

A bark beetle infestation predictive model based on satellite data in the frame of decision support system TANABBO

Renata Ďuračiová⁽¹⁾,
Milan Muňko⁽¹⁾,
Ivan Barka⁽²⁾,
Milan Koreň⁽³⁾,
Karolina Resnerová⁽⁴⁾,
Jaroslav Holuša⁽⁴⁾,
Miroslav Blaženec⁽⁵⁾,
Mária Potterf⁽⁶⁾,
Rastislav Jakuš⁽⁴⁻⁵⁾

The European spruce bark beetle *Ips typographus* L. causes significant economic losses in managed coniferous forests in Central and Northern Europe. New infestations either occur in previously undisturbed forest stands (*i.e.*, spot initiation) or depend on proximity to previous years' infestations (*i.e.*, spot spreading). Early identification of newly infested trees over the forested landscape limits the effective control measures. Accurate forecasting of the spread of bark beetle infestation is crucial to plan efficient sanitation felling of infested trees and prevent further propagation of beetle-induced tree mortality. We created a predictive model of subsequent year spot initiation and spot spreading within the TANABBO decision support system. The algorithm combines open-access Landsat-based vegetation change time-series data, a digital terrain model, and forest stand characteristics. We validated predicted susceptibility to bark beetle attack (separately for spot initiation and spot spreading) against beetle infestations in managed forests in the Bohemian Forest in the Czech Republic (Central Europe) in yearly time steps from 2007 to 2010. The predictive models of susceptibility to bark beetle attack had a high degree of reliability (area under the ROC curve – AUC: 0.75-0.82). We conclude that spot initiation and spot spreading prediction modules included within the TANABBO model have the potential to help forest managers to plan sanitation felling in managed forests under pressure of bark beetle outbreak.

Keywords: Spatial Predictive Model, Bark Beetle Infestation, GIS, ROC Curve, Norway Spruce

Introduction

Bark beetles cause severe damage to coniferous forests in the Northern Hemisphere (Raffa et al. 2015). Spatio-temporal analyses of bark beetle population dynamics are the subject of study for several research teams in Europe (Kärvemo et al. 2014, Havašová et al. 2017, Mezei et al. 2017) and in North America (Thatcher et al. 1980, Simard et al. 2012, Meddens & Hicke 2014, Senf et al. 2015). Spatially explicit models predicting the risk of tree mortality are needed to efficiently mitigate damage related to bark beetle attacks (Kärvemo et

al. 2014).

Predictive modelling creates a model of future behaviour or occurrence of a phenomenon. In geographic information systems (GIS), spatial predictive modelling aims to predict the occurrence of a certain phenomenon in unknown areas based on its occurrence in space. As such, spatial predictive modelling relies on a multi-layered spatial database that contains attribute information (slope, aspect, distance from important objects, etc.) and their distribution in space (Gomarasca 2009). The spatial database is transformed into an ag-

gregation of attribute values to obtain an index of suitability (or of the contrasting propensity) of the landscape to an occurrence of the modelled phenomenon. Various mathematical and machine learning methods have been applied to create spatial predictive models (Kuhn & Johnson 2013), such as multiple linear regression or logistic regression (Pontius & Schneider 2001, Jakuš et al. 2005, 2017), multi-criteria decision-making using crisp or fuzzy set theory (Lieskovsky et al. 2013, Caha et al. 2014), Dempster Shafer theory, Bayesian theory, and support vector machines (Karell et al. 2017). The result of spatial predictive modelling is usually a raster with cell values between 0 and 1, where a high value indicates a high probability of occurrence of the modelled phenomenon.

Several databases or statistical models based on a GIS approach have also been developed to estimate bark beetle population dynamics in order to predict future damage or optimize forest protection measurements (Thatcher et al. 1980, Simard et al. 2012, Kärvemo et al. 2014, Meddens & Hicke 2014), as operational and public data-based models are missing.

Satellite remote sensing offers the opportunity for repeated mapping of tree mortality, and Landsat-resolution (30 m) imagery is sufficient to map the dynamics of tree mortality due to bark beetles within stands during the epidemic phase of a bee-

