TIME SERIES ANALYSIS OF NOAA AVHRR DERIVED VEGETATION COVER AS A MEANS TO EXTRACT PROPORTIONS OF PERMANENT AND SEASONAL COMPONENTS AT PIXEL LEVEL

Christof J. Weissteiner[1], Stefan Sommer[1] and Peter Strobl[2]

1. Joint Research Centre (JRC), Institute for Environment and Sustainability – Rural, Water and Ecosystem Resources, Ispra, Italy; {christof.weissteiner / stefan.sommer}(at)jrc.it
2. Joint Research Centre (JRC), Institute for Environment and Sustainability – Land Management and Natural Hazards Unit, Ispra, Italy: peter.strobl(at)jrc.it

ABSTRACT

The scope of this study was to find a simple and robust technique to analyse a 16-year time series (totalling 576 decades) of NOAA-AVHRR derived Green Vegetation Fraction (GVF). The bio-physical properties of the observed vegetation canopy are described as a function of its composition in terms of a seasonally changing vegetation component and a permanent vegetation component. The principal idea behind the analysis is to use a simple model of an annual vegetation growing cycle per pixel which is fitted against the available time sequence of data, and interpret the parameters of the fit on the one hand, and the residuals of the original versus the fitted data on the other hand. For simplicity reasons this part is represented by a sine curve with a fixed wavelength of one year. This model allows splitting of the timely resolved vegetation signal into two components in vegetation appearance. One represents a "permanent background" throughout the year which is the off-set between the 0-level representing the absence of vegetation cover and the minimum of the modelled seasonal change. The second represents the difference between the maximum and the minimum vegetation cover modelled every year. This technique was applied to the entire Mediterranean region covered by a NOAA-AVHRR time series. The derived proportions of permanent and seasonal vegetation components were finally interpreted in relation to the European CORINE land cover class 'Olive grove', assessing the variation of permanent and seasonal vegetation components as a function of management intensity. This led to a distinction of different olive grove management intensity classes within the limits of the CORINE class. The olive class was chosen as test case because of its well known linkages between the evergreen component represented by the olive trees and the more or less pronounced presence of annual herbaceous understorey.

INTRODUCTION

The value of a time series lies along the time axis. When assessing vegetation by remote sensing, both, intra-annual changes and inter-annual changes can be observed, providing an implicit link to vegetation phenology and vegetation dynamics. Since the 1970s, monitoring and analysis of vegetation phenology is of high interest to research fields, such as ecology, agriculture land degradation and more recently climate change (1,2). There are various key parameters for investigation of remote sensing time series of vegetation. Depending on the focus of investigation, classical vegetation indices (e.g. Normalized Difference Vegetation Index *NDVI*), but also bio-physical variables such as *FAPAR* (Fraction of Absorbed Photosynthetically Active Radiation) or green vegetation fraction are being used. Also, the temporal resolution of time series is varying; with time steps ranging from daily data to aggregated values of weekly, monthly or even yearly intervals. Typically the available remote sensing time series for vegetation monitoring of large areas are based on relatively coarse data in terms of spatial resolution, e. g. NOAA AVHRR, SPOT VEGETATION or MODIS. While offering only low spatial resolution but high observation frequency, these data bases contain valuable information, especially when provided for a large number of years. This information needs to be extracted in a systematic and (semi-)automatic way by appropriate tools accounting for the numerous data inherent noise factors interfering with the relevant bio-physical information content of the recorded data.

Important work about vegetation dynamics parameterisation was done by Reed (3) and Hill & Donald (4). Examples how to characterise and map land cover properties by *NDVI* derived products can be found in literature (5,6). While many works have been carried out analysing the onset and end of the growing season (3,4), phenology analysis as such, for this investigation, represents only a side product. Instead, the main focus of this development was laid on vegetation dynamics.

Dynamics of a vegetation time series with a regular seasonal run can be characterised by the quantification of two main components inherent in the growing cycle: One component, in a descriptive graphical-mathematical sense, representing a "permanent background" base of the growing cycle, which is constant throughout the year. The other component, on top of the first, stands for a seasonally changing part with a maximum and a minimum period every year. Since image pixels with a strong dynamic throughout the year stand for different vegetation types than image pixels with a strong permanent component, this should allow the distinction of several biomes. In other words, assessing the degree of seasonality and/or permanence of vegetation, quantified for each year, should allow for a gradual distinction of 'vegetation functioning', related to vegetation dynamics throughout a year.

The aim was the development of a robust and simple tool, which allows extraction of parameters describing vegetation dynamics. Robustness was an important requirement, in order to overcome noisy and oscillating data, while preserving important dynamic features. The tool handles remote sensing data sensor-independently and is able to deal with different spatial resolutions. However, for a number of reasons (e. g. data availability), more emphasis was put on coarse spatial data processing. Independent from input data timely resolution, as a consequence of required robustness, but also for data handling reasons, all outcomes are on the base of full single years. The tool is simple to use, with a minimum of parameterisation required and produces an easily interpretable and understandable outcome. Nevertheless, in spite of the high degree of robustness of the model, input data should be pre-processed according to the requirements of the envisaged analysis (e.g. trend removal if required, etc.).

