

Temporal Optimisation of image acquisition for land cover classification with Random Forest and MODIS time-series

Ingmar Nitze^a, Brian Barrett^a, Fiona Cawkwell^a

^a Department of Geography, University College Cork, Ireland

Abstract:

The analysis and classification of land cover is one of the principal applications in terrestrial remote sensing. Due to the seasonal variability of different vegetation types, the ability to discriminate different land cover types changes over time. Multi-temporal classification can help to improve the classification accuracies, but different constraints, such as financial restrictions or atmospheric conditions, may limit their application. The optimisation of image acquisition timing and frequencies can help to increase the effectiveness of the classification process. For this purpose, the Feature Importance measure of the state-of-the-art machine learning method Random Forest was used to determine the optimal image acquisition periods for a general (Grassland, Forest, Water, Settlement, Peatland) and grassland specific (Improved Grassland, Semi-Improved Grassland) land cover classification in central Ireland on a 9-year time-series of MODIS Terra 16 day composite data (MOD13Q1). Feature Importances for each acquisition period of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) were calculated for both classification scenarios. In the general land cover classification, the months December and January showed the highest, and July and August the lowest separability for both VIs over the entire nine-year period. This temporal separability was reflected in the classification accuracies, where the optimal choice of image dates outperformed the worst image date by 13% using NDVI and 5% using EVI on a mono-temporal analysis. With the addition of the next best image periods to the data input the classification accuracies converged quickly to their limit at around 8 to 10 images. The binary classification schemes, using two classes only, showed a stronger seasonal dependency with a higher intra-annual, but lower inter-annual variation. Nonetheless anomalous weather conditions, such as the cold winter of 2009/2010 can alter the temporal separability pattern significantly. Due to the extensive use of the NDVI for land cover discrimination, the findings of this study should be transferrable to data from other optical sensors. However, the high impact of outliers from the general climatic pattern highlights the limitation of spatial transferability to locations with different climatic and land cover conditions. The use of high-temporal, moderate resolution data such as MODIS in conjunction with machine-learning techniques proved to be a good base for the prediction of image acquisition timing for optimal land cover classification results.

Keywords: Land cover classification, Random Forest, MODIS, Machine-learning

1. Introduction

Analysis of land cover is one of the principal applications in terrestrial remote sensing, however, the timing and frequency of image acquisition, among other factors, can limit the accuracy with which different land cover types and changes can be distinguished (Carrão et al., 2008; Lunetta et al., 2004). If the optimal number and dates of images for land cover classification can be identified in advance, the efficiency with which the image processing is undertaken can be enhanced. In practice, the optimal acquisition date(s) are often determined following assessment of the variability in spectral signature of classes of interest for site specific applications. With the systematic examination of remote sensing time-series and feature selection techniques this process can be automated.

While feature selection processes are important for improved image classification, or for an increased understanding of the classification rules (Liu et al., 2005), their most frequent field of application is in bio-informatics; where predictive variables are extracted from gene sequences (Díaz-Uriarte and Alvarez de Andrés, 2006). To date, only a limited number of studies have been published which explore methods of advanced feature selection within multi-spectral or multi-temporal imagery. Van Niel et al. (2005) used the Jeffries-Matusita (JM) distance for crop classification in Australia, Carrão et al. (2008) utilized a median Mahalanobis distance for general land cover classification in Portugal and Lhermitte et al. (2011) compared time-series of different Vegetation Indices with several mathematical distance metrics. In contrast, hyper-spectral remote sensing approaches to feature selection have been widely used for the reduction of dimensionality (e.g. Serpico and Bruzzone, 2001; Pal, 2006; Guo et al., 2008; Pal and Foody, 2010). Analogous to bioinformatics, different embedded approaches like Recursive Feature Elimination (RFE) for Support-Vector-Machines (SVM) (Pal and Foody, 2010) and the Feature Importance measure within Random Forest (Stumpf and Kerle, 2011), as well as distance based measures, such as Bhattacharyya or JM (Guo et al., 2008) have been widely used.

Hyper-spectral and hyper-temporal datasets have a very similar structure; containing a high number of correlated spectral bands. The difference lies in their variable dimension, namely the spectral domain for hyper-spectral data and the temporal domain for hyper-temporal data. Therefore, methodologies for the reduction, or selection, of the data used in the former can be transferred to the optimisation of data acquired over a period of time. Hyper-temporal data are generated from a platform which offers a high temporal resolution and a sufficiently long time-series such as AVHRR (Hill et al. 1999; Hermance et al., 2007), MERIS (Zurita-Milla et al., 2009; Carrão et al., 2010; O'Connor et al., 2012) or MODIS (Lunetta et al., 2006; Carrão et al., 2008; Wardlow and Egbert, 2008; Pringle et al., 2012). Typically, Normalized Difference Vegetation Index (NDVI) or other sensor-specific vegetation indices (VI) serve as the base for land-cover and crop type classification or phenological analysis from regional to

global scales. Although universally applied, NDVI has its drawbacks. It is susceptible to soil and atmospheric influences and saturates in dense vegetation canopies (Huete et al., 2011). Sensor specific VIs such as the MODIS Enhanced Vegetation Index (EVI) captures vegetation phenology more accurately and can demonstrate sharper growing season peaks and greater sensitivity to canopy structure differences (Xiao et al., 2003; Huete et al., 2011).

Due to cloud contamination, atmospheric variability and bi-directional effects, the signal of the time-series of hyper-temporal datasets can be severely affected by noise, demonstrating highly volatile behavior in their raw state. In countries like Ireland, where cloud cover and atmospheric attenuation are persistent, time compositing procedures such as extraction of maximum values over a given period are widely applied (O'Connor et al., 2012), but noise can still prevail (Chen et al., 2004). To overcome the random spikes in the composited time-series, different smoothing techniques have been used, ranging from Fourier based filters (Sellers et al., 1994; Chen et al., 2004) to curve fitting functions (Jönsson and Eklundh, 2002; Chen et al., 2004; Beck et al., 2006; Atzberger and Eilers, 2011).

In this paper we present an analysis of the separability of different land-cover classes at different times of year using a MODIS Terra 16-day composite time-series and the internal Feature Importance measure of the machine-learning Random Forest algorithm (Breiman, 2001). The performance of NDVI and EVI vegetation index time-series are compared, and single year fluctuations as well as long-term trends are analysed over a 9-year period.

2. Data and Methods

2.1 Study area

The study area is centred on County Longford in the central Midlands Region of Ireland, with parts of the surrounding counties of Galway, Roscommon, Leitrim, Cavan, Westmeath, Meath and Offaly also included (Figure 1). This region is characterised by lowlands to the south and a strongly undulating terrain in the north, interspersed with lakes and peatbogs of varying size. Grassland is the dominant land-cover, occupying 75.4% of the entire area, with peatland (10.9%), forest (7.0%) and water (3.4%) the other key land cover types (CORINE 2006). Due to its low population density of approximately 35 inhabitants per km² (CSO, 2011), settlements and built-up areas cover only 1.1% of the study area and crops are only sparsely cultivated in the Midlands region, covering around 2.2% of the study area. The maritime location of the island of Ireland, and the strong influence of the North Atlantic Drift, gives rise to a climate with a small temperature range, frequent cloud-cover and annual precipitation of approximately 1000 mm in this region (Met Éireann, 2013).

