CLIMATOLOGICAL CHARACTERISTICS OF NDVI TIME SERIES: CHALLENGES AND CONSTRAINTS

Stéphanie Horion*1, Bernard Tychon2 and Yves Cornet3

Abstract
Many studies already investigated the impact of climate change and climate variability on vegetation at global and continental scales. Low resolution satellite imagery is one of the main sources of information. In this paper, we describe a strategy to improve the quality of 10-daily time series of Normalized Difference Vegetation Index derived from SPOT-VEGETATION. A specific methodology was also defined in order to identify optimal test sites for the analysis of climate control on intra-annual dynamic of croplands. Finally two cases studies are presented to illustrate this research and in particular the non linear relationship between NDVI and meteorological parameters during the growing season.

Keywords
Vegetation dynamic, climate variability, low resolution imagery, geostatistics, time series analysis

INTRODUCTION

Many studies already investigated the impact of climate change and climate variability on vegetation at global and continental scales. Using time series of remote sensing and climate data, Nemani et al. (2003) analysed trends in Net Primary Production in relation with changes in climate and showed that, between 1982 and 1999, primary productivity increased by 6% globally in response to climate change. This study also stressed the need to take into account the spatial variability of climatic constraints on plant growth when analysing the climate change impact on vegetation. Others authors described different phenomena linked with climate change, such as increases of seasonal Normalized Difference Vegetation Index amplitude and of growing season duration in the Northern high latitudes (Myneni et al., 1998) or changes in circumpolar photosynthetic activities (Bunn et al., 2005).

Understanding the interactions between climate and vegetation is also a key issue in the PhD research of the main author (S. Horion) of this article. However, unlike other studies, we do not consider the vegetation globally but we focus on two specific types of vegetation: croplands and grasslands. One of our main objectives is to identify the meteorological factors that limit the development of croplands and grasslands in relation with their geographical localisation.

Because they provide both spatial and temporal data in large amount, low resolution satellites are customarily used as primary source of information on vegetation status and on meteorological conditions. However their coarse spatial resolution is greatly limiting their potential use. Indeed many vegetation types, land covers or natural processes can co-exist in a 1 km² pixel.

In this paper, we describe our strategy to improve the quality of 10-daily time series of Normalized Difference Vegetation Index derived from SPOT-VEGETATION. A specific methodology was also defined in order to identify optimal test sites for our research. Finally two
cases studies are presented to illustrate the non linear relationship between NDVI and meteorological parameters during the growing season.

2. DATA

2.1. Remote sensing Normalized Difference Vegetation Index

The Normalized Difference Vegetation Index (NDVI) is widely used as a proxy of the vegetation status and health. It has proven its efficiency in many research fields such as the monitoring of land cover changes and land degradation (Lambin et al., 2001; Lambin, 2000), the crop growth monitoring and yield forecasting (Balaghi et al., 2008; Zhang et al., 2005), the study of vegetation phenology and its evolution over time (White et al., 2006; Verstraete et al., 2007), the interaction between climate and vegetation (Yu et al., 2003; Vogt et al., 2000; White et al., 1997) and more specifically the impact of climate change on vegetation (Nemani et al., 2003; Zhou et al., 2001; White et al., 2005; Bunn et al., 1998).

The NDVI is derived from the red and the near infrared bands following the equation:

\[ \text{NDVI} = \frac{(\text{NIR} - \text{RED})}{(\text{RED} + \text{NIR})} \]  
\[ \text{eq. (1)} \]

Where NIR and RED are the amount of near infrared and red light reflected by the spectral object and recorded by the sensor (here: VEGETATION 1 and 2 sensors). This index ranges from -1 to +1 and relies on the absorption in the red part of the light spectrum due to the chlorophyll contained in the leaves, and on the scattering of the NIR by the mesophyll cells of the leaves. The more green and turgid are the leaves, the closer to 1 is the NDVI. A null value represents more or less the threshold between the presence and the absence of vegetation.

For our research, ten years of 10-daily SPOT-VEGETATION NDVI were acquired from April 1998 till May 2008 for the entire globe. The original image data set has a spatial resolution of 1 km². However, before using it for time series analysis, it went through a 3-steps procedure during which pixel values were aggregated to administrative entities (fig. 1a).

The first step is the pre-processing of raw images, i.e. geometric and atmospheric corrections and the computation of ten-days synthesis images, carried out by the Image Processing and Archiving Centre (CTIV, Mol) (Passot, 2000).