For a first product testing, the tool was applied for investigation of a specific topic related to olive groves. Aim of this test was, on the base of known 'Olive grove' sites, to distinguish different types of them, according to the seasonal behaviour of the associated ground flora. Since the type and amount of associated ground flora of olive groves is related to the cultivation type, as it was shown in other works (7,8), this can be used as a proxy for management intensity. For the land cover class 'Olive grove' which has particularities in the botanical and agronomical sense, parameters describing vegetation dynamics represent crucial descriptive elements to derive the level of management intensities.

DATA

This study was based on the MEDOKADS data set (Mediterranean Extended Daily One Km AVHRR Data Set), provided by the Free University of Berlin (9). It comprises a fully intercalibrated, radiometrically pre-processed time series of NOAA AVHRR at a resolution of 0.01 degrees, starting in 1989. The MEDOKADS map extent was subset to the range 27°N to 46°N and 10°W to 42°E (see Figure 5). For this data set the Green Vegetation Fraction (GVF) was computed by linear spectral mixture modelling for 10-day composites of *NDVI* and surface temperature (*Ts*). The spectral unmixing strategy was implemented based on the inverse relationship between the vegetation index *NDVI* and the land surface temperature. The method is based on the assumption that vegetation cover should predominantly control the position of an AVHRR land surface pixel within the feature space spanned up by *NDVI* and surface temperature (10,11).

Ancillary data comprised CORINE land cover 2000 (CLC) (12) and aerial photography (13) for validation purposes. CORINE land cover 2000 represents a consistent pan-European dataset with a land cover class 'Olive grove'.

METHODS

The goal was to model the permanent and seasonal vegetation component of a time series. The model (called SINFIT hereafter) chosen for this purpose is based on the assumption of two components in vegetation multi-temporal appearance. One represents the "permanent background" of vegetation, the base which is constant throughout the year, and the other is a seasonally changing part with a maximum and a minimum period every year. For simplicity reasons this part is represented by a sine curve with a fixed wavelength of one year (see Figure 1).

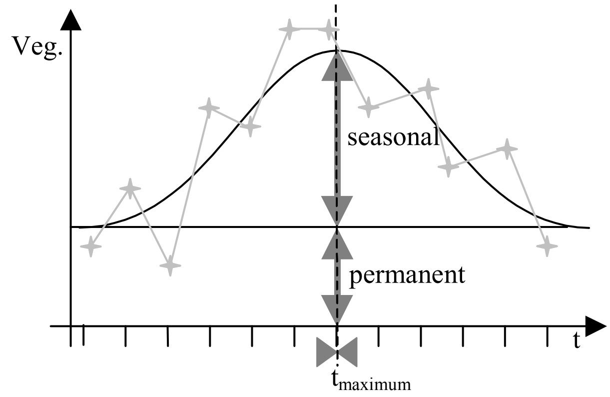

Figure 1: Observed and modelled permanent and seasonal component of a typical seasonal run of green vegetation.

Technically, this is achieved by fitting the sinusoidal model against the GVF run of the remote sensing data (MEDOKADS); the parameters of the fit and of the residuals are used in the following. The best fit of the model to the data is found in a two step procedure: First the cross-correlation between a 3-year time window of the data and the model is calculated whereby the model is subsequently shifted against the data in steps of the time interval. This is repeated over a full period (1 year). The time shift at which the model cross-correlates best with the data is fixed as the maximum period. In the second step, a linear regression is performed between the weighted 3-year time window of the data and the normalised model at the maximum cross-correlation shift. The offset found is recorded as the value of the permanent fraction, while the seasonal fraction is determined from the slope multiplied by the number of time slices per year (see Figure 2). The correlation coefficient is stored as a measure for the reliability of the seasonal parameters. The fit is executed separately for every available year of observation by using a 3-year moving time window, the target year of which is weighted 10 times higher than both surrounding marginal years. This weight was empirically tested for providing the best compromise in terms of root mean square error (*RMSE*) minimisation while best approximating the seasonal run. Obviously, by using a three- year time window, the first and last year of the time series cannot be considered, as there is a neighbouring (year 1-1 and year *n*+1) year missing for a correct calculation.

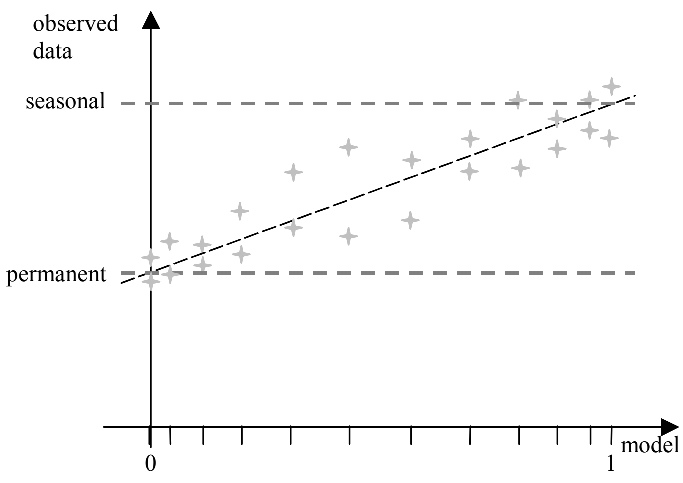

Figure 2: Correlation of observed and modelled data with offset for permanent vegetation.

