Irish Geography, 2013
Vol. 46, Nos. 1–2, 25–49, http://dx.doi.org/10.1080/00750778.2013.845489

The implications of cloud cover for vegetation seasonality monitoring across the island of Ireland using the MERIS Global Vegetation Index (MGVI)

Brian O'Connor[a,b]*, Edward Dwyer[a] and Fiona Cawkwell[b]

[a]*Coastal and Marine Research Centre (CMRC), Environmental Research Institute, University College Cork, Cork, Ireland;* [b]*Department of Geography, University College Cork, Cork, Ireland*

(*Received 16 July 2012; final version received 15 July 2013*)

Seven years of the MERIS Global Vegetation Index (MGVI) data have been obtained for a national-scale study of vegetation seasonality in Ireland from 2003 to 2009. The selection of an appropriate composite period for the daily MGVI data was guided by *in situ* observations of vegetation spring greening and cloud cover from two representative point locations across the island. A period of 10 days was selected as an optimum, minimising the amount of cloud cover across the island while still capturing vegetation seasonality change. Short-term variation in the MGVI time series after time-compositing had been applied was found to be unrelated to vegetation dynamics, suggesting that external factors, such as cloud cover compromise the quality of daily MGVI values. A verification study, using the METEOSAT Cloud Mask (CLM), was conducted to validate this hypothesis. The results suggest that for 7 out of 10 MGVI images over half the values may be in error due to the presence of cloud cover, indicating that the MERIS cloud screening approaches are sub-optimal for conditions experienced over Ireland. A review of the MGVI atmospheric model indicates that MERIS atmospheric corrections may only partially correct for scattering by aerosols or absorption by water vapour.

Keywords: MERIS Global Vegetation Index; vegetation seasonality; cloud; satellite monitoring; METEOSAT; Ireland

Introduction

Both plant-specific observations and regional-scale studies of vegetation by multi-annual time series of satellite imagery have been combined to give a synoptic view of land surface vegetation (Reed *et al*. 1994, Reed and Brown 2005, Studer *et al*. 2007). The most commonly used satellite-derived measure of vegetation growth to date has been the Normalised Difference Vegetation (NDVI) which has found a variety of applications in studies of the biosphere (Stöckli and Vidale 2004, Xiao and Moody 2005, Julien and Sobrino 2009). Other commonly used vegetation indices include the Soil Adjusted Vegetation Index (Huete 1988), the Global Environmental Monitoring Index (Pinty and Verstraete 1992) and the Perpendicular Vegetation Index (Campbell 2002) which vary in performance and application. However, vegetation indices have now advanced beyond the red/near infrared difference ratio on which the NDVI and other indices are based. One such measure derived using a physically based approach by means of radiative transfer

*Corresponding author. Email: Brian.O'Connor@unep-wcmc.org

modelling is the MERIS Global Vegetation Index (MGVI) which has improved geometric error correction, atmospheric interference reduction and greater sensitivity to seasonal vegetation dynamics than empirically based indices such as the NDVI (Govaerts *et al*. 1999). Gobron *et al*. (1999) describe the physical and mathematical basis of the MGVI algorithm. The Medium Resolution Imaging Spectrometer (MERIS) was an oceanic and terrestrial monitoring sensor aboard the Envisat platform, which operated from 2002 to 2012 and achieved global coverage in three days. Reflectance was measured in 15 spectral bands, of which, bands 5 (green), 8 (red) and 13 (near infrared) are used in vegetation monitoring. Band 2 (blue) is sensitive to Rayleigh scattering in the atmosphere and is used for atmospheric corrections. The MGVI is calculated from reflectance data in the blue, red and near infrared MERIS bands and has been optimised to estimate the Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), a biogeophysical measure of vegetation growth (Gobron *et al*. 1999).

Noise in NDVI time series and its impact on the extraction of metrics for the study of vegetation seasonality has been well documented (Hird and McDermid 2009). Multiple factors affect the satellite-measured reflectance creating noise in time series. Molecules of nitrogen and oxygen gas in the atmosphere cause Rayleigh scattering, the intensity of which increases at shorter wavelengths, e.g. in the blue spectrum of visible light (Campbell 2002). Oxygen, ozone, other trace gases and water vapour result in signal absorption while aerosols can cause both scattering and absorption (Holben 1986). In addition, instrument precision, calibration and off-nadir viewing create deviations in observations unrelated to vegetation dynamics (Goward *et al*. 1991). For these reasons, satellite-measured reflectance can be altered even in a cloudless atmosphere (Holben 1986).

Optical satellite sensors measure reflectance from the Earth in the same spectral region as clouds absorb and scatter radiation, i.e. the visible to near-infrared spectrum (Gomez-Chova *et al*. 2007). Therefore, cloud cover poses a significant obstacle for monitoring vegetation growth from space-borne sensors. This is of particular relevance to Ireland, where cloud cover is more frequent and persistent than in continental Europe due to the regular flow of warm humid air from the Atlantic Ocean (Rohan 1986). Both ground-based and satellite observations show that monthly average cloud cover amounts vary seasonally over Ireland, but that minimum cloud cover is in the April-May period (Rohan 1986, Pallé and Butler 2001).

Long time series of continuous satellite-derived geophysical products are important for monitoring vegetation dynamics over seasonal to annual time scales. However, time series are frequently punctuated by data gaps when clouds prevent a valid measurement and are affected by noise due to atmospheric components like aerosols (Colditz *et al*. 2008). Furthermore, atmospheric conditions show considerable spatio-temporal variation making it difficult to acquire daily data over the mid-latitude temperate areas on a consistent basis. Time-compositing overcomes this problem by selecting a measure which best represents the state of the surface over the compositing interval. The length of the interval is determined by the size of study area, the frequency of cloud cover and the surface phenomena being monitored. For example, while a 4-day interval may be appropriate for local-scale applications, a global study may employ a 4-week period to attain cloud-free conditions worldwide (Justice *et al*. 1985, Yang *et al*. 2006).

The MGVI undergoes two distinct cloud screening stages; firstly the MERIS cloud mask removes cloud pixels from the raw MERIS imagery from which the remaining cloud-free land pixels are screened by the MGVI cloud-detection algorithm (Gobron *et al*. 2004). The influence of atmosphere on the MERIS reflectance measurements has been lessened due to the optimisation of band widths for atmospheric windows and dynamic atmospheric correction of their values using blue band data. Despite these safeguards, uncertainty in the MERIS measurements is still likely to occur. This is evident from a visual analysis of the MGVI time series where the transition between adjacent data points is not smooth. An example of the instability in MGVI time series is given in Figure 1 (blue line). This feature of the time series is not restricted to pixels of particular landcover types, location or any particular year, but is common to all pixels across the image. Although a smoothing function can be fit to the raw time series data (black line), noise in the underlying time series may still detract the ability of accurately determining vegetation seasonality parameters (Eklundh and Jönsson 2010).

The main aim of this study is to investigate what composite period would be appropriate to monitor vegetation seasonality across the island of Ireland using the reduced spatial resolution MGVI product. This study aimed to guide the selection of an appropriate composite period using ground-based observations of vegetation seasonality and cloud cover at representative point locations on the island. A second aim is to investigate the cause of noise in the MGVI time series following the application of a time composite algorithm. It was first hypothesised that undetected cloud cover was present in the MERIS reflectance data used to construct the MGVI. A validation study using an independent source of satellite-derived cloud data was used to test this assumption. Secondly, other components of the atmosphere over Ireland which might affect the quality of the MGVI were considered. For this, the MGVI atmospheric model was reviewed in order to understand the link between the various absorption and scattering processes of the atmosphere and noise in the MERIS reflectance data used to construct the MGVI.

Methodology

The selection of an appropriate compositing period

A strategy for composite period selection was designed to determine the minimum number of days needed to monitor vegetation seasonality across the island by sufficient cloud-free data. The following two criteria have been used as a guide for selecting the composite period:

(1) The length of the composite period should be minimised to capture the vegetation start of season which corresponds to the beginning of leaf unfolding, a ground-based measure routinely observed in phenological observatories. This

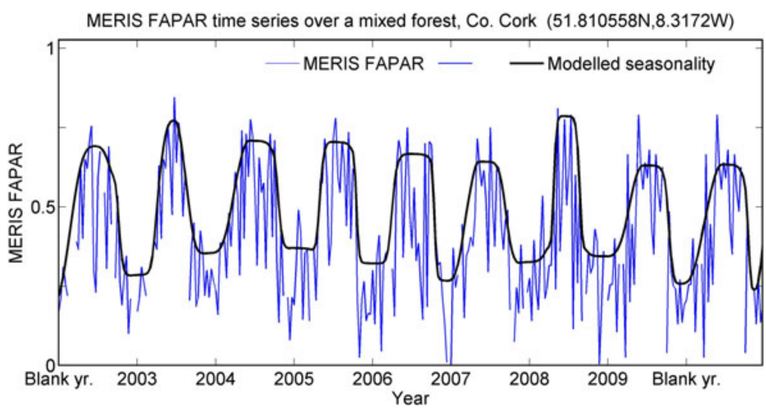

Figure 1. The blue line represents noisy MGVI time series and the black line shows the fitted smoothing function which models the seasonality in the time series.

was supported by observations of the rate of spring greening in a selection of tree species at Currabinny Wood, Co. Cork whose location can be seen in Figure 2 (51.812°N, 8.300342°W).