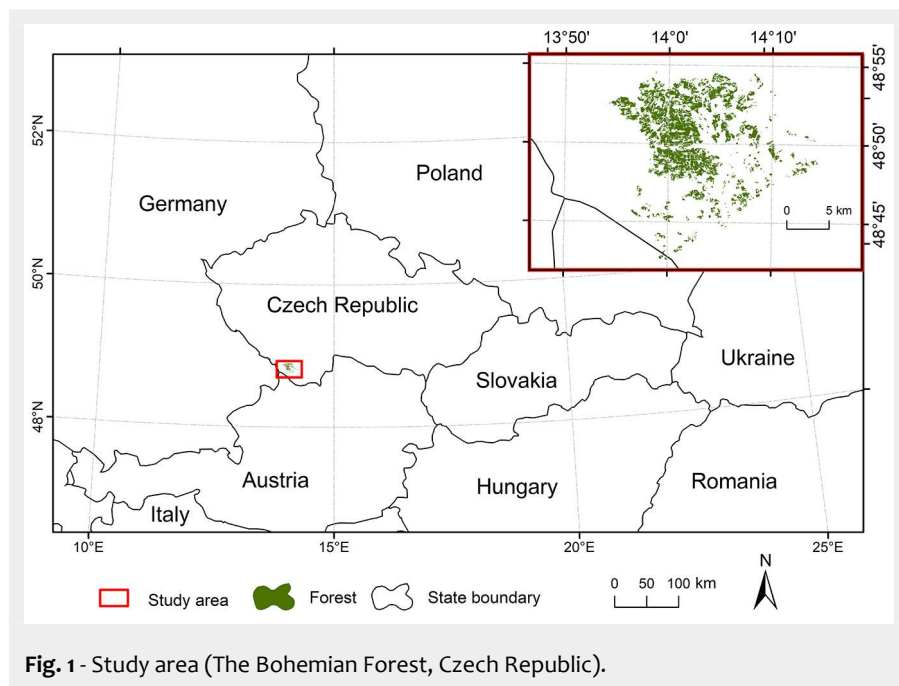
□ (1) Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Bratislava (Slovakia); (2) National Forest Centre, Forest Research Institute, Zvolen (Slovakia); (3) Faculty of Forestry, Technical University in Zvolen, Zvolen (Slovakia); (4) Faculty of Forestry and Wood Sciences, Czech University of Life Sciences, Prague (Czech Republic); (5) Institute of Forest Ecology, Slovak Academy of Sciences, Zvolen (Slovakia); (6) Department of Biological and Environmental Sciences, University of Jyväskylä, Jyväskylä, Finland

@ Rastislav Jakuš (rasti.jakus@gmail.com)

Received: Oct 17, 2019 - Accepted: Apr 05, 2020

Citation: Ďuračiová R, Muňko M, Barka I, Koreň, M, Resnerová K, Holuša J, Blaženec M, Potterf M, Jakuš R (2020). A bark beetle infestation predictive model based on satellite data in the frame of decision support system TANABBO. *iForest* 13: 215-223. - doi: [10.3832/ifor3271-013](https://doi.org/10.3832/ifor3271-013) [online 2020-06-06]

Communicated by: Massimo Faccoli



tle outbreak (Wulder et al. 2006, Meddens & Hicke 2014, Havašová et al. 2015). Open-access Landsat-based classifications record the removal of canopy tree cover (Hansen et al. 2014) or monitor forest health (Barka et al. 2018, Rossi et al. 2019). The monitoring of forest defoliation (Bucha & Barka 2010) and parametrized country-wide main tree species can even reveal subtle canopy changes due to beetle attack (Havašová et al. 2017). Satellite data are reliable for the estimation of bark-beetle-caused tree mortality, especially in the later stages of bark beetle outbreak (Havašová et al. 2015, 2017). However, bark-beetle-caused tree mortality interacts with wind damage (wind-caused tree mortality – Mezei et al. 2014a). This means that, despite many advantages, the ability to use satellite data for the mapping of bark-beetle-caused tree mortality has certain limitations.

Satellite data, especially Landsat products, together with relevant forest stand characteristics, are used to model tree mortality (Kissiyar et al. 2005, Simard et al. 2012, Kärvmö et al. 2014, Meddens & Hicke 2014). The bark-beetle-induced tree mortality predictors are usually various stand characteristics obtained from forestry databases or from remote sensing sources (Simard et al. 2012, Kärvmö et al. 2014, Meddens & Hicke 2014) and various landscape characteristics (Kissiyar et al. 2005, Simard et al. 2012).

The TANABBO (Tatra Mountains Bark beetle infestation prediction model – Kissiyar et al. 2005, Jakuš et al. 2005) decision support system combines freely available satellite imagery with stand characteristics to predict the occurrence of beetle-induced tree mortality in subsequent years. TANABBO is also a GIS-based system for the evaluation of forest stand predisposition to bark beetle attacks (Jakuš et al. 2017). The system is based on the known

causal links between bark beetle outbreaks and environmental parameters. A module on the prognosis of a bark beetle stand infestation is also a part of the system. The system captures vegetation change over time using a time-series of Landsat images, NDVI (the normalized difference vegetation index), a digital terrain model at the resolution of Landsat pixel size, and available stand characteristics (stand age, volume, average diameter, and average height) at the stand resolution from forest management databases (Jakuš et al. 2003). A system of partial models is used, where each partial model produces an output which is later used as an input for the main model. The main output from the forecasting part of TANABBO is a subsequent-year prediction of bark-beetle-caused tree mortality. This forecast is important for the planning of forest protection measures in the subsequent year. The one-year forecast of bark-beetle-caused tree mortality is based on the combination of the two processes related to the spread of tree mortality over the landscape: the initiation of bark beetle spots and the spots spreading (Jakuš et al. 2003, Robertson et al. 2007). Spot initiation means the creation of new bark beetle infestations in areas without previous attacks, which is difficult to predict and relates to long-distance beetle dispersal occurring even during beetle epidemics (Jakuš et al. 2005). Spot spreading represents the expansion of an existing bark beetle spot into a neighbouring forest stand (Jakuš et al. 2005). Spot spreading becomes dominant over the course of the beetle outbreak, and in later phases, the majority of new infestations are adjacent to previous years' infestations (Kautz et al. 2011, Potterf et al. 2019).

