Development of forest cover mask to monitor the health condition of forests in Poland using long-term satellite observations

Maciej Bartold

University of Warsaw, Faculty of Geography and Regional Studies, ul. Krakowskie Przedmieście 26/28, 00–927 Warszawa, Poland
Tel. +48 22 5520654, email: maksh1983@gmail.com

Abstract. The work presented here aims at developing cover mask for monitoring forest health in Poland using remote sensing data. The main objective was to assess the impact of using the mask on forest condition monitoring combined with vegetation indices obtained from long-term satellite data. In this study, a new mask developed from the CORINE Land Cover 2012 (CLC2012) database is presented and its one-kilometer pixel size matched to low-resolution data derived from SPOT VEGETATION satellite registrations. For vegetation mapping, only pixels with a cover ≥ 50% of broad-leaved and mixed forests defined by CLC2012 were taken into account. The masked pixels were used to evaluate spatial variability in eight Natural-Forest Regions (NFRs). The largest coverages by masked forests were obtained in Sudetian (65.7%), Carpathian (65.9%) and Baltic (51.3%) regions. For other forest regions the coverage was observed to be around 30–50%.

Time-series of the Normalized Difference Vegetation Index (NDVI) comprising SPOT VEGETATION images from 1998 until 2014 were computed and cross-comparison analyses on ≥ 50% and < 50% forest cover masks brought up frequent differences at a level higher than 0.05 in seven out of eight NFRs. An exception is the Sudetian region, where the data was highly consistent. Furthermore, the Mann-Whitney U non-parametric test revealed statistically significant differences in two regions: Baltic and Masurian-Podlasie NFR. The comparative analysis of NDVI confirmed that there is a need for additional investigation of the quality of newly developed forest mask combined with vegetation and meteorological data.

Keywords: CORINE Land Cover 2012, forest condition, NDVI, satellite images, vegetation mapping

1. Introduction

Information systems on the condition of forests and tree stands are a good and essential source of knowledge for central government administration, local governments, forestry services and society. The systems developed from research projects, such as the Data Bank on Forests and the National Forest Inventory, are a response to the growing demand for forest monitoring in Poland. Information is collected on the basis of continuous or periodic observations and measurements taken at sample sites and permanent observation plots. Annual reports prepared by the Forest Research Institute, commissioned by the Directorate General of the State Forests National Forest Holding, are comprehensive sources of data about the health of the forests.

A wide range of satellite products can support forest monitoring programmes and systems. Satellites placed on circumpolar orbits are able to conduct regular, daily observations of the earth’s surface. The mission of the American NOAA meteorological satellites has continued since the 1970s. The environmental Terra and Aqua satellites, developed under NASA, have been orbiting continuously since 1999 and 2000 respectively. The European program of environmental satellites, such as SPOT VEGETATION, was implemented in 1998 to 2014 and continued by Proba-V since 2013.

Satellites are classified as high, medium and low resolution depending upon the ground resolution of the pixel, i.e. the actual size of the smallest unit of an image. The above-mentioned satellites belong to the category of low-re-
solution satellites. AVHRR and MODIS sensors, mounted on the NOAA meteorological satellites and Terra and Aqua environmental satellites, scan the surface of the Earth with a spatial resolution of 1 km × 1 km and 250 m × 250 m. The newest environmental Proba-V satellite views an area at a minimum spatial resolution of 100 m × 100 m. The sensors of these satellites capture the surface within the visible light range, as well as in the near, mid- and far infrared.

As a result of access to long-term observations from different satellites, many scientific papers have appeared on monitoring the forest cover and the health condition of the forest environment. Analysing the percentage of forests on the surface of each pixel of satellite images taken over a long period of time has enabled the development of a series of forest maps to determine the rate of deforestation in Europe’s forests in 2012, with a resolution of 250 meters, which was developed on the basis of images recorded by MODIS sensors (website: https://lpdaac.usgs.gov/database_discovery/modis/modis_products_table/mod44b). However, the HRL and VCF products have certain temporal and spatial limitations. HRL databases provide information for only one year and do not cover the entire territory of Poland, whereas VCF databases include the last 15 years, but do not have information about the types of forests.

