

# Monitoring Vegetation Dynamics Inferred by Satellite Data Using the PhenoSat Tool

Arlete Rodrigues, *Student Member, IEEE*, Andre R. S. Marçal, and Mário Cunha

**Abstract**—PhenoSat is an experimental software tool that produces phenological information from satellite vegetation index time series. The main characteristics and functionalities of the PhenoSat tool are presented, and its performance is compared against observed measures and other available software applications. A multiyear experiment was carried out for different vegetation types: vineyard, low shrublands, and seminatural meadows. Temporal satellite normalized difference vegetation index (NDVI) data provided by MODerate resolution Imaging Spectroradiometer and Satellite Pour l’Observation de la Terre VEGETATION were used to test the ability of the software in extracting vegetation dynamics information. Three important PhenoSat features were analyzed: extraction of the main growing season information, estimation of double growth season parameters, and the advantage of selecting a temporal region of interest. Seven noise reduction filters were applied: cubic smoothing splines, polynomial curve fitting, Fourier series, Gaussian models, piecewise logistic, Savitzky–Golay (SG), and a combination of the last two. The results showed that PhenoSat is a useful tool to extract NDVI metrics related to vegetation dynamics, obtaining high significant correlations between observed and estimated parameters for most of the phenological stages and vegetation types studied. Using the combination of SG and piecewise logistic to fit the NDVI time series, PhenoSat obtained correlations higher than 0.71, except for the seminatural meadow start of season. The selection of a temporal region of interest improved the fitting process, consequently providing more reliable phenological information.

**Index Terms**—Normalized difference vegetation index (NDVI), phenology, PhenoSat, time series.

## I. INTRODUCTION

**T**EMPORAL vegetation profiles based on remotely sensed data provide valuable information for understanding land cover dynamics, generally interpreted by vegetation phenological events. Over the last years, the number of remote sensing sensors capable of acquiring phenological appropriate temporal data has been increasing considerably. Most of these sensors

[Advanced Very High Resolution Radiometer, Satellite Pour l’Observation de la Terre (SPOT) VEGETATION (SPOT VGT), MODerate resolution Imaging Spectroradiometer (MODIS), and MEdium Resolution Imaging Spectrometer] provide data related to a specific vegetation index (VI), such as the normalized difference vegetation index (NDVI) or the enhanced VI [1]. The VI temporal profile (time series) can be analyzed through time and used to extract information about the vegetation dynamics. Currently, access to satellite vegetation data is widespread, with low or no costs. However, the analysis and extraction of relevant information can be a difficult and time-consuming process due to the large amounts of data and the presence of noise.

The VI time series obtained by Earth Observation Satellite (EOS) images generally include various noise components such as atmospheric disturbances, viewing and solar illumination variability, cloud cover, and others. In the SPOT VGT (ten days) and MODIS (16 days) composites, some of this noise is reduced by the maximum value compositing process [2], where only the highest VI value in a predefined period (10 or 16 days according to the sensor) is retained. However, additional noise may be also introduced by the process of overlaying several images (for example, due to image registration). Noise reduction or fitting a model to the observed data is thus necessary before the extraction of vegetation dynamics information. It is important to carefully evaluate the time series and the noise present in order to choose the best fitting algorithm, one capable of smoothing the data without introducing artifacts or suppressing natural variations of vegetation [3]. Various time-series fitting algorithms have been developed [2], [4]–[10], and comparison studies have been conducted [11], which concluded that asymmetric Gaussian (AG) [6] and double logistic functions [4] present the best results.

During the last years, more fitting algorithms (e.g., [12]–[14]) and software products were developed to extract phenology from satellite VI time series. These software include TIMESAT [7], TimeStats [15], enhanced TIMESAT [16], PPET [17], and the software developed by United States Geological Survey Earth Resources Observation and Science Center [18]. All these software tools have important functionalities for the extraction of phenological information, but none of them allow the selection of an in-season window of interest, which is fundamental for application to crops with discontinuous canopy. Moreover, except for TIMESAT, none of them have a specific option to determine a double growth season phenology. To address these limitations, the authors of this paper developed PhenoSat, a new tool that is flexible to detect the number of growth seasons in each year and has an option that permits

Manuscript received February 29, 2012; revised July 12, 2012 and September 12, 2012; accepted September 30, 2012. Date of publication November 26, 2012; date of current version March 21, 2013. This work was supported by the Fundação para a Ciência e a Tecnologia under Doctoral Grant SFRH/BD/62189/2009.

A. Rodrigues and M. Cunha are with the Department of Geosciences, Environment and Spatial Planning and the Centro de Investigação em Ciências Geo-Espaciais, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal (e-mail: dr.arlete@gmail.com; mcunha@mail.icav.up.pt).

A. R. S. Marçal is with the Department of Mathematics and the Centro de Investigação em Ciências Geo-Espaciais, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal (e-mail: andre.marcal@fc.up.pt).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2012.2223475

to define, manually or automatically, an in-season window of interest.

The aim of this paper is to evaluate the capability of NDVI metrics obtained from PhenoSat to identify the phenology *in situ*. Different NDVI time series provided by MODIS and SPOT VGT satellite sensors, from different geographical locations and vegetation types, were used to address the objectives of this paper. The following section describes PhenoSat functionalities and insights into the algorithm performance. Sections III and IV present an experiment carried out to test the PhenoSat ability to detect accurate phenology with various vegetation types, and Section V presents the concluding remarks.

## II. PHENOSAT SOFTWARE DESCRIPTION

### A. Functionalities and Implementation

PhenoSat is a simple-to-use tool that extracts satellite VI metrics related to vegetation phenology. PhenoSat was implemented in Matlab, receiving as input a temporal VI data set and as outputs the phenological information and the data from the various fitting steps.

PhenoSat extracts a number of phenological parameters for the main growth period such as the following: start of the season, maximum vegetation growth, or senescence. Some vegetation types can have a regrowth in the same year, which is related to some factors such as the grass-cut-animal pastures or particular weather conditions. The beginning and maximum of the regrowth period can be determined by PhenoSat.

Some VI data sets available online (e.g., MODIS and SPOT VGT) are already preprocessed in order to remove many of the disturbances provided by different factors such as atmospheric conditions or geometry and illumination variability [19]–[22]. Although this preprocessing is generally effective, the VI data sets still retain enough problems (punctual outliers or abrupt changes) that require additional processing. The elimination of these artifacts can be achieved by the application of noise reduction filters, which permits the researcher to conduct a better subsequent analysis and to obtain more reliable vegetation dynamics information.