Each predictive model requires quality determination before its use. One of the most widely used methods to validate the re-

sults of predictive modelling is the Receiver Operating Characteristic (ROC) curve analysis, first introduced in signal detection theory (Egan 1975). In ROC curve analysis, the area under the ROC curve (AUC) is a very popular metric of prediction quality (Metz 1978, Fawcett 2006). This method is also useful to measure the quality of spatial predictive models and maps (Pontius & Schneider 2001, Lieskovsky et al. 2013), as well as to compare them.

The aim of our work is (i) to introduce an algorithm of an updated module to predict landscape-level spot bark beetle spot initiation and spot spreading within the TANABBO decision support system methodology, and (ii) to validate the model outcomes against recorded proxies of bark beetle mortality data from public resources in conditions of managed forests in 2008, 2009, and 2010 using ROC curve analysis.

Materials and methods

Study area and main disturbances

The case study area is a Division of “Vojenské lesy a statky” (VLS, Military Forests and Properties), s.p., Horní Planá in the Czech Republic (Fig. 1). VLS is a special-purpose organization managing military training areas. The forest enterprise manages 19,960 ha of land. Forests represent 16,569 ha of this area, and a water reservoir covers 203 ha; the rest of the area is covered by grassland under intensive military activity (<https://www.vls.cz/en>). The rugged hills with an altitude change of 100-150 m are a landscape typical of the region. Most of the area is of altitude between 600 and 800 m, with the highest point at 1228 m a.s.l. (Lysá Mt). The area partly belongs to the Bohemian Forest (the Šumava Mts.) and partly to the Šumavské podhůří Mts. The mean annual temperature varies between 5 and 7 °C, and annual precipitation totals vary between 700 and 800 mm (Culek 1996).

Norway spruce (*Picea abies* L. Karst) is the dominant tree species in the forests (69%), very often growing with silver fir (*Abies alba* Mill., 6%) and European beech (*Fagus sylvatica* L., 5%). Scots pine (*Pinus sylvestris* L., 12%) is abundant at lower altitudes. Forest stands younger than 40 years cover 23.5% of the area, and 32% of stands are older than 100 years (<https://www.vls.cz/en>). Natural conditions are represented by *Querceto-Fagetum* to *Fageto-Piceetum* forest altitudinal zones. Forests form a large complex in higher altitudes surrounded by meadows and pasturelands.

Wind, snow, and bark beetles are the main disturbances to spruce forests. Windthrows (föhn is a common phenomenon in the study area) are the most important disturbance factor. In recent years, incidental felling due to being windthrown and bark beetle infestations have accounted for approximately 50% of total felling. Bark beetle outbreaks in the study area follow trends of bark beetle populations in the entire

Czech Republic, with peaks in the mid-1980s and mid-1990s. A long-term outbreak started in 2003 following a strong drought in the whole of Central Europe (Rouault et al. 2006). This outbreak was triggered by windstorm Kyrill in 2007 (Fink et al. 2009). Since 2015, a new bark beetle outbreak started as a result of drought and warm vegetation seasons.

Prediction framework

The predictive model involved known parameters controlling the occurrence of tree mortality by spruce bark beetle (see more details in Input data). We used yearly data from 2007 to 2010 to predict and validate the outcomes. The study area was hit by a large windstorm, Kyrill, in January 2007 (Fink et al. 2009); therefore, we chose the three years after the storm (2008, 2009, and 2010) to capture newly created beetle infestations during the beetle outbreak.

Input data to predict beetle infestations

Age structure, percentage of spruce, stand density, and wood volume values were taken from the forest management plan; potential solar radiation (Hofierka & Suri 2002) was derived from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM, 30 m resolution – <http://srtm.csi.cgiar.org/>). NDVI was calculated from Landsat 5 and 7 satellite data (<https://landsat.gsfc.nasa.gov/>). For 2008, the NDVI raster layer was based on Enhanced Thematic Mapper Plus (ETM+) scenes from July 28 and June 10; for 2009, on the TM (Thematic Mapper) scene from August 24; and for 2010, on the TM scene from July 10. The selection of data was to a large extent limited by the public availability of satellite time-series and GIS data. The

selection of stand characteristics from forest management databases was based on our previous work (Mezei et al. 2014a, 2014b).

Ground-truth data on bark-beetle-caused tree mortality

To validate the predictive model, model quality parameters (performance metrics) were calculated using maps of bark-beetle-caused tree mortality from 2008, 2009, and 2010 (Fig. 2).