(2) The composite period should be long enough to maximise the number of cloud-free days per pixel within the composite. The MGVI pixel cloud flags were used as a means for assessing the number of potential cloud-free days in a given composite period.

In-situ phenological data

The main aim of the field campaign was to investigate the temporal dynamics of spring greenup in mixed woodland vegetation. In particular, it was intended to investigate how

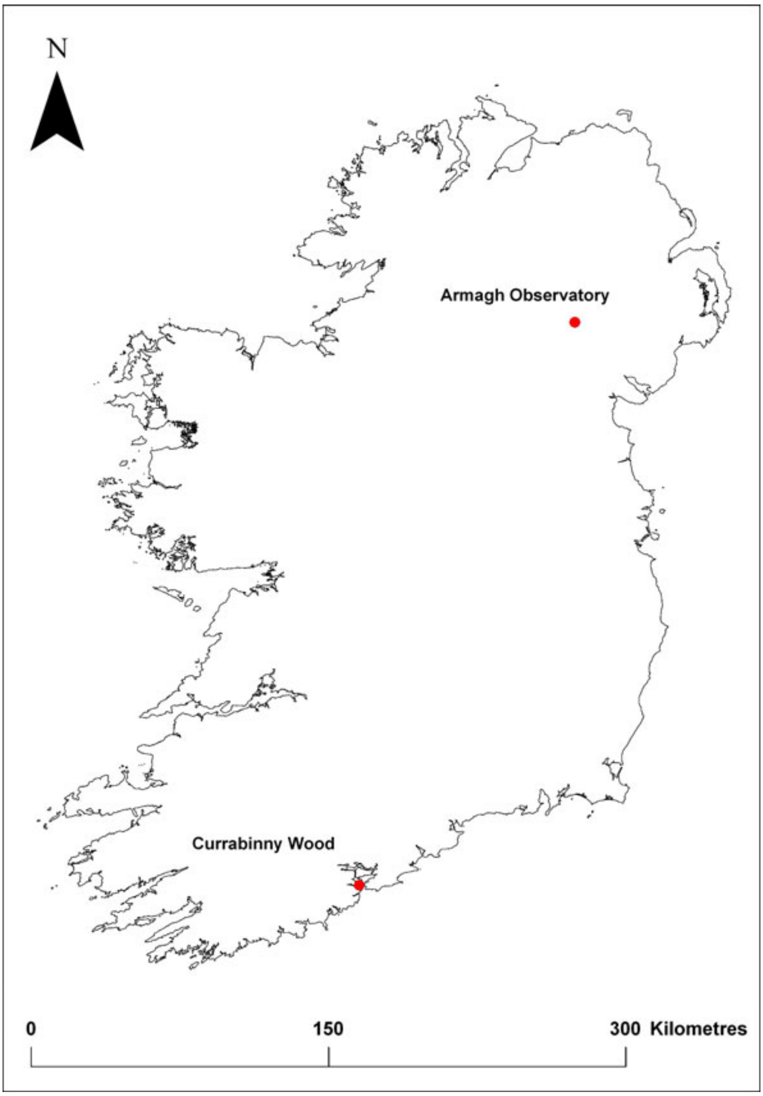

Figure 2. The location of point observations of cloud cover (Armagh) and spring greening (Currabinny) used to guide the selection of an appropriate composite period.

frequently observations need to be made, so that the rate of change of phenological stages from dormancy to budburst to leaf out could be monitored across different tree species. In order to achieve this aim, both the budburst and beginning of leaf unfolding phenological stages were observed from late winter to spring in an area of mixed forest at Currabinny Wood.

The Beginning of Leaf unfolding (BO) is one of the most commonly recorded phases at the International Phenological Gardens and is recorded with the appearance of several leaves across the tree canopy (IPG 2005). However, such a small amount of leaf unfolding would not be comparable to satellite-derived measures of the beginning of growing season. Therefore, the start of the growing season was recorded when 50% of the tree crown was estimated to have unfolded leaves. The date of site visit when 50% of the leaves had unfolded in each tree was noted, as well as the percentage of leaves unfolded in the canopy on every subsequent site visit. Site visits continued after the event had occurred until the canopies had reached maximum leaf cover. A sample of eighteen trees from six different species, three trees per species, was selected and monitored on a weekly basis from March to May of 2008 and from late February to May 2009. The tree species sampled were Oak (*Quercus robar*), Ash *(Fraxinus excelsior*), Horse Chestnut (*Aesculus hippocastanum*), Sycamore (*Platanus occidentalis*), Beech (*Fagus Sylvatica*) and Birch (*Betula pubescens*). The deciduous species were chosen for monitoring as the budburst and leaf out phenological stages are more readily observed than the equivalent phenological stages in coniferous species (Kubin *et al*. 2007). Photographs of the greening stages of some of the observed trees were taken during every site visit for supporting the estimates of the timing of budburst, leaf unfolding and the estimate of percentage leaf out. Photographs were taken of the same tree from the same position each week for further consistency.

Armagh Observatory cloud data

Human observer networks have been used to record a number of cloud variables, such as percentage cloud cover, to validate satellite observations of cloud cover, as well as a means to understanding daily cloud cycles (Chambers *et al*. 2004). Therefore, ground observations of cloud cover were used to investigate inter-annual trends in cloud cover during the growing season (February to October), as well as cloud cover variation during the growing season. The cloud data were obtained free of charge from the Armagh Observatory (Armagh Observatory 2011). The Armagh data were not intended to represent cloud conditions at the site of phenological observations but were considered a means to understand temporal variation in cloud cover on an annual and seasonal basis. The inland location of the Observatory, shown in Figure 2, is away from any coastal and mountainous influences in cloud formation. The observations taken daily at 9 a.m. have been recorded for the entirety of the twentieth to twenty-first century and are archived online, where they are available to download freely for general use (Armagh Observatory 2011). Cloud cover is estimated on a scale from zero to eight okta, where eight okta represents a fully-clouded sky. The current observer has been recording observations since 1998.

The daily observations were first averaged over 7-day intervals (equivalent to those used for field visits) for three full growing seasons from 2005 to 2007. Secondly, the averaging interval was extended to 10 days in order to observe the effect of using a longer averaging period on annual cloud trends.

Cloud cover analysis

The daily Armagh cloud observations were averaged to investigate the effect of extending the averaging interval on temporal cloud patterns. The year on which to base the selection of the composite period was chosen from a visual inspection of the averaged cloud patterns in 2005, 2006 and 2007, and monthly weather summaries of synoptic weather station data around the country (Met Éireann 2009). This assisted in refining the appropriate composite period for the daily satellite data. A time series of the MGVI was obtained by using the FAPAR time-composite algorithm which selects a daily value closest to the mean of daily values in the composite period (Pinty *et al*. 2002). The algorithm has been integrated with the processing capabilities of the European Space Agency's Grid-Processing On Demand (G-POD) service to produce spatially continuous gridded MGVI data over areas specified by the user (European Space Agency 2011). Seven years of reduced spatial resolution (1.2 km) MGVI from 2003 to 2009 were obtained within a geographical window over Ireland, defined by latitude 51–56°N and longitude 5–11°W, from the G-POD service. To conclude the analysis, the MGVI imagery was composited in the same intervals used for averaging the Armagh data. The percentage of image pixels that was cloud-covered in the different compositing periods was compared and used to select the most appropriate composite period. To show the spatial pattern in data loss due to cloud cover across the island, cloud composite images were also generated from the cloud flag data.

The verification of valid MGVI using the METEOSAT Cloud Mask (CLM)

It was hypothesised that the MGVI time series was noisy because undetected cloud created anomalous pixel values. This assumption was supported by previous concerns which have been raised about the effectiveness of MERIS cloud screening, e.g. thin clouds are not consistently detected over land, while ice and snow are frequently mistaken for cloud (MERIS 2006). These shortcomings potentially create noisy time series of MERIS-derived products, such as the MGVI (Baret *et al*. 2006). In order to examine the extent to which these issues may affect the quality of daily MGVI values composited over Ireland, and thus the reliability with which phenological information can be extracted, the valid MGVI pixels were compared to the METEOSAT Cloud Masks (CLM). The METEOSAT SEVIRI (Spinning Enhanced Visible Infra-Red Imager) instrument was chosen as it has been designed solely for meteorological applications, and therefore has the potential to detect types of cloud which the MERIS sensor cannot. For instance, the MERIS spectral range (0.4–0.9 μm) covers the visible to near infrared spectrum in narrow bandwidths, while the SEVIRI instrument covers the same range, as well as the water vapour absorption region (5.7–7.1 μm) but in wider intervals. Furthermore, the SEVIRI cloud detection algorithm uses a combination of thirty-four threshold tests, combined with weather forecast data and radiative transfer model outputs (EUMETSAT 2007). This is a more complex cloud-detection strategy than the simple threshold test on which the MGVI cloud-detection algorithm is based.