The outcomes of the SINFIT model are:

1) The value of the permanent background (absolute *GVF* units, absolute permanent component - *APC*)
2) The amplitude of the seasonal changing portion (absolute *GVF* units, absolute seasonal component - *ASC*)
3) The month or decade for the maximum of the seasonal portion

Two more parameters of interest can be derived from (1) and (2).

1a) The ratio of permanent (1) background to the overall *GVF*, expressed as fraction in %. This is referred to as the normalised permanent component (*NPC*). It expresses the permanent (relative) proportion or fraction of the overall *GVF*, the base or minimum recorded value of GVF throughout the year.

2a) The ratio of the seasonal (2) changing portion to the overall *GVF*, expressed as fraction in %. This is referred to as the normalised seasonal component (*NSC*). This represents the changing (relative) proportion or fraction of *GVF*, expressing seasonality, mathematically being the complementary element of (1a).

Since complementary, the normalised elements (1a) and (2a) are not independent from each other, while this is true for the absolute elements (1) and (2).

A number of additional parameters can be calculated using the compliance obtained by the fitted model. These parameters help to interpret the results and allow statements about systematic deviations of individual years. The following parameters are calculated:

4) The yearly correlation coefficient for the seasonal variation versus the sine model
5) The yearly mean deviation of the model versus the original data
6) The yearly *RMS* deviation of the model versus the original data

SINFIT output classification

This classification was implemented with the purpose of deriving olive grove sub-classes of different management intensity levels.

The spatial olive area data is based on CLC 2000 with a spatial resolution of 100 m grid size. The classification considers all those pixels which contain a minimum of 1% olive area coverage within a coarse (0.01 degree) MEDOKADS pixel. Olive area fraction according to CLC was calculated for the MEDOKADS pixels by superimposing both data sets. Of those pixels complying with the mentioned requirements a further extraction is made which separates pixels of an area coverage of more or less than 40 % olive. With decreasing fraction of olive areas, the uncertainty of derived products regarding olives is decreasing; hence pixels of less than 40% olive area are extracted as separate class (class 6) and are not further considered. This threshold was chosen as the area of remaining olive pixels was approximately corresponding to the olive area of a re-sampled CLC map of 1 km^2 pixels. The principle of the further rules of classification is to distinguish between vegetation of high or low dominance of seasonal and/or permanent vegetation components. These rules were created in view of a subsequent class assignment to Beaufoy's olive farming types, expressing different management intensities (8). The approach is designed as a decision tree applying thresholds which are set by expert knowledge and/or are empirically derived from known olive sites. It takes into account the 3 already described SINFIT outputs *NPC*, *APC* and *ASC*, averaged over the years 1998-2002. The decision tree is depicted in Figure 3.

