Remote sensing-based forest health monitoring systems – case studies from Czechia and Slovakia

Ivan Barka1*, Petr Lukeš2,3, Tomáš Bucha1, Tomáš Hlášny1,4, Radim Strejček2, Marek Mlčoušek2, Štěpán Křístek2

1National Forest Centre - Forest Research Institute Zvolen, T. G. Masaryka 2175/22, SK – 960 92 Zvolen, Slovak Republic
2Forest Management Institute Brandýs nad Labem, branch Frýdek-Místek, Nádražní 2811, CZ – 738 01 Frýdek-Místek, Czech Republic
3Global Change Research Institute, Academy of Sciences of the Czech Republic, Bělidla 4a, CZ – 60 300 Czech Republic
4Czech University of Life Sciences, Faculty of Forestry and Wood Sciences, Kamýcká 129, CZ – 165 01 Praha 6 – Suchdol, Czech Republic

Abstract

Aim of this paper is to present the remote sensing-based systems of forest health assessment in the Czech Republic and Slovakia, and to analyse both their strengths and weaknesses. Nationwide assessment of forest health in the Czech Republic is based on the interpretation of Sentinel–2 satellite data using novel approaches for cloud-free image synthesis based on all available satellite observations. A predictive statistical model to yield time series of leaf area index (LAI) from satellite observations is developed above extensive in-situ data, including LAI and forest defoliation assessment. Forest health is evaluated for each pixel from yearly changes of forest LAI, while the country-wise assessment of the health status is performed at the cadastral level. Methodology developed for Slovakia is based on a two-phase regression sampling. The first phase of the procedure provides an initial fast estimate of forest damage using only satellite observations (visible and infrared channels from Landsat or Sentinel–2 systems). The second phase refines the result of the first phase using data from a ground damage assessment (site-level defoliation from ICP Forests database). Resulting forest health assessment over the whole forest area is presented in 10 defoliation classes. The Czech Republic shows 1.6% of heavily damaged forests, 12.5% of damaged forests, 79.2% of forests with stable conditions, 6.3% of regenerated forests and 0.4% of strongly regenerated forests. In Slovakia, the total share of damaged stands (i.e. with defoliation higher than 40%) increased from 6 – 8% in 2003 – 2011 to 13 – 15% in 2012 – 2017. Both methodologies conduct nationwide assessment of forest health status in a fast and automatized way with high accuracy and minimal costs. The weaknesses are, for example, a high computational demands for production cloud free mosaics, inability to identify initial phases of forest health decline, exclusion of stands older than 80 years (in the Czech Republic) and inability to differentiate between harvested and severely damaged stands (in Slovakia). Finally, the paper outlines future development of both methodologies.

Keywords: forest health status; defoliation; satellite scenes; Landsat; Sentinel–2

1. Introduction

Monitoring of forest health status has gained increasing importance, especially in recent years when influence of climatic and other stressors adversely affected forest conditions in many regions (Lindner 2014; Millar & Stephenson 2015; Hlášny et al. 2012). While terrestrial evaluation of forest health is time consuming and is usually conducted on monitoring plots, remote sensing (particularly satellite based) allows for cost- and time-efficient, continuous evaluation of forest conditions over large areas (Franklin et al. 2002; Frolking et al. 2009). Optical remote sensing of forests is based on the spectral response of vegetation. Vegetation spectral characteristics result from interaction of solar radiation with cell structure, chlorophyll and other pigments. The amount of pigments relates to damage level and with increasing damage level the amount of chlorophyll decreases (Szekielda 1988). Extensive damage causes deterioration of chloroplasts, recognised as yellowing of leaves and needles, while maximal reflectance is shifted from green to red spectral band. This principle led to development of
several so-called vegetation indices used in forest disturbance monitoring, e.g. Red-Green Index (RGI, Wulder et al. 2008). Rather than calculation of single vegetation indices (band ratios) and their classifications, current trend in this field lies in applications of time series and identification of trajectories or functions in datasets. A comprehensive utilisation of eight recently used forest change detection algorithms was published by Healey et al. (2018).

Landsat satellite scenes are used in forestry for decades (e.g. Goldberg et al. 1985; Coleman et al. 1990), however, their usage increased greatly after deployment of free data policy in 2008 (Wulder et al. 2012). This policy and long-term mission since 1972 enabled cheap development of long time-series of satellite products, suitable also for change detection based on trend evaluation (Kennedy et al. 2010; Meigs et al. 2011; Zhu et al. 2012; Banskota et al. 2014). The same data policy was adopted by ESA for Sentinel data within Copernicus programme. Sentinel optical and radar data acquired by satellites operating in tandem, and with even higher spatial resolution, represent unique data source suitable for forest change detection.

All Landsat satellites scan systematically the Earth’s surface roughly once every 16 days in a spatial resolution of 30 – 60 m per pixel in several spectral channels in visible, near infrared and medium infrared wavelengths (Roy et al. 2014). Thanks to the relatively wide swath (185 km), the data are suitable for large-scale systematic mapping of the entire continents. The availability of multispectral “Landsat based” satellite sensors have recently significantly expanded with the launch of the satellite system of the European Space Agency (ESA) called Sentinel–2 (Drusch et al. 2012). It shares similar features as Landsat (10 and 20 m pixel spatial resolution), but its distinctive advantage is the remarkable revisit time of 5 days and the presence of several bands in the red-edge area for the retrieval of vegetation quantitative parameters associated with leaf pigments and biomass. Besides Landsat and Sentinel–2 satellites, there are numerous commercial satellites to capture very high spatial resolution data, or scientific satellites to monitor global trends in vegetation, but these are beyond the scope of this paper.