The wide range of low-resolution, long-term data and their cost-free, widespread availability enable complex analyses to be made of the condition of vegetation. Even the limited spatial resolution allows the environment to be monitored at the national as well as the regional level. However, the low spatial resolution databases, most commonly offered at 1 km × 1 km resolution, and the complexity of the geometric correction of satellite images (Bychawski W. 1988) may result in errors in the accurate identification of forested areas, which may be perceived as extensions of abutting farmland or built-up areas. And this could lead to errors in measuring the status and condition of forests.

Studies are conducted on the growth and development of forests based on many years of satellite observations. Vegetation indices, derived from satellite images, allow us to determine when broad-leaved forests start turning green or shed their leaves in winter. The methods based on these indices allow us to evaluate the state of growth and development of forests. Therefore, in order to monitor the health condition of plants using satellite remote sensing, an invaluable role is played by the thematic layers of forest areas, which serve as a mask for narrowing the research only to forested areas. For example, in studies of the temporal and spatial variability of vegetation in Svalbard, test polygons were prepared with precise information about plant species (Karlsen S. et al. 2014). In turn, the mapping parameters that describe the phenology of forests in India used digital resources on forests to distinguish it from other vegetation complexes (Prabakaran C. et al. 2013). The widely available digital database of land cover, the CORINE Land Cover 2006 inventory, was used in testing methods to detect deforestation in Poland between 2000 and 2006 (M. Bartold 2013).

Reports on the state of forests (State of Europe’s Forests 2015) and in publications on changes in the phenology of plants (Garamvolgyi A., 2013) emphasise the particular role of monitoring the growth and development of forests in the context of climate change. Therefore, in this article, we propose the use of masks of forest areas to monitor their health condition. The aim of this study is to develop a layer of representative pixels of forest areas obtained from the CORINE Land Cover 2012 inventory. The method of selecting pixels is based on an analysis of the percentage of forest area in a given pixel. The thematic layer was then validated by analyzing vegetation index time-series data over forest areas derived from low-resolution satellite data. Temperate zone broad-leaved forests are characterised by a distinct seasonality, and thus are a good test area to analyze the variation of vegetation indices for monitoring the condition of tree stands from satellites.

2. Methodology

The study was conducted in three stages. Figure 1 presents the schematic development and testing of broad-leaved forest masks. In the first stage of the work, the borders of forested areas occurring in Poland were determined using the CORINE Land Cover (CLC) database. As part of the CORINE program, areas with a minimum of 25 ha and a width of at least 100 m were mapped on the basis of high-resolution satellite data (Ciolkosz, Bielecka 2005). The current, fourth edition of the CLC2012 was established within the framework of the European Copernicus GIO Land Monitoring program (Hościlo, Tomaszewska 2015). Because the purpose of the mask is to monitor the health condition of forests from the CLC2012 semi-natural and forest ecosystem clas-
ses, only those areas covered by broad-leaved forests were selected. This condition is fulfilled by classes 311 and 313, defined by the CORINE nomenclature as broad-leaved and mixed forests, with no breakdown by age and species composition of trees. Classes 311 and 313 include areas of forest with over 30% tree cover or a density of 500 trees per hectare. A detailed definition of class 311 includes forest areas in which a minimum of 75% is dominated by broad-leaved forests. Class 313 describes areas of mixed forests, in which the percentage of broad-leaved or coniferous forests does not exceed 25% (CORINE Land Cover nomenclature illustrated guide). As a result of the first phase of the work, we obtained a vectored primary thematic layer of forest areas as of 2012.

Next, an analysis of the percentage of forest area on the surface of a given pixel was conducted (Fig. 2). For this purpose, a grid of 1 km squares was made, corresponding to the pixel grid of SPOT VEGETATION satellite images. We adopted a criterion threshold of 50% of the area occupied by forest within a pixel. This stage of the work was done to minimise errors in measuring the condition of the forest that may arise as a result of the presence of several land cover classes within a single pixel.