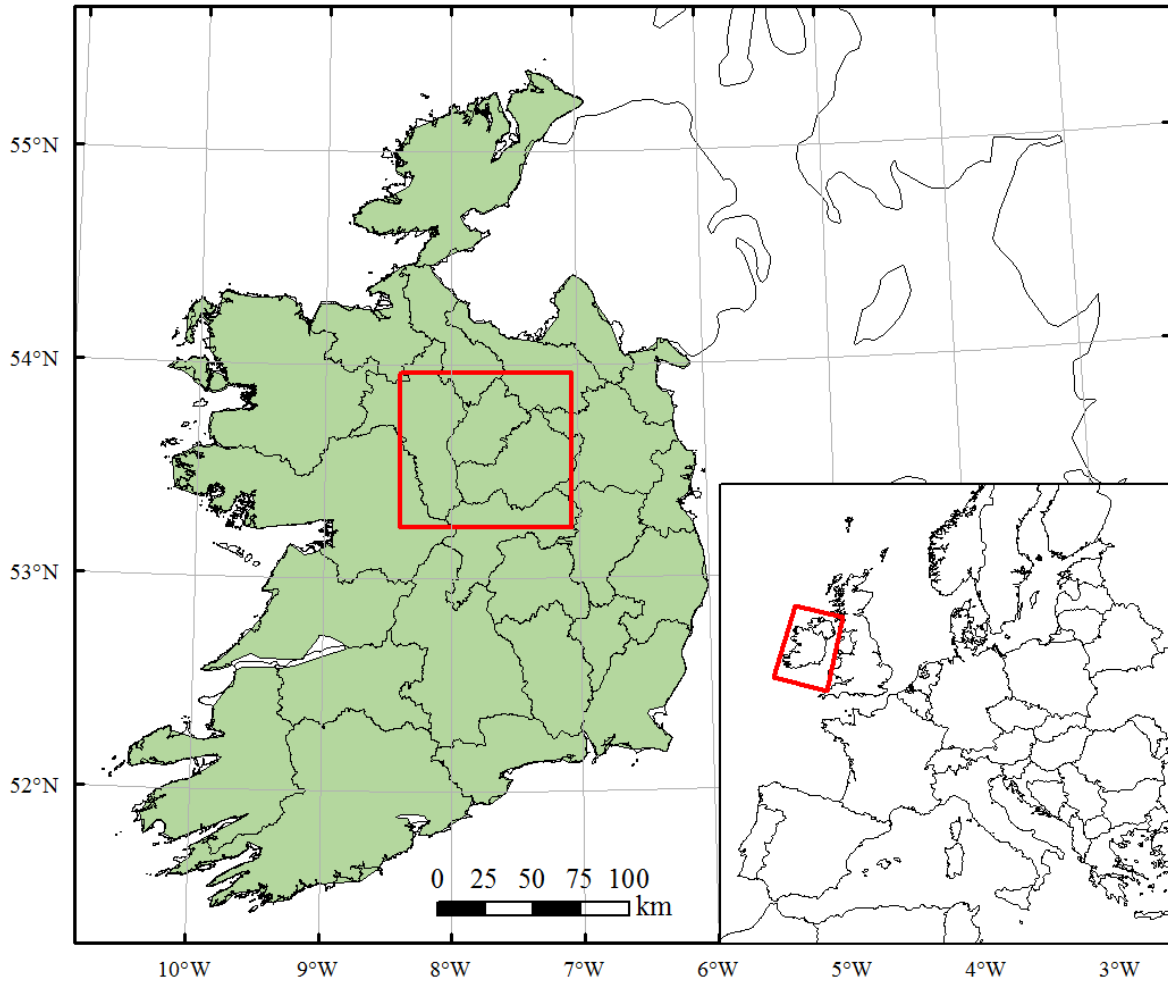


Figure 1: Geographical location and extent of the study area.

2.2 MODIS Data

A large amount of low quality data caused by cloud cover is inherent in daily datasets acquired over Ireland, and as O'Connor et al. (2012) demonstrated, a compositing period of at least 10 days is required. Thus, the MODIS Terra 16-day composites (MOD13Q1), with a spatial resolution of 250m were chosen as the major data source for this study, with the optimal image date inside the fixed 16-day period determined by a constrained view-angle maximum-value-composite (CV-MVC) algorithm (Huete et al., 2011). The MOD13Q1 product contains two vegetation indices (NDVI and EVI); as well as four spectral bands: blue, red, near-infrared (NIR) and mid-infrared (MIR). In addition to the spectral data, pixel-based supplementary information, such as VI-quality, pixel-reliability (PR), and shadow- and cloud-masks, is available in these datasets. The PR value refers to the reliability of the spectral measures on the pixel-level, where a value of 0 represents a highly reliable value, 1 represents reduced quality, 2 represents low quality due to snow or ice and 3 represents a non-visible target. The MODIS datasets used in this study cover an 11-year period from 2001 to 2011, consisting of 23 images per year, giving 253 datasets in total.

2.3 Land Cover Classes and Ground Truth

Due to the coarse spatial resolution of the sensor and the heterogeneity of the Irish landscape, the land cover classes were restricted to Grassland (G), Forest (F), Water (W), Settlement (S) and Peatland (P). Grassland was split into the sub-classes Improved (GA) and Semi-Improved (GS) in recognition of the ecological and economic importance of pasture in Ireland, where this particular land-cover type accounts for 55% of the total land surface and 90% of the agricultural areas (Eaton, 2008). Cropland was omitted due to its small parcel sizes and scarcity in the study area.

The Random Forest classifier required ground truth data for training the classifier and evaluating the products. A vector grid of the outlines of the MODIS pixels was overlaid on Ordnance Survey Ireland (OSi) cadastre boundary vector data as well as high-resolution imagery including OSi aerial photography, Google Earth and Bing Maps to identify pure pixels of the five land cover classes. For the Settlement class, the requirement of using pure pixels was relaxed in order to facilitate a sufficient number of training pixels and to accommodate the rural character of this region. Grassland ground-truth pixels may include a small fraction of hedgerows or walls, which are typical features for the separation of agricultural fields in Ireland. Due to the small field- and large pixel size, these objects are inevitably included in a ground-truth selection of agricultural lands.

During the ground-truth selection process, all Grassland pixels were assigned to the Improved or Semi-Improved classes based on supplementary grassland survey data (O'Neill et al., 2009) and visual interpretation of aerial imagery. In total, 1060 training points were selected for the land cover classification (Table 1).

Table 1: Number of ground-truth samples per land cover class

Land Cover Type	# of Pixels
G – Grassland	246
<i>GA - Improved Grassland</i>	124
<i>GS - Semi-Improved Grassland</i>	122
F - Forest	102
W - Water	320
S - Settlement	73
P - Peatland	319

2.4 Random Forest

Random Forest (RF, Breiman, 2001) is a machine learning ensemble based on CART decision tree classifiers (Breiman et al., 1984). One of its greatest strengths, in addition to its high accuracy and robustness to outliers and noise, is the calculation of different internal quality measures, such as out-of-bag error or feature importance (FI) (Breiman, 2001).

RF builds a user-specified number of decision trees, using a random, bootstrapped subset of around 2/3 of the dataset (in-the-bag/itb) for training while keeping the remaining third (out-of-bag/oob) for the error estimation of the training process. Each single tree is built with the randomised itb-data and is fully grown. At each tree node a random selection of variables

(features) is chosen and tested for the best split, which is calculated by the Gini index of impurity. The final class assignment is determined through a voting process choosing the majority of the single tree classification output. The FI measure, which represents the contribution of a specific feature/variable towards the classification accuracy, is calculated on single trees of the Random Forest using the decrease in error-rate for a specific variable at a split node. The FIs of the single trees are averaged to the final FI of the Random Forest Classifier. Finally, the FI values are normalised in a range of 0 to 1, where the sum of all FI equals 1.

3. Data Processing

3.1 Data pre-processing

The NDVI, EVI and PR values were extracted for each MODIS 16-day composite dataset, and re-distributed to create time-series data stacks. For each of the 1060 ground truth samples the NDVI and EVI time-series from 2001 to 2011 were selected from the data stacks and prepared for analysis. As the time-series exhibited noisy behaviour, a cleaning process based on the MODIS Pixel Reliability in conjunction with a Hodrick-Prescott (HP) time-series filter was applied. An adaptive, weighted temporal filter was chosen to decrease the influence of low-quality values by 1) suppressing outliers, 2) reducing the influence of low-quality data and 3) filtering noise.