Initially, the VI time-series values that are substantially different (with an NDVI difference above 0.2) from the left and right spatial neighbors, and from a median of a window ( $M_w$ ), are considered outliers, and their values are replaced by the  $M_w$  value. PhenoSat has an option that permits the application of an upper envelope [4], enhancing the spring and summer periods. Despite these actions, some noise might still remain in the time series. Further improvements in the analysis can be obtained by PhenoSat using seven smoothing algorithms: cubic smoothing splines (CSSs), polynomial curve fitting (PCF), Gaussian models (GMs), Fourier series (FS), piecewise logistic (DL), Savitzky-Golay (SG), and, also, the sequential combination of the SG and DL.

The CSS algorithm fits a spline or smooth piecewise-polynomial algorithm, and a smoothing parameter determines just how closely the smoothing spline adheres to the given data [23].

The PCF [24] applies a polynomial of a given degree to fit the data. The higher the degree, the closer the fitting curve will be to the given data.

The FS [25] is a sum of sine and cosine functions of different period that describes a periodic signal. It is represented in either the trigonometric form or the exponential form

$$y = a_0 + \sum_{i=1}^n [a_i \cos(nwx) + b_i \sin(nwx)] \quad (1)$$

where  $a_0$  models a constant (intercept) term in the data and is associated with the  $i = 0$  cosine term,  $w$  is the fundamental frequency of the signal, and  $n$  is the number of terms (harmonics) in the series ( $1 \leq n \leq 8$ ).

The GM [26] fits peaks and is given by

$$y = \sum_{i=1}^n a_i e^{-\left(\frac{x-b_i}{c_i}\right)^2} \quad (2)$$

where  $a$  is the amplitude,  $b$  is the centroid (location),  $c$  is related to the peak width, and  $n$  is the number of peaks to fit ( $1 \leq n \leq 8$ ).

The SG is a particular type of low-pass filter, well adapted for data smoothing [27]. This filter

$$g_i = \sum_{n=nL}^{nR} c_n f_i + n \quad (3)$$

replaces each data value  $f_i$ ,  $i = 1, \dots, N$ , by a linear combination  $g_i$  of nearby values in a window defined by the number of points used “to the left” ( $nL$ ) and “to the right” ( $nR$ ) of a data point  $i$ , respectively. The application of the SG in PhenoSat uses the simplest case ( $nL = nR$ ), where the same number of points is used “to the left” and “to the right.” Some tests were done to estimate the best SG moving window and polynomial order, and a first-degree polynomial with a frame size of five proved to be capable of removing the undesirable data from NDVI time series, capturing efficiently all the transitions related to the main and double growth seasons.

The DL, defined by (4), uses seven parameters to fit the VI data

$$VI_t = VI_w + \frac{k}{1 + \exp[-c(t-p)]} - \frac{k + VI_w - VI_{w1}}{1 + \exp[-d(t-e)]} \quad (4)$$

where  $t$  represents the time,  $VI_t$  is the VI value at time  $t$ ,  $k$  is related with the VI asymptotical value,  $c$  and  $d$  are the slopes at “left” and “right,” and  $p$  and  $e$  are the dates of the inflection points.  $VI_w$  and  $VI_{w1}$  are the VI values before the start of growing season and after the leaf fall, respectively. The continuity between the vegetation growth and senescence parts is assured by the  $k$  parameter, even when they differ in shape [28]. The parameters of the DL method were estimated by the Levenberg–Marquardt [29] algorithm that requires some reasonable initial values.

Fig. 1 presents a schematic representation of the DL parameters, using two consecutive years of NDVI SPOT VGT data. The continuity between the two years is assured by the  $NDVI_w$  and  $NDVI_{w1}$ , being the  $NDVI_w$  for the second year (beginning of the time series) the same as the  $NDVI_{w1}$  for the first year (final of the time series).

A new fitting method using a combination of SG and DL (SG+DL) was also implemented. It consists of the application

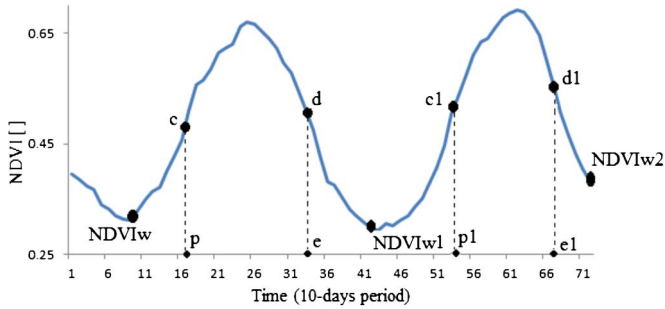


Fig. 1. Schematic representation of the DL parameters, using two consecutive years of NDVI SPOT VGT data.

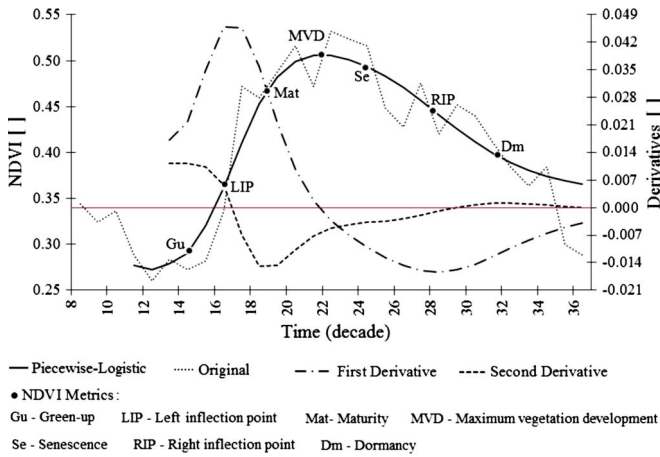


Fig. 2. Representation of NDVI metrics (Gu, LIP, Mat, MVD, Se, RIP, and Dm) calculated based on maximum and minimum rates of curvature change. Global maximum and minimum of the curvature estimate the inflection points.

of the SG filter to the NDVI time series, followed by the DL fitting process.