As a proxy for the localization of bark-beetle-caused tree mortality, we used Landsat-based and expert-classified forest health data available as Web map service (WMS) layers from the Czech Institute of Forest Management (*Ústav pro hospodářskou úpravu lesů Brandýs nad Labem – ÚHÚL*); these data were available until 2016. The portal has published classified Landsat time-series data (forest stands health status) since 1994 (Stoklasa 2003, Barka et al. 2018). These data were accepted as a good estimation of forest health in the Czech Republic and in the study area, where they corresponded to areas infested by bark beetles. The forest health status layers contain the following categories: 0, healthy stands; 0/I, first signs of damage; I, small damage; II, medium damage; IIIa, strong damage; IIIb/IV, very strong damage. We transformed the forest health data to bark-beetle-caused tree mortality. We considered pixels classified to the classes “IIIa, strong damage” and “IIIb/IV, very strong damage” as bark-beetle-caused tree mortality, based on expert knowledge of the study area. The pixels classified as bark-beetle-caused tree mortality are considered as such only in the first year of occurrence. The yearly combi-

nation of beetle occurrences resulted in a time series of bark-beetle-caused tree mortality.

For the next step, we measured the distances between new and previous years’ infestation spots (Euclidean distance tool in ArcGIS® v. 10.5 – Esri, Redlands, CA, USA). Based on the proximity to the previous year’s infestation, we considered pixels more distant than 30 m (Landsat resolution, delineating neighbouring pixels) from the previous year’s infestation as spot initiation; otherwise, they were considered spot spreading.

Predictive maps of susceptibility to a bark beetle attack

In our study, the prediction of bark beetle attack probability was based on the combination of susceptibility to bark beetle spot initiation and susceptibility to the spreading of existing bark beetle spots.

Prediction of susceptibility to bark beetle spot initiation

Based on the TANABBO methodology (Jakuš et al. 2017) we applied logistic regression (Hair et al. 2010) to estimate susceptibility to the initiation of bark beetle spots over the landscape. Logistic regression is a special form of multivariate regression analysis in which the dependent variable is a non-metric, dichotomous (binary) variable. It is also the case with spatial predictive models, because we predict the occurrence of phenomena, *i.e.*, whether or not they are present at a given site. This is also the reason why the better-known linear regression is not suitable for this type of classification problem.

Tree mortality (presence of pixels with dead trees) caused by spot initiation was

Fig. 2 - Map of spruce bark beetle infestations in the study area (The Bohemian Forest, Czech Republic) in 2008-2010.

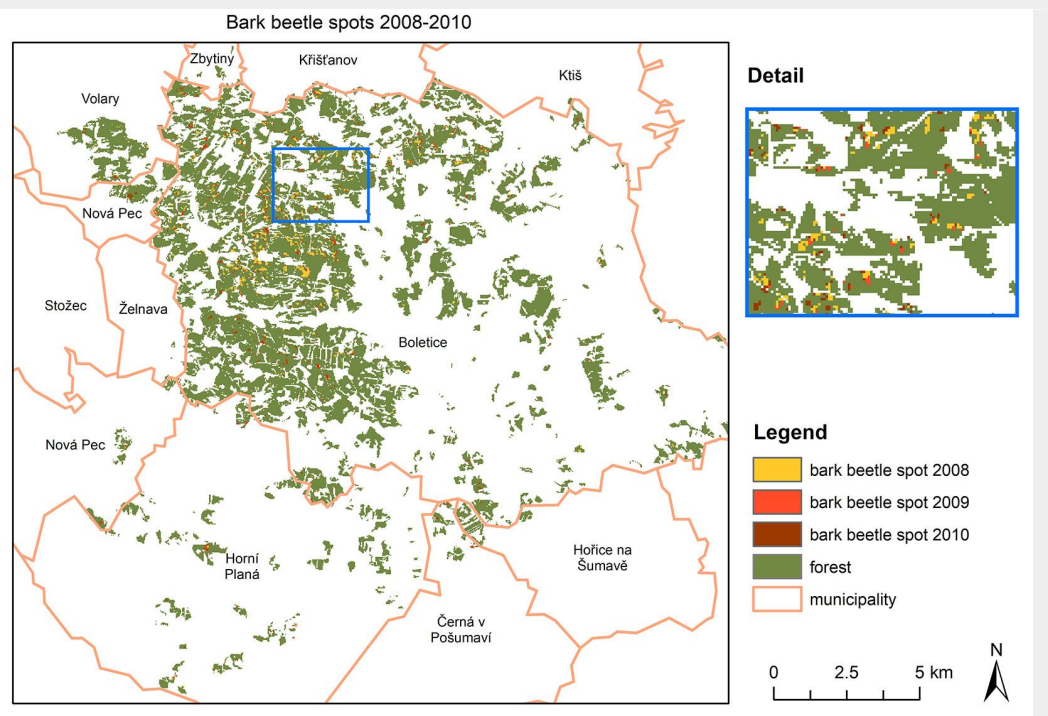
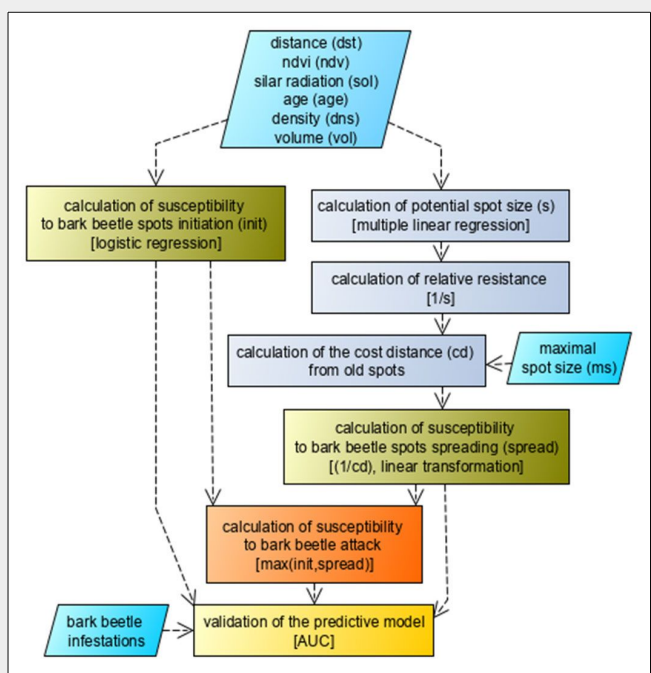


