

EXTRACTION OF PHENOLOGICAL PARAMETERS FROM TEMPORALLY SMOOTHED VEGETATION INDICES

A. Klisch, A. Royer, C. Lazar, B. Baruth, G. Genovese

Joint Research Centre, Institute for the Protection and Security of the Citizen, Agrifish Unit, 21020 Ispra (VA), Italy
anja.klisch@jrc.it

Commission VIII, WG VIII/10

KEY WORDS: Agriculture, Remote Sensing, Vegetation Indices, temporal

ABSTRACT:

Within the MARS Crop Yield Forecasting System (MCYFS; Royer and Genovese, 2004) of the European Commission vegetation indicators like NDVI, SAVI and fAPAR are operationally derived for daily, decadal and monthly time steps. Besides low resolution sensors as SPOT-VGT and NOAA-AVHRR, medium resolution data from TERRA/AQUA-MODIS or ENVISAT-MERIS are used at pan-European level. In case of available time series, esp. NOAA-AVHRR (since 1981) and SPOT-VGT (since 1998) difference values of the indicators (e.g. relative or absolute differences) and frequency analysis of the indicators (e.g. position in historical range or distribution) are calculated. The exploitation of the data is performed at full resolution, at grid level of the MCYFS or regional unmixed means (C-indicators) are used. Therefore a database has been set-up in order to provide the indicators based on a weighted average for each CORINE land cover class within an area of interest.

The current work aims to develop a strategy for an optimal use of the different sensors and thus derived indicators at different aggregation levels for the ingestion into the MCYFS. As a first step smoothing algorithms have to be applied to the time series to diminish noise effects. Therefore, existing methods as simple sliding windows, piecewise linear regression or fitting of polynomial functions are employed and compared. Thereafter, the time series analysis is performed with the aim to establish relationships between indicators profile features and the crop phenology.

1. INTRODUCTION

Simple vegetation indicators like the Normalised Difference Vegetation Index (NDVI) or the Soil Adjusted Difference Vegetation Index (SAVI) are used since a long time to establish empirical relationships between spectral reflectance and biophysical parameters (e.g. biomass, leaf area or vegetation cover) or the spectral reflectance is used to estimate the amount of absorbed photo-synthetically active radiation (APAR).

Within the MARS Crop Yield Forecasting System (MCYFS) of the European Commission the indicators NDVI, SAVI, Dry Matter Productivity (DMP) and fraction of APAR (fAPAR) are operationally derived for daily, decadal and monthly time steps. Besides low resolution sensors as SPOT-VGT, NOAA-AVHRR and MSG-SEVIRI, medium resolution data from TERRA/AQUA-MODIS or ENVISAT-MERIS are used at pan-European level. In case of available time series, esp. NOAA AVHRR (since 1981) and SPOT-VGT (since 1998) difference values of the indicators (e.g. relative or absolute differences) and frequency analysis of the indicators (e.g. position in historical range or distribution) are calculated.

These vegetation indicators are intended to be employed at three ingestion levels :

- independent source of information to confirm crop growth indicators and identify area with anomalies,
- ingestion in the crop model to calibrate agrophenological parameters like leaf area index or development stages,
- crop yield estimation by means of scenarios and regression analysis.

According to the different ingestion levels, the exploitation of the data is performed at full resolution, at grid level of the Crop Yield Forecasting System (50 km) (Lazar and Genovese, 2004) or regional unmixed means (C-indicators) are used. For the latter a specific aggregation scheme has been set-up based on a weighted average for each CORINE land cover class (De Lima, 2005) within an area of interest as administrative boundaries (Genovese et al., 2001).

The overall objective of the study is to develop a strategy for an optimal use of the different sensors and thus derived indicators at pixel and aggregation level for the ingestion into the MCYFS. It can be separated into the following tasks:

1. Identification of an optimal strategy for the preparation of the satellite based time profiles (noise filtering and curves features extraction).
2. Identification of vegetation parameters to be used at the different ingestion levels.
3. Analysis of the impact of different sensors according to the spatial resolution and different time periods on the different ingestion levels.