The application of modern methods of remote sensing for the study of the health state of forests in the Czech Republic was initiated by Rock et al. (1986, 1988). The authors used for the first time the moisture stress index (MSI) and the blue shift phenomena. These principles were further elaborated by Lambert et al. (1995). Using logistic regression, field survey and TM1, TM3, TM4 and TM7 spectral bands they classified forest health status in 3 classes with 71 – 75% accuracy. Campbell et al. (2004) studied the potential of hyperspectral airborne data and were able to distinguish four classes of forest damage by means of indexes in the range of 673 – 724 nm. Hais et al. (2009) focused on the detection of forest disturbance. By studying Landsat’s time series, the authors demonstrated the potential of the Tasseled cap analysis and Disturbance Index. Mišurec et al. (2012) observed a close relationship between several hyperspectral vegetation indices and the amount of chlorophyll in leaves. A similar approach was used by Malenovský et al. (2013). With the use of hyperspectral airborne data, the accuracy of the chlorophyll a + b retrieval over 90% was achieved without the use of field data for model calibration. Lhotáková et al. (2013) tested the use of laboratory methods to measure the reflectance of the pine needles by varying the level of stress. To distinguish the stress level, they recommended studying the water content of the tissues, the amount of leaf pigments and phenols as stress indicators. Kopačková et al. (2015) demonstrated the relationship between the two toxic elements – aluminium (Al) and arsenic (As), which were redistributed from soil to needles and affected the amount of soluble phenols and the ratio of chlorophylls to carotenoids. Ability of Tasseled transformation to predict the bark beetle attack was demonstrated by Hais et al. (2016) using retrospective analysis of Landsat data. The wetness components and the slope of its time curve had the greatest prediction potential. The combination of the Landsat time series and hyperspectral airborne data (ASAS and APEX) was studied by Mišurec et al. (2016) in order to monitor changes in the physiological state of spruce stands. An analysis of the time series of the Disturbance index (similar to Hais et al., 2009) demonstrated its potential to detect previous crop disturbances. From the hyperspectral indices, the highest predictive potential for damage identification had the Vogelmann red-edge index, which is sensitive to changes in photosynthetic leaf pigments. The advantage Sentinel–2 satellite data was recently demonstrated by Homolová et al. (2016). Inversion of Discrete Anisotropic Radiative Transfer Model (DART) was applied to retrieve three products related to forest health – chlorophylls a + b, leaf water and leaf index for two types of ecosystem – beech and spruce forests.

Satellite remote sensing-based assessment of forest health in Slovakia was developed in relation with aerial forest monitoring (Račko 1996). Scheer (1997) presented advantages of two-phase regression sampling in classification of forest health from satellite scenes in a case study from northern part of Slovakia. The same methodology was later used and developed in studies Barka & Bucha (2010), Bucha & Barka (2014) and Barka & Bucha (2017). Since 2009, classifications of forest health status (including historical assessments from 1990 onward) based on Landsat and Sentinel–2 data are published via web map applications (http://www.nlcsk.sk/stales) on a yearly base.

In both Slovakia and the Czech Republic, the main motivation to conduct and improve satellite-based remote sensing of forests is the need to monitor forest health trends, particularly in Norway spruce (Picea abies) forests suffering from a compound of stressors, including windstorms and biotic agents. Large area of managed forests with relatively high share of spruce stands in both
countries (2.02 mill. ha with 23% of spruce in Slovakia and 2.67 mill. ha with 44% of spruce in the Czech Republic) requires efficient monitoring of forest health status capable to evaluate the whole territory at least once a year. Secondary Norway spruce stands started to show worsening of health status in late 1990s. In Slovakia, the process accelerated rapidly after a windstorm in the High Tatra Mts. in November 2004 as well as numerous smaller events, leading to subsequent bark beetle outbreaks and an increase of sanitary fellings up to 65% of the total annual harvest. In the Czech Republic, the most severe large-scale dieback of spruce stands occurred between 1992 and 2010 in Šumava National park and from 2012 up to now in two regions – North Moravia and Olomouc. In 2016, almost 2 mil. m$^3$ of spruce wood infested by bark-beetle was harvested in this two regions from the total of 3 mil. m$^3$ of total volume of spruce wood. This trend further accelerated in 2017 and 2018 with expected sanitary loggings reaching unprecedented volume of 20 mil. m$^3$. Similar remote sensing based monitoring systems are developed in the neighbouring countries (Bartold 2016; Somogyi et al. 2018).

Objective of this study is to provide systematic description of remote sensing-based forest health monitoring systems in both countries, which is the information that has been so far available in numerous fragmented sources and often in national languages only. Assessment of strengths and weaknesses of the two systems might support their further development as well as their broader use in forest management and research.

2. Material and methods

2.1. Czechia

2.1.1 In-situ data

The assessment of forest health status in Czechia relies primarily on two in-situ datasets: 1) leaf area index (LAI) assessment on the plots of National Forest Inventory, 2) plot-level defoliation assessment as performed within ICP Forests initiative. Sampling scheme in both datasets is well suited for satellite data interpretation, having plot size of 30 × 30 m for LAI plots and 50 × 50 m (2 500 m$^2$) respectively for ICP Forests plots.

In total 189 plots were sampled for LAI using method of digital hemispherical photography (DHP), from which 79 stands were dominated by coniferous trees (42% of plots) and 110 by deciduous trees (58%). Details on DHP data analysis can be found in Leblanc et al. (2005). Sampling of DHP was designed to meet the Sentinel–2 spatial resolution of 20 m per pixel as square with side of 20 m, while the number of photos and their position was selected according to Majasalmi et al. (2012) as cross-shape with 13 sampling locations in 3 m distance.

