ALTERRA
WAGENINGEN UR

Phenology related measures and indicators at varying spatial scales

Investigation of phenology information for habitat classification using SPOT VGT and MODIS NDVI data

Alterra Report 2259
ISSN 1566-7197

Nicola Clerici, Christof J. Weissteiner, Andrej Halabuk, Gerard Hazeu, Gerbert Roerink and Sander Mücher

ALTERRA
WAGENINGEN UR

SEVENTH FRAMEWORK PROGRAMME

EBONE
EUROPEAN BIODIVERSITY OBSERVATION NETWORK

Phenology related measures and indicators at varying spatial scales

This research has been carried out in the framework of the EC FP7 project EBONE (EC-FP7 Contract ENV-CT-2008-212322)

Phenology related measures and indicators at varying spatial scales

Investigation of phenology information for habitat classification using SPOT VGT and MODIS NDVI data

Nicola Clerici[1], Christof J. Weissteiner[1], Andrej Halabuk[2] , Gerard Hazeu[3], Gerbert Roerink[3], Sander Mücher[3]

(Edited by C.A. Mücher)

[1] European Commission, Joint Research Centre, Ispra, Italy
[2] Institute for Landscape Ecology, Slovak Academy of Sciences
[3] Alterra Wageningen UR

AlterraReport 2259

Alterra, part of Wageningen UR
Wageningen, 2012

Abstract

Nicola Clerici, Christof J. Weissteiner, Andrej Halabuk, Gerard Hazeu, Gerbert Roerink, Sander Mücher, 2012. *Phenology related measures and indicators at varying spatial scales. Investigation of phenology information for habitat classification using SPOT VGT and MODIS NDVI data*. Wageningen, Alterra, Alterra-Report 2259. 112 pp.; 33 figs.; 15 tab.; 34 ref.

Abstract: The main objective here is to investigate if leaf phenology indicators as derived from SPOT and MODIS NDVI time series can provide useful information for the detection, characterization and mapping of habitats, with specific reference to the General Habitat Category and Annex I (Natura 2000) schemes. The report is divided into three main parts. Part I "Extraction and analysis of phenology indicators", using the Phenolo model of the Joint Research Centre (JRC, Ispra), and a phenological characterization and classification of test sites using Random Forest Classification, and an intercalibration of GHCs with MODIS-derived phenometrics. A set of 31 leaf phenology indicators (phenometrics) was extracted using JRC Phenolo model from time series of NDVI 10 day Maximum Value composites with 6 years of MODIS satellite data and 11 years of SPOT data. Classifications to discriminate deciduous and coniferous forest were performed in selected regions using MODIS satellite data. The main sources identified for low classification accuracy are both the large heterogeneity allowed by the GHC scheme for forests (tree cover proportion), and the low number of training points currently available from field survey. Part II "Multi-temporal analysis of NDVI for grassland mapping and classification", focusses on two specific case study areas for grassland mapping. Part III 'EO time series analysis to identify Annex I habitat types' describing the processing of MODIS medium resolution time series with the HANTS algorithm and the use of it's mathematical derived components for the identification of specific Annex I habitats, such as H9150 and H4060. A major limitation to identify those habitats is that the used times series have a minimum spatial resolution of 250 meters which can still be considered as too coarse for most habitats. Next to the fact that not all Annex I habitats have their unique phenology. A GHCs intercalibration strategy integrating EO-based phenology with life forms information, e.g. from LiDAR, would be potentially more effective than a purely phenology-based approach.

Keywords: Phenology, Phenolo model, HANTS algorithm, MODIS, NDVI, classifications, intercalibration, General Habitat Categories, Natura 2000.

Foto omslag: High Tatra (Slovakia); R.H.G. Jongman

ISSN 1566-7197

The pdf file is free of charge and can be downloaded via the website www.alterra.wur.nl (go to Alterra reports). Alterra does not deliver printed versions of the Alterra reports. Printed versions can be ordered via the external distributor. For ordering have a look at www.rapportbestellen.nl.

Alterra Report 2259
Wageningen, January 2012

Contents

Summary

The main goal of the present work within the context of the EBONE objectives is to investigate if leaf phenology indicators as derived from SPOT and MODIS NDVI time series can provide useful information for the detection and mapping of forest habitats, with specific reference to the General Habitat Category scheme.

The report is divided into three main parts. Part I "Extraction and analysis of phenology indicators", describing the Phenolo model of the Joint Research Centre (JRC, Ispra), and a phenological characterization and classification of Forest Phanerophytes in selected test sites. Part II "Multi-temporal analysis of NDVI for grassland mapping and classification", focusses on two specific case study areas for grassland mapping. Part III 'EO time series analysis to identify Annex I habitat tyes' describes the processing of MODIS medium resolution time series with the HANTS algorithm and the use of its mathematical derived components for the identification of specific Annex I habitats.

PART I focuses on the use and description of the Phenolo model. This includes the pre-processing and processing steps applied to extract leaf phenology indicators from SPOT and MODIS data, a short analysis of the spatial distribution of a selection of phenometrics in test areas, and classification of GHC Forest Phanerophytes using field data and MODIS-derived phenometrics. It introduces two pilot habitat classification tests using the Random Forests™ classification approach and SPOT NDVI data. The last part focuses on investigating the intercalibration of GHCs with MODIS-derived phenometrics. Random Forests classifications were tested in a variety of configurations and accuracy checked using the JRC Forest Map 2006.

A set of 31 leaf phenology indicators (phenometrics) was extracted using JRC Phenolo model from a time series of NDVI ten day Maximum Value composites of six years (MODIS) and eleven years (SPOT).The Phenolo model considers an annual cycle of vegetation leaf phenology as represented by one permanent component, or 'background' and a variable component, function of seasonal dynamics. Pre-processing involved substitution of no data, outlier analysis and filtering. NDVI time series processing involved the extraction of date and productivity phenometrics. The model, coded in IDL, provided fast calculations in a stable environment.

The performance of the Random Forests classifications and the contribution of individual phenometrics were tested through the calculation of the Mean Decrease Accuracy parameter (MDA). Overall, the results suggest date phenometrics to be more important for forest habitat classification than productivity phenometrics, especially indicators defined around the Peak of Season point and the NDVI curve minima. Apart from areas with spatially and spectrally homogeneous forest habitat classes (Coniferous forests in Austria), the overall classification accuracy achieved with the Random Forests approach using MODIS-based phenology indicators is generally not satisfactory. We identified three main factors influencing these result: the spatial/spectral heterogeneity present in the GHC forest polygons and subsequently in the training pixels associated to these classes: the low number of training pixels available and the use of an independent dataset to calculate accuracy which was built uniquely on spectral information. The introduction of artificial data gaps within the MODIS NDVI time series did not influence significantly classification accuracy.

On the basis of the investigation results, the following remarks were made: 1) the spatial scale of current EO-based phenology data (250 m) is at the edge of an adequate resolution for effective habitat classification with respect to GHC categories and field data; 2) It is recommended to build a large dataset of GHC training pixels in order to take into account the high spectral variability present within single GHC classes and 3) Adequate

classification accuracy assessment should be based on a reference dataset which takes into account as much as possible the elements of heterogeneity typical of the GHCs.

The structural (height) characteristics of the life forms types considered in the General Habitat Category scheme are very valuable information which should be taken into account when using EO-derived information. For this reason, for the purpose of GHCs classification a strategy that integrates EO-based phenology indicators with EO derived information on vegetation structure, from e.g. LiDAR or high resolution radar, could potentially be more effective than only a phenology-based approach.

Part II "Multi-temporal analysis of NDVI for grassland mapping and classification", focusses on two specific case study areas for grassland mapping. Grassland seasonal pattern of NDVI can vary substantially reflecting not only the differences in vegetation type but also land use, management practices or site hydrology. This fact mainly limits mapping of grassland as a single land cover class especially when above factors are evenly present across the study area. This was demonstrated in Slovakia by the supervised image classification where one "grassland signature" was confused with many forests, and arable crops. In fact, different land covers may have similar seasonal patterns of productivity (for example, some shrubs and the unmanaged grasslands); conversely, the same land-cover type may have different NDVI dynamics (for example, the intensive grasslands and extensive grasslands. On the other hand in unsupervised approach clusters with similar seasonality were merged together and misclassification was introduced by our attempt to attribute distinct seasonal information to respective land cover class. We demonstrated the limitation of using supervised approach for a full coverage classification at broader scale which required more effort in training compared to the relatively easy labelling of an unsupervised product. It seems that PCA approach were more suitable for exploring distinctive characteristics of NDVI temporal profile across different land cover classes.

Combined approach with HR sensors is suggested in heterogeneous landscape. In our examples we tried to briefly demonstrate capabilities and limitations of multi-temporal approaches for grassland mapping and classification. It seems that for the mapping of grasslands in heterogeneous landscape, specific approaches need to be further explored for increasing of mapping capabilities of multi-temporal analyses. These approaches (including seasonal based classification of grasslands) need to be tested in near future across contrasted landscapes. Anyway, explanatory analysis and classification of grasslands using available sample data revealed that specific features of grasslands can be detected and reasonable classification made what proves that multi-temporal analysis should represent a valuable tool mainly in the assessment and monitoring component of the proposed biodiversity observation system.

Part III 'EO time series analysis to identify Annex I habitat types' describes the processing of MODIS medium resolution time series with the HANTS algorithm and the use of its mathematical derived components for the identification of specific Annex I habitats H9150 and H4060. As the objective here is to enhance the spatial distribution of European Annex I habitats, read Natura 2000, based on their phenology. For this reason NDVI-time series have been analysed that could be processed for the whole of Europe. There is a demand for a high temporal resolution together with a spectral resolution that allows the calculation of the Normalized Difference Vegetation Index (NDVI). The best suitable sensor, concerning a high temporal resolution and adequate spectral and spatial resolution, is MERIS (300m) or MODIS (250m). Since the latter is easy and freely downloadable, we used MODIS satellite data for our purpose. MODIS has a daily revisit time with a spatial resolution of 250 meter. The seasonal cycle of the NDVI can be approximated by a limited number of frequency components derived from a Fourier analyses. This principle is implemented in the HANTS algorithm (Harmonic Analysis of NDVI Time-Series) which employs an iterative routine to filter out poor NDVI estimates due to cloud cover or other disturbances from the NDVI cycle. The basic concept behind the algorithm is that the vegetation development as indicated by the NDVI has a strong seasonal effect in most parts of the world (apart from the tropics) which can be described using a series of low frequency sine functions with different phases, frequencies and amplitudes. Based on the HANTS analysis of the MODIS time series with a 250 meter

spatial resolution, classifications were made separately for each environmental zone as the vegetation development between zones is very different due to biophysical conditions. The HANTS results were classified with the Maximum Likelihood parametric rule based on a signature file. The signature file contained two groups of training sets: i) general land cover signatures and ii) specific signatures related to the selected habitats. For the specific habitat-training sites the in situ data (vegetation relevés) were used in combination with the CORINE Land Cover database.

The distribution of the habitats H4060 and H9150 were classified on basis of the HANTS results for the Alpine South and Continental Environmental zone. The assessment of the classification results suggests an overestimation of the H9150 habitats in both environmental zones. An underestimation is present in the case of H4060. However, the classification of habitats on basis of satellite imagery needs improvement. The main limitation is the lack of more detailed (higher resolution) HANTS vegetation phenology product, next to the fact that many habitats do not have a unique phenology. The present spatial resolution of the times series analysis was 250 meter based on the MODIS satellite imagery, while most of the Natura 2000 habitats are still very fragmented at this scale. The quality of the classification results differ between habitats and between environmental zones. Generic classification parameters valid for all kinds of habitat-environmental zones combinations will be an utopia. Habitats differ in reflectance from each other and differ between environmental zones as the biophysical conditions and the phenology development is different.

Introduction

The following investigation is part of the research activities of the European Biodiversity Observation NEtwork project (EBONE). The project's main aim is to establish a framework for an integrated biodiversity monitoring and research system based on key biodiversity indicators, and implemented at institutional European level.

EBONE will:
- Design a biodiversity observation network based on current national capabilities.
- Develop techniques for upscaling between sites, habitats and remote sensing data to detect and understand changes in indicators and ecosystems.
- Recommend refinements to current observation systems, and for the implementation of the proposed system in Europe.
- Integrate measurements and data structures within existing data management systems.
- Develop and test the worldwide compatibility of the system in regions outside Europe.

The EBONE project is implementing a habitat classification scheme based on General Habitat Categories, GHCs (Bunce et al., 2010). The main goal of the present investigation in the context of the EBONE objectives is to investigate if phenology information (phenology indicators) as derived from NDVI time series can provide valuable information for the identification and mapping of forest habitats as defined in the GHCs scheme (intercalibration of EO phenology data with GHC forest phanerophytes).

More detailed objectives are:
- Derive phenology indicators for continental Europe using SPOT Vegetation satellite data with a 1 km spatial resolution and MODIS satellite data with a 250 meter spatial resolution.
- Investigate advantages and limitations of phenology based indicators for identification for forest, heathland and grassland habitats
- Provide recommendations for the intercalibration of EO-derived phenology indicators and General Habitat Categories.

The report is divided into three main parts. Part I "Extraction and analysis of phenology indicators", describing the Phenolo model of the Joint Research Centre (JRC, Ispra), and a phenological characterization and classification of test sites. Part II "Multi-temporal analysis of NDVI for grassland mapping and classification", focusses on two specific case study areas for grassland mapping. Part III 'EO time series analysis to identify Annex I habitat types' describes the processing of MODIS medium resolution time series with the HANTS algorithm and the use of its mathematical derived components for the identification of specific Annex I habitats.

PART I: EXTRACTION AND ANALSYSIS OF PHENOLOGY INDICATORS

Nicola Clerici and Christof J. Weissteiner

European Commission, Joint Research Centre, Ispra, Italy

1 Remote Sensing Data

This section focused on NDVI time series from the SPOT VGT and MODIS sensors, provided by the MARS Unit of the IPSC institute (EC JRC). The spatial resolution is 1km for SPOT VGT data and 250 meters for MODIS. The time series are composed of 10-days Maximum Value Composites (MVC) of NDVI, covering the following temporal windows:

- SPOT VGT: from 1999 to 2009 (11 full years).
- MODIS: from 2004 to 2009 (6 full years).

NDVI data was prepared for EC JRC by the Flemish Institute for Technological Research (VITO NV) and includes atmospheric correction, cloud detection, and calibration (Rahman and Dedieu, 1994; Paola and Schowengerdt, 1994; Klisch et al., 2005). Both NDVI datasets cover the whole of Europe. An example of SPOT NDVI mosaic for Europe is shown in Figure 1. Data pre-processing and processing is fully discussed in next sections.

Figure 1

Example of SPOT NDVI decade data for Europe (light green: high NDVI; dark green: low NDVI; white: water/no data). Third 10 day-period in June 2007.

2 Extraction of Phenology indicators

2.1 Background

It is a frequent assumption in phenology time series analysis to assume regular patterns that describe the vegetation leaf seasonal cycles (Jonson and Eklundh, 2004). An annual season cycle can be described in general terms as represented by a) one component which is the permanent signal, or 'background' and b) a variable component which is function of the seasonal dynamics (e.g. Weissteiner et al., 2007). The latter is generally characterized by a growing period, during which the vegetation signal intensity increases, it reaches at a certain time a peak of maximum signal (t_{MAX}), and it decreases towards the background level (senescence period). An illustrative scheme is shown in Figure 2.

This behaviour is ideally reflected in the NDVI signal pattern. In the real case this pattern is influenced by a number of variables which shape and modify the NDVI signal, such as meteorological factors (snow, frosts, cloud coverage) or changes in the vegetation (*e.g.* land cover change processes, health status, drought effects, etc.).

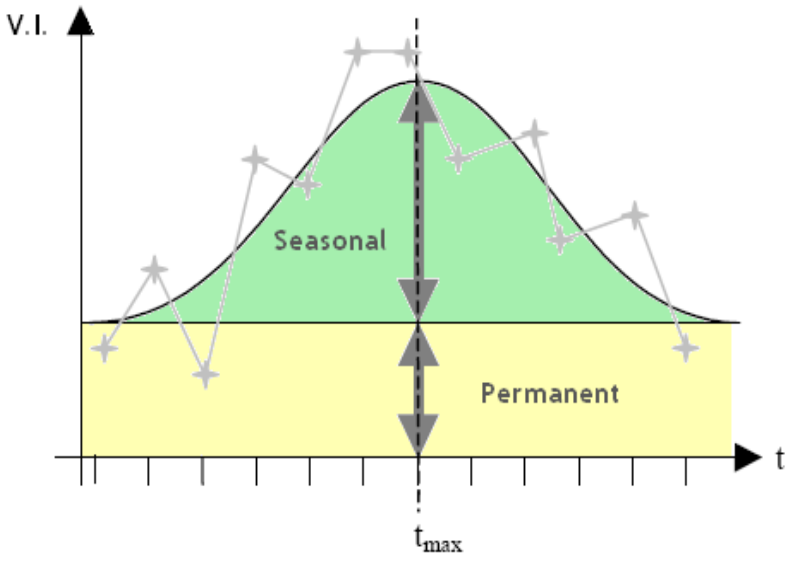


Figure 2
Observed (grey crosses) and seasonal/permanent components of a theoretical vegetation cycle (modified after Weissteineret al., 2007).

The modelling approach, developed by EC JRC (Ivits et al., 2009) and on which the phenometric extraction is based, relies on the use of smoothing and moving average algorithms as a basis to extract a large set of phenology indicators. This model, called *Phenolo*, is explained in detail in Section 2.3.The version of Phenolo model used in this investigation refers to 2009; a new version with substantial changes in amount and types of indicators is currently under development.

2.2 Pre-processing

2.2.1 SPOT VGT NDVI Data

A single layer stack of SPOT NDVI data was prepared containing only those years for which full time series were available. The data series ranges from 1999-2009, comprises 396 10-days maximum value composites (hereafter called decades) of SPOT NDVI (built according to Holben, 1986). Missing NDVI values were already flagged as missing, clouds, snow and out rocks in the original raw data. Missing values (flagged pixels) were substituted by seasonal means when available (mean of NDVI for that decade for available years). Pixels with no seasonal mean (e.g. snow-covered throughout the same time periods of each year) were flagged. Outliers detected according to Chebychev's theorem (95% confidence interval) were substituted by seasonal means (Lohninger, 1999). Remaining missing data were given a linear interpolated NDVI value using the nearest existing data points in time. Finally, NDVI data were filtered using a Savitzky-Golay filter (Chen et al., 2004), temporal window size of 6 decades, see Figure 3. An aggregated mask was also created adding up all single decadal masks (36) for which no seasonal mean could be calculated, and combined with a water mask (Figure 4).

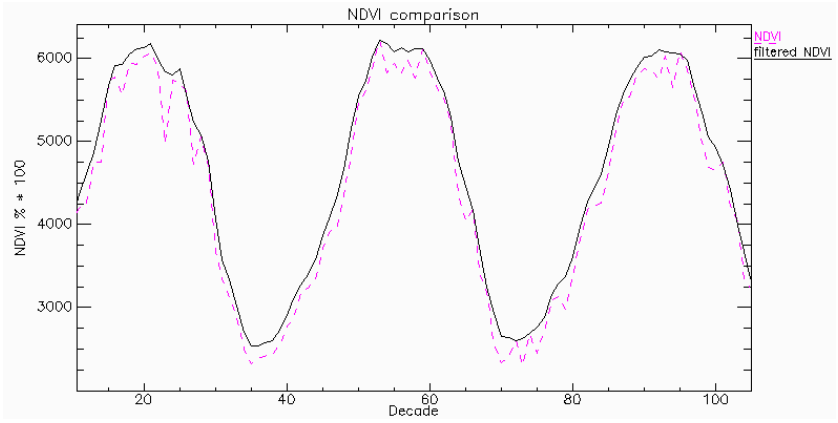


Figure 3

Example of NDVI time series before and after filtering operations with Savitzky-Golay filter (SPOT VGT data).

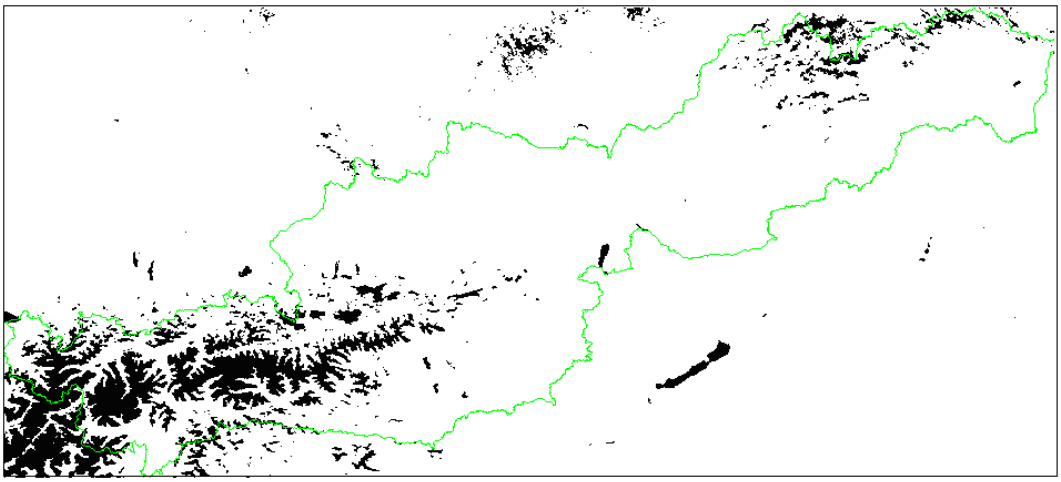


Figure 4

No data and water mask (black) for Austria and Slovakiatest areas (SPOT VGT data).

2.2.2 MODIS NDVI Data

Also for MODIS NDVI data a single layer stack was prepared comprising six years with full data availability. For 2004-2009, the time series contained 216 10-days Maximum Value Composites (MVCs) of MODIS NDVI. Missing values and outliers were substituted by seasonal means, applying the same criteria as for SPOT data. The remaining missing data were given an interpolated value between the nearest existing data points in time. Also in this case the data were filtered using a Savitzky-Golay filter (Chen et al., 2004) using the same specifications described in Section 3.2.1, and an aggregated mask was created synthesizing all single decadal masks (36) for which no seasonal mean could be calculated. A 'No data' flag layer was also derived by counting the number of decadal data gaps that are interpolated (see Figure 30). The following chart resumes the entire pre-processing chain.

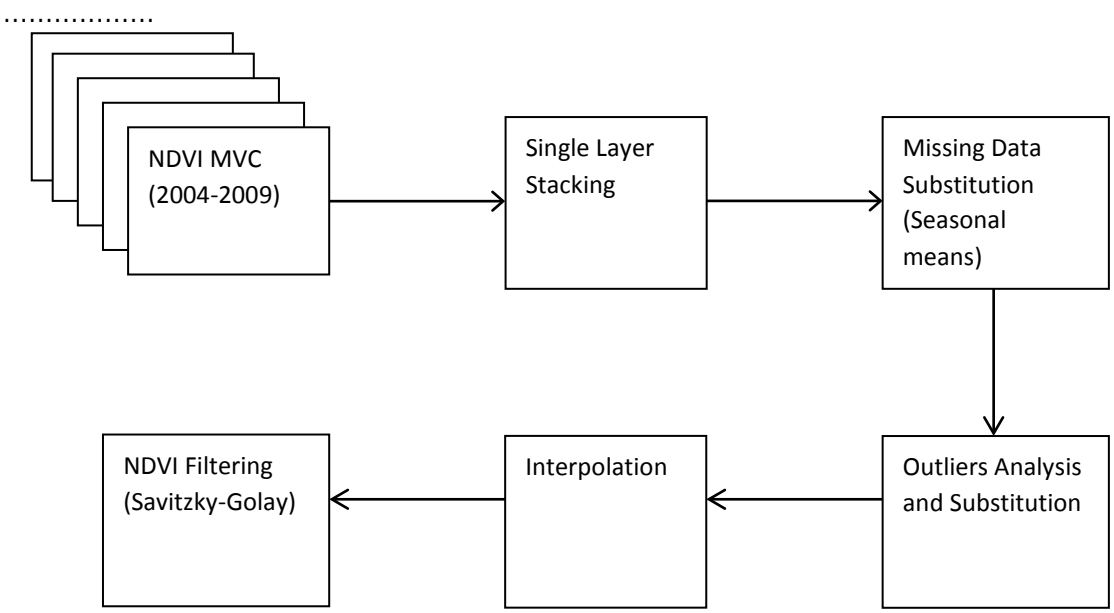


Figure 5
Pre-processing chain applied to SPOT NDVI and MODIS NDVI data.

2.3 The phenological model

The purpose of the Phenolo model, developed by EC JRC (Ivits et al., 2009), is to extract leaf phenological indicators from time series of vegetation indices. In the first step the model applies a median filter on a sliding temporal window of five successive time points. This is followed by the calculation of one forward and one backward lagging curve using a moving average algorithm. For example, for a forward lag, an *x*-day moving average value of time point *p* is calculated as the average of values for the *x* time points from *(p-x)* to *p*. The resulting averaged values will always reach similar magnitudes as the original p values later in time. The lag distance (defined in terms of the number of successive time points *x*) is defined by applying one standard deviation from the bary centre of the integral surface of the curve (Figure 6,Ivits et al., 2008). This value can be changed according to analysis needs.

Phenology Indicators

A number of studies investigated vegetation dynamics using information from date phenometrics (e.g. Reed et al., 1994; Hill and Donald, 2003); all of them considered a start and end of growing season. Following Reed et al. (1994), the start of the growing season (SOS in Figure 6) is defined as the crossing point between the smoothed curve and the forward lagged curve. The same criterion applies for the end of season (EOS), defined as the intersection between the backward curve and the smoothed one. The point corresponding to the maximum value of the vegetation signal is the Peak of Season (POS in Figure 6). The Growing Season End (GE) is defined as the point when, following the EOS, the forward lagged curve that defines SOS intersects the signal curve. The EOS, SOS, POS and GE points define two phenology indicators each: the correspondent day and NDVI value (see Table 1). The time interval in days between SOS and EOS defines the Season Length (SL), while the time interval between the minima in the phenology curve is referred in the model as Total Length (TL). For further details on the phenology indicators construction see Ivits et al. (2008, 2009).

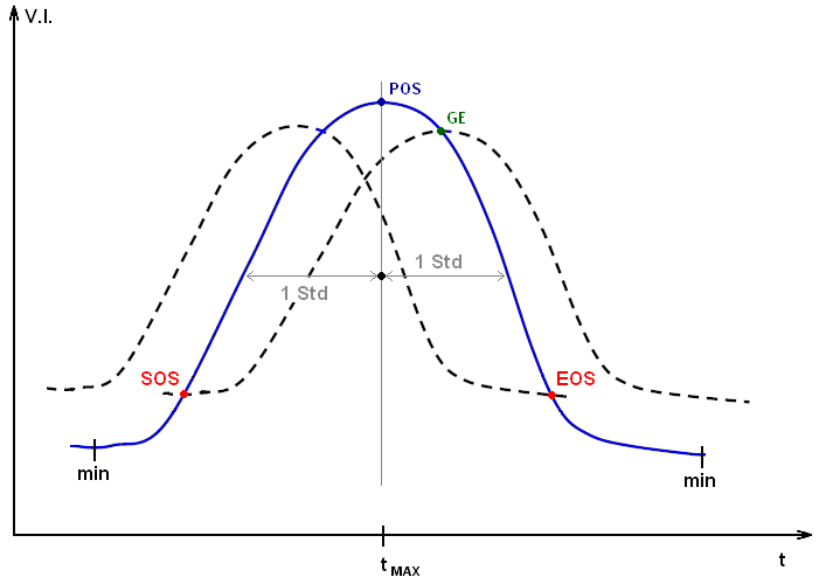


Figure 6

Smoothed curve (blue) and forward and backward lagging curves (dotted) defining date phenometrics in Phenolo ver. 2009 (after Ivits et al., 2008).

From the phenology curve it is also possible to define a series of productivity phenometrics (e.g. in Figure 7), a selection of which are:

- *Seasonal Permanent Fraction* (SPF), defined by the area between the line connecting Start and End of season and the x axis.
- *Season Integral* (SI), is the integral under the vegetation signal curve delimited by the start and the end of season.
- *Total Permanent Fraction* (TPF),is the area between the timeline connecting the vegetation signal minima and the x axis.
- *Total Integral* (TI), is the integral under the vegetation signal curve delimited by the two vegetation signal minima. TI is a proxy that represents an approximation of the Net Primary Productivity.

The GE point defines the Growing season Integral (GI) and derived integrals. Other phenology indicators, derived mathematically from the phenometrics listed above, are also extracted by Phenolo (see Table 1, indicators list and short explanation). The development of Phenolo is still in evolution, consequently all derived parameters' description and their use are related to the version in time which was the only available at the beginning of this research (v.2009); for this reason the calculation of certain variables is not guaranteed in future versions.

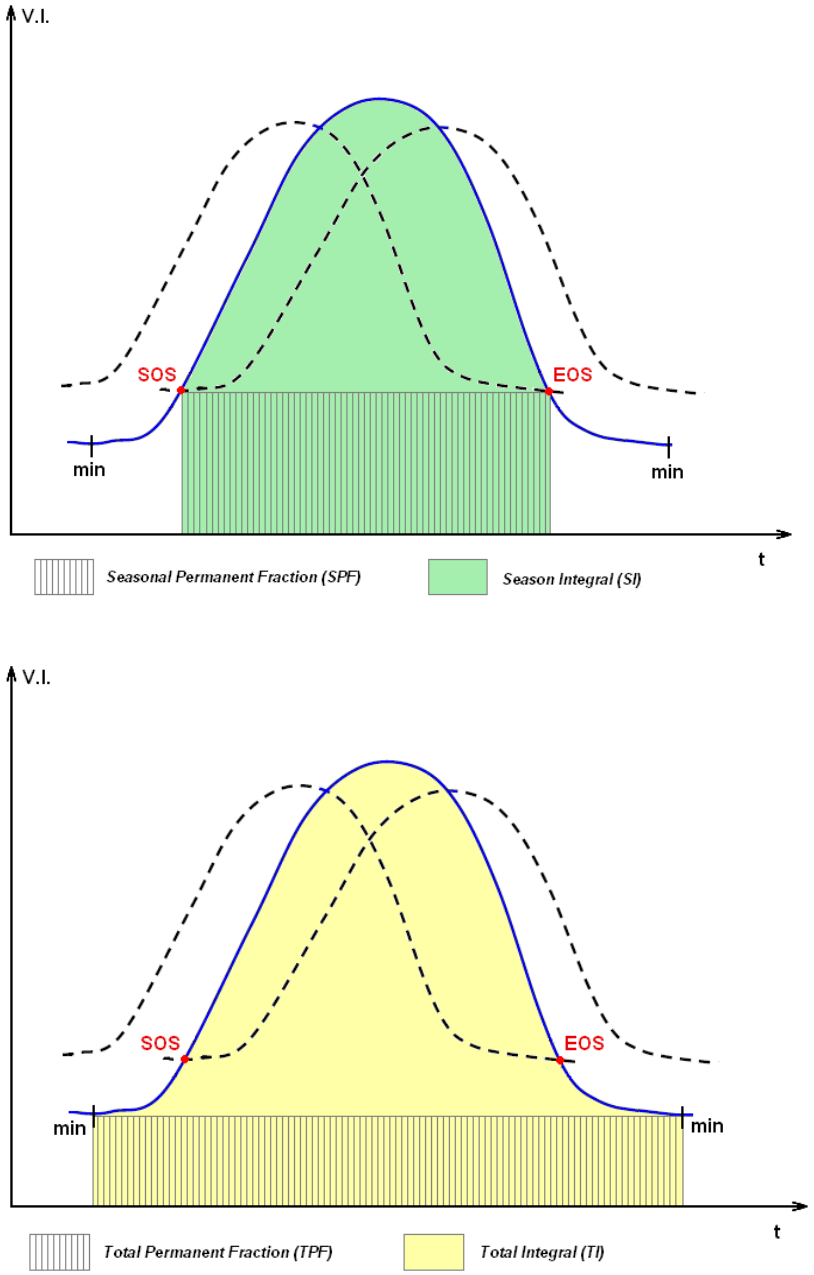


Figure 7

A selection of productivity phenometrics extracted by Phenolo v.2009 (modified, after Ivits et al., 2009).

Table 1

*Phenometrics extracted by Phenolo (2009 version), with short explanation and acronyms defined in the model. (*GTR was finally discarded).*

Acronym used in text	Phenometric extracted and short explanation	Acronym in Phenolo
SOS	Start of season (Day)	SBD
"	Start of season (Value)	SBV
EOS	End of season (Day)	SED
"	End of season (Value)	SEV
SL	Length of season (EOS-SOS)	SL
SI	Season Integral: the integral under the vegetation signal curve delimited by EOS and SOS	SI
SNI	Normalized SI	SNI
SPF	Seasonal Permanent Fraction: the area below the line connecting SOS with EOS, and the x axis.	SPI
STR	Season Total Ratio [SI/(SI+SPF)]	STR
GE	Growing season End (day)	GED
GE	Growing season End (value)	GEV
GL	Growing season Length	GL
GI	Growing season Integral	GI
GNI	Normalized GI	GNI
GTR*	Growing season Total ratio [GI/(GI+SPF)]	GTR
GPI	Growing season Permanent Fraction: the permanent area fraction below the curve connecting SOS with End of growing season	GPI
MBD	Minimum before SOS (Day)	MBD
MBV	Minimum before SOS (Value)	MBV
MED	Minimum after EOS (Day)	MED
MEV	Minimum after EOS (Value)	MEV
TL	Total Length: Length in time between minima (Days)	ML
TI	Total Integral: the area under the vegetation signal curve delimited by the two minima.	MI
NTI	Normalized TI	MNI
MTR	Above Minima Total Ratio: above minima integral over TI	MTR
TPF	Total Permanent Fraction: the area below the line connecting the vegetation signal minima and the x axis.	MPI
SEI	Season Exceeding Integral: (TI-SI)	SEI
GEI	Growing Season Exceeding Integral: (TI-GI)	GEI
SBC	Season Barycentre	SBC
SSD	Standard Deviation of the Season vegetation curve	SSD
POS	Peak of Season (Day)	MXD
"	Peak of Season (Value)	MXV
OMI	Output minus Input Length (365 - GL)	OMI

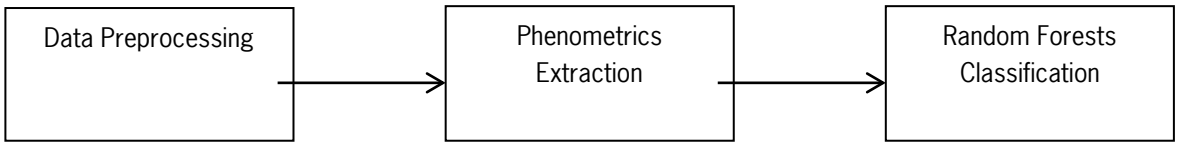
3 Analyses

3.1 Exploring the use of NDVI derived phenometrics using SPOT Data

3.1.1 Introduction

Extraction of phenometrics was performed for both SPOT and MODIS NDVI time series. Pilot tests were carried out using SPOT NDVI data, as the processing of SPOT data is less computationally intense. Austria and Slovakia were selected as study areas. The pilot study involved: 1) testing the feasibility of methods applied and implemented throughout the different stages from phenometrics extraction to forest habitat / species classification; 2) deriving measures of reliability of the Random Forests classifier, using Natura2000 forest habitat data and AFOLU Trees distribution data (Koebleand Seufert, 2001).

The three main sequential blocks of data processing carried out were as follows:



A short phonological characterization of the study area is shown using date and productivity phenometrics. A formal classification accuracy assessment was not performed for the SPOT analysis, as the main aim of this task was to test the phenology model and the Random Forests classification routines. Instead, a systematic accuracy assessment was performed in the MODIS NDVI classification in the second part of the analyses work(Section 3.2).

3.1.2 Phenological Characterization of test sites

To provide a general phenological characterization of the test area, the average and standard deviation of the phenometrics extracted were calculated using as spatial units the *Environmental Strata (EnS)* proposed by Metzger et al.(2005). In other words, the mean value of the phenometric for the NDVI time series was averaged over all locations belonging to each EnS. Metzger's bio-geographical division, together with the upper hierarchy level (Environmental Zone, EnZ) is considered as an up-to-date and appropriate stratification for environmental modelling exercises and reporting in the European region. Figure 8 shows the Environmental Zones and Environmental Strata present in the study area. The region is dominated by the Alpine South (ALS), Continental (CON) and Pannonian (PAN) EnZ, with small areas representing the Mediterranean Mountains (MSM) and Alpine North (ALN). A further sub-division in Environmental Strata is also shown. For detailed construction of this stratification see Metzger et al. (2005).