The next step is the computation of the NDVI regional unmixed statistics. Within a given region (fig.1a), pixels covered totally by cropland are used for the calculation of the regional NDVI mean and standard deviation (Eerens et al., 2004; Genovese et al., 2001) (fig. 2). As reference for the land cover types, we used the Global Land Cover 2000 (Bartholomé et al., 2005; Mayaux et al., 2006) generalized in 6 classes at the original 1 km² resolution, GLC2000-6C (fig. 1b). This aggregation step is really important as it allows us to focus on a certain type of vegetation with a specific phenology.

Finally the last processing step is the temporal smoothing of the time series. During this step residual noise due to mainly undetected cloud/snow is filtered out.

2.2. Global meteorological data set

The global meteorological dataset used in this study has been acquired freely through the FOODSEC portal of the Joint Research Centre (http://cidportal.jrc.ec.europa.eu/home/idp/thematic-portals/foodsec-imageserver/). Out of the twelve indicators available, six
were downloaded for this study (table 1). A seventh indicator was computed using the precipitation and the potential evapotranspiration. It corresponds to a proxy of the climatic water balance. Time series are composed by 10-daily global images at 0.5 degree resolution. They are derived from 2 different sources, (1) the ERA40 re-analyses (Uppala et al., 2005) for the period starting in January 1990 till August 2002 and (2) the operational ECMWF (European Centre for Medium-Range Weather Forecast) atmospheric model for the period from September 2002 till May 2008. Regional statistics of the 7 meteorological indicators were also extracted per administrative region following a similar procedure than for the NDVI images, i.e. they were calculated using strictly the values of grid cells covered totally by agricultural areas in the GLC2000-6C (fig. 1b).

Table 1. 10-daily meteorological indicators.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Unit</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum temperature</td>
<td>°C</td>
<td>Tmin</td>
</tr>
<tr>
<td>Maximum temperature</td>
<td>°C</td>
<td>Tmax</td>
</tr>
<tr>
<td>Mean temperature</td>
<td>°C</td>
<td>Tmean</td>
</tr>
<tr>
<td>Cumulated precipitation</td>
<td>mm</td>
<td>Rain</td>
</tr>
<tr>
<td>Cumulated potential evapotranspiration</td>
<td>mm</td>
<td>ET0</td>
</tr>
<tr>
<td>Cumulated global radiation</td>
<td>KJ/m²*day</td>
<td>Rg</td>
</tr>
<tr>
<td>Climatic water balance</td>
<td>mm</td>
<td>P-ETP</td>
</tr>
</tbody>
</table>

3. SITE SELECTION

Optimal test sites for the modelling of vegetation-climate interactions using the NDVI should ideally correspond to regions where the major part of the signal variation can be attributed to the climate and not to other phenomena such as land cover changes. For this reason our set of study cases has been established using a group of criteria which describe the spatial and temporal heterogeneities of the land covers and of the signal retrieved by the sensor (Horion et al., 2007). The criteria and their implementation strategies are defined hereafter.

3.1. Spatial heterogeneity of land covers

The spatial heterogeneity introduced by the occurrence of different land covers in a region is evaluated with 2 different indices, the Relative Area and the Fragmentation index, both extracted from the generalized GLC2000-6C. The relative area is calculated for each land cover class existing in the administrative regions and gives information on the relative importance of land covers within the region. This indicator is important to assess the reliability of the regional NDVI statistics. Indeed the unmixing process used for the extraction of these statistics already improves the quality of the dataset by considering only pixels completely covered by croplands. But the regional statistics might nevertheless not be reliable or representative of the regional vegetation state if they are derived from too few pixels.

The Fragmentation index refers to the spatial pattern in the surroundings of a pixel (3x3 window) (Eq. 2). The overall fragmentation of a region is evaluated by averaging the Fr values of each pixel contained in that region.

\[
F_r = \frac{(c_r - 1)}{(n_r - 1)} \quad \text{(eq. 2)}
\]

Where \(c_r\) is the number of different classes observed in a 3x3 window and \(n_r\) is the number of cells in the kernel (always 9 cells in this case).