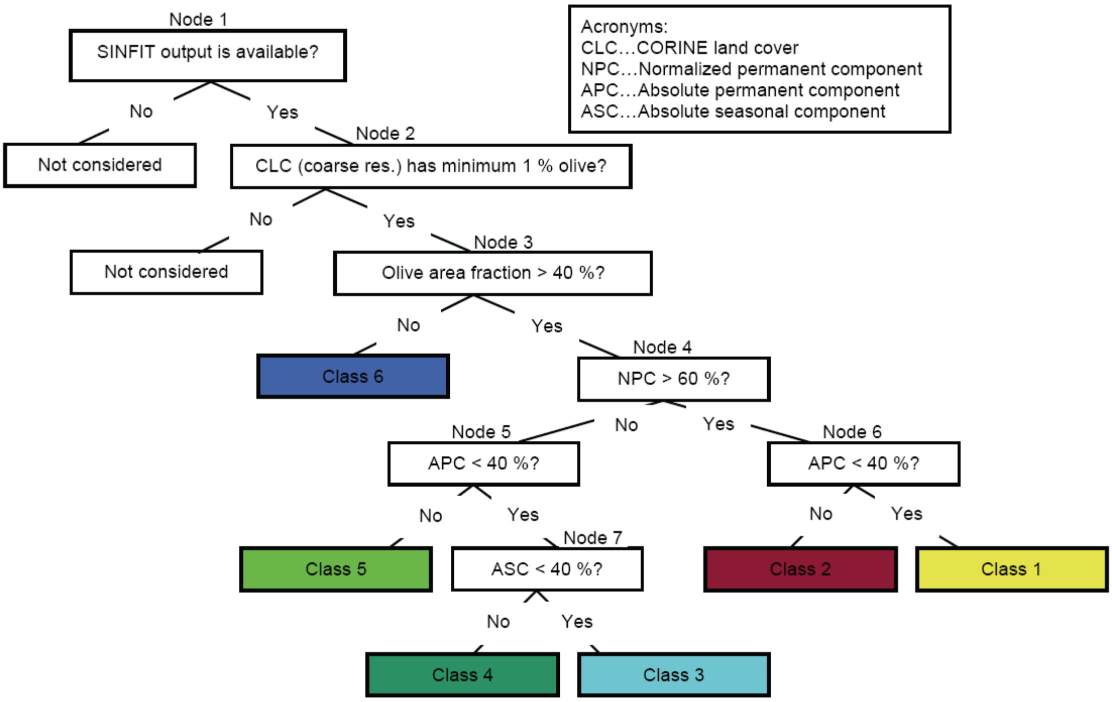

Figure 3: Decision tree for Olive grove management intensity classification. Description of the classes is given in the section 'Interpretation and discussion'.

RESULTS

An example for the seasonal run, as modelled with SINFIT, is depicted in Figure 4; with a permanent vegetation component of olive groves ranging in average between 22 and 30 % GVF. Depending on the year, the remaining seasonal vegetation component accounts for 15 – 30 % GVF in this example.

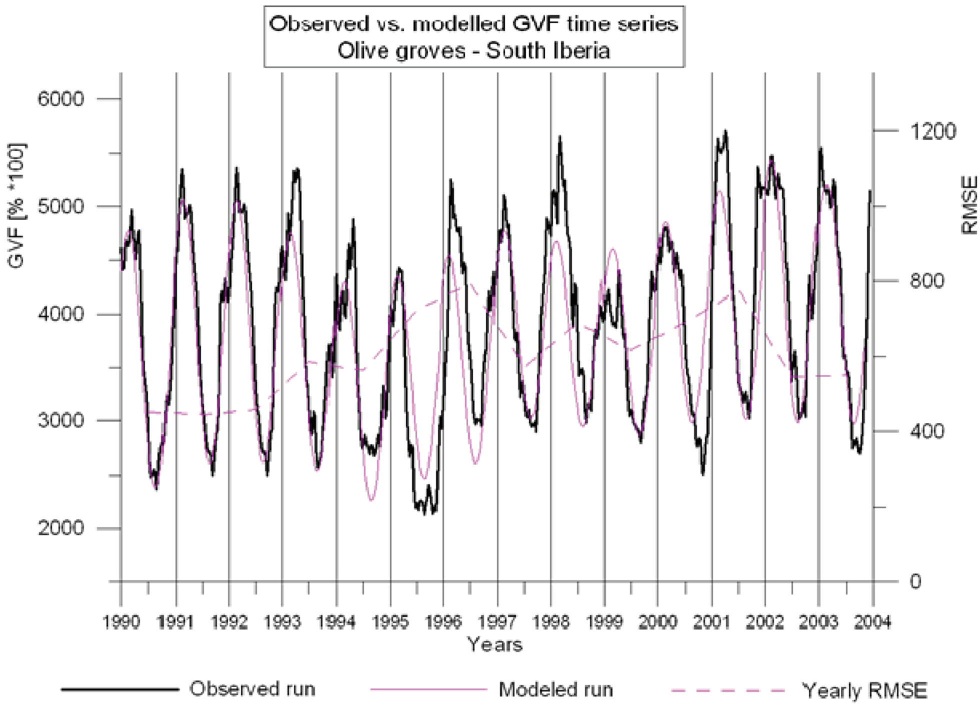

Figure 4: Observed GVF vs. SINFIT modelled GVF, averaged above the pixel aggregated CLC2000 class 'Olive groves' for Southern Iberia (corner points: 9.995° W, 40.715° N, 0.505 ° W, 35.025° N). Yearly RMSE displayed at right y-axis.

The RGB colour coded image in Figure 5 is depicting the prevailing vegetation components. Green pixels stand for predominantly permanent vegetation, while pink pixels stand for predominantly seasonal vegetation. Where both components are uniformly distributed, a greyish colour prevails. Strong seasonality takes place in mountainous regions but also in regions of strong agricultural activity with seasonal changes. Irrigated agricultural land which is covered by vegetation almost throughout the year (thus 'permanent' although annual vegetation) appears in intensive green.