**B. Phenological Information Extraction**

PhenoSat extracts information (VI value and time of occurrence) of seven phenological stages in the main growing season: green-up (Gu) or start of season (SOS), left and right inflection points (LIP and RIP), maturity (Mat), maximum vegetation development (MVD), senescence (Se), and dormancy (Dm). To estimate these stages, the fitted data derivatives are used. Initially, the LIP and RIP, corresponding to the maximum growth and maximum senescence rates, are calculated using the maximum and minimum of the fitted data first derivative. Then, the MVD is determined as the maximum VI value observed in the data interval delimited by the inflection points. The Gu and Dm correspond to the maxima of the fitted data second derivative to the left and right of the MVD, respectively. Furthermore, the maturity [beginning of ripening stage (fruit vegetation) or full canopy (nonfruit vegetation)] and senescence correspond to the minima of the fitted second derivative to the left and right of the MVD. As an illustration, Fig. 2 shows the NDVI values optimized after DL application and a schematic procedure for the calculation of the NDVI metrics. The NDVI values were obtained from SPOT VGT sensor for a vineyard region in Portugal, for the year 2000.

In addition to the phenological information extracted by PhenoSat for the main growing season, it is also possible to record information for a double growth season or regrowth occurrence. This option allows obtaining information about the VI value and time occurrence for the start and maximum of this period. To determine the regrowth stage occurrence, the use of two successive years of data is necessary: the data for the year in analysis and the data for the following year. The start occurs when an increase of three or more consecutive points is verified after the Dm stage. The maximum of the double growth happens when, after the start occurrence, a decreasing period is verified.

Another feature of PhenoSat is the possibility of selecting an annual time-series subinterval, reducing the volume of data to be processed and thus improving the fitting process. This temporal region of interest can be selected manually, by the user, or automatically, by PhenoSat. The manual selection can be done by introducing the initial and final time positions in a specific application defined for this purpose. This selection must be based on the knowledge of the vegetation behavior, at normal growth conditions. For the automatic detection of the temporal region of interest, the PhenoSat first finds the maximum value of the data and then searches for the point where a significant increase (or abrupt decrease) is verified to the left of the maximum. That point corresponds to the initial position of the region of interest. Afterward, to determine the final position, the program proceeds in a similar way but evaluates the data to the right of the maximum.

**III. EXPERIMENTAL SETUP**

**A. Satellite Data and Study Areas**

In order to evaluate the performance of PhenoSat, various tests were carried out using NDVI time series acquired from different sensors, representing a variety of vegetation types and geographical locations. Satellite vegetation data covering Portugal from Aqua MODIS NDVI (16-day composite and 250-m resolution) and NDVI SPOT VGT (ten-day composite and 1-km resolution) were downloaded from Reverb [30] and Vision on Technology (VITO) [31], respectively. Three different vegetation types, namely, vineyard (VIN), seminatural meadows (SNM), and low shrublands (LSL), were used for the experimental test carried out (Table I). These land cover types were selected due to the availability of VIN ground-based information as well as for its different growth patterns: 1) The grapevines in Portugal have a long dormancy period with intense understory vegetation growth and a discontinuous canopy [28]; 2) SNM presents a vegetation regrowth by August whose intensity and date of occurrence are mainly dependent on climatic conditions [32]; and 3) LSL, composed mainly of shrubs and permanent herbaceous species [33], is an extensive grazing area with minimal agronomic human intervention, where vegetation development occurs later than in the SNM.

For each test site selected, the median NDVI value was recorded from each image. The median was used instead of the mean values as it has lower sensitivity to erroneous or outlier values. Annual NDVI time series were created for VIN, SNM, and LSL.

TABLE I  
DESCRIPTION OF STUDY AREAS AND SATELLITE DATA SETS USED

Land Cover Type	Code	Location	Coordinates (Lat/Long WGS84)	Satellite product	Time-series period	Size (pixels)
Vineyard (*)	VIN	Portugal- Douro region	UL: 7d25'09W, 41d15'21N BR: 7d23'37W, 41d13'53N	SPOT VGT (1km; 10 days)	2000 to 2005	9
Low Shrublands	LSL	Portugal- Montalegre	UL: 7d55'25W, 41d50'11N BR: 7d54'53W, 41d48'35N	SPOT VGT (1km; 10 days)	1999 to 2006	48
Semi-Natural Meadows	SNM	Portugal- Montalegre	UL: 7d57'19W, 41d37'58N BR: 7d56'58W, 41d37'34N	AQUA MODIS (250m; 16 days)	2003 to 2007	6

(\*) Ground-based measurements available.

### B. PhenoSat-Derived Phenology

For each test site, the SOS and MVD phenological stages were extracted. If the crop shows a double growth season, the start and maximum of this period were also recorded. The PhenoSat-derived phenology, for each vegetation type, was compared against observed measures and the results produced by TIMESAT. Both software tools extract phenological information from satellite VI time series. TIMESAT was chosen for this experiment because it is a freely available software tool and has time-series fitting algorithms similar to PhenoSat.

TIMESAT uses three distinct methods to fit the NDVI time series: AG, DL, and SG. Previous studies showed that AG and DL algorithms present similar results [11], [34]. As described in Section II-A, PhenoSat permits the use of seven different methods to smooth VI time series. However, for consistency, in this experiment, we only used DL, SG, and SG+DL as they are implemented in both software tools.

The SG filter was implemented in both tools in a similar way, replacing each data value by a linear combination of nearby values in a window [27]. The implementation of the DL differs in the number of parameters: PhenoSat uses seven parameters (see (4) in Section II-A), and TIMESAT uses only four parameters, which are related to the inflection points and the respective rates of change.

To determine the phenological metrics, TIMESAT uses a threshold approach: The start time of the season corresponds to the time for which the left edge has increased by 20% (default value used by TIMESAT) of seasonal amplitude, measured from the left minimum level. The time for the middle of the season is obtained by calculating the mean value of the times for which the left edge has increased to the 80% level and the right edge has decreased to the 80% level. On the other hand, PhenoSat uses an algorithm based on derivatives avoiding thresholds or empirical constants, providing a method that can be applied globally and is capable of identifying multiple growth periods within a single year.