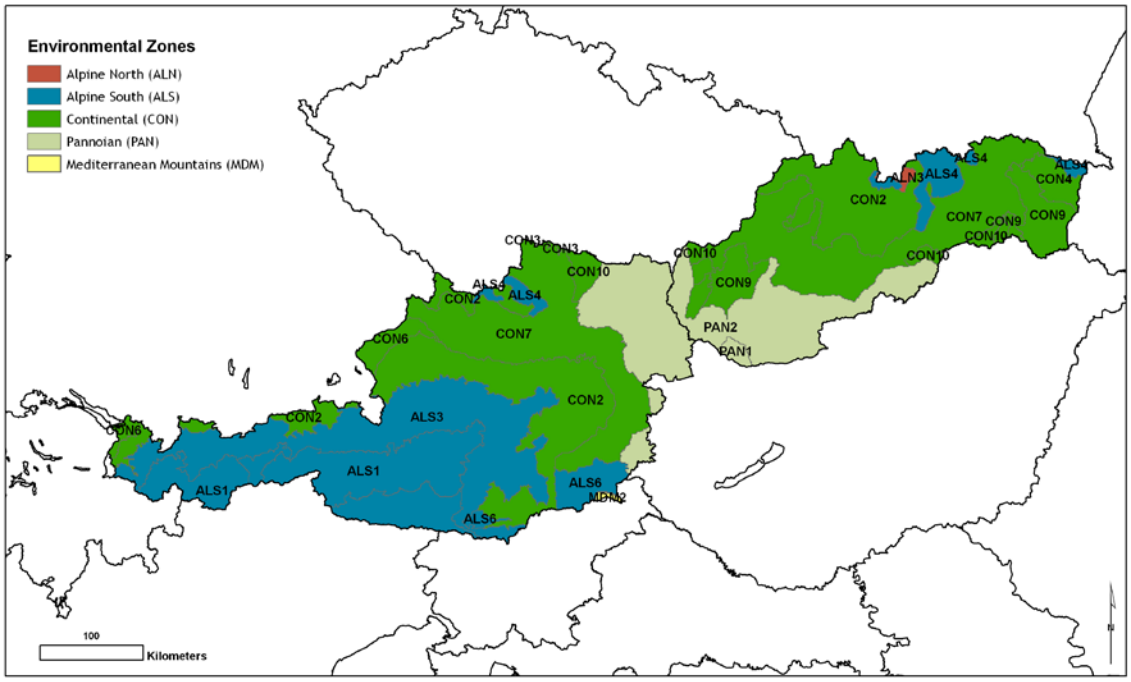


Figure 8

Environmental Stratification (EnZ, EnS) in the study area (after Metzger et al., 2005).

The Corine 2000 dataset (EEA, 2010) was used to mask out non-vegetated areas from the mean and standard deviation statistics. This operation eliminates all pixels representing a) water bodies, b) built areas and c) bare rocks.

To get an interpretation key regarding the distribution of phenology indices, altitude statistics from the GTOPO30 dataset and land-cover information from the GLC2000 map (Bartholomè and Belward, 2005) were extracted and summarised per EnS (Table 5.1, Table 5.2).

The following representative phenology indices were selected to illustrate their distribution throughout the study area (see Section 2.3 for more detail on the phenology parameters):

- Seasonal Permanent Fraction (SPF).
- Total Integral (TI).
- Peak of season value (MXV).

Seasonal Permanent Fraction

The Seasonal Permanent Fraction (SPF) is a phenological parameter defined as the area between the line connecting Start and End of season and the x axis. It refers approximately to the part of the vegetation signal which is permanent throughout the season.The average SPF and standard deviation values calculated for the two data series are presented in Table 5.3.The SPF averages from SPOT and MODIS reveal consistent differences in magnitude with the SPOT derived SPFs being lower than MODIS derived SPFs (linear correlation Pearson's $\rho = 0.89$). In other words, the resulting spatial pattern found across the study area of the average SPF or average level of permanent greenness is similar for both data series. As an example, the contiguous Environmental Strata PAN1,PAN2, CON9 and CON10 belong to the lower SPF classesfor both data series. The land cover distribution of these environmental strata show a dominance of the 'Cultivated and managed areas'which is likely to have strongly influenced the vegetation signal: the low SPF is possibly due to prolonged periods of no vegetation in the crop fields, occurring on and off throughout the year.

Table 2

Altitude statistics per Environmental Strata as derived from GTOPO30 data.

EnS	ALN3	ALS1	ALS3	ALS4	ALS5	ALS6	CON10	CON2	CON3	CON4	CON6	CON7	CON9	MDM2	**PAN1**	**PAN2**
Mean height (m)	1498	2038	1279	733	1463	376	317	817	487	386	467	470	177	231	111	192
Min height (m)	852	648	423	318	225	230	160	326	424	109	316	126	93	201	106	102
Max height (m)	2348	3713	2805	1409	3179	914	541	2031	562	965	1373	1312	537	296	124	584

Table 3

Percentage of land cover type (GLC2000) per Environmental Strata (in bold the dominant classes in the correspondent EnS). Masked classes are not reported.

EnS	Tree cover, broadleaved, deciduous, closed (%)	Tree cover, needle-leaved, evergreen (%)	Tree cover, mixed leaf type (%)	Shrub cover, closed-open, deciduous (%)	Herbaceous cover, closed-open (%)	Sparse herbaceous or sparse shrub cover (%)	Mosaic: Cropland / shrub and/or grass cover (%)	Cultivated and managed areas (%)
ALN3	3.44	75.40	0.79	0	0	19.05	0	0.79
ALS1	5.05	30.67	2.12	0.34	30.87	25.41	0	0
ALS3	14.35	54.55	12.31	0.44	11.00	6.47	0.01	0.40
ALS4	23.24	43.06	7.71	0	0.35	0	0	25.33
ALS5	15.52	54.64	9.08	0	12.98	6.69	0	0.41
ALS6	64.79	11.23	10.26	0	0.64	0	0	11.62
CON10	17.90	4.83	1.56	0	0	0	0	74.57
CON2	23.43	52.57	11.07	0	3.48	0.08	0.02	8.58
CON3	0	15.96	5.32	0	0	0	0	78.72
CON4	87.22	0.36	0.16	0	0.12	0	0	11.74
CON6	29.23	9.61	9.02	0	17.67	0	1.48	28.44
CON7	44.50	14.62	5.00	0	1.00	0	0	33.84
CON9	15.99	2.83	0.71	0	0.09	0	0	78.83
MDM2	22.97	12.92	8.61	0	0	0	0	55.50
PAN1	3.35	1.80	1.16	0	0	0	0	91.51
PAN2	11.23	3.30	0.41	0	0.23	0	0	81.30

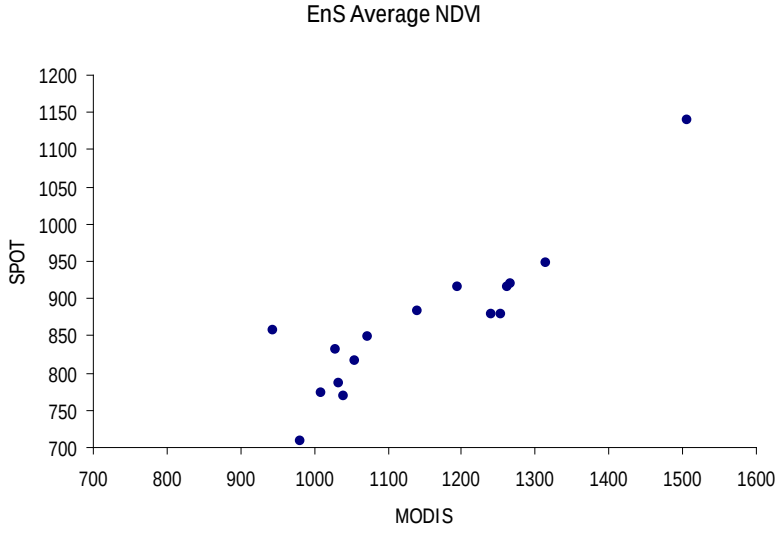
The SPF Standard deviations are much higher when MODIS data is used; this is possibly a consequence of MODIS pre-processing algorithms, which required the substitution of a higher number of no data than SPOT. The different number of years that characterize the SPOT and MODIS time series (eleven yrs and six yrs respectively) is likely to produce an effect on the mean values, thus comparisons between datasets are only indicative.

The spatial distribution of the mean SPF per Environmental Strata is presented in Figure 9 for SPOT and in Figure 10 for MODIS data.

Table 4

Mean SPF and standard deviation values per Environmental Strata (SPOT and MODIS data) and associated scatterplot (SPOT vs. MODIS NDVI EnS average).

EnS	SPOT		MODIS	
	SPF mean	SPF std	SPF mean	SPF std
ALN3	830.67	86.22	1028.90	261.42
ALS1	856.74	95.73	944.437	288.82
ALS3	915.42	116.87	1194.30	288.84
ALS4	877.94	102.09	1239.66	180.93
ALS5	882.12	129.24	1139.58	302.48
ALS6	915.17	104.02	1262.37	216.00
CON10	785.68	80.19	1032.96	216.45
CON2	947.19	130.61	1313.87	227.49
CON3	848.14	110.85	1072.64	300.91
CON4	878.47	38.77	1252.46	91.530
CON6	1140.66	119.01	1507.28	218.05
CON7	919.61	109.30	1267.42	186.46
CON9	769.78	84.26	1040.98	197.73
MDM2	817.08	57.03	1055.76	188.80
PAN1	709.57	53.42	981.828	152.95
PAN2	772.59	92.49	1010.06	207.19



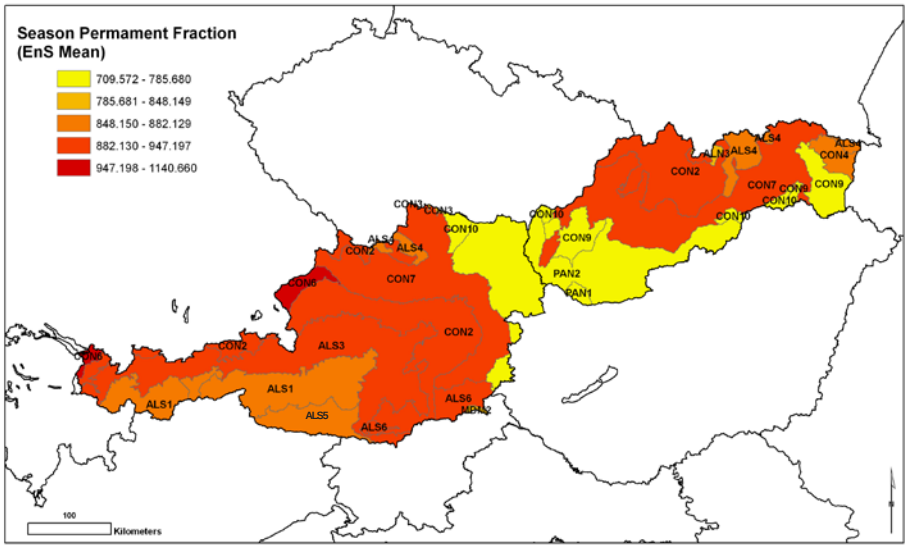


Figure 9

Distribution of mean SPF per Environmental Strata (SPOT NDVI data).

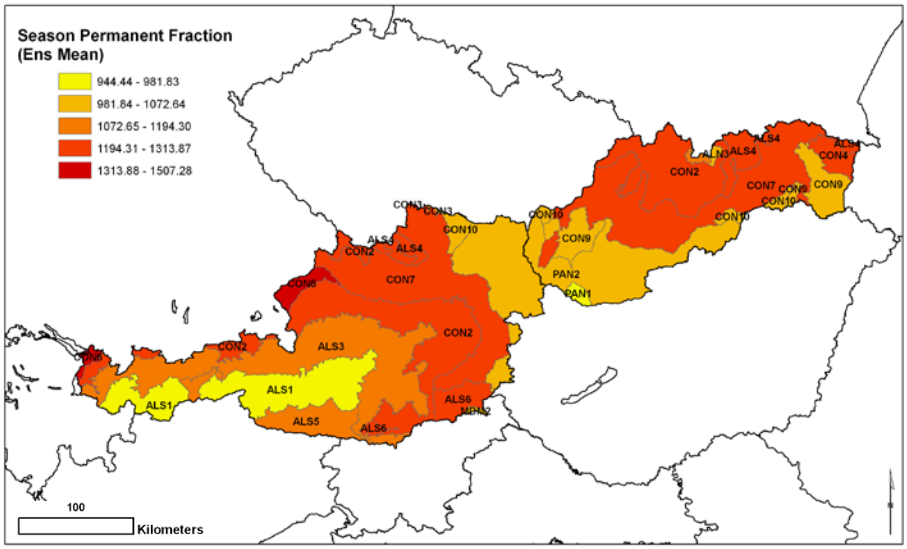


Figure 10

Distribution of mean SPF per Environmental Strata (MODIS NDVI data).

Total Integral

The Total Integral (TI) phenometric represents the area under the NDVI curve delimited by the two vegetation signal minima. TI is an important proxy representing an approximation of the Net Primary Productivity (NPP), see Reed et al. (1994). Mean TI and standard deviation values calculated for the two data series are presented in Table 5.

Again, the mean and standard deviation show differences in magnitude between the two data series with TI mean and standard devation values from SPOT being consistently lower than those from MODIS. The Continental environmental strata (CON), which in the study area are characterized mainly by natural broadleaf forests (GLC2000), showed in both data series the highest values of TI. This suggests that in this study area the CON stratum is one of the more productive. Lower productivity is related to the Alpine (ALS1) and the Pannonian strata (PAN1, PAN2), characterized by lower productive vegetation, such as, respectively, natural herbaceous vegetation and cultivated fields.

The spatial distribution of mean TI is presented in Figure 11 for SPOT data and in Figure 12for MODIS data. Data correlation between the two mean series is very high (ρ = 0.96).

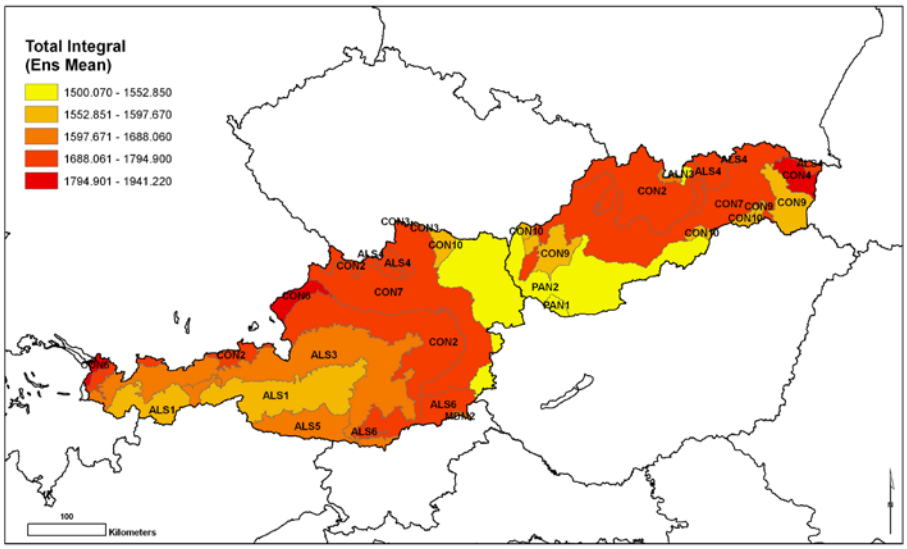


Figure 11
Distribution of Mean TI per Environmental Strata (SPOT NDVI data).

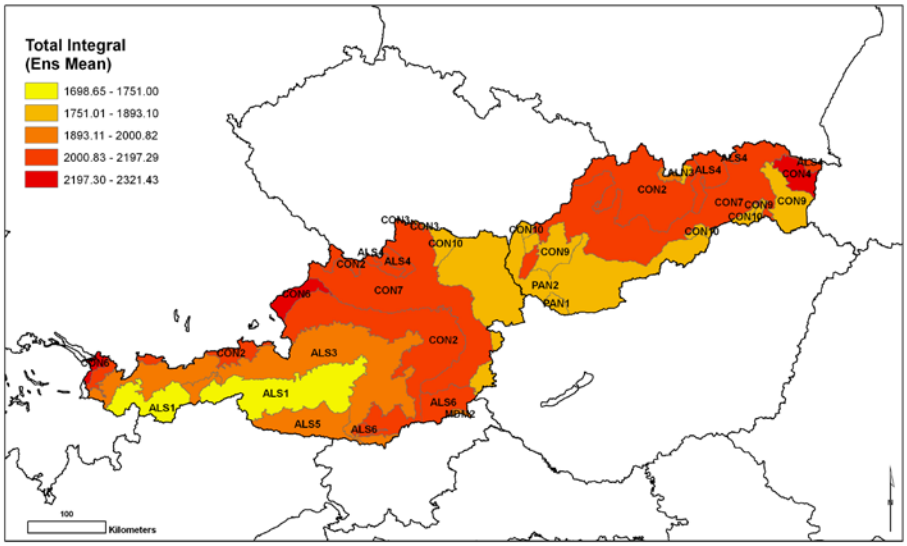
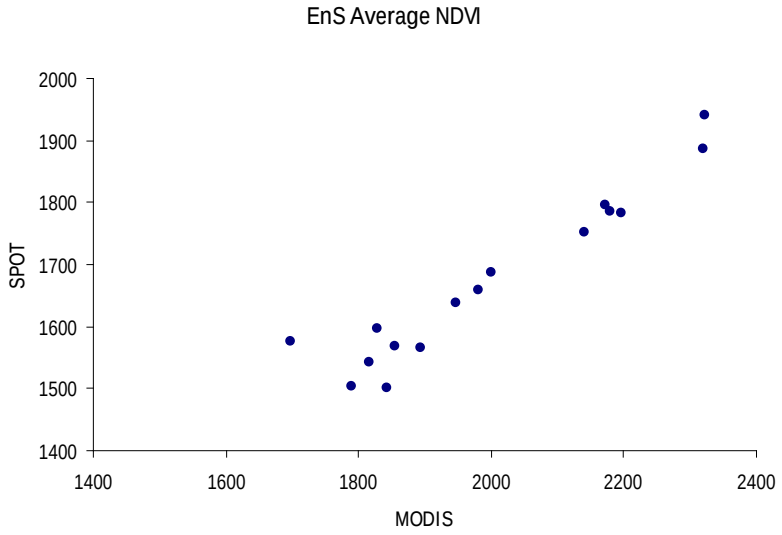


Figure 12
Distribution of Mean TI per Environmental Strata (MODIS NDVI data).

Table 5

Mean TI and standard deviation values per Environmental Strata (SPOT and MODIS NDVI data) and associated scatterplot (SPOT vs. MODIS NDVI EnS average).

EnS	SPOT		MODIS	
	TI mean	TI std	TI mean	TI std
ALN3	1502.85	127.72	1790.38	325.77
ALS1	1576.64	158.39	1698.65	383.90
ALS3	1688.06	174.91	2000.82	352.39
ALS4	1751.05	121.98	2141.75	219.22
ALS5	1638.95	183.11	1947.87	349.48
ALS6	1783.27	113.31	2197.29	224.51
CON10	1567.54	142.39	1854.80	340.98
CON2	1785.30	145.01	2179.95	258.83
CON3	1597.67	84.57	1827.71	416.88
CON4	1886.28	91.26	2321.17	184.20
CON6	1941.22	146.94	2321.43	247.66
CON7	1794.90	136.49	2172.62	250.95
CON9	1564.38	152.33	1893.10	313.95
MDM2	1657.81	67.80	1981.04	246.08
PAN1	1500.07	87.02	1843.05	217.48
PAN2	1541.79	148.27	1815.30	330.15



Peak of Season (MXV Value)

The peak of season corresponds to the point of maximum NDVI along the growing season. We demonstrate here results of the MXV phenometric (NDVI value at peak of season). Although MXV mean values differ significantly between the two data series (see Table 6), they present similar spatial distribution (Figures below).

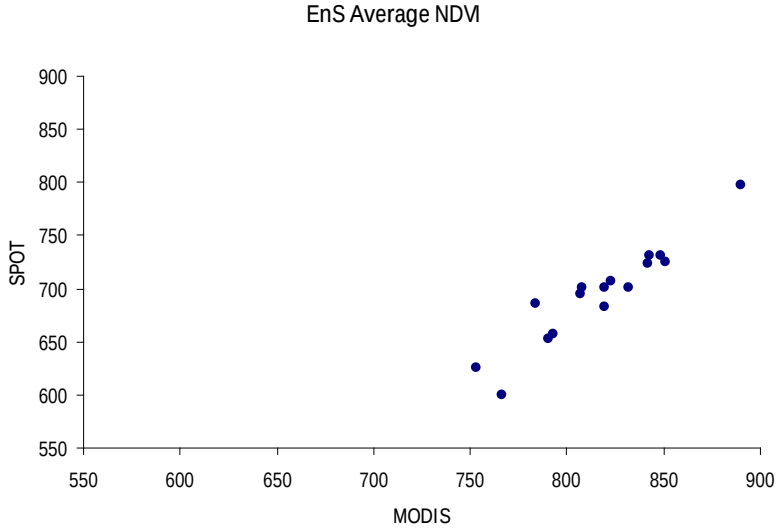
Low MXV values were detected for environmental strata ALN3 and ALS1 (Alpine areas), which are characterized by a higher proportion of bare soil/rocks, grass and sparse vegetation. These small values are possibly related to the amount and type of vegetation characterizing these strata (low density, higher proportion of herbaceous plants). Overall, higher MXV values belong to the Continental (CON) strata, dominated

by broadleaf forests and on a second instance by agricultural areas. Data correlation between mean values of the MODIS and SPOT series is found to be high, with $\rho = 0.94$.

Table 6

Mean MXV and standard deviation values per Environmental Strata (SPOT and MODIS NDVI data) and associated scatterplot (SPOT vs. MODIS NDVI EnS average).

EnS	SPOT		MODIS	
	MXV mean	MXV std	MXV mean	MXV std
ALN3	599.33	46.23	766.59	58.47
ALS1	625.12	52.33	752.79	69.09
ALS3	657.06	57.35	793.24	59.10
ALS4	723.75	48.12	841.97	39.55
ALS5	652.44	52.51	790.59	58.47
ALS6	730.33	21.20	848.39	22.68
CON10	701.48	47.15	808.07	55.74
CON2	700.68	43.79	832.26	37.24
CON3	686.39	30.03	783.56	27.35
CON4	798.10	37.15	889.98	35.07
CON6	706.46	27.68	822.67	33.70
CON7	731.50	50.94	843.08	48.16
CON9	700.93	45.77	819.85	48.01
MDM2	725.00	19.74	850.82	25.35
PAN1	683.17	28.24	819.53	42.38
PAN2	695.36	45.75	806.93	56.38



It should be noted that differences in wavelength centres and spectral windows of the bands relative to the SPOT HRV and MODIS sensors used in the calculation of NDVI values (VNIR and Red bands) produce differences in the raw NDVI values of the two sensors (NDVI MODIS always higher than SPOT). This means that

an inter-comparison can be done, for example, relatively to the spatial patterns, but not on comparing single absolute values.

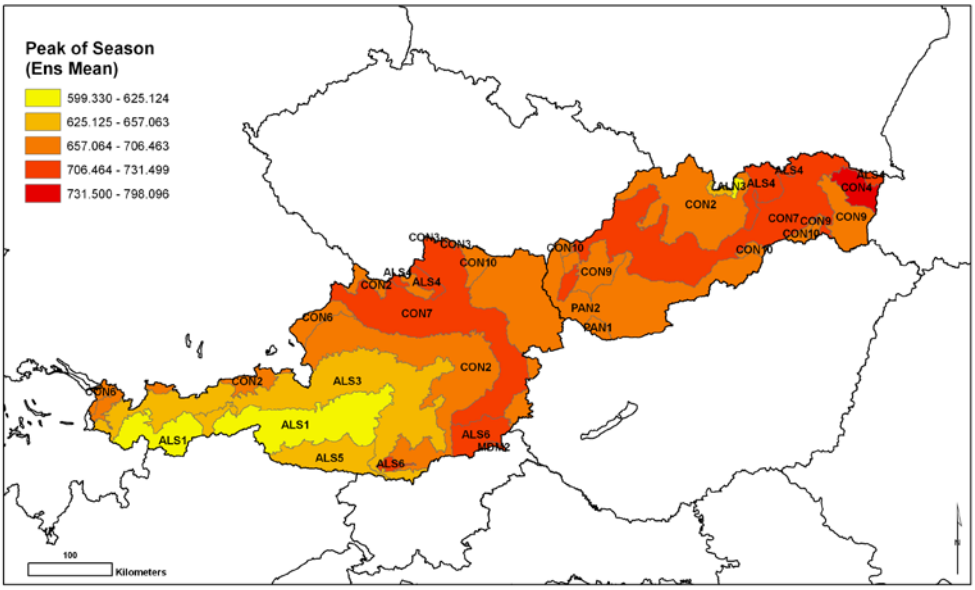


Figure 13

Distribution of mean MXV per Environmental Strata (SPOT NDVI data).

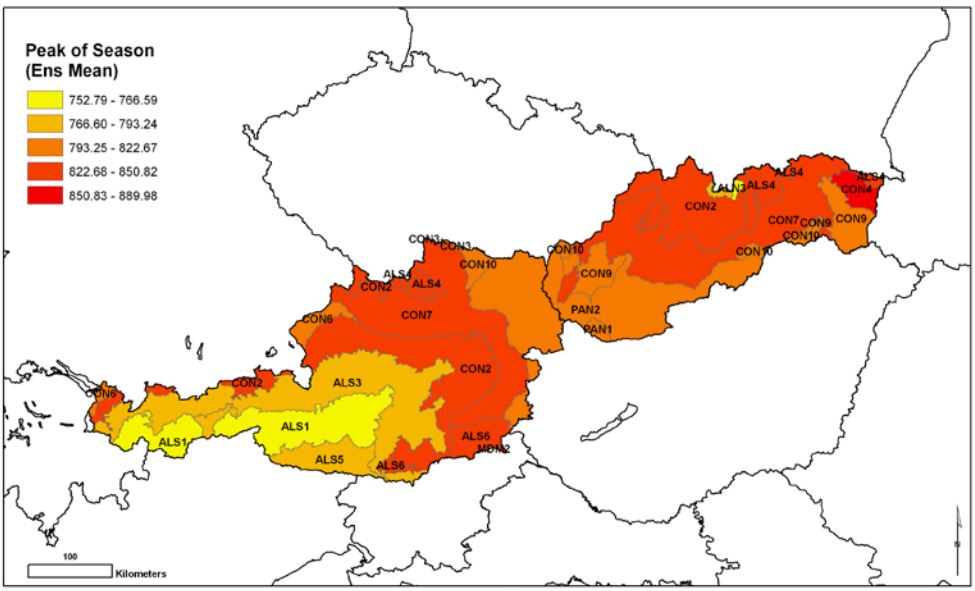


Figure 14

Distribution of mean MXV per Environmental Strata (MODIS NDVI data).

3.1.3 Forest Habitat classification using SPOT NDVI

Phenology information as extracted from remotely sensed VIs time series can provide valuable information to classify vegetation (see Steenkamp et al., 2008; Geerken et al., 2005; Geerken, 2009). The following task was undertaken to investigate the use of Phenolo phenometrics to classify forest habitats. In this first section of the work pilot classification tests were performed based on time series of SPOT NDVI data at a 1km resolution. Random Forests (Breiman, 2001) was selected as classification technique, having multiple advantages like being accurate, not sensitive to noise and computationally lighter than other classification methods. This approach has also produced promising results when applied to the classification of multisource remote sensing and geographical data (see Gislason et al., 2004).

The thematic level at which forest vegetation can be effectively mapped from medium-low resolution phenometrics is still an issue under investigation in the scientific community. Significant results were obtained mapping vegetation at the biome level using 1km AVHRR data (Wessels et al., 2011; Steenkamp et al., 2008). In the following methodology tests we focused on two lower ecological hierarchy levels for which data are available: *habitat* and *species*.

Mapping Natura2000 forest habitats
At the habitat level two Natura2000 habitat classes (see //www.eea.europa.eu) were selected: the Medio-European limestone beech forests of the *Cephalanthero-Fagion* (Natura2000 code: 9150) and the Acidophilous Picea forests of the mountain to alpine levels (*Vaccinio-Piceetea*), code 9410 each of which represent respectively one of the most common broadleaved and needle leaved forest types in the area. This choice would enable us to establish whether the phenology of deciduous broadleaved and evergreen needleleaved forest is distinct enough to allow for a reliable classification based on phenometrics. No direct correspondence is present between Natura2000 and GHC habitat classification schemes; the focus of this section is a general testing of the methodology and processing chain.

Breiman (2001) defines the Random Forests as:
A classifier consisting of a collection of tree structured classifiers $\{h(\mathbf{x}, \Theta_k), k=1,...\}$ *where the* $\{\Theta_k)\}$ *are independent identically distributed random vectors, and each tree casts a unit vote for the most popular class at input* $\mathbf{x}$.The collection of trees ('forest') classifier finally determines the most popular class by combining all the 'votes' from the trees. Split within tree is based on a CART algorithm (Steinberg and Colla, 1995; Breiman et al., 1984).

Each tree is grown as follows:
- If the number of cases in the training set is N, then sample N cases at random with replacement.
- If there are M input variables, then a number m<<M is specified such that, at each node, m variables are selected at random out of the M and the best split on these m is used to split the node (value of m kept constant during the forest growing).
- Each tree is grown to the largest extent possible. There is no pruning.

In our case, the M input variables were 31indicators extracted for the study area. The Random Forests technique needs as input a number of reference samples, which are then internally split into a set of *training* samples and a set of*test* samples. The former provides the 'truth' information regarding the classes of selected habitats (two in our case), while the latter is a set of points used to provide an estimate of error in the classification trees ('OOB Error'). In the Random Forests technique, there is no need for cross-validation or a separate test to get an unbiased estimate of the test set error, which is estimated internally during the run. For further details see Breiman (2001) and Gislasonet al. (2004).

The Natura 2000 GIS dataset is based on polygons with attributes related to proportion (%) of priority habitats. First, it was tested if the share within delineated habitat site geometries (polygons) was sufficient to take the location into consideration for further classification and choice of training samples. Ideally, the polygon training samples had to be homogeneously covered with the one forest habitat class of interest. To achieve an adequate level of 'purity' in the reference polygon samples, three selection criteria were applied: 1)the habitat has an area share >60% within the polygon; or 2) the habitat has a percentage of <60% but >30% and it is the habitat with highest proportional share in the polygon. Finally,3) all the extracted locations need to have a minimum share of 60% of CLC2000 (EEA, 2010) class Coniferous forest for Habitat 9410, and 60% class Deciduous Forest for Habitat 9150.This latter criterion was added as the Natura2000 site geometry polygons share their percentage criteria attribute on large areas instead than on a pixel basis, data structure on which the classification and remote sensing data are based.
After selecting the initial reference samples, a Random Forests classification was performed using routines developed by Liaw and Wiener (2002) in R language.

3.1.4 Results and discussion

The RF classifier was used to predict the distribution of the two selected priority habitat types in Slovakia (Austria was finally excluded due to scarcity of Natura2000 data). The parameters for RF were 31 variables (phenology indicators), number of trees=500, m_{try}=3 (initial number of randomly selected input variables, constant every node, determined using an in-built optimization step).

The extraction of training pixels led to an equally distributed population between Habitat 9150 (n=214) and Habitat 9410 (n=256). The inbuilt accuracy assessment, which is calculated on approximately ⅓ of independent data (test samples), resulted in an error estimate of 0.65% in the choice of the two habitat classes (9410 and 9150); in other words, pure pixels were classified correctly in 99.35% of the cases.
A Confusion Matrix is presented in Table 7. It should be noted that the training and test samples (1km pixels) are located within the Natura2000 polygons, which covers a small percentage of the study area. This could influence the accuracy results.

A significant feature of the RF technique is to provide an estimate of the importance of the phenology variables used in the classification, in other words RF calculates which phenometrics contributed more to the classification result and which one were less influent. This is estimated by looking at how much the prediction error increases (OOB) when data for that variable is permuted while all others are left unchanged. Figure 15 shows the phenometrics used in the classification sorted for importance - Mean Decrease Accuracy - (see Table 1 for variable ID and meaning). In this test the more influent phenometrics in the classification accuracy resulted: Peak of Season-value (MXV), Normalized Season and Growing Integral (SNI, GNI) and value of Growing Season End (GEV).

The RF classification results showing the predicted probability of a pixel being classified as habitat 9150 (*Acidophilous Picea* forest) or habitat 9410 (Limestone beech forest) was mapped in a GIS environment (Figure 16). The spatial distribution of the inferred habitats was compared qualitatively with the results obtained from the Habitat maps produced by Mucher et al. (2009). These authors created a series of predicted habitats maps based on the integration of ecological knowledge with available European datasets of land cover and site conditions.

Table 7

Confusion matrix of Natura2000 Beech and Picea pure pixels classification.

	Beech	Picea	OOB Error
Beech	213	1	0.0046
Picea	2	248	0.0080

OOB estimate of error rate: 0.65%

The maps shown in Figure 16 indicate generally a good agreement between the Random Forest classification and the Mücher et al. (2009) predicted distributions, showing similar general patterns for both maps at regional scale. However, Habitat 9150 is slightly underestimated in the RF classification with respect to the results of Mucher et al (2009), and a South western patch of *Pinus sylvestris* was misclassified as *Picea* forest. A more rigorous comparison between the two maps would need a better harmonization between compared data (e.g. regarding threshold probabilities, see legend of Figure 16). The map generated using the RF classification uses only remote sensing time series and no ancillary data such as climatic, geological/soil data or topographic information as in Mücher et al. (2009) model.

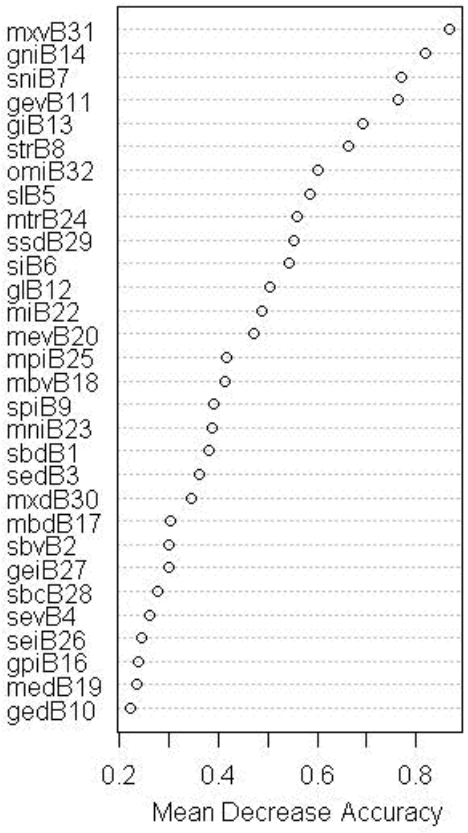


Figure 15

Phenometrics sorted by the importance in contributing to the classification accuracy (Acronyms in Table 1).

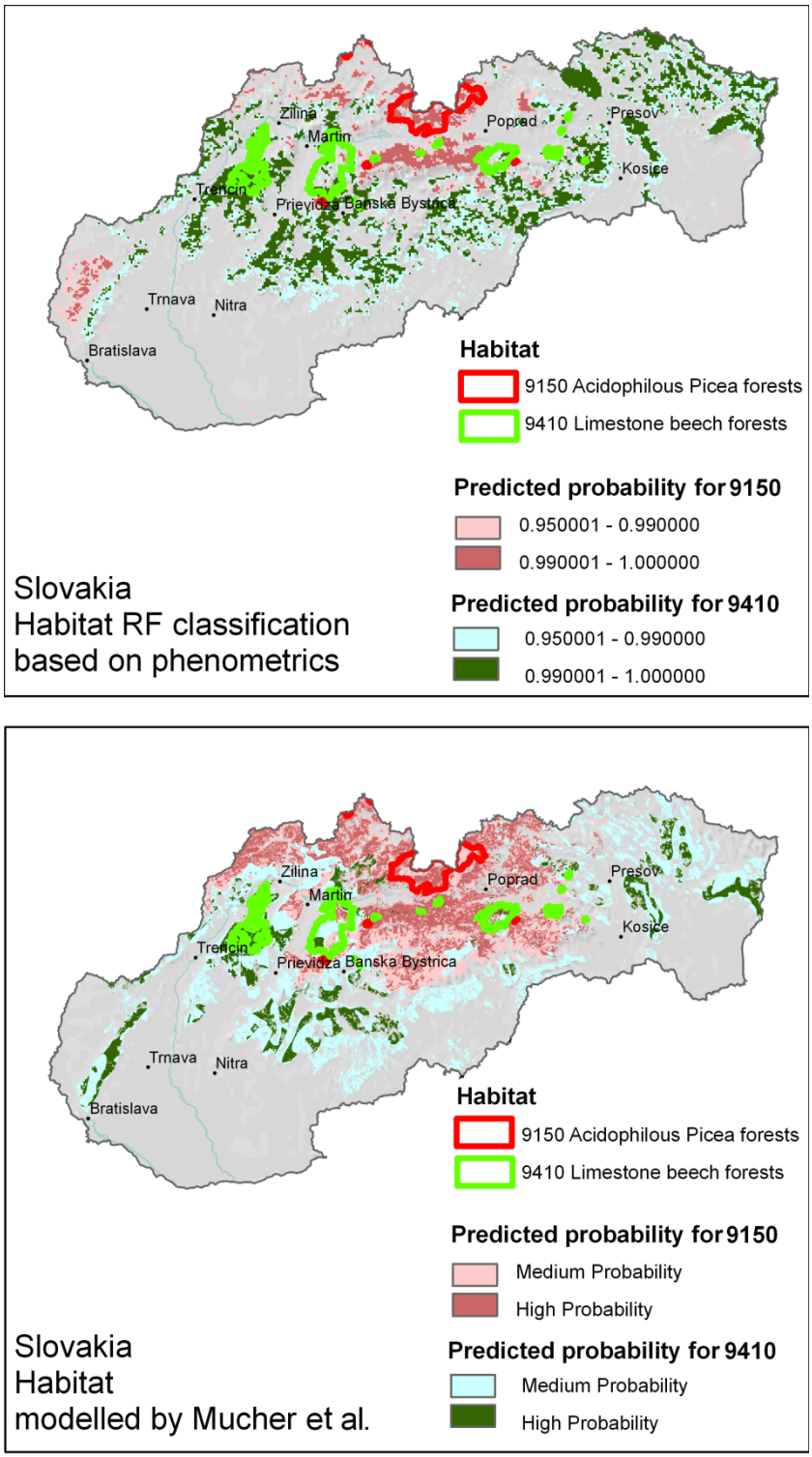


Figure 16

Random Forests habitat classification based on phenometrics (upper figure) and the habitat prediction model by Mücher et al. (2009), lower figure. Habitat polygons source for training: Natura2000.