First, a one-dimensional Gaussian Filter with a standard deviation (σ) of 2 was convolved with the time-series of each ground-truth pixel. The mean and standard deviation of the residuals between the raw and the filtered time-series were calculated and raw values of more than 1σ below the filtered curve discarded and replaced with the filtered value, whereas for potential upper outliers, we used a less constrained σ of 5. An asymmetric solution was chosen due to the nature of possible outliers in VI-time series, where negative outliers are usually caused by snow or unfavourable atmospheric conditions.

Next a local running Gaussian filter with σ of $10^{\text{PR}-1}$ was applied, where the internal pixel reliability measure (PR) determined the size of the filter. Where PR equals zero, and the pixel value was assumed to be true, the σ was reduced to 0.1 so that the original value was nearly unchanged. With a PR of 1, the σ of the Gaussian kernel was set to 1, thus including adjacent values for the recalculation of the local value. A PR of 2 or 3 increased the kernel size significantly, so that the influence of the local value was further reduced and the weight of adjacent observations was increased.

The final smoothing process was performed with the Hodrick-Prescott-(HP) filter (Hodrick and Prescott, 1997) with its parameter λ , which controls the magnitude of the smoothing, set to 20 after testing different configurations. After the curve-smoothing process, the first and last years of the time-series (2001 and 2011) were removed from the dataset in order to avoid inaccurate results caused by edge-effects of time-series filtering, thus leaving 2002 to 2010 inclusive for the analysis. Finally, the smoothed time-series datasets were split into single years, each containing 23 values per VI.

3.2 Random Forest Feature Importance

A Random Forest Classifier, built-up of 5000 single classification trees (as recommended by Diaz-Uriarte and Alvarez de Andrés, 2006 and Stumpf and Kerle, 2011), was first initialized and trained with the 23 features (dates) of the 2002 NDVI dataset. During the training stage, the FI and error estimates (out-of-bag error) were calculated and stored internally. The FI values were then exported and prepared for analysis. The initialisation and training stages were repeated for each year, for both VIs, and for the group of general land cover classes and for Grassland sub-classes (i.e. 36 times in total). The data processing was executed using the scikit-learn package (Pedregosa et al., 2011) in python 2.7.

3.3 Analysis of Feature Importance values and classification outputs

The 23 calculated FI values for each year were averaged to obtain the mean FI value for each 16-day composite period of the 9-year time-series. This allowed long-term trends to be analysed and the periods of high and low class separability to be identified. Additionally, single year results were extracted and compared to the general trend as well as to each other, to assess the variability of single years and to highlight potential factors that influence the seasonal separability.

Finally every single acquisition period was classified independently using the NDVI or EVI smoothed values of that period with just 200 trees. This low number, compared to the extraction of FI values, was determined by testing different forest sizes, where the classification results stabilised quickly with more than 100 trees. The RF internal out-of-bag accuracy was used as the primary evaluator, while a randomised stratified five-fold cross-validation further confirmed the classification results. In a second step, a successive classification was performed, where the data of the most separable acquisition period were used first followed sequentially by the next best in order to find the minimum number of necessary acquisitions for the maximum classification accuracy. This process was repeated, using a reversed order, with the least informative image dates first.

4. Results and Discussion

4.1 Vegetation Index Time-Series

The averaged seasonal NDVI curves of all five land cover types exhibit a typical seasonal behaviour, with annual minima during winter and maxima during summer (Figure 2). Despite smoothing, the curves still include sporadic small-scale fluctuations, predominantly during summer. Grassland is more strongly affected by this behaviour than the other classes, which could be a result of local management practices such as cutting and grazing.

Some class combinations indicate a higher potential for misclassification, due to a significant overlap in their NDVI values. Grassland and Forest display similar NDVI values and characteristics, with the greatest differences during winter, but nearly identical values during spring and early summer. Grass and Forest NDVI values are typically in the range of 0.7-0.8, and are thus comparable with other studies such as Lunetta et al. (2006) and Colditz (2007). The Settlement and Peatland classes also have comparable average NDVI values although Settlement has a larger standard deviation. High variability in urban structures and of the

amount of vegetation within some pixels, caused by the loosened selection criteria for Settlement ground-truth values, may have an influence on this high intra-class variation. The NDVI curves of Water are below those of each other class and should be therefore highly separable. Nonetheless, during summer its average NDVI frequently exceeds 0.5, occasionally reaching values of higher than 0.6. These comparably high values might potentially be caused by algal blooms or high biomass content in the predominantly shallow lakes.

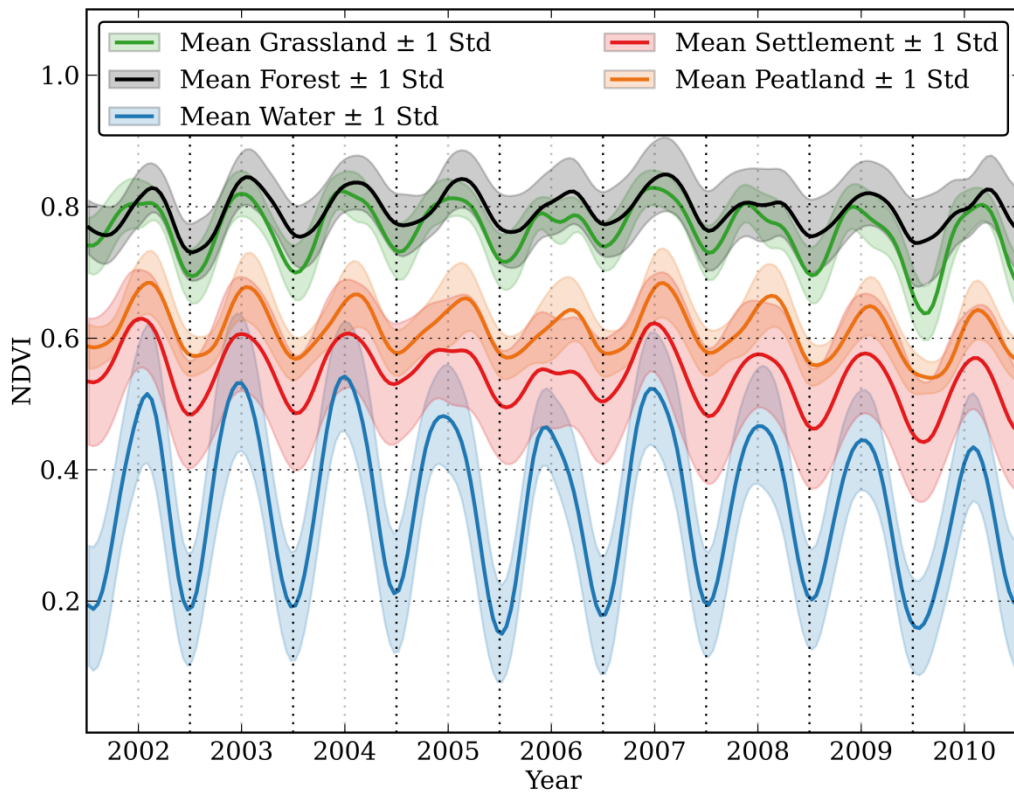


Figure 2: Seasonal mean land cover class NDVI curves with one standard deviation above and below the mean shown in the shaded region.

Compared to the NDVI, the EVI time-series exhibits a more well-defined curve with distinct seasonal peaks and few deviations (Figure 3). Grasslands and Forest exhibit a smaller overlap, with Grassland having a generally stronger EVI signal in the range 0.5-0.7. Settlement and Peatland however exhibit almost identical behaviour in terms of both their mean and standard deviation, while Water is completely separated from the other classes, having very low values generally less than 0.2.

The differences in behaviour between the NDVI and EVI curves, such as the saturation of NDVI in dense vegetation and its higher sensitivity to less intense vegetation such as that of Peatlands in addition the more defined growing season peaks of EVI, is in agreement with the observations of Huete et al. (2011).