2.1.2 Satellite data used for interpretation

The key to assessing the health status of forests from remote sensing data is the availability of high quality (i.e. cloud-free) mosaic generated from all-available Sentinel–2 data. This is a basic pre-requisite for any remote sensing data interpretation. Czech methodology for forest health assessment proposed a novel processing chain for automated cloud-free image synthesis based on the analysis of all available Sentinel–2 satellite data for selected sensing period (e.g. the vegetation season from June to August). The processing chain is implemented in three follow-up processes – 1) batch downloading, 2) atmospheric corrections of raw images (so-called L2 process) and 3) automated synthetic mosaic generation (so-called L3 process, or spatio-temporal image synthesis...
Fig. 2. Pre-processing workflow of Sentinel-2 satellite data from L1C to L3 product.
Due to its high computational and data storage requirements, the processing chain is implemented on IT4Innovations supercomputer facility (© 2018 VŠB-TU Ostrava), which enables for distributed computing on many computational nodes.

In the first step, Sentinel–2 scenes are automatically downloaded from Copernicus Open Access Hub (global Copernicus data access point) and CESNET (Collaborative ground segment of Copernicus implemented in Czech Republic). Next, the atmospheric and topographic corrections are performed for each Sentinel–2 image using Sen2Cor tool by ESA. Then, each pixel in image mosaic is evaluated independently in the time series of images. Selection of the highest quality pixel, having lowest cloud cover and being in vegetation growing season, is based on decision tree using the values of vegetation index sensitive to biomass (e.g. the Normalized Difference Vegetation Index, or NDVI). In addition to highest NDVI value, several other rules are applied in the form of decision tree: these include cloud masking and a-priori assumptions on reflectance range in visible and near infrared region. An example of the synthetic cloud free mosaic and the individual dates used for its construction is shown on Fig. 3.

2.1.3 Forest health assessment system

In the presented Czech methodology, health status is not assessed as absolute amount of leaf biomass (having LAI as proxy for leaf biomass), but on its change over time. The basic premise is that the health status can be objectively determined only by observing the relative change in LAI over time. In the first step, we calculate selected vegetation indexes (e.g. Normalized Difference Vegetation Index – NDVI, Red Edge Inflection Point – REIP and Normalized Difference Infrared Index – NDII), and image transformations (e.g. components of Tasseled Cap transformation) and compare their sensitivity against in-situ data from sampled plots (e.g. LAI and ICP Forests plots). For each dataset, a linear regression models between in-situ data and Sentinel–2 indices were calculated and evaluated. For indices yielding best linear fit, neural network was trained and applied per-pixel to retrieve prediction LAI maps (Fig. 4).

Classification system of forest health is proposed, which will evaluate forest health on pixel level as a change in LAI values over time and classify each pixel in the following five categories:

I. significant increase: increase in LAI by 1.5 and higher
II. moderate increase: increase in LAI from 0.5 to 1.5
III. stable conditions: change of LAI between −0.5 and 0.5
IV. moderate decrease: decrease in LAI from −1.5 to −0.5
V. significant decrease: decrease in LAI higher than −1.5

The countrywide assessment of forest health is carried out on cadastral level, where the area of forest stands of classes IV and V are evaluated for the total forest area of cadastral for stands of age between 0 and 80 years. This condition is put due to the fact that it is not possible to distinguish between sanitary logging and planned logging for old-grown forests – both will be reflected by a sharp decrease in LAI values. Each cadastral is assigned in one of the following categories:
The methodology is based on a two-phase regression sampling. The first phase estimates forest health directly from satellite images. The second phase refines the initial assessment using data from a ground defoliation assessment. Defoliation was set as the main indicator of a visual symptom of health of trees. Even though the value of spectral reflectivity recorded with sensors and the defoliation are not of an identical phenomenal base, most works proved a close relation between them (e.g. Gross 2000; Joria & Ahearn 1997).

2.2.1 Ground data and forest mask

The ground data on defoliation were collected by visual inspection on 112 permanent monitoring plots, distributed regularly over the afforested territory of Slovakia in a grid 16 × 16 km, using methodology of the ICP Forests program. Each plot has a squared shape with 50 m long sides (ICP Forests 2010; Pavlenda et al. 2014). Defoliation, defined as needle/leaf loss in the assessable crown as compared to a reference tree, is observed regardless of the cause of foliage loss and assessed in a 5% interval. Visual inspection is carried out on yearly basis. Defoliation expresses the percentage of missing assimilation organs to a fully leafed tree, i.e. 0% – healthy tree, 100% – dead tree. To identify forest area and distribution of main tree species in Slovakia, a raster map prepared by Bucha (1999) on the base of Landsat scenes is used. This source is preferred as it contains information on tree species distribution within forestry management units which is the main advantage against the more actual data taken from forest management plans.

2.2.2 Phase 1

Alternative 1

Two components are derived from the original satellite bands using an orthogonal transformation; the components are referred to as new synthetic channels (NSC1 and NSC2). Only red (R), near infrared (IR) and shortwave infrared (SWIR) bands are used. The second SWIR band (band 7 in Landsat products) was utilized in analyses for 2015 – 2017. Mathematic apparatus for the components derivation follows Jackson (1983). There is a difference in the way of defining the physical meaning of the derived components. In Jackson’s approach, the first component represents brightness. It is derived from points with low and high reflectivity, ordinarily represented with moist soil on one side and dry soil on the other side. The second component (greenness) represents the amount of green vegetation. In our study, the NSC1 represents a spectral variability of tree species. The NSC2 component is optimized for defoliation estimation.