This methodology has already been verified during the development of the thematic layer of agricultural land used in forecasting yields based on NOAA AVHRR images (Turlej et al. 2013). The authors of this work compromised between obtaining an optimal number of pixels and preserving their agricultural representativeness. They analysed the percentage of CORINE Land Cover 2006 agricultural classes in the pixel areas of NOAA-AVHRR as well as in the pixel areas of 600 m wide buffer zones. Studies on an advanced agricultural mask showed a slight improvement in the growth index readings for monitoring crops in relation to the index readings using the standard mask threshold of 50%. The 50% criterion is also used in the environmental CORINE programs and GLOBCOVER 2009 products to distinguish dominant forms of land cover in the area of a

![Flowchart of steps to create and test broad-leaved and mixed forests mask](image-url)
given pixel (CORINE Land Cover nomenclature installation guide, GLOBCOVER 2009 Product Description Manual). Therefore, this study also retained a 50% threshold criterion in developing masks of forest areas.

The third stage of the work involved verifying the impact of applying the new mask of forest areas on the measurement quality of the normalised difference vegetation index (NDVI) calculated by using many years of satellite observations. This index, introduced in the 1970s, is widely recognised as one of the most universal measures of the health and resilience of vegetation (CJ Tucker 1979). It is calculated by using the spectral characteristics of chlorophyll in plants contained in the red range of the electromagnetic spectrum and cellular structures in the near infrared. The range of NDVI values is from –1.0 to 1.0, wherein the closer values to 1 indicate a higher intensity of the growth phase and better condition of the plants, and thus a higher amount of biomass and greater percentage of broad-leaved species in the forests. The measurements of the vegetation index for forests in the period 1998–2014 was based on SPOT VEGETATION satellite data. A total of 581 NDVI 10-day temporal composites were used for the measurements, which consisted of daily satellite observations after the maximum value of the index observed in a given 10-day period. This method of combining the daily images in a composite solves the problem of the frequent cloud cover over Poland. 10-day composites, designated as a s10 product, are produced and made available free of charge by the Flemish Institute for Technological Research (VITO) in Belgium.

The NDVI values of broad-leaved and mixed forests were averaged within eight natural-forest regions, i.e. the Baltic region (I), Masurian-Podlasc region (II), Greater Poland-Pomerania region (III), Mazovia-Podlasc region (IV), Silesian region (V), Lesser Poland region (VI), Sudeten region (VII) and the Carpathian region (VIII). Because the forest masks are to be used to monitor the health condition of forests in Poland, we divided the regions in this study to take into account the climatic factors affecting forest growth. The spatial division of the regions was developed by the Faculty of Forestry at the Warsaw University of Life Sciences – SGGW as part of the work on regionalising natural-forest areas in 2010 (R. Green, Kliczkowska A. 2012). Natural-forest regions are hierarchically the highest units of regionalisation. The boundaries were determined on the basis of differences in geology, geomorphology, climate and types of natural landscapes. Such a division also takes into account the variation in the percentage of forest species and their suitability for forest management. For the purposes of this study, the regional boundaries were generalised to a scale of 1:20,000,000.

The assessment of the usefulness of the new mask was based on NDVI measurements performed for both the areas that met the 50% criterion of forest within the 1 km pixel (≥ 50% mask) and for the remaining forest areas, which did not meet this criterion (< 50 % mask). Figure 3 shows the spatial distribution of the percentage of forest area using the example of the Carpathian region. To determine the statistical significance of differences in the 1998–2014 NDVI measurements for the two separate masks of ≥ 50% and < 50% forest area, we used the non-parametric Mann–Whitney U test. Selecting the non-parametric test is justified by the lack of a normal distribution of the NDVI index for forests during the growing season and the juxtaposition of pairs of observed index measurements from two independent populations of forest areas masked with the threshold of ≥ 50% and < 50%. The null hypothesis assumes no difference in the NDVI measurements and thus no impact of the developed forest mask on their quality. The alternative hypothesis assumes that differences in the vegetation index are statistically significant and the use of a new

![Figure 2. Methodology of producing forests mask: a) broad-leaved and mixed forests in CORINE Land Cover 2012 database, b) SPOT-VEGETATION 1 kilometer grid c) pixels with 50% threshold of occurred broad-leaved and mixed forests](image-url)
forest mask influences the results obtained when monitoring the health condition of forests. The level of significance was established at 0.05. If a test result of $p < 0.05$ was obtained, it was assumed that the distributions vary significantly and the alternative hypothesis was adopted. However, if the result was $p > 0.05$, there was no basis to reject the null hypothesis.