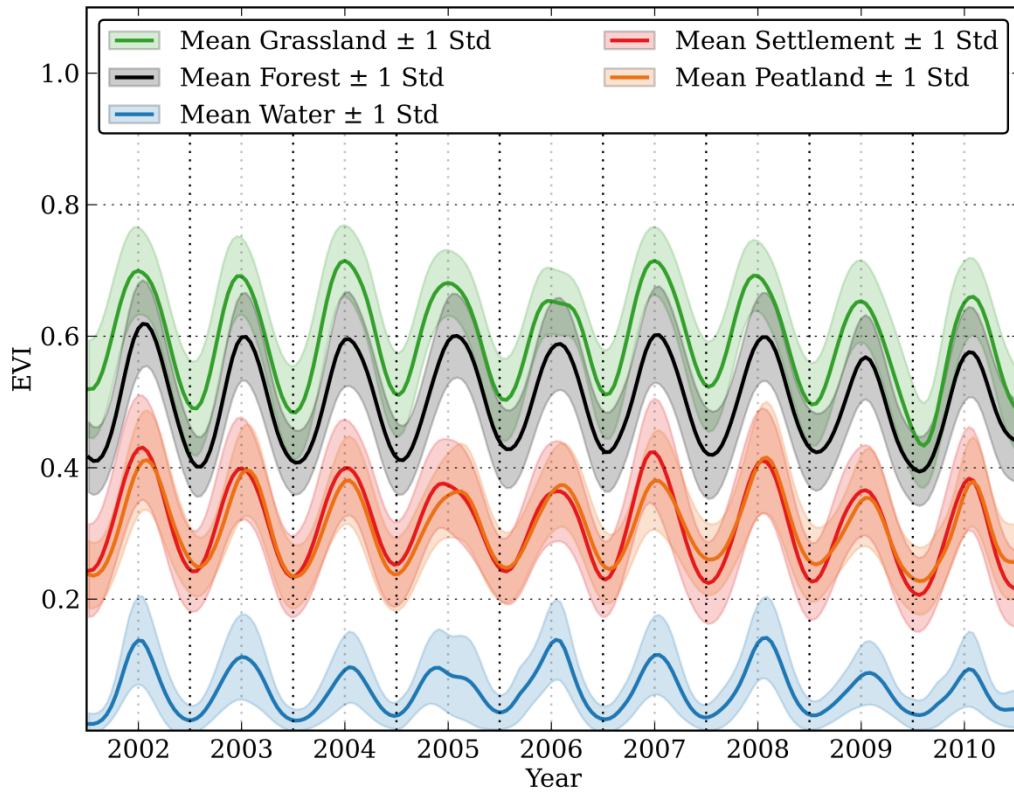


Figure 3: Seasonal mean land cover class EVI curves with one standard deviation above and below the mean shown in the shaded regions.

These effects are also evident when the two different Grassland types are distinguished. Using the NDVI, their temporal profiles show a high degree of overlap, with nearly identical values during the summer months (Figure 4) when NDVI reaches its saturation level. The small-scale fluctuations identified previously during the summer are more prevalent in Improved Grassland (GA) as opposed to the Semi-Improved Grassland (GS). The EVI time-series, by contrast, exhibits a better distinction of both classes, where GA is characterised by constantly higher values than GS and a slightly more pronounced asymmetric seasonality having a rapid green-up phase, with some small-scale fluctuations evident.

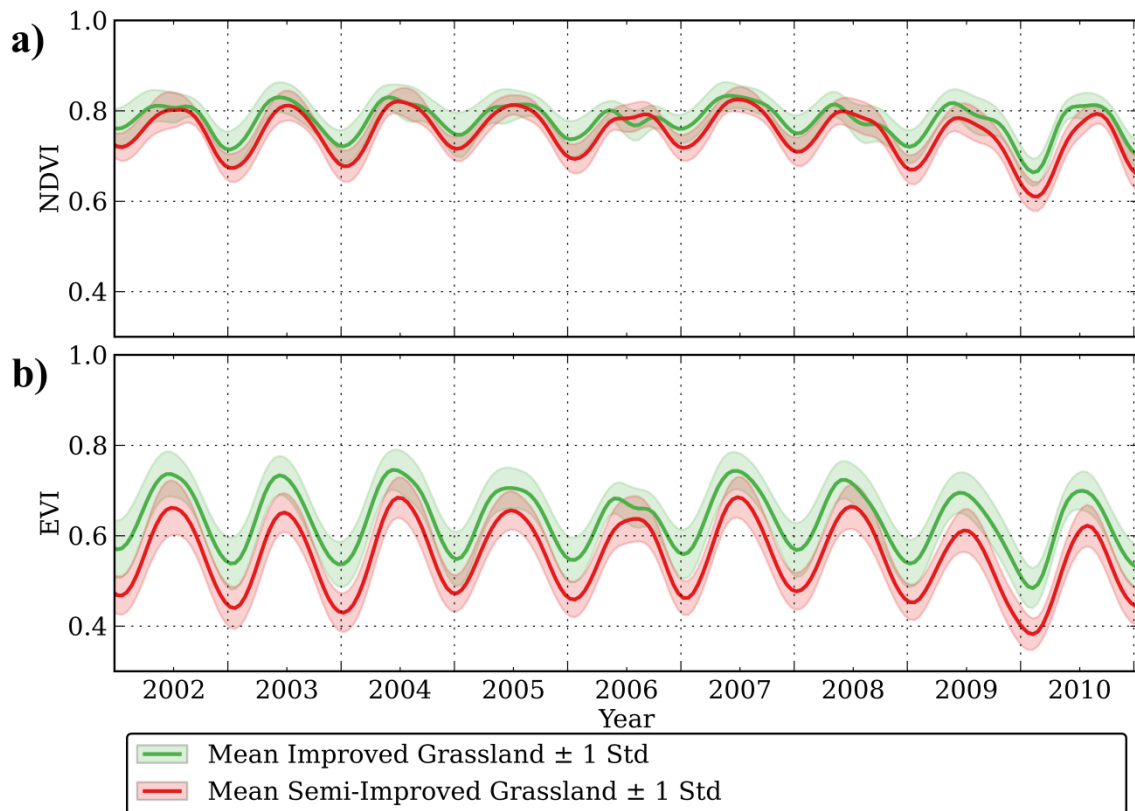


Figure 4: Seasonal mean Grassland classes VI curves with one standard deviation above and below the mean shown in the shaded region – a) NDVI, b) EVI

4.2 Feature Importances

4.2.1 All classes

a) NDVI

The averaged FI values derived from smoothed single-year NDVI time-series exhibit a seasonal pattern with a distinct peak during autumn and winter (late October to January), a pronounced summer minimum (June to August), and little variability throughout spring (March and April) (Figure 5). Hence, the best average separability between the Grassland, Forest, Water, Settlement and Peatland classes occurs during the winter months, which should result in the highest classification accuracies from images acquired during this period.

The single year FIs are characterised by the two same extremes as the whole dataset. During each year, the maximum separability occurs from November to January, while the summer months exhibit decreased separabilities (Figure 5). The time of largest inter-annual variation appears to be the spring period from March to May, where some years (e.g. 2004, 2005 and 2009) feature a minor FI maximum, which varies in length and intensity. During most other years this period shows stagnant or steadily decreasing FI values.

b) EVI

The average FI of the EVI for the five land cover classes exhibits a similar pattern to that of the NDVI, with a maximum in winter and minimum during the summer months. EVI in contrast features a slightly higher spring-time separability (Figure 5), whereas NDVI values are more separable towards the end of the calendar year.

The single year EVI FI measures have a rather high year to year stability with only a few outliers. With the exception of 2009 and 2010, all years exhibit maximum separability from December to February, followed by a very slow decline of FI values to the summer minimum. 2009 and 2010 have a significantly different temporal separability with strong peaks in April-May and September-October (in 2010) accompanied by winter minima (Figure 5). The reason behind this anomalous behaviour can be attributed to a harsh and long lasting cold spell during the winter 2009/2010 in Europe, with average temperatures in central Ireland of 1.9°C, which lies 2.4°C below the mean winter temperature (Met Éireann, 2010). In December 2010, similar harsh winter conditions prevailed, which again may have suppressed the separability at the end of that year.