Fully foliaged stands with very different reflectivity (coniferous and broadleaves) were chosen and the differences of their reflectance were calculated as:
\[ b_j = (X_s - X_b) \]

where: \( i = R, IR \) and SWIR bands; \( X_s \) and \( X_b \) represent reflectance values of spruce (coniferous) and beech (broadleaves) stands.

The low reflectance values represent spruce stands; high reflectivity is typical of beech. These values (low and high) were calculated as mean value from several plots identified manually on the satellite scene.

A standardization of vector \( b = (b_1, b_2, b_3) \) was carried out into the form of unit vector by dividing each of vector elements by a normalization factor \( B \):

\[ A_{ij} = b_j / B \quad \text{where} \quad B = \left( \sum_{i=1}^{3} b_i^2 \right)^{1/2} \]

and \( A_{ij} \) are coefficients used to derive the first component NSC1:

\[ NSC1 = A_{1,1} * X_1 + A_{1,2} * X_2 + A_{1,3} * X_3 \]

where \( X_i \) are reflectance values of pixel in \( i \)-th band.

To derive the second component, the reflectance values representing maximum defoliation (dead stands \( - X_{\text{dead tree}} \)) are calculated as well as the difference to the first component:

\[ g_i = (X_{\text{dead tree}} - X_i) - D_{2,1} * A_{1,1} \]

\[ D_{2,1} = \sum_{i=1}^{3} (X_{\text{dead tree}} - X_i) * A_{1,1} \]

Such a process ensures orthogonality of \( b \) and \( g \) vectors. Similarly, as in the first step, the \( g \) vector is standardized by means of the normalization factor \( G \), and transformation coefficients are calculated:

\[ A_{ij} = g_i / G \]

Calculation of NSC2 is as follows:

\[ NSC2 = A_{2,1} * X_1 + A_{2,2} * X_2 + A_{2,3} * X_3 \]

Calculated values of NSC2 component represent a perpendicular distance from the line of reflectance values of component NSC1 defined by the fully foliated trees of spruce and beech on its borders. The distance from the line is proportional to the extent of the tree damage in the given pixel.

**Alternative 2**

An alternative approach takes benefit of a moderate correlation in defoliation between subsequent years (Bucha & Barka 2014). This allowed simplifying the 1st phase of regression sampling. Instead of calculating the NSC2 component, a defoliation was estimated in a given year \( t \) on the basis of defoliation in the previous (or next) assessment \( t-1 \) (or \( t+1 \)), i.e. \( \text{def}_{t} = f(R_t, IR_t, SWIR) \) where independent variables are the reflectance values of satellite bands from year \( t \). Multiple linear regression is used to analyse the relationship. The defoliation predicted in this way can be considered to be a proxy of NSC2. In 2017 assessment, derivation of NSC2 component from actual satellite scenes was used as an input for correlation analysis instead of classification from previous year (Table 2).

### 2.2.3 Phase 2

The 2nd phase is based on the evaluation of defoliation ICP Forest dataset, complemented with data from the Forest Protection Service Centre and other field measurements, e.g. known dead stands with defoliation 100% to cover full range of defoliation values (0 – 100%, see Fig. 1). For the needs of health classification, the average defoliation on plot was calculated. The total number of plots for each year is given in Table 2.

Mean spectral reflectance and standard deviation are found for each monitoring plot. This spectral characteristic is paired with respective defoliation value. Then, a regression model is derived by means of a simple linear regression analysis between the output of the first stage (NSC2) and defoliation. Finally, the defoliation for each pixel of the NSC2 raster is assessed by means of the derived regression equation.

The result of forest stand classification is the percent defoliation. For the presentation purposes, the forest stands were considered as damaged if their defoliation exceeded 40%. It is noteworthy that forest stands after felling (both regular and accidental felling) are included in this category. Stands with defoliation below 40% were classified as healthy.

### 3. Results

#### 3.1. Czechia

**3.1.1 Developing forest health assessment system**

Scattergrams and coefficients of determination of linear regression indicated that NDVI and REIP vegetation indices calculated from the Sentinel–2 data were not found to be related to LAI values measured in the field. The determination coefficients are very low for both indices \( R^2 = 0.01 \) and the relationships are distinctive for coniferous and deciduous forest stands. The Tasseled Cap transformation of the DI’ index (Wetness – Greenness component) has a moderate sensitivity to field data \( R^2 = 0.29 \), but the relationships are species-specific. The highest potential for prediction of ground-based LAI values have indices NDII \( R^2 = 0.57 \), DI \( R^2 = 0.63 \) and Wetness component of Tasseled Cap transformation \( R^2 = 0.58 \). All three products also yield identical relationships for coniferous and deciduous stands (Fig. 6).

For the development of LAI prediction model, NDII, DI and Wetness indices were found to be the most suitable ones. In the training phase we designed a neural network with good theoretical performance (correlation coefficient \( R = 0.77 \), Fig. 7a) and very good prediction of LAI from Wetness values \( R = 0.86 \), Fig. 7b). Neural network was stored as Matlab function and applied per pixel to Wetness values for Sentinel–2 cloud free mosaics of 2015, 2016 and 2017. Fig. 8 shows an example of the retrieved LAI map.
3.1.2 Validation of leaf area index maps against in-situ data

Leaf area index

First validation was based on the comparison of field measurements of LAI against per pixel retrieval of LAI from Sentinel–2 mosaic of 2017. This was done by plotting the measured LAI against retrieved LAI values and evaluating its root mean square error (RMSE). We found a good correspondence of measured and retrieved LAI, with values close to the 1:1 line. The RMSE of LAI retrieval was 0.53 and mean absolute error (MAE) was 0.41 (Fig. 9). The validation is, however, not truly independent as some of the validation data were used in model development.

