

PHENOSAT – A TOOL FOR VEGETATION TEMPORAL ANALYSIS FROM SATELLITE IMAGE DATA

Arlete Rodrigues, André R.S. Marcal, Mário Cunha

Centro de Investigação em Ciências Geo-Espaciais, Faculdade de Ciências, Universidade do Porto

ABSTRACT

The availability of temporal satellite image data has increased considerably in recent years. A number of satellite sensors currently observe the Earth with high temporal frequency thus providing a tool for monitoring/understanding the Earth-surface variability more precisely, for several applications such as the analysis of vegetation dynamics. However, the extraction of vegetation phenology information from Earth Observation Satellite (EOS) data is not easy, requiring efficient processing algorithms to properly handle the large amounts of data gathered.

The purpose of this work is to present a new, easy-to-use software tool that produces phenology information from EOS vegetation temporal data – PhenoSat. This paper describes PhenoSat, focusing on two new features: the determination of the beginning and maximum of a double growth season, and the selection of a temporal sub-region of interest in order to reduce and control the data evaluated.

Index Terms— Remote Sensing; Vegetation Index; Time-series; Phenology; Vegetation Dynamics.

1. INTRODUCTION

The number of satellite sensors with capability to acquire temporal image data has considerably increased in the last few years. The most important of these sensors are NOAA AVHRR, SPOT VEGETATION, ENVISAT MERIS, Metop-AVHRR and MODIS. With the increased number of sensors and improved data availability, it is now considerably easier to collect high temporal resolution image datasets, allowing for the monitorization of the Earth-surface to be done more frequently and with increased precision.

Phenological studies are very important to understand ecosystem functioning and associated seasonal patterns of carbon, water and energy fluxes, as well as to quantify crop response to climate variability and agronomic practices. The easier access to vegetation temporal data provides a continuous high frequency range of observations, and thus a more detailed crop and vegetation monitoring is now possible. However, the adequate processing of the large data volumes obtained is a difficult and time consuming process. Therefore, the development of software tools that allow for an efficient, semi-automatic processing is highly necessary.

There are a number of software tools available to process vegetation temporal data, such as TiSeG [1], TSPT [2], TIMESAT [3] and PPET [4]. These tools enable the extraction of phenology

information, but in general do not permit the extraction of phenology for a double growth season, neither a selection of the data interval to process.

The most common approach to extract vegetation parameters (phenological phases) uses deterministic models based on regression analysis. In [5] it is shown that the best methods were developed by [6] and by [7]. Although most of these models provide satisfactory parameters estimation, they present limitations when applied to different years, areas and vegetation types.

The main objective of this work is to present the software tool PhenoSat for the extraction of phenology information from vegetation indices temporal data, addressing the flaws encountered in the existing software packages.

This paper presents the methodologies used for processing temporal series of vegetation index imagery and the main features included in the software tool developed.

1.1. Data

The Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) are remote sensed vegetation indices (VI) commonly used in vegetation dynamics studies. NDVI is calculated from the reflectance in RED and near-infrared (NIR) spectral bands [8] using Eq.1.

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)} \quad (1)$$

$$EVI = G \cdot \frac{(NIR - RED)}{(L + NIR + C_1 RED + C_2 BLUE)} \quad (2)$$

The EVI improves upon the quality of the NDVI product, correcting some distortions in the reflected light caused by particles in the air as well as the ground cover below the vegetation. When viewing rainforests and other areas of the Earth with large amounts of chlorophyll, the EVI has the advantage that it does not become saturated as easily as the NDVI. EVI is calculated using Eq.2, where *NIR*, *RED* and *BLUE* are the surface reflectance in near-infrared, red and blue spectral bands, *G* is a gain factor, *L* is a canopy background adjustment term, and *C₁* and *C₂* are the coefficients of the aerosol resistance term, which use the blue band to correct for aerosol influences in the red band [9].

For agricultural and vegetation condition monitoring, clouds are partially screened from NDVI and EVI images by producing Maximum Value Composites (MVC) over a specific period. For this period, the NDVI (EVI) values are analyzed, on a pixel-by-

pixel basis, and the highest value is retained for each pixel location. When all pixels have been evaluated, a MVC image is obtained. In [10] it is shown that the MVC imagery is highly related to green vegetation dynamic, and the common problems present in single-date remote sensing studies, as cloud contamination, atmospheric attenuation and observation geometry are minimized using MVC.