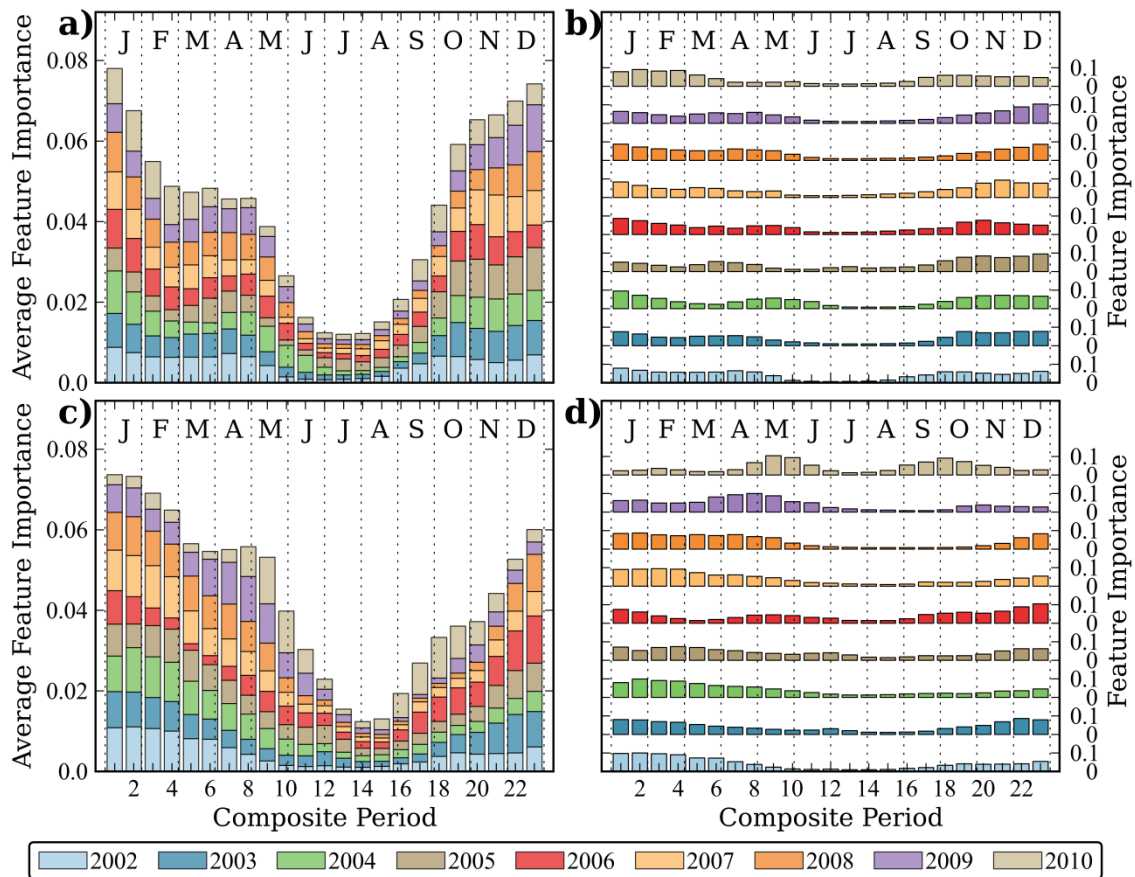


Figure 5: a) Mean Feature Importances NDVI – all classes; b) Single year Feature Importance NDVI – all classes; c) Mean Feature Importances EVI – all classes; d) Single year Feature Importance EVI – all classes

4.2.2 Single Classes

The results presented to date were achieved by classifying all classes simultaneously; however, classifying specific land-cover types can further exploit their temporal variability.

Spectrally similar classes were tested for their temporal separability, applying the same methodology as using all classes at once. Meaningful class combinations are considered as those, with similar VI-profiles and therefore high potential of misclassification, namely Grassland-Forest and Improved Grassland to Semi-Improved Grassland.

Grassland vs. Forest

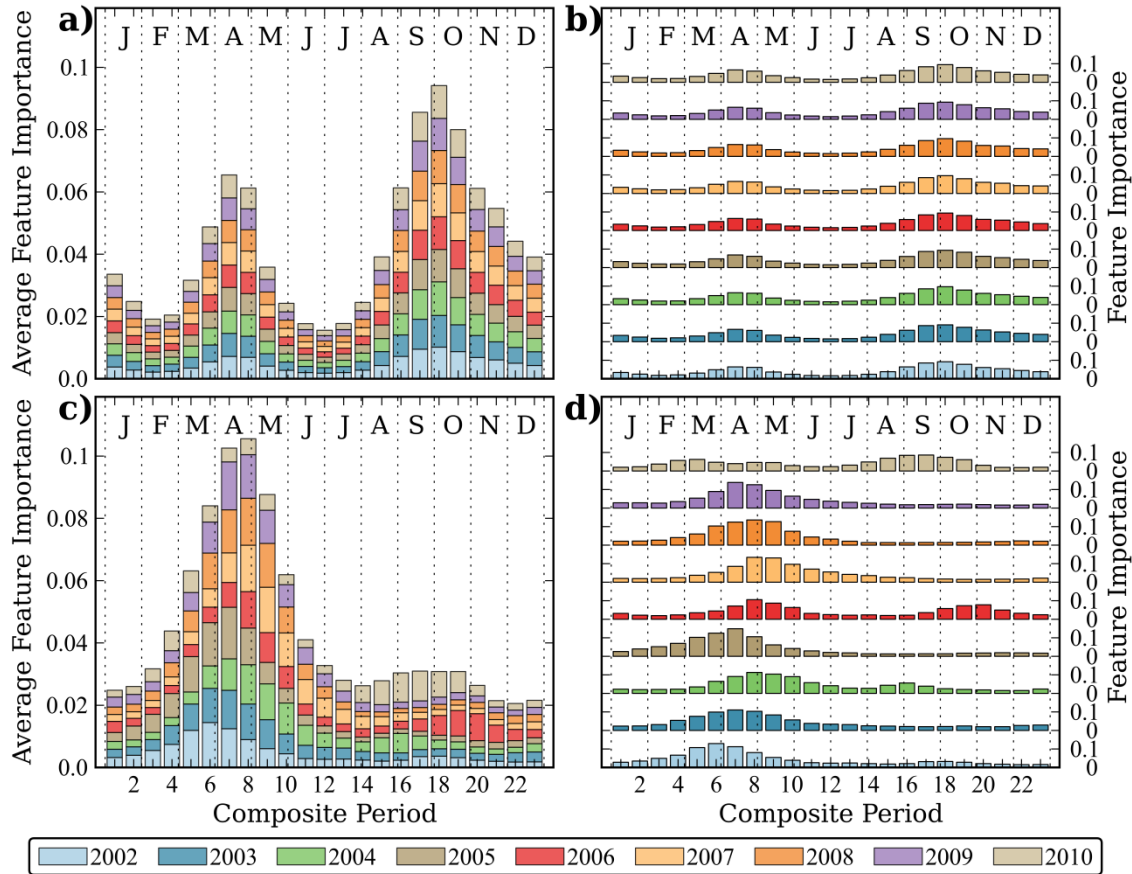


Figure 6: a) Cumulative Feature Importances NDVI - G vs F; b) Single Year Feature Importance NDVI - G vs F; c) Cumulative Feature Importances EVI - G vs F; d) Single Year Feature Importance EVI - G vs F.

The NDVI dataset indicates that the best separability between Grassland and Forest pixels occurs between September and January within all years (Figure 6). The specific timing and magnitude of the peak however, varied in its intensity from one year to another, a result that has already been discussed for the FIs using all classes. Between the winter maximum and the summer minimum, a second minor peak at the end of April is present within the accumulated FIs, caused by anomalous high spring separability periods in 2002 and 2004.