The present paper covers mainly issue 1 and partly issue 2. As a first step smoothing algorithms have to be applied to the time series to diminish noise effects (e.g. clouds, snow, bidirectional effects). Therefore, existing methods as simple sliding windows, piecewise linear regression or fitting of polynomial functions are employed and compared. These methods have been reviewed by several authors recently (e.g. Pettorelli et al., 2005; Swets et al., 1992).

Thereafter the analysis of the time series is performed with the aim to extract features of the temporal profiles (e.g. start of growth, max NDVI, end of cycle) and to establish regional relationships with the biomass development.

2. METHODS

2.1 Smoothing of NDVI Time Series

The available remote sensing time series are pre-processed in terms of sensor calibration, degradation, atmospheric conditions and geometry. However noise still exists in the data and biases the data exploitation. It can be grouped as follows (Pettorelli et al., 2005):

- decrease of NDVI values due to clouds, water, snow, shadow, bidirectional effects,
- less frequent false highs due to high solar or scan angles, transmission errors.

In the case of NDVI time series, smoothing methods have to meet special requirements due to the fact, that most errors are decreasing the NDVI and the error structure can vary in time and space (Pettorelli et al., 2005). On the other hand they have to keep the general shape of the profile or the low frequency characteristics in terms of intensity and phasing. It should improve the biophysical relation of the NDVI profiles, which is an important precondition for the extraction of curve features and derivation of phenological parameters.

2.1.1 Smoothing Methods

A review of the most common smoothing methods as well as their main advantages and disadvantages is given by Swets et al., 1992; Pettorelli et al., 2005; Curnel and Oger 2007.

Within this study, the following methods have been analysed:

- Best Index Slope Extraction – ‘BISE’ (Viovy et al., 1992),
- Weighted least-squares linear regression – ‘SWETS’ (Swets et al., 1999),
- Modified Swets approach – ‘modSWETS’,
- Savitzky-Golay filter – ‘SAV-GOL’ (Press et al., 1992),
- ‘4253H, twice’ (Velleman and Hoaglin, 1981).
- Case sensitive filter – ‘CASE-SENS’.

The case sensitive filter detects different configurations of consecutive NDVI values, which are represented by lows, highs and slopes depending on the window size. It applies linear interpolation to remove decreased NDVI values.

The Swets approach performs a linear regression for a moving window. It includes the weighting of points according to their configuration (lows, highs and slopes). The resulting family of regression lines for each point is averaged. The modification of the ‘SWETS’ algorithm presented by ‘modSWETS’ has been introduced to avoid too high resulting values before and after NDVI slopes (correction of peaks) and to keep the original high values due to the fact, that false highs are less frequent. The principle of the corrections is shown in figure 1.

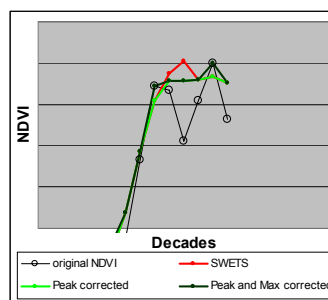


Figure 1. Principle of the modifications of the ‘SWETS’ approach (‘modSWETS’)

Moreover, the herein analysis was performed on different window sizes for each method as shown in table 1.

Methods	Time window size			
	3	4	5	6
BISE	x	x	x	x
SWETS	x	x	x	
modSWETS	x	x	x	x
SAV-GOL	-	x	x	x
‘4253H, twice’	- *	- *	- *	- *
CASE-SENS	x	x	x	-

Table 1. Overview of the analysed methods and time window sizes; * window size is predefined

2.1.2 Criteria for Evaluation of Results

In order to compare the different smoothing methods and the respective time window sizes, following criteria are defined, which are specific to the requirements given by the objective of ingestion in the MCYFS and the application to the pan-European level for arable land and grassland:

1. High frequency noise should be diminished to the highest possible degree.
2. Overall intensity given by an area-specific NDVI profile (low frequency feature of profile) has to be retained.
3. High values should be retained to highest possible degree due to the low frequency of false highs.
4. Timing of profile features, e.g. start of cycle (see 2.2.1) should not be influenced by the smoothing or to the lowest possible degree.