VI satellite data are easily available and can be searched and ordered with reduced or no cost over the internet. These datasets are organized in two-dimensional images, with the time acting as a 3rd dimension in the data cube (Fig.1). For an observation t , for pixel location (j,k) , $N_t(j,k)$ gives the VI value at time t [11]. The set of observations $N_t(j,k)$, with $i=1, \dots, N$, is the time-series for pixel (j, k) .

Time-series of VI can be used to obtain information on seasonal vegetation development. This information is essential to understand the characteristics and behavior of vegetation dynamics, thus allowing a more precise knowledge and monitorization [12] [13].

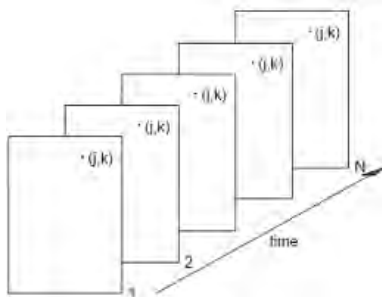


Fig.1. Vegetation indices satellite data organized in two-dimensional images.

2. METHODOLOGY

The software presented here – PhenoSat – was developed using Matlab [14]. It receives as input a VI temporal data cube (a set of VI images). It is important to understand that VI values are not only dependent of vegetation density and type, they are also influenced by observation geometry, atmospheric and illumination conditions, although the MVC technique used minimizes the influence of these undesired factors.

The final temporal VI data has nevertheless some noise due to the contribution of some parameters that cannot be fully eliminated by the MVC process. To reduce this noise and its influence in the results, a moving median filter [15] is after applied to remove rare events such as rapid shocks or other anomalies. Usually the application of the median filter is not sufficient to properly remove the noise present in temporal VI data. To remove occasional outliers, a Savitzky-Golay (SG) filter [16] can be applied. The SG filter coefficients are derived by performing an unweighted linear least square fit using a polynomial of a given degree. In general, higher degree polynomials can more accurately capture the heights and widths of narrow peaks, but perform poorly at smoothing wider peaks.

Field-based ecological studies have proved that vegetation phenology tends to follow relatively well-defined temporal patterns. The beginning of growth is followed by a rapid growth,

stabilizing in the maximum leaf area. The transition between senescence follows a similar trend, but in a reverse pattern. Thus, it is possible to represent the vegetation growth by a logistic function [17] [18], such as the period between senescence and dormancy.

Temporal variations of satellite NDVI data can be modeled by a double logistic function of the type presented in Eq.3, where t is the Julian day, VI_t is the VI value at time t , k is related to high asymptotical value of VI, c and d denote the slopes at the “left” and “right” inflection points, respectively, and p and e are the dates of these inflection points. VI_w and VI_{wl} are the VI values before the budbreak and during the dormancy after the leaf fall, respectively. This double logistic function is similar to others already developed ([17] [6]) but the k parameter insures the continuity between the vegetation growth and senescence parts, even when they differ in shape [19] [20].

$$VI_t = VI_w + \frac{k}{1 + \exp[-c(t-p)]} - \frac{k + VI_w - VI_{wl}}{1 + \exp[-d(t-e)]} \quad (3)$$

As an illustration, Fig.2 shows the NDVI values optimized after a double logistic function application and a schematic procedure for the calculation of the transition dates. As shown in Fig.2, it is possible to identify the principal phenological phases (Green-up (Gu), Maturity (M), Senescence (Se) and Dormancy (Dm)) from the rate of curvature change of Eq.3, using its minima and maxima. To identify the inflexion points (IP_L and IP_R) the global maximum and minimum of the curvature are used.