3.1.5 Tree species classification

A Random Forests classification based on SPOT NDVI phenometrics was carried out using forest tree species distribution data. The exercise was performed to test the performance of the classification technique using phenology information to deliver a thematic classification level which is more detailed than the habitat level, i.e. species. The data used belongs to the AFOLU Tree Species distribution maps, developed by EC JRC (Koeble and Seufert, 2001). The Dataset is composed by 137 tree species distribution maps at a 1km resolution. The data was derived by modelling information from four Pan-European datasets: 1) Land cover from CORINE; 2) Land cover from PELCOM (Pan European Land Cover Mapping) for areas where no CORINE data was available; 3) Species information from ICP Forest Level I; and 4) statistical data from the FAO-TBFRA2000 (Temperate and Boreal Forest Resources Assessment 2000). For this analysis, the study area included both Slovakia and Austria.

The assessment focused on three common species in the study area: *Fagus sylvatica, Picea abies* and *Pinus sylvestris.* The tree species reference data were derived from the AFOLU database by applying the following criteria for the extraction of 'pure' pixels:

– the pixel is classified as one of the selected tree species;
– one of the three species is present in more than 60% of the pixel area (attribute in the AFOLU database).

A subset of pixels was then randomly chosen to be used as classifiers. Their number varies depending on the abundance of the tree species in the study area (Table 8). For this classification the number of phenology indicators used in the RF was reduced to 24. This number was chosen after performing a first run of the classification with 31phenometrics and eliminating from the second run the eight phenometrics with the least impact on the classification accuracy results. Doing this led to a lower OOB error. The number of trees was set equal to 500 and m_{try}=4 (initial number of randomly selected input variables at each node).

Results

The accuracy assessment of the classification resulted in an overall OOB error of 17.6 % on the determination of the three tree species. A confusion matrix is produced and reported below (Table 8).The higher classification error for *Pinus sylvestris* is likely due to a high occurrence of 'low purity' pixels, as observed in Figure 18, left. The RF predicted probability of tree species presence was mapped in a GIS environment, and the spatial distribution compared qualitatively with the AFOLU data (Figure 19).

Table 8

Confusion matrix of the tree species RF classification, and number of training points used (n).

	Pinus sylvestris	*Picea abies*	*Fagus sylvatica*	*OOB Error*	*n (training points)*
Pinussylvestris	233	174	93	0.534	500
Piceaabies	56	1860	84	0.070	2000
Fagussylvatica	27	182	791	0.209	1000

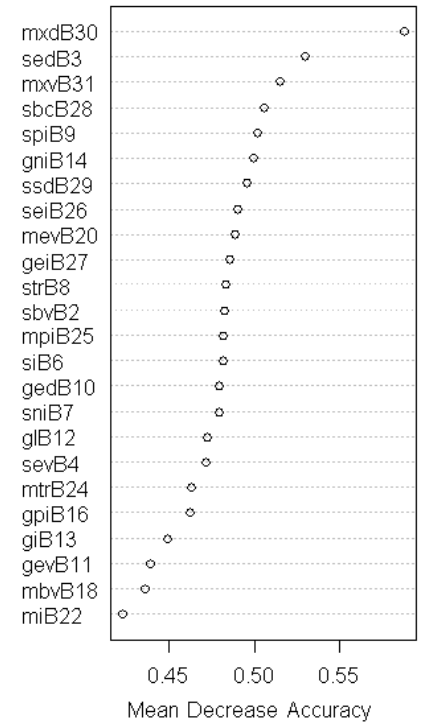


Figure 17

Phenometrics sorted by the importance in contributing to the RF classification (Acronyms in Table 1).

A comparison between the two classification data was produced using scatter plots (Figure 18). In order to reduce noise, the original maps of 1km grid size have been aggregated to 3 x 3km grids and linear correlation coefficients calculated. Pearson's correlation coefficient (R) is calculated between the predicted probabilities of belonging to a class versus the observed area share of the pixel (AFOLU) for that class. Provided this assumption is a simplification (assuming the area share is equivalent to a presence probability in the pixel, see Hill et al., 2007), the coefficient reveals some accordance between the datasets, especially for *Pinus* and *Fagus.* A good spatial agreement is also qualitatively observed, especially for the Alpine strata -AL- and the CON2 stratum (Continental 2).

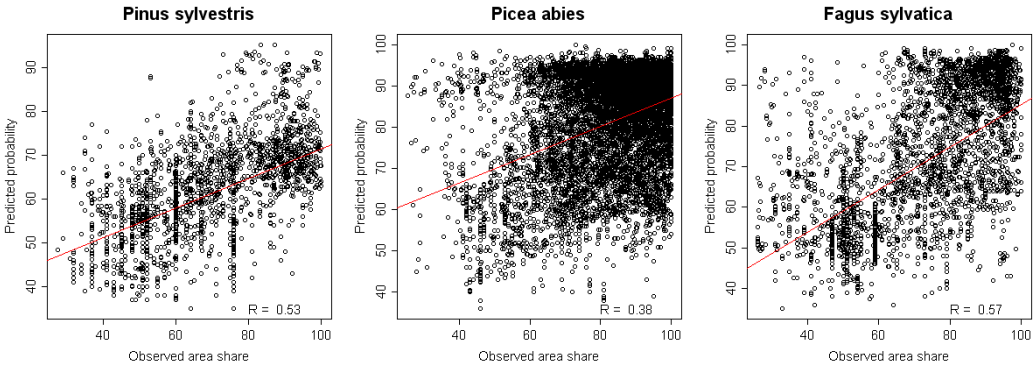


Figure 18

Scatter-plots and R coefficients for mean aggregated (3 by 3km) values of AFOLU data (area share) and predicted tree species probability of RF classification. Only pixels of dominant species probabilities and those greater than 25% of area share are shown.

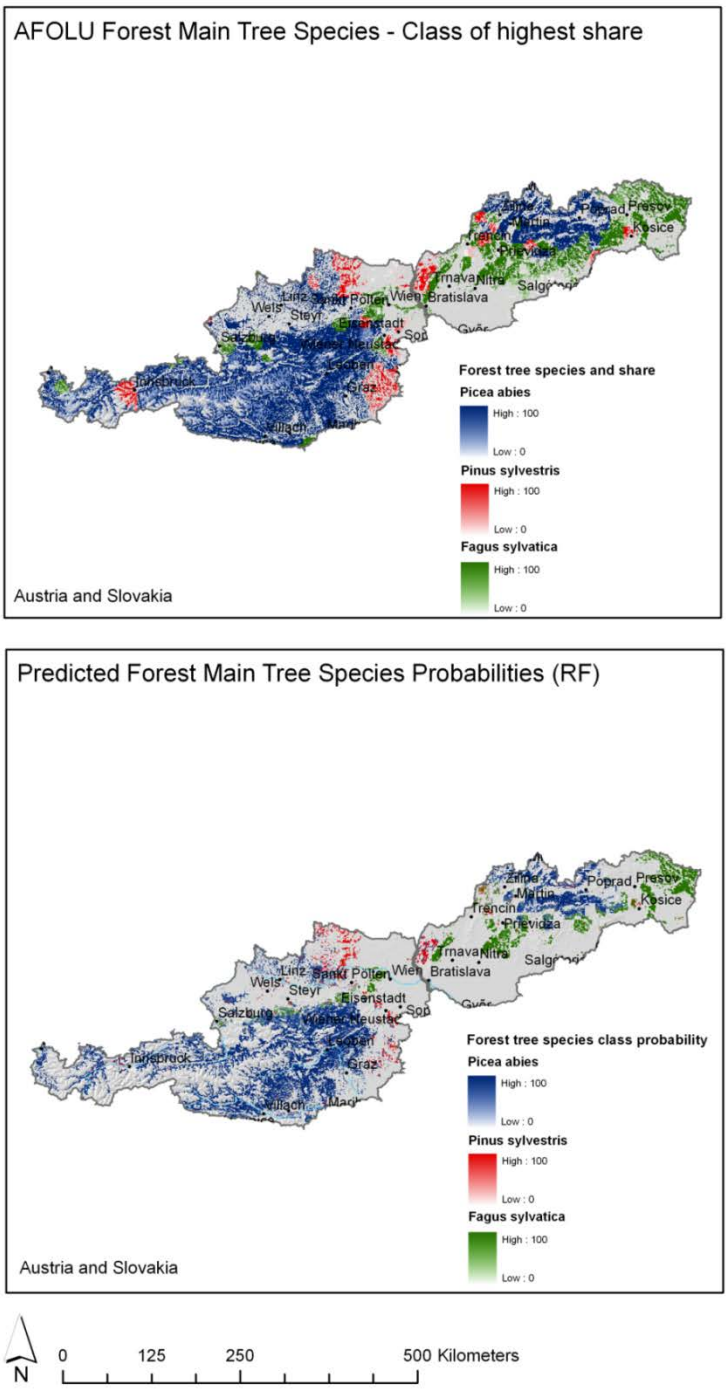


Figure 19

AFOLU map of the selected tree species-highest share class- (Koeble and Seufert, 2001) and RF tree species classification based on phenometrics (lower figure).

3.2 Classification using MODIS NDVI Data

3.2.1 Introduction

After pilot testing the general methodology, the analysis focused on MODIS NDVI data as this data set was deemed more adequate because of its spatial resolution (250m): the higher spatial resolution was expected to reduce the number of mixed ('not pure') pixels and thus lead to better classification results. This time Austria was selected as the initial study area, which in a second step was extended to a selected Environmental Zone (Metzger et al., 2005).This time the thematic level tested was the habitat level and the adopted classification system determining the EO derived forest habitat classes was the *General Habitat Category classification system* (Bunce et al., 1998; 1996), which is the classification scheme adopted by EBONE.

Forests in the GHC system are categorised as the Forest Phanerophytes (FPH) class, which is defined as woody vegetation with a minimum height of 5m. The following (leaf) forms allow for a further subdivision: Forest coniferous (FPH/CON), Forest deciduous (FPH/DEC) and Forest evergreen (FPH/EVR) classes. Detailed information on the rule based system adopted to establish which habitat and phyto-sociological vegetation association is grouped in FPH is found in Bunce et al. (2010).

3.2.2 Field data

As discussed, the Random Forests technique needs in input a set of training pixels to run the classifier ('pure pixels'). This information is extracted from 1km x 1km plots sampled and classified in the field by EBONE partners, or from existing field samples, translated into the GHC scheme (see Table 9).Available field plot data are located in Austria, Italy, Southeast France and Sweden. Field data from Sweden were in some cases discarded from the analysis as the NDVI time series in this regions affected by large periods of missing data (prolonged cloud coverage, snow, etc).

The plot data were provided as ESRI shapefiles, classified under the GHC scheme. Polygons with FPH/CON, FPH/DEC and FPH/EVR attributes were selected and overlapped with the 250m raster grid structure of MODIS NDVI data. A pixel was considered as 'pure' if the proportion of CON, DEC or EVR was greater than or equal to 70% within the MODIS pixel (see Figure 20).A very limited total of 81 pure pixels were identified(51 CON and 29 DEC). No pure pixels for the GHC EVR were found.

Table 9
GHC field plots description.

Country	#Plots	Organization/Project
Austria	48	University of Wien/SINUS
France	11	EBONE/PYRODIV
Sweden	25	NILS
Italy	15	BioHab
Sum	99	

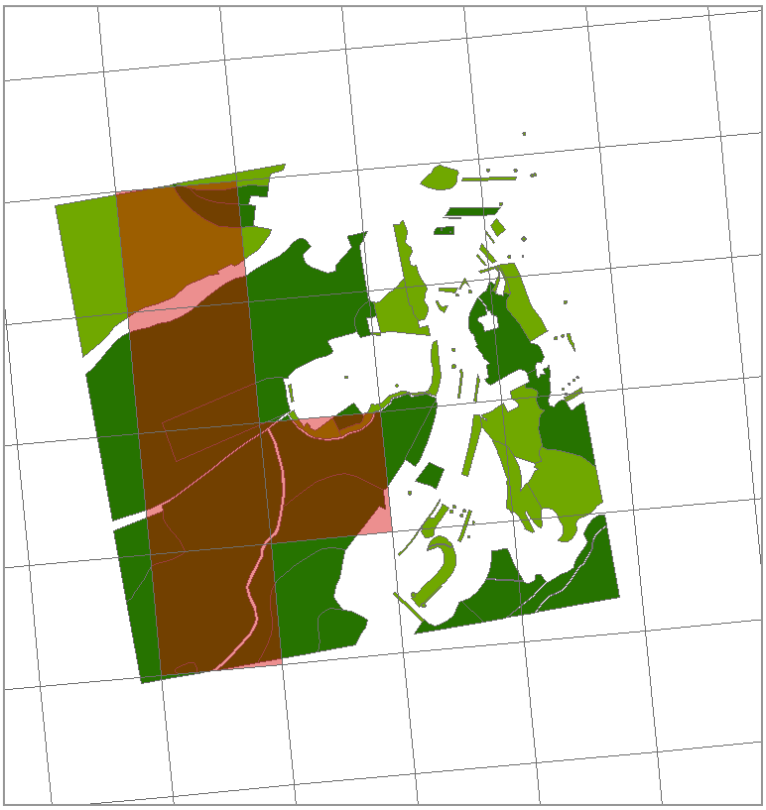


Figure 20

FPH polygons (dark and light green) were overlapped with the MODIS NDVI grid to extract 'pure' MODIS pixels (in red).

3.2.3 Classification of coniferous and deciduous forests (GHC)

Austria

The first part of the MODIS data analysis was to investigate the performance of the RF classification using NDVI phenometrics to detect forests under the GHC scheme. Two forest types were considered: coniferous (CON) and deciduous (DEC).The test area was initially limited to Austria. To understand the relationships between the number of phenometrics and the ability of the RF classifier to discriminate Coniferous and Deciduous forest, 29 recursive classification tests were launched using the 48 'pure pixels' reference plots. At every cycle (1+n)phenometrics were excluded, where *n* varies from 0 to 28 and the Mean Decrease Accuracy (MDA) calculated. MDA is a measure provided by the RF algorithm that quantifies the decrease in classification accuracy when eliminating one of the phenometrics from use in the classification. The MDA is used to determine the relative contribution of each phenometric to the classification. The type and number of phenometrics were selected by ranking the phenometrics according to MDA (Figure 21). For each of these classification tests 100 runs were performed.

In this test the OOB error decreases by performing classifications with the phenometrics listed at the top of the MDA graph (Figure 21), while adding phenometrics listed at the bottom of the MDA graph led to decreased ability of the classifier to discriminate CON and DEC pixels in the training set (monotonic relationship statistically significant based on Spearman's rank correlation coefficient: Figure 22). The four most relevant phenometrics in this classification test are all *date* phenometrics (MXD, MEV, MBV, SBV).

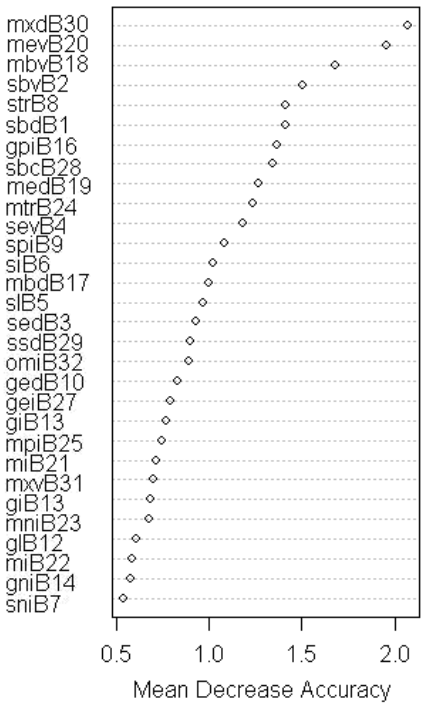


Figure 21

Mean Decrease Accuracy using all phenometrics for CON and DEC classification (Austria).

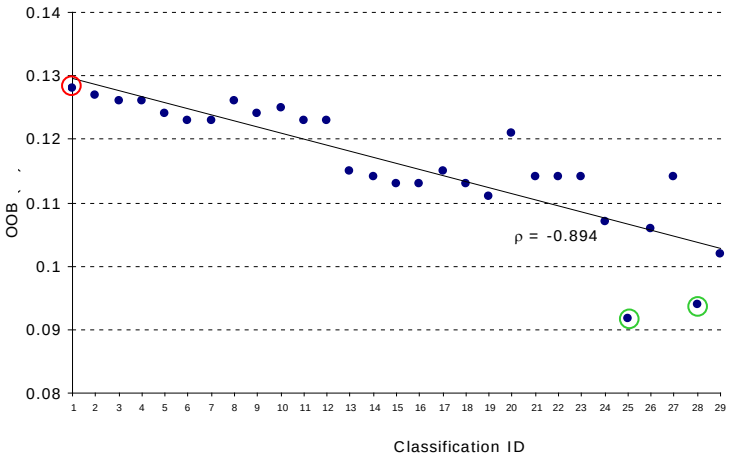


Figure 22

Ability of the RF in discriminating pure CON and DEC pixels among MDA-ordered phenometrics configurations (red circle: higher OOB; green, lower).

The lowest OOB errors were obtained using the phenometrics configuration in Ph28 (OOB=0.094) and Ph25 (OOB=0.092), while the highest error was obtained using all phenometrics except one: Ph1 (OOB=0.128). These three phenometrics configurations (Ph1, Ph25, Ph28) were used to run RF classifications in Austria to compare the classification performance for GHC Coniferous and Deciduous forests. The classifications were

carried out on a subpopulation of the 250m MODIS NDVI pixels defined by extracting within the 25m JRC Forest Mask 2006 (Kempeneers et al., 2010) those pixels that have at least 70% share of CON or DEC (the latter correspondent to 'broadleaf' in the Forest Mask 2006dataset). This operation (i.e. removing pixels not classified as forest in any of the two datasets) was performed to get a fully comparable dataset when evaluating classification accuracy. *Figure 23* shows a map of the RF classification using the Ph28 configuration and the JRC Forest mask 2006 data up-scaled at 250m. The outcome of the RF classifications is turned into pixel based maps, containing for each forest pixel the probability of it being CON or DEC.

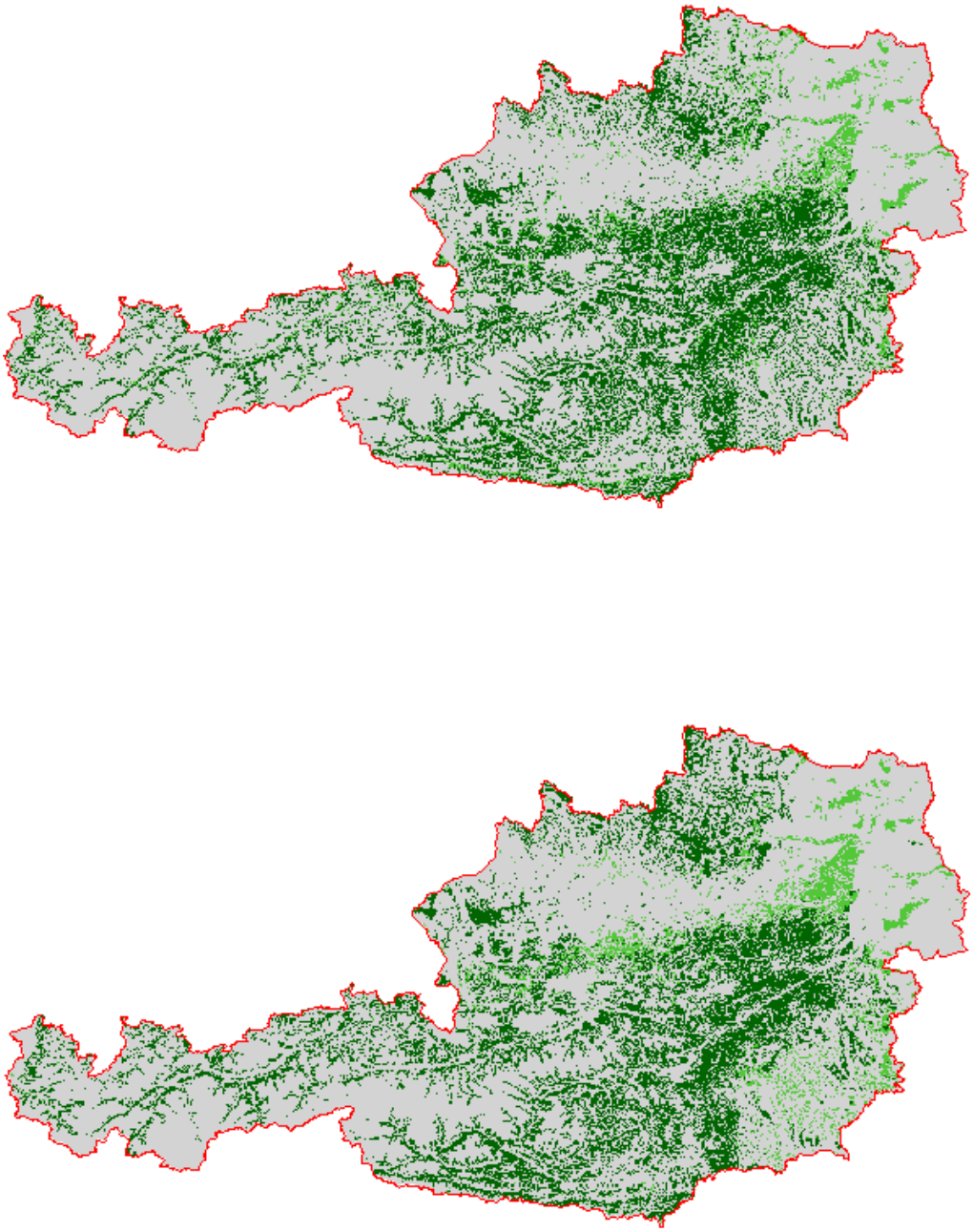


Figure 23

Forest Classification based on Random Forests using the Ph28 configuration (upper image), and the Forest layer derived from the JRC Forest Mask 2006 (lower image). Both images are at 250 m.

Accuracy Assessment

An assessment of the overall accuracy of the three RF classifications (Ph1, Ph25, Ph28) was performed using the JRC Forest Cover Map 2006 (Kempeneers et al., 2010) aggregated to 250m. The JRC Forest Cover Map 2006 provides coverage for Europe at a 25m resolution. It was derived using IRS-P6 LISS-III, SPOT4 (HRVIR) and SPOT5 HRG imagery for the years 2005-2007. The overall accuracy of the JRC Forest Cover Map 2006 was reported to be higher than 85% (Kempeneers et al., 2010). The GHC categories considered by the RF classifications are comparable with the ones defined by the JRC Forest Cover Map 2000 (see Table 10;Pekkarinen et al., 2009). Two forest classes are present in the JRC Forest Cover dataset: *broad-leaved* and *coniferous*; the broad-leaved class contains Deciduous and Evergreen types.

The JRC Forest Cover Map 2006 dataset was adopted for multiple reasons: 1) it has a pan European coverage, thus allowing inter-comparisons across a wide range of study areas in Europe, 2) it covers a period included in the MODIS NDVI time series, and 3) it is the only recent European dataset holding Broadleaved/Coniferous forest type information.

The validation dataset was derived from the JRC Forest Map 2006 as follows:

- The JRC Forest data was summarized to match the spatial resolution of the MODIS NDVI grid by calculating the proportion of 25 m forest class pixels present within each 250 m pixel.
- The validation set was created by selecting the 250 m pixels with a proportion of either Coniferous or Broadleaved forest ≥ 70% and classifying those as CON and DEC respectively.

Table 10

Categories included in the forest class of the Forest Cover Map 2000 (from the Forest Action website, 2010).

Vegetation types
Broad-leaved:
– broad-leaved forest with more than 30% crown cover
– plantations of e.g. eucalyptus, poplars
– evergreen broad-leaved woodlands composed of sclerophyllous trees (mainly *Quercus ilex, QuercusSuber, Quercus Rotondifolia*)
– arborescent mattoral with sclerophyllous species
– olive-carob forests dominated by *Olea europaea* sylvestris, *Ceratonia siliqua*
– palm groves woodlands, *tamarix* woodlands, holly woodlands
– broad-leaved wooded dunes
– sub-arctic broad-leaved forests not reaching the 5m height
– transitional woodland areas when the canopy closure of trees cover more than 50% of the area and if their average breast height diameter is at least 10cm
Coniferous:
– coniferous forest with more than 30% crown cover
– non-evergreen coniferous trees woodland composed of *Larix* species
– arborescent mattoral with dominating *Juniperus oxycedrus/phoenica*
– Christmas trees plantations
– coniferous wooded dunes
– sub-arctic coniferous forest, not reaching the 5m height
Mixed:
– mixed forest, the share of coniferous or broad-leaved does not exceed 25% in the canopy closure
– mixed wooded dunes

The accuracy assessment was performed by carrying out a pixel based comparison between the validation data set and the Random Forests CON and DEC class assignments of Ph28, Ph25, Ph1 which were determined by assigning to the pixels the class with the highest *p* value (e.g. if p_{CON} = 0.51 the pixel is classed as CON). Figure 24 summarizes the accuracy assessment scheme.

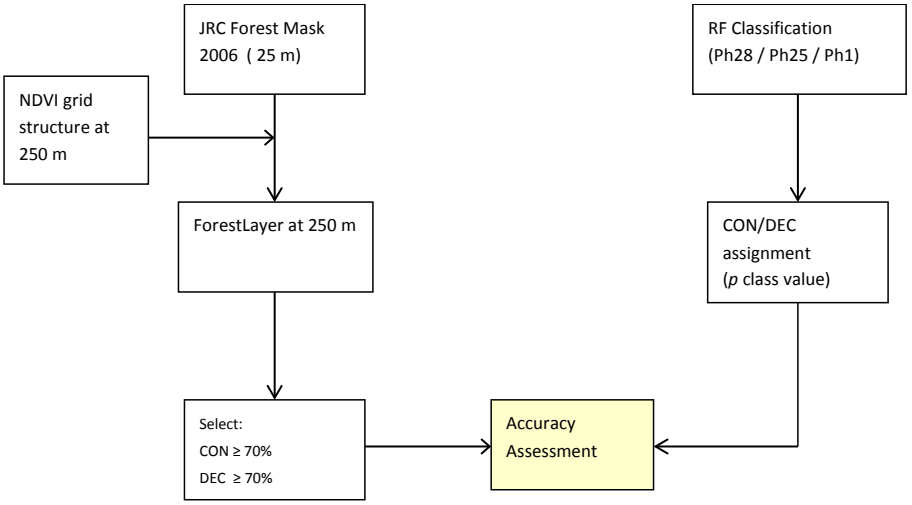


Figure 24

Flow chart of the Accuracy Assessment scheme used.

Class and overall mapping accuracy for the three RF classification configurations are reported in Table 11.

Table 11

Class and overall classification accuracy for Ph28/ Ph25 / Ph1 configurations.

Class Accuracy (%)	Ph28	Ph25	Ph1
Coniferous - CON	82.18	82.39	76.70
Deciduous - DEC	37.31	36.55	29.65
Overall Accuracy (%)	83.89	84.01	78.78

Low accuracy values in the deciduous class can be due to a series of reasons, among them the vegetation type heterogeneity of the class which will cause the phenometrics values of the class to vary substantially.

The area with the highest discordance between the RF classifications and the reference dataset is located in Southeast Austria (Graz region). Here forest types are characterized by *mixed* formations, as observed using CLC2000 data and regional maps (vegetation map from Austrian *Institut für Wald inventur, http://bfw.ac.at/rz/bfwcms.web?dok=4636*). To investigate if the GHCs forest classification accuracy can be improved we introduced the presence of a mixed class in the RF scheme. The mixed class is defined with the following rule: a pixel should have a proportion of FPH/CON < 70% or FPH/DEC <70% but their sum is greater or equal to 70% of forest-FPH (the adopted definition of forest pixel is when it is included at least 70% share of forest type).This defines a new population of 'pure pixels' for the mixed forest class. The Ph25 phenometrics

configuration was selected for comparison, having achieved the best overall accuracy in the previous assessment.
A Random Forests classification with the three forest classes was run for Austria. The accuracy assessment is performed using CLC2006 data (from EEA, www.eea.europa.eu) at 250m as reference data set, including the land cover classes broad-leaved forest (class 311), coniferous forest (class 312) and mixed forest (class 313). Accuracy results are shown in the following table.

Table 12

Class and overall accuracy for Ph25 configuration after introduction of the mixed class.

Class Accuracy (%)	Ph25
Coniferous - CON	79.03
Deciduous - DEC	44.54
Mixed - MIX	21.31
Overall Accuracy (%)	75.77

Discussion

The RF classification using pure forest classes in Austria revealed satisfactory results for the classification of coniferous vegetation (FPH/CON), while low performance is found in classifying deciduous forest (FPH/DEC). Low accuracy values of the latter are likely to be caused by a higher variability in the time series characteristics of this class, which is reflected in the phenology indicators variance. Figure 25A shows for example, the presence of higher fluctuations of NDVI signal in winter decades for the pure deciduous points. In the contrary, coniferous forests show generally more homogeneous inter-annual runs. The satisfactory value of the overall mapping accuracy is due to the unbalanced proportions of coniferous and deciduous pixels, the former being much more abundant in Austria, and increasing the overall accuracy.

The introduction of a mixed class decreased the overall classification accuracy and the accuracy of the FPH/CON class. The FPH/DEC showed a slight increase in accuracy. The construction of this class based on theoretical assumptions of CON-DEC mixtures could have introduced a form of class 'noise' negatively affecting the classification accuracy. Also, unexpected trends in the time series of the pure training pixels for the MIX forest class were present: NDVI values in the summer decades are occasionally higher than the CON and DEC classes (Figure 25A). This may suggest that the training data set may not be fully suitable. A control profile using a random set of points in Austria from CLC2000 forest classes (Coniferous, broad-leaved and mixed) showed a different and more reasonable trend (Figure 25B).

Two remarks can be drawn:

– - RF classification accuracies showed inverse proportionality with Out Of Bag (OOB) errors, calculated on training (pure) pixels. The OOB error is an indication which helps to discriminate worst and best phenometric configurations. In the presence of correlated variables and well distinguished pure pixels, the OOB error can help to address a more efficient classification. This is not to be taken as a general rule, but it should be checked with respect to 1) site specifications and 2) it can vary depending on phenometrics correlations.
– - Mean Decrease Accuracy (MDA) can help understand which phenology indicator would better contribute to discriminate the two forest classes. On the other hand, MDA is calculated with respect to the training pixel group. The variance present in no pure data can potentially provide different MDA behaviour. Inference of these conclusions to the general case needs further investigation at pan-European level.

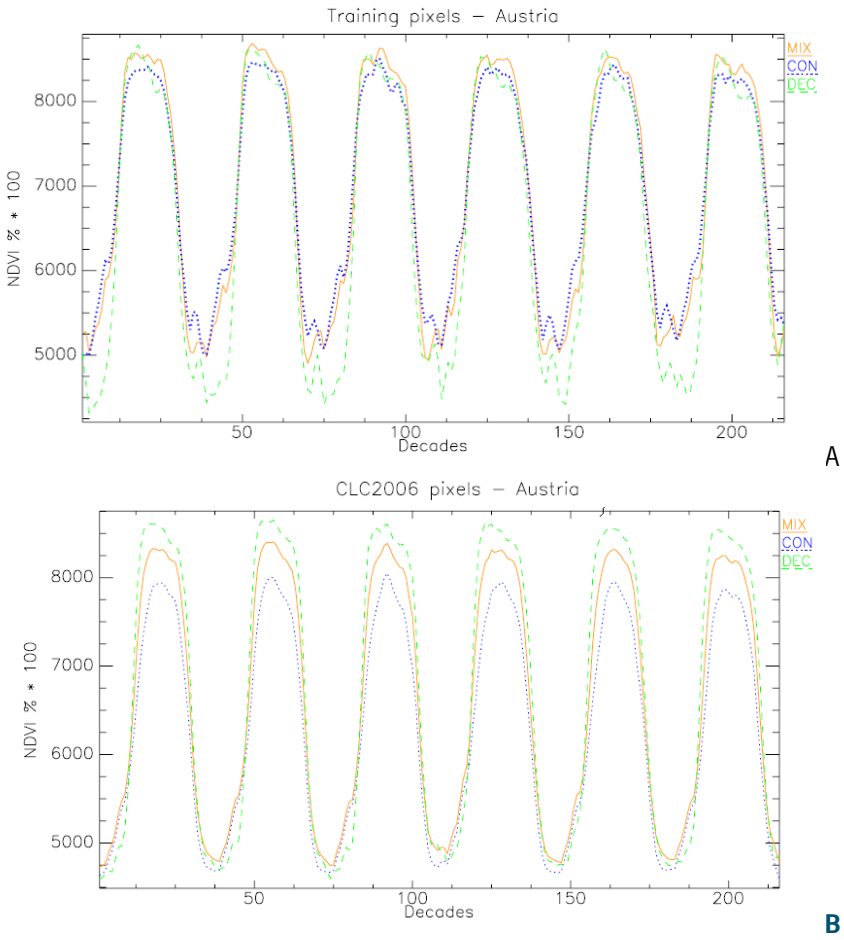


Figure 25

NDVI interannual run for pure pixels of CON, DEC and MIX classes (A: GHC training pixels. B: pure pixels from CLC 2000).

Mediterranean Environmental Zone

The same classification steps described in the previous section where performed for a region dominated in this case by broadleaf forest (Mediterranean Zone, MDN in *Metzger* classification). The three phenometrics configurations chosen are Ph28 (excluding 28 indicators), Ph14 and Ph0. These were selected to represent 'extreme' situations in terms of phenometrics number, plus a middle configuration. In this way we wanted to test a configuration criterion independent to the analysis of OOB errors. A total of 30 training pixels were selected within the MDN zone. Only FPH/CON and FPH/DEC pure pixels were available, while no pure FPH/EVR were found in the available data (DEC and EVR were considered as a single class). Three RF classifications of FPH Coniferous and Deciduous forests were performed (no mixed classes), with trees number=500 and m_{try}=4. Also in this case, the subpopulation of pixels on which the classification was applied was defined by selecting those 250m pixels that have at least a 70% share of coniferous or 70% of broadleaved forest (the within pixel proportion of coniferous and broadleaved forest was calculated using the 25m JRC Forest Mask 2006, as described in Section 3.2.2). As discussed, this operation is added to eliminate mixed pixels from the validation dataset, focusing on 250m pixels representing a dominant vegetation type. Figure 26 shows the classified image using all phenometrics (Ph0).

An accuracy assessment following the same processing chain described in Figure 24 was performed. Class and overall mapping accuracy for the three classification configurations are reported in Table 13.

Table 13

Class and overall classification accuracy for three phenometrics configurations in the MDN Environmental Zone.

Class Accuracy (%)	Ph28	Ph14	Ph0
Coniferous - CON	34.84	38.65	47.65
Deciduous - DEC	52.96	54.30	58.01
Overall Accuracy (%)	62.41	64.52	69.62

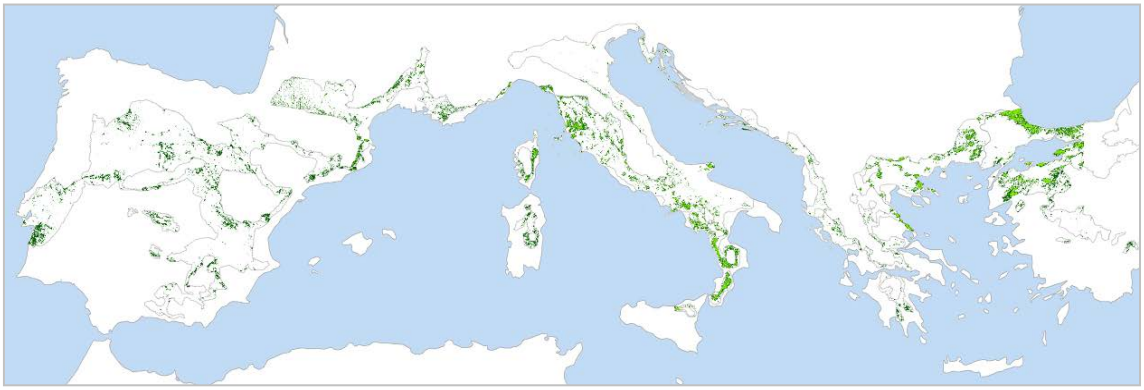


Figure 26

RF classification map of forest types using the Ph0 configuration (MDN Environmental Zone). Dark green Coniferous (FPH/CON) and light green Deciduous (FPH/DEC).

Discussion

All RF classifications performed in the MDN Environmental Zone showed low classification accuracy, which are much inferior to the levels achieved for Austria (Table 11). These results are potentially produced by a series of factors, among them the pixels employed to train the RF classifier, and the validation dataset used. In the classification accuracy assessment we used as reference the JRC Forest map 2006 (Kempeneers et al., 2010). This dataset is derived using a spectral classification approach, while the classes identified in the GHC field plots are built based on field observation. The former made use of medium resolution (30/250 m) imagery information, classified using training pixels extracted from the CLC2000 map; the latter is based on *in situ* assessment of forest type, their density and life forms.

Two important considerations should be made about the two different approaches:

1) 'Pure pixels' from GHCs field plots were defined based on thresholds of class share within the MODIS pixel (see Figure 20).Visual observation of the FPH/CON and FPH/DEC training pixels using the *GoogleEarth®* interface showed that a GHC forest class often shows differences in spatial characteristics of the vegetation cover. This is due to local vegetation characteristics (e.g. forest density, species), together with the influence of the non-tree background. Such differences are likely to have an impact in the spectral signature of the training pixels. As illustrated in Figure 28A, GHC FPH/DEC polygons overlapped on high resolution images show heterogeneity in terms of tree cover and background extension (see for example the zoomed inset). Figure 27 shows large NDVI interannual differences for the winter period associated to two pure FPH/DEC points in the same field plot (1 and 2 in Figure 28A). Coniferous stands in Austria, at the contrary, showed a high degree of homogeneity in the very high resolution images. Additionally, another factor potentially lowering classification accuracy is the absence of any pure pixel related to the FPH/EVR class (Evergreen). This is due to the small number of currently available GHC field plot data. FPH/EVR is relatively abundant in Mediterranean habitats, thus their influence in classification is possibly not negligible.