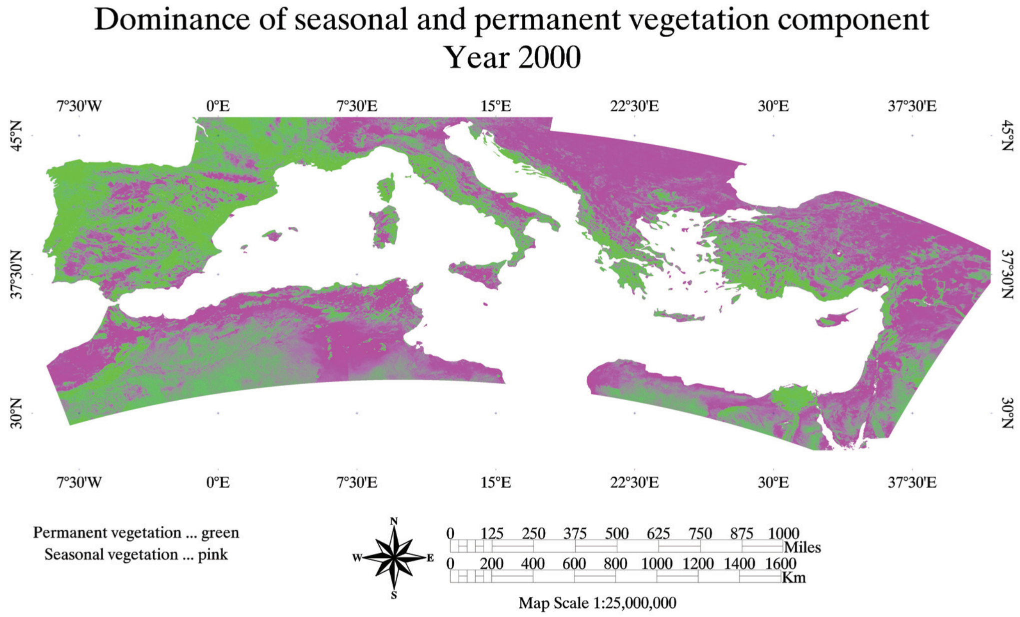

Figure 5: SINFIT output: Dominance of (normalised) seasonal and permanent vegetation components for the year 2000 in the Mediterranean area.

In Figure 6 the time intervals with the seasonal vegetation maximum have been colour coded. In general, it shows what would be expected in terms of dependence of growing season with geographical latitude, elevation and other climate influencing factors. It should be noticed that these results are meaningful only where reasonable amounts of periodical vegetation can be detected; hence this maximum variable is significant neither in the desert regions nor in regions of bi- or multimodal growing cycles (e.g. Nile delta). High altitude regions with snow cover almost throughout the year are not meaningful either, but this is due to the derivation of the input data itself.

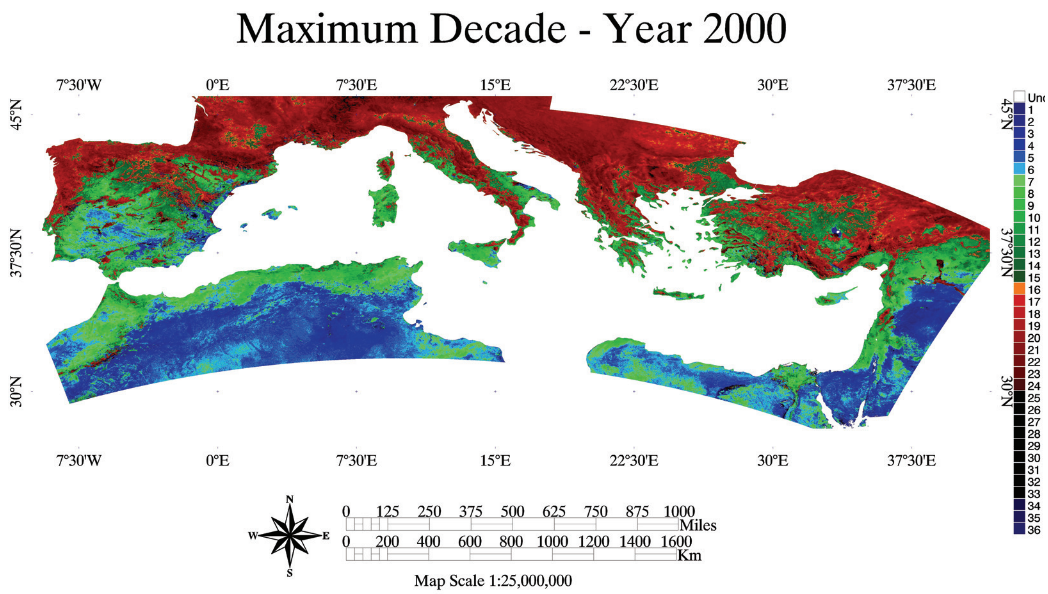

Figure 6: SINFIT output: Maximum decade for the Mediterranean in the year 2000.

Figure 7 shows the average of absolute deviation values of the model versus the observation over all available years in percent deviation. The error is generally low in vegetated areas (ranges around 10 %) and higher in areas with almost no vegetation (desert).