### C. Software Performance Evaluation

To evaluate the performance of the software in the extraction of phenological metrics, a Spearman's rank correlation (e.g.,

TABLE II  
STATISTICS OF OBSERVED PHENOLOGICAL MEASURES OBTAINED FOR VINEYARD (VIN), LOW SHRUBLAND (LSL), AND SEMINATURAL MEADOW (SNM) VEGETATION TYPES. NOTE THAT THE OBSERVED MEASURES WERE OBTAINED FROM FIELD MEASUREMENTS (VIN) OR BY THE ANALYSIS OF THE NDVI ANNUAL PROFILES (LSL AND SNM)

Phenological Stage	Statistics	VIN	LSL	SNM
<i>Start of Season (SOS)</i>	Mean (DOY)	82.16	88.75	102.40
	Maximum (DOY)	91.00	100.00	128.00
	Minimum (DOY)	79.00	70.00	96.00
	Standard deviation (days)	4.44	9.91	14.31
<i>Maximum Vegetation Development (MVD)</i>	Mean (DOY)	206.66	177.50	150.40
	Maximum (DOY)	215.00	180.00	176.00
	Minimum (DOY)	201.00	170.00	128.00
	Standard deviation (days)	6.21	4.62	18.24

Note: The observed measures were obtained from field measurements (VIN) or by the analysis of the NDVI annual profiles (LSL and SNM).

[35]) and a root-mean-square error (rmse) analysis were performed to compare the estimated and observed phenological measures. Table II presents the statistics of observed phenological measures obtained for each study area. For VIN, the observed measures were collected in the field, according to the Baggiolini scale [36]; for LSL and SNM, they were derived by the analysis of the original NDVI time series, taking into account the knowledge of the vegetation behavior in the field under normal conditions (regular atmospheric conditions, without human interventions that affect the normal growth of vegetation). The SOS was determined as the first point where a continuous (five or more points) vegetation growth was verified (March/April). The MVD was identified as the maximum NDVI value in the annual time series, which generally occurs in June or early July. The SNM regrowth was related to the first significant (three or more points) vegetation growth after the MVD (around August).

Two other important points to evaluate are the ability of TIMESAT and PhenoSat to detect a double growth season and

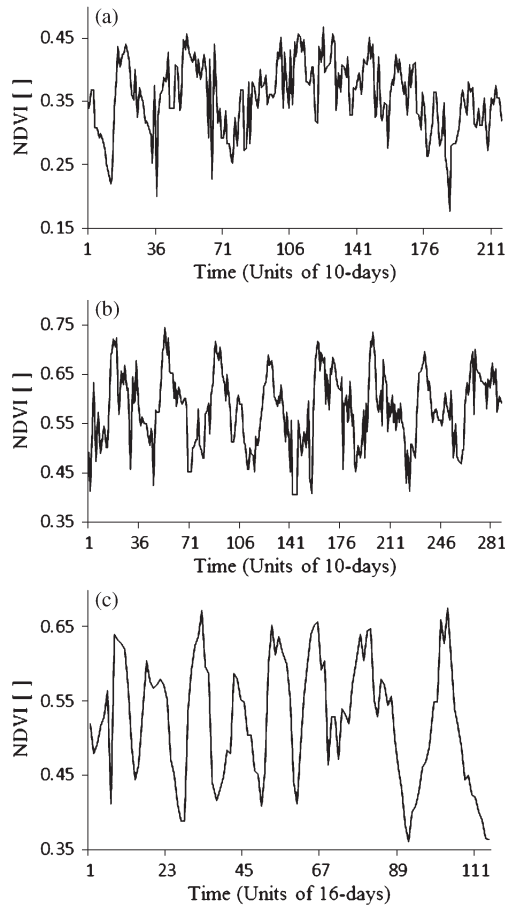


Fig. 3. Original NDVI temporal profiles for three land cover types: (a) Vineyard, (b) low shrubland, and (c) seminatural meadows. The  $x$ -axis scale of the plot (c) is different from the other two as the sensor used was MODIS (16 days), and for plots (a) and (b), the sensor used was SPOT VGT (ten days).

the advantage of the option of selecting a temporal region of interest (available in PhenoSat but not in TIMESAT).

#### IV. RESULTS AND DISCUSSION

The original NDVI temporal profiles produced for each test site are shown in Fig. 3. As described, these plots were produced from the median NDVI of the test area available for each land cover type. The general shape of these temporal plots is very much what could be expected for the vegetation types. The SNMs are irrigated all year round, and in general, they present a SOS in March/April, an MVD in June/July, and a regrowth by August. LSL presents a similar behavior to SNM; however, as the LSLs are not irrigated, the regrowth is less frequent and with less magnitude.

For the VIN vegetation type, the annual profile presents high variability among years and is not so well defined as for SNM and LSL. Grapevine (VIN) budbreak occurs in March/April (Julian Day 79 to 91), followed by a period of about four weeks of intensive growth and then a steady decrease until “veraison” (change of color of the grapes) that occurs between Julian Day 201 and 215 (Table II; Fig. 3). Due to the large interrow space and discontinuous canopy, it is possible to see, particularly in the annual winter season, the effect of the soil vegetation growth on the VI temporal profiles.

#### A. PhenoSat Fitting Methods

To evaluate the ability of the PhenoSat fitting algorithms (SG, DL, SG+DL, CSS, PCF, GM, and FS) in estimating phenology, a comparison was made between derived and observed phenological information.

The performance of CSS, PCF, GM, and FS algorithms depends on a smoothing parameter, which affects the adherence of the fitting curve to the original data. To test the sensitivity of the algorithms to the smoothing parameter, PhenoSat was executed eight times for each algorithm: For PCF, GM, and FS, the parameter was varied between one and eight, and for CSS, it was varied from 0.1 to 0.8.

Table III presents the Spearman’s correlation rank ( $r_s$ ) between PhenoSat-estimated phenological parameters and observed values, for the VIN region. The SOS and MVD obtained from PhenoSat were compared with grapevine field measures of budbreak and veraison (“change of color of the grapes”), respectively. The SG+DL algorithm presented correlations above 0.70 for both phenological stages evaluated. For CSS, FS, GM, and PCF algorithms, high smoothing parameters allowed a more accurate detection of the SOS, presenting correlations of 0.71, 0.83, 0.89, and 0.64, respectively. However, for the MVD, intermediate values obtained the best correlations except for CSS, where the lowest parameter (0.1) produced the highest correlation (0.40).

#### B. PhenoSat and TIMESAT Results

Table IV shows the Spearman’s rank correlation ( $r_s$ ) and rmse results obtained between observed and estimated phenological SOS and MVD, for VIN, LSL, and SNM vegetation types.

Analyzing the fitting algorithms, SG+DL was the best for PhenoSat with correlations above 0.70 for all tests, except for the seminatural meadow SOS. In TIMESAT, SG and SG+DL were the algorithms that obtained the better and the worst correlations, respectively, between derived and observed phenologies.