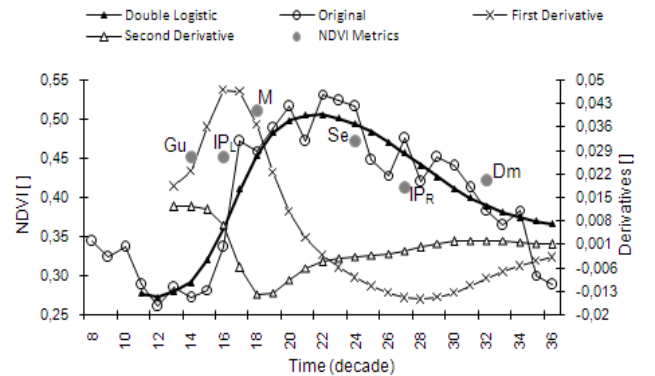


Fig.2. Schematic representation of NDVI Metrics calculated based on maxima and minima rate of curvature change. Global maximum and minimum of the curvature estimate the inflexion points.

$$\Delta = VI_{t+1} - VI_t \quad (4)$$

$$VI_{t+1} < VI_t \ \& \ VI_{t+2} < VI_{t+1} \quad (5)$$

The PhenoSat program permits to deal with double-crop growth season and/or vegetation re-growths during the seasons. To estimate the time occurrences of this re-growth, for a specific year N , the data after dormancy in year N and all data in year $N+1$ are considered. Eq.4 and Eq.5, where t is the time occurrence and VI_t is the VI value at time t , allow for the calculation of the beginning

of re-growth and its maximum value. The beginning occurs when $\Delta > 0$ for three or more distinct points and the maximum VI value occurs at a time t that verifies Eq.5.

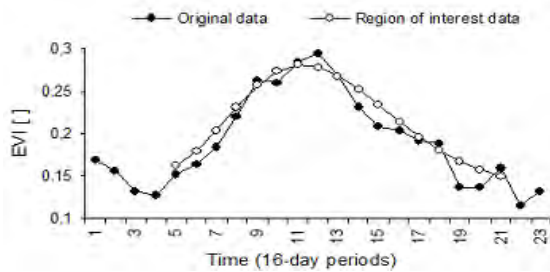


Fig.3. Example of an EVI temporal series for a year: original data and adjusted to a subset of the data.

Two more options are available in PhenoSat: the execution of an upper envelope [6] to enhance the spring and summer periods, and the selection of a region of interest allowing a reduction of the temporal data volume to be processed.

The region of interest can be selected manually by the user, or automatically by PhenoSat. The automatic selection of the temporal region of interest is done by PhenoSat, by first finding the maximum value of the data and searching for the point where a significant increase (or abrupt decrease) is verified at the left of that maximum. That point corresponds to the initial position of the region of interest. Afterwards, to determine the final position the program proceeds in a similar way, but evaluates the data at the right of the maximum (Fig.3). The data between initial and final positions are used to compute the phenological phases, with the remaining data being discarded.

The outputs of PhenoSat are two text (on MSExcel) files containing information about data processing (original and fitting data) and the vegetation dynamic information (phenology phases such as green-up, maturity, senescence, dormancy, and beginning and maximum of a double growth season if it occurs).

3. TEST RUN

PhenoSat was tested in a vineyard and semi-natural mountain meadows both in the North of Portugal. AQUA MODIS EVI satellite images, with 16-days temporal (23 images per year) and 250 meters spatial resolutions, were acquired from Land Processes DAAC User Services, USGS Earth Resources Observation and Science (EROS) Center. All available data for this site were used – 6 years (2003-2008).

A sub-image representing a specific test site in Régua (latitude 41°.167, longitude -7°.783) region was selected. The EVI values were extracted for all years for this sub-image, thus obtaining the EVI time-series.

As the EVI values are not distributed in a [0-1] range, the first step was to scale those values using the information provided by MODIS [21] image information. The VI data was exported to a text file table with 6 columns (years) and 23 rows (16 day MVCs). This text file was selected as input file in the PhenoSat interface. Median and SG filters were applied as an upper envelope which enhances the rise and fall of the curve.

Table1. Time occurrences (in 16-day periods) of the phenological phases obtained from EVI time-series in Régua vineyard region, between 2003 and 2008.