2) Using reference data derived uniquely from spectral information to calculate classification accuracy (JRC Forest Map, 2006) can lead to misleading accuracy results and a higher degree of 'mismatches' between the compared datasets. This is especially true when the background component is strong, and forest density low (such as in the MDN zone).Figure 28B shows the overlap between the JRC Forest Map 2006 (darker green, broad-leaved) within the previous GHC FPH/DEC polygons. Mismatches between the two are evident (bary centres of MODIS pure pixels shown as red points). This highlights one important issue with continental or global scale EO derived products: there is a widespread lack of available high quality and independent data suitable for validation.

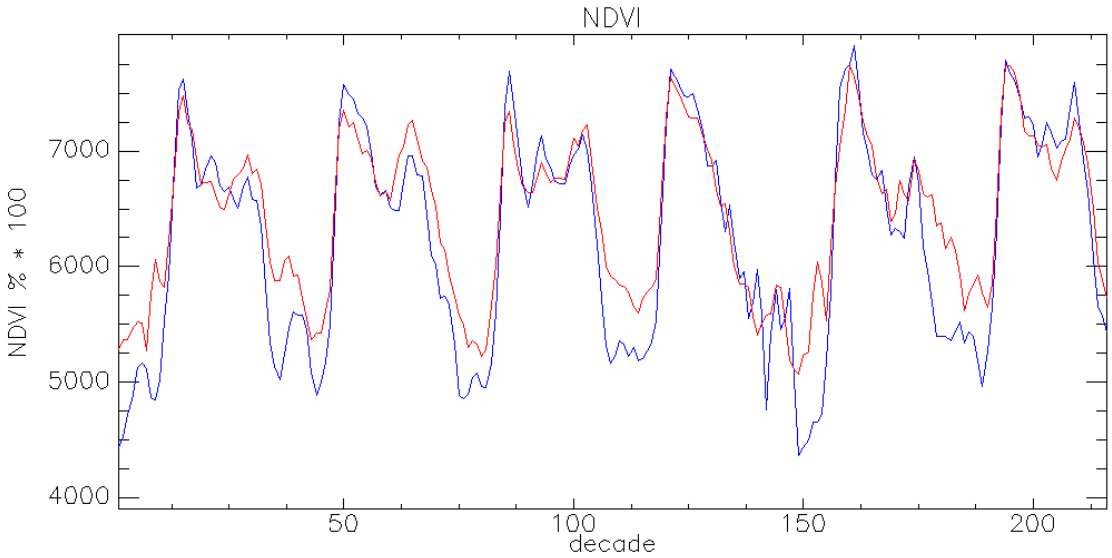


Figure 27

NDVI interannual run for two FPH/DEC pure points in the same field plot (see next Figure in the upper part for locations of Point 1-red- and Point 2-blue-).

It is clear from the two classification tests in coniferous and deciduous dominated sites (Austria and MDN zone), how the classification accuracy depends on site specific conditions. Especially when the tree density is low, revealing a large amount of background (e.g. under storey of shrubs, herbs, herbaceous, forbs or soil), the inter-annual NDVI run in the same forest class can appear significantly different, hence influencing phenology indicators values. Selecting the best performing phenometrics did not lead to the expected accuracy improvement, appearing to be a secondary factor of importance with respect to the site specific forest characteristics.

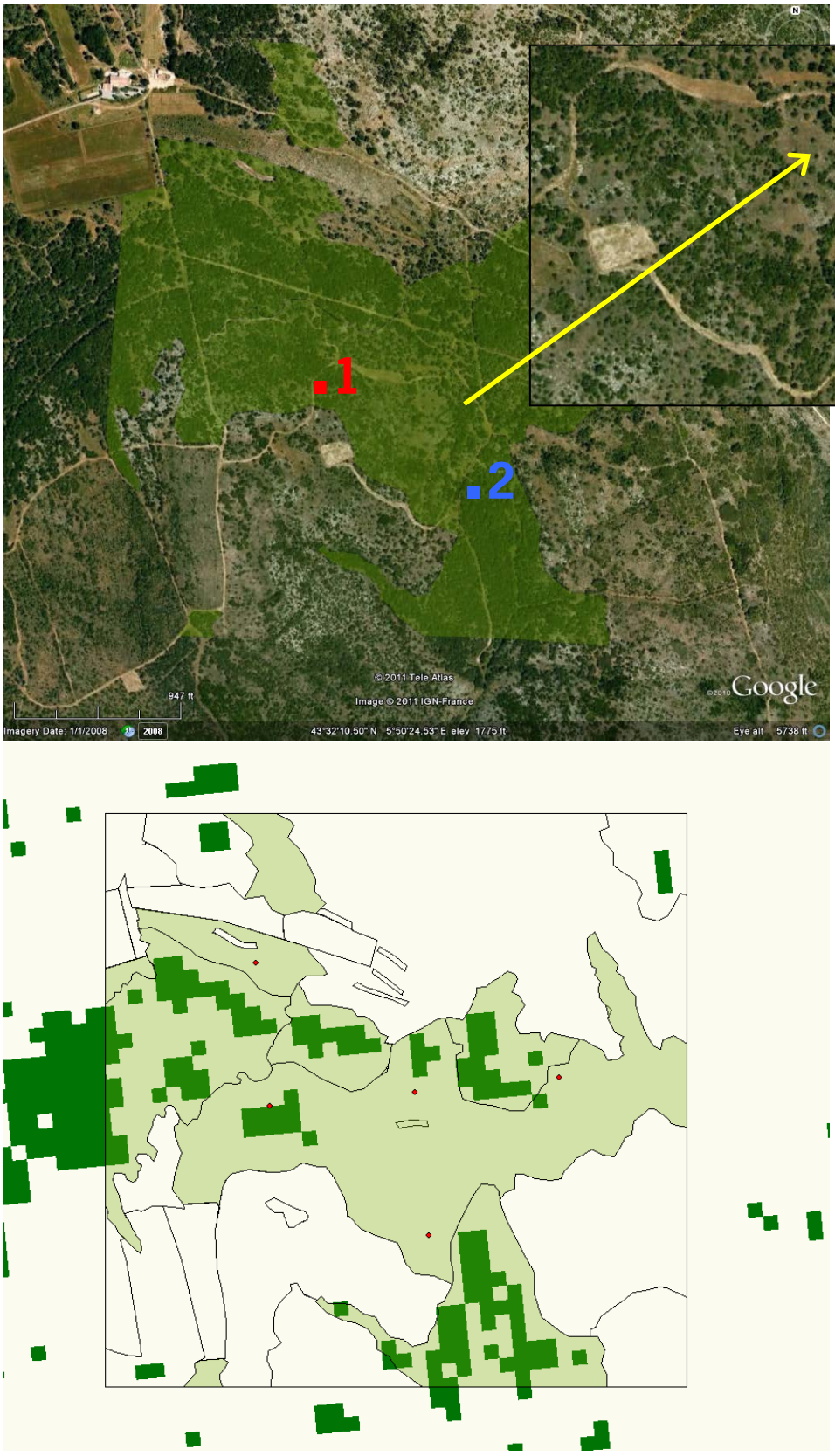


Figure 28

FPH/DEC polygons on (A)Google Earth® images and (B) on JRC Forest Mask 2006.

3.2.4 Influence of NDVI data gaps in Classification

Time series of vegetation indices are often characterized by the presence of data gaps. These are due to atmospheric events, like cloud coverage, snow and smoke which do not allow adequate detection of the vegetation spectral signal. Both number and sequential arrangement of these gaps can have an influence on the performance of a classification by altering the data quality. Here we assessed the potential impact these data gaps could have on the performance of phenology based GHC RF classifications.

A set of MODIS representing pure pixels of FPH/CON and FPH/DEC showing no data gaps (no interpolated values in the series of NDVI MVC decades) were selected. The 'purity' criterion is the same as applied in Section3.2.2. All pixels were chosen with the condition of being located within the Mediterranean North zone - MDN- (Metzger et al., 2005), but also to have a correspondent pixel in the JRC Map 2006 validation dataset at 250m. The spatial constraint was adopted to minimize the influence of bio-geographical variations in forest composition, focusing on areas with homogeneous bioclimatic conditions. The number of training points was limited to a total of 10.

The following processing steps were followed:
- Introduction of ten consecutives 'data gaps' (ten contiguous data gaps decades) per year within the full six years MODIS NDVI time series.
- Extraction of the FPH/CON and FPH/DEC training (pure) pixels from the modified NDVI time series (series with data gaps added).
- RF Classifications using all the phenometrics for the NDVI time series with data gaps added.
- Accuracy assessment and comparison with classification accuracy using the original NDVI data.

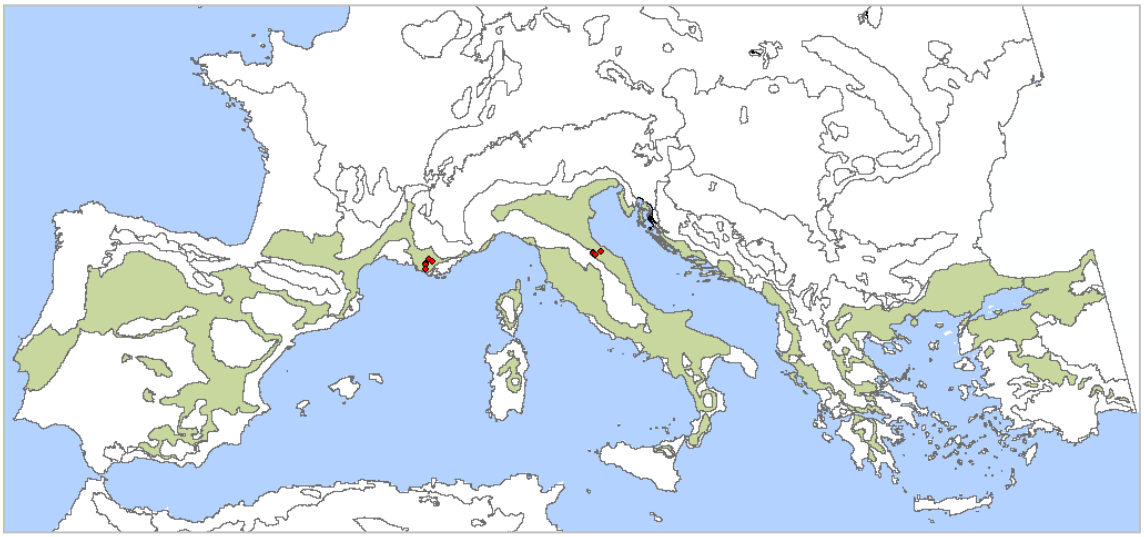


Figure 29
Location of pure pixels (red) selected within the MDN Environmental Zone (light green).

Figure 30 shows the count of data gaps, i.e. number of decades available to build a seasonal mean, present within the MODIS NDVI time series after pre-processing operations (Section*2.2*). Data gaps occur mainly in the Northern areas and in mountainous regions, and range between 0 and 25 in the European territory. To

investigate their influence on GHCs classification accuracy, a series of data gaps were artificially introduced in the NDVI time series of FPH/CON and FPH/DEC pure pixels previously used to train the RF classifier.

The pre-processing chain applied to the NDVI data is partly able to cope with data gaps, especially if they are of short lengths. When seasonal means are available, the data gap is replaced by this value. When seasonal means cannot be calculated, data is linearly interpolated between the existing gap-contiguous data points. For longer data gaps this may imply serious problems, especially if significant vegetation dynamics are expected within the missing time gap. The impact is generally attenuated due the fact that gaps are occurring for the majority during winter time, when vegetation activity is low. But still, the majority of phenological indicators (e.g. minimum, start and end of season, etc.) rely on parameters extracted in this period.

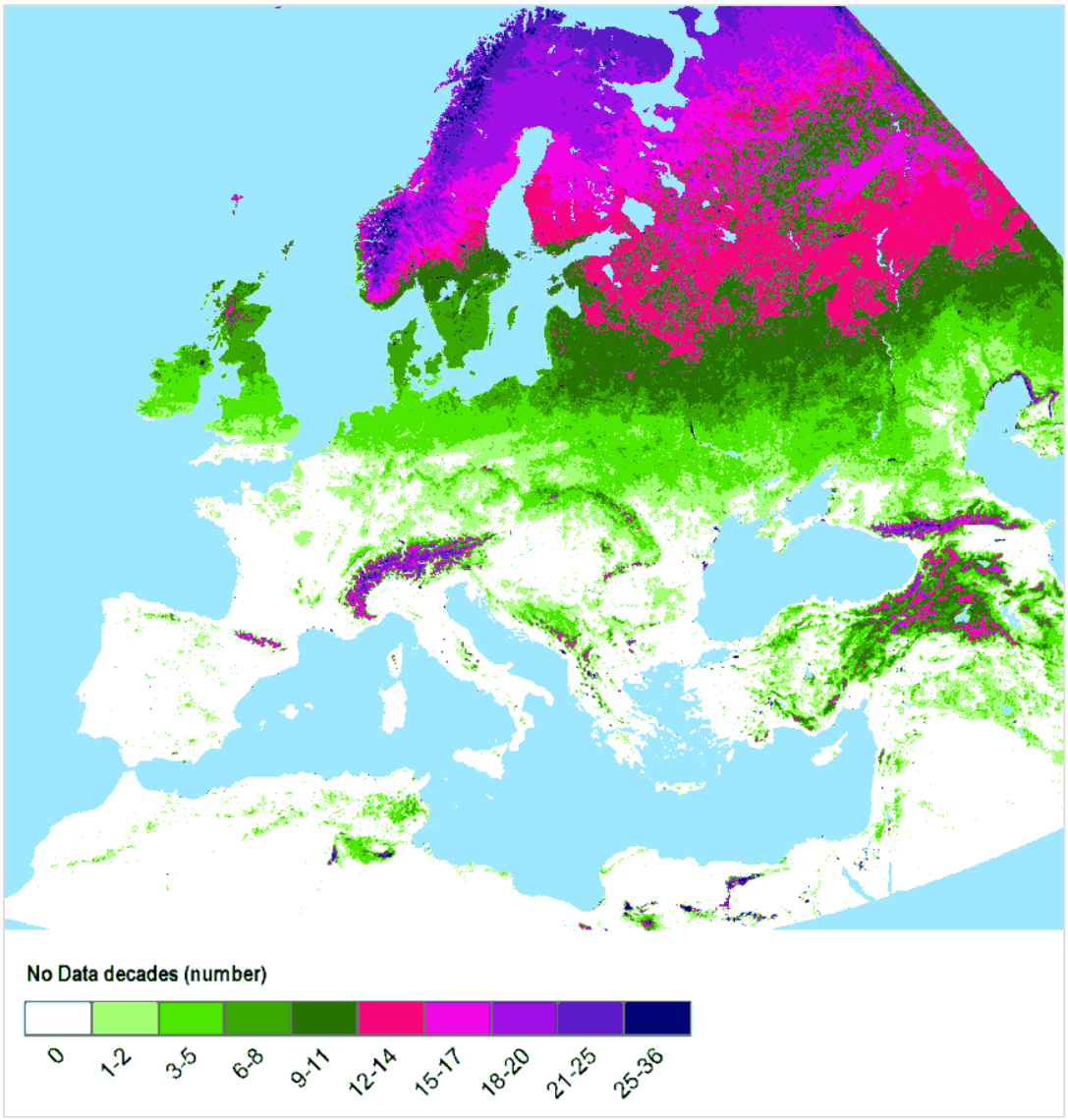


Figure 30

Count of no data decades in the MODIS NDVI time series after pre-processing operations.

Five pure pixel points from GHC field plots in South Sweden are shown to explore the effect of pre-processing on NDVI time series in regions largely affected by data gaps. The common number of missing decades for all five points is eight decades for the winter 2004/2005 and ten, ten, nine and eleven for the following winters.

Smaller gaps outside the winter period do not exceed three decades. The gaps are generally located symmetrically around the turn of the year (Figure 31A) . Missing data and outliers are substituted by seasonal means (Figure 31B). The outlier criterion was chosen according to Chebychev's theorem (95% confidence interval, see Lohninger, 1999). The '0' value visible around the turn of the year was used to identify missing data when using integer type. Finally (Figure 31C), missing data which could not be substituted with the seasonal mean are given an interpolated value using the nearest existing points in time and subsequently filtered using Savitzky-Golay algorithms (Chen et al., 2004).

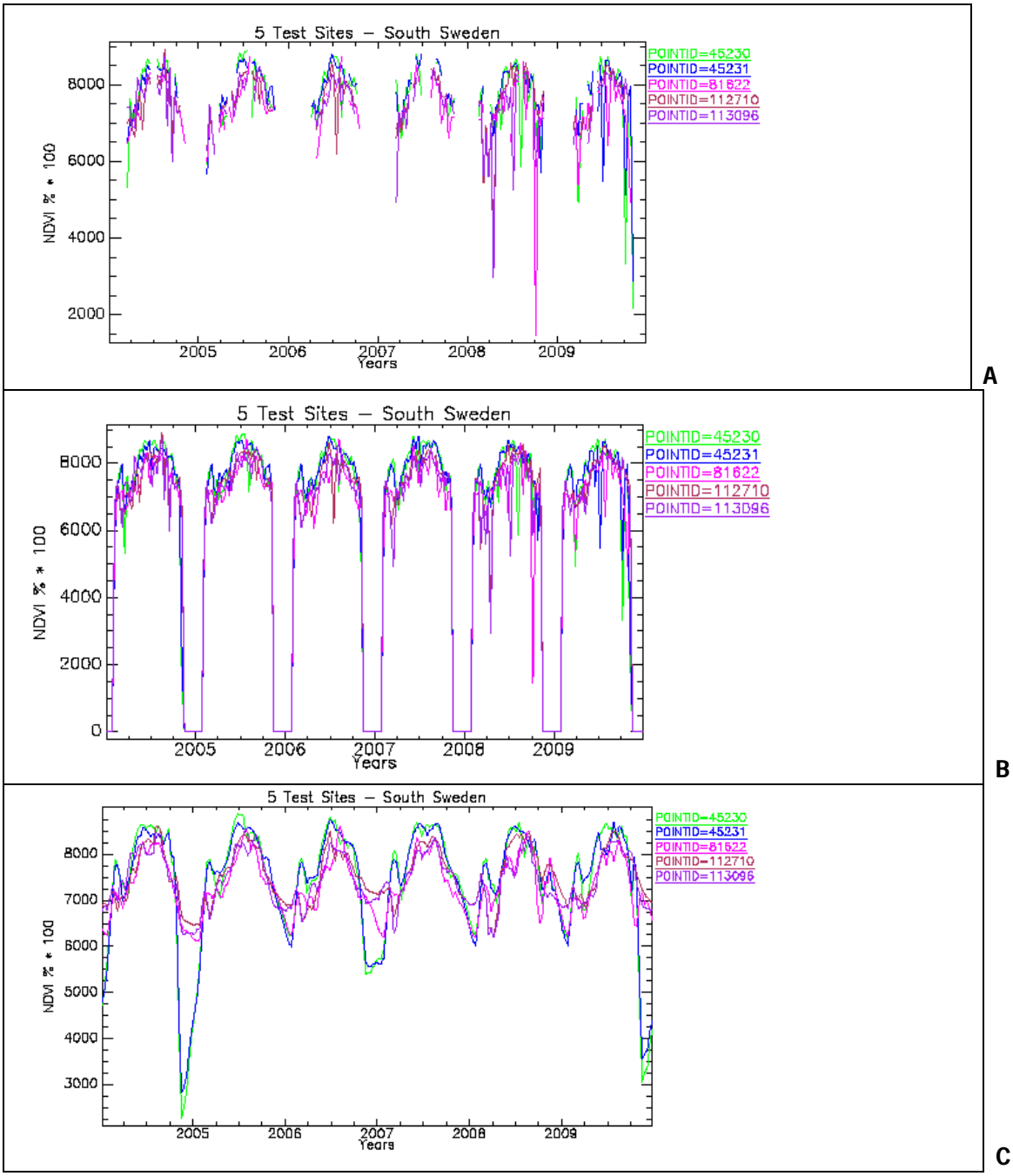


Figure 31

Time series of NDVI for five selected points for three pre-processing steps.

Filtering delivers a better representation of the real run of the vegetation cycle, attenuating or eliminating artefacts, as the ones visible in April and October 2008 in Figure 31A. However, some other artefacts (e.g. around November 2004) are not eliminated by the filter and recognized as valid data. Also, due to a missing reference, it is not clear which values would in reality correspond to proper winter minima for the presented vegetation cycles. In Figure 31C the lowest values calculated (excluding the drop in autumn 2004) assumes NDVIs on average around 0.6. As observed, the original NDVI data, in order to be adequately prepared for the extraction of phenometrics, are manipulated through a large number of processing steps that make use of a series of different assumptions and processing algorithms.
Statistics of single NDVI decades revealed that the majority of no data flags occur in winter period, January and December being the most affected months. This time interval was chosen as the more adequate to introduce the artificial no data sequences in order to simulate a real-like situation; the length of no data segments introduced is equal to ten decades each. A comparison of the processing steps for a sample point for the original NDVI time series and the same with added data gaps is shown in Figure 32 (A, original NDVI data series; B, substitution with seasonal means and outlier analysis; C, interpolate no data values between nearest existing points in time; D, Savitzky-Golay filtering).

After modification of the original NDVI series, the extraction of phenometrics using Phenolo was launched again for the MDN zone. The values of the phenology indicators for the CON and DEC training pixels were extracted. A RF classification is then launched using the same configuration as used for the data series with no gaps added (all phenometrics). An accuracy assessment, using the JRC Forest Map 2006, represented the final step. Results are reported in the following table, together with class accuracy values of the classification with original NDVI data.

Table 14

Class accuracies for the Ph0 phenometric configuration (31 phenometrics) using original NDVI values and with added data gaps.

Class Accuracy (%)	Ph0 with data gaps added	Ph0 original Data	Accuracy Decrease
Coniferous - CON	57.12%	57.40%	<1%
Deciduous - DEC	67.32%	67.63%	<1%

Classification accuracies using NDVI data with and without added gaps differ less than 1%. Insertion of data gaps, contiguous and located in the same temporal windows, did not produced remarkable effects in the extraction of date and productivity phenometrics. Pre-processing operations dealt effectively with data gaps: substitution/filtering operations generally produced plausible continuous NDVI time series (e.g. Figure 32D). Due to time constraints it was not possible to test other data gaps configurations, like non continuous gaps, no data segments longer than ten decades or with random distribution, etc. As a consequence, in order to infer general conclusions on data gaps influence on classification further tests are needed to cover a variety of data gaps distributions and applied to other regions (e.g. Boreal). Nevertheless, when the NDVI data series is characterized for a large part by data with few substituted values (as in the MDN zone) and gaps are located in the NDVI time series minima, data gaps appear to influence classification accuracy less than the site specific characteristics of forest stands.

An analysis to test the statistical significance of change in phenometrics extracted from original data and from data with added gaps was performed using a random subset of 4,085 points. The majority of phenometrics show a statistically significant change (Table 15). Boxplots of paired phenometrics are illustrated in Figure 33.

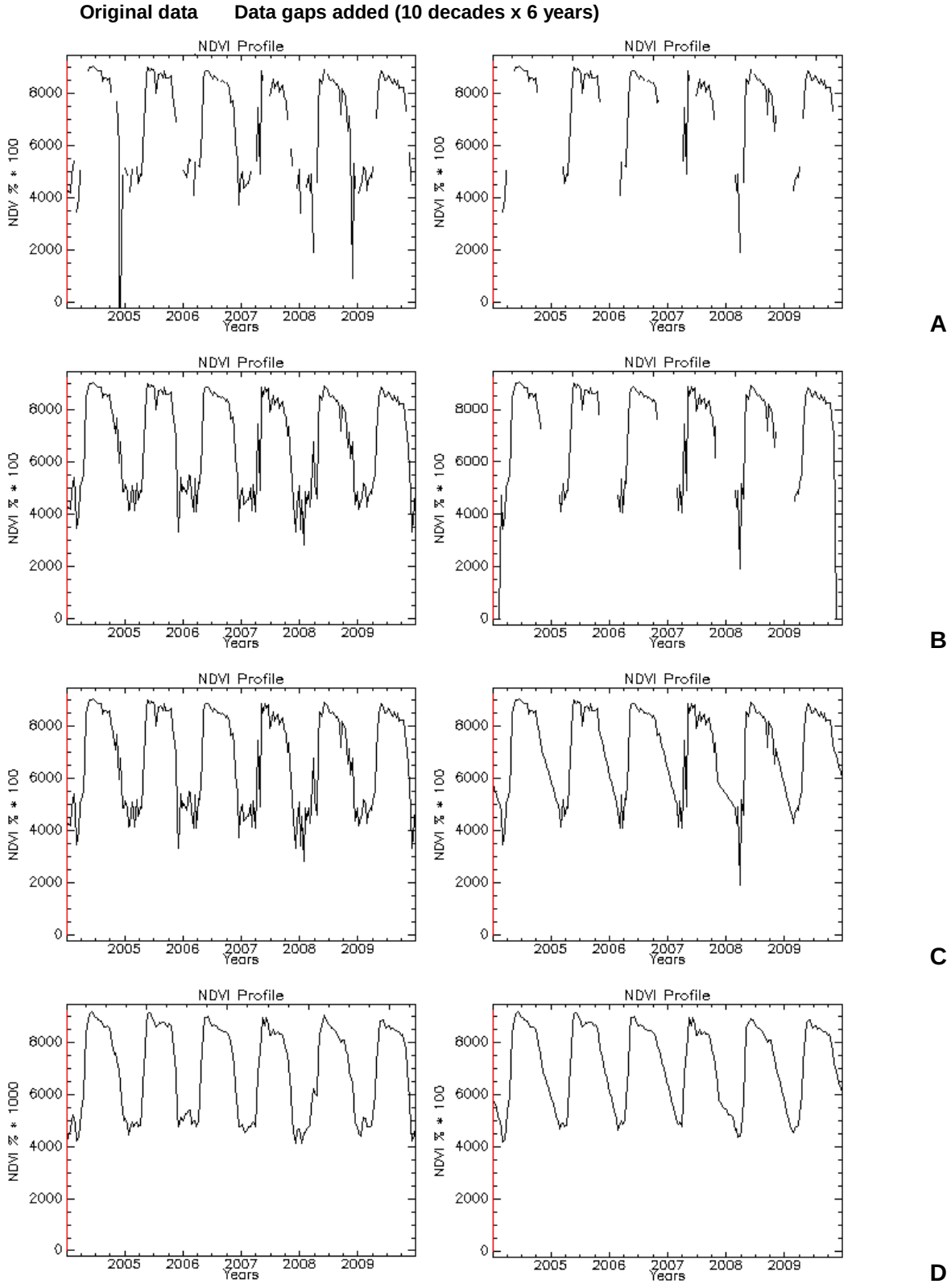


Figure 32

NDVI time series at different steps of the pre-processing chain for the original data (left column) and with added data gaps (right column); sample point of Deciduous forest.

Table 15

*Significance levels (stating that samples means are statistically different) for paired t-test between phenometric data without and with gaps. Significance levels: n.s: not significant (or p-value >= 0.01), *: 0.005 < p <= 0.01, **: 0.001 < p <= 0.005, ***: p <= 0.001. Samples number = 4,085.*

B1_sbd	***	B17_mbd	***
B2_sbv	***	B18_mbv	***
B3_sed	***	B19_med	***
B4_sev	***	B20_mev	***
B5_sl	***	B21_ml	***
B6_si	***	B22_mi	***
B7_sni	***	B23_mni	***
B8_str	***	B24_mtr	***
B9_spi	***	B25_mpi	***
B10_ged	***	B26_sei	n.s.
B11_gev	***	B27_gei	n.s.
B12_gl	n.s.	B28_sbc	n.s.
B13_gi	*	B29_ssd	***
B14_gni	***	B30_mxd	n.s.
B15_gtr	excluded	B31_mxv	***
B16_gpi	n.s.	B32_omi	***

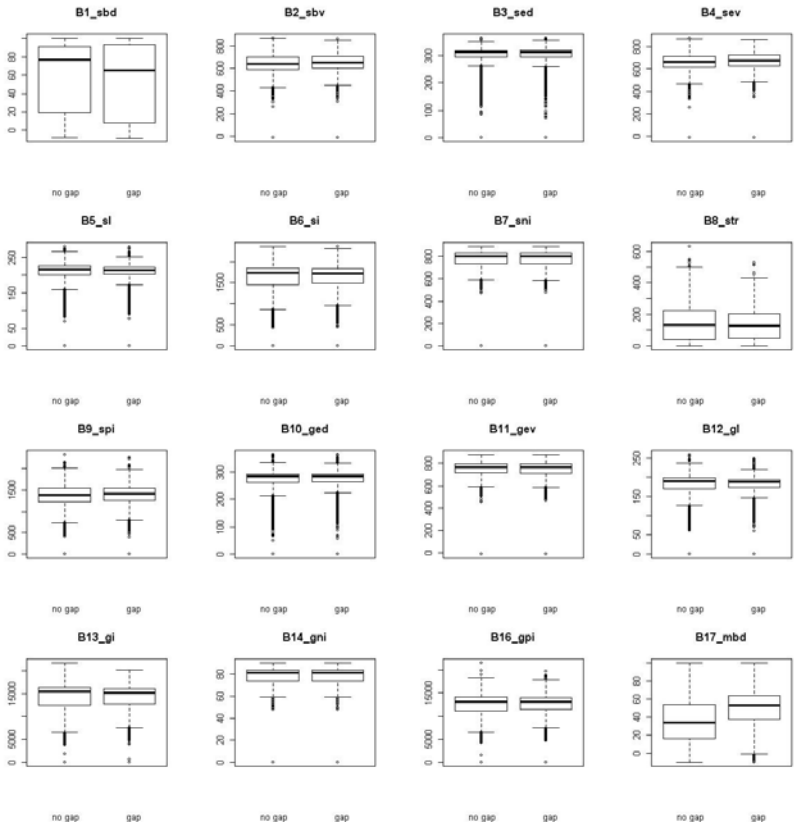


Figure 33

Boxplots of compared phenometrics values extracted by MODIS data of original and with no data gaps added (4,085 random points).

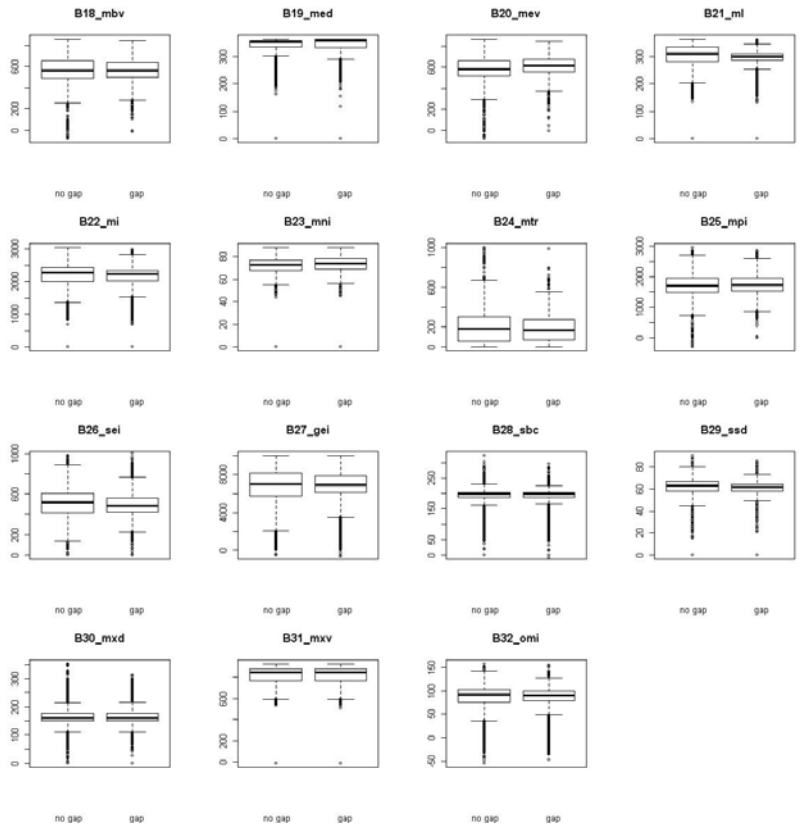


Figure 33 (continued)

Boxplots of compared phenometrics values extracted by MODIS data of original and with no data gaps added (4,085 random points).

4 Conclusions and recommendations

The JRC Phenolo model (version 2009) allowed the extraction of a large set of date and productivity phenology indicators from SPOT and MODIS NDVI time series. Model coded in IDL provided fast calculations within a stable environment. The degree of information redundancy (based on calculations of correlation matrix) present among the 31 Phenolo phenometrics suggests it is possibly to focus on smaller sets of indicators instead than a large set of metrics without reducing the effectiveness of a classification.

We have demonstrated that the Random Forests classification technique is an attractive method for classifying remotely sensed data because of the following reasons: 1) it is very fast in training large datasets, 2) it provides an error measure based on the set of training pixels (OOB), and more importantly 3) the RF algorithm gives an indication of variables importance in the classification (Gislason et al., 2004). In the tests performed, the Mean Decrease Accuracy (MDA) calculation generally indicated *date phenometrics* as more important for classification than productivity phenometrics. The most recurrent phenology indicators (top of MDA graphs) were located around the Peak of Season point (MXV, MXD) and the curve absolute minima (MBV, MEV). Nevertheless, further analyses are needed to infer more general rules on single phenometrics importance, as defined by Phenolo, for habitat classification. The OOB error did not show a recurrent pattern. In our tests, in presence of correlated phenometrics and well differentiated training pixels among classes, the use of a small selected set of phenology indicators produced higher classification accuracy. This trend can be different when these conditions are not respected, such as using noisy training datasets.

Apart from spatially and spectrally homogeneous classes (FPH/CON in Austria), the overall classification accuracy achieved based on Random Forests and MODIS-based phenology indicators is generally not satisfactory. The following factors were considered to negatively influence the intercalibration exercise:

1) The GHC scheme makes use of general categories that allow degrees of heterogeneity in the classification of the same habitat category. For the GHC forest category (FPH) the proportion of treed vegetation covers ranges in the wide interval from 30% to full coverage. This heterogeneity is reflected in remarkable variance associated to the NDVI trends of the training pixels, and consequently in the RF classification. In Mediterranean or semi-arid environments this is possibly more evident, due to the characteristics of the different bare or scarcely vegetated soils.
2) The number of GHC field plots data currently available did not allow retrieving highly populated sets of pure pixels for the classifier training. This would limit the possibility to take into account the large variability of GHC forest vegetation signal over different environmental zones. Moreover, no pure pixels were obtainable for the FPH/EVR class, thus introducing an additional noise component.
3) The accuracy assessment was performed using information from the JRC Forest Cover Map 2006. This continental dataset, built uniquely on spectral information, has no rigid correspondence with the GHC forest classes. Hence, increased 'mismatches' could have been measured when comparing the datasets.

The introduction of artificial data gaps within the MODIS NDVI time series was not very relevant towards classification accuracy (test for the MDN environmental zone). The applied pre-processing operations effectively dealt with no data decades by reconstructing reasonable NDVI curves. Changes in artificial gaps configuration (number, time location, etc.) and zone of analysis could potentially provide more noise in the signal, and affect more significantly the classification accuracy.

On the basis of the above results our concluding remarks are as follows:

- The spatial scale of current EO-based phenology data (250m) is at the edge of an adequate resolution for effective habitat classification with respect to the GHC categories. MODIS 250m grid overlapped on high resolution GHC field plots provide polygons with a variety of mixed classes, which are difficult to classify and unmix.
- For the intercalibration of GHCs with EO-based phenology indicators, the production and use of a large dataset of GHC training pixels ('pure pixels') is recommended to take into account the high spectral variability present within single GHC classes. This can be achieved by the sampling of several field plots in different Environmental Zones with a variety of local conditions.
- An adequate classification accuracy assessment should be based on a reference dataset which is not processed uniquely using spectral information, but that is built taking into account as much as possible the elements of heterogeneity typical of the General Habitat Categories. This can be possibly addressed using regional or national habitat map and datasets.

To conclude, the elements characteristics of the life forms types considered in the General Habitat Category scheme (e.g. height of stand) are very valuable information to be taken into account in intercalibration using EO-derived information. For this reason and for the purpose of GHCs classification we believe that a strategy which integrates EO-based phenology indicators with other remotely sensed information, such as typically LiDAR or high resolution radar, can be potentially more effective than a purely phenology-based approach.

Acknowledgements

The authors thank Dr S. Kay and the MARS Unit of the EC JRC for providing the raw NDVI SPOT/MODIS data, Dr E. Ivits and team (EC JRC) for providing Phenolo software/material and Dr S.Mucher (Alterra, the Netherlands) for supplying the modelled habitat data for Austria and Slovakia. Thanks are due also to Dr F. Gerard (CEH) for insightful discussions and document revision, and Dr P. Zuccolotto, Dr M. Sandri (University of Brescia) for their suggestions with the Random Forests algorithms.

5 References

Bartholomé, E. and A.S. Belward, 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing, Volume 26, Number 9, pp. 1959-1977(19).

Breiman, L., J.H. Friedman, R.A. Olshen and C.J. Stone, 1984. Classification and regression trees. Monterey, Calif., U.S.A.: Wadsworth, Inc.

Breiman, L., 2001. Random forests. Machine Learning, 45, 5-32. Elsevier.