Furthermore, it should be outlined, that it is not clear for every situation, up to which extent the smoothing has to be performed in order to get the best signal/noise ratio. Thus, the final decision can only be based on expert knowledge through visual comparison, which is to a certain degree subjective.

Local characteristics can appear as unimodal or bimodal distributions within a cycle due to the cultivation of different crops, as different intensities of NDVI for different countries / areas or as local maxima before the real start of growth due to winter cover (see figure 2).

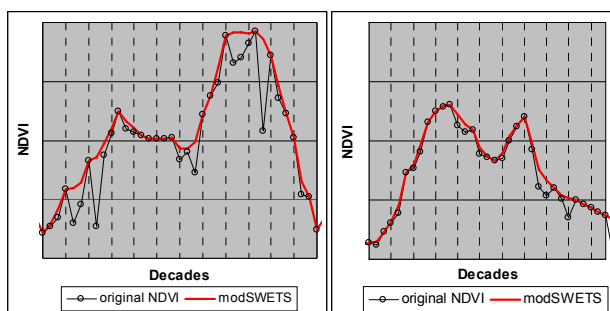


Figure 2. Examples for local profile crop characteristics in Germany with winter cover (left) and in Romania with bimodal NDVI profile (right)

2.2 Extraction and Derivation of Parameters

2.2.1 NDVI Profile Features/Characteristics

Typical features that describe the annual cycle of the NDVI profile are given below (Pettoelli, 2005; Joensson and Eklundh, 2004):

- date of start and end of growth,
- maximum NDVI cycle and date of maximum NDVI,
- rates of increase and decrease of NDVI within a cycle,
- cumulated NDVI: sum of NDVI for a given time period (e.g. for whole year, from start to end of cycle).

Moreover, the above mentioned profile features of the current year are compared with the one of the available historical years for the use of satellite based indicators as independent source of information. It is realised by calculation of differences or by ranking of the current year feature within the whole time series.

The extraction of the profile features can be performed by means of empirical equations (Moulin et al., 1997), thresholds (Joensson and Eklundh, 2004) or derivatives (Zhang et al., 2003). The latter method is possible, if the NDVI profiles are smoothed by polynomial fitting or logistic curves. Within this analysis an approach is employed, which detects increases of NDVI country-specifically.

2.2.2 Agrophenological Parameters for Calibration

The agrophenological parameters are meant to be directly coupled with the crop growth monitoring system (CGMS) to realise the second ingestion level. The CGMS is based on the WOFOST model adapted to the European scale and LINGRA used for pastures. The main processes being modelled within WOFOST are phenological development, CO₂-assimilation, transpiration, respiration, partitioning of assimilates and dry matter formation (Lazar and Genovese, 2004).

It is therefore the objective to derive several agrophenological parameters, which are used in the CGMS. These can be the leaf area index (LAI), above ground biomass or crop development stages. There are several approaches, mostly empirical ones, which link vegetation indicators (e.g. NDVI, SAVI) to the LAI or to the biomass (DMP) (e.g. Baret and Guyot, 1991; Turner et al., 1999). The linkage between NDVI and possible development stages, which has to be set-up region-specific, are shown in figure 3.