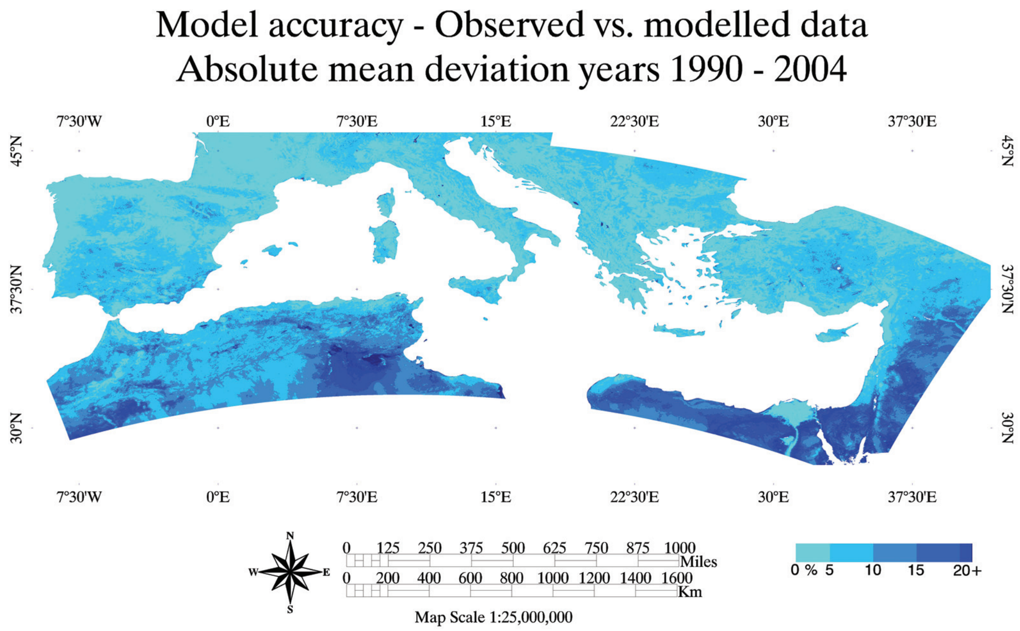

Figure7: SINFIT output: Correlation coefficient of observed GVF vs. modelled GVF.

Class map and statistics

Figure 8 displays the modelled output which is related to the CLC 2000 map and to the MEDOKADS data of the surrounding years (1998 – 2002), respectively. According to the relative class distribution of the whole Mediterranean, most olive areas are assigned to class 3 (40%), followed by class 1 (27%) and class 2 (20%). Classes 4 and 5 sum up to 12% (9% and 3%, respectively). Class interpretation is discussed in the next section.

INTERPRETATION AND DISCUSSION

Interpreting the modelled classes according to their biophysical properties (class 1 to 5) is essential for the assignment to intensity classes as described by Beaufoy (8). This author distinguishes the three classes 'intensive modern plantations', 'intensified traditional plantations' and 'low-input traditional plantations'. Intensity classes are, in the special case of olive groves, direct proxies of vegetation heterogeneity and vegetation dynamics represented by a single pixel. This heterogeneity is, however, not extracted from a spatial neighbourhood, but from the time domain of the time series.

Class 1 and 2 pixels are characterised by a high permanent vegetation component (> 60%) in combination with a weakly developed seasonal fraction. Class 1 represents all those pixels not reaching an absolute permanent component of 40% *GVF*. This class is assigned to 'intensive modern plantations'. Main indicator for the assignment is, besides the high permanent vegetation component ranging on a low level, the almost missing ground cover, reflected in low vegetation seasonality. Ground cover is, by physical or chemical means, almost fully removed in highly intense cultivations. Class 2 is assigned to 'intensified traditional plantations'. These pixels dispose of a higher level of permanent ground vegetation, often due to higher precipitation (and often located in mountainous areas).