Defoliation levels from ICP Forest database

In the next step, we compared the ICP Forest site-level defoliation values from 2015 against the LAI values obtained by LAI prediction model for Wetness component of the Sentinel–2 mosaic from 2015 (see 3.1). We extracted the defoliation values for 194 ICP Forests plots and compared them against LAI values obtained by the prediction model. Here we observe a strong linear relationship between site-level defoliation and retrieved LAI ($R^2 = 0.58$), with increase in defoliation with decreasing LAI (Fig. 10).

3.1.3 Developing forest health assessment system – a synthesis

The final concept of the monitoring system is as follows:

1) Sentinel–2 data processing chain has been developed, which produces time series of high-quality, cloud-free mosaics for the Czech Republic and for a user-selected time period (see 2.2)

2) A predictive model for LAI retrieval from the Wetness component of the Tasseled Cap image transformation of Sentinel-2 reflectance.
The model has been successfully validated against in-situ data (see 3.1.2). There was a strong relationship between LAI values and site-level defoliation observed from ICP Forests data. Finally, we also found a good correspondence of the predicted LAI with forest health assessment based on independent airborne hyperspectral dataset with a very high spatial and spectral resolution.

Prediction model was applied on cloud-free image mosaics of 2016, 2017 and 2018 to yield maps of LAI and its change in the studied time period. Difference LAI maps were then used to classify each forested pixel in one of the five classes of forest health.Cadastre-level assessment of forest health is then performed based on the percentage share of class IV and V forests (high and medium LAI decrease) to total forested area of cadastre (Fig. 11). In Table 1 total areas of different forest health classes for the area of Czech Republic is given.

### 3.2. Slovakia

#### 3.2.1 Models for forest health status classifications

Statistical models for estimation of forest health condition based on the two-phase approach are shown in Table 2. Pearson correlation coefficients between reflectance in actual satellite composition and defoliation in previous year (1st phase of the assessment) varied from 0.47 to 0.70. Correlations between 1st and 2nd phase were in the range 0.85 – 0.97. Standard error ($s_y$) of regression lines for the 2nd phase varied from 2.1 to 11.0 % of defoliation. This measure could be used for interpretation of accuracy. For example, $s_y = ± 8.9%$ in 2013 means that if defoliation is determined at 30%, it can actually range from 21.1 to 38.9% at 68% confidence, respectively from 12.2 to 47.8% at 95% confidence.

#### 3.2.2 Evaluation of forest health status

The total share of damaged stands (i.e. with defoliation higher than 40%) increased from 6–8% in 2003–2011 to 13–15% in 2012–2017 (Fig. 12 and 13). Situation in coniferous stands was worse by ca 5–10% compared to broadleaved tree species for most of the studied period. The most damaged tree species in Slovakia in the period 2003–2017 was spruce, with the share of damaged stands reaching as much as 22.8% of the total spruce stands area in 2013 (Fig. 12 and Table B1 in Appendix B). Pine (*Pinus sylvestris*) showed higher damage than other tree species as well, with the peak at 30.2% in 2017. The most damaged broadleaved tree species was hornbeam (16.8% in 2013).

### 4. Discussion

Presented methodologies strongly depend on the quality of input satellite data. While Slovak methodology combine scenes from Sentinel–2 and Landsat satellites in the input mosaics and thus takes benefit of longer data archive, Czech methodology is based on the Sentinel–2 data dating back to 2015 only. Optical satellite data are particularly sensitive to cloud cover, atmospheric aerosol content and low solar illumination. While low solar...
illuminations is not a problem in Central Europe during the peak of vegetation season (July – August) and aerosol contamination can be partly removed by atmospheric correction or by a careful selection of input satellite scenes, both presented methodologies rely on accurate cloud and shadow identification.

Despite automatic cloud cover detection algorithms have recently significantly improved (Zhu & Woodcock 2012, 2014; Zhu et al. 2015), the best results can be obtained by manual identification of clouds and shadows, especially in case of low cloud cover and low number of scenes (in case of Slovakia there are 7–8 scenes a year, what makes the manual detection feasible). Still, improved automated identification of clouds and shadows would be a great benefit for the whole monitoring system, future research should focus on this issue. A novel approach to cloud and shadow detection has been developed in the frame of the Czech methodology, where number of input scenes is considerably higher (with respect to the size of territory). Despite the use of more frequent observations acquired by the pair of Sentinel–2 satellites, and use of advanced methods for automatic synthesis of cloud free mosaics using the “big data” approach, some areas still remain contaminated by clouds. This leads to potentially incorrect assessment in some pixel in time series of LAI change.

In addition to the predictive power of developed model, its resistance to atmospheric effects also highlights its efficiency. As each of the indices uses different wavelength, their resistance indeed differs. Specifically, the Wetness component of Tasseled Cap transformation yielded significantly better results compared to NDII and DI (results not shown here).

Sentinel–2 system has a very good specification for high-quality automated atmospheric corrections. The system contains several spectral bands, which are able to capture instantaneous atmospheric parameters and to detect different types of clouds, including cirruses, which are usually difficult to detect. Still, automated cloud detection and atmospheric correction are not perfect. For
methodology evaluates forest health conditions via vegetation indices. The latter approach strongly benefits from the prior suitability analysis of selected indices for prediction of LAI.

Satellite data are generally not able to interpret the cause of the observed change in forest health status. While the Slovak methodology does not distinguish between standing forests and harvested compartments (planned or sanitary harvests), the Czech methodology addressed this issue by analysing stands up to 80 years of age, where planned forest logging is not expected to occur. This gives that the forest health assessment on cadastral level may not be representative if cadastre unit is dominated by mature stands with worsened health conditions.