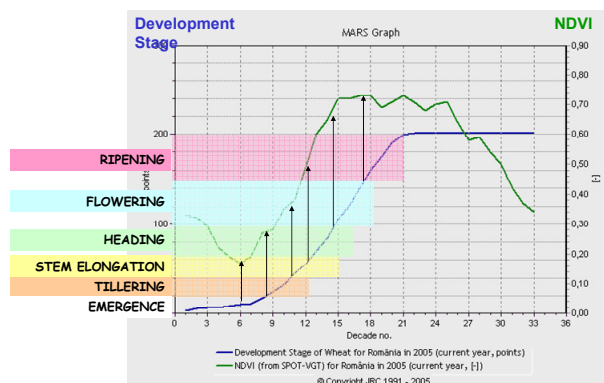


Figure 3. Regional linkage between NDVI profile and crop specific development stages

3. DATA BASES

The remote sensing data employed within the MCYFS for vegetation analysis are listed in table 2.

Platform	NOAA/ METOP	SPOT	ENVISAT	TERRA/ AQUA
Sensor	AVHRR 2 AVHRR 3	VGT 1 VGT 2	MERIS	MODIS
Pixel size (m)	1000	1000	300/1200	250/500
Existing from	1981	1998	2002	2000
MARS archive	1981	1998	2006	2004
Time period	d, t, m	t, m	t, m	d, t, m
NDVI	x	x	x	x
SAVI	x	x	x	x
fAPAR	x	x	x	x

Table 2. Overview of remote sensing data within MCYFS (d - daily, t - 10-daily, 30 - monthly)

The herein used data are decadal NDVI values calculated from the reflectances of the SPOT-VGT satellite. The pre-processing takes place in three steps (Royer and Genovese, 2004):

- spectral/radiometric: calibration and atmospheric correction; labelling of snow, cloud, sea,
- spatial: geometric correction to template with European extent and INSPIRE – Lambert Azimuthal Equal-Area Projection,
- temporal: synthetic images for different time periods by maximum value compositing (MVC).

The decadal NDVI profiles are corrected for snow and cloud values by linear interpolation between available values before and after snow or cloud occurrence. The extraction of arable land pixels for the analysis is carried out by means of the CORINE land cover 2000 (De Lima, 2005). The CLC data has been resampled to the extent and resolution of the SPOT-VGT data containing the percentage of class arable land according to the pixel size of 1km. Only pixels with more than 50% have been included in the analysis.

4. RESULTS

4.1 Comparison of Smoothing Methods

The method comparison has been carried out for sample pixel in Romania, Germany, France, Italy, Spain and Poland. In a

first step for each location all methods for each window size have been analysed. At this level a pre-selection of methods has been done. Figure 4 shows the results of the different methods using time window size 5 for Germany.

The methods are first checked according to their ability to reduce high frequency noise (see 2.1.2 criteria 1). As shown in figure 4, the approaches ‘4253H,twice’, ‘SWETS’ and ‘modSWETS’ fulfil this criteria very well. Whereas ‘SAV-GOL’ very often tends to insufficiently smooth the NDVI values. In contrary, the methods ‘BISE’ and ‘CASE-SENS’ appear to oversmooth the NDVI time series with increasing time window size.

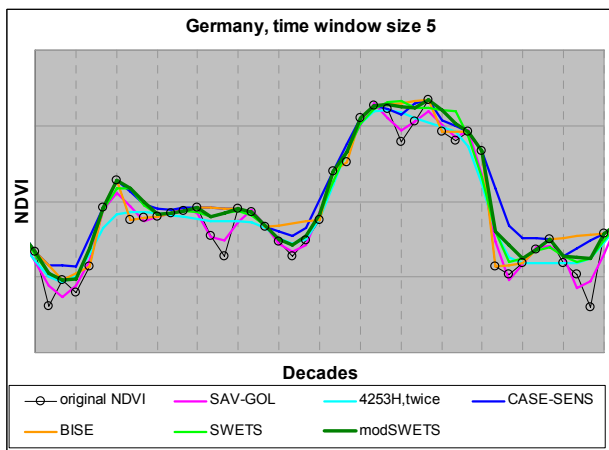


Figure 4. Smoothing methods for example Germany

The second criteria, which is keeping the overall intensity, is well fulfilled by all approaches. In figure 4 it is clearly visible, that the methods ‘SAV-GOL’ and ‘4253H,twice’ (magenta and turquoise) manage to smooth high frequency noise.