In order to determine if the spiked values occurred because of undetected cloud cover, the METEOSAT Cloud Mask (CLM) was acquired from the Meteorological Archive and Retrieval Facility (EUMETSAT 2011). The CLM was selected for this work as it provides the required information on the presence or absence of cloud within a pixel, has the finest grid resolution of the METEOSAT cloud products (3 km at equatorial latitudes, approximately 6 km at Irish latitudes) and, since the launch of Meteosat-8 in 2006, data are acquired every fifteen minutes (according to D. Richards, personal communication, 3 November 2009).

In order to investigate whether the variation in the MGVI time series was caused by undetected cloud cover, ten MGVI pixels, spread across the image grid, were selected from five MGVI composites in 2006. This produced a sample of fifty MGVI values, which were assessed as being noisy if their values produced spikes in the time series. This assessment was carried out as a process of visual inspection whereby a margin of variability between adjacent MGVI values was permitted to accommodate natural variation in the vegetation photosynthetic cycle. Nevertheless, a limit had to be defined in order to identify noisy spikes from this natural variation. A positive (upward) spike was identified as a value that exceeded the mean of the preceding and subsequent value by more than 20%, and a negative (downward) spike identified when the pixel value was less than the mean by more than 20%. A value was considered as normal when it was within the mean ±20% range. This threshold was established based on a pragmatic approach to spike identification coupled with an understanding of the natural versus artificial variability in the MGVI time series following visual inspection. Furthermore, there has not been any previous documented attempt to threshold for noise in MGVI time series. This process of spike identification is illustrated graphically in Figure 3.

The five MGVI composite periods were selected at various times in 2006 in order to capture the annual variation in cloud cover across the island. The MGVI composite dates were 21–30 January, 21–30 April, 20–29 June, 28 September–7 October and 27 November– 6 December 2006. A contemporaneous sample of fifty daily CLM grids was also acquired. The pixel values of the composite were derived only from daily values identified as cloud free, and if no pixels within the composite period met that criterion a no-data value was returned. Daily images of cloud free pixels were created for each day within the composite. They were compared with the CLM pixels, which consisted of land, ocean and cloud, for the same day and time. In order to ensure consistency in timing between the MERIS and METEOSAT observations, the exact over-pass time of the MERIS sensor was obtained for a pixel in central Ireland. This information was derived from the solar zenith and azimuth angles provided for each MGVI pixel and

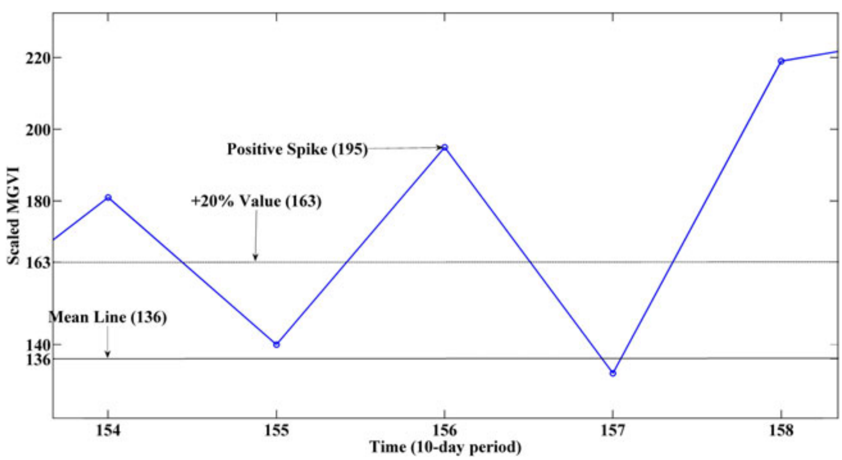

Figure 3. The composite value at period 156 is anomalous with respect to the left and right neighbour forming an upward spike. The mean line is calculated from the left (155) and right (157) neighbouring values while the +20% value is also indicated.

corresponding time of day (UTC) calculated using the NOAA solar calculator (Cornwall *et al*. 2010). With METEOSAT images acquired every 15 minutes, the most coincident image was selected, with the difference in time between the two compared images not greater than 7.5 minutes. The METEOSAT cloud mask was remapped and resampled to a 1.2-km grid, equivalent in size to the MGVI grid. For each day within the composite period, both the valid and invalid MGVI values were extracted and compared with their CLM counterparts. A resulting image was produced indicating which of the five possibilities of pixel pairs was present, shown in Table 1.

To summarise the results of the MGVI verification, numbers were assigned to the results of each comparison test. A positive test result was generated for pixels with a valid MGVI value coinciding with a METEOSAT land value (1 in the comparator image) and pixels with an invalid MGVI value coincident with a METEOSAT cloud value (2 in the comparator image). Other results were deemed negative, that is, an MGVI value is produced for a cloud covered pixel (3 in the comparator image) or no MGVI is produced for a clear-sky pixel (4 in the comparator image). Pixels for which the MGVI and METEOSAT pixels were over the ocean were assigned a value of 5 in the comparator image. Figure 4 (a)–(b) illustrates an example of the remapping output and the results of the comparison tests for 21 January 2006. This date was selected, as it was a relatively clear day across the island with a high number of the time-composited values for the 10-day period, 21–30 January, chosen from this day. The clear sky values extracted from the MERIS data are shown in Figure 4(a), and the comparison with the remapped CLM is displayed in Figure 4(b). A small number of MGVI pixels remain outside the extent of the CLM grid, and therefore have a value of 0. All 50 of the METEOSAT dates were compared with the MERIS data and the proportion of positive and negative test results for each date calculated.

Results and discussion

The selection of the MGVI compositing period

Weekly field visits were timed to monitor changes in the phenological stages of trees as there was no noticeable change on visiting twice a week, while on a 14-day interval, a phenological event, e.g. budburst, could go unrecorded and large increases in the percentage leaf out remaining unobserved. Changes in the leaf canopy, in terms of the percentage of leaf unfolding, as well as specific phenological stages, were easily observed on a weekly basis. Key phenological stages can be seen in Figure 5(a)–(c) from the weekly photographs of a Birch *(Betula pubescens)* tree at the field site from budburst (a), 50% leaves unfolded (b) and 100% leaves unfolded (c).

There was sufficient change in the leaf canopy on a weekly basis to suggest that a 7-day composite period would also be optimal to track the start of the growing season from MGVI time series data. However, seasonal and inter-annual differences in the cloud

Table 1. New pixel values assigned based on the comparison of coincident METEOSAT and MGVI pixels (note that an ocean mask prevents MGVI being calculated over open water).

		METEOSAT CLM		
		Ocean	Land	Cloud
MGVI	Valid		1	3
	Invalid	5	4	2

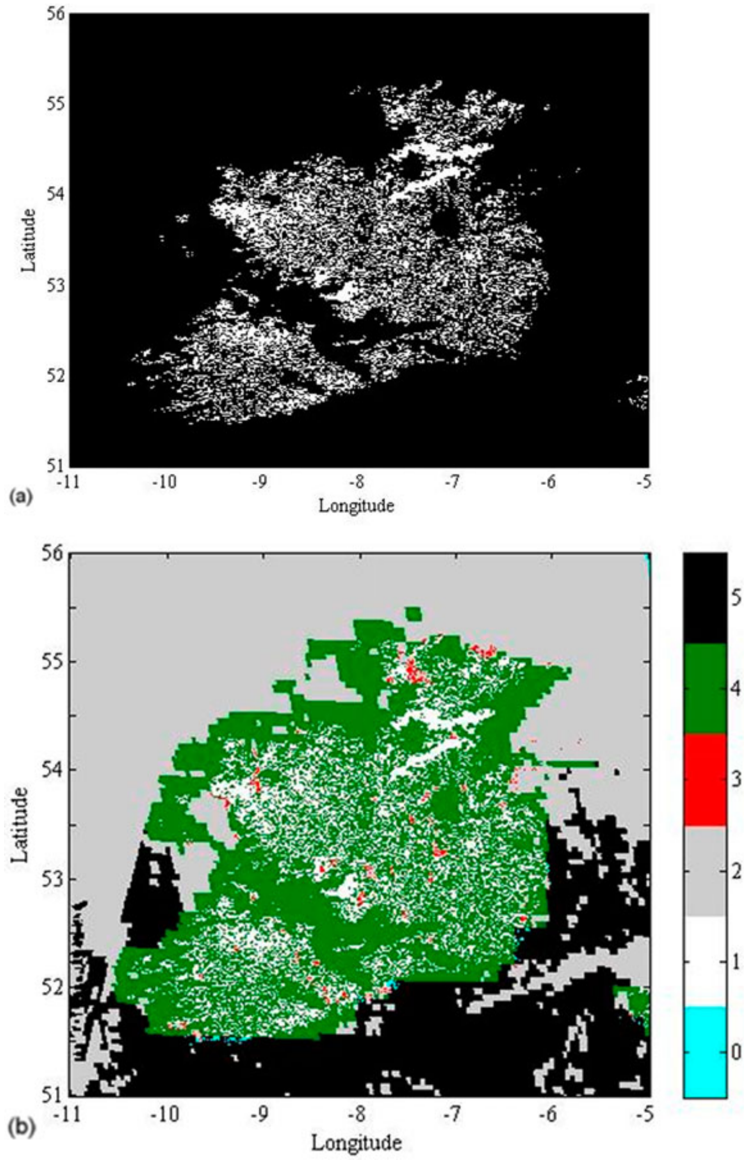

Figure 4. (a) The MERIS data from 21January 2006 with clear sky MGVI pixels shown in white and invalid pixels in black. (b): The 21January 2006, result from the comparison of coincident MGVI and METEOSAT data. The values correspond to the pixel validity and surface types as defined in Table 1. For full colour version of this figure please refer to the online version of the article at http://www.tandfonline.com/rigy

cover pattern at Armagh were indistinguishable from short-term variability when the daily values were averaged over 7-day intervals. Therefore, the averaging period was extended to ten days due to the excess of cloud cover in a 7-day interval. The amount of cloud cover at Armagh, averaged over both 7-day and 10-day intervals in 2005, 2006 and 2007 is shown in Figure 6(a)–(b). Lower than average cloud amounts were present across the country from the beginning of April to the beginning of June 2007. This made it an unsuitable year for basing the selection of a composite period. Therefore, the year 2006