For VIN, PhenoSat showed the best results in all fitting methods, providing rmse values between 15 and 23 days. In TIMESAT, some VIN years are not considered in the correlation estimation as the software was unable to predict the metrics for those years. The poor adherence of the TIMESAT fitting algorithms to the VIN original data can explain the abnormal rmse (above 100 days) results obtained.

For all vegetation types, TIMESAT and PhenoSat proved to be useful tools to detect phenological events, obtaining, in some cases, correlations above 0.70. TIMESAT obtained better results for the MVD except using the SG+DL algorithm, for which the PhenoSat always was the best with mean correlations of about 0.70 and rmse values between 5 and 23 days.

Table V presents the observed and computed regrowth season parameters. In 2007, there was no regrowth, which was correctly verified by both software tools. For the other years, TIMESAT had some difficulty in determining the double growth season using the different algorithms; the SG+DL was the worst case, as TIMESAT was unable to determine the parameters for any year tested. The TIMESAT 1-season option does not report any regrowth; thus, the results presented in

TABLE III  
SPEARMAN’S CORRELATION RANK (*rs*) BETWEEN PHENOSAT-ESTIMATED PHENOLOGICAL PARAMETERS AND OBSERVED MEASURES, FOR THE VINEYARD TEST SITE

Algorithm	Start of Season (SOS)			Maximum Vegetation Development (MVD)		
	Best Parameter	<i>rs</i>	<i>p-value</i>	Best Parameter	<i>rs</i>	<i>p-value</i>
DL	—	-0.89	0.009	—	-0.09	0.44
SG	—	-0.15	0.39	—	0.15	0.39
SG+DL	—	-0.77	0.04	—	-0.71	0.06
CSS	0.8	-0.71	0.06	0.1	-0.40	0.22
PCF	8	-0.64	0.09	4	0.62	0.10
FS	8	-0.83	0.02	6	-0.43	0.20
GM	8	-0.89	0.01	5	-0.60	0.10

DL, SG, SG+DL, CSS, PCF, FS and GM represent the 7 fitting algorithms, respectively, piecewise-logistic, Savitzky-Golay, combination of Savitzky-Golay and piecewise-logistic, cubic smoothing splines, polynomial curve fitting, Fourier series and Gaussian models.

TABLE IV  
SPEARMAN’S RANK CORRELATION (*rs*) AND RMSE BETWEEN OBSERVED AND ESTIMATED PHENOLOGIES, OBTAINED FROM TIMESAT AND PHENOSAT, FOR VINEYARD (VIN), LOW SHRUBLAND (LSL), AND SEMINATURAL MEADOW (SNM) VEGETATION TYPES

Statistics	DL		SG				SG+DL					
	TIMESAT		PhenoSat		TIMESAT		PhenoSat		TIMESAT		PhenoSat	
	SOS	MVD	SOS	MVD	SOS	MVD	SOS	MVD	SOS	MVD	SOS	MVD
<i>VIN</i>												
n	5	5	6	6	4	4	6	6	3	3	6	6
<i>rs</i>	0.10	-0.60	-0.89	-0.09	-0.40	-0.80	-0.15	0.15	-0.50	--	-0.77	-0.71
<i>p-value</i>	0.44	0.14	0.009	0.44	0.30	0.10	0.39	0.39	0.33	--	0.04	0.06
RMSE	*	*	18	23	*	*	15	19	*	*	15	23
<i>LSL</i>												
n	8	8	8	8	8	8	8	8	8	8	8	8
<i>rs</i>	0.43	-0.13	0.56	-0.13	0.55	0.76	0.43	0.61	0.12	0.38	0.79	0.76
<i>p-value</i>	0.14	0.38	0.07	0.38	0.08	0.02	0.14	0.06	0.39	0.18	0.01	0.02
RMSE	*	11	9	17	11	11	13	8	11	10	5	5
<i>SNM</i>												
n	5	5	5	5	5	5	5	5	5	5	5	5
<i>rs</i>	-0.71	0.87	0.71	0.05	-0.35	0.98	-0.41	0.92	-0.35	0.41	0.35	0.95
<i>p-value</i>	0.09	0.03	0.09	0.47	0.28	0.002	0.25	0.01	0.28	0.27	0.28	0.007
RMSE	8	11	10	19	10	14	7	7	10	28	17	10

SOS represents the start of season and MVD the maximum vegetation development.

*rs* is the Spearman’s rank correlation coefficient and *p-value* the probability associated. *n* is the number of observations.

DL, SG and SG+DL represent the 3 fitting algorithms, respectively, piecewise-logistic, Savitzky-Golay and the combination of Savitzky-Golay with piecewise-logistic.

RMSE is the Root Mean of Square Error and \* represents a RMSE value above 100 days.

Table V were obtained using the option 2-seasons. Although this option might work for some years, it is not sufficiently flexible for the years in which there is no regrowth. On the other hand, PhenoSat determined the parameters with high precision, only failing on the maximum for the year 2005. The PhenoSat algorithm, to derive the regrowth, uses the preprocessed time series before the application of the fitting process, being independent from the fitted data, and thus provides identical results for all fitting algorithms tested.

Fig. 4 shows the original and SG, DL, and SG+DL fitted data, obtained by PhenoSat and TIMESAT software tools, for

the SNM region in 2003. From the original data, it is possible to verify a regrowth occurrence around the Julian Day 240 to 256. The TIMESAT SG algorithm cannot detect efficiently the regrowth, smoothing this period, contrary to PhenoSat that can detect with high precision the beginning of the regrowth. For DL, the TIMESAT and PhenoSat fitted results present slight differences for the main growing season; however, for the regrowth period, PhenoSat tends to smooth this period. The SG+DL algorithm presents a similar performance to that verified in SG for the TIMESAT tool; for PhenoSat, the SG+DL fitted results show a “mix” of the SG and DL results.

TABLE V  
OBSERVED AND ESTIMATED DOUBLE GROWTH SEASON PARAMETERS, FROM TIMESAT (2-SEASONS)  
AND PHENOSAT, FOR A SEMINATURAL MEADOW REGION

Year	DL				SG				SG+DL				OBSERVED	
	TIMESAT		PhenoSat		TIMESAT		PhenoSat		TIMESAT		PhenoSat			
	Start	Max	Start	Max	Start	Max	Start	Max	Start	Max	Start	Max		
2003	238	291	252	304	290	333	252	304	n.d.	n.d.	252	304	252	304
2004	229	283	252	288	n.d.	n.d.	252	288	n.d.	n.d.	252	288	252	288
2005	262	321	252	336	264	326	252	336	n.d.	n.d.	252	336	252	328
2006	235	305	252	320	244	300	252	320	n.d.	n.d.	252	320	252	320
2007	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

n.d. signifies that no regrowth is verified on this year.