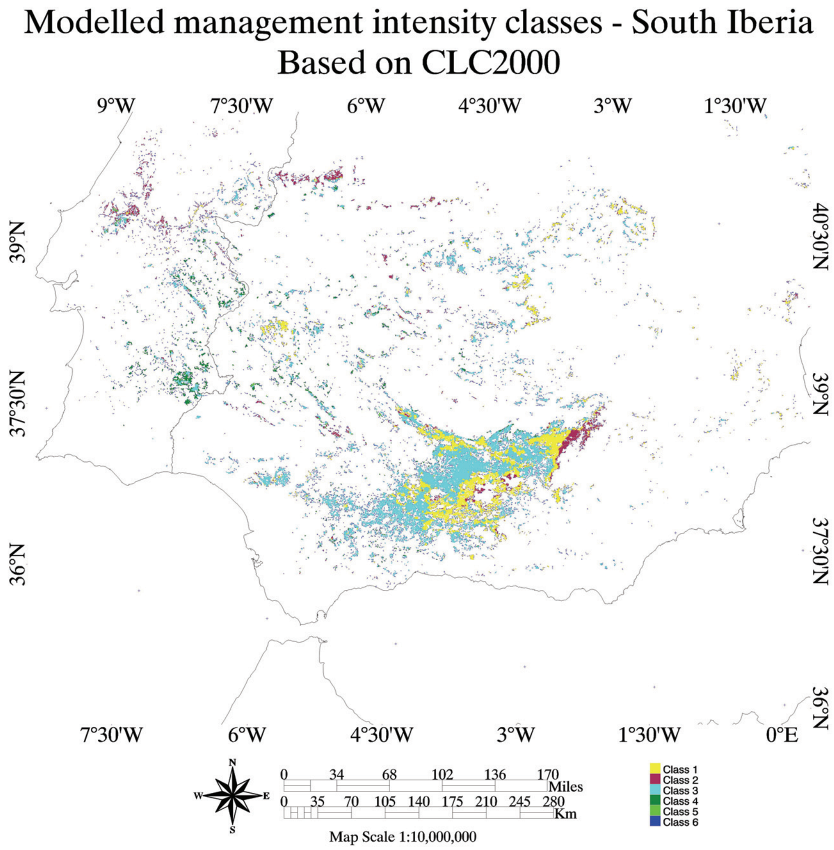

Figure 8: Detailed modelled management intensity classes for Southern Iberia.

Class 3, 4 and 5 are predominantly governed by a strong seasonal vegetation component. This group is further divided: Pixels with a low absolute permanent (< 40% *GVF*) and a low absolute seasonal component (< 40%) are the closest ones to pixels of the former classes 1 and 2; they are forming class 3. This class, representing a sort of intermediate class, with distinct dominance of the seasonal component, but still on a low absolute level, is assigned to 'intensified traditional plantations'. Class 3 was the most frequently occurring class. Class 4, on the other hand, represents a class of high seasonality for both normalised and absolute components. The high seasonality factor is typical for less intense cultivation forms, where grazing and/or mowing still occurs, or where a second crop (e.g. cereals) is cultivated together with olives. Class 4 is therefore most probably associated with 'low-input traditional plantations'. Also the rarely occurring class 5, which does not have limitations for the strength of the absolute permanent component, is associated with 'low-input traditional plantations'.

However, these thought classes represent ideal cases and in reality we are dealing with a continuum. Thresholds for the classes are derived empirically and/or set by expert knowledge. Obviously, the setting of thresholds may lead to different results. Also, mixed pixel information is influencing the SINFIT output. The limitations of the system are strongly linked with the spatial resolution and quality of MEDOKADS but also CLC data. The use of CLC data bears certainly problems which are mainly due to the consequences of the mapping approach by individual photo-interpreters and the minimum mapping unit of 25 ha. Nevertheless, a priori land cover knowledge was absolutely required for the underlying application, and CLC data presented the only available consistent data set on a pan-European extent.

Although the methodological approach using a sine curve for the modelling of the growing cycle is leading to a positive effect regarding data noise reduction for the applied data, it bears also some critical aspects. For example, it implies a symmetrical growing cycle, which might not be desirable in certain cases. However, the phenological product, which might be affected most by this shortcoming, did not find application in the presented product. Also, undershooting and overshooting of modelling, as visible for some years in Figure 4, might lead to misclassifications. However, this is certainly attenuated by averaging several years for classification (here 5 years), by use of normalised values in the classification. Also, robustness was considered when determining thresholds.

A first validation was done with data from the Spanish online Olive cadastre (13). High resolution aerial imagery allowed a basic assessment of olive plantations, taking into account criteria like extension, tree density, crown shape, ground cover, spatial arrangement of trees and roads, the 'mix' of crops, land cover, topography, etc. Although there was no systematic validation possible, it could be proved by random tests that the modelled classes fitted well with most probable management intensity level.

CONCLUSIONS

Deriving permanent and seasonal vegetation components of a long term time series represents a value added data product, which enriches the data content of the coarse spatial resolution data base. Permanent and seasonal vegetation components are key parameters of the growing cycle and when quantified, allow for a distinction of different vegetation types or biomes, or the derivation of trends. Quantification of 'permanence' and 'seasonality' occurs in the amplitude domain, and, additionally, in the time domain for 'seasonality' when deriving the maximum vegetation peak.

The SINFIT tool provides an automatic way to extract vegetation dynamics and to a certain extent phenology parameters, while only a minimum of parameterisation is required. The model SINFIT is appropriate for the analysis of relatively noisy NOAA AVHRR data, accounting for the data inherent dynamics while simplifying the growing cycle as much as possible. However, within this application not all of the potentialities of SINFIT have been used (e. g. phenology such as the growing cycle peak). An explicit trend or change detection analysis is not yet carried out either, but will be covered in a future publication.

The peculiarities of olive groves retrieved in a coarse remote sensing time series analysis allow a coherent assignment of the three management intensity classes described by Beaufoy (8). This is especially due to the fact that important management intensity characteristics of olive groves are linked to soil and ground flora status and not to the olive trees themselves. A conclusive validation of the outcome was difficult, as ground data was rarely available. However, statistics and orthophotography allowed limited validation which led to promising results.

SINFIT outputs offer a vast range of applications for remote sensing time series, which are becoming longer and more. Especially when coupled with trend analysis, changes of vegetation performance become a major point of interest. Applications for other land use classes or even groups of land use classes reveal important information about occurring dynamics and changes and represent the link to the socio-economic part of land use change. In this context, SINFIT allows derivation of important information for researchers but also policy decision makers.