Figure 5. (a)–(c) Three stages in the seasonality of a Birch tree (*Betula pubescens*), Currabinny wood, Co. Cork, 2009- (a) budburst, (b) 50% leaf out and (c) 100% leaf out.

Figure 5. (*Continued*)

was chosen and the period from 1 March to 29 June selected as it corresponded to the period of leaf unfolding as shown by fieldwork investigations.

A threshold of 10% cloud-covered land pixels was chosen as a simple benchmark to assess the performance of the two composite periods in minimising the amount of data lost due to cloud in national-scale imagery. Although a subjective threshold, it was intended as a relative measure of composite performance and selecting a lower or higher threshold would not be expected to greatly influence the outcome of the comparison exercise.

In Figure 7, less than 10% cloud cover was achieved on 77% of the 10-day composites, compared with 50% of the 7-day composites. This suggests that there will be fewer data gaps due to cloud in 10-day imagery than in 7-day imagery.

To illustrate the spatial distribution of data loss due to cloud in each composite period, binary images of the cloud and non-cloud pixels for each image composite were created using the cloud flag information. The number of cloud flags per pixel in the time-series of eighteen 7-day and thirteen 10-day composite images from 1 March to 29 June 2006 was then calculated. The results can be seen in Figure 8(a) and (b).

Both images have a maximum of six cloud flags per pixel over the test period, that is, the pixel contained no valid data for six composite periods. However, the percentage of cloud-covered periods per pixel was quite different. A frequency distribution of the number of cloud flags reported per pixel during the test period was derived from the cloud composites in Figure 8(a) and (b) and is shown in Figure 9.

Figure 9 illustrates the fact that shorter composite periods incur greater data loss due to cloud cover. For example, 52% of the 10-day cloud composite image pixels had no cloud-covered periods compared to only 32% of the pixels in the 7-day image composite.

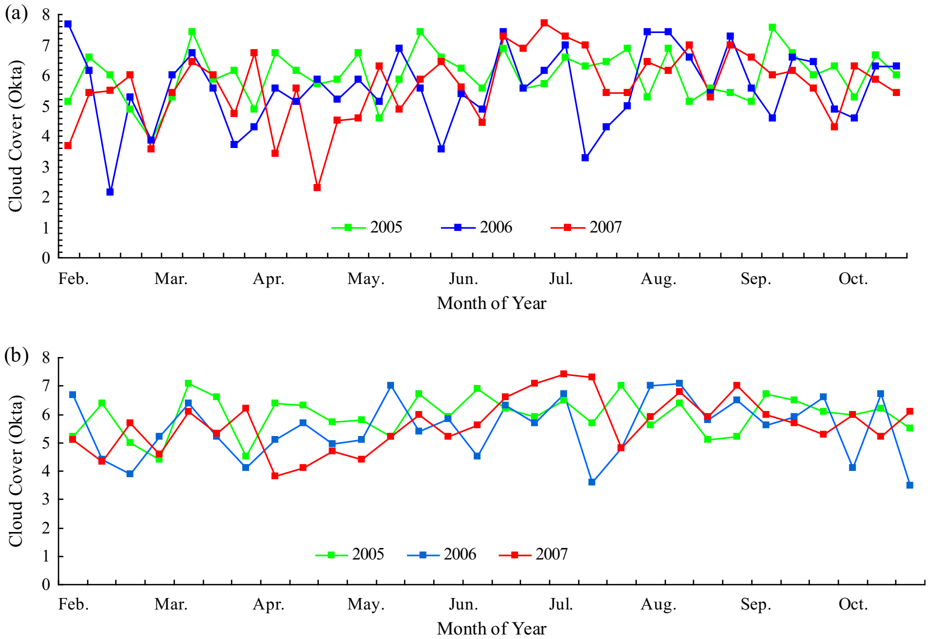

Figure 6. (a) 7-day and (b) 10-day average cloud cover amounts (Okta) observed at Armagh Observatory for the growing season of 2005 to 2007.

Furthermore, 22% of the 7-day cloud composite pixels had two cloud-covered periods compared to only 12% of the 10-day cloud composite pixels. Overall, there were fewer pixels with one or more cloud-covered periods in the 10-day image composites than there were in the 7-day composites. Therefore the 10-day composite period contained more cloud-free periods than the 7-day period from March to June, 2006 and was selected as an appropriate composite period for the study. Extending the period would undoubtedly generate further cloud-free data; however, this would potentially compromise the aim of

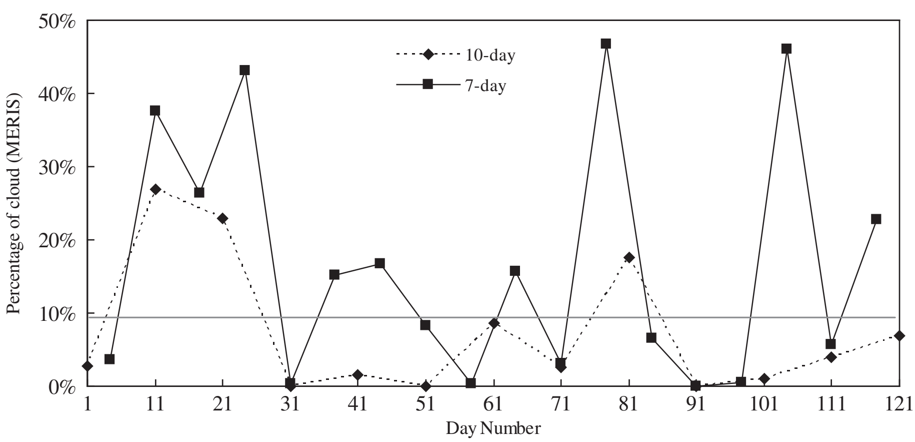

Figure 7. Line graph of the percentage of cloud-covered MGVI pixels over the island of Ireland per 7-day and 10-day composite period during spring 2006 (Day 1 represents March 1st and day 121, June 29th). The grey line marks the 10% threshold of acceptable data loss due to cloud.

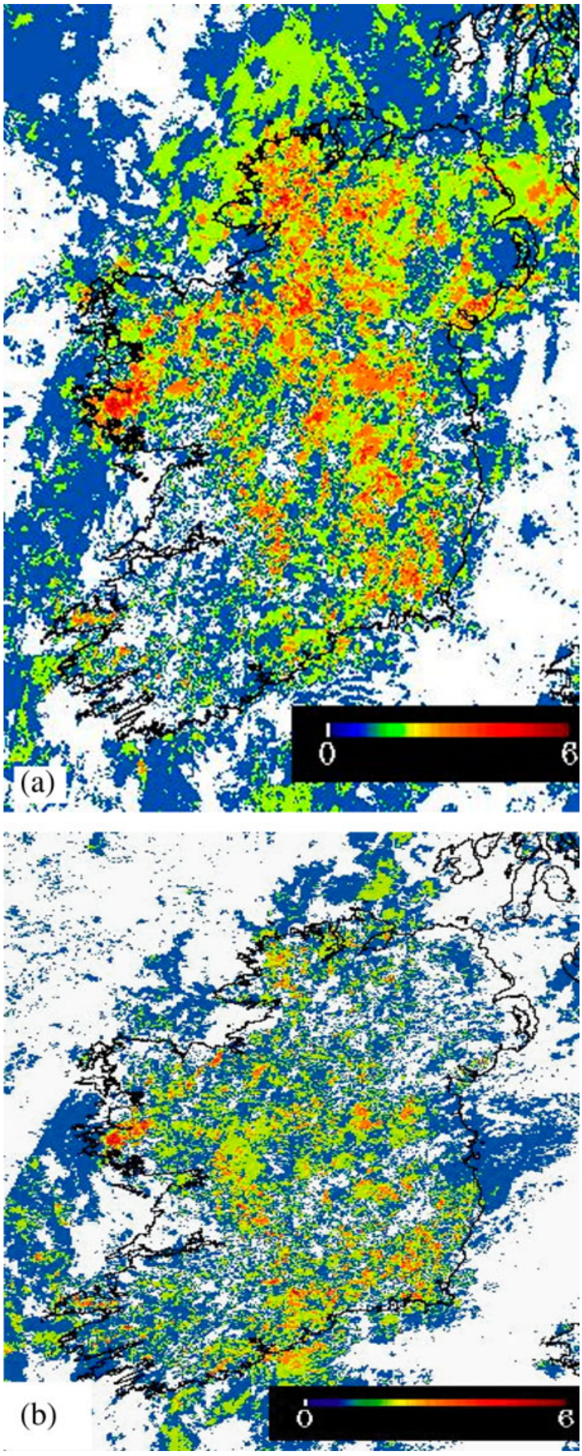

Figure 8. Cloud composite images showing the number of cloud-covered periods per pixel over the island of Ireland in the period, March 1st to June 29th, 2006 (a) using 7-day composites and (b) using 10-day composites. For full colour version of this figure please refer to the online version of the article at http://www.tandfonline.com/rigy