DL, SG and SG+DL represent the 3 fitting algorithms, respectively, piecewise-logistic, Savitzky-Golay and the combination of Savitzky-Golay with piecewise- logistic.

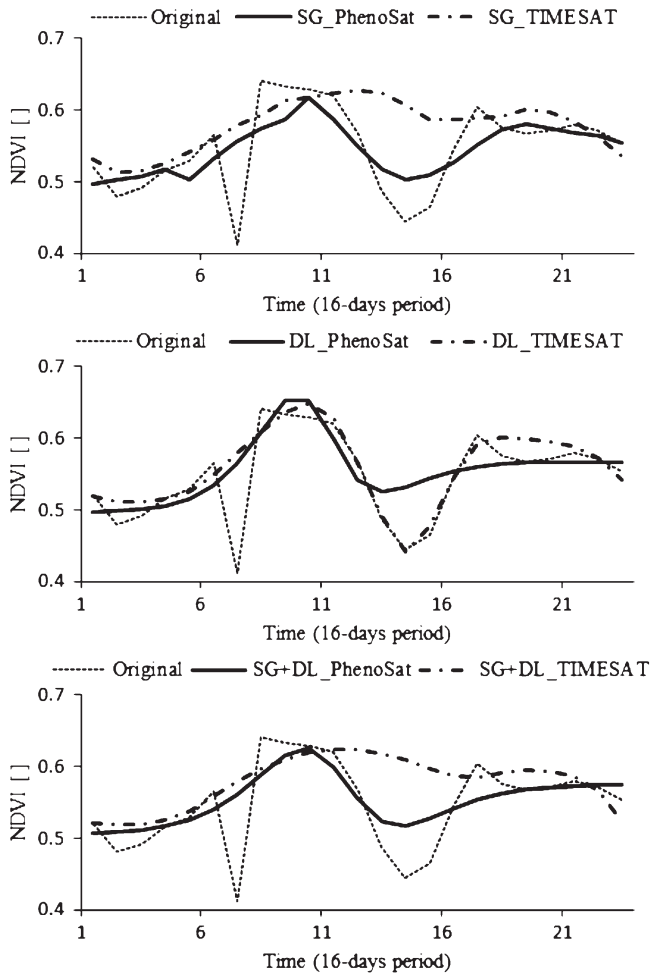


Fig. 4. Original and fitted results obtained from PhenoSat and TIMESAT using SG, piecewise logistic (DL), or both (SG+DL) algorithms, for the seminatural meadow vegetation type in 2003.

The selection of an in-season window is of particular interest for crops like VIN, with a long period of winter dormancy, and crops with discontinuous canopies. The VIN, in general, presents a budbreak in March/April (Julian Day 79 to 91) and dormancy stage in October/November (Julian Day 300 to 310). Thus, the period comprising the Julian Day 79 to 310 is usually

used to define a manual selection of the region of interest. However, the manual selection, based on the knowledge of the vegetation behavior in the field under normal conditions, is not flexible to adapt to the different vegetation dynamics that can occur over the years due to some factors, such as adverse weather conditions. The automatic selection can be a good option to eliminate these limitations and detect more accurately the temporal region of interest.

Fig. 5 shows the original VIN data for 2000 and 2003 and SG+DL fitted data obtained using all range of observations and using a region of interest automatically selected by PhenoSat. For the year 2000, PhenoSat selects the range Julian Day 110 to 310 as the temporal region of interest, which is easily recognized as the VIN dynamics period in the original data. However, for the year 2003, only 15 observations (Julian Day 100 to 240) are selected. The difference in the two years can be explained by the fact that, in 2003, there was an irregular precipitation from August 30 to September 5, which was not verified in 2000. The weather conditions in 2003 caused an early growth of the soil vegetation, which can be seen graphically around the Julian Day 250.

Although, for the year 2000, no advantage was verified in selecting a subtemporal region of interest, this is not verified in 2003, where the fitted data provided by all 36 observations tend to oversmooth the original data and are therefore unable to detect the regrowth occurrence around Julian Day 250. This difference in the fitted results leads to different phenological stage occurrences. Using all 36 observations ( $r^2 = 0.41$ ), the SOS and the MVD were obtained on Julian Day 140 and Julian Day 160, respectively; however, using the 15 observations ( $r^2 = 0.96$ ) provided by the region of interest, an earlier SOS (Julian Day 120) and a later MVD (Julian Day 180) were verified, these results being more consistent with the original data. This demonstrates that the selection of the temporal region of interest produces an improvement in the fitted process, thus obtaining more reliable results.

Table VI presents the effect of the temporal region of interest selection on the Spearman's correlations between observed and PhenoSat-derived phenologies, for the VIN region (2000–2005). Using the temporal region of interest, the correlations range between  $-0.70$  and  $-0.77$  being always higher than

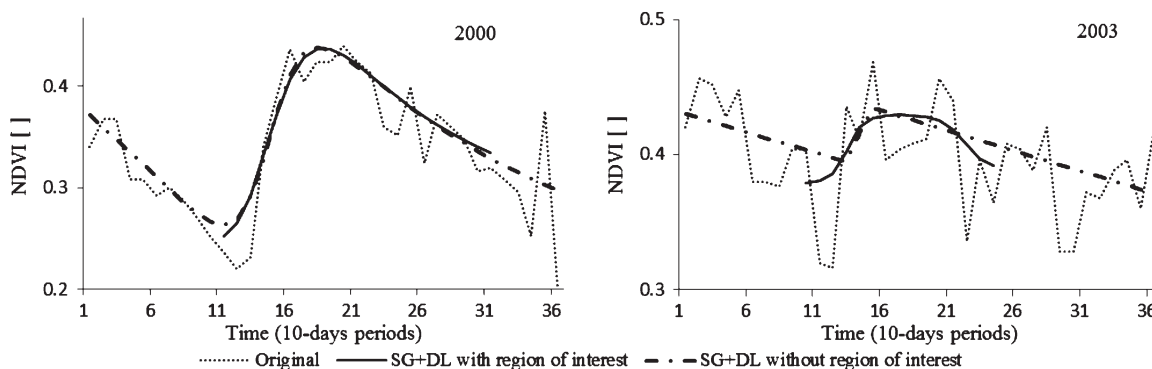


Fig. 5. (Dotted line) Original vineyard data for 2000 and 2003, (solid line) SG+DL fitted data obtained using a subtemporal region of interest, and (dashed line) all range of observations.