Fig. 3 - A flowchart of the methodology of creation and validation of the predictive models of the susceptibility of an area to bark beetle attack.



Prediction of susceptibility to a bark beetle spot spreading

The relationship between active bark beetle spot size and independent factors (the same as in the case of modelling susceptibility to bark beetle spot initiation) was the basis for calculations of the potential of a bark beetle spot spreading (Jakuš et al. 2005). In the frame of raster GIS, it could be understood as the estimation of the likelihood that a given cell would be added to an existing spot, thus increasing the spot's size.

First, for each bark beetle spot, the spot size was computed. This means that for each raster cell, the spot size containing that cell was calculated. Cells not affected by bark beetle spots have a value of 0. The potential for the bark beetle spot size was calculated using linear regression. The size of the active bark beetle spot was considered as the dependent factor, and the other parameters (distance, NDVI, solar radiation, age structure, stand density, and wood volume per hectare) were considered as the independent factors. All factors were normalized before entering the calculation, i.e., their values were transformed to values ranging from 0 to 1. The resulting multiple linear regression function was used for the calculation of the potential bark beetle spot size for each pixel. The raster of the relative resistance was computed from the potential bark beetle spot size by the function $1/s$, where s is a potential spot size. Based on a GIS algorithm of movement of a friction surface (Eastman 1999), the model of relative resistance was used as a friction surface in the computation of cost distances from existing bark beetle spots. The maximum spreading distance was derived from the maximum spot

the dependent factor and was represented as presence (1) or absence (0) of new (initial) bark beetle spots in the analysed year. The distance from existing spots (dst), NDVI (ndv), solar radiation (sol), age structure (age), stand density (dns), and wood volume per hectare (vol) were the independent factors in the statistical analyses. Before calculating the logistic regression, we statistically analysed the significance of the input parameters by calculating the Gini index (Gini 1912) and mutual independence by calculating the correlation coefficients. The Gini coefficient, sometimes called the Gini index or Gini ratio (Gini 1912), is a mea-

sure of statistical dispersion used primarily in economics, but it is also useful in geostatistics (Wackernagel 2003). The Gini index measures the inequality among the values of a variable, in our case between the values of input factors in the healthy forest and the values in the bark beetle spots. We applied it to calculate the significance of input factors in spatial predictive modelling.

The result of the logistic regression in each raster cell (corresponding to the input raster data) forms a prediction model of spot initiation.