3.2. Temporal heterogeneity of land covers (or land cover changes)

MODIS Land Cover products, MODIS LC, have been acquired in order to evaluate the error introduced by the use of a static land-cover map such as the GLC2000. These products are computed annually at a 1-km² resolution (Friedl et al., 2002). The temporal trajectory of each pixel evaluated with MODIS LC for the period 2001-2004 is analyzed in order to identify land cover changes. Aberrant behaviours are identified using simple logical rules and such pixels are excluded from the land cover change analysis. A temporal trajectory is considered as aberrant or incoherent if one of the following rules is not respected: (1) a maximum of two land cover changes is acceptable in the temporal trajectory of the pixel; (2) the ‘Shrublands’ class is the only accepted phase prior to ‘Forest’; (3) pixels with a temporal trajectory oscillating between water and another class are removed.

3.3. Spatial heterogeneity of the signal recorded by the sensor

The Coefficient of Variation CV measures the relative dispersion of a variable (Eq. 3). It is used here to esti-
mate how much the spectral signature of croplands is varying inside a same region.

\[ CV_r = \frac{1}{d} \sum_{i=1}^{d} \left( \frac{\text{Std}}{\text{Mean}} \right) \]  

(eq. 3)

Where \( \text{Std} \) and \( \text{Mean} \) are the standard deviation and average values calculated using the NDVI values of pixels covered by cropland and pertaining to region \( r \). \( d \) is the 10-daily period (decade). The final \( CV_r \) corresponds to the average of all 10-daily \( CV_r \).

4. RESULTS AND DISCUSSION

4.1. Optimal test sites for croplands

The analysis of the indicators presented in the previous section allowed us to identify candidate regions for our research. Figure 3a presents the relative area (RA) of cropland for our area of interest which comprises Europe and the part of Africa situated in the Northern Hemisphere. In Europe, regions mostly dominated by croplands (RA > 60%) are quite numerous and spread over the entire area, notably in northern France, in Germany, in Poland, in Ukraine and in Spain. By contrast, in Africa, they are more concentrated on the Sahelian belt. Like the relative area, the annual Fragmentation index per region \( Fr \) is also a good indicator of the spatial heterogeneity caused by mixed land covers. Most regions in our area of interest present a very low level of fragmentation, with \( Fr \) equal or below 0.125, meaning that in the 3x3 kernel (9 km²) only two different land covers are observed in average (fig. 3b).

The temporal trajectories of the MODIS LC images from 2001 to 2004 have been analyzed to identify incoherencies and to locate the pixels where land cover changes occurred from 2001 to 2004. The identification of stable regions is crucial as we are working on the response of vegetation to climatic events. Using regions with a high land cover change dynamic or incoherent changes would affect our analysis by introducing extra-variability in the vegetation response retrieved from the sensor. Croplands are probably the most stable land cover in Europe, followed by the forest (fig.3c). The agricultural regions of France (Nord Pas-de-Calais, Picardie, Ile de France, etc) have recorded few land cover changes as well as the Pô region, the Garonne plain and Ukraine. On the contrary, regions with prevailing cropland stability are almost absent in Africa. Indeed land cover changes are important especially in the cropland and shrubland areas. Very few African regions can be

Figure 3. (top left) Relative Area of cropland derived from GLC2000, (top right) Spatial fragmentation of land covers, (bottom left) Land cover changes between 2001 and 2004, (bottom right) Coefficient of variation of NDVI for cropland (in grey, no data).
selected as potential agricultural regions for our study. The coefficient of variation is globally low (CV < 0.2), except in some areas in the Middle East and in Egypt (fig.3d). This suggests that the signal recorded at the sensor for the pixels covered by croplands can be considered homogeneous at the region scale.

4.2. Two cases studies

After evaluation of the preceding results, 15 regions were selected as test cases for our research, 10 in Europe and 5 in Africa. In this section we present two test cases, one in Ethiopia (fig.4a) and one in France (fig.4b).

The first case study corresponds to the region of Amhara in the Northern highlands of Ethiopia. Its size is about 165,000 km$^2$, of which more than 70 % are covered by cropland. The Coefficient of Variation of NDVI for cropland and the land cover fragmentation are both quite small ($C{V}'_r = 0.22; Fr = 0.01$). However the percentage of stable cropland is rather low (about 11 %) and incoherent changes recorded between 2001 and 2004 by the MODIS Land Covers are important (12 %). Figure 5 (left) shows the typical landscape of Amhara. The agricultural fields are quite small (between 50 and 100 m in length) and mainly located on the plateau areas. These observations and the importance of the relief in this area can explain the relatively high score for incoherent changes.