At the same time these methods tend to depress the intensity of the local maximum NDVI (criteria 3). This is due to the fact, that they are not adapted to the special situation that higher NDVI values are more reliable than lower ones. Whereas the other methods (‘CASE-SENS’, ‘BISE’, ‘modSWETS’) are better able to keep the maximum values. The ‘SWETS’ method (light green) seems to be a transition between these two groups of approaches, because it fits better in the NDVI situation due to its point weighting for the linear regression. But it is several times below the maximum values (see figure 4).

For the further analysis, the methods ‘CASE-SENS’, ‘BISE’, ‘modSWETS’ are taken into consideration. They are now compared for different time window sizes. In figure 5 the different behaviour for some critical decades are shown.

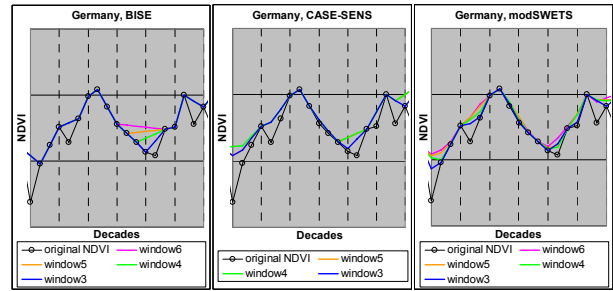


Figure 5. Pre-selected smoothing methods for different time window sizes

Considering the fourth criteria, the method ‘BISE’ (left) tends to smooth valleys too roughly depending on the window size and thus changes the phasing or timing of profile features like the start of cycle. A similar but less severe situation is given by ‘CASE-SENS’ (middle), whereas window size 5 behaves very similar to 4. The approach ‘modSWETS’ (right) apparently keeps the timing of the profile feature in a better way throughout all window sizes. Here it is of importance to find the optimal extent, up to which smoothing has to be performed.

A survey of the previous defined criteria and their fulfilment by the different methods is given in table 3.

Methods	Evaluation criteria (see 2.1.2)			
	1	2	3	4
BISE	xx	xx	xxx	x
SWETS	xxx	xx	xx	xx
modSWETS	xxx	xx	xxx	xx
SAV-GOL	x	xx	x	xx
‘4253H, twice’	xxx	xx	x	x
CASE-SENS	xx	xx	xxx	x

Table 3. Overview of the analysed methods and time window sizes; degree of fulfilment: x poor, xx good, xxx very good

4.2 Extracted Curve Parameters

As a first step, start of crop cycle and maximum NDVI value have been extracted from the NDVI time profiles. The SPOT-VGT data have smoothed with ‘modSWETS’ and a window size of 5 decades. The results are presented for a sample area in Central Europe in figure 6. Light to dark green colors indicate in most of the cases the start of regrowth for winter crops, which is initialised mainly between January and March. Light blue colours (April to June) are related to the start of growth for prevailing summer crops.

The above mentioned phasing behaviour of NDVI profiles after application of the smoothing methods can cause differences in the estimated start of growth. The impact of the differences strongly depends on the used time composite, which is in the present analysis 10-daily. It will be less severe for daily composites but more severe for monthly NDVI.

Start of Growth / Regrowth for Arable Land

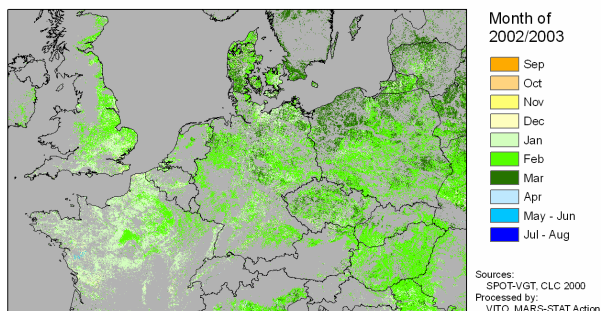


Figure 6. Estimated month of start of growth / regrowth based on 'modSWETS' smoothing for 2002/2003

Furthermore, in future the approaches will be applied to different indicators like SAVI and different sensors. According to the latter point medium resolution sensors (MODIS, MERIS) will play an important role to improve the mixed pixel situation.