**Figure 3.** Mask of one kilometer pixels with ≥50 and <50% broad-leaved and mixed forests in Carpathian Natural-Forest Region (CNFR)

**Figure 4.** Broad-leaved and mixed forests mask at ≥50% overlaid natural-forest regions
3. Results and discussion

Figure 4 shows a map of the spatial arrangement of broad-leaved and mixed forest areas, whose percentage in 1 km² is equal to or greater than 50%. The map also shows the boundaries of the eight natural-forest regions. The criterion threshold of 50% allowed us to maintain a certain representativeness of the forest pixel for studies on monitoring forest growth and development. The percentages of the masked broad-leaved and mixed forest areas in the regions are summarised in Table 1. The largest area of forests in the study was recorded in the Carpathian region, accounting for 27.9% of the region. Next on the list were the forests of the Sudeten and the Baltic regions, 16.7% and 13.5% respectively. Masurian-Podlasie also had more than 10% of its area covered by forests. The least amount of masked forests was detected in the Greater Poland-Pomerania region, where they accounted for 5.2% of the area.

Table 1 also shows the percentages of the area of broad-leaved and mixed forests classified in the CORINE Land Cover 2012 database. The percentage of identified broad-leaved and mixed forests was in the Sudeten and Carpathian region, representing, respectively, 65.7% and 65.9% of these areas. The forests of the Baltic region also had a high result, over 50% of the area. The percentage of forests in the remaining regions ranged from 30–50% of the areas.

Juxtaposing the percentage of broad-leaved and mixed forests classified by CLC2012 and the ≥50% mask is aimed at demonstrating the amount of information about forests that is lost when producing a new mask to monitor the health condition of forests. The largest loss of information on forest areas, 49% (from 65.7% to 16.7%), was recorded for the Sudeten region. This result indicates that virtually every other analysed 1-km pixel did not exceed the required criterion threshold of 50% of broad-leaved and mixed forests. This may be due to the presence of numerous small areas of forest with scattered tree stands that survived intensive industrial activities and the gradation of insect pests (Bocheneck Z. et al. 1997). Large losses of information about forest areas were also observed in the Baltic region – 37.8% (from 51.3% to 13.5%) and the Carpathian Region – 38.0% (from 65.9% to 27.9%). In other regions, the size of the forest area decreased to 25–35%.

To summarise, an average of about one-third of the areas of broad-leaved and mixed forests identified in the CLC2012 database were rejected due to not meeting the criterion threshold of 50% for establishing a new forest mask. These 1-km areas, where forests are not the dominant form of land cover, were rejected due to: (1) the need to maintain representative forest pixels and (2) the need to obtain the most reliable index measurements and to minimise interference from other land cover types (e.g. agricultural areas, areas being built up in the vicinity of forests). To answer the question of how the rejected forest areas could affect the quality of the NDVI measurements, a comparative analysis was made of the index obtained using the time series for 1998–2014. Figures 5 and 6 show the differences in the NDVI measurements, obtained on the basis of many years of SPOT VEGETATION satellite observations for the ≥ 50% mask of broad-leaved and mixed forests and the <50% mask of these forests. The results in the context of monitoring vegetation were sorted by the decade of the year.

To evaluate the preliminary results of the NDVI comparison, a confidence level of 95% with a rejection threshold for measurement error of 0.05 was applied. As the scale of NDVI values is from −1.0 to 1.0, as is the scale of differences for this index, the thresholds of measurement errors of the differences in NDVI are also at −0.05 and 0.05. The