ACKNOWLEDGEMENTS

Part of this work was performed under the IP DESURVEY project (IP Contract FP6 GCR Programme Contract No 003950) funded by the EC DG RTD 6th Framework Programme. In this context the MEDOKADS data set was prepared and made available by Dirk Koslowsky and colleagues from the Free University of Berlin. Within the same framework Marion Stellmes of Trier University contributed with software solutions to automated endmember extraction and *NDVI-Ts* unmixing. The contribution of the JRC HNV-team around M.L. Paracchini is greatly acknowledged and appreciated.

REFERENCES

1 Badeck F-W, A Bondeau, K Böttcher, D Doktor, W Lucht, J Schaber & S Sitch, 2004. Responses of spring phenology to climate change. New Phytologist, 162(2): 295-309

2 Justice C O, J R G Townshend, B N Holben & C J Tucker, 1985. Analysis of the phenology of global vegetation using meterological satellite data. International Journal of Remote Sensing, 6(8): 1271-1318

3 Reed B C, J F Brown, D VanderZee, T R Loveland, J W Merchant & D O Ohlen, 1994. Measuring phenological variability from satellite imagery. Journal of Vegetation Science, 5(5): 703-714

4 Hill M J & G E Donald, 2003. Estimating spatio-temporal patterns of agricultural pro-ductivity in fragmented landscapes using AVHRR NDVI time series. Remote Sensing of Environment, 84(3): 367-384

5 Azzali S & M Menenti, 2000. Mapping vegetation-soil-climate complexes in Southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. International Journal of Remote Sensing, 21(5): 973-996

6 Malingreau J-P, 1986. Global vegetation dynamics: satellite observations over Asia. International Journal of Remote Sensing, 7(9): 1121-1146

7 Allen H D, R E Randall, G S Amable & B J Devereux, 2006. The impact of changing olive cultivation practices on the ground flora of olive groves in the Messara and Psiloritis regions, Crete, Greece. Land Degradation & Development, 17(3): 249-273

8 Beaufoy G, 2001. The environmental impact of olive oil production in the European Union: Practical options for improving the environmental impact, http://ec.europa.eu/environment/agriculture/pdf/oliveoil.pdf (last date accessed: 8 Feb 2008) In: Studies and Reports on Agriculture. (European Commission, DG Environment, Brussels, Belgium) 74 pp. http://ec.europa.eu/environment/agriculture/studies.htm

9 Koslowsky D, H Billing & K Friedrich, 2006. Inter-comparison of NOAA/NASA Pathfinder and Medokads AVHRR based NVDI time series, http://ubt.opus.hbz-nrw.de/volltexte/2006/362/pdf/03-rgldd-session2.pdf (last date accessed: 8 Feb 2008) In: Proceedings of the 1st International Conference on Remote Sensing and Geoinformation in the Assessment and Monitoring of Land Degradation and Desertification, http://ubt.opus.hbz-nrw.de/volltexte/2006/362/, edited by A Röder & J Hill (Remote Sensing Department, University of Trier, Germany) 191-198

10 Stellmes M, S Sommer & J Hill, 2005. Regional vegetation monitoring based on the NOAA AVHRR time series in the Mediterranean area, http://www.ggrs.uni-goettingen.de/ggrs2004/CD/Applications_in_Geography/GGRS2004_Stellmes_G124.pdf (last date accessed: 8 Feb 2008) In: GGRS 2004 - 1st Göttingen GIS & Remote Sensing Days - Environmental Studies, http://www.ggrs.uni-goettingen.de/ggrs2004/. Remote Sensing & GIS for Environmental Studies: Applications in Geography, edited by S Erasmi, B Cyffka & M Kappas. Göttinger Geographische Abhandlungen, 113: 124-130

11 Stellmes M, S Sommer & J Hill, 2005. Use of the NOAA AVHRR NDVI-Ts feature space to derive vegetation cover estimates from long term time series for determining regional vegetation trends in the Mediterranean, http://ubt.opus.hbz-nrw.de/volltexte/2006/362/pdf/03-rgldd-session2.pdf (last date accessed: 8 Feb 2008) In: Proceedings of the 1st International Conference on Remote Sensing and Geoinformation in the Assessment and Monitoring of Land Degradation and Desertification, http://ubt.opus.hbz-nrw.de/volltexte/2006/362/, edited by A Röder & J Hill (Remote Sensing Department, University of Trier, Germany) 231-238

12 EEA, 2005. Corine land cover 2000 (CLC2000) 100 m - version 8/2005 http://dataservice.eea.europa.eu/dataservice/metadetails.asp?id=822 (European Environment Agency, Copenhagen, Denmark) (last date accessed: 8 Feb 2008)

13 Ministero de Agricultura, Pesca y Alimentación; 2006. Sistema de información geográfica oleícola, http://www.mapa.es/es/sig/pags/sig/intro2.htm (last date accessed: 08 Feb 2008)