Bunce R.G.H., M.M.B. Bogers, D. Evans and R.H.G. Jongman, 2010. D4.2: Rule based system for Annex I habitats. EBONEDeliverable report.www.ebone.wur.nl/NR/rdonlyres/DADAAB1E-F07C-4AA3-8621-20548A9B7DE6/106197/EBONED42KeyAnnex1.pdf

Bunce, R.H.G., M.J. Metzger, R.H.G. Jongman, J. Brandt, G. de Blust, R. Elena Rossello, G.B. Groom, L. Halada, G. Hofer, D.C. Howard, P. Ková□, C.A. Mücher, E. Padoa-Schioppa, D. Paelinx, A. Palo, M. Perez-Soba, I.L. Ramos, P. Roche, H. Skånes and T. Wrbka, 2008. A Standardized Procedure for Surveillance and Monitoring European Habitats and provision of spatial data. Landscape Ecology, 23:11-25.

Bunce R.G.H., C.J. Barr, R.T. Clarke, D.C. Howard and A.M.J. Lane, 1996. Land Classification for Strategic Ecological Survey. J. of Environmental Management 47:37-60.

Chen J., P. Jönsson, M. Tamura, Z. Gu, B. Matsushitaand L. Eklundh,2004.A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter, Remote Sensing of the Environment91, pp. 332-344.

EEA -European Environment Agency-, 2010. Corine Land Cover 2000 seamless vector data (Publish date: 27 May 2010).URL: http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2000-clc2000-seamless-vector-database-2

Geerken, R., B.Zaitchikand J.P. Evans, 2005. Classifying rangeland vegetation type and fractional cover of semi-arid and arid vegetation covers from NDVI time-series, International Journal of Remote Sensing26(24), pp. 5535–5554.

Geerken, R., 2009. An algorithm to classify and monitor seasonal variations in vegetation phenologies and their inter-annual change. ISPRS Journal of Photogrammetry and Remote Sensing. Volume 64, Issue 4, July 2009, Pages 422-431.

Gislason, P.O, J.A. Benediktsson and J.R. Sveinsson, 2004. Random Forests for land cover classification, Pattern Recognition Letters, Volume 27, Issue 4, Elsevier.

Hill, M. and 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

Hill, R.A., K. Granica, G.M. Smith and M. Schardt, 2007. Representation of an Alpine Treeline Ecotone in SPOT HRG Data. Remote Sensing of Environment, 110 (4), pp. 458-467.

Holben, B.N., 1986. Characteristics of maximum-value composite images from temporal AVHRR data. International Journal of Remote Sensing, 7(11):1417-1437, pp. 562-571.

Ivits, E., M. Cherlet, W. Mehl and S. Sommer, 2008. Spatial Assesment of the Status of Riparian Zones and Related Effectiveness of Agri-Environmental Schemes in Andalusia, Spain , EUR Report, EUR 23299 N1018-5593, 80p, JRC Report, Luxembourg.

Ivits, E., M. Cherlet, W. Mehl and S. Sommer, 2009. Estimating the ecological status and change of riparian zones in Andalusia assessed by multi-temporal AVHHR datasets,. Ecological Indicators, Volume 9, Issue 3, May 2009, Pages 422-431, DOI: 10.1016/j.ecolind.2008.05.013.

Jönsson, P. and L. Eklundh, 2002. Seasonality extraction by function-fitting to time series of satellite sensor data. IEEE Transactions on Geoscience and Remote Sensing, 40, 1824–1832.

Jönsson, P. and L. Eklundh, 2004. TIMESAT - a program for analysing time-series of satellite sensor data, Computers and Geosciences 30, 833-845.

Klisch, A., A. Royer, C. Lazar, B. Baruth and G. Genovese, 2005. Extraction of phenological parameters from temporally smoothed vegetation indices; ISPRS Archives XXXVI-8/W48 Workshop proceedings: Remote sensing support to crop yield forecast and area estimates.

Kempeneers, P., F. Sedano, L. Seebach and J. San Miguel Ayanz, 2010. A high resolution PAN-European forest type map based on multispectral and multi-temporal remote sensing data. Proceedings of ForestSat2010 Conference: Operational Tools in Forestry using Remote Sensing Techniques. 7-10 Semptember 2010, Lugo y Santiago de Compostela, Spain.

KoebleR. and G. Seufert, 2001. Novel Maps for Forest Tree Species in Europe. http://afoludata.jrc.ec.europa.eu/index.php/public_area/tree_species_distribution

Liaw, A. and M. Wiener,2002. Classification and Regression by RandomForests. R News 2(3), 18-22.

Liaw, A. and M. Wiener, RandomForest: Breiman and Cutler's RandomForests for Classification and Regression, 2009. URL: http://cran.r-project.org/package=randomForest. R package version 4.5-30.

Lohninger, H., 1999.Teach/Me Data Analysis, Springer-Verlag, Berlin-New York-Tokyo.

Metzger, M.J. and R.G.H. Bunce et al.,2005. A climatic stratification of the environment of Europe. Global Ecology & Biogeography. 14: 549-563.

Mücher, C.A., S.M. Hennekens, R.G.H. Bunce, J.H.J. Schaminée and M.E. Schaepman, Modelling the spatial distribution of Natura 2000 habitats across Europe. 2009. Landscape and Urban Planning, 92 (2), pp. 148-159.

Paola, J.D. and R.A. Schowengerdt, 1995. A Review and Analysis of Backpropagation Neural Networks for Classification of Remotely Sensed Multispectral Imagery. International Journal of Remote Sensing, vol. 16, no. 16, pp. 3033-3058, November 10, 1995.

Pekkarinen, A., L. Reithmaier and P. Strobl, 2009. Pan-European forest/non-forest mapping with Landsat TM+ and CORINE Land Cover 2000 data. ISPRS Journal of Photogrammetry and Remote Sensing 64: 171-183.

Rahman, H. and G. Dedieu, 1994: SMAC: A simplified method for the atmospheric of satellite measurements in the solar spectrum. International Journal of Remote Sensing, 15, 123-143

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

Steenkamp, K., K.J. Wessels, S.A. Archibald and G. von Maltitz, 2008. Long-term phenology and variability of Southern African vegetation. IEEE International GeoSciences and Remote Sensing Society (IGARSS) Symposium. Boston, Massachusetts, U.S.A., 6-11 July 2008.

Steinberg, D. and P. Colla, 1995. CART: Tree-structured non-parametric data analysis. SanDiego, Calif., U.S.A.: Salford Systems.

Weissteiner, C., S. Sommer and P. Strobl, 2007. Time series analysis of NOAA AVHRR Green vegetation fraction as a means to derive permanent and seasonal vegetation component.Proceedings from the EARSeL Workshops in the framework of the 27th Symposium. June 7-9, 2007. Bozen (Italy).

Wessels, K., K. Steenkamp, G. von Maltitz and S. Archibald, 2011. Remotely sensed vegetation phenology for describing and predicting the biomes of South Africa. Applied Vegetation Science 14:49-66. Wiley.

PART II: MULTI-TEMPORAL ANALYSIS OF NDVI FOR GRASSLAND MAPPING AND CLASSIFICATION

Andrej Halabuk[2]

Institute for Landscape Ecology, Slovak Academy of Sciences

6 Introduction

Since the first temporal composite data of NDVI from AVHRR has become available in early 80-ies, a massive utilization of these data series emerged. This lead to the development of novel approaches for analysing time series of satellite data with high temporal resolution for land cover mapping and assessment. Within this deliverable, we have briefly reviewed the main approaches available and through case studies assessed their performance for grassland mapping and monitoring. Hence, we considered studies that used data from sensors with high temporal resolution, namely: AVHRR, SPOT VEGETATION, MODIS and MERIS. Grassland ecosystems were selected as land cover class with special interest within the EBONE team as these ecosystems along with their importance as an agricultural landscape, also bear high biodiversity values and there is still lack of precise information of their spatial extent and status at pan European scale. The importance of precise grassland monitoring is proven by their incorporation as a special class in the recently established land monitoring component of the GMES where specific approaches are being developed in order get more precise information on specific features such as grasslands, wetlands, forests. First synthesized results of utilization of multi-temporal AVHRR NDVI data for grassland vegetation assessment were reported from semiarid Africa in the special issue of Intern. Journal of Remote Sensing in 1986 proving that coarse data (8 km spatial resolution) with high temporal resolution could be used for grassland vegetation monitoring at regional scale. This issue discussed the first time series based approaches for detecting the different phenologies of grassland vegetation types by satellite data, their interrelation to regional climatic conditions, estimating of biomass, detecting droughts, etc. Since that time, and even more after the Terra and Aqua satellites with the MODIS sensor on board have been launched, a wide range of possible applications have emerged in grassland remote sensing that could be broadly classified into three groups:

1. grassland mapping
2. grassland type classification and monitoring
3. grassland change analysis

Grassland mapping includes those applications when there was no or poor knowledge about grassland occurrence. These are mainly relevant at broader scale and grassland mapping actually occurred within the global and regional land cover mapping initiative using multi-temporal analyses. Grassland type classification and monitoring includes a broad variety of studies which location of particular grasslands was known, focus on multi-temporal analyses to increase the knowledge on grassland status e.g. increasing thematic detail of grassland types, biomass estimation, stress detection, fire risk, etc. Grassland change analysis involve those studies that utilize longer time series in order to assess changes in grassland occurrence and status, e.g. overgrowing, drying, change of Land management practice, and studies of trend and seasonal change, phenology shifts and prediction of future grassland change including climate change studies. In this deliverable we were mainly focused on the first two categories: grassland mapping and grassland type classification and monitoring.

6.1 Multi-temporal analysis of NDVI Time series

Multi-temporal approaches comprise a broad group of techniques that explore and analyse dynamic changes of the land surface. In particular, seasonal change and long-term change of vegetation greenness are mainly studied using time series of vegetation indexes such as EVI or NDVI. NDVI time-series belong to the most widespread and commonly used to characterize seasonality, inter-seasonal variability, phenology, overall productivity, biomass, etc. (see Pettorelli et al., 2005 for review). Since coherent time series of NDVI at global

scale have been available, specific methods have been developed using information of distinct seasonality or phenology of land cover classes for land cover classification and mapping (Reed et al., 1994; DeFries et al., 1995). These approaches would use multi-date NDVI images that reflect main seasonal differences in a specific region (Wessels et al. 2004); various multi-temporal metrics (DeFries et al., 1998; Sedano et al., 2005; Samson et al., 1993; Paruelo et al., 2001); phenometrics (Reed et al., 1994); or time series analysis of whole annual NDVI series (Evans and Geerken, 2006; Jakubauskas et al., 2001; Hall-Bayer et al., 2003). All variety of image classification methods using supervised, unsupervised or multi-temporal segmentation (here later referred as "multi-temporal classification") were used to produce a full coverage land cover map (Mucher et al., 2000); classify specific ecosystems (Hill et al., 1999; Paruelo et al., 2001); or to extract phenology-based information of sample sites (Paruelo et al., 1998; Fontana et al., 2008). Different methods for pre-processing the noisy raw time series data have also emerged, and have been thoroughly tested and validated across the different biomes and scale (Hill and McDermid, 2009). In our case studies we tested different approaches using mainly MODIS 16 day NDVI and 8 day surface reflectance composite data with 250m spatial resolution, which are of the most commonly used products for multi-temporal land cover mapping at regional scale.

6.2 Grassland mapping

Grassland mapping and detection is mainly incorporated in land cover classification initiatives at global, national and regional scales where grasslands are represented as 1 or more land cover classes. Grasslands in different classification legends are defined quite consistent usually as the land cover class with prevailing open herbaceous or grassy vegetation with less than 10% tree cover. Coarse spatial resolution satellite data with global coverage stimulated the development of many large scale land cover mapping products that use information of specific temporal behaviour of the land cover classes for their classification. The main global land cover maps that exist (Discover, Vega200, UmD, Globcover, Modis GLC) proved variable validation statements for grasslands however validation of global products is tricky and validation statements cannot be reasonably compared with regional studies. It is obvious, that despite of the extremely valuable contribution to global studies, these global coarse resolution products fail to capture land cover and land use variability at finer scales (national and regional) especially in heterogeneous landscape where mixed pixels are prevalent. As a result coarse resolution maps have to use complex mosaic classes in their classification system. For example, Mucher et al. (2000) stated that almost 27 % of Europe occurred in complex vegetation mosaics of the IGBP Discover product. These complex classes can include grassland ecosystems and thus contribute to underestimation and uncertainty in grassland distribution. Sedano et al. (2005) by using 250m MODIS product demonstrated that the 1 km MODIS GLC product does not capture the spatial land cover variability of the study area, providing a very broad classification in which most of the region is covered by only two classes (savannah and woody savannah). These two classes account for 59.33% and 36.1%, respectively of the study area, misrepresenting the heterogeneous distribution pattern of the land over and land use types in the study area. This clearly showed that the small-scale agriculture, dominating the region, could not be separate at 1 km spatial resolution. In addition, due to their similar phenology and structure, grasslands and agriculture crops are difficult to separate at 1 km resolution. Although we did not carry out that quantitative validation of the global grassland products at regional scale, Fig. 1 illustrate that both the distribution and coverage of grasslands differ a lot in our study area and global products will probably not be suitable for grassland monitoring in Europe.

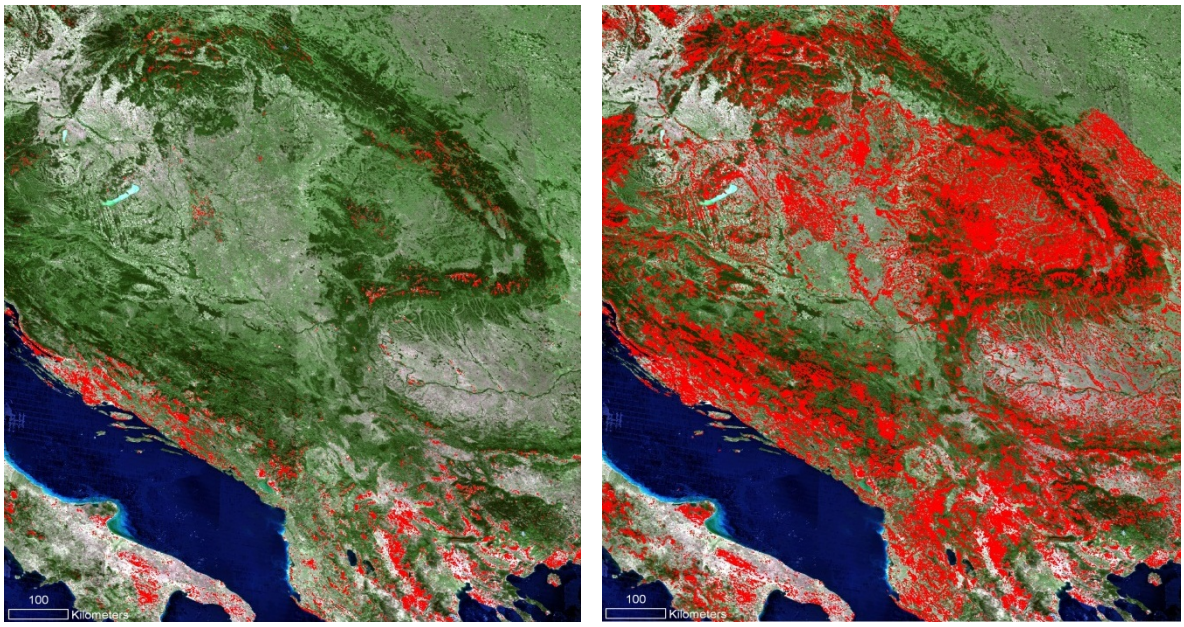


Figure 1

Comparison of distribution grassland land cover class in MODIS GLLC product (left) and CLC class 231 product (right).

At continental scale there were also many initiatives using coarse resolution sensors for multi-temporal classifications (see Franklin and Wulder, 2002 for overview). For example Stone et al., (1994) used a multistep combined approach where seasonality was also used to map the land cover in South America. They achieved a good accuracy for the grassland classes ranging from 81% (savannah/grassland) to 87.5 % (mountain grasslands). However due to the small area of flooded grassland in the validation sample these grasslands only achieved 36% accuracy. Giri and Jenkins (2005) used a multi-date classification of MODIS surface reflectance product at 500 m resolution to produce a land cover map for greater Mesoamerica and reported a grassland producer's accuracy of 60%; a user's accuracy of 65% and a substantial misclassification with shrub lands. In Europe, the PELCOM initiative (Mücher et al., 2000) delivered a land cover map with grassland user's accuracy of 42% and producer's accuracy of 36 % using a heterogeneous validation data set and 53% and 52% using homogenous validation data set. At regional scale, regionally specific approaches can benefit from specific knowledge on drivers and effects that determine the distinct seasonal profile of different land covers. Mucher et al., (2000) succeed in mapping grassland at national scale in the Netherlands by exploiting these distinct seasonal features between grassland, arable land, forest and water: In early spring, there is a significant difference in spectral reflectance between grassland and arable land because most arable land is still bare. Exceptions are areas covered with winter wheat, which will be confused with grassland. However, as winter wheat is harvested around the end of July, wheat can be separated from grassland using data acquired after July. Similarly Moody et al. (2001) described regional specific distinct temporal profiles of NDVI for main vegetation types (including grasslands) and the use of the Fourier terms (amplitudes and phases of a series of harmonics) derived from the NDVI time series of single year for the distinction of the main vegetation types in southern California. The mean NDVI, or 0th-order harmonic, indicated overall productivity, allowing the differentiation of unproductive, moderately productive, and highly productive sites. The amplitude of the he first harmonic indicated the variability of productivity over the year as expressed in a single annual pulse of net primary production. This summarized the relative dominance of evergreen vs. deciduous or annual habit. The phase of the first harmonic summarized the timing of green-up relative to the timing of winter and spring rains. This differentiated rapidly responding annual grasslands from slowly responding evergreen life-forms, and irrigated agriculture. The second harmonic indicated the strength and timing of any biannual signal. This provided information on secondary vegetation types, such as subcanopy grasses beneath evergreen

woodlands or mixtures of annual grasslands and irrigated agriculture. The authors reported an overall 77% producer's and 46% user's accuracy with high levels of misclassification of the woody and open shrub savannah. Jakubauskas et al. (2001) used similar approach in cropland dominated landscape in southwestern Kansas in the High Plains. The authors reported the best classification accuracy for grasses (72 producer's, 66% user's accuracy) and the weak accuracy for wheat. The small patches were the worst classified demonstrating how mapping is often confounded by the spatial resolution of the fields being smaller than the resolution of the scanner. Sedano et al. (2005) used multi-temporal analysis of MODIS 250m for a rapid, operable and low-cost land cover mapping product protocol developed for natural resources and biodiversity. The monitoring concluded that complementary to existed global land cover product their regional specific approach based on knowledge of specific phenology responses of main vegetation types can effectively map and monitor regional ecosystems. Using different multi-temporal spectral indexes and 250m multi-temporal NDVI they were able to capture grasslands at around 85% accuracy. Huang and Siegert (2006) optimized a land cover classification of North China to detect areas at risk of desertification. By exploiting the seasonal information from the 1km NDVI time series product of Spot VGT they increased the thematic detail of the grasslands class by subcategorizing it into – "sparse", "dense" and "mixed with agriculture", achieving producer's accuracy of 66 %, 91 %, 86 % and user's accuracy of 89 %, 75 % and 79 % respectively. The authors proved that different vegetation types and land use practices show distinct temporal shape of the NDVI. Geerken et al. (2009) developed shape based classification system for Middle East study area matching pre-defined class specific temporal profiles of NDVI with the temporal response of respective pixels. The authors defined 32 classes in the total with inclusion of many mixed and mosaic classes such as "dense bushes and grass", "shrubs, grasslands, low productive", "shrubs, grass, scattered fields", "sparse grasslands" with high values of accuracy in all mapped classes. However, when there is no a-priori knowledge on the temporal behaviour of specific grasslands (e.g. drivers and effects that affect temporal NDVI profiles) or when grassland seasonality varies a lot in space and between years, the mapping accuracy of grasslands is greatly reduced even at regional scale. For example Gao et al. (2009) using object-oriented multi-temporal classification for Mexico achieved disappointing grassland mapping results (22 % user's accuracy and 59% producer's accuracy), the main misclassification being with temperate forests. Similarly, Matsuoka et al. (2004) struggled to sufficiently discriminate agricultural fields from grasslands in the Yellow river basin which they attributed towards the lack of distinction between the multi-temporal metrics of grasslands and single cropped agricultural fields. Mucher et al. (2000) using a classification approach, that was very successful in The Netherlands (80% accuracy for grasslands), do not obtain satisfactory results in Eastern Spain which was explained mainly by a variable eco-climate and a heterogeneous landscape. Ratana et al. (2005) showed that the MODIS seasonal - temporal VI profiles are highly useful in mapping converted pasture in a highly complex and diverse cerrado landscape, however grassland and shrub cerrado formations were difficult to separate using their seasonal profile. To conclude, grasslands can be relatively successfully mapped in cases when they represent relatively large compact homogenous areas of similar grassland types or when a regional adaptive approach benefits from a-priori knowledge of the distinct seasonality of the respective mapped land cover classes. Taking into account all the above mentioned issues, we tested different approaches (including data pre-processing, classification strategies, training and validation data set) in order to assessed the added value of multi-temporal analysis of MODIS 250m NDVI time series for mapping grasslands in heterogeneous landscape in Slovakia.

6.3 Grassland classification and monitoring

Grassland classification aims to bring more precise knowledge on grasslands in terms of types or condition when the location of grasslands is already known. Grasslands with similar physiognomy may have different temporal pattern of NDVI affected by a broad range of natural or human driven factors. Multi-temporal analysis of NDVI time series iteratively explores the main determinants of seasonality and uses this information for the subsequent classification of grasslands to determine their vegetation type, status and functioning. Grassland

classification may include both full coverage classification of grassland areas or exploration and classification of main grassland types of sampled areas. For example Aragon and Oesterheld (2008) used a combined approach using information of spatial arrangement (from single date HR Landsat TM image) and information on functional properties (NDVI dynamics derived from multi-temporal MODIS 250m NDVI series) to map grassland vegetation communities in Argentinean flooded Pampa grasslands. The authors successfully classified 5 grassland vegetation types with an overall accuracy of 76% and documented that grassland vegetation communities significantly differ in their seasonal and interseasonal pattern of NDVI. Wen et al. (2008) identified 6 main grassland types in Tibet using time series analysis of MODIS NDVI profile with estimated accuracy ranges from 21 % (high cold meadow steppe) to 68 % (high cold typical steppe). Hill et al. (1999) classified a pastoral landscape in eastern Australia resulted into 8 broad categories: sown perennial pastures, sown perennial pastures with woodland, sown annual pastures, mixed pasture and cropping, native pastures, native pastures with woodland, degraded or re-vegetated areas and forest. Boles et al., (2004) distinguished 4 main types of grasslands in one of the world's largest grassland region of temperate East Asia: typical steppe, eadow/meadow steppe, meadow steppe/typical steppe, shrubland / grassland. Although they do not report a validation statement of each grassland type, the aggregated grassland class was mapped with an estimated 60 % user's and 46 % producer's accuracy. Paruelo et al. (1995, 1998) characterized seasonal patterns of the NDVI in 49 grasslands in North America with low human impact and derived indicative metrics of NDVI curve in order to describe climatic controls on grassland structure and functioning. Later Paruelo et al. (2001) used NDVI dynamics as a descriptor of ecosystem functioning and widely applied this approach for mapping and classifying ecosystem functional types. They used three measures calculated from the seasonal curve of NDVI: annual integral of NDVI as an estimate of primary production, relative annual range of NDVI and date of maximum NDVI both of which were used to capture the seasonality of primary production. There are many applications that using NDVI time series for grassland assessment. Perhaps the most widely used approach represents the analyses of the NDVI temporal profiles (and derived surrogates) to explore climate-grasslands relationships. This represents a long time effort to detect climate driven threats on grassland ecosystems. These grassland phenology studies include trend and seasonal change detection as a tool for climate change impact assessments (Tao et al., 2008). Land use change impacts on the NDVI profile can also be analysed at regional scale (Paruelo, 2001) and indicators derived for direct or indirect biodiversity assessment (Coops et al., 2009; John et al., 2008; Huang et al., 2009). In our case study we explored the use of temporal NDVI profiles of grasslands to characterise grassland structure and functioning including their specific phenology, management, hydrology, and overgrowing.

7 Case study 1 - Grassland mapping

In our mapping case study we attempt to map the grassland land cover class in Slovakia by exploiting differences in the seasonal NDVI dynamics. The main objective was to test whether the grassland land cover class (as defined in CLC classification system) can be discriminated from other land cover classes, especially arable lands and forests by their specific seasonal NDVI profile. As mentioned above, temporal information based on seasonal profile of vegetation greenness (indicated by NDVI) has been used for rapid large scale land cover mapping. However the coarse spatial resolution of satellite data with high temporal resolution is considered as the main limitation for mapping of grasslands within the heterogeneous landscape. In this case study we tested the possibilities of multi-temporal classification for full coverage mapping of grasslands and aimed to identify the main limitations of this approach for mapping of grasslands at regional scale in highly heterogeneous landscapes such as Slovakia.

Figure 2
Grassland distribution in Slovakia (CLC2000, class 231 – pastures).

7.1 Study site

Grasslands in Slovakia are formed mainly as small scattered patches with diverse spatial arrangement (Fig. 2). From the agricultural statistics the total coverage of grasslands (excluded natural alpine grasslands) is estimated to be 690 000 ha which represents approx. 15% of the Slovakia and 30% of the agricultural landscape. CLC grassland class 231 clearly underestimates the total coverage of grasslands, perhaps because of the minimal CLC mapping unit (25 ha) and aggregated classes (242 and 243). Grassland vegetation types vary broadly based on the nutrition, geology substrate, hydrology and elevation. Land use represents an important driver of grasslands including grazing with different intensities on pastures, cutting on meadows or both (spring cutting and autumn pasturing). A substantial proportion of grasslands have been abandoned and overgrown after the changes in socioeconomically condition in early 90-ties. Recently, agro

environmental subsidies have introduced special management in the most valuable semi natural grasslands in Slovakia.

7.2 Data and processing

We used MOD13Q1 (250m 16 day NDVI) and MOD09Q1 (250m 8 day surface reflectance) downloaded from Land Processes distributed active archived centre (LPDAAC) for the area covered by the MODIS grid tile h19v4 (Fig. 12). Usefulness index and quality assurance layer of MOD09A1 were used to minimize negative effects of clouds, cloud shadows, aerosols, sun-sensor geometries and snow. Missing data were later interpolated and smoothed using Savitsky – Golay filter incorporated in the TimeSat software (Eklundh and Jonson, 2004) in order to produce complete NDVI time series from 2001 to 2010. Those areas for which two subsequent 16 day observation were missing were masked and omitted from later analyses. In our mapping case study we used annual time series of NDVI for 2009. In order to explore the effect of temporal resolution (16 day vs. 8 day composite series) and time period using for classification (whole annual series vs. vegetation growth period) we carried out the following analyses:

1. Principal component analysis (PCA) of the whole annual 16d NDVI series
2. PCA of the vegetation growth period (22nd March - 1st November) of the 8d NDVI
3. Fourier analysis (TFA) of the whole annual 16d NDVI series
4. TFA of the whole annual 8d NDVI series

Results from these analyses e.g. PCA component scores, amplitudes and phases of Fourier transform were later explored and used for the image classification. Two basic approaches for the image classification were used. The first approach was a maximum-likelihood supervised classification which was applied to map 5 broad CLC classes:

Deciduous forests (CLC 311)
Arable land (CLC 21*)
Water areas (51*)
Grasslands (231)
Urban areas (1**)

We used the Corine land cover map from 2006 to learn the classification procedure (i.e. signature development). In order to minimize the number of mixed pixels (in multi-temporal space) within the CLC grassland polygons we first carried out a multi-temporal segmentation using 4PCA components (Fig. 3).

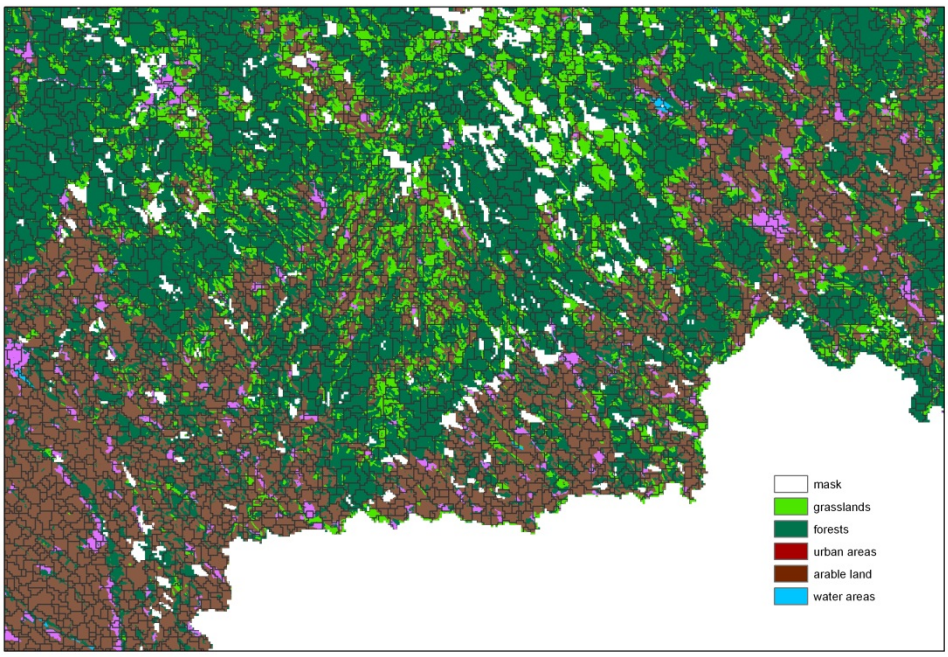


Figure. 3
Multi-temporal segments for signature development.

The second approach was an unsupervised classification approach. Here we used an ISODATA algorithm to produce multi-temporal distinctive clusters. We set the maximum number of classes to 15 based on an expert judgement (i.e. based on the exploration of PCA components and Fourier terms) with the expectation of classifying the following classes:

- winter crops (wheat, oil rape)
- annual crops (barley)
- annual crops (corn, sunflower)
- deciduous forests
- shrubs
- coniferous forests
- cut grasslands
- uncut grasslands
- intensive grasslands
- overgrown grasslands
- alpine meadows
- water areas
- urban areas
- bare lands
- occasionally flooded areas

The class assigning of the resulting clusters was based on expert judgement by evaluating temporal profiles of the cluster centres and by iterative validation of clusters with reference temporal profiles. Two kinds of data were used for the validation. The first set represented randomly selected grasslands identified through the Slovakian land parcel information system and resampled to match the MODIS resolution using a simple majority rule (totally

80000 pixels). This validation sample set was considered to be heterogeneous. The second validation sample set was a subset of the first set representing only the pure homogenous grasslands. These were identified through a thorough inspection of Google Earth. Totally 3758 homogenous grassland pixels were selected for the validation (Fig. 4). Here we need to point out that these validation dataset represent only grasslands in agricultural land (excludes natural grasslands or pastures in forest landscape). A simple cross validation was used to estimate classification accuracies. Producer's accuracy describes how many of actual grasslands have been mapped. User's accuracy describes how many of mapped grasslands are actual grasslands. Users accuracy however was difficult to estimate as we did not have a reliable knowledge on the extent and distribution of grasslands and we did not have a "true zero" validation data set. Therefore we reported two Users accuracies. The first uses CLC class 231 and 321 as the mask for "real" grassland distribution and the second uses CLC 231, 321 and 243, 242 as the mask.

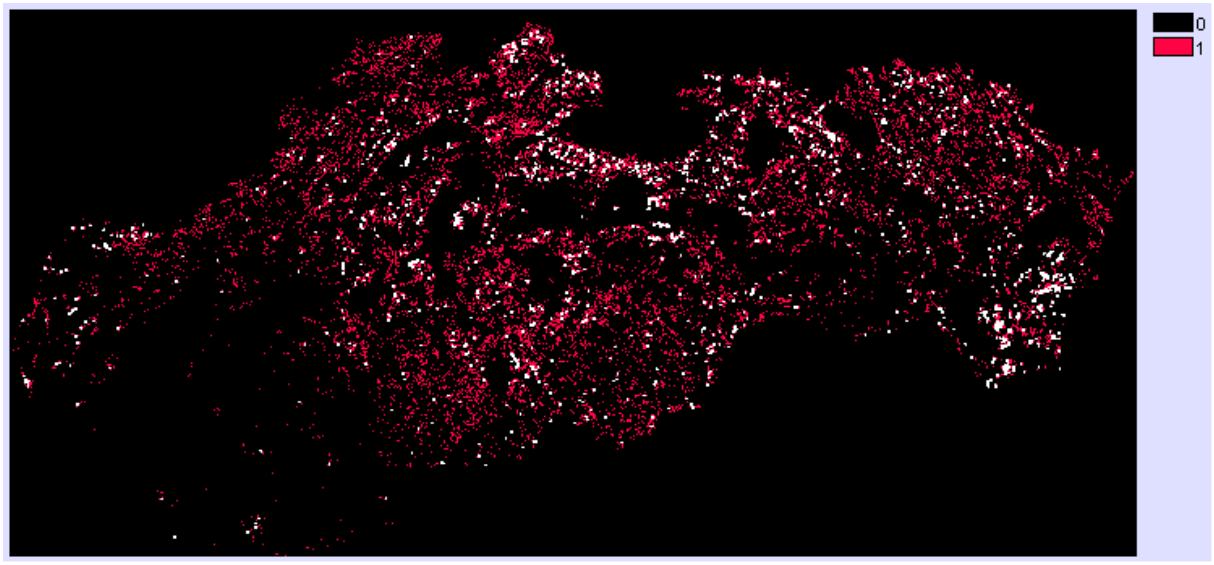


Figure 4
Grassland validation sites (red – all; white – homogenous pixels).

7.3 Results

PCA of 16 day NDVI composite of the whole 2009 period 4 PCA components were extracted that explain almost 80 % of the seasonal variation of the data. The first component mainly characterises the start of the season (spring green up) and start of the senescence period in autumn. In the Slovak case, this component is related to deciduous forests, shrubs and grasslands as these massively green up during April and May and remain relatively green until October compared to agricultural areas (Fig. 5).

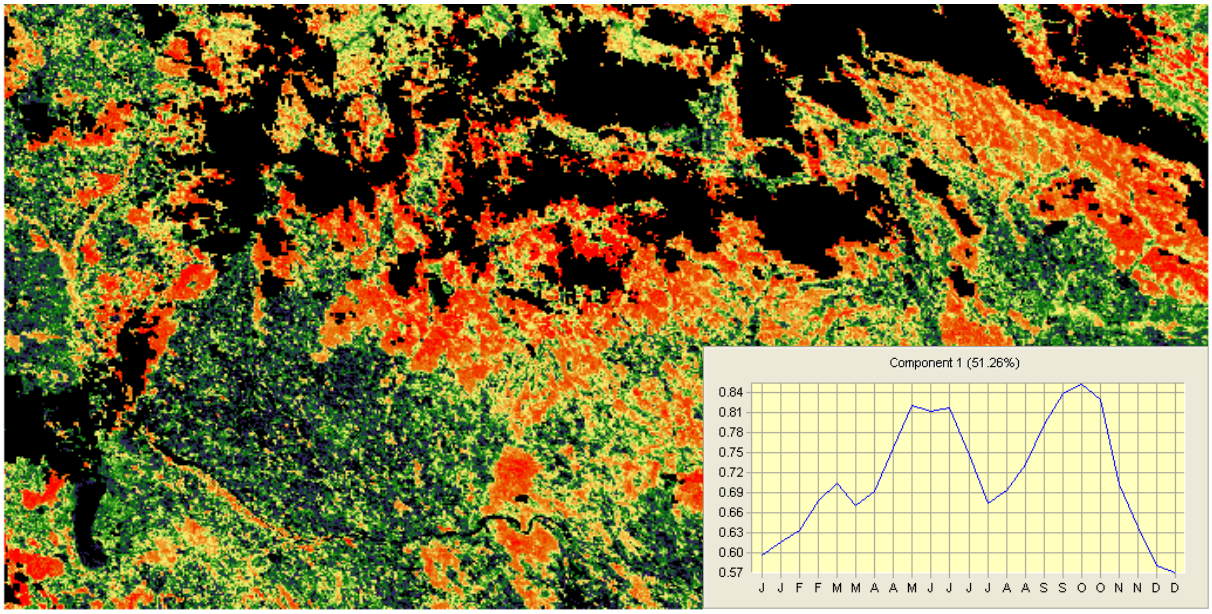


Figure 5

PC1 scores derived from PCA of the whole period (red – high, blue – low, black – mask).

The second component characterises the variation within the peak vegetation season e.g. July, August and explains mainly the annual crops (as barley, corn, and sunflower) and part of the intensive grasslands (Fig. 6).

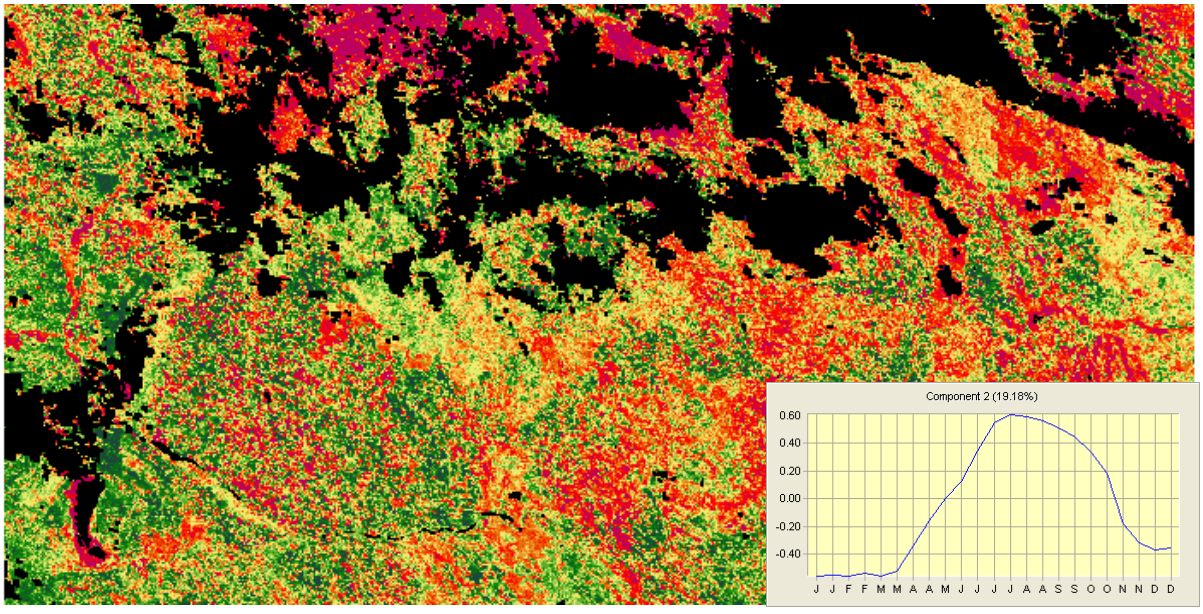


Figure 6

PC2 scores derived from PCA of the whole period (red – high, blue – low, black – mask).