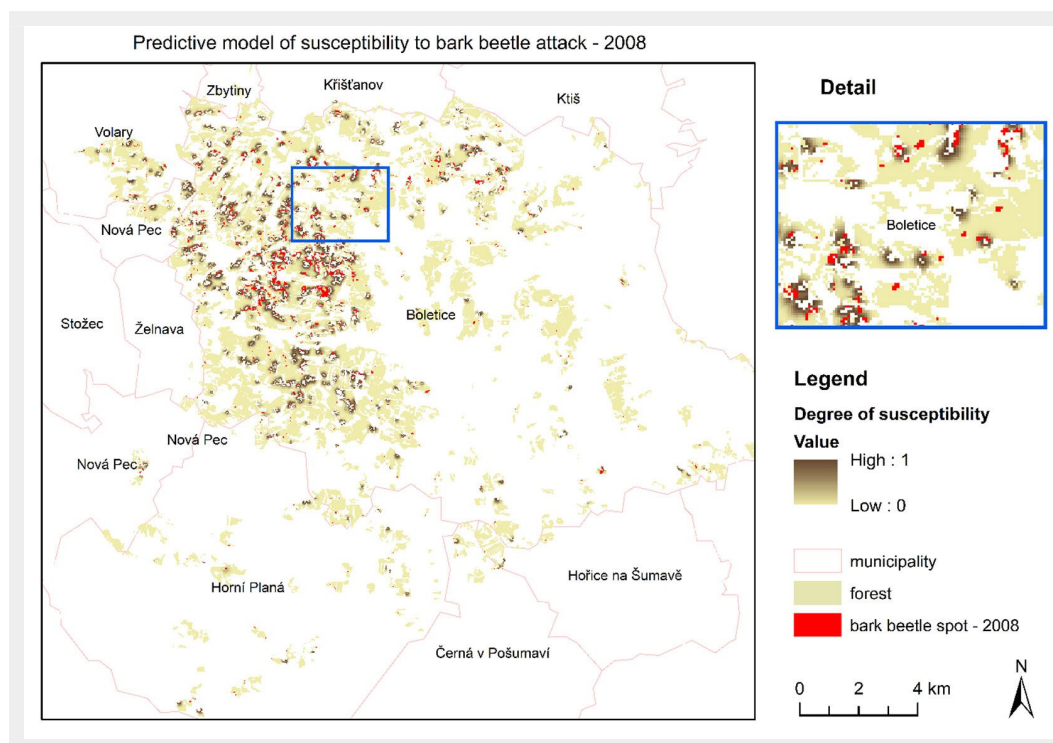


Fig. 4 - The predictive model of the susceptibility of the study area (The Bohemian Forest, Czech Republic) to bark beetle attack in 2008.

size in the study area. The cost distances were linearly resized to 0-1 to create a predictive model of the probability of bark beetle spot spreading (maximum cost distance corresponds to 0, distance 0 m corresponds to 1).

A combined predictive model of susceptibility to bark beetle attack

A model of susceptibility to bark beetle attack was created based on the model of bark beetle initiation and the model of bark beetle spreading using the maximum operation (the MAX() function in map algebra). In other words, each value of the model of susceptibility to bark beetle attack represented the higher corresponding value from the two predictive models described above. The flowchart of the methodology of the creation of the predictive models is shown in Fig. 3. Details of the complete methodology of the creation of the predictive models are described in previous papers (Jakuš et al. 2005, 2017).

Validation of the predictive models and predictive maps

In this paper, we propose a methodology and new module of the TANABBO decision support system for the validation of all predictive models. We applied ROC curve analysis according to Pontius & Schneider (2001), where AUC is computed using the trapezoidal rule of integral calculus. To validate the predictive model for each year, we used reference bark beetle spots detected in the subsequent year (Fig. 4). We proposed the application of two types of validation. In the first case (combined), we validated all models, i.e., models of susceptibility to the initiation, spreading, and attack, using all bark beetle spots detected in the subsequent year. In the second (separated) case, we validated the models of susceptibility to bark beetle attack initiation only using the initialized bark beetle spots and the models of susceptibility to bark beetle spots spreading only using the spread spots.

Data processing and analysis

We performed most of the data preparation and processing in the TANABBO v. 1.9 software system (Jakuš et al. 2017 – <http://www.tanabbo.org>). The TANABBO 1.9 enables time-series analysis of data on spruce stand mortality. Statistical analysis of input parameters, logistic regression, and ROC curve analysis were performed with the Python programming language and the libraries NumPy, SciPy, Matplotlib, and scikit-learn.

All predictive models were visualized by predictive maps created in the ArcGIS® v. 10.5 software environment.

Results

In the case study, we created predictive models of the susceptibility of an area to a bark beetle attack for 2008, 2009, and 2010. For each model, we calculated pa-

Tab. 1 - Gini index of all input factors for 2008, 2009, and 2010, which expresses the difference in the distribution of individual factors in bark beetle spots compared to the distribution in the forest (the larger the absolute value of the index, the greater the difference).

Input factor	Year		
	2008	2009	2010
Distance from spots	-0.45	-0.64	-0.48
NDVI	-0.16	-0.24	-0.19
Solar radiation	-0.12	-0.12	-0.15
Age structure	0.19	0.21	0.15
Stand density	-0.13	-0.07	-0.04
Wood volume ha ⁻¹	0.16	0.16	0.15

Tab. 2 - Correlation matrix of input factors.

Input factor	Distance from spots	NDVI	Solar radiation	Age structure	Stand density	Wood volume ha ⁻¹
Distance from spots	1.00	-	-	-	-	-
NDVI	0.11	1.00	-	-	-	-
Solar radiation	0.00	-0.03	1.00	-	-	-
Age structure	-0.15	-0.01	-0.02	1.00	-	-
Stand density	0.12	-0.10	-0.02	-0.18	1.00	-
Wood volume ha ⁻¹	-0.10	0.01	0.00	0.44	0.17	1.00

rameters of its accuracy based on the proposed methodology.

Statistical analysis and predictive model creation

The results of preliminary statistical analysis (correlation matrix and Gini index calculation) are presented in Tab. 1 and Tab. 2.