	2003	2004	2005	2006	2007	2008
Green-up	5.72	6.81	5.98	6.02	7.15	7.18
Inflexion1	7.63	8.14	7.26	7.94	8.32	8.68
Maturity	10.04	9.64	8.64	11.32	9.54	10.41
Senescence	11.54	11.73	11.15	12.84	14.46	12.21
Inflexion2	15.34	15.29	15.01	14.69	16.41	15.72
Dormancy	18.45	18.27	17.86	16.17	18.30	19.25
Maximum	11.2	11.04	10.46	11.73	11.63	11.68

To minimize the processing time and focus on the relevant data, a region of interest was defined with initial and final positions as 5 (March) and 21 (November). The 7 parameters of the double logistic function were estimated using the Levenberg-Marquardt method [22], and the result was used to fit the data in the interval (5-21). Finally, the fitted data obtained were used to compute the phenological phases (green-up, inflexion1, maturity, senescence, inflexion2 and dormancy).

The process described here was applied to all available years (2003-2008). Table1 presents the time occurrences of the phenological phases obtained from PhenoSat for all 6 years. The mean value of R-square, for all years, was estimated in 0.99 which proves the efficiency of the method to fit the EVI time-series used.

As the vineyard crop does not verify a double growth season in the same year, a semi-natural mountain meadows (“lameiros”) crop, in Northwest Portugal was selected for testing. In some years this type of crop verifies a re-growth, so it is ideal to test the ability of the software in its detection. If the re-growth does not occur in the same year, the value of the beginning and maximum refers to the following year, and the maximum of the re-growth in a year can be coincident to the maximum vegetation growth in the following year. Fig.4 shows the NDVI data acquired in 2006, 2007 and 2008 in a “lameiros” crop, where it is possible to see the values of re-growth obtained in two years (2006 and 2007).

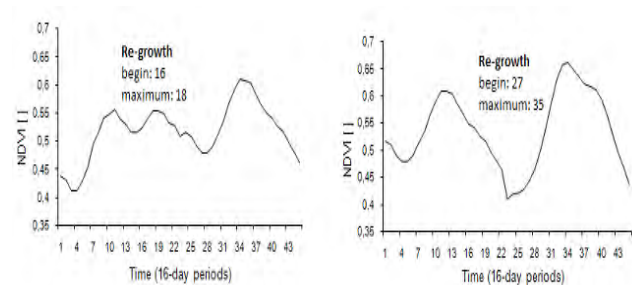


Fig.4. NDVI data after SG filter application, from “lameiros” crop in 2006 and 2007 (left) and in 2007 and 2008 (right), and the re-growth in 2006 and 2007, respectively.

To evaluate the Phenosat software and its results, different tests using various test sites, various VI time-series and various sensors were done, and a comparison was made between the phenology information obtained in the field and those computed by the software. The results in the field and obtained by PhenoSat were found to be very similar for vineyard [20] and for lameiros [23].

5. CONCLUSIONS

This work presents a new software tool (PhenoSat), developed in Matlab, which allows for information about phenology phases to be obtained from vegetation index (VI) temporal data, using an easy-to-use semi-automatic process. This software tool focuses in two new and relevant features: the ability to obtain information about a double growth season and the option available to select a temporal sub-region of interest.

PhenoSat was tested for different VI time-series acquired from different sensors, for different years and in different geographical areas. A test case using MODIS AQUA data was presented here. The results obtained from PhenoSat proved to be very similar to measurements made in field. Thus, satellite VI data processed by PhenoSat has high capability to predict phenology phases, and has the advantage of doing it in a short time and in a relatively simple way.

The PhenoSat software tool is available for public use freely at http://www.fc.up.pt/LamSat_XXI.

6. ACKNOWLEDGMENTS

Arlete Rodrigues has a PhD grant (SFRH / BD / 62189 / 2009) by FCT (Fundação para a Ciência e a Tecnologia, Portugal). This work is also part of the FCT project LamSatXXI, FCOMP-01-0124-FEDER-006996.