In the Slovak methodology, entire classifications for the whole afforested area is published as well as input satellite mosaics. In the Czech approach, raster classification is further processed and the interpretation of the difference maps LAI is done for the cadastral areas which differ in area and forest cover. Considering that the methodology classifies the cadastres into individual classes based on the share of Class IV and V forest stands to total forest area, the classification may not be relevant for cadastres with little forest cover.

Evaluation of forest conditions using optical remote sensing is often based on spectral (vegetation) indices. In such a case, the results strongly depend on index evaluated. It is obvious that selection of an appropriate index and knowledge of its properties is crucial for proper use of the method (Pickel et al. 2015). The two methodologies differ in their approach to the use of spectral data. While the Slovak methodology is based on direct relationship between defoliation and spectral reflectance, the Czech methodology evaluates forest health conditions via vegetation indices. The latter approach strongly benefits from the prior suitability analysis of selected indices for prediction of LAI.

Satellite data are generally not able to interpret the cause of the observed change in forest health status. While the Slovak methodology does not distinguish between standing forests and harvested compartments (planned or sanitary harvests), the Czech methodology addressed this issue by analysing stands up to 80 years of age, where planned forest logging is not expected to occur. This gives that the forest health assessment on cadastral level may not be representative if cadastre unit is dominated by mature stands with worsened health conditions.

In the Slovak methodology, entire classifications for the whole afforested area is published as well as input satellite mosaics. In the Czech approach, raster classification is further processed and the interpretation of the difference maps LAI is done for the cadastral areas which differ in area and forest cover. Considering that the methodology classifies the cadastres into individual classes based on the share of Class IV and V forest stands to total forest area, the classification may not be relevant for cadastres with little forest cover.

Methodology developed by Stoklasa Tech was previously used to assess the health status of the Czech forests. Since it has never been presented in detail and published,
we can only assume that it uses the time series of the MSI or NDII vegetation index from the Landsat data. The experience of previous authors (e.g. Lambert et al. 1995; Campbell et al. 2004) also leads to doubts about the possibility of re-evaluating the absolute value of 10 classes of defoliation of stands (from 0 to 100%) based on Landsat multispectral data without performing an annual land defoliation survey to calibrate the prediction model.

In addition to a high share of damaged spruce stands, high defoliation is visible in pine stands in Slovakia. However, these results could be affected by lower density of tree crowns and state of undergrowth vegetation. Particularly pine stands are sensitive, from the view of selected methodology, to the conditions of grass vegetation in the understory. If satellite scene used for evaluation comes from relatively dry period, damage rate in the stands with lower density can be overestimated compared to terrestrial measurements.

ICP Forest data usually lack stands with defoliation higher than 50 or 60% (Fig. 1). As inclusion of such data indicates higher than 50% defoliation, the proportion of damaged stands tends to be lower than expected. This is especially true for the methodology, to the conditions of grass vegetation in the understory. If satellite scene used for evaluation comes from relatively dry period, damage rate in the stands with lower density can be overestimated compared to terrestrial measurements.