The analysis was performed on 734,608 (937 × 784) raster cells representing the study area. The results show the statistical significance of the input parameters (Tab. 1) and their mutual independence (no or low correlation – Tab. 2).

The coefficients of regression functions for the creation of the predictive model of

Tab. 3 - Coefficients of logistic regression functions for the creation of the predictive models of susceptibility to the bark beetle spot initiation for 2008, 2009, and 2010.

Parameters	Year		
	2008	2009	2010
Distance from spots	-36.7099	-26.9146	-18.9125
NDVI	-1.6011	-0.0525	-0.5652
Solar radiation	-0.0240	-0.2789	-0.2934
Age structure	1.0524	0.6254	1.0125
Stand density	-0.0004	-0.0021	0.0001
Wood volume ha ⁻¹	-0.3194	-0.4225	-0.3119

Tab. 4 - Coefficients of linear regression functions for potential bark beetle spot spreading calculations for 2008, 2009, and 2010 (all parameter values were normalized).

Parameters	Year		
	2008	2009	2010
Constant	0.732	0.598	0.731
Distance	-0.535	-0.622	-0.443
NDVI	-0.607	-0.564	-0.591
Solar radiation	-0.303	-0.212	-0.178
Age structure	0.47	0.419	0.416
Stand density	0.192	0.267	0.043
Wood volume ha ⁻¹	-0.572	-0.545	0.521

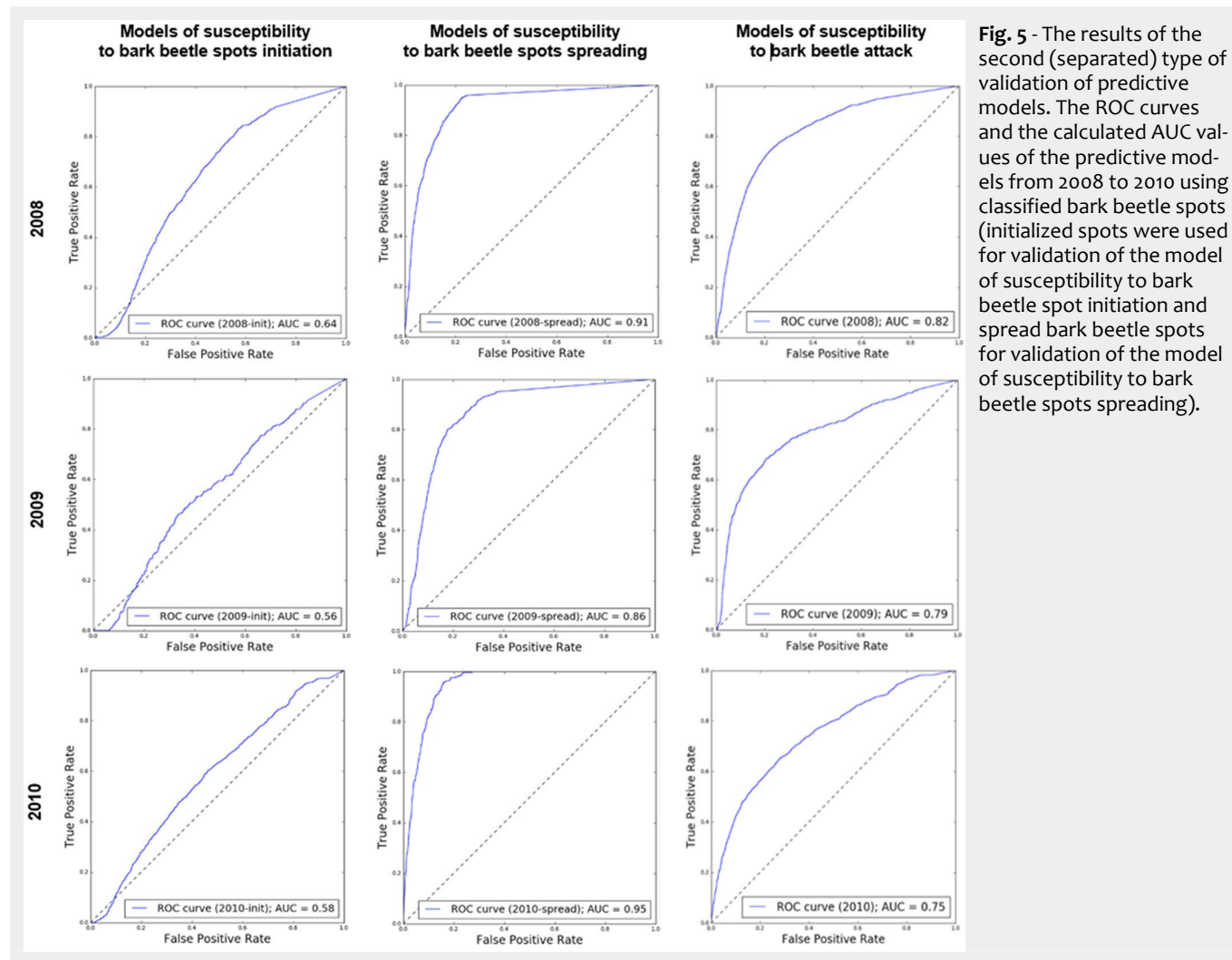


Fig. 5 - The results of the second (separated) type of validation of predictive models. The ROC curves and the calculated AUC values of the predictive models from 2008 to 2010 using classified bark beetle spots (initialized spots were used for validation of the model of susceptibility to bark beetle spot initiation and spread bark beetle spots for validation of the model of susceptibility to bark beetle spots spreading).