The 3rd component explains mainly winter crops (wheat, oil rape) because they are relatively green in winter and the arable land is usually ploughed up and left bare the next autumn and winter (Fig. 7).

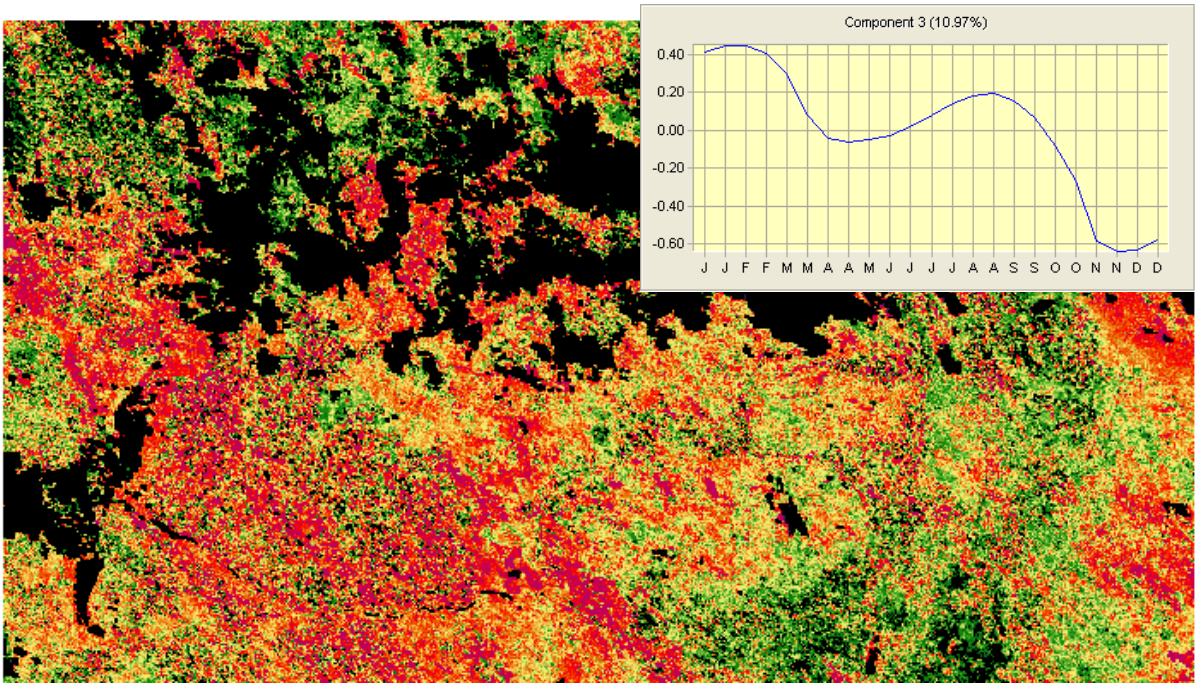


Figure 7

PC3 scores derived from PCA of the whole period (red – high, blue – low, black – mask).

The fourth component probably represents a specific example of those annual crops (e.g. sunflower and corn) where crop residuals from the previous season have been left on the land and are ploughed up in April (Fig. 8).

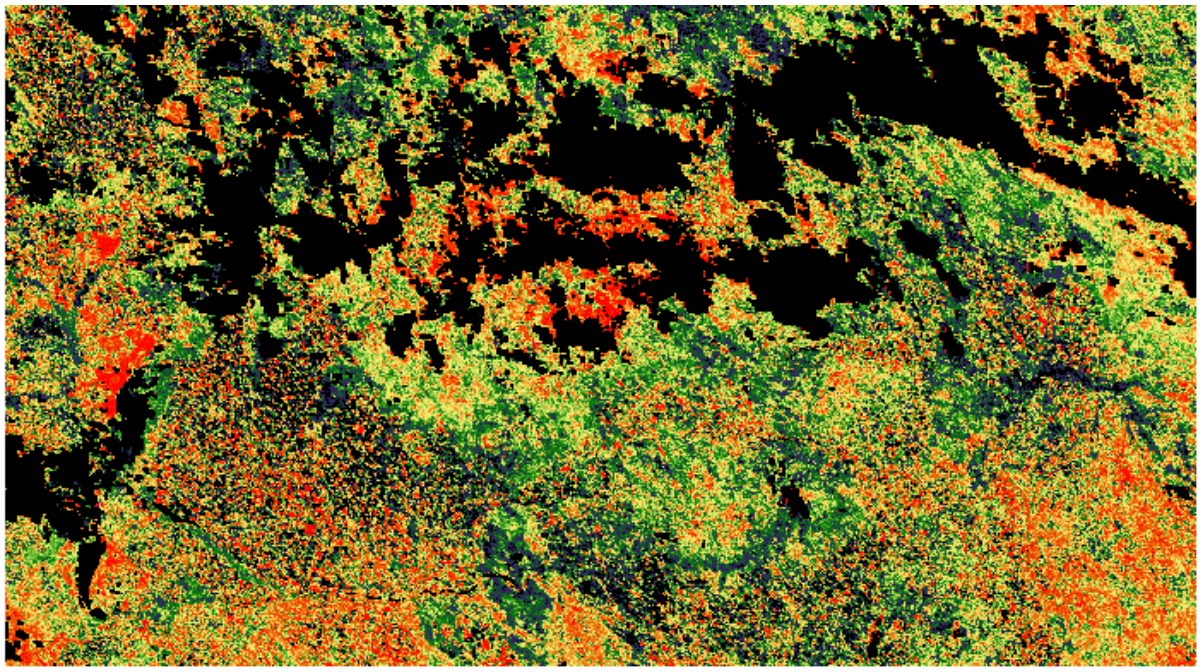


Figure 8

PC4 scores derived from PCA of the whole period (red – high, blue – low, black – mask).

PCA of 8 day NDVI composite during vegetation season period 2009

Six components were extracted which explain 89% of the total variance. The first 4 components represent same vegetation types as described by components derived for the 16 day full year time series. However, two additional components that represent higher sub-seasonal variability were extracted. Component 5 probably reflects a one cut management practice and component 6 two cut management practices in agricultural landscape which could also represent cutting meadows (Fig. 9).

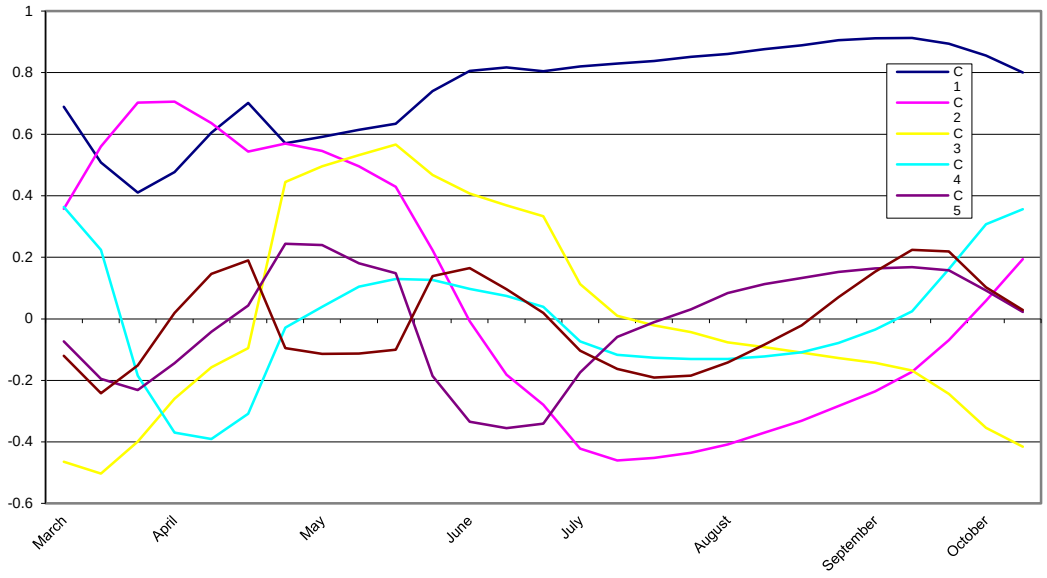


Figure 9

Loadings of PC derived from 8 day NDVI composite during the vegetation growth period (March – October).

Temporal Fourier Analysis

Three harmonics were extracted with associated phases in order to reasonably interpret and label the classification clusters. Preliminary tests revealed that higher order Fourier terms may introduce more noise in classification results. Here first harmonic represents the annual amplitude of the NDVI curve being the highest in alpine meadows, deciduous forests and annual crops, and the lowest in coniferous forests, urban and water areas. Associated phase distinguishes winter and summer crops. The second amplitude reflects bimodal seasonality visible in agricultural crops and cut grasslands. The associated phase should reflect timing of cutting. We left the third harmonics because we believed this could reveal two-cut meadows, however this seems to be difficult to validate as for the many uncertainties depended mainly on quality of the data series and applied smoothing techniques.

Classification

Multi-temporal signatures of forests, arable lands and grasslands (using 4 PCA components of the whole annual series of NDVI) indicate quite good separability although high variability was visible mainly for the arable land class. This is obvious because of different temporal response of individual agricultural crops. Cross tabulation of training samples with classified product indicates the main misclassification between grasslands and arable lands, however substantial disagreement was revealed between forests and grasslands (Tab. 1). Distribution of grasslands based on the supervised classification is shown on Figure 9. The results do not substantially differ when either PCA of 8 day vegetation growth period or Fourier terms were used for signature development.

Table. 1

Error matrix of training sample (column) and classified results (rows).

	G	F	U	A	W	Total	ErrorC
G	345	44	65	54	0	508	0.32
F	32	196	27	17	0	272	0.27
U	16	0	59	29	0	104	0.43
A	12	12	38	413	19	494	0.16
W	11	55	2	3	75	146	0.48
Total	416	307	191	516	94	524	
Error	0.17	0.36	0.69	0.19	0.20		

Expert judgement of 15 clusters derived from unsupervised classification indicates using 6 PCA components of 8 day NDVI composite for vegetation growth performed the most reasonable results. The main source of bad classification results across the all analyses represents erroneous classification of grasslands as forests. Therefore the main visible criterion for the best analysis was the minimum erroneous classification of grasslands as deciduous forests. The full coverage map of the PCA based classification is illustrated on Fig. 10 and derived grassland distribution map on Fig. 11.

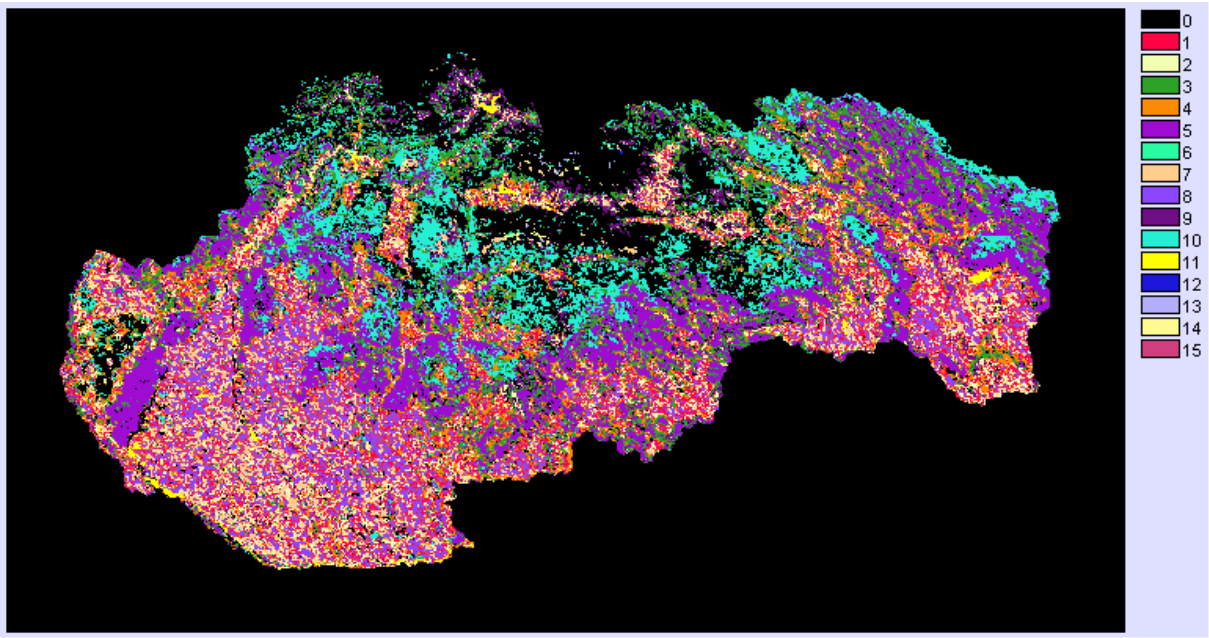


Figure 10

Full coverage unsupervised product based on 6 PC of the vegetation growth period (Mar-Oct 2009).

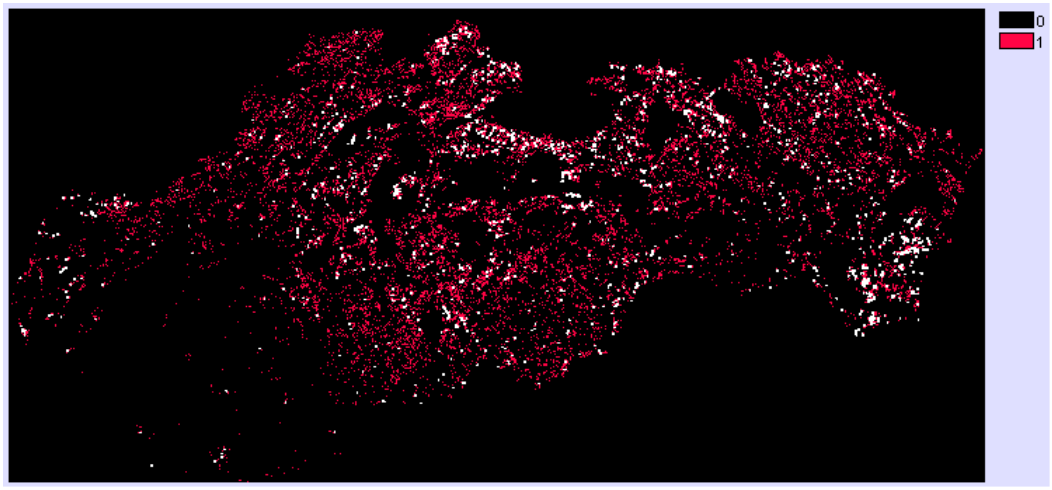


Figure11

Grassland distribution based on unsupervised approach using 6 PC of the vegetation growth period (Mar-Oct 2009).

Validation results of all products are presented in Table 2. Based on the accuracy and reliability of the products and expert visual check of the full coverage map of clustering results (Fig. 10) we found that the PCA based approach using 8 day vegetation growth period outperformed the other tested approaches. Estimated users accuracies are fairly low across all the products. However, we need to comment here that our reference data is not necessarily correctly classified either and this needs to be considered in the interpretation of the users accuracies. In future studies we need to create true non-grass validation pixels evenly distributed across different land cover classes in order to provide more confident results on user's accuracy. Interesting findings were revealed in validation of CLC grassland product. Here, only a 38 % producer's accuracy was achieved which is relatively low number. This could be partly explained by minimal mapping unit of CLC approach, however as we used a majority rule for the spatial resampling of the validation data, we think that the validation pixels should be homogenous enough for detection in the CLC product. As for the total coverage area, unsupervised products clearly overestimate and CLC and supervised approach clearly underestimate the total coverage of grasslands.

Table 2

Total coverage and accuracy estimation of different analyses.

Validation set			PCA based	PCA based	Fourier based	Supervised
		CLC	8d	16d		
Heterogenous	Area(ha)	223993	955100	922500	842500	433500
	Prod. acur.	38%	58%	51%	28%	16%
	Users acur.	58%	30%	30%	25%	38%
Homogenous	Prod. acur.	39%	68%	53%	38%	17%
	Users acur.	77.5	61%	60%	54%	53%

To conclude mapping of grassland land cover at regional scale in Slovakia based on solely multi-temporal classification seems to be quite difficult not only because of the coarse resolution of MODIS data but, as it is later analysed, also because of the fairly variable seasonal pattern of the Slovak grasslands. However we also documented that even the HR Landsat based CLC product in Slovak case is not very successful in grassland mapping.

8 Case study 2 - Grassland classification.

Here it is assumed that it is possible to first map the location and extent of grassland reliably, for which we explored the variability in the seasonal pattern of NDVI of grasslands across different study areas. We also explored possibilities of using temporal profiles of NDVI for extracting specific features of interests (e.g. cut management) or for phenology-based grassland type classification.

8.1 Study sites

Two grassland datasets were used to identify the location and extent of grasslands (Table 3).

Table 3
Data sets used for this case study.

ID	Area	Grassland types	NDVI data applied	Temporal coverage	Source	Scale
1	Slovakia	All grasslands on agricultural land (excluding alpine meadows)	16 day; 8 day	2001-2010	National dataset – land parcel information system	National
2	Hungarian lowlands	N2000 grasslands	8 day	2006-2010	EEA	Regional

Data set 1
We used homogenous pixel derived from Slovak national land parcel information system that mainly covered grasslands on agricultural land. Therefore natural grasslands (alpine meadows) and some pastures in forest landscape are not covered in this dataset. These sites were visually inspected on Google Earth. Totally 3758 pure pixels were included in this dataset (Fig. 12).

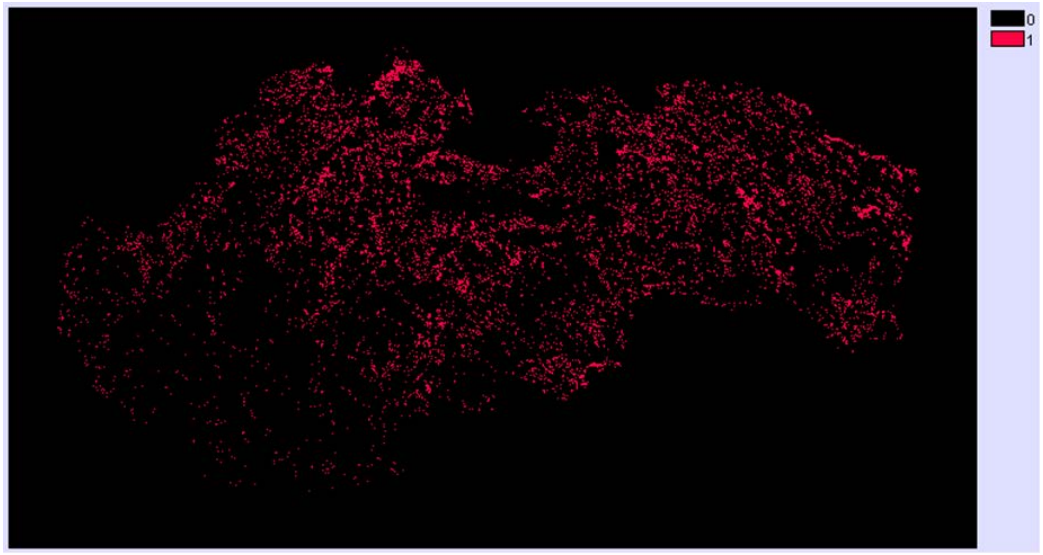


Figure 12
Grasslands in Slovakia and 3750 selected homogenous pixels (white).

Data set 2

Natura 2000 grassland sites were extracted from EEA dataset. Only those Natura 2000 sites that contained 80 % majority of grassland habitats were included in data set. From there, 2800 pixels were randomly selected for the analysis (Fig. 13).

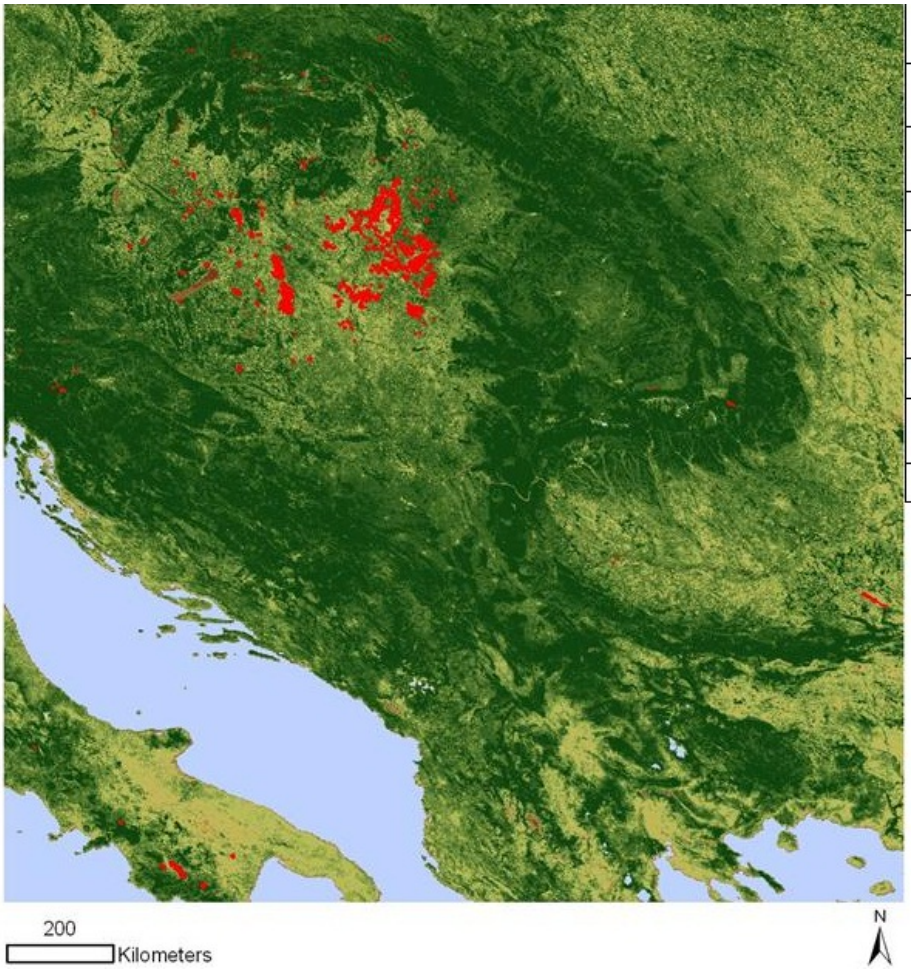


Figure 13
Hungrarian grasslands (data set 2).

8.2 NDVI data and processing

We used MOD13Q1 (16 day MVC NDVI with 250 m spatial resolution) and MOD09A1 (8 day surface reflectance at 250 m spatial resolution) downloaded from LP DAAC distribution centre for the area covered by MODIS tile grid h19v4. Usefulness index and quality assurance layers of MOD09A1 were checked in order to minimize negative effects of clouds, cloud shadows, aerosols, sun-sensor geometries and snow. Missing data were interpolated and smoothed using Savitsky – Golay filter within the TimeSat software (Eklundh and Jonson, 2004) in order to get complete NDVI time series from 2001 to 2010. As a first step we used PCA of the 2009 annual series in order to explored main season-driven variability in grasslands. The extracted components were later used to produce a broad classification of grasslands based on their seasonality. Finally we used several specific multi-temporal indexes (e.g. variability in peak season, spring negative anomaly, etc.) to test the potential for detecting grassland specific temporal features (e.g. cut management, flooding regime, etc).

8.3 Results

Using field observations and local expert knowledge we explored many NDVI time profiles that explained a wide variety of drivers of grassland status and functioning. Here we briefly described the following examples:

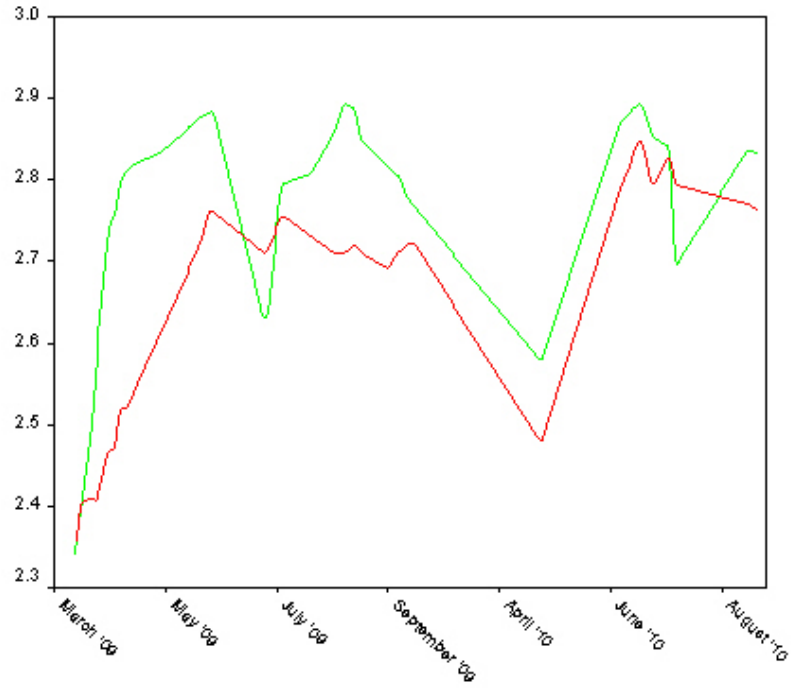


Figure 14

Cut (green) vs. uncut (red) meadows, both flooded in spring 2010 (original scale of NDVI -1/+1 rescaled to 1/3).

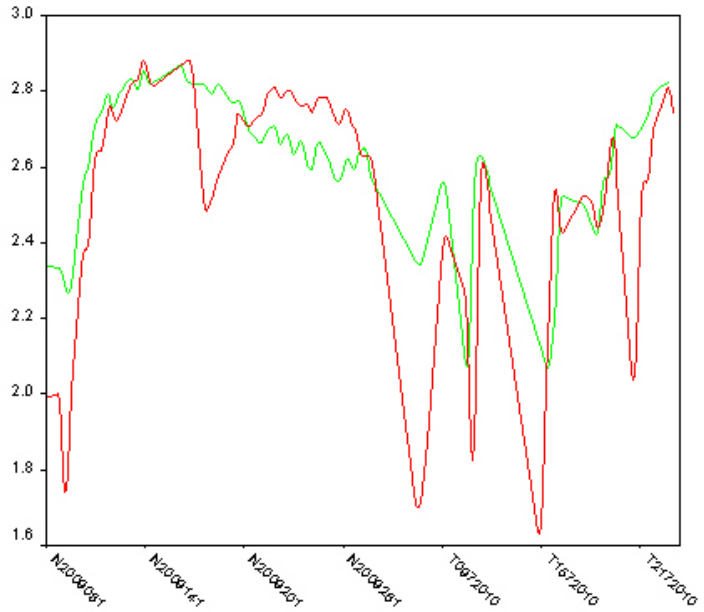


Figure 15

Cut meadows (red) and pasture (green) both flooded several times at different rate during spring 2010 (original scale of NDVI -1/+1 rescaled to 1/3).

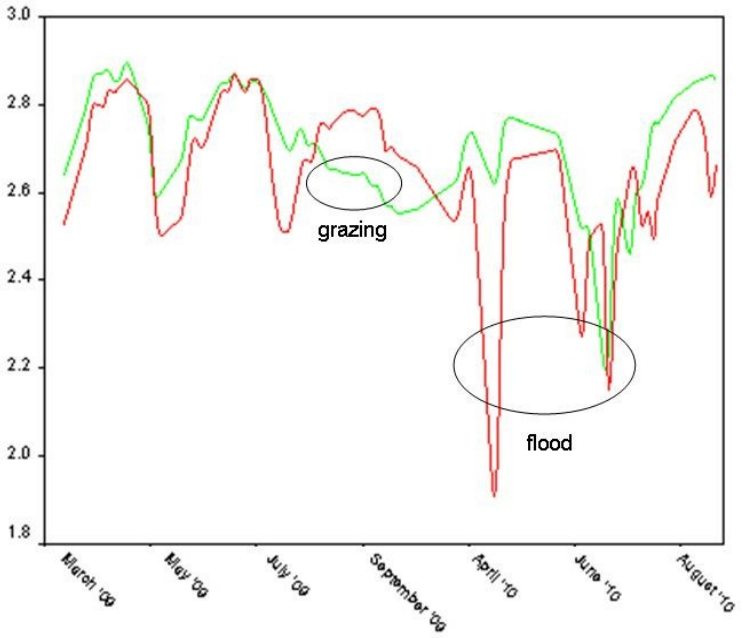


Figure 16

Twice cut meadows with grazing in autumn (green), both sites flooded once (green) and two times (red) in 2010 (original scale of NDVI -1/+1 rescaled to 1/3).

The PCA of 16 day 2009 NDVI composite of Slovak grasslands (data set 1) extracted 5 components with a total explained variance of 91 % (Fig. 17). This demonstrates that the seasonal NDVI pattern of grasslands varies substantially. Factor loadings show that it is mainly the different timing of peak season, cutting and different greenness in spring and autumn determine seasonal variability in Slovak grasslands. Classification of these 5 components using k-means clustering resulted in 5 main types of grasslands with specific temporal profile of NDVI (Fig. 18).

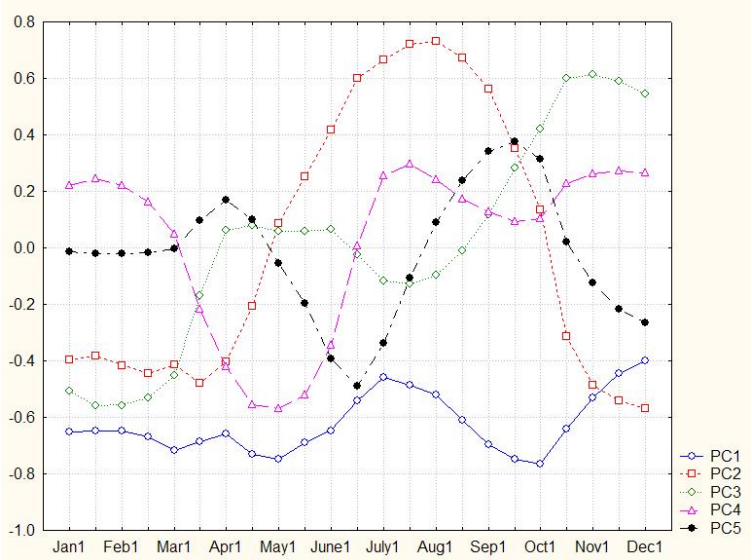


Figure 17

Factor loadings resulted from PCA of 16 day 2009 NDVI composite of Slovak grasslands.

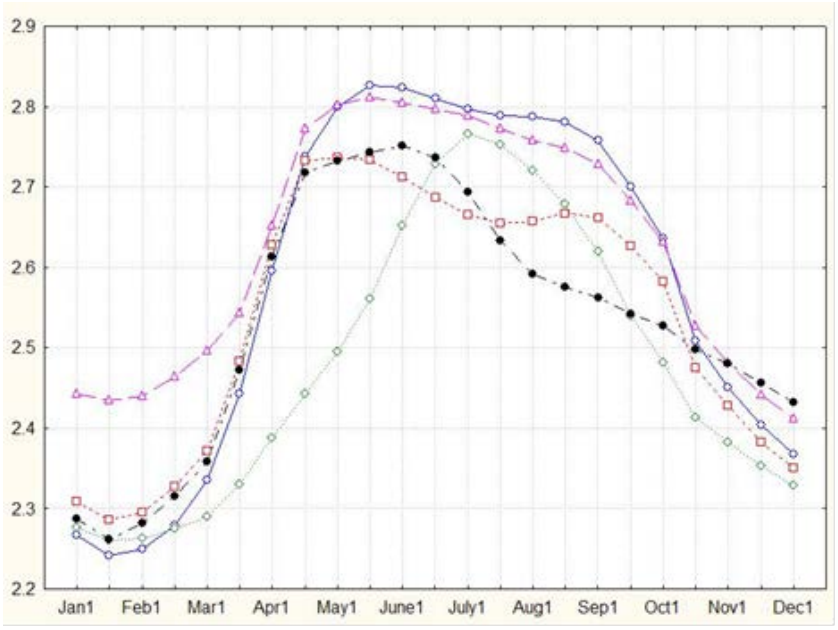


Figure 18

Average temporal profile of NDVI of resulted clusters (16 day NDVI composite of 2009) (original scale of NDVI -1/+1 rescaled to 1/3).

It is visible that mainly productivity (blue and purple clusters), different management (red and black clusters) described main differences in managed grasslands. A special case represents green cluster that can be attributed to mountainous grasslands typical with shorter vegetation season and delayed vegetation peak.

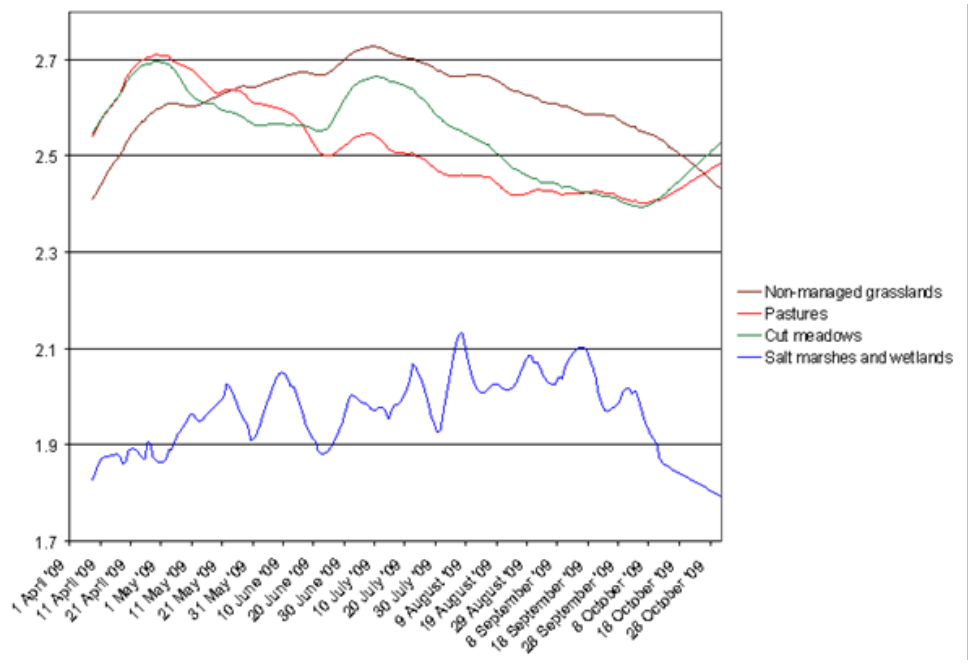


Figure 19

Classification of grasslands using 8 day NDVI composite of vegetation growth period in 2009 (original scale of NDVI -1/+1 rescaled to 1/3).

When we used the same approach with an increased temporal resolution (8day) of NDVI composite data from vegetation growth period (May-September) we obtained slightly different results (Fig. 19) with better distinction of flooded grasslands, pastures and cut grasslands. The same approach and data were used for classifying Hungarian grasslands in N2000 areas (Fig. 20). These grassland types exhibit distinct seasonal variability: wetland and salt marshes, cut meadows, pastures and non-managed natural grasslands.

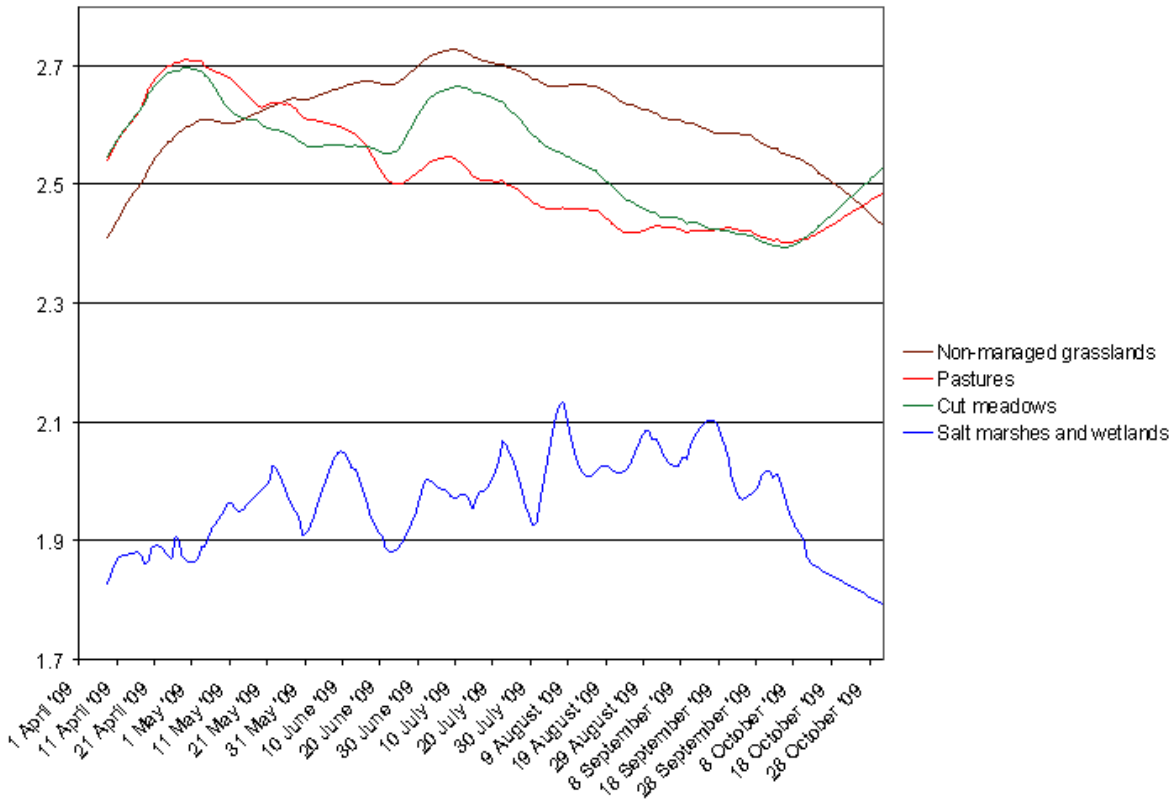


Figure 20

Classification of Hungarian grasslands in Natura 2000 areas using 8 day NDVI composite of vegetation growth period in 2009 (original scale of NDVI -1/+1 rescaled to 1/3).