**Table 2. Derived models of two-phased sampling method used for damage classification in 2003 – 2017.**

<table>
<thead>
<tr>
<th>Year</th>
<th>Phase</th>
<th>Model</th>
<th>( r )</th>
<th>( \text{st. error} )</th>
<th>Size of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2003} = 16.93 + 0.83 \text{IR}</em>{2003} - 0.52 \text{IR}<em>{2003} - 0.46 \text{SWIR}</em>{2003} ]</td>
<td>0.55</td>
<td>±11.0</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2003</td>
<td>[ \text{SAO}<em>{2003} = -18.16 + 1.34 \text{SAO}</em>{2003} ]</td>
<td>0.86</td>
<td>±7.4</td>
<td>n2: 24 plots</td>
</tr>
<tr>
<td>2006</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2006} = 26.8 + 0.28 \text{IR}</em>{2006} - 0.47 \text{IR}<em>{2006} + 0.59 \text{SWIR}</em>{2006} ]</td>
<td>0.48</td>
<td>±5.0</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2006</td>
<td>[ \text{SAO}<em>{2006} = -38.26 + 2.34 \text{SAO}</em>{2006} ]</td>
<td>0.97</td>
<td>±7.7</td>
<td>n2: 94 plots</td>
</tr>
<tr>
<td>2007</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2007} = 30.50 + 0.54 \text{IR}</em>{2007} + 0.56 \text{IR}<em>{2007} + 0.57 \text{SWIR}</em>{2007} ]</td>
<td>0.57</td>
<td>±7.4</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2007</td>
<td>[ \text{SAO}<em>{2007} = -17.85 + 1.57 \text{SAO}</em>{2007} ]</td>
<td>0.87</td>
<td>±7.4</td>
<td>n2: 94 plots</td>
</tr>
<tr>
<td>2009</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2009} = 33.70 + 0.06 \text{IR}</em>{2009} - 0.61 \text{IR}<em>{2009} + 0.72 \text{SWIR}</em>{2009} ]</td>
<td>0.48</td>
<td>±2.1</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2009</td>
<td>[ \text{SAO}<em>{2009} = -46.34 + 2.58 \text{SAO}</em>{2009} ]</td>
<td>0.85</td>
<td>±7.7</td>
<td>n2: 98 plots</td>
</tr>
<tr>
<td>2010</td>
<td>1st phase: Derivation of NSC1 and NSC2 components:</td>
<td>[ \text{NSC2} = 0.3631 \text{R}_{2010} ]</td>
<td>0.68</td>
<td>±0.0</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment 2009 (SPOT scenes classified into 10 classes – spruce forests only)</td>
<td>[ \text{SAO}_{2010} = 3.06 + 0.197 \text{NSC2} ]</td>
<td>0.68</td>
<td>±0.0</td>
<td>n2: all spruce pixels</td>
</tr>
<tr>
<td>2011</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2011} = 28.9 + 0.14 \text{IR}</em>{2011} - 0.84 \text{IR}<em>{2011} + 0.99 \text{SWIR}</em>{2011} ]</td>
<td>0.60</td>
<td>±9.0</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2011</td>
<td>[ \text{SAO}<em>{2011} = -12.55 + 1.74 \text{SAO}</em>{2011} ]</td>
<td>0.88</td>
<td>±8.9</td>
<td>n2: 104 plots</td>
</tr>
<tr>
<td>2012</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2012} = 26.9 + 0.14 \text{IR}</em>{2012} - 0.83 \text{IR}<em>{2012} + 0.97 \text{SWIR}</em>{2012} ]</td>
<td>0.70</td>
<td>±9.7</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2012</td>
<td>[ \text{SAO}<em>{2012} = 31.66 + 2.11 \text{SAO}</em>{2012} ]</td>
<td>0.95</td>
<td>±9.7</td>
<td>n2: 24 plots</td>
</tr>
<tr>
<td>2013</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2013} = -25.27 + 0.53 \text{IR}</em>{2013} - 0.54 \text{IR}<em>{2013} + 1.08 \text{SWIR}</em>{2013} ]</td>
<td>0.65</td>
<td>±8.9</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2013</td>
<td>[ \text{SAO}<em>{2013} = -26.77 + 1.78 \text{SAO}</em>{2013} ]</td>
<td>0.86</td>
<td>±8.9</td>
<td>n2: 100 plots</td>
</tr>
<tr>
<td>2015</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2015} = 1.346 + 0.036 \text{IR}</em>{2015} - 0.027 \text{IR}<em>{2015} + 0.068 \text{SWIR}</em>{2015} + 0.037 \text{SWIR}_{2015} ]</td>
<td>0.65</td>
<td>±10.45</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2015</td>
<td>[ \text{SAO}<em>{2015} = -24.41 + 16.64 \text{SAO}</em>{2015} ]</td>
<td>0.86</td>
<td>±9.7</td>
<td>n2: 97 plots</td>
</tr>
<tr>
<td>2016</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2016} = 1.2846 + 0.0001 \text{IR}</em>{2016} - 0.0131 \text{IR}<em>{2016} + 0.0027 \text{SWIR}</em>{2016} + 0.0024 \text{SWIR}_{2016} ]</td>
<td>0.59</td>
<td>±10.41</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2016</td>
<td>[ \text{SAO}<em>{2016} = -15.13 + 1.5239 \text{SAO}</em>{2016} ]</td>
<td>0.87</td>
<td>±10.41</td>
<td>n2: 90 plots</td>
</tr>
<tr>
<td>2017</td>
<td>1st phase: Estimation by classification</td>
<td>[ \text{SAO}<em>{2017} = -189.8 + 2.746 \text{SAO}</em>{2017} ]</td>
<td>0.60</td>
<td>±3.4</td>
<td>n1: all forest pixels</td>
</tr>
<tr>
<td></td>
<td>2nd phase: Assessment on monitoring plots in 2017</td>
<td>[ \text{SAO}<em>{2017} = -189.8 + 2.746 \text{SAO}</em>{2017} ]</td>
<td>0.88</td>
<td>±3.4</td>
<td>n2: 86 plots</td>
</tr>
</tbody>
</table>
It is planned that both methodologies will further develop in the future. Data from Synthetic Aperture Radar (SAR) satellite sensors are increasingly being used for monitoring of forest cover and its changes (Schmullius et al. 2015). The research is often driven by effort to evaluate growing stock (e.g. Santoro et al. 2011, 2013), carbon stock (Thurner et al. 2014) or forest disturbances (Ranson et al. 2014). SAR data can be used alone or in combination with optical data (Reiche et al. 2018). Combining data from different sensors in detection algorithms seems to have an increasing trend, and attempts to combine outputs from several detection algorithms are conducted as well (Healey et al. 2018). Much of the newly developed algorithms is based on trend analysis in longer time periods (Zhu et al. 2015).

Cloud-based processing of time series of high temporal and spatial resolution data like SAR, Sentinel-1 and optical Sentinel–2 will further enhance capabilities for incorporating time domain in the analyses of forest health status. This will allow not only to increase capacity for cloud-free image generation, but also to increase the frequency of assessment. Temporal trends in forest reflectance may be also used e.g. to improve species classification and include more detailed species maps.

With the availability of new remote sensing datasets acquired, for example by multispectral and hyperspectral UAV cameras or very high temporal resolution micro-satellites (e.g. Planet Labs), an entirely new type data analysis in forestry emerges. Whereas very high spatial resolution data from UAVs may be applied on regional level e.g. to identify pre-visual stress on individual tree level, daily revisit time of Planet system allows for near-real time monitoring of forest operations in very high spatial resolution. Our capacity for incorporation of this type of data is however limited due to the additional costs related to data acquisition (UAVs) and access (commercial satellite systems such as Planet Labs). Therefore, these advances are unlikely to be included in regular forest health assessment in the Czech Republic and Slovakia in the near future.

5. Conclusions

Nationwide assessment of the forest health status in the Czech Republic is based on the analysis of synthetic cloud-free Sentinel–2 satellite observations using a novel analysis of all-available satellite observations (so called “spatio-temporal synthesis”) in a big data environment. Using several independent datasets containing information on health status (e.g. site-level defoliation assessed within ICP Forests plots, in-situ measurements of LAI on NFI plots, airborne hyperspectral data) an advanced machine learning techniques are applied to predict LAI from Sentinel–2 transformed images. Forest health status is assessed at pixel-scale as a change in LAI for a given time interval. Final assessment of the state of health is applied at the cadastral level, where each cadastral unit is classified into four health classes based on the fraction of the lowest health status stands with significant LAI decrease out of the total forest cover for stands up to 80 years of age.