Compared to the NDVI time series, the EVI has a completely different pattern in its temporal separability behaviour of Grassland and Forest (Figure 6). One dominant peak during spring, from late March to early May is accompanied by a nearly level period from July to January. The timing of the peak period is very stable in its length and magnitude with the exception of 2010 where the very cold winter led to more spectral similarity between these classes accompanied by a peak during August and September, which is not evident in any other year. A secondary peak in autumn 2006 can be probably attributed to particularly warm

temperatures, more than 1°C above normal and potential drought stress or an altered management regime of Grassland during the summer, which led to a decreased EVI.

Improved Grassland vs. Semi-Improved Grassland

The seasonal separability of both Grassland types exhibits similar patterns to other class combinations with the summer minimum preceded by a strong spring peak and a second local maximum during November/December (Figure 7). Highly managed Improved Grasslands have a slightly higher photosynthetic activity level than their non- or less intensely managed counterparts, with an earlier start to their growing season and senescence occurring later. The NDVI and EVI both follow the same trend, although the degree of separability in the spring months is higher for the EVI, with a distinct peak evident every year, which is not always apparent in the NDVI data.

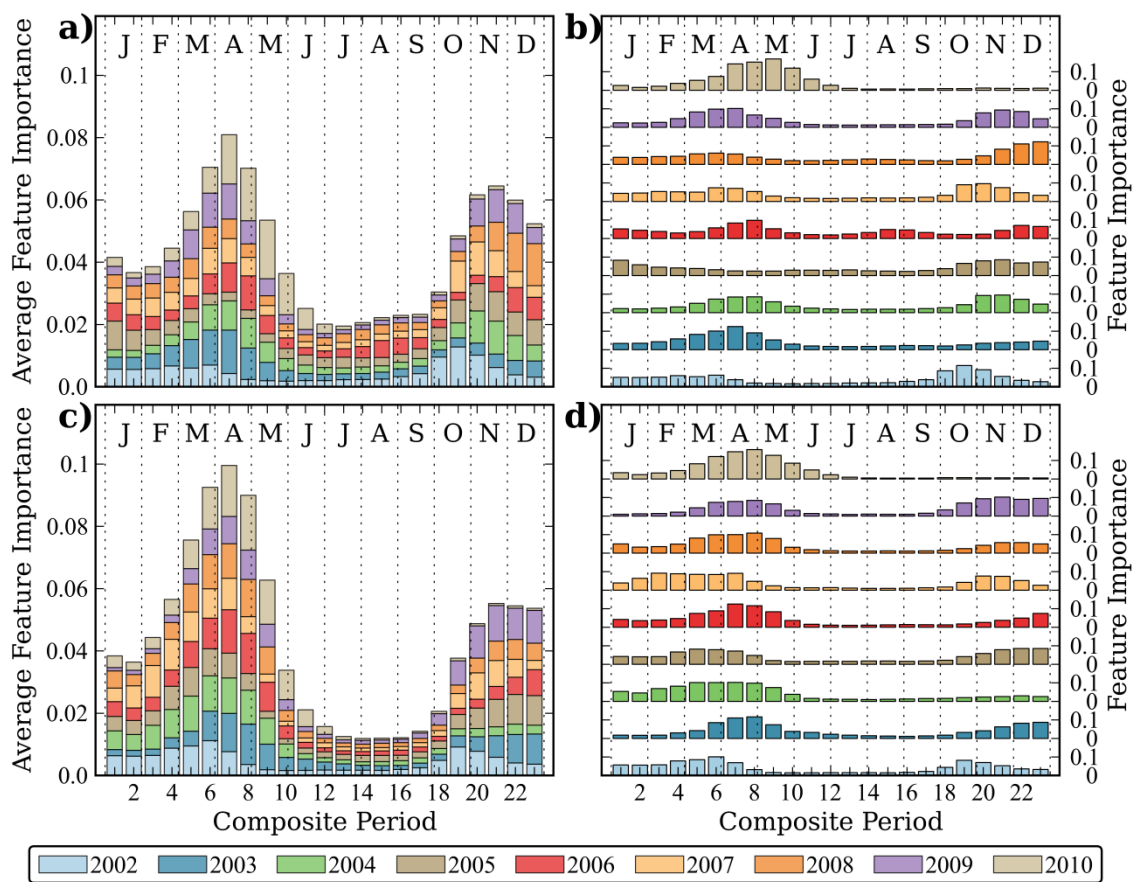


Figure 7: Cumulative Feature Importances NDVI – GA vs GS ; b) Single year Feature Importance NDVI – GA vs GS; c) Cumulative Feature Importances EVI – GA vs GS; d) Single year Feature Importance EVI – GA vs GS.

4.3. Classification results

4.3.1 All Classes

Figure 8a shows the average overall classification accuracies (oob-score) based on NDVI values for single acquisition periods in solid green with single year classification accuracies in dotted green. Unsurprisingly the overall accuracy trends resemble the observed

autumn/winter maxima and summer minima of the Feature Importances very well. The average overall accuracy ranges from between 79 %, from images acquired from October to January, to 67 % at the end of June. Single date maxima of 84 % in March 2010 and a minimum of 59 % in June 2004 represent the extreme values of the entire nine-year period. Independent of any specific year, the winter acquisitions generally outperform all summer images. Inter-annual differences are the highest during the early summer months May to July, where accuracies exhibit a range of up to 15 %, whereas the classification accuracies in winter are much less varied.

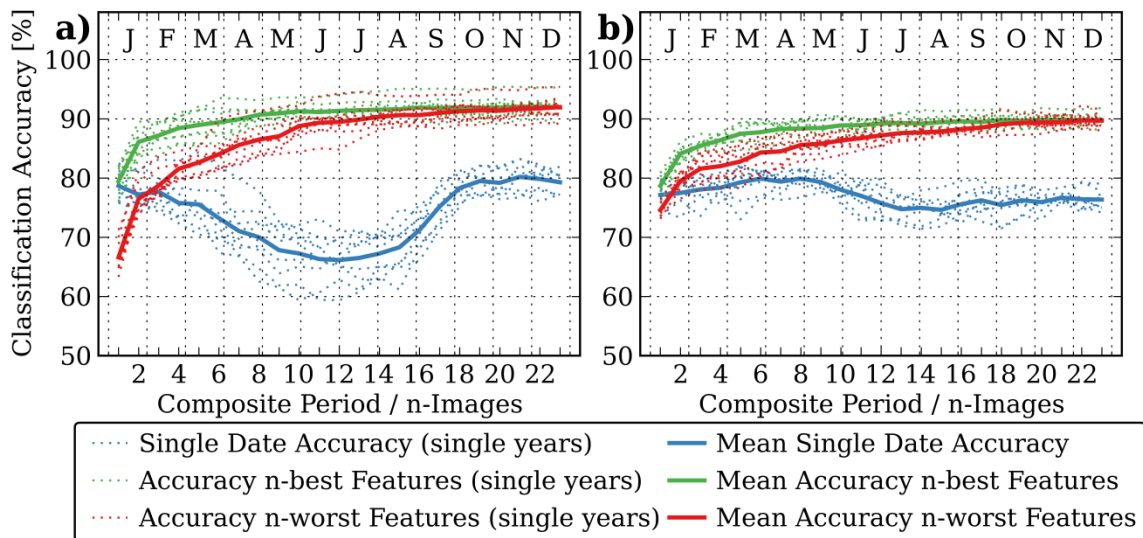


Figure 8: Classification accuracies of general land cover classes. Mono-temporal (green), Multi-temporal Classification accuracies – best number of features first (red), worst number of features first (blue). a) NDVI; b) EVI

The cumulative accuracies of NDVI, using the best images first, reach their maximum classification potential of just over 90 % with the 8th image after which the classification accuracy converges to a maximum of around 92 %. Using the least informative images first, the average classification accuracy is significantly decreased at the beginning compared to the optimal image choice (66.6 % vs. 79.4 %). The addition of further images improves the classification results; however this curve converges very slowly to the classification optimum, reaching it after about 20 images. The results for the EVI are somewhat similar, although the separability derived from the images acquired late in the year is less apparent. With a range of only 5.3 % between the best (79.9 %) and the worst (74.6 %) single date mean accuracy in April and July respectively, the influence of the acquisition timing on classification accuracy is much lower than for NDVI. Nonetheless, despite its slightly inferior single date classification performance, NDVI proves to be the better choice in a multi-temporal classification, where it outperforms EVI by up to 2.5 %. The higher volatility and therefore smaller spectral correlation of the specific class combination might have caused the superiority of NDVI in a multi-temporal analysis.