9 Discussion and conclusion

Grassland seasonal pattern of NDVI can vary substantially reflecting not only the differences in vegetation type but also land use, management practices or site hydrology. This fact mainly limits mapping of grassland as a single land cover class especially when above factors are evenly present across the study area. This was demonstrated in Slovakia by the supervised image classification where one "grassland signature" was confused with many forests, and arable crops. In fact, different land cover types may have similar seasonal patterns of productivity (for example, some shrubs and the unmanaged grasslands); conversely, the same land-cover type may have different NDVI dynamics (for example, the intensive grasslands and extensive grasslands. On the other hand, in the unsupervised classification approach, the clusters with similar seasonality were merged together and misclassification was introduced by our attempt to attribute distinct seasonal information to respective land cover class. We demonstrated the limitation of using a supervised approach for a full coverage classification at broader scale which would require much more effort in training compared to the relatively easy labelling of an unsupervised product.

It seems that PCA approach were more suitable for exploring distinctive characteristics of NDVI temporal profile across different land cover classes. Although Fourier harmonics revealed some specific features (e.g. alpine meadows) and better reflect phenology we got more noisy classification results leading to more difficult labelling of the clusters. However as PCA is fully data dependent, and a better stratification (e.g. forested, agricultural landscape) could increase the explanatory power of PCA. NDVI temporal profiles of the majority of grasslands in Slovakia were similar to the ones of deciduous forest, which cause the main misclassifications. It seems that the exclusion of winter months from the analysis and using 8 day NDVI composites within the vegetation growth period would also increase the explanatory power of PCA. Wen et al. (2008) also reported that using of vegetation growth period was better than the whole year time series for multi-temporal analysis of Tibetian grassland classification. In fact, analysis of the whole annual period can be negatively influenced with winter periods and noise in data that is not primarily connected to vegetation seasonality resulting in producing of unmeaning full classes (Mucher et al., 2000). To conclude, the first reason of our poor success in grassland mapping relates to classification problem of land cover.

Different approaches for classification legend of land cover may lead to diverse final products resulting in uncertainties of information about status and changes of the landscape. Harmonization of the different approaches is very difficult because of diverse user community across the many disciplines and applications e.g. biodiversity conservation, hydrology, nature resource management, climatology, etc.). At regional scales broadly used CLC product faced many criticisms for inconsistent approach for classification and researches call for harmonized approaches for the European land cover classification (Mucher et al., 2000). As for the biodiversity observation initiative within EBONE project, usefulness of land cover maps for biodiversity assessment and monitoring needs to be further evaluated what will be done in near future. Going back to the classification problem, grasslands appeared to be relatively consistent in different classification legends as grassland or pastures, sometimes with different proportion of trees, based mainly on the physiognomy of grasslands. However, grasslands can vary a lot reflecting land use (meadow, pasture), management practice (one cut, two-cut, improved, semi-natural), hydrology regime (seasonally flood meadows), degradation (sparsely vegetated, overgrazed), abandon (overgrown). All these information are crucial as the indicators of their biodiversity values. Therefore, in their simple merging in one grassland land cover class, a poor knowledge would be extracted for biodiversity assessment. Hence, we suggest incorporating also information on temporal behaviour and seasonality into the classification system in order to reflect functioning of grasslands as proposed by Paruelo et al. (2001). A new approach within the EBONE project has been fostered

in order to make classification of land cover more suitable for consistent large scale habitat mapping and biodiversity assessment. Furthermore the classification approach would reflect current and future possibilities and limitation of remote sensing what gives this approach more chance to be applicable and operational. At the broader scale (European and Global), this process is not easy and many case studies need to be done across the scale in order to test its applicability. In this deliverable we reflect mainly regional and local scale which allows reasonable validation of the results. GHC are think to reflect several features of land surface (including phenology) what all should help to precisely define habitats for biodiversity assessment. Loveland et al. (2000) defines logic of land cover class definition based on their similar physiognomy and seasonality. Paruelo et al. (2001) using rather ecosystem function types than land cover classes and stated that these better reflect status and functioning of ecosystems allowing better detection of threats and trends. In fact, when landscape is monitored, (for example by using CLC1990 and CLC2000 products) the same area could be identified as CORINE class 231 after 10 years stated for no negative trend in grassland change. However analysis of the landscape based on the ecosystem functioning (e.g. temporal pattern of productivity) could reveal gradual degradation of the area (e.g. abandonment followed by successive overgrowing of meadows). This information is crucial for biodiversity observation and can serve as early warning system for change detection.

The second limitation of the multi-temporal image classification is coarse spatial resolution of the sensors. This is obvious mainly in diverse heterogeneous landscape what broadly demonstrated by many authors (Wessels et al. 2004).

Mücher et al. (2000) produced relatively accurate classification of grasslands in Netherland what could be caused by their large quasi homogenous status and extent across the landscape. On the other hand when grassland landscape is very diverse with patchy mosaic of small patches and with different land use and management practices, classification could became very difficult. Here we need to add that this is also problem of the HR based approaches. For example, accuracy of HR resolution based CLC grassland product in Slovakia was estimated only to 38%. This is could be due to the fact that many grasslands are incorporated within aggregated classes which can underestimate their total coverage estimates. For example 242 category of CLC represents almost 30 % of Slovakia indicating that we have a really poor knowledge about the grassland distribution. Unlike HR physiognomy based classification it is difficult to classified aggregated classes by multi-temporal approach although some successful examples exist (Geerken et al., 2009). We think that combined approach using HR satellite data for spatial arrangement with functional information derived from coarser satellite data could be a good approach for mapping diverse heterogeneous landscape. In our example we did not document a big difference in accuracies when we validated results with pure homogenous pixels, what suggests that both high seasonal grassland variability and coarse spatial resolution were proportionally responsible for low capability of multi-temporal approach for grassland mapping in Slovakia. However a relative importance of spatial resolution vs. within grassland temporal variability as main limits of multi-temporal approach for grassland mapping needs to be further explored by comparative studies in contrasting grassland landscapes (e.g. Hungary vs. Slovakia). To conclude full coverage grassland mapping together with other land cover classes by using only multi-temporal approach seems to be difficult. Combined approach with HR sensors is suggested in heterogeneous landscape. Furthermore, in order to deliver reasonable product of grassland mapping designated for biodiversity observation system, grassland classification (based on the NDVI temporal profile) needs to be reasonably defined reflecting information with added value for biodiversity monitoring and assessment. This classification should be regionally specific considering eco climatic conditions and different grassland types and land use practices.
Explanatory analysis of NDVI temporal profiles across the different regions identified specific features of grasslands in temporal space. In our second case study we demonstrated that when some knowledge about grassland occurrence exists, classification based on temporal NDVI profile brings reasonable information on grassland functioning. In general, productivity (which relate to amplitude of NDVI curve) and seasonality (range or variance within season) represent main distinctive characteristics of grasslands. Productivity and

seasonality derived from NDVI temporal curve are broadly used in ecosystem classification studies (Paruelo et al., 2001). NDVI composite data at finer temporal resolution (8 day) better distinguishes grasslands. In more detailed examples we documented that also specific timing of NDVI peak, rate of increase of NDVI in spring, bimodal shape of NDVI, or negative anomaly in spring can be used for distinguishing of pastures, mountainous meadows, intensive, extensive meadows, flooded meadows or salt marshes. However, more regional specific knowledge from grassland experts needs to be used in order to derive consistent grassland classification system for broad scale mapping and classification.

In our examples we tried to briefly demonstrate capabilities and limitations of multi-temporal approaches for grassland mapping and classification. It seems that for the mapping of grasslands in heterogeneous landscape, specific approaches need to be further explored for increasing of mapping capabilities of multi-temporal analyses. These approaches (including seasonal based classification of grasslands) need to be tested in near future across contrasted landscapes. Anyway, explanatory analysis and classification of grasslands using available sample data revealed that specific features of grasslands can be detected and reasonable classification made what proves that multi-temporal analysis should represent a valuable tool mainly in the assessment and monitoring component of the proposed biodiversity observation system.

10 References

Aragon, R. & Oesterheld, M. Linking Vegetation Heterogeneity and Functional Attributes of Temperate Grasslands Through Remote Sensing. Applied Vegetation Science 11[1], 117-130. 2008.

Coops, N. C., Waring, R. H., Wulder, M. A., Pidgeon, A. M. & Radeloff, V. C. Bird Diversity: a Predictable Function of Satellite-Derived Estimates of Seasonal Variation in Canopy Light Absorbance Across the United States. Journal of Biogeography 36[5], 905-918. 2009.

De Fries, R. S., Hansen, M., Townshend, J. R. G. & Sohlberg, R. Global Land Cover Classifications at 8 Km Spatial Resolution: the Use of Training Data Derived From Landsat Imagery in Decision Tree Classifiers. International Journal of Remote Sensing 19[16], 3141-3168. 1998.

Defries, R., Hansen, M. & Townshend, J. Global Discrimination of Land Cover Types From Metrics Derived From Avhrr Pathfinder Data. Remote Sensing of Environment 54[3], 209-222. 1995.

Fontana, F., Rixen, C., Jonas, T., Aberegg, G. & Wunderle, S. Alpine Grassland Phenology as Seen in Avhrr, Vegetation, and Modis Ndvi Time Series - a Comparison With in Situ Measurements. Sensors 8[4], 2833-2853. 2008.

Franklin, S. E. & Wulder, M. A. Remote Sensing Methods in Medium Spatial Resolution Satellite Data Land Cover Classification of Large Areas. Progress in Physical Geography 26[2], 173-205. 2002.

Gao, Y., Mas, J. F. & Navarrete, A. The Improvement of an Object-Oriented Classification Using Multi-Temporal Modis Evi Satellite Data. International Journal of Digital Earth 2[3], 219-236. 2009.

Geerken, R. A. An Algorithm to Classify and Monitor Seasonal Variations in Vegetation Phenologies and Their Inter-Annual Change. Isprs Journal of Photogrammetry and Remote Sensing 64[4], 422-431. 2009.

Giri, C. & Jenkins, C. Land Cover Mapping of Greater Mesoamerica Using Modis Data. Canadian Journal of Remote Sensing 31[4], 274-282. 2005.

Hall-Beyer, M. Comparison of Single-Year and Multiyear Ndvi Time Series Principal Components in Cold Temperate Biomes. Ieee Transactions on Geoscience and Remote Sensing 41[11], 2568-2574. 2003.

Hill, M. J., Vickery, P. J., Furnival, E. P. & Donald, G. E. Pasture Land Cover in Eastern Australia From Noaa-Avhrr Ndvi and Classified Landsat Tm. Remote Sensing of Environment 67[1], 32-50. 1999.

Huang, S., Siegert, F., Goldammer, J. G. & Sukhinin, A. I. Satellite-Derived 2003 Wildfires in Southern Siberia and Their Potential Influence on Carbon Sequestration. International Journal of Remote Sensing 30[6], 1479-1492. 2009.

Huang, S. & Siegert, F. Land Cover Classification Optimized to Detect Areas at Risk of Desertification in North China Based on Spot Vegetation Imagery. Journal of Arid Environments 67[2], 308-327. 2006.

Jakubauskas, M. E., Legates, D. R. & Kastens, J. H. Harmonic Analysis of Time-Series Avhrr Ndvi Data. Photogrammetric Engineering and Remote Sensing 67[4], 461-470. 2001.

John, R., Chen, J. Q., Lu, N., Guo, K., Liang, C. Z., Wei, Y. F., Noormets, A., Ma, K. P. & Han, X. G. Predicting Plant Diversity Based on Remote Sensing Products in the Semi-Arid Region of Inner Mongolia. Remote Sensing of Environment 112[5], 2018-2032. 2008.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L. & Merchant, J. W. Development of a Global Land Cover Characteristics Database and Igbp Discover From 1 Km Avhrr Data. International Journal of Remote Sensing 21[6-7], 1303-1330. 2000.

Matsuoka, M., Hayasaka, T., Fukushima, Y. & Honda, Y. Land Cover Classification Over Yellow River Basin Using Satellite Data. Igarss 2004: Ieee International Geoscience and Remote Sensing Symposium Proceedings, Vols 1-7 , 231-234. 2004.

Moody, A. & Johnson, D. M. Land-Surface Phenologies From Avhrr Using the Discrete Fourier Transform. Remote Sensing of Environment 75[3], 305-323. 2001.

Mücher, C. A., Steinnocher, K. T., Kressler, F. P. and Heunks, C. (2000) 'Land cover characterization and change detection for environmental monitoring of pan-Europe', International Journal of Remote Sensing, 21:6, 1159 — 1181

Mücher, C.A., 2000 (eds.). PELCOM project. Development of a consistent methodology to derive land cover information on a European scale from remote sensing for environmental modelling. Final report. 236 pp.

Paruelo, J. M., Jobbagy, E. G. & Sala, O. E. Current Distribution of Ecosystem Functional Types in Temperate South America. Ecosystems 4[7], 683-698. 2001.

Paruelo, J. M., Jobbagy, E. G., Sala, O. E., Lauenroth, W. K. & Burke, I. C. Functional and Structural Convergence of Temperate Grassland and Shrubland Ecosystems. Ecological Applications 8[1], 194-206. 1998.

Paruelo, J. M. & Lauenroth, W. K. Regional Patterns of Normalized Difference Vegetation Index in North-American Shrublands and Grasslands. Ecology 76[6], 1888-1898. 1995.

Ratana, P., Huete, A. R. & Ferreira, L. Analysis of Cerrado Physiognomies and Conversion in the Modis Seasonal-Temporal Domain. Earth Interactions 9. 2005.

Reed, B. C., Brown, J. F., Vanderzee, D., Loveland, T. R., Merchant, J. W. & Ohlen, D. O. Measuring Phenological Variability From Satellite Imagery. Journal of Vegetation Science 5[5], 703-714. 1994.

Samson, S. A. 2 Indexes to Characterize Temporal Patterns in the Spectral Response of Vegetation. Photogrammetric Engineering and Remote Sensing 59[4], 511-517. 1993.

Sedano, F., Gong, P. & Ferrao, M. Land Cover Assessment With Modis Imagery in Southern African Miombo Ecosystems. Remote Sensing of Environment 98[4], 429-441. 2005.

Stone, T. A., Schlesinger, P., Houghton, R. A. & Woodwell, G. M. A Map of the Vegetation of South-America Based on Satellite Imagery. Photogrammetric Engineering and Remote Sensing 60[5], 541-551. 1994.

Tao, F., Yokozawa, M., Zhang, Z., Hayashi, Y. & Ishigooka, Y. Land Surface Phenology Dynamics and Climate Variations in the North East China Transect (Nect), 1982-2000. International Journal of Remote Sensing 29[19], 5461-5478. 2008.

Wen, Q. K., Liu, S., Zhang, Z. X. & Qiao, W. Classification of Grassland Types in Tibet by Modis Time-Series Images. 2008 International Workshop on Earth Observation and Remote Sensing Applications , 249-254. 2008.

Wessels, K., Steenkamp, K., Von Maltitz, G. & Archibald, S. Remotely Sensed Vegetation Phenology for Describing and Predicting the Biomes of South Africa. Applied Vegetation Science 14[1], 49-66. 2011.

PART III: EO TIME-SERIES ANALYSIS TO IDENTIFY ANNEX 1 HABITATS

Gerard Hazeu, Gerbert Roerink, Sander Mücher

ALTERRA, Wageningen UR, the Netherlands

11 Introduction

The objective here is to enhance the identification of the spatial distribution of European Annex I habitats, read Natura 2000 habitats, based on their phenology, compared to the prior spatial modelling of habitats based on satellite derived land cover information in combination with environmental data sets and species distribution data (European Commission, 2007; Mücher et al., 2004; Mücher et al., 2009; Mücher, 2009). For this reason NDVI-time series have been analysed that could be processed for the whole of Europe. There is a demand for a high temporal resolution together with a spectral resolution that allows the calculation of the Normalized Difference Vegetation Index (NDVI). The best suitable sensor, concerning a high temporal resolution and adequate spectral and spatial resolution, is MERIS (300m) or MODIS (250m). Since the latter is easy and freely downloadable, we used MODIS satellite data for our purpose. MODIS has a daily revisit time with a spatial resolution of 250 meter.

11.1 MODIS satellite data

The analysis is based upon multi-annual time series of MODIS images. MODIS has the highest resolution (250 m) of freely available remote sensing images with a daily revisit time. MODIS has an additional advantage, since it also provides composite images from the 16-days-maximum-NDVI product (MOD13Q1). More precisely, this MODIS product is the NDVI 16-Day L3 Global 250m. We used the version 5 MODIS product. Version-5 MODIS / Terra NDVI products have "Validated Stage 2", meaning that accuracy has been assessed over a widely distributed set of locations and time periods via several ground-truth and validation efforts. Although there may be later improved versions, these data are ready for use in scientific publications. The 16-days-maximum-NDVI composite images have the advantage to exclude cloud affected data. Only when all 16 days in a composite are cloudy, the image is cloud affected. Unfortunately this is a regular occurrence in The Netherlands, so screening is necessary to filter out the cloud affected data. Additionally, the 16-day composites use a mask to filter out the large surfaces of open water, like sea and lakes. This mask is known to be inconsistent, which makes the results near the shoreline (within 15 km) unreliable. In spite of the shortcomings of the composite images, the 16-days-maximum-NDVI product also saves a large amount of time in pre-processing and provides annually 23 images which is sufficient to be used in time series analyses. Therefore the NDVI composite product has been downloaded and used during this research.

11.2 HANTS algorithm

The seasonal cycle of the NDVI can be approximated by a limited number of frequency components derived from a Fourier analyses. This principle is implemented in the HANTS algorithm (Harmonic Analysis of NDVI Time-Series) (Roerink et al., 2000) which employs an iterative routine to filter out poor NDVI estimates due to cloud cover or other disturbances from the NDVI cycle. In the current analyses only the zero (mean), first (annual cycle) and second (half-yearly cycle) frequency components of the Fourier analysis were used to describe the NDVI cycle.

The HANTS algorithm was originally developed by Wout Verhoef from NLR (Netherlands Space Laboratorium) in the Netherlands. The idea behind the algorithm was to have a fast method for smoothing and reconstructing NDVI time-series at continental scales. For various purposes it was desirable to have a version of HANTS which was easy to use but with similar functionality. This has led to the implementation of HANTS using the remote sensing software package IDL-ENVI.

The basic concept behind the algorithm is that the vegetation development as indicated by the NDVI has a strong seasonal effect in most parts of the world (apart from the tropics) which can be described using a series of low frequency sine functions with different phases, frequencies and amplitudes. Cloud cover and other disturbing effects are usually more or less randomly occuring “spikes” in the NDVI time-series and can be considered as high frequency “noise”. The working of the HANTS algorithm is therefore based on a Fourier analysis.

In contrast to the standard fast Fourier transform (FFT), the HANTS algorithm works in an iterative manner (see next figure). The algorithm starts in the upper left block with the raw NDVI time series. These are used as input in the FFT and the relevant frequencies (usually mean, annual and half-year signal) are selected from the fourier spectrum. The inverse FFT (iFFT) then transforms the spectrum back into a filtered NDVI time-series.

Next, a comparison is made between the filtered NDVI time-series and the original NDVI time-series. The difference is calculated between the filtered and the original NDVI time-series. Any points in the original NDVI time-series that are below a user-defined treshold are considered ‘cloudy’ and are replaced with the value of the filtered NDVI time-series. However, by replacing values in the NDVI time-series the average of the entire profile changes (becomes larger). Therefore a next iteration is needed and again the NDVI time-series is searched for possible cloud contaminated NDVI observations. This process continues until no new points are being found.

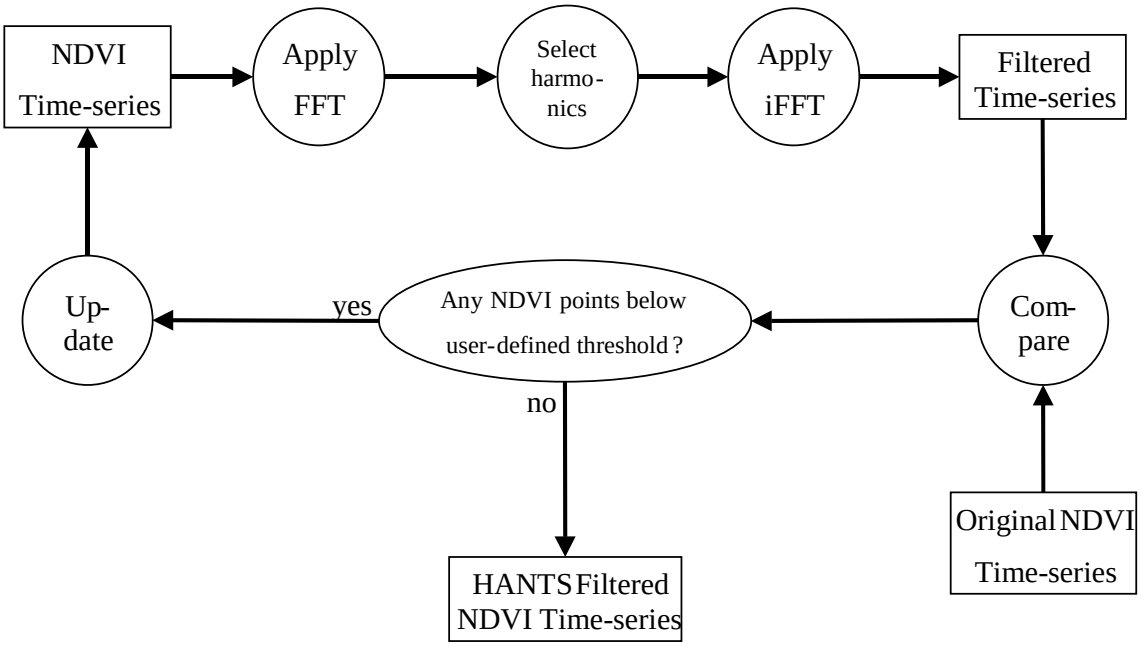


Figure 1

Iterative workflow of the HANTS algorithm. See text for explanation of the algorithm and acronyms.

The use of time series analysis on remote sensing images offers great opportunities for year-to-year monitoring of the earth surface. However, two serious drawbacks have to be dealt with, namely (i) time series analysis on remote sensing data produces huge amounts of data that needs to be processed and analysed and (ii) the presence of erroneous data, like cloud affected or missing pixels.

The selected HANTS (Harmonic ANalysis of Time Series) algorithm (Roerink et al., 2000) deals with the latter mentioned drawbacks pretty well, and has three major benefits (Roerink & Danes, 2010):

- Large data reduction. The method allows reducing the amount of data by a factor of at least 5 without loss of information. In the example in the Figure 2 the individual NDVI values from 36 decades (10 day periods) are reduced into 5 HANTS components (3 amplitude and 2 phase values).
- Exclusion of deviant data exclusion. The method is able to exclude cloud affected and missing pixels in the analysis
- Vegetation dynamics. Objective and quantitative characterisation of plant phenology. The time series of NDVI remote sensing images are described by the Fourier components (amplitude and phase)

Because of its benefits, HANTS has been used successfully in various applications, such as cloud screening, removal and replacement (Roerink et al., 2000), land cover classification (Zhang et al., 2008), plant phenology characterisation (White et al., 2009) and climate variability assessment (Roerink et al., 2003).

HANTS is a least squared curve fitting procedure, based on harmonic components (cosine-functions), and considers only the most significant frequencies expected to be present in the time profiles. In an iterative process, input data values that have a large positive or negative deviation from the current estimated curve are excluded from the procedure. This process is repeated until the maximum error is acceptable or the number of remaining values becomes too small. The entire curve fitting procedure is controlled by 5 parameters, which have to be set at the beginning of each HANTS run:

- Number Of Frequencies (NOF). A curve is described by mean of its average (frequency zero) and a number of cosine functions with different frequencies. By this control parameter the user defines how many cosine functions are used and what the frequency (time period) of each cosine function is. This results in 2 x NOF - 1 output parameters (an amplitude and phase value for each frequency), where NOF includes a frequency zero (time series average), which has no phase.
- High/Low Suppression Flag (SF). This flag indicates whether high or low values (outliers) should be rejected during curve fitting.
- Invalid Data Rejection Threshold (IDRT). In some cases one might know that digital numbers below or above a certain threshold should be considered invalid.
- Fit Error Tolerance (FET). During curve fitting the absolute difference in the Hi/Lo direction of the remaining (i.e. not rejected) data values with respect to the current curve is determined after each iteration. The iteration stops when the difference of all remaining values becomes smaller than the FET. The FET value should not be set too low, as otherwise the fit might be based on too few values, which gives unreliable results.
- Degree of OverDeterminedness (DOD). The number of valid observations must always be greater than or equal to the number of parameters that describe the curve (2 x NOF – 1). In order to get a more reliable fit the user can decide to use more data values than the necessary minimum. The minimum number of extra data values, which have to be used in the ultimate fit, is given by the DOD value.

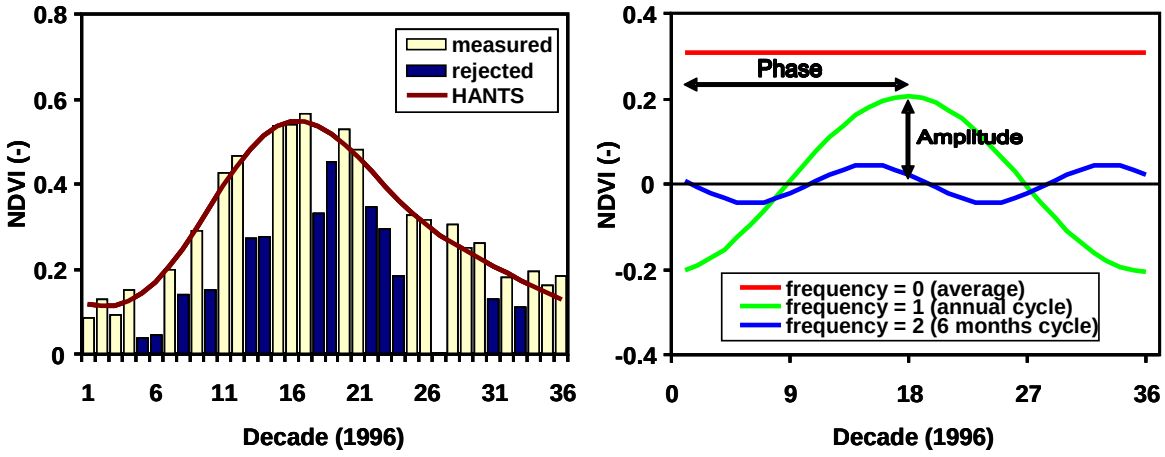


Figure 2

Visualisation of the HANTS algorithm (after Roeirnk et al., 2000).

The basic principle how HANTS works is visualised in Figure 2, where an original and HANTS reconstructed NDVI time series for a pixel of arable farming in Northern France are shown. The number of frequencies used by HANTS was set at 3: the average NDVI (frequency = 0), the yearly amplitude (frequency = 1) and the amplitude of 6 months (frequency = 2). The iteration stopped when 14 out of the 36 original NDVI values were rejected, i.e. classified as cloud affected data points. In this case, the remaining 22 values are allowed to have a maximum negative deviation from the curve of 0.05 NDVI units (=FET). The right graphs in Figure 2 shows the harmonic components of the 3 different frequencies, from which the cloud-free profile is reconstructed. Frequency zero (straight line) is represented only by amplitude and no off-set, while the other remaining frequencies (cosine functions) are defined by an amplitude and a phase value.

11.3 Example HANTS Results

HANTS has been applied now on the MODIS 16-day maximum NDVI composites for the years 2001 & 2006. The curve fitting process is controlled by 5 control parameters, which have to be set at the beginning of each HANTS run (Roerink and Danes, 2010). In the framework of this study the control parameters are set at:

- NOF = 3, where ;
  - Frequency 0: NDVI average
  - Frequency 1: Phase and amplitude of the annual cosine function
  - Frequency 2: Phase and amplitude of the six monthly cosine function
- SF = Low; which means only low values (outliers) should be rejected during curve fitting, as they correspond to cloud affected data.
- IDRT = 0; as missing data in the original NDVI composites have a value 0.
- FET = 0.05; as the NDVI values ranges from 0 to 1, a FET of 0.05 means that 5% deviation from the fitted curve is tolerated.
- DOD = 8; the maximum number of data points that may be rejected is 10 out of 23 available values.

Since this HANTS operation includes three frequencies, the output will consist out of five images:

Image 1. The amplitude of frequency 0, which is the average NDVI value, or the average amount of vegetation over a year (frequency 0 has no phase as it is a constant value over the year, i.e. has no starting point)

Image 2. The amplitude of frequency 1, which reflects the seasonal vegetation difference between summer and winter

Image 3. The phase of frequency 1, which describes when exactly the peak vegetation takes place.

Image 4. The amplitude of frequency 2, which is the amplitude of the 6 months cosine function. As in most cases vegetation dynamics have only one growing season during a year, the six months amplitude has no physical meaning like the annual amplitude, but is necessary for a smooth curve fitting procedure.

Image 5. The phase of frequency 2 which is phase of the 6 months cosine function; like the amplitude of the six months cosine function its physical meaning is limited, but is necessary for a smooth curve fitting procedure.

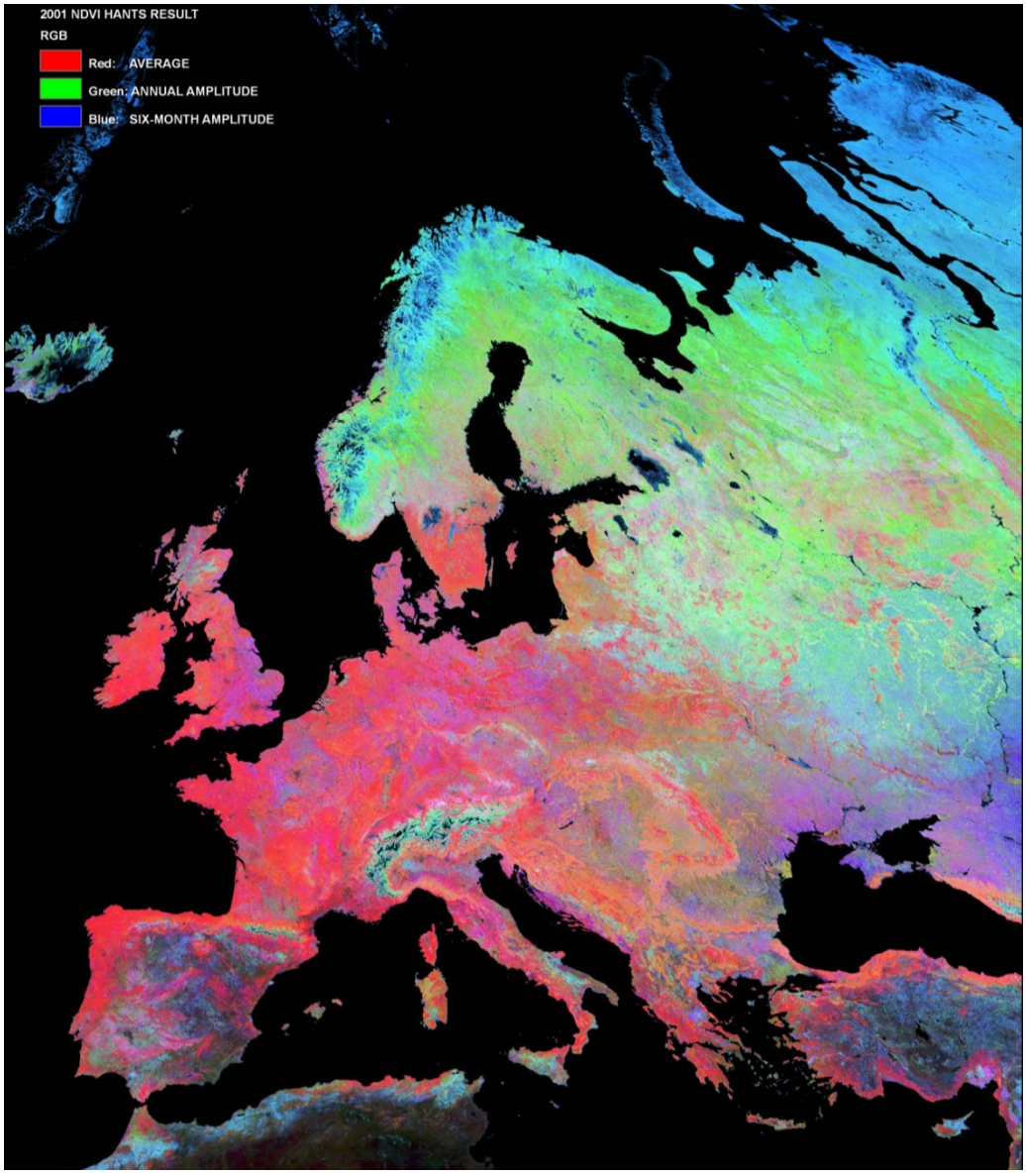


Figure 3

False colour European composite image of the HANTS results of the year 2001, where red colour indicates the NDVI average, green colour indicates the amplitude of annual frequency, and the blue colour indicates the amplitude of the six months frequency contacts: matthijs.danes@wur.nl or gerbert.roerink@wur.nl).

11.4 Classifications

Many land cover classes have a typical phenological cycle, as shown in the figure below. This is only partly true for habitat classes that are moreover most of the time extremely fragmented in their spatial distribution over Europe. The first classification experiments using MODIS time series at a 1km spatial resolution failed completely. In other words, these data are too coarse to spatially identify different habitats as described in the former sections.

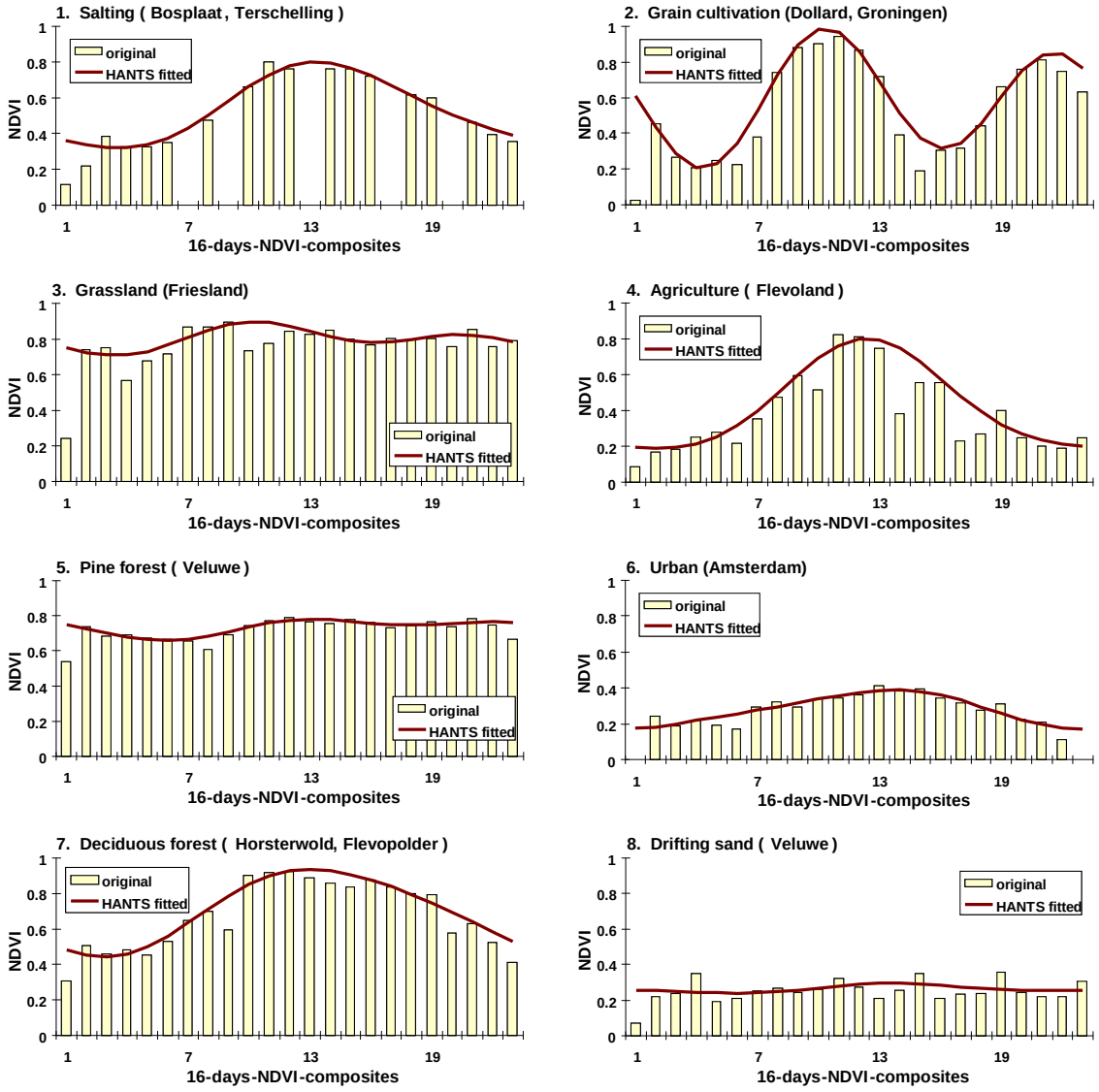


Figure 4

Yearly profiles of original NDVI values and the HANTS fitted curve of typical land cover types in the Netherlands.