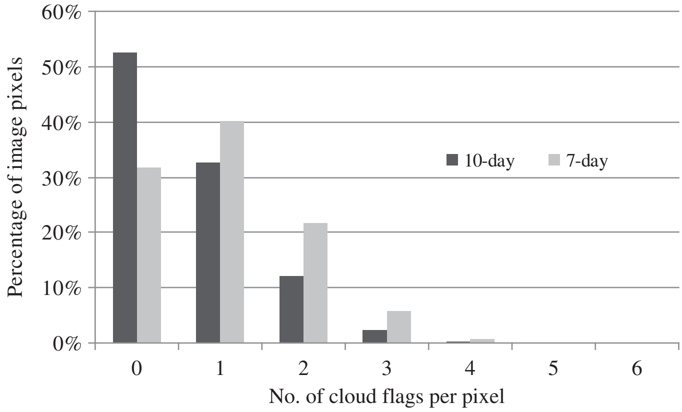

Figure 9. A frequency distribution of the number of cloud flags per pixel in a time series of eighteen 7-day and thirteen 10-day composite images in the period, March 1st to June 29th, 2006.

the study, to establish the minimum composite period required to track the seasonality of terrestrial vegetation across Ireland.

In order to investigate the spatio-temporal pattern in cloud-covered MGVI across the whole island on an annual basis, the same method was applied, as described previously, to annual time series of the 10-day image composites, i.e. thirty-six 10-day periods per year, which are shown in Figure 10(a)–(g).

The number of cloud-covered periods per pixel indicates the extent of image data lost due to a cloud and may be useful in selecting areas for vegetation seasonality monitoring which are the most consistently cloud-free. For example; the Northwest, West and Southwest, as well as the Wicklow Mountains tend to be characterised by areas of frequent cloud cover and so would not be suitable for local-scale monitoring. All these areas are at higher altitudes suggesting that orographic cloud (generated by moist air rising and cooling over mountains) is routinely detected in MGVI pixels. The surrounding ocean has relatively fewer cloud periods but this is an artefact of MERIS processing rather than any meteorological occurrence, as cloud pixels over the ocean had already been removed from the scene before the MGVI cloud-detection algorithm was applied across the image (European Space Agency 2006).

The regional variation in cloud cover across the island was further demonstrated by selecting individual pixels well distributed across the country and calculating the percentage of 10-day periods which were cloud-covered per year. These results are summarised in the column chart in Figure 11.

The highest proportions of cloud-covered periods occur in the Wicklow Mountains followed by the Northwest and West, hence where the most image data are lost due to cloud. In contrast, in the South, Southeast and East there is a much smaller cloud-covered proportion of the annual total. These differences in cloud-cover have implications for the selection of local sites for vegetation monitoring using satellite data at a comparable spatial resolution. For example, the composite period could be shortened in areas of less frequent cloud cover, while it may need to be extended for monitoring vegetation in areas of more frequent cloud cover. This is particularly the case in mountainous regions of the western seaboard, along the northern coast and over the Wicklow Mountains.

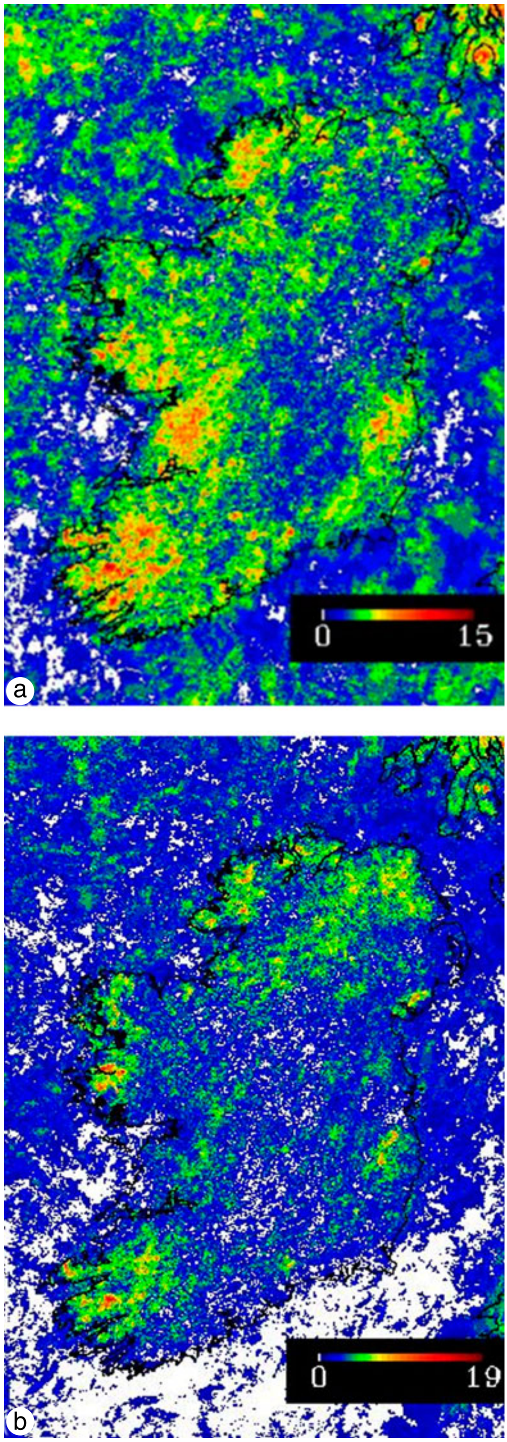

Figure 10. Annual 10-day cloud composites generated over Ireland using both cloud flags derived from the MERIS cloud mask and the MGVI cloud-detection algorithm (a) 2003, (b) 2004, (c) 2005, (d) 2006, (e) 2007, (f) 2008, (g) 2009. The colour bar indicates the number of cloud-covered 10-day periods out of a total of 36. For full colour version of this figure please refer to the online version of the article at http://www.tandfonline.com/rigy

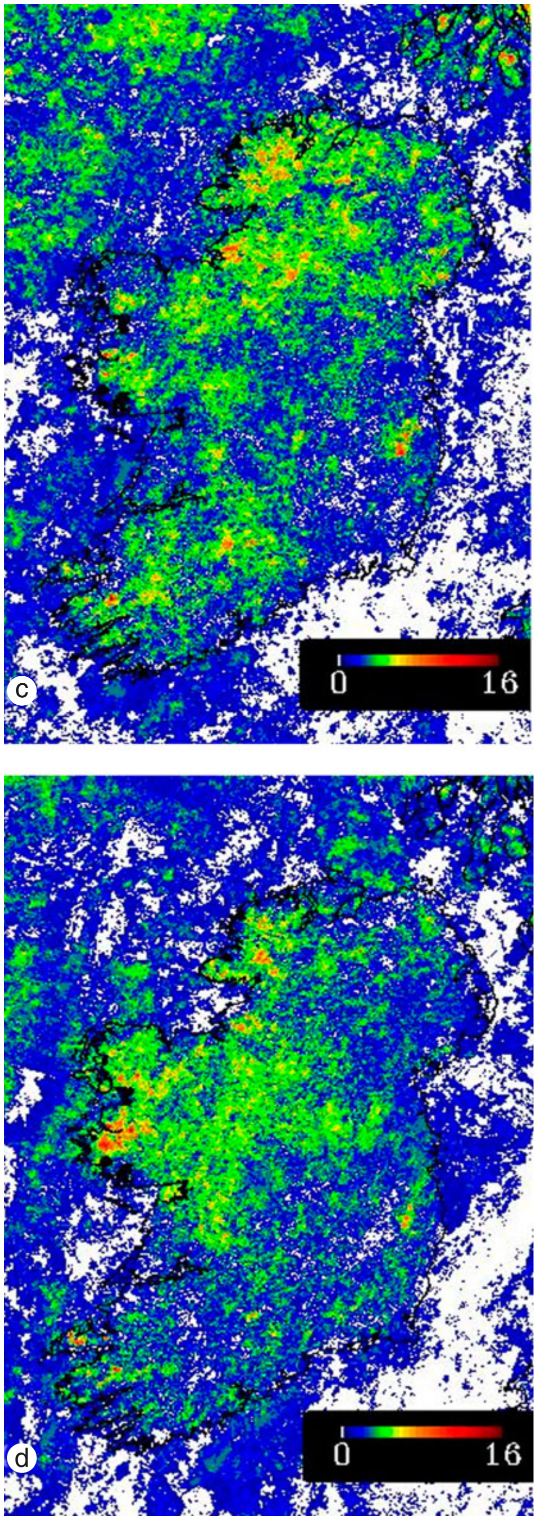

Figure 10. (*Continued*)