The second case corresponds to the region of Picardie in northern France. Its size is about 20,000 km$^2$, of which more than 85 % are covered by stable croplands. The medium field size, much bigger than in Amhara, is about 500x100 m (fig.5b). Incoherent changes are negligible (< 2 % of the total area) and $C{V}'_r$ is also really low (~0.11), which indicates that the NDVI values recorded for all VEGETATION pixels covered by cropland in the region are similar.

Figure 6 presents the 10-daily evolution of NDVI for the two case studies in parallel with the 10-daily evolution of a limiting meteorological factor: 10-daily cumulated rainfall in the case of Amhara and mean temperature in the case of Picardie. The seasonality linked to the growing cycle of plants is obvious in both cases. In Ethiopia, a double crop cycle can be observed: from April to May-June (Bleg season) and from July to October (Meher season). In Picardie, the NDVI green-up is generally observed in March and the end of the growing season in October. The identification of a double crop season is less clear even though a small increase of the NDVI is often observed in August-September.
A good correlation between the NDVI cycle and the limiting meteorological parameter is also recorded in both regions. In Amhara, the rainfall increase precedes that of NDVI while in Picardie, cycles of NDVI and mean temperature are shifted in the opposite way. Scatterplots between NDVI and the limiting meteorological parameter show also interesting results regarding the type of relation between both variables (fig. 7a and 7b). We see clearly that fitting a linear regression line on the cloud of points is not a good solution. This demonstrates that a meteorological factor does not have the same impact on the NDVI evolution throughout the year and even through the growing season. In order to find statistically significant linear relation which is also meaningful from an agro-meteorological point of view, we need to split the growing season into separate (phenological) phases. For example, in the case of Picardie, a good linear relation can be found between March and June ($R^2 = 0.65$), which corresponds respectively to the greening phase and to the beginning of the maturing process for winter crops (USGS crop calendar, http://www.usda.gov/oce/weather/CropCalendars/index.htm, consulted in December 2009).

Following those preliminary results, a detailed investigation of interrelations between meteorological events and the vegetation response is currently undergone with specific focuses on (1) identifying key phenological stages within the growing season, (2) considering the cumulated impact of meteorological events (e.g. rainfall deficit during several decades, heat wave, etc) and (3) considering the possibility of a delayed response of vegetation.

5. CONCLUSION AND PERSPECTIVES

We have presented here some results of the PhD research of S. Horion concerning the impact of climate on croplands using two low resolution global datasets, respectively 10-daily NDVI estimated for cropland and 10-daily meteorological parameters combining ERA-40 and ECMWF data (table 1). A specific strategy was set up to improve these data-
Climatological characteristics of NDVI time series

sets and to focus only on agricultural areas. The Global Land Cover 2000 was used to identify VEGETATION pixels completely covered by cropland. These pixels were then used to compute regional statistics of NDVI and meteorological parameters.

The next step was the selection of optimal test cases for the analysis of climate control on croplands. Ideally, NDVI seasonal signal of such regions should be mainly influenced by the variation of meteorological parameters. Therefore regions with high fragmentation in landscapes, high land cover change dynamics and high spatial variation of NDVI with respect to the mean signal were excluded. The final choice was based on the percentage of stable cropland cover.

Further work will be dedicated to the analysis of cross-correlations between the monthly NDVI and the 7 meteorological parameters. As the two case studies showed, interactions between meteorological parameters and NDVI can not be modelled using a simple linear regression over the year or the growing season. Indeed they need to be studied at a time scale smaller than the growing season in order to identify properly the limiting factors on plant growth. Like the limiting factors, which are variable from a region to another, the time scale for the analysis needs also to be adjusted for each region, taking into account the phenology of the vegetation under consideration. Moreover, we also considered in our analysis the possibility of a delayed response of the vegetation and/or a cumulated effect of meteorological events (up to 3 months).

Acknowledgments

The PhD research of S. Horion was founded by the F.R.I.A (Fonds pour la formation à la Recherche dans l’Industrie et dans l’Agriculture). We would like also to thank the Agriculture Unit (European Commission, Joint Research Centre) and Herman Eerens (VITO) for providing us with regional unmixed statistics of SPOT-VEGETATION NDVI.

References


Lambin, E.F., Turner, I.L.B., Geist, H., Agbola, S., Angelsen, A., Bruce, J.W., Coomes, O., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E.F.,