4.3.2 Single Classes

Grassland vs. Forest

The corresponding single-date classification of the NDVI images confirms the seasonal FI patterns, with their peak in accuracy around acquisition period 18 in early October. The minor local FI maximum in spring was not evident in the average classification accuracies, since only single years exhibit this trend. 2010 stands out with unusually high classification accuracies, in excess of 91 % from a single composite image, and a long phase of very good separability. During early 2010 the NDVI values of Grassland were unusually low compared to other years, while Forest did not suffer from such extremely depressed NDVI values, and as a result classification at this time period outperformed the best result of each other year by more than 5%. The successive classification process shows very similar results to the classification of all classes together. Starting from the image with highest separability, the classification accuracy converges to its natural classification limit of 90 % at around 10 images.

The single date EVI images produced better or equal results compared to NDVI during the first eight months of the year where mean classification accuracies of up to 81 % in acquisition period 8 are achieved, however images from the later part of the year result in poorer classification outputs (Figure 9). In all but one year (2010) the best mono-temporal accuracy exceeds 79.9 %, and unlike the NDVI dataset the spectral difference between the grass and forest is not decreased in this year. As with all the classes classified together, the EVI slightly under-performs the NDVI when multiple images are input to the classifier, however the difference is only of the order of 2-3% and 8 to 10 images appear sufficient to reach the classification accuracy plateau. Nonetheless, the final choice of the number of images has to be chosen by the user, where different factors, such as number of available images and image costs, have to be considered.

Improved Grassland vs. Semi-Improved Grassland

This seasonal distribution of the NDVI-FI is mirrored in the classification results, with average mono-temporal classification accuracies of between 52 % and 72 % taking all years into consideration, and a maximum single-date classification accuracy of 93.9 % in spring 2010). With a range of 20 %, the seasonal differences are much more accentuated than for other class combinations. The EVI achieves superior average classification accuracies, outperforming NDVI by 9.5 % to 17.6 %, which could be attributed to the inability of the NDVI to discriminate between highly photosynthetically active vegetation types due to saturation (Figure 9).

Increasing the number of images using NDVI improves the classification accuracy. When the images with the highest separability are selected first, the average classification accuracies converge quickly towards the maximum accuracy. With only three images the average accuracy reaches 82.6 %, only 3.4 % less than the maximum of 86 %, which is reached with 17 different input images. Between different years, the results vary widely, ranging between 80 % and 95 % maximum accuracy. By comparison, when the least separable images are selected first the classification results are typically 20 % below the optimal image choice for

a small number of images and converge slowly to the reach a maximum accuracy when approximately 20 images are included.

With the usage of EVI the classification accuracies are already on a high level of more than 80 % and benefit less by the addition of additional images. The multi-temporal classification accuracies outperform those of NDVI by up to 10.4 % on a mono-temporal analysis, which decreases quickly to 5 % with three images.

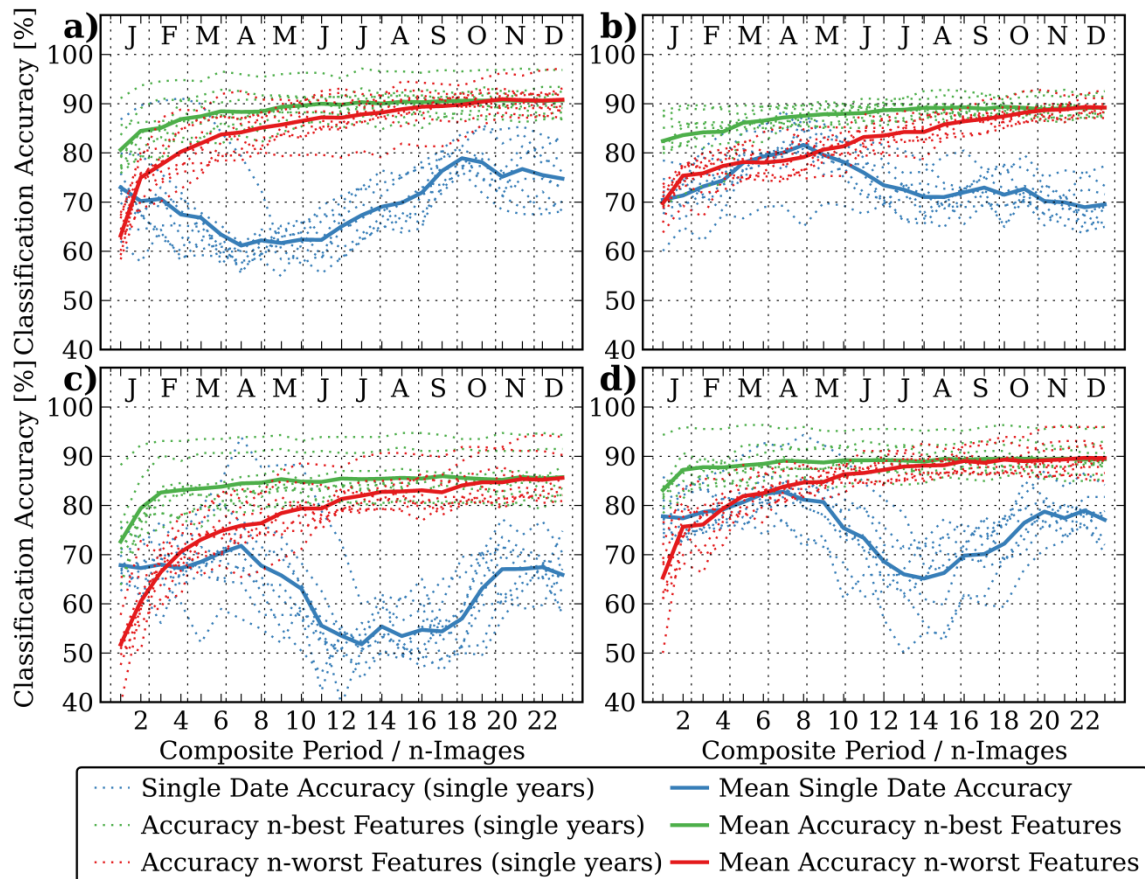


Figure 9: Classification accuracies of class combinations Grassland (G) vs. Forest (F) and Improved Grassland (GA) vs. Semi-Improved Grassland (GS). Mono-temporal (green), Multi-temporal Classification accuracies – best number of features first (red), worst number of features first (blue). a) NDVI – G vs. F; b) EVI G vs. F; c) NDVI – GA vs. GS; d) EVI – GA vs. GS

5. Conclusion

The optimisation of image acquisition dates for classification of vegetated landscapes in Ireland using the internal Feature Importance measures of the state-of-the art machine-learning method Random Forest proved to be a useful tool for removing redundancy within an annual time series and maximising the classification accuracy with minimal image input. Their temporal development was mostly reflected in the seasonal variability of the complementary classification accuracies. However, their normalised values only indicate the relative separability.

By exploiting variations in the annual phenological cycle of vegetation in Ireland it is apparent that NDVI data from late autumn/early winter are optimal for generic land-cover classification, whereas EVI data performed best from winter to spring (January-May). Summer images, by contrast, showed the weakest separability for all class combinations and both vegetation indices. Using NDVI data, October proved to be the best acquisition period for discrimination of Grassland from Forest, but for EVI the maximum separability for this pair occurs around April. Typically the NDVI values were higher for forest regions than grass although with a high degree of overlap, however EVI demonstrated a more distinct range of values, with the grassland consistently higher than the forest. During periods of typical Irish maritime meteorological conditions, EVI consistently produced more accurate, stable classification results for this class-combination. Following a harsh winter more typical of continental climatic regimes, a strongly reduced separability for EVI was evident, with NDVI displaying a highly increased distinction between grassland and forest, caused by the strongly decreased NDVI of Grassland.