7. REFERENCES

- [1] R.R. Colditz, C. Conrad, T. Whermann, M. Schmidt and S. Dech, "TiSeG: A flexible software tool for time-series generation of MODIS data utilizing the quality assessment science data set", *IEEE Transactions on Geoscience and Remote Sensing* 46 (10), pp. 3296-3308, 2008.
- [2] R. McKellip, D. Prados, R. Ryan, J. Spruce, G. Gasser and R. Greer, "Remote sensing time-series analysis, a vegetation monitoring tool", *NASA Tech Briefs*, 2008.
- [3] P. Jonsson and L. Eklundh, "TIMESAT – a program for analyzing time-series of satellite sensor data", *Computers & Geosciences* 30 (8), pp. 833-845, 2004.
- [4] K.W. Ross, G. Gasser and B. Spiering, "Feasibility of estimating relative nutrient contributions of agriculture using Modis time-series", *NASA Technical Reports Server, Stennis Space Center, Mississippi*, 2008.
- [5] J.N. Hird and G. McDermid, "Noise reduction of NDVI time-series: An empirical comparison of selected techniques", *Remote Sensing of Environment* 113 (1), pp.248-258, 2009.
- [6] P.S.A. Beck, C. Atzberger, K.A. Hogda, B. Johansen and A.K. Skidmore, "Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI", *Remote Sensing of Environment* 100 (3), pp.321-334, 2006.
- [7] P. Jonsson and L. Eklundh, "Seasonality extraction by function-fitting to time-series of satellite sensor data", *IEEE Transactions on Geoscience and Remote Sensing* 40 (8), pp.1824-1832, 2002.
- [8] A. Huete, C. Justice and W.V. Leeuwen, "MODIS vegetation index (MOD13): algorithm theoretical basis document", 1999.
- [9] Z. Jiang, A.R. Huete, K. Didan and T. Miura, "Development of a two-band enhanced vegetation index without a blue band", *Remote Sensing of Environment* 112 (10), pp.3833-3845, 2008.
- [10] C.J. Tucker, J.U. Hielkema and J. Roffey, "The potential of satellite remote sensing of ecological conditions for survey and forecasting desert-locust activity", *International Journal of Remote Sensing of Environment* 6 (1), pp.127-138, 1985.
- [11] J.W.J. Rouse, R.H. Haas, J.A. Schell and D.W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS", *Third Symposium, NASA SP-351, Washington DC*, pp.309-317, 1974.
- [12] C.J. Tucker, C.O. Justice and S.D. Prince, "Monitoring the grasslands of the Sahel 1984-1985", *International Journal of Remote Sensing* 7 (11), pp.1571-1581, 1986.
- [13] J.P. Malingreau, "Global vegetation dynamics: satellite observations over Asia", *International Journal of Remote Sensing* 7 (9), pp.1121-1146, 1986.
- [14] D.J. Higham, N.J. Higham, *Matlab Guide*, SIAM, Philadelphia, 2000.
- [15] P. Gonçalves, H. Carrão and M. Caetano, "Parametric model for intra-annual reflectance time series", *Proceedings of the 26th EARSeL Symposium, New Developments and Challenges in Remote Sensing*, edited by Z. Bochenek, Millpress, pp.367-375, 2007.
- [16] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing, third edition*, pp.766-772, 2007.
- [17] X. Zhang, M.A. Friedl, C.B. Schaaf, A.H. Strahler, J.G.F. Hodges, F. Gao, B.C. Reed and A.R. Huete, "Monitoring vegetation phenology using MODIS", *Remote Sensing of Environment* 84 (3), pp.471-475, 2003.
- [18] D. Villegas, N. Aparicio, R. Blanco and C. RoYo, "Biomass accumulation and main stem elongation of durum wheat grown under Mediterranean conditions", *Annals of Botany* 88 (4), pp.617-627, 2001.
- [19] A. Fischer, "A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters", *Remote Sensing of Environment* 48 (2), pp.220-230, 1994.
- [20] M. Cunha, A.R.S. Marçal and A. Rodrigues, "A Comparative Study of Satellite and Ground-Based Vineyard Phenology", *Proceedings of the 29th EARSeL Symposium, Imagin [e,g] Europe*, pp.68-77, 2010.
- [21] MODIS Website, <https://lpdaac.usgs.gov/lpdaac/products/> (available at 27May2011).
- [22] D.W. Marquardt, "An algorithm for least-squares estimation of non-linear parameters", *Journal of the Society for Industrial and Applied Mathematics* 11 (2), pp.431-441, 1963.
- [23] M. Cunha, I. Poças, A.R.S. Marçal, A. Rodrigues and L. Pereira, "Evaluating MODIS vegetation indices using ground based measurements in mountain semi-natural meadows of Northeast Portugal", *Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS)*, pp.1525-1528, 2010.