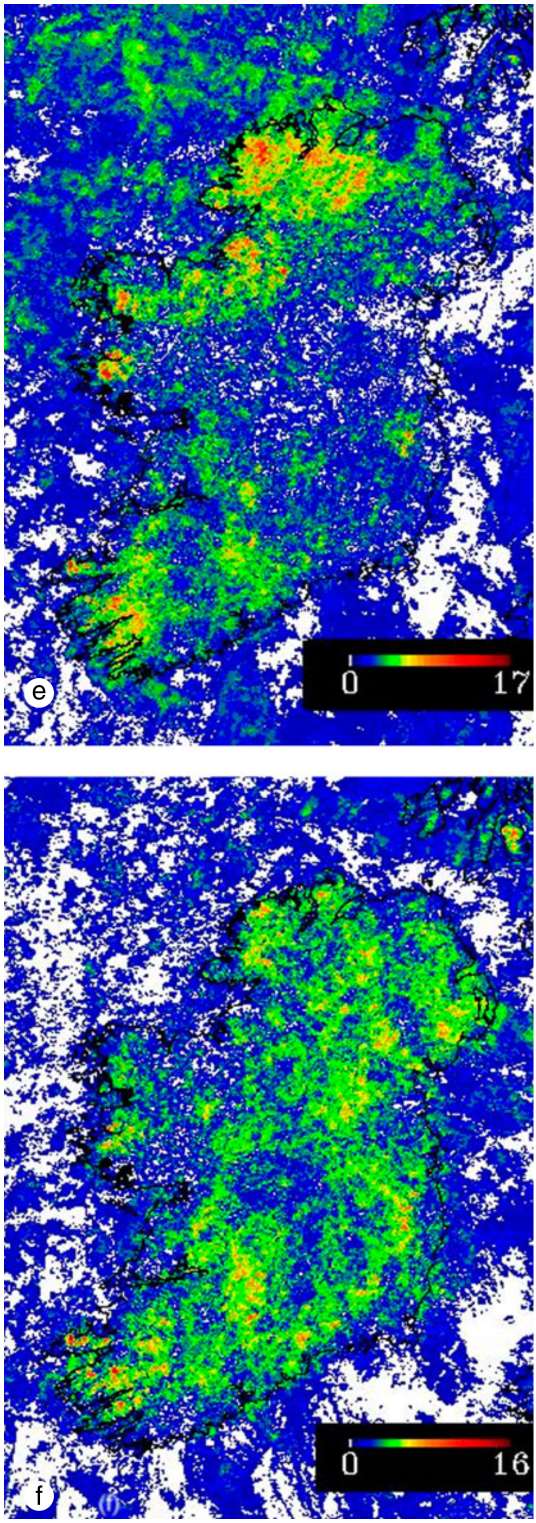

Figure 10. (*Continued*)

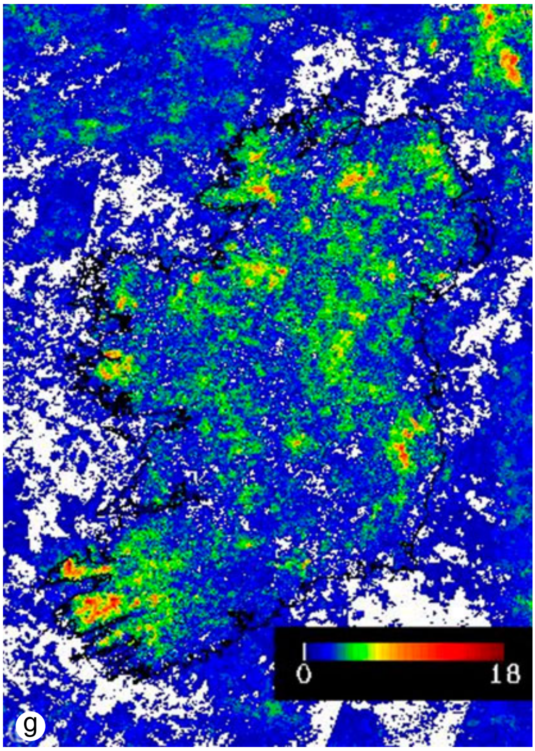

Figure 10. (*Continued*)

Summary of composite period selection

Cloud cover in the daily MGVI grids was the main justification for compositing the daily data, while ensuring the interval was also suited to tracking spatio-temporal variability in vegetation seasonality. However, there were considerable implications of shortening or extending the period. In extending the period, there are spatial gains in terms of coverage of cloud-free imagery. Yet, the period could not be extended indefinitely due to the reduced-sensitivity of the composite value to spatio-temporal variation in vegetation growth. Minimising the time period would have the opposite effect of maximising the sensitivity of the VI measure at the expense of cloud-free imagery. This was evident even in a 10-day period when data gaps still appeared in the MGVI time series. The combined effects of time compositing and temporal smoothing of VI time series inevitably add uncertainty to studies of vegetation seasonality from remotely-sensed data.

Verification of the valid MGVI values using the METEOSAT CLM

For 21 January 2006, almost all (95%) of the valid MGVI pixels were found to be derived under clear sky conditions as defined by METEOSAT, with only 4% of valid MGVI pixels defined as being cloud covered by METEOSAT, mostly under scattered cloud (see Table A1). Of the invalid MGVI pixels, nearly a third of them (31%) occurred in areas identified as cloud-free land by METEOSAT, while a greater number were under cloud (46%). The remaining invalid MGVI values (23%) were in ocean pixels. By comparison on 23 January, 100% of the valid MGVI pixels were determined under potentially cloudy

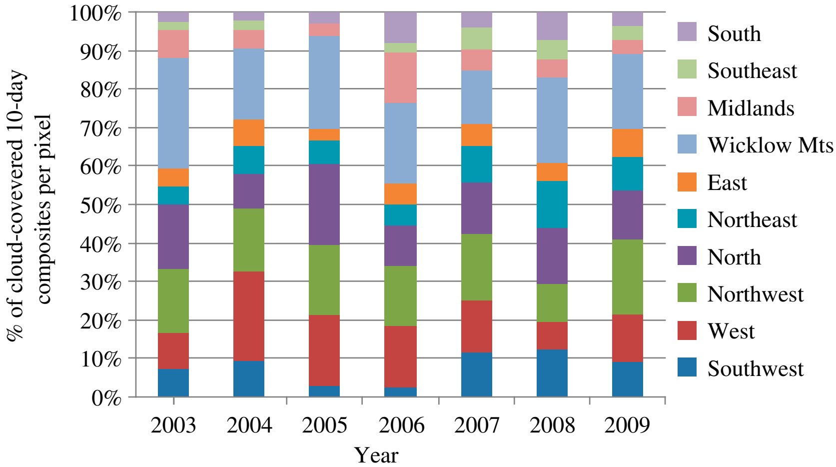

Figure 11. The percentage of cloud-covered 10-day periods per pixel from 2003 to 2009. The pixels were selected from regions of the country to show variability in cloud cover across the island per year.

conditions (however such was the amount of cloud cover on this day that less than 1% of the MGVI grid contained valid MGVI pixels). Such daily analysis is of limited value on days of extensive cloud cover when there were few valid MGVI values, indeed over the 50 days of the five composites 12 days showed that no valid MGVI values were calculated. Therefore, only 38 daily MGVI grids were used in the final verification study.

Overall, it was seen that valid MGVI pixels were frequently cloud-covered, e.g. in 27 of the 38 MGVI grids (71% of the sample), 50% or more of the valid MGVI pixels were deemed to be cloud-covered in the CLM. For easier comparison therefore, the MGVI values from each composite period were amalgamated, as they are for extraction of phenological information. As the focus of attention concerns MGVI pixels which were cloud-free, only those pixels with valid values are considered further. The average percentage of pixels within each composite derived under conditions deemed cloud-free and cloudy by METEOSAT is shown in Table 2, while the daily percentage results are shown in Table A1. The addition of test results does not equal 100% in some cases due to the presence of un-compared MGVI pixels outside the extent of the CLM grid.

There is an increase in the amount of valid MGVI values determined to be cloudy as the year progresses except in the November composite when there are equal numbers of cloudy and cloud-free values. Monthly meteorological reports confirm this as

Table 2. The percentage of valid MGVI pixels derived under conditions defined by METEOSAT as cloudy or cloud-free. The daily values were averaged over the number of days in the composite period.