The procedure for spruce forest status assessment based on two phase sampling was described and applied in Slovakia. The procedure uses R, IR and SWIR satellite bands. The first phase is based either on the component (NSC2) optimized for the damage estimation or on classification from the previous year. The second phase refines estimations by ground data and multi-regression analysis. The high correlations and low standard errors between 1st and 2nd phase of regression sampling proved the method’s suitability for consistent classification of forest status and time series analysis. Compositions of satellite scenes as well as classifications of forest status were made available at http://www.nlcks.sk/stales where forest health status maps can be explored using different web map tools. The boundaries and identification of forestry management units can be overlaid in these applications to localize visible changes. The applications can be particularly useful for forest managers, state administration as well as for a wider public.

The ability for rapid and cost effective nationwide assessment of forest health status is the main strength of both presented methodologies. On the other hand, their weaknesses stem in demanding procedure for creation of cloud free mosaics of satellite scenes, exclusion of stands older than 80 years in the Czech approach and mixing of harvested stands and severely damaged stand in the same categories in the Slovak approach. Data fusion with Sentinel 1 and other radar datasets seems to be the most promising way of future methodological development in both countries.

Acknowledgements

This work was supported by the Slovak Research and Development Agency under Grants APVV-15-0413, APVV-0439-12 and APVV-16-0325. Partial support was provided also by grant “EVA4.0”, No. CZ.02.1.01/0.0/0.0/16_019/0000803 financed by OP RDE. The project was implemented also with the support of Horizon 2020 “DataBio” project and Czech Science Foundation project “Radiation balance of forest stands of the Czech Republic", GJ17-05608Y. We thank the Forestry and Game Management Research Institute for providing data on deforestation from the ICP Forests project and IT-Innovations for providing access to computing infrastructure within the pilot implementation of the Sentinel–2 satellite data processing chain. We also thank the staff of Forest Management Institute from the Frýdek-Místek, Kroměříž, Brno, Jablonec nad Nisou and Plzeň branches for helping to collect digital hemispherical photographs on the National Forest Inventory plots.
References


Bartold, M., 2016: Development of forest cover mask to monitor the health condition of forests in Poland using long-term satellite observations. Leśne Prace Badawcze / Forest Research Papers, 77:141–150.


Appendix figures and tables

**Fig. A1.** Sampling scheme for acquisition of digital hemispherical photographs. The images were taken with a Nikon D5500 digital SLR camera with a Sigma 4.5 mm circular fisheye lens. The camera was placed on a Vanguard Espod CX203 AP tripod and aligned horizontally with a two-axis level. All photos were shot with lens facing north and taken as RAW uncompressed images.

**Fig. A2.** Location of National forest inventory plots sampled for digital hemispherical photography. All field plots were visited during the period of maximum vegetation foliage, for 2016 and 2017 in June to August, while in 2015 was the test period, where photos were taken only for evergreen coniferous plots, mostly in October. All DHP photos were analysed in Hemisfer software (WSL, Switzerland). The software uses the LAI value inversion from angular distribution of canopy gaps for a set of statistically representative set of images.

**Table A1.** Basic descriptive statistics of retrieved leaf area index (LAI) values for sampled plots.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of sampled sites</th>
<th>Average LAI</th>
<th>Maximum LAI</th>
<th>Minimum LAI</th>
<th>Standard deviation of LAI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>32</td>
<td>2.52</td>
<td>3.50</td>
<td>1.75</td>
<td>0.13</td>
</tr>
<tr>
<td>2016</td>
<td>45</td>
<td>3.12</td>
<td>5.22</td>
<td>0.49</td>
<td>1.03</td>
</tr>
<tr>
<td>2017</td>
<td>112</td>
<td>2.04</td>
<td>3.48</td>
<td>0.72</td>
<td>0.26</td>
</tr>
<tr>
<td>2015–2017</td>
<td>189</td>
<td>3.15</td>
<td>5.22</td>
<td>0.49</td>
<td>0.61</td>
</tr>
</tbody>
</table>

**Table A2.** Share of damaged forest stands in 2003–2017 out of the total areas occupied by a given tree species (Slovakia).

<table>
<thead>
<tr>
<th>Forest type</th>
<th>Area [ths. ha]</th>
<th>Share of damaged stands [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spruce</td>
<td>427.554</td>
<td>7.9</td>
</tr>
<tr>
<td>Fir</td>
<td>78.455</td>
<td>3.5</td>
</tr>
<tr>
<td>Pine</td>
<td>100.595</td>
<td>17.7</td>
</tr>
<tr>
<td>Other conifers</td>
<td>43.042</td>
<td>6.9</td>
</tr>
<tr>
<td>Conifers</td>
<td>649.647</td>
<td>8.8</td>
</tr>
<tr>
<td>Beech</td>
<td>585.803</td>
<td>6.2</td>
</tr>
<tr>
<td>Oaks</td>
<td>297.157</td>
<td>6.1</td>
</tr>
<tr>
<td>Hornbeam</td>
<td>113.577</td>
<td>8.2</td>
</tr>
<tr>
<td>Other broadleaves</td>
<td>122.171</td>
<td>12.0</td>
</tr>
<tr>
<td>Broadleaves</td>
<td>1,118.708</td>
<td>7.0</td>
</tr>
<tr>
<td>Total</td>
<td>1,768.355</td>
<td>7.7</td>
</tr>
</tbody>
</table>