<table>
<thead>
<tr>
<th>Natural-Forest Region (NFR)</th>
<th>NFR Area [km²]</th>
<th>CLC2012 broad-leaved and mixed forests share [%]</th>
<th>Masked broad-leaved and mixed forests share [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baltic</td>
<td>40193</td>
<td>51.3</td>
<td>13.5</td>
</tr>
<tr>
<td>Masurian-Podlasie</td>
<td>27660</td>
<td>41.6</td>
<td>10.7</td>
</tr>
<tr>
<td>Greater Poland-Pomerania</td>
<td>71439</td>
<td>30.8</td>
<td>5.2</td>
</tr>
<tr>
<td>Mazovia-Podlasie</td>
<td>53162</td>
<td>37.4</td>
<td>6.0</td>
</tr>
<tr>
<td>Silesian</td>
<td>27125</td>
<td>43.7</td>
<td>8.9</td>
</tr>
<tr>
<td>Lesser Poland</td>
<td>68520</td>
<td>39.0</td>
<td>8.2</td>
</tr>
<tr>
<td>Sudetian</td>
<td>5070</td>
<td>65.7</td>
<td>16.7</td>
</tr>
<tr>
<td>Carpathian</td>
<td>19342</td>
<td>65.9</td>
<td>27.9</td>
</tr>
</tbody>
</table>

Table 1. CLC2012 and ≥50% masked broad-leaved and mixed forests share in Natural-Forest Regions areas
Figure 5. NDVI differences calculated from satellite SPOT-VEGETATION 1998–2014 data for ≥ 50 and < 50% forest mask in Natural-Forest Region: a) Baltic, b) Masurian-Podlasie, c) Greater Poland-Pomerania, d) Mazovia-Podlasie
Figure 6. NDVI differences calculated from satellite SPOT-VEGETATION 1998-2014 data for ≥ 50 and < 50% forest mask in Natural-Forest Region: a) Silesian, b) Lesser Poland, c) Sudetian, d) Carpathian
rejection thresholds are depicted in Figs. 5 and 6, where the differences in NDVI measurements of greater than 0.05 and less than 0.05 indicate a significant change in the vegetation index readings. NDVI differences greater than 0.05 indicate higher values of the vegetation index based on the ≥ 50% mask. Measurements recording a NDVI of less than 0.05 indicate higher values of the vegetation index of the < 50% mask.

NDVI measurements were made for 581 ten-day composites (observations) from SPOT VEGETATION, starting from the first decade of April 1998 to the second decade of May 2014, both months inclusive. The most frequent differences in the vegetation index were recorded in the Baltic region, where out of 581 observations, as many as 76 periods were found to have exceeded the NDVI measurement error of 0.05 (Fig. 5a). This represents 13% of all measurements of the vegetation index made in 1998–2014. Masurian-Podlasie and Greater Poland-Pomerania followed, registering 51 and 48 cases of significant differences of NDVI (Figs. 5b and 5c). On the other hand, the results of the measurements for the Sudeten region were at the opposite extreme. There was only one case of the occurrence of a difference exceeding 0.05 of the NDVI in broad-leaved and mixed forests with masks of ≥ 50% and < 50%.

We then examined how often the differences in NDVI measurements occurred during periods of tree stand growth and development, and the implications this may have for monitoring the vegetation of forests. For this purpose, the NDVI measurements in Figs. 5 and 6 were sorted by the decades. In the case of the Baltic region, having the most frequently recorded measurement differences, up to 99% of them were observed in decades 18–30, which corresponds to the period of June–October (Fig. 5a). This is a very important piece of information, highlighting the importance of carrying out vegetation measurements of forests. Figures 5c for the Greater Poland-Pomerania region, 6a for the Silesian region and 6b for the Lesser Poland region also confirm the existence of significant differences in NDVI measurements only in the period of plant phenology, from the 18th to the 30th decade. On the other hand, Figs. 5b and 5d show cases of differences in the index values in winter and summer-autumn. Variations in the measurements greater than 0.05 observed during the winter months may indicate imperfections related to masking snow cover. These differences have no significant impact on studying forest vegetation during spring and autumn.

At the end of the analysis, we performed nonparametric Mann–Whitney U tests. The measurement statistics of the regions are as follows: (a) Baltic $p = 0.021$, (b) Masurian-Podlasie $p = 0.038$, (c) Greater Poland-Pomerania $p = 0.095$, (d) Mazovian-Podlasie $p = 0.221$, (e) Silesian $p = 0.072$, (f) Lesser Poland $p = 0.105$, (g) Sudeten $p = 0.389$ and (h) Carpathian $p = 0.083$. The test results confirm that the measurements of forests with a mask of ≥ 50% and < 50% in the regions of northern Poland, i.e. the Baltic and Masurian-Podlasie regions, had statistically significant differences in NDVI values. The greatest risk of rejection of the null hypothesis, which assumes no difference, was found for the Sudeten (38.87%), Mazovian-Podlasie (22.12%) and Lesser Poland (10.45%) regions.