MGVI composite start date (number of valid days)	METEOSAT cloudy (mean%)	METEOSAT cloud-free (mean%)
20060121 (6)	25	73
20060421 (7)	54	43
20060621 (9)	83	16
20060928 (10)	90	9
20061127 (6)	50	50

progressively cloudier conditions were reported as the year progressed (Met Éireann 2010). This suggests that even though cloud presence over Ireland was confirmed by weather reports and detected by METEOSAT, it was not detected by either MERIS cloud screening or the MGVI cloud-detection algorithm. The number of cloud-covered daily MGVI values was unexpectedly high. However, daily valid MGVI values were also absent even in areas of clear sky. This latter result would suggest that the FAPAR time-composite algorithm is conservative in the selection of a representative daily value. However, there are some external factors which may have influenced this analysis that require discussion.

Spatially, a geostationary orbit ensures that the METEOSAT SEVIRI sensor's optimal spatial resolution of 3 km is maintained at equatorial latitudes while at Irish latitudes, it is approximately 6 km. Therefore, the presence of cloud in 50% or more of the 36 km^2 area covered by the METEOSAT pixel would render that pixel cloud-covered. As a consequence, when remapped to the MGVI grid size, any cloud-free areas will be incorrectly labelled as cloud. The observation angle of the instrument at Irish latitudes is shallower than at equatorial latitudes causing the cloud surface to be observed obliquely rather than vertically as viewed by polar-orbiting MERIS. This would result in all cloud fields observed over Ireland by the SEVIRI sensor to have different characteristics than those observed by MERIS. Temporally, the compared data were not exactly coincident; with six image pairs being seven minutes or more apart, a sufficient time difference for clouds to move.

The cloud-detection strategies of both sensors are dependent on their spectral ranges, which have been designed for very different applications. The MGVI cloud-detection strategy is comparatively weaker owing to its narrower spectral range with a much smaller number of threshold tests conducted. Consequently, the amount and type of cloud detected by both sensors is almost certain to vary.

Despite the differences between the SEVIRI and MERIS sensors, METEOSAT CLM data were useful for MGVI validation. The high temporal frequency of observations and sophisticated methods of cloud retrieval, combined with a freely available data archive make it a very valuable source of cloud information. Currently, no other sensor product is capable of providing such high temporal frequency cloud data over Ireland and ground based observations are of limited spatial extent.

The relationship between cloud-covered values and time series noise

Of the 50 MGVI time series values inspected for the presence of spikes, two were data drop-outs. This meant that the valid sample for the spike analysis contained 48 MGVI values of which 17 were cloud-covered and 31were cloud-free. Of the 17 cloud-contaminated values, 53% of them produced a spiked value while 47% did not. Of the 31 MGVI values that were validated as land by the CLM, 45% of them produced a spike while 55% produced no spike. These results are illustrated in the column chart in Figure 12 and show that anomalous spikes did not consistently occur where the values were cloud-covered, as there were almost equal proportions of spikes present on occasions when the METEOSAT cloud masks indicate clear and cloudy skies.

This would suggest that anomalous variations in the MGVI time series could be potentially caused by cloud cover, but it is not the sole cause of their occurrence. An examination of the literature suggested that factors other than cloud such as poor radiometric correction, uncorrected scattering by aerosols, or increased absorption by water vapour over Ireland resulting from its maritime climate might be likely causes in

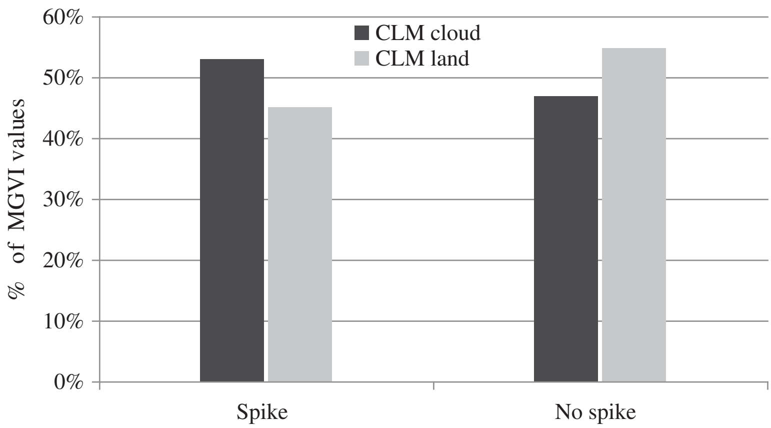

Figure 12. A column chart showing the percentage of cloud-covered and cloud-free MGVI values that produced spikes in the MGVI time series values.

producing anomalous values in the MGVI time series. Others who have used the MGVI data have not reported such widespread noise when working at lower latitudes and in more continental climates (according to N. Gobron, personal communication, 18 June 2010).

MGVI atmospheric corrections

Poor atmospheric corrections are a possible contributory factor to time series instability (Pinty *et al*. 2002). However, a rigorous analysis of the quality of atmospherically-corrected MERIS measurements was not within the scope of this study. The MGVI algorithm uses the 6S atmospheric model to simulate the absorption and scattering processes of the atmosphere and determine their effect on the MERIS reflectance data used in the construction of the index (Gobron *et al*. 1999). The 6S atmospheric model (Vermote *et al*. 1997) limits but does not completely remove the impacts of the absorption and scattering processes of the atmosphere on the MERIS reflectance.

There is no evidence from the MGVI documentation that the 6S atmospheric model is geographically or temporally tuned for variability in the atmosphere on a local scale. This is unlike the dynamic correction for Rayleigh scattering using the MERIS blue band. In fact, only three possible parameter values for optical thickness are tested in the 6S model for MGVI, accounting for only a limited set of atmospheric conditions (see Gobron *et al*. 2004). In contrast, the MERIS sensor would be expected to encounter a wide range of atmospheric conditions on a global scale. Like clouds, water vapour strongly absorbs radiation in the optical domain and is highly variable in vertical profiles of the atmosphere (Schroedter-Homscheidt *et al*. 2008). Over Ireland, water vapour levels tend to vary on a daily and seasonal basis (Rohan 1986). Therefore, errors in the atmospheric correction of MERIS reflectance data would be expected without tuning the parameters of the atmospheric model for such variability.

Conclusions

Although the use of time series satellite-derived vegetation index products for vegetation monitoring has become widespread, the importance of composite period selection is rarely emphasised. In this study, the selection of an appropriate composite period has been guided by ground-based observations of seasonality in surface vegetation and cloud cover. The appropriate compositing interval was defined by an acceptable limit of cloud as completely cloud-free imagery is rarely attainable in cloudy climates like Ireland. However, the compositing interval was also at a sufficient temporal resolution to detect changes in the state of the vegetation surface which could be related to seasonal events on the ground.

The ESA-ESRIN G-POD service was an important tool to determine a composite period appropriate for the application of the dataset, in contrast to many satellite-derived land surface products which are provided at fixed temporal resolutions. The integration of the FAPAR time composite algorithm with the processing capabilities of the G-POD tool allowed long time series of the MGVI product to be generated at a temporal resolution requested by the user.

Generally, a good compositing method ensures that the value selected over the compositing interval is the most representative of the surface state during that interval (Pinty *et al*. 2002). Therefore, composited values should be consistent enough from one period to the next to ensure that robust time-series analysis can be conducted (Huete *et al*. 2002). However, as was found in this study, time series satellite data are not always smooth due to various sensor disturbances. The presence of undetected cloud cover in the daily valid MGVI values selected by the compositing algorithm was suggested by the METEOSAT cloud mask. Although spatial and temporal inconsistencies between the sensor products may have been a factor in the large amounts of cloud-covered MGVI values, MERIS cloud screening may be sub-optimal for cloud cover experienced over Ireland while the MGVI cloud-detection algorithm is possibly too simple a threshold test to identify every cloud type. These assumptions would need to be verified by comparison of the METEOSAT CLM with the pixels identified by MERIS cloud screening and the MGVI cloud-detection algorithm separately. The rigour of both cloud detection strategies could be compared in this way.

Undetected cloud cover was not the sole cause of noise in the MGVI time series, although it was most certainly one of the main causes, as demonstrated by the spike analysis. Further work is required on identifying the cause of spikes in MGVI values validated as cloud-free by the METEOSAT CLM. One possible cause of instability in a cloud-free pixel is the partial atmospheric correction of the MERIS reflectance data used to construct the MGVI. Atmospheric parameters, such as water vapour and aerosols vary in atmospheric composition depending on location and time of year. As the MGVI atmospheric model was designed for a global vegetation index it is possible that there is no temporal or geographical tuning of the model parameters for spatio-temporal variation in atmospheric constituents on a local scale.

While this does not negate the use of MGVI for vegetation monitoring in Ireland, single data imagery should be used with caution as any one value could be anomalous for that particular day. However, a carefully selected time compositing and temporal smoothing method can minimise the influence of anomalous daily variation in time series data. The cause of local-scale variation in daily MGVI value requires further exploration.