The classification results highlighted interesting characteristics of the different vegetation indices and classes. For classes with high photosynthetic activity, such as Improved and Semi-Improved Grassland, EVI proved to be better able to distinguish the subtle differences in land cover, due to a higher sensitivity and stability in this range. However, the very similar seasonal trends for Settlement and Peatland in the EVI dataset suggest that the NDVI values are more suitable for discriminating between these less photosynthetically active classes. For optimal classification results therefore, consideration needs to be given to the photosynthetic activity of the classes, and potentially both indices used. Finally, the number of images required for achieving the optimal classification accuracies, in excess of 90%, proved to be between 6 and 10, after which the value gained from additional images becomes marginal.

Land cover classification from images acquired on multiple dates has become widely adopted as best practice for distinguishing between different vegetation habitats, with recognition that the phenological signal is the driver for selection of those dates. This work highlights the time periods when the phenological signals are most distinct for different land cover classes and using different vegetation indices, with the results having implications for all land cover mapping applications at regional or global scales requiring medium-low spatial resolution, and for countries with limited temporal data due to extensive cloud cover or other restrictions.

6. References

- Atzberger, C. & Eilers, P. H., 2011. A time series for monitoring vegetation activity and phenology at 10-daily time steps covering large parts of South America. *International Journal of Digital Earth*, 4(5), 365-386.
- Beck, P. S., Atzberger, C., Høgda, K. A., Johansen, B. & Skidmore, A. K., 2006. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment*, 100(3), 321–334.

- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J., 1984. Classification and regression trees. Wadsworth & Brooks. Monterey, CA.
- Breiman, L., 2001. Random Forests. *Machine Learning*, 45, 5-32.
- Carrão, H., Gonçalves, P. & Caetano, M., 2008. Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment*, 112(3), 986-997.
- Carrão, H., Araujo, A., Gonçalves, P. & Caetano, M., 2010. Multitemporal MERIS images for land-cover mapping at a national scale: a case study of Portugal. *International Journal of Remote Sensing*, 31(8), 2063-2082.
- Central Statistics Office Ireland, 2011. Population of each Province, County and City, 2011. < <http://www.cso.ie/en/statistics/population/populationofeachprovincecountyandcity2011> > , accessed on 24th June 2013.
- Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B. & Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment*, 91(3–4), 332-344.
- Colditz, R., 2007. Time Series Generation and Classification of MODIS Data for Land Cover Mapping, PhD thesis.
- Díaz-Uriarte, R. & Alvarez de Andrés, S., 2006. Gene selection and classification of microarray data using random forest. *BMC bioinformatics*, 7(1), 3.
- Eaton, J.; McGoff, N.; Byrne, K.; Leahy, P. & Kiely, G., 2008. Land cover change and soil organic carbon stocks in the Republic of Ireland 1851–2000. *Climatic Change*, 91, 317-334.
- Guo, B.; Damper, R.I.; Gunn, S.R. & Nelson, J.D.B., 2008. A fast separability-based feature-selection method for high-dimensional remotely sensed image classification. *Pattern Recognition*, 41(5), 1653–1662.
- Hermance, J. F., Jacob, R. W., Bradley, B. A. & Mustard, J. F., 2007. Extracting Phenological Signals from Multiyear AVHRR NDVI Time Series: Framework for Applying High-Order Annual Splines with Roughness Damping. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10), 3264-3276.
- Hill, M. J., Vickery, P. J., Furnival, E. & Donald, G. E., 1999. Pasture Land Cover in Eastern Australia from NOAA-AVHRR NDVI and Classified Landsat TM. *Remote Sensing of Environment*, 67(1), 32-50.
- Hodrick, R. J. & Prescott, E. C., 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, 29(1), pp. 1-16.

Huete, A., Didan, K., van Leeuwen, W., Miura, T., & Glenn, E., 2011. MODIS vegetation indices. In *Land Remote Sensing and Global Environmental Change* (pp. 579-602). Springer New York.

Jonsson, P., & Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *Geoscience and Remote Sensing, IEEE Transactions on*, 40(8), 1824-1832.

Lhermitte, S., Verbesselt, J., Verstraeten, W. & Coppin, P., 2011. A comparison of time series similarity measures for classification and change detection of ecosystem dynamics. *Remote Sensing of Environment*, 115(12), 3129-3152.

Liu, H., Dougherty, E., Dy, J., Torkkola, K., Tuv, E., Peng, H., Ding, C., Long, F., Berens, M., Parsons, L., Zhao, Z., Yu, L. & Forman, G., 2005. Evolving feature selection. *Intelligent Systems, IEEE* 20(6), 64-76.

Lunetta, R. S., Johnson, D. M., Lyon, J. G., & Crotwell, J., 2004. Impacts of imagery temporal frequency on land-cover change detection monitoring. *Remote Sensing of Environment*, 89(4), 444-454.

Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G. & Worthy, L. D., 2006. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment*, 105(2), 142-154.

Met Éireann, 2010. Monthly Weather Bulletin No. 286 February 2010.

Met Éireann, 2013. Climate of Ireland. < <http://www.met.ie/climate-ireland/climate-of-ireland.asp>>, Accessed on 24th June 2013.

O'Connor, B., Dwyer, E., Cawkwell, F. & Eklundh, L., 2012. Spatio-temporal patterns in vegetation start of season across the island of Ireland using the MERIS Global Vegetation Index. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68, 79-94.

O'Neill, F.H., Martin, J.R., Perrin, P.M., Delaney, A., McNutt, K.E. & Devaney, F.M., 2009. Irish Semi-natural Grasslands Survey - Annual Report No. 2: Counties Cavan, Leitrim, Longford & Monaghan.

Pal, M., 2006. Support vector machine based feature selection for land cover classification: a case study with DAIS hyperspectral data. *International Journal of Remote Sensing*, 27(14), 2877-2894.

Pal, M. & Foody, G., 2010. Feature Selection for Classification of Hyperspectral Data by SVM, *IEEE Transactions on Geoscience and Remote Sensing*, 48(5), 2297-2307.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D.,

- Brucher, M., Perrot, M. & Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Pringle, M., Denham, R. & Devadas, R., 2012. Identification of cropping activity in central and southern Queensland, Australia, with the aid of MODIS MOD13Q1 imagery. *International Journal of Applied Earth Observation and Geoinformation*, 19, 276-285.
- Sellers, P. J., Tucker, C., C. J., Collatz, G. J., Los, S. O., Justice, C., C. O., A., D. D. & Randall, D. A., 1994. A global 1° by 1° NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *International Journal of Remote Sensing*, 15(17), 3519-3545.
- Serpico, S. & Bruzzone, L., 2001. A new search algorithm for feature selection in hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), 1360-1367.
- Stumpf, A. & Kerle, N., 2011. Object-oriented mapping of landslides using Random Forests. *Remote Sensing of Environment*, 115(10), 2564-2577.
- Van Niel, T. G. V., McVicar, T. R. & Datt, B., 2005. On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification. *Remote Sensing of Environment*, 98(4), 468-480.
- Wardlow, B. D. & Egbert, S. L., 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment*, 112(3), 1096-1116.
- Xiao, X., Braswell, B., Zhang, Q., Boles, S., Frohling, S. & III, B. M., 2003. Sensitivity of vegetation indices to atmospheric aerosols: continental-scale observations in Northern Asia. *Remote Sensing of Environment*, 84(3), 385-392.
- Zurita-Milla, R., Kaiser, G., Clevers, J., Schneider, W. & Schaepman, M., 2009. Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sensing of Environment*, 113(9), 1874-1885.