TABLE VI  
CORRELATIONS ( $r_s$ ) BETWEEN OBSERVED AND PHENOSAT-DERIVED PHENOLOGY PARAMETERS FOR VINEYARD (2000–2005)

Phenological stages	In-season selected region	
	Not selected	Selected
Start of Season (n=6)	-0.51	-0.77
Flowering (n=6)	0.26	-0.70
Maximum Vegetation Development (n=6)	0.47	-0.71

Significant level of the coefficient correlation for 6 observations:  $r_s > 0.676$  (0.10);  $r_s > 0.759$  (0.05).

those obtained using the complete range of 36 observations. Thus, the PhenoSat feature to select an in-season region of interest proved to be a valuable tool for vineyard monitoring, allowing for high significant correlations between estimated and observed values to be obtained. Adverse weather conditions such as unseasonal snow, extreme heat, or irregular precipitation could result in a false vegetation regrowth. The selection of a subtemporal region of interest based on vegetation dynamics knowledge could help to deal with the false report of regrowth in particular natural environments (high latitudes or boreal regions), leading to a better analysis and more consistent results.

### V. CONCLUSION

The experiment carried out to evaluate the performance of the PhenoSat showed that it can produce accurate and consistent results, compared with observed measures. This tool proved capable of solving some limitations present in other software tools, such as the following: the detection of a double growth season, with the extraction of phenological parameters for this period, and the possibility to select, manually or automatically, an in-season region of interest.

The extraction of phenological parameters using an algorithm based on derivatives allows PhenoSat to avoid thresholds or empirical constants, providing a method that can be applied globally and that is capable of identifying multiple crops or regrowth within a single year. The tool proved to be very efficient in detecting the double growth period. The independency of the fitted results leads to a more realistic time-series profile over the year and, thus, more accurate regrowth-derived results.

The option to select, automatically or manually, a temporal region of interest of the VI time series provides an improvement in the fitting process, leading to more reliable results. The PhenoSat feature to select an in-season region of interest proved to be a valuable tool for vineyard monitoring and can enlarge the PhenoSat application to crops with discontinuous canopy, like forestry and deciduous fruit trees.

PhenoSat is a freely available software tool, with a preliminary version currently available at [http://www.fc.up.pt/LamSat\\_XXI](http://www.fc.up.pt/LamSat_XXI). This version is able to run a single data set file (with data for a single vegetation type, for different years) at a time. However, a new version is being developed to improve and automate the process in order to receive various data set files at once. Another improvement consists in creating an application that permits the extraction of phenological information directly from the EOS images. This option eliminates the laborious and manual preprocessing steps that involve the extraction of VI data from the images to create the data set input file.

### ACKNOWLEDGMENT

The authors would like to thank VITO for the SPOT VEGETATION and Reverb for the MODIS images and also to the authors of TIMESAT for providing access to their software tool.

### REFERENCES

- [1] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sens. Environ.*, vol. 83, no. 1/2, pp. 195–213, Nov. 2002.
- [2] B. N. Holben, "Characteristics of maximum-value composite images from temporal AVHRR data," *Int. J. Remote Sens.*, vol. 7, no. 11, pp. 1417–1434, Nov. 1986.
- [3] F. Fontana, C. Rixen, T. Jonas, G. Abereg, and S. Wunderle, "Alpine grassland phenology as seen in AVHRR, VEGETATION and MODIS NDVI time series—A comparison with *in situ* measurements," *Sensors*, vol. 8, no. 4, pp. 2833–2853, Apr. 2008.
- [4] 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 Sens. Environ.*, vol. 100, no. 3, pp. 321–334, Feb. 15, 2006.
- [5] B. A. Bradley, R. W. Jacob, J. F. Hermance, and J. F. Mustard, "A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data," *Remote Sens. Environ.*, vol. 106, no. 2, pp. 137–145, Jan. 30, 2007.