The time-series analysis experiment performed here is based on MODIS 250 meter NDVI time series and aimed to explore the value of the HANTS results in helping to better define areas in which specific Natura 2000 habitats may occur.

12 Materials and methodology

12.1 Materials

The material used in these experiments are 1) HANTS results based on MODIS 250m 16 days-maximum-NDVI time series for the year 2006; 2) Environmental zones (EnZ) (Metzger et al., 2005); 3) CORINE Land Cover 2006; 4) habitat information and 5) Natura2000 database.

Ad.1. The MODIS 250m 16 days-maximum-NDVI time series for the year 2006 were processed by the HANTS algorithm i.e. a time series analysis of vegetation development by Fourrier analysis (Roerink et al., 2000). The fitting process was controlled by 5 control parameters (see section 7.4). The amplitude and phase of the vegetation development is described by a cosines function. The amplitude and phase are recalculated into xy-coordinates which were used for the classification.

Ad.2. The experiment focussed on the Alpine South (ALS) and Continental (CON) Environmental zones as for these zones sufficient in situ habitat measurements were available.

Ad.3. The CORINE Land Cover (CLC 2006) dataset was used as ancillary data to define suitable training sets for the classification of the different habitats selected.

Ad.4. In-situ data (vegetation relevés) from countries (see section 3), results from habitat distribution modelling (section 5 and Mücher et al., 2009) were used to select training sets for classifying a specific habitat on the HANTS results. The classification focussed on H4060 '*Alpine and Boreal heaths*' and H9150 *'Medio-European limestone beech forest of the Cephalathero-Fagion*'.

H4060 '*Alpine and Boreal heaths*' are small, dwarf or prostrate shrub formations of the alpine and sub-alpine zones of the mountains of Eurasia dominated by ericaceous species, *Dryas octopetala*, dwarf junipers, brooms or greenweeds (European Commission, 2007).

H9150 *'Medio-European limestone beech forest of the Cephalathero-Fagion'* are Xero-thermophile Fagus sylvatica forests developed on calcareous, often superficial, soils, usually of steep slopes, of the medio-European and Atlantic domaines of Western Europe and of central and northern Central Europe, with a generally abundant herb and shrub undergrowth, characterized by sedges (*Carex digitata, Carex flacca, Carex montana, Carex alba*), grasses (*Sesleria albicans, Brachypodium pinnatum*), orchids (*Cephalanthera spp., Neottia nidus-avis, Epipactis leptochila, Epipactis microphylla*) and thermophile species, transgressive of the Quercetalia pubescentipetraeae. The bush-layer includes several calcicolous species (*Ligustrum vulgare*, *Berberis vulgaris*) and *Buxus sempervirens* can dominate (European Commission, 2007).

Ad.5. The Natura2000 database was used in the assessment of the classification results.

12.2 Methodology

So far the identification of Annex I habit types has been based on predictive habitat modelling using remotely sensed land cover information in combination with environmental data sets, such as climatic zonations, altitude and soil distribution, and if available information of the specific distribution of individual species (European

Commission, 2007; Mücher, et al., 2004. Mücher et al., 2009; Mücher, 2009). The use of satellite derived times series that can describe the phenology of the vegetation could be an extra valuable tool to discriminate specific Annex I habitat types.

Based on the HANTS analysis of the MODIS time series with a 250 meter spatial resolution, classifications were made separately for each environmental zone as the vegetation development between zones is very different due to biophysical conditions. The HANTS results were classified with the Maximum Likelihood parametric rule based on a signature file. The signature file contained two groups of training sets: i) general land cover signatures and ii) specific signatures related to the selected habitats. For the specific habitat-training sites the in situ data (vegetation relevés) were used in combination with the probability maps resulting from the habitat distribution modelling exercise (section 5) and the CORINE Land Cover database. The signatures were created by region growing in which seeding properties were adapted and more restricted for the specific habitat signatures (e.g. spectral euclidean distance was smaller/ narrow band width). Signatures for the habitats were compared and outliers were removed. The habitat classification was repeated several times with different number of training sets for a specific habitat. Finally, the ultimate classification was based on a limited number of signatures for that specific habitat (preferably the signature should be a pure end member, but this is difficult to realise at a spatial resolution of 250 by 250 meters). Mainly training sites were used for areas that have a high probability that the specific habitat is present, located on the site where an in-situ measurement indicated the presence of the habitat and a 'logical' land cover type in the CORINE database.

13 Results and assessment

13.1 Results

The distribution of the habitats H4060 and H9150 were classified on basis of the HANTS results for the Alpine South and Continental Environmental zone. Figure 5 and 6 show the results for the Alpine South environmental zone for H4060 and H9150. Comparing the classification results with the results from the habitat distribution modelling (probabilities) the habitat is more widely distributed and the presence of H9150 is largely overestimated. The contrary is the case for H4060. See also section 7.5.4. 'Assessment of results'.

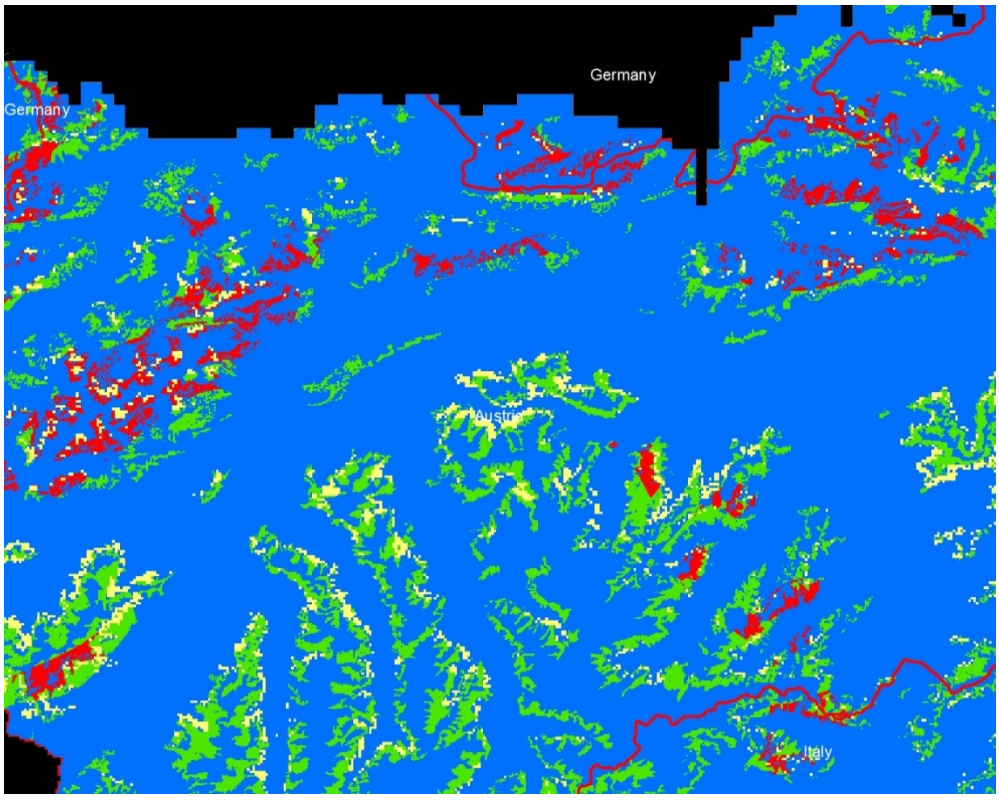


Figure 5

The Alpine environmental zone (blue) with in red (high) and green (low) probabilities that H4060 is present in this part of Austria. The yellow colour indicates the result of the HANTS classification. The classification result shows a limited presence of H4060 mainly distributed over areas with high and low probabilities.

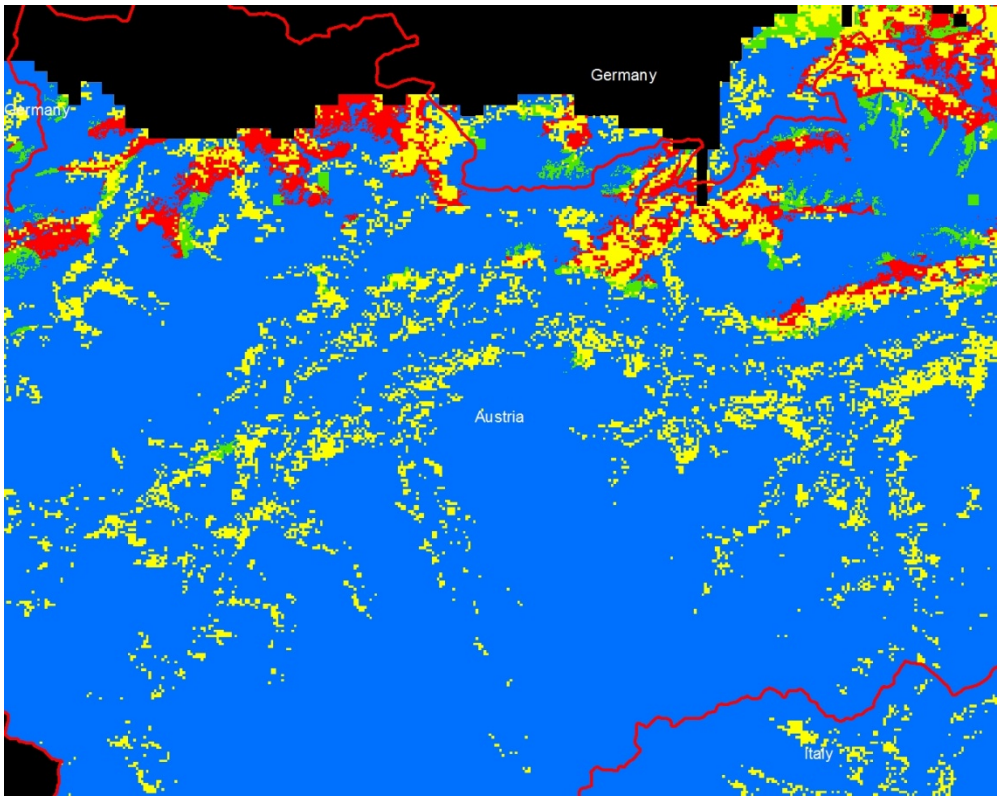


Figure 6

The Alpine environmental zone (blue) with in red (high) and green (low) probabilities that H9150 is present in this part of Austria. The yellow colour indicates the result of the HANTS classification. The classification result shows an overestimation of H9150 that is distributed all over the environmental zone. The distribution is not restricted to areas with high and low probabilities.

The results for the Continental zone are limited to the H9150 habitat, since there were no in situ H4060 habitat data available. The distribution of H9150 in the Continental zone shows an similar/intermediate position as in the Alpine south environmental zone: a large overestimation of H9150 and a scattered distribution all over the environmental zone (see Figure 7). The overestimation for H9150 in both environmental zones is, when comparing with CLC, partly due to the classification of mixed and coniferous forest as H9150.

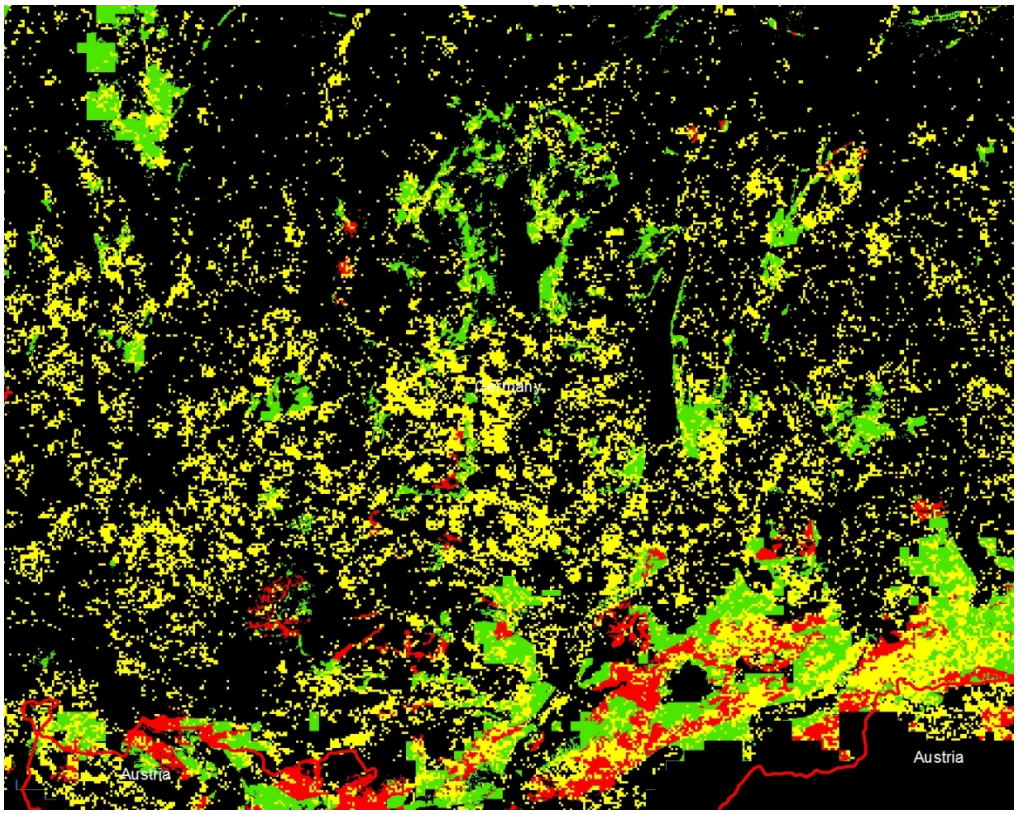


Figure 7

The Continental environmental zone (black) with in red (high) and green (low) the probabilities that H9150 is present in this part of Germany. The yellow colour indicates the result of the HANTS classification. The classification result shows an overestimation of H9150 that is distributed all over the environmental zone. The distribution is not restricted to areas with high and low probabilities.

13.2 Assessment of results

Table 1 indicates the number of pixels (as percentage of the total number of pixels) that agree between the different probability classes and the classification results.

For all three combinations (H4060-environmental zone ALS, H9150-environmental zone ALS and H9150-environmental zone CON) only small proportions of pixels are classified as H4060/H9150 and having low or high probability distributions. This is to be expected as the majority of pixels fall within the classes having no probability or within the classes not having that specific habitat. The table presents a possible overestimation of pixels classified as H9150 as the number of pixels falling in the probability class 0 is higher than the total number of pixels classified as not having that specific habitat (e.g. 86.0% versus 81.0% for the H9150/ALS combination). For H4060 this comparison may conclude that there is an underestimation of pixels classified as H4060 (93.0% versus 99.3%).

In case of the H4060/ALS combination a relatively high percentage of "correctly" classified pixels of 92.8% (92.5 + 0.2 + 0.1%) compared to the other two combinations with 75.4% (71.2 + 3.3 + 0.9%) and 86.3% (85.6 + 0.5 + 0.2%) can be found. The H4060 classification agrees better with the probability mapping of the habitat distribution modelling results than for the other two combinations.

Table 1

A cross table of probability versus the classification results for the combinations H4060/ALS, H9150/ ALS and H9150/CON (areas as percentages of total number of pixels). Classification values 0 and 1 mean that the area is not classified respectively is classified having that specific habitat. Probability values 0, 2 and 3 mean that the specific habitat is probably not, with low respectively high probability present in the area based on the modelling in chapter 5.

H4060/ALS		Probability			
		0	2	3Total	
Classification	0	92.5%	5.7%	1.1%	99.3%
	1	0.5%	0.2%	0.1%	0.7%
		93.0%	5.9%	1.1%	100.0%
H9150/ALS		Probability			
		0	2	3Total	
Classification	0	71.2%	8.3%	1.5%	81.0%
	1	14.8%	3.3%	0.9%	19.0%
		86.0%	11.6%	2.4%	100.0%
H9150/CON		Probability			
		0	2	3Total	
Classification	0	85.6%	4.1%	1.2%	90.9%
	1	8.4%	0.5%	0.2%	9.1%
		94.0%	4.7%	1.4%	100.0%

Table 2

Number of Natura2000 sites with specific habitat for the specific habitat-environmental zone combinations that contain MODIS pixels classified as having that specific habitat. The Natura2000 sites are grouped according to the amount of pixels falling within the Natura2000 site.

				Number of pixel counts			
Habitat	EnZ	Total sites with specific habitat	Number of site within EnZ	1-5	5-25	>25	Total
4060	Als	157	91	13	13	3	29
9150	Als	47	10	2	1	-	3
9150	Con	47	9	1	1	7	9

Of the 91 Natura2000 sites with H4060 falling in the environmental zone ALS, 29 sites match the H4060 classification (see Table 2). Only 3 sites contain more than 25 MODIS pixels classified as having H4060. The H9150/CON combination shows that all 9 Natura2000 sites with H9150 have also pixels classified as H9150. The majority of these sites even contain more than 25 pixels with that specific habitat. One important remark to be made is that the number of pixels classified as having a specific habitat that fall within a Natura2000 site

depends on the area of the Natura2000 site. So it is difficult to make conclusions based on the number of pixels falling within a site without weighting for the surface areas of the Natura2000 sites.

Table 3

All Natura2000 sites for a specific habitat-environmental zone combination and the number of sites that contain MODIS pixels classified as having that specific habitat. The Natura2000 sites are grouped according to the amount of pixels falling within the Natura2000 site.

		Number of pixel counts			
Habitat/EnZ	Number of sites within EnZ	1-5	5-25	>25	Total
H4060/ALS	1367	72	60	27	159
H9150/ALS	1367	222	283	531	1036
H9150/CON	6375	1623	1291	1084	3998

The majority of all Natura2000 sites within the ALS and CON environmental zones contain pixels that are classified as having H9150 MODIS pixels (see Table 3). This suggest an overestimation of the H9150 classification in both environmental zones.

Table 4

Error matrices between Natura2000 sites and the classification of HANTS phenology product. The number of Natura2000 polygons with or without MODIS pixels classified for a specific habitat/Environmental Zone combination (H4060/ALS, H9150/ALS and H9150/CON).

H4060/ALS		Natura2000 (N2K)		
		present	not present	Total
classification RS	present	29	130	159
	not present	62	1146	1208
	Total	91	1276	1367

H9150/ALS		Natura2000 (N2K)		
		present	not present	Total
classification RS	present	3	1033	1036
	not present	7	324	331
	Total	10	1357	1367

H9150/CON		Natura2000 (N2K)		
		present	not present	Total
classification RS	present	9	3989	3998
	not present	0	2377	2377
	Total	9	6366	6375

In Table 4 pivot tables are presented between the Natura2000 database and the remote sensing classification results for each habitat/environmental zone combination. The number of Natura2000 polygons with a specific habitat that has a match with pixels classified as having that specific habitat, but also the polygons without that specific habitat and not containing pixels classified as having that habitat are an indication of the quality of the classification. So the summation of the number of Natura2000 polygons on the axis present/present and not present/not present as percentage of the total number of polygons in that specific Environmental Zone are an indication of the quality of the classification. The H4060/ALS classification with 86% showed up as best, while the H9150/ALS and H9150/CON classifications with 23.9% and 37.4% have relatively low levels of agreement between the reference data (Natura2000 sites) and the classification results.

14 Discussion

The assessment of the classification results suggests an overestimation of the H9150 habitats in both environmental zones. An underestimation is present in the case of H4060, although there are indications that this classification is better than in the case of the H9150 classifications. However, the classification of habitats on basis of satellite imagery needs improvement. The main limitation is the lack of more detailed (higher resolution) HANTS vegetation phenology product, next to fact that many habitats do not have a unique phenology. The present spatial resolution of the times series analysis was 250 meter based on the MODIS satellite imagery, while most of the Natura 2000 habitats are still very fragmented at this scale. So for most habitats it will be impossible to find good training samples. Therefore, the resolution/scale of habitat observations and the classification results should be more in line with each other to make them comparable. Remote sensing imagery with higher spatial and/or spectral resolution would improve the possibilities of habitat classification with remote sensing. Next to the fragmented character of the habitats, also the agreement between temporal profiles of for example different forest habitats shows the limitations of the classification methodology.

The quality of the classification results differ between habitats and between environmental zones. Generic classification parameters valid for all kinds of habitat-environmental zones combinations will be an utopia. Habitats differ in reflectance from each other and differ between environmental zones as the biophysical conditions and the phenology development is different. As a consequence the classification of habitats is partly subjective as it depends on the selection and delimitation of training sites.

15 References

European Commission, 2007. Interpretation manual of European Habitats – EUR27. Published by the European Commission, DG Environment, Nature and biodiversity.

Mücher, C.A., S.M. Hennekens, R.G.H. Bunce and J.H.J. Schaminée, 2004. Mapping European Habitats to support the design and implementation of a Pan-European Ecological Network. The PEENHAB project. Alterra report 952, Wageningen.

Mücher, C.A., Hennekens, S.M., Bunce, R.G.H., Schaminée, J.H.J., Schaepman, M.E., 2009. Modelling the spatial distribution of Natura 2000 habitats across Europe. Landscape Urban Plan. 92 (2), 148-159. doi:10.1016/j.landurbplan.2009.04.003.

Mücher, C.A., 2009. Geo-spatial modelling and monitoring of European landscapes and habitats using remote sensing and field surveys. PhD thesis Wageningen University, Wageningen, The Netherlands, ISBN 978-90-8585-453-1, 269 pp.

Roerink, G.J., Menenti, M. and Verhoef, W., 2000. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. International Journal of Remote Sensing, 21(9): 1911-1917.

Roerink, G.J., Menenti M., Soepboer W., Su Z. (2003), Assessment of climate impact on vegetation dynamics by using remote sensing, Physics and Chemistry of the Earth, Vol. 28, pp. 103-109;

Roerink, G.J.; Danes, M.H.G.I. (2010) General surveillance of genetically modified crops : possibilities of GIS and remote sensing

White, M.A., de Beurs, K.M., Didan, K., Inouye, D.W., Richardson, A.D., Jensen, O.P., O'Keefe, J., Zhang, G., Nemani, R.R., van Leeuwen, W.J.D., Brown, J.F., de Wit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A.S., Kimball, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., Lauenroth, W.K. (2009), Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006, Global Change Biology, Vol. 15, pp. 2335-2359;

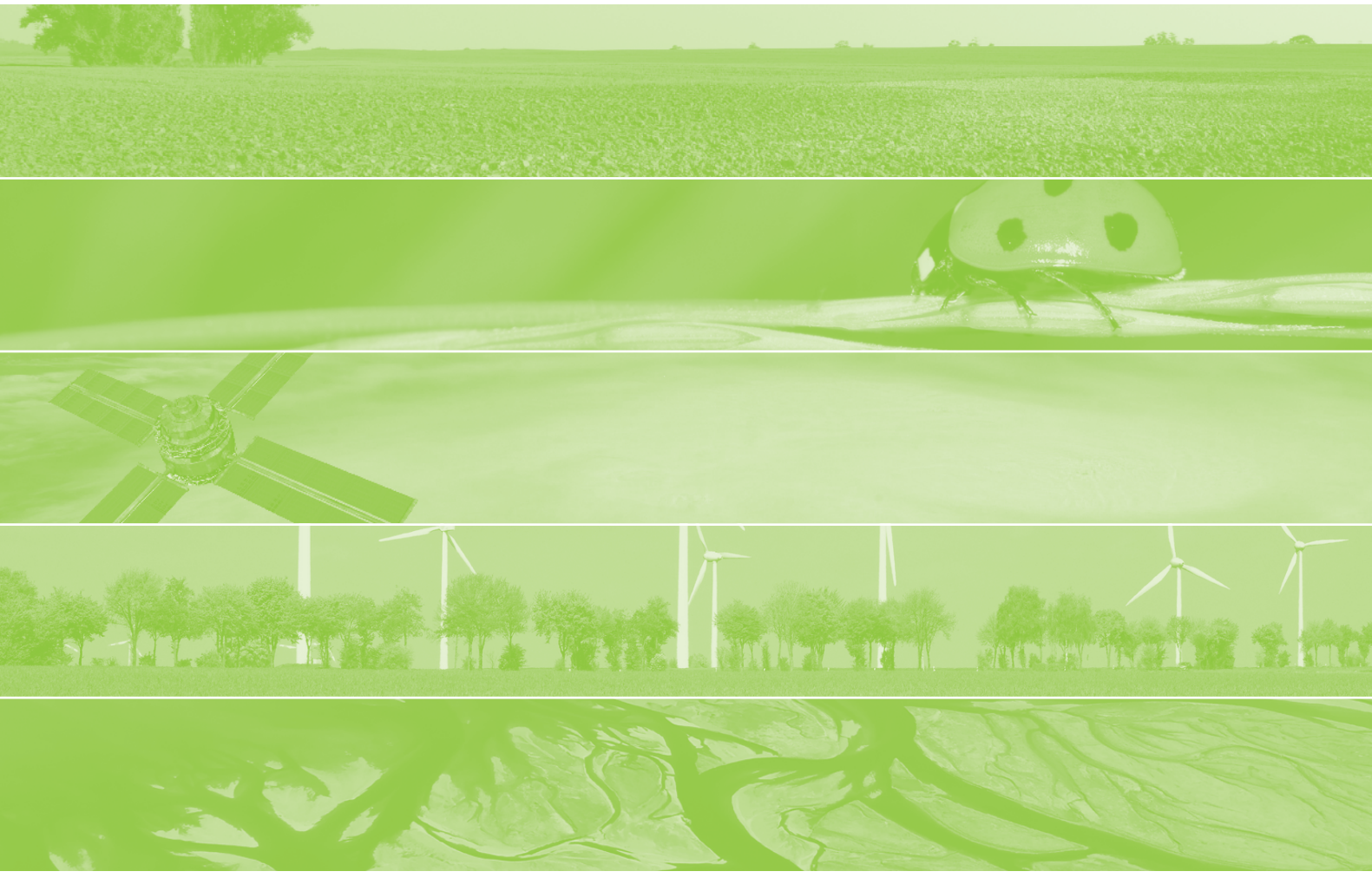
Zhang, X., Sun, R., Zhang, B., Tong, Q. (2008), Land cover classification of the North China Plain using MODIS_EVI time series, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 63, pp. 476-484;

Annex 1

Table A1

Pearson's correlation matrix among Phenolo indicatorsextracted from the MODIS NDVI data for Europe (395,584,800 points).In grey are highlighted correlations greater or equal than 0.9.

	SBD	SBV	SED	SEV	SL	SI	SNI	SPI	STR	GED	GEV	GL	GI	GNI	GPI	MBD	MBV	MED	MEV	ML	MI	MNI	MTR	MPI	SEI	GEI	SBC	SSD	MXD	MXV	OMI
SBD	1	0.19	0.77	0.26	-0.15	0.21	0.26	-0.02	0.18	0.82	0.24	-0.10	0.20	0.26	0.17	0.82	0.18	0.60	0.31	-0.35	0.18	0.28	-0.06	0.18	0.08	0.12	0.93	-0.33	0.84	0.26	-0.15
SBV	0.19	1	0.41	0.99	0.39	0.89	0.94	-0.22	0.95	0.41	0.97	0.42	0.88	0.93	0.93	0.07	0.97	0.36	0.96	0.23	0.92	0.97	-0.21	0.95	0.82	0.87	0.29	0.17	0.23	0.88	0.39
SED	0.77	0.41	1	0.49	0.51	0.60	0.52	0.02	0.54	0.96	0.50	0.49	0.58	0.52	0.53	0.50	0.32	0.84	0.46	0.18	0.54	0.52	0.01	0.43	0.31	0.41	0.90	0.18	0.76	0.53	0.51
SEV	0.26	0.99	0.49	1	0.41	0.91	0.96	-0.21	0.96	0.48	0.97	0.43	0.90	0.95	0.93	0.13	0.95	0.43	0.97	0.23	0.93	0.98	-0.20	0.95	0.81	0.87	0.37	0.18	0.28	0.90	0.41
SL	-0.15	0.39	0.51	0.41	1	0.65	0.45	0.06	0.59	0.39	0.46	0.90	0.63	0.46	0.61	-0.32	0.26	0.50	0.29	0.75	0.60	0.44	0.11	0.44	0.38	0.47	0.14	0.74	0.06	0.47	1.00
SI	0.21	0.89	0.60	0.91	0.65	1	0.96	-0.08	0.96	0.56	0.96	0.66	0.99	0.95	0.97	0.00	0.78	0.56	0.81	0.47	0.98	0.95	-0.08	0.86	0.77	0.84	0.39	0.37	0.28	0.94	0.65
SNI	0.26	0.94	0.52	0.96	0.45	0.96	1	-0.10	0.92	0.51	0.99	0.49	0.95	1.00	0.93	0.09	0.86	0.48	0.88	0.31	0.95	0.99	-0.12	0.87	0.78	0.84	0.38	0.19	0.28	0.99	0.45
SPI	-0.02	-0.22	0.02	-0.21	0.06	-0.08	-0.10	1	-0.19	0.01	-0.13	0.05	-0.07	-0.09	-0.14	-0.06	-0.27	0.05	-0.25	0.10	-0.11	-0.14	0.23	-0.23	-0.17	-0.17	-0.01	0.03	-0.02	-0.05	0.06
STR	0.18	0.95	0.54	0.96	0.59	0.96	0.92	-0.19	1	0.51	0.94	0.60	0.94	0.91	0.98	0.02	0.89	0.49	0.90	0.39	0.96	0.94	-0.16	0.95	0.81	0.88	0.35	0.36	0.26	0.87	0.59
GED	0.82	0.41	0.96	0.48	0.39	0.56	0.51	0.01	0.51	1	0.50	0.49	0.58	0.51	0.54	0.54	0.33	0.78	0.45	0.09	0.51	0.51	0.00	0.42	0.31	0.35	0.92	0.10	0.82	0.51	0.39
GEV	0.24	0.97	0.50	0.97	0.46	0.96	0.99	-0.13	0.94	0.50	1	0.50	0.95	0.99	0.95	0.07	0.89	0.46	0.90	0.32	0.96	0.99	-0.14	0.90	0.81	0.86	0.36	0.21	0.27	0.97	0.46
GL	-0.10	0.42	0.49	0.43	0.90	0.66	0.49	0.05	0.60	0.49	0.50	1	0.69	0.48	0.67	-0.30	0.29	0.43	0.31	0.68	0.62	0.47	0.10	0.46	0.42	0.43	0.18	0.68	0.15	0.48	0.90
GI	0.20	0.88	0.58	0.90	0.63	0.99	0.95	-0.07	0.94	0.58	0.95	0.69	1	0.95	0.98	-0.01	0.76	0.55	0.79	0.47	0.97	0.93	-0.07	0.84	0.76	0.81	0.38	0.37	0.28	0.94	0.63
GNI	0.26	0.93	0.52	0.95	0.46	0.95	1.00	-0.09	0.91	0.51	0.99	0.48	0.95	1	0.92	0.09	0.84	0.49	0.86	0.31	0.95	0.99	-0.11	0.86	0.77	0.83	0.38	0.18	0.27	0.99	0.45
GPI	0.17	0.93	0.53	0.93	0.61	0.97	0.93	-0.14	0.98	0.54	0.95	0.67	0.98	0.92	1	-0.03	0.84	0.48	0.85	0.43	0.97	0.94	-0.13	0.91	0.80	0.84	0.34	0.38	0.27	0.89	0.60
MBD	0.82	0.07	0.50	0.13	-0.32	0.00	0.09	-0.06	0.02	0.54	0.07	-0.30	-0.01	0.09	-0.03	1	0.10	0.46	0.22	-0.66	-0.06	0.13	-0.09	0.00	-0.20	-0.13	0.73	-0.54	0.66	0.07	-0.32
MBV	0.18	0.97	0.32	0.95	0.26	0.78	0.86	-0.27	0.89	0.33	0.89	0.29	0.76	0.84	0.84	0.10	1	0.27	0.97	0.12	0.83	0.91	-0.27	0.95	0.80	0.83	0.25	0.06	0.20	0.78	0.26
MED	0.60	0.36	0.84	0.43	0.50	0.56	0.48	0.05	0.49	0.78	0.46	0.43	0.55	0.49	0.48	0.46	0	1	0.39	0.36	0.53	0.47	0.03	0.42	0.37	0.45	0.78	0.33	0.65	0.50	0.50
MEV	0.31	0.96	0.46	0.97	0.29	0.81	0.88	-0.25	0.90	0.45	0.90	0.31	0.79	0.86	0.85	0.22	0.97	0.39	1	0.10	0.84	0.93	-0.25	0.95	0.78	0.83	0.37	0.07	0.30	0.81	0.29
ML	-0.35	0.23	0.18	0.23	0.75	0.47	0.31	0.10	0.39	0.09	0.32	0.68	0.47	0.31	0.43	-0.66	0.12	0.36	0.10	1	0.51	0.26	0.12	0.35	0.53	0.52	-0.11	0.84	-0.15	0.34	0.75
MI	0.18	0.92	0.54	0.93	0.60	0.98	0.95	-0.11	0.96	0.51	0.96	0.62	0.97	0.95	0.97	-0.06	0.83	0.53	0.84	0.51	1	0.95	-0.11	0.92	0.88	0.93	0.35	0.40	0.25	0.93	0.59
MNI	0.28	0.97	0.52	0.98	0.44	0.95	0.99	-0.14	0.94	0.51	0.99	0.47	0.93	0.99	0.94	0.13	0.91	0.47	0.93	0.26	0.95	1	-0.16	0.91	0.79	0.85	0.39	0.18	0.29	0.96	0.44
MTR	-0.06	-0.21	0.01	-0.20	0.11	-0.08	-0.12	0.23	-0.16	0.00	-0.14	0.10	-0.07	-0.11	-0.13	-0.09	-0.27	0.03	-0.25	0.12	-0.11	-0.16	1	-0.23	-0.17	-0.17	-0.03	0.11	-0.03	-0.08	0.11
MPI	0.18	0.95	0.43	0.95	0.44	0.86	0.87	-0.23	0.95	0.42	0.90	0.46	0.84	0.86	0.91	0.00	0.95	0.42	0.95	0.35	0.92	0.91	-0.23	1	0.90	0.94	0.30	0.28	0.23	0.80	0.44
SEI	0.08	0.82	0.31	0.81	0.38	0.77	0.78	-0.17	0.81	0.31	0.81	0.42	0.76	0.77	0.80	-0.20	0.80	0.37	0.78	0.53	0.88	0.79	-0.17	0.90	1	0.96	0.20	0.39	0.15	0.74	0.38
GEI	0.12	0.87	0.41	0.87	0.47	0.84	0.84	-0.17	0.88	0.35	0.86	0.43	0.81	0.83	0.84	-0.13	0.83	0.45	0.83	0.52	0.93	0.85	-0.17	0.94	0.96	1	0.25	0.40	0.17	0.80	0.47
SBC	0.93	0.29	0.90	0.37	0.14	0.39	0.38	-0.01	0.35	0.92	0.36	0.18	0.38	0.38	0.34	0.73	0.25	0.78	0.37	-0.11	0.35	0.39	-0.03	0.30	0.20	0.25	1	-0.08	0.85	0.37	0.14
SSD	-0.33	0.17	0.18	0.18	0.74	0.37	0.19	0.03	0.36	0.10	0.21	0.68	0.37	0.18	0.38	-0.54	0.06	0.33	0.07	0.84	0.40	0.18	0.11	0.28	0.39	0.40	-0.08	1	-0.13	0.18	0.74
MXD	0.84	0.23	0.76	0.28	0.06	0.28	0.28	-0.02	0.26	0.82	0.27	0.15	0.28	0.27	0.27	0.66	0.20	0.65	0.30	-0.15	0.25	0.29	-0.03	0.23	0.15	0.17	0.85	-0.13	1	0.28	0.05
MXV	0.26	0.88	0.53	0.90	0.47	0.94	0.99	-0.05	0.87	0.51	0.97	0.48	0.94	0.99	0.89	0.07	0.78	0.50	0.81	0.34	0.93	0.96	-0.08	0.80	0.74	0.80	0.37	0.18	0.28	1	0.47
OMI	-0.15	0.39	0.51	0.41	1.00	0.65	0.45	0.06	0.59	0.39	0.46	0.90	0.63	0.45	0.60	-0.32	0.26	0.50	0.29	0.75	0.59	0.44	0.11	0.44	0.38	0.47	0.14	0.74	0.05	0.47	1



Alterra is part of the international expertise organisation Wageningen UR (University & Research centre). Our mission is 'To explore the potential of nature to improve the quality of life'. Within Wageningen UR, nine research institutes – both specialised and applied – have joined forces with Wageningen University and Van Hall Larenstein University of Applied Sciences to help answer the most important questions in the domain of healthy food and living environment. With approximately 40 locations (in the Netherlands, Brazil and China), 6,500 members of staff and 10,000 students, Wageningen UR is one of the leading organisations in its domain worldwide. The integral approach to problems and the cooperation between the exact sciences and the technological and social disciplines are at the heart of the Wageningen Approach.

Alterra is the research institute for our green living environment. We offer a combination of practical and scientific research in a multitude of disciplines related to the green world around us and the sustainable use of our living environment, such as flora and fauna, soil, water, the environment, geo-information and remote sensing, landscape and spatial planning, man and society.

More information: www.alterra.wur.nl/uk