5. CONCLUSIONS AND PROSPECTS

The ingestion of satellite based data into the MARS crop yield forecasting system requires an adequate preparation and approaches for parameter extraction. Within the present study different smoothing methods have been analysed. It could be shown, that existing methods are no panacea. Their limits mainly consist in the capability to adjust to local characteristics due to agricultural practices. Within this study the 'modSWETS' approach was selected as the best method according to the given requirements. It is an efficient approach in reducing high frequency noise, but it retains the general curve shape, intensity as well as high values. At the same time 'modSWETS' is able to retain the phase of curve features. Moreover it appears, that the time window sizes 4 and 5 present the best compromise.

The smoothed NDVI profiles have been successfully used to extract curve features like the start of growth and the maximum NDVI for European countries. These curve features have to be linked to agrophenological parameters contained in the MARS Crop Yield Forecasting System. In order to establish this link a number of ground truth data sources will be employed, which is available for main crops in Europe.

REFERENCES

- Baret, F., Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, 35, pp. 161-173.
- Curnel, Y., Oger, R., 2007. Agrophenology indicators from remote sensing: state of the art. In: *Workshop proceedings Remote Sensing support to crop yield forecast and area estimates. ISPRS Archives XXXVI-8/W48*.
- De Lima (ed), 2005. *CORINE Land Cover updating for the year 2000. IMAGE200 and CLC2000. Products and Methods*. EC, DG JRC, Institute for Environment and Sustainability, ISBN 92-894-9862-5.

Genovese, G., Vignolles, C., Nègre, T., Passera, G., 2001. A methodology for a combined use of normalised difference vegetation index and CORINE land cover data for crop yield monitoring and forecasting. A case study on Spain. *Agronomie*, 21, pp. 91-111.

Joansson, P., Eklundh, L., 2004. TIMESAT – a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30, pp. 833-845.

Lazar, C., Genovese, G. (eds), 2004. Agrometeorological data collection, processing and analysis. Methodology of the MARS Crop Yield Forecasting System. Vol 2, Publication Office of the EC, ISBN 92-894-8181-1.

Moulin, S., Kergoat, L., Viovy, N., Dedieu, G., 1997. Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements. *Journal of Climate*, 10(6), pp. 1154-1170.

Pettorelli, N., Vik, J.O., Mysterud, A. Gaillard, J.-M., Tucker, C.J., Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. *TRENDS in Ecology and Evolution*, 20(9), pp. 503-510.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. *Numerical recipes in Fortran77. The art of scientific computing*. Second edition. Vol. 1, Cambridge University Press, pp. 644-649.

Royer, A., Genovese, G. (eds), 2004. Remote Sensing information, data processing and analysis. Methodology of the MARS Crop Yield Forecasting System. Vol 3, Publication Office of the EC, ISBN 92-894-8182-X.

Swets, D.L., Reed, B.C., Rowland, J.D., Marko, S.E., 1999. A weighted least-squares approach to temporal NDVI smoothing. In: *Proceedings of the 1999 ASPRS Annual Conference*, Portland, Oregon, pp. 526-536.

Turner, D.P., Cohen, W.B., Kennedy, R.E., Fassnacht, K.S., Briggs, J.M., 1999. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, 70, pp. 52-68.

Velleman, P., Hoaglin, D.C., 1981. Applications, basics and computing of exploratory data analyses. Duxbury Press.

Viovy, N., Arino, O., Belward, A.S., 1992. The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *Int. J. Remote Sensing*, 13(8), pp. 1585-1590.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84, pp. 471-475.