- [6] P. Jonsson and L. Eklundh, "Seasonality extraction by function fitting to time-series of satellite sensor data," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1824–1832, Aug. 2002.
- [7] P. Jonsson and L. Eklundh, "TIMESAT—A program for analyzing time-series of satellite sensor data," *Comput. Geosci.*, vol. 30, no. 8, pp. 833–845, Oct. 2004.
- [8] M. G. Ma and F. Veroustraete, "Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China," *Adv. Space Res.*, vol. 37, no. 4, pp. 835–840, 2006, Natural Hazards and Oceanographic Processes from Satellite Data.
- [9] A. Moody and D. M. Johnson, "Land-surface phenologies from AVHRR using the discrete Fourier transform," *Remote Sens. Environ.*, vol. 75, no. 3, pp. 305–323, Mar. 2001.
- [10] G. J. Roerink, M. Menenti, and W. Verhoef, "Reconstructing cloudfree NDVI composites using Fourier analysis of time series," *Int. J. Remote Sens.*, vol. 21, no. 9, pp. 1911–1917, Jun. 15, 2000.
- [11] J. N. Hird and G. J. McDermid, "Noise reduction of NDVI time series: An empirical comparison of selected techniques," *Remote Sens. Environ.*, vol. 113, no. 1, pp. 248–258, Jan. 15, 2009.
- [12] W. Zhu, Y. Pan, H. He, L. Wang, M. Mou, and J. Liu, "A changing-weight filter method for reconstructing a high-quality NDVI time series to preserve the integrity of vegetation phenology," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1085–1094, Apr. 2012.
- [13] H. Carrao, P. Goncalves, and M. Caetano, "A nonlinear harmonic model for fitting satellite image time series: Analysis and prediction of land cover dynamics," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1919–1930, Apr. 2010.
- [14] J. M. Chen, F. Deng, and M. Z. Chen, "Locally adjusted cubic-spline capping for reconstructing seasonal trajectories of a satellite-derived surface parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2230–2238, Aug. 2006.
- [15] T. Udelhoven, "TimeStats: A software tool for the retrieval of temporal patterns from global satellite archives," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 2, pp. 310–317, Jun. 2011.
- [16] B. Tan, J. T. Morisette, R. E. Wolfe, F. Gao, G. A. Ederer, J. Nightingale, and J. A. Pedelty, "An enhanced TIMESAT algorithm for estimating vegetation phenology metrics from MODIS data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 2, pp. 361–371, Jun. 2011.
- [17] R. D. McKellip, K. W. Ross, J. P. Spruce, J. C. Smoot, R. E. Ryan, G. E. Gasser, D. L. Prados, and R. D. Vaughan, "Phenological parameters estimation tool," NASA Tech. Briefs, New York, Sep. 30, 2010.
- [18] K. W. Ross, B. A. Spiering, and M. T. Kalcic, "Monitoring phenology as indicator for timing of nutrient inputs in northern gulf watersheds," in *Proc. OCEANS*, 2009, pp. 1–4, vols 1–3.
- [19] J. M. B. Carreiras, J. M. C. Pereira, Y. E. Shimabukuro, and D. Stroppiana, "Evaluation of compositing algorithms over the Brazilian Amazon using SPOT-4 VEGETATION data," *Int. J. Remote Sens.*, vol. 24, no. 17, pp. 3427–3440, Sep. 2003.
- [20] B. Holben and R. S. Fraser, "Red and near-infrared sensor response to off-nadir viewing," *Int. J. Remote Sens.*, vol. 5, no. 1, pp. 145–160, Feb. 1984.
- [21] G. G. Gutman, "Vegetation indexes from AVHRR—An update and future-prospects," *Remote Sens. Environ.*, vol. 35, no. 2/3, pp. 121–136, Feb./Mar. 1991.
- [22] X. W. Li and A. H. Strahler, "Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy—Effect of crown shape and mutual shadowing," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 2, pp. 276–292, Mar. 1992.
- [23] C. H. Reinsch, "Smoothing by spline functions," *Numerische Math.*, vol. 10, no. 3, pp. 177–183, 1967.
- [24] J. Verschelde, "Introduction to symbolic computation: A maple/MATLAB course," in *Proc. Maple Conf.*, Waterloo, ON, Canada, 2005, pp. 500–509.
- [25] *Curve Fitting Toolbox: Fit Curves and Surfaces to Data Using Regression, Interpolation, and Smoothing*, MathWorks, Natick, MA, 2011.
- [26] A. Goshtasby and W. D. Oneill, "Curve-fitting by a sum of Gaussians," *CVGIP—Graph. Models Image Process.*, vol. 56, no. 4, pp. 281–288, Jul. 1994.
- [27] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.
- [28] M. Cunha, A. R. S. Marçal, and A. Rodrigues, "A comparative study of satellite and ground-based vineyard phenology," in *Proc. 29th Symp. EARSeL*, Chania, Greece, 2010, pp. 68–77.
- [29] D. Montgomery, E. Peck, and G. Vining, *Introduction to Linear Regression Analysis*, 4th ed. Hoboken, NJ: Wiley, 2006.
- [30] Reverb—The Next Generation Earth Science Discovery Tool, Jan. 13, 2012. [Online]. Available: <https://reverb.echo.nasa.gov>
- [31] Vision on Technology, Jan. 13, 2012. [Online]. Available: <http://free.vgt.vito.be/>
- [32] I. Pocas, M. Cunha, and L. S. Pereira, "Dynamics of mountain semi-natural grassland meadows inferred from SPOT-VEGETATION and field spectroradiometer data," *Int. J. Remote Sens.*, vol. 33, no. 14, pp. 4334–4355, Jul. 2012.
- [33] I. Pocas, M. Cunha, and L. S. Pereira, "Remote sensing based indicators of changes in a mountain rural landscape of Northeast Portugal," *Appl. Geograph.*, vol. 31, no. 3, pp. 871–880, Jul. 2011.
- [34] F. Gao, J. T. Morisette, R. E. Wolfe, G. Ederer, J. Pedelty, E. Masuoka, R. Myneni, B. Tan, and J. Nightingale, "An algorithm to produce temporally and spatially continuous MODIS-LAI time series," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 60–64, Jan. 2008.
- [35] J. L. Myers and A. D. Well, *Research Design and Statistical Analysis*, 2nd ed. Mahwah, New Jersey: Lawrence Erlbaum Assoc. Inc., 2003.
- [36] M. Baggiolini, "Les stades repères dans le développement annuel de la vigne et leur utilisation pratique," *Rev. romande Agric. Vitic. Arboric.*, vol. 8, pp. 4–6, 1952.



**Arlete Rodrigues** (S'12) was born in Avelar, Leiria, Portugal. She received the B.Sc. degree in mathematics from the Universidade de Aveiro, Aveiro, Portugal, in 2004 and the M.Sc. degree in remote sensing from the Faculdade de Ciências, Universidade do Porto, Porto, Portugal, in 2008, where she is currently working toward the Ph.D. degree.



Since 2007, she has been a Researcher with the Centro de Investigação em Ciências Geo-Espaciais (CICGE), Faculdade de Ciências, Universidade do Porto, with a research topic of temporal remote sensing data analysis for vegetation studies. Her scientific interests include image processing, pattern recognition, remote sensing, and time-series analysis.

**Andre R. S. Marçal** was born in Luanda, Angola. He received the B.S. degree in physics (solid-state physics) from the Universidade do Porto, Porto, Portugal, in 1991 and the M.Sc. and Ph.D. degrees in remote sensing and image processing from the University of Dundee, Dundee, U.K., in 1994 and 1998, respectively.

He is currently an Assistant Professor with the Department of Mathematics, Universidade do Porto. His research interests include various topics in remote sensing and image processing.

Prof. Marçal was a recipient of the prize for the Best Ph.D. from the Remote Sensing Society (U.K.) in 1999. From 2005 to 2010, he was a bureau member of the European Association of Remote Sensing Laboratories, as the Secretary General and later as the Vice-Chairman.



**Mário Cunha** was born in Barcelos, Porto, Portugal. He received the B.S. degree in agronomic engineering from the Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal, and the Ph.D. degree in agrarian sciences from the Faculdade de Ciências, Universidade do Porto (FCUP), Porto, Portugal.

He is currently an Assistant Professor with the Department of Geosciences, Environment and Spatial Planning, FCUP, where he is also a Researcher with the Remote Sensing, Image Processing and GIS Group, Centro de Investigação em Ciências

Geo-Espaciais.