour is closely linked to rainfall, accurate climatic data would have been of immense help in understanding and identifying NDVI temporal profiles representative of crops. It should therefore be of interest to relate temporal rainfall data with the NDVI temporal series. Moreover, this would allow, for example, to significantly improve the identification of the start of the vegetative cycle, since the latter is so closely linked to the amount of precipitation.

At the same time, we need to better understand multi-temporal data at low spatial resolutions through its relationship with high spatial resolution images. While we have used a visual comparative analysis in this study, it should be interesting to explore a more automated method by measuring, for example, the degree of homogeneity of low-resolution pixels in terms of the proportions of the classes of the high-resolution images.

Finally, we must not lose sight of the fact that there will soon be available on the market images at a high temporal resolution with spatial resolutions approaching those of high-resolution ones. Data that will then become available will, in one go, facilitate in-depth research into the structure and dynamics of the cultivated domain and, more generally, into studies on the evolution of vegetation covers. It is therefore necessary to continue to use these new techniques and to introduce new spatial and temporal indicators.

REFERENCES

Bertrand, R., Killian, J., Raunet, M., Guillobez, S. et Bourgeon, G.
La connaissance des systèmes de paysages naturels, un préalable à la protection du milieu. L'approche morphopédologique.

(1985) Bull. Rech. Agro. Gembloux, 20(3/4), pp 545-559.

DOCUMENT TYPE: Colloque

Bruneau, M., Kilian, J.

L'apport des données satellitaires dans l'établissement de cartes utilisées par le développement rural

(1986) CNES SPOT Image, Colloque international Spot 1

DOCUMENT TYPE: Colloque

Gond, V

Contribution des données NOAA-AVHRR à la caractérisation des savanes africaines : approche locale, approche continentale.

(1995) Centre d'Etudes Spatiales sur la Biosphère, CESBIO, Toulouse, France, 151 page

<http://kourou.cirad.fr/umr/perso/gond/index.htm>

DOCUMENT TYPE: Thèse

Gond, V., Fontès, J., Loudjani, P.

Dynamique des biomes africains par l'analyse de séries temporelles satellitaires

(1997) Comptes Rendus de l'Académie des Sciences, Sciences de la vie / Life Sciences, pp 320: 179-188.

<http://kourou.cirad.fr/umr/perso/gond/index.htm>

DOCUMENT TYPE: Article

Houlès, M.

Indicateurs spatiaux et temporels pour l'identification du domaine cultivé en milieu soudano-sahélien

(2007) Université de Toulouse Le Mirail – Institut National Polytechnique de Toulouse, 85pp.

DOCUMENT TYPE: Mémoire de Mastère 2

Houlès, M.

Document techniques : i) Cahier annexe A : Note sur le téléchargement et l'utilisation d'images basse résolution gratuites : Modis / Terra et Spot Végétation, ii) Cahier annexe B : Glossaire des fonctions IDL utilisées – Codes et fonctionnement des outils IDL pour ENVI, iii) Cahier annexe C : Images Modis / Terra Vegetation Indices 16 jours 250 m – Présentation des images

Maselli, F., Rembold, F.

Analysis of GAC NDVI data for cropland identification and yield forecasting in Mediterranean African countries

(2001) Photogrammetric Engineering and Remote Sensing 67 (5), pp. 593-602. Cited 6 times.

<http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0035004435&partner=40&rel=R4.0.0>

DOCUMENT TYPE: Article

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., and Oho, H.

(2005). A crop phenology detection method using time-series MODIS data.

Remote Sens. Environ., 96, pp. 366-374.

DOCUMENT TYPE: Article

Soumaré, P.B., Martini, M.

La Veille environnementale au Centre de suivi écologique (CSE). Utilisation des images d'estimation des pluies dans le suivi de la campagne agricole.

DOCUMENT TYPE: Rapport

Tappan G.C., Sall M., Wood E.C., Cushing M.
Ecoregions and land cover trends in Senegal.
(Nov 2004), *Journal of Arid Environments* Vol 59 No 3, pp. 427-462.
DOCUMENT TYPE: Article

LINKS

MODIS:

<http://lpdaac.usgs.gov/main.asp>
<http://glcfapp.umiacs.umd.edu:8080/esdi/index.jsp>

Spot -Végétation:

<http://www.spot-vegetation.com/vegetationprogramme/Pages/TheVegetationSystem/userguide/userguide.htm>
<http://www.vgt.vito.be/>
<http://free.vgt.vito.be/>

Landsat:

<http://glcfapp.umiacs.umd.edu:8080/esdi/index.jsp> :
<http://glovis.usgs.gov/>

Senegal:

<http://www.senegalaisement.com>
<http://www.au-senegal.com>
<http://www.roselt-oss.org/coord/index.php?la=fra>
http://medias.obs-mip.fr/sud_sahel/
<http://www.ibimet.cnr.it/Case/ap3a/>
<http://www.cse.sn/>
<http://www.fao.org/>
<http://www.fao.org/giews/workstation/page.jsp>
<http://www.cartographie.ird.fr/SenegalFIG/index.html>

ACRONYMS

AMMA: Analyse Multidisciplinaire de la Mousson Africaine

CNES: Centre National d'Etudes spatiales

EVI: Enhanced Vegetation Index

HDF-EOS: Hierarchical Data Format - Earth Observing System

HRVIR: Haute Résolution dans le Visible et l'Infrarouge

IDL: Interactive Data Language

MODIS: Moderate Resolution Imaging Spectroradiometer

MRT: MODIS Reprojection Tool

NDVI: Normalized Difference Vegetation Index