In summary, the new forest mask of ≥ 50% had an influence on obtaining higher NDVI measurements by more than 0.05 in the growth periods of forests in seven natural-forest regions: Baltic, Masurian-Podlasie, Greater Poland-Pomerania, Mazovian-Podlasie, Silesian, Lesser Poland and Carpathian. Because we obtained relatively reliable NDVI values, the results of these measurements could imply the validity of using masks. The developed forest mask of ≥ 50% partially eliminates the problem of measurement errors of the indicator for “pure forest” areas and areas with no other form of land cover. However, it is advisable to check the classes of land cover other than forests that occur in the pixels and to check the surrounding areas of each pixel, thereby creating buffer zones of a width of over ½ of the 1 km pixel, i.e. 600 meters. The experiences of working on an agricultural mask to forecast yields (Turlej K. et al. 2013) showed no unequivocal, clear and universal criteria for mask and a buffer zone thresholds. On the other hand, the assumption adopted in this paper of an authoritative threshold of 50% is used in environmental programs and products to classify land cover, such as CORINE Land Cover or GLOBCOVER 2009.

The results of the Mann–Whitney U test showed that among all of the eight regions surveyed, only two, Baltic and Masurian-Podlasie, obtained statistically significant differences. These are regions where the most numerous differences were observed of NDVI measurements for 1998–2014 – 76 and 51 respectively out of 581 observations. We found no justification to reject the null hypothesis for the remaining six regions, and the risk of rejecting the true hypothesis was at a level of 7.19% for the Silesian to 38.87% for the Sudeten region.

The forests of the Sudeten region are a special case, where we noted absolutely no impact of the new mask on the quality of the readings (one erroneous observation for all 581 possible observations). It is worth noting that in terms of area, this is the smallest natural-forest region of all eight studied, having a mountain landscape dominated by coniferous forests in the upper layer and mixed forests in the lower layer. Coniferous forests, which are very close or identical to the readings of broad-leaved and mixed forests in the NDVI measurements, dominate in the 1-km pixel grid. Therefore, the mask threshold of 50% may not
have had any effect on the quality of the vegetation index measurements.

To conclude the discussion, we recommend that an additional assessment be made of the NDVI measurements by checking the correlation with meteorological data or satellite data from a different source. A supplemental verification of research results could also be the juxtaposition of index measurements based on low- and high-resolution images, for example, SPOT VEGETATION and Landsat. In this case, due to the nature of registering high-resolution images, the research can be performed only for selected test polygons within a specified time interval.

**Conclusions**

1. The developed forest mask of ≥ 50% influenced the results of vegetation index measurements for forests in seven natural-forest regions. Use of the new forest masks had no effect on the measurement results of the profiles of the vegetation index for the Sudeten region’s forests.

2. In total, 99% of the differences in NDVI larger than the 0.05 were recorded during the period of forest growth and development, from the 18th to the 30th decade. This points to the need to take into account the developed mask in research on monitoring the health condition of the forest environment.

3. The results of NDVI measurements using a new forest mask of ≥ 50% and a mask of < 50% showed statistically significant differences for the Baltic and Masurian-Podlasie regions, accounting for 13.1% and 8.8%, respectively, of all observations in 1998–2014. The risk of the rejection of the null hypothesis, which assumes that the measurements agree, is less than 2.14% and 3.80%, respectively.

**Conflict of interest**

The author declares no potential conflicts of interest.

**Acknowledgements and source of funding**

This study was conducted within the framework of Doctoral Studies at the Faculty of Geography and Regional Studies at University of Warsaw. The author would like to thank the staff of the Remote Sensing Centre at the Institute of Geodesy and Cartography in Warsaw for their help and valuable advice.

**References**


Translated by: Barbara Przybylska