Overall, this paper has outlined an approach to composite period selection using *a priori* information on cloud cover and vegetation seasonality. A validation of the quality of daily MGVI values, representative of a time-composite period, has shown that cloud cover can affect many of these valid values, but cannot be conclusively linked to the occurrence of spikes in time series generated from the cloud-covered data. Further work on validating the MGVI cloud-screening steps separately with METEOSAT cloud mask data could be useful to assess their performance. The study recommends further local-scale tuning of atmospheric corrections employed in global-scale land surface products such as the MGVI. This will ensure more consistent time series data appropriate to monitoring the local to regional scale surface changes.

Acknowledgements

The Environmental Protection Agency (EPA), Ireland, has provided funding for this PhD project under the STRIVE initiative, 2007–2010 (Project 2007-S-ET-4). The authors are grateful to the staff at the Institute for Environment and Sustainability, JRC, Ispra, Italy, especially Nadine Gobron for advice given. The G-POD technical team at the European Space Agency's ESRIN facility provided the MGVI data and the Armagh Observatory provided the cloud data.

References

Armagh Observatory, 2011. *Climate research and meteorology* [online]. Armagh. Available from: http://www.arm.ac.uk/ [Accessed 22 February 2011].

Baret, F., *et al*., 2006. *Report on the validation of MERIS TOC_VEG land products*. European Space Agency.

Campbell, J.B., 2002. *Introduction to remote sensing*. London: Taylor & Francis.

Chambers, L.H., *et al*., 2004. Students as ground observers for satellite cloud retrieval validation. *In:* E.M. Prins, *et al*., eds. *13th Conference on Satellite Meteorology and Oceanography*. Norfolk, VA: American Meteorological Society. Paper no. 6.2.

Colditz, R.R., *et al*., 2008. TiSeG: a flexible software tool for time-series generation of MODIS data utilizing the quality assessment science data set. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 3296–3308.

Cornwall, C., Horiuchi, A., and Lehman, C., 2010. *NOAA ESRL solar position calculator* [online]. Available from: http://www.srrb.noaa.gov/highlights/sunrise/azel.html [Accessed 18 March 2010].

Eklundh, L. and Jönsson, P., 2010. *TIMESAT 3.0 Software Manual*. Lund: Department of Earth and Ecosystem Sciences, Lund University; Malmö, Sweden: Sweden and Center for Technology Studies, Malmö University.

EUMETSAT, 2007. *Cloud detection for MSG - algorithm theoretical basis document*. EUM/MET/REP/07/0132 ed.: EUMETSAT, Am Kavalleriesand 31, D-64295 Darmstadt, Germany.

EUMETSAT, 2011. *Earth observation portal* [online]. Available from: https://eoportal.eumetsat.int/userMgmt/gateway.faces?goto=http://eoportal.eumetsat.int:8080/userMgmt/protected/dataCentre.faces [Accessed 25 February 2011].

European Space Agency, 2006. *MERIS product handbook*. European Space Agency.

European Space Agency, 2011. *Earth observation G-POD (Grid Processing on Demand)* [online]. Available from: http://gpod.eo.esa.int/ [Accessed 2 February 2011].

Gobron, N., *et al*., 1999. The MERIS Global Vegetation Index (MGVI): description and preliminary application. *International Journal of Remote Sensing*, 20, 1917–1927.

Gobron, N., *et al*., 2004. *An optimized FAPAR algorithm, theoretical basis document*. 3rd ed. Ispra: Institute for Environment and Sustainability, European Commission Joint Research Centre.

Gomez-Chova, L., *et al*., 2007. Cloud-screening algorithm for ENVISAT/MERIS multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 4105–4118.

Govaerts, Y.M., *et al*., 1999. Designing optimal spectral indices: a feasibility and proof of concept study. *International Journal of Remote Sensing*, 20, 1853–1873.

Goward, S.N., *et al*., 1991. Normalized difference vegetation index measurements from the advanced very high resolution radiometer. *Remote Sensing of Environment*, 35, 257–277.

Hird, J.N. and McDermid, G.J., 2009. Noise reduction of NDVI time series: an empirical comparison of selected techniques. *Remote Sensing of Environment*, 113, 248–258.

Holben, B.N., 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7, 1417–1434.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25 (3), 295–309

Huete, A., *et al*., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83, 195–213.
IPG, 2005. Phenological observation guide of the international phenological gardens. Revised version of the observation guide of 1960 ed.
Julien, Y. and Sobrino, J.A., 2009. Global land surface phenology trends from GIMMS database. *International Journal of Remote Sensing*, 30, 3495–3513.
Justice, C.O., *et al*., 1985. Analysis of the phenology of global vegetation using meteorological satellite data. *International Journal of Remote Sensing*, 6, 1271–1318.
Kubin, E., *et al*., 2007. *Monitoring instructions of the Finnish national phenological network*. Finnish Forest Research Institute (Metla), Muhos Research Unit.
MERIS, 2006. *MERIS products quality status report (MEGS7.4 and IPF 5)*. Paris, France: The MERIS Quality Working Group.
Met Éireann, 2009. *Met Éireann Monthly Weather Bulletins* [online]. Available from: http://www.met.ie/climate/monthly-summary-archive.asp [Accessed 15 July 2010].
Met Éireann, 2010. *Met Éireann Monthly Weather Summary* [online]. Available from: http://www.met.ie/climate/monthly-summary-archive.asp [Accessed 15 July 2010].
Pallé, E. and Butler, C.J., 2001. Sunshine records from Ireland: cloud factors and possible links to solar activity and cosmic rays. *International Journal of Climatology*, 21, 709–729.
Pinty, B., *et al*., 2002. *A time composite algorithm for FAPAR products, theoretical basis document*. EUR Report No. 20150 EN ed.: Institute for Environment and Sustainability.
Pinty, B. and Verstraete, M.M., 1992, GEMI: a non-linear index to monitor global vegetation from satellites. *Journal of Plant Ecology*, 101 (1), 15–20.
Reed, B. and Brown, J.F., 2005. Trend analysis of time-series phenology derived from satellite data. *IEEE Transactions on Geoscience and Remote Sensing*, 166–168.
Reed, B.C., *et al*., 1994. Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5, 703–714.
Rohan, P.K., 1986. *The climate of Ireland*. Dublin: Meteorological Service.
Schroedter-Homscheidt, M., Drews, A., and Heise, S., 2008. Total water vapor column retrieval from MSG-SEVIRI split window measurements exploiting the daily cycle of land surface temperatures. *Remote Sensing of Environment*, 112, 249–258.
Stöckli, R. and Vidale, P.L., 2004. European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset. *International Journal of Remote Sensing*, 25, 3303–3330.
Studer, S., *et al*., 2007. A comparative study of satellite and ground-based phenology. *International Journal of Biometeorology*, 51, 405–414.
Vermote, E.F., *et al*., 1997. Second simulation of the satellite signal in the solar spectrum, 6S: an overview. *IEEE Transactions on Geoscience and Remote Sensing*, 35, 675–686.
Xiao, J., and Moody, A., 2005. Geographical distribution of global greening trends and their climatic correlates: 1982–1998. *International Journal of Remote Sensing*, 26, 2371–2390.
Yang, W., *et al*., 2006. Analysis of leaf area index products from combination of MODIS Terra and Aqua data. *Remote Sensing of Environment*, 104 (3), 297–312.

Appendix Table A1. The percentage of the daily MGVI grids that were validated as cloud-flagged (test 3) and cloud-free (test 1) by the METEOSAT CLM

Date	Total daily valid pix	Cloud-covered valid pix	Land and valid pix	Test 3 (%)	Test 1 (%)
20060121	16384	598	15637	4	95
20060123	1130	1130	0	100	0
20060124	1545	141	1404	9	91
20060126	2110	445	1599	21	76
20060127	1956	106	1798	5	92
20060129	33630	4515	28838	13	86
Comp. Mean				25	73
20060421	5	5	0	100	0
20060423	13048	7892	4937	60	38
20060425	7465	5821	1473	78	20
20060426	2551	2055	491	81	19
20060427	259	38	192	15	74
20060428	35841	1543	34032	4	95
20060429	2522	1035	1410	41	56
Comp. Mean				54	43
20060621	10573	8432	2132	80	20
20060622	6713	5726	923	85	14
20060623	3435	2483	949	72	28
20060624	2080	2018	48	97	2
20060625	21641	14893	6651	69	31
20060626	6723	5438	1230	81	18
20060627	1346	1160	138	86	10
20060628	5960	5911	37	99	1
20060629	1164	939	198	81	17
Comp. Mean				83	16
20060928	6360	6164	195	97	3
20060929	18523	15423	2853	83	15
20060930	45	35	10	78	22
20061001	9934	9311	566	94	6
20061002	10736	9311	978	91	9
20061003	1287	1047	226	81	18
20061004	9932	8971	949	90	10
20061005	222	222	0	100	0
20061006	1711	1507	204	88	12
20061007	1541	1536	5	100	0
Comp. Mean				90	9
20061127	6039	1003	5023	17	83
20061201	4225	2475	1731	59	41
20061202	22064	2353	19518	11	88
20061203	6	5	1	83	17
20061204	183	178	5	97	3
20061205	7193	2345	4785	33	67
Comp. Mean				50	50