susceptibility to bark beetle spot initiation from 2008 to 2010 are shown in Tab. 3, and the coefficients of linear regression functions for potential spot size calculation are shown in Tab. 4. The size of the coefficients also indicates the relative importance of the input factors.

The maximal spot size during the analysis period in the study area was 52,200 m²; thus, we used the value 228 m (calculated as a square root) as the maximal distance of bark beetle spot spreading in constructing the predictive model of susceptibility to a bark beetle spot spreading (cells beyond a distance of 228 m from existing spots were ineligible).

The predictive models of the susceptibil-

ity of an area to bark beetle attack were created using both the above-described models and the MAX() function. All predictive models are shown as predictive maps in Fig. 4 and Figs. S1-S2 in Supplementary material.

Validation of predictive models

Fig. S3 (Supplementary material) shows the validation of the combined spot initiation and spread predictive models (e.g., 2008, AUC=0.82). The validation for separate spot initiation and spot spreading is shown in Fig. 5. The results of the validation of all predictive models created within our case study are summarized in Tab. 5, which contains the mean of all predictive

values (m_a), mean of predictive values in bark beetle spots (m_s), and the AUC for each model. For the prediction of susceptibility to bark beetle spot initiation, the AUC values ranged from 0.71 to 0.82. The AUC values for the model of susceptibility to a bark beetle spot spreading ranged from 0.72 to 0.76.

The combined type of validation shows relatively high accuracy in the prediction of all infested areas. Separated types of validation show better results for bark beetle spot spreading.

Discussion

The use of public data derived from the Landsat time series is sufficient for monitoring beetle outbreaks during epidemic phases (Havašová et al. 2015). However, mid-resolution satellite scenes represent an important data source. We used a national web service (Stoklasa 2003, Barka et al. 2018), with limited references. Havašová et al. (2017) showed that the national web service data could be more precise in bark beetle damage identification than data from the well-referenced Global Forest Watch service (Hansen et al. 2014).

The main limitation of our reference tree mortality data was the difficulty in distinguishing between wind damage and bark-

Tab. 5 - Results of the validation of predictive models from 2008 to 2010. (m_a): mean of all predictive values of the model; (m_s): mean of predictive values in bark beetle spots; (AUC): area under ROC curve. The higher the m_s value compared to the m_a value, the more reliable the model.

Year of prediction	m_a	m_s	AUC
2008	0.110	0.435	0.82
2009	0.063	0.242	0.79
2010	0.182	0.631	0.75

beetle-caused tree mortality. However, our study was performed in a period after extensive windstorm damage, when bark beetles were the dominant cause of tree mortality, which was the subject of the study. This allowed us to use public data about forest health as a proxy of bark beetle damage.

Our models show relatively good results in the early stages of the bark beetle outbreak. According to Økland et al. (2016), this is a period when an attack is relatively hard to predict. Our study provides testing and extension of the TANABBO decision support system and also the design and implementation of a module for its validation.

Statistical analysis and predictive model creation

The predictors used herein are similar to predictors applied by other authors (Simard et al. 2012, Kärvmemo et al. 2014). We used Landsat-based NDVI as an important predictive variable. Its significance was also confirmed by statistical analysis (Gini index and coefficients of regression equations). Meddens & Hicke (2014) also used several Landsat-based indices, including NDVI, as explanatory variables to predict bark-beetle-caused tree mortality. The use of data on tree mortality from the previous year is in agreement with the approach used by Kärvmemo et al. (2014). During an outbreak, infestations are usually spatially concentrated and show a high spatial and temporal autocorrelation pattern (Simard et al. 2012). Kautz et al. (2011) and Simard et al. (2012) used the distance to the closest local infestation as a bark beetle pressure variable. This parameter also proved to be significant in our study. Other important factors are wood volume per ha, age, and potential solar radiation. On the contrary, the least significant parameter in our analysis was density. The statistical significance of the solar radiation parameter was not so high in our analysis, but this parameter alone has the potential for more accurate determination based on more accurate digital models created by aerial laser scanning. Nevertheless, the results showed that the predictive models of susceptibility to bark beetle attack from the TANABBO decision support system provided predictions of subsequent years' infestations with relatively high accuracy ($AUC = 0.75-0.82$).

Validation of whole predictive models

Modelling complex spatiotemporal phenomena such as forest susceptibility to bark beetle attack is a very challenging problem, which deals with complex nonlinear conditioning parameters that are often dynamic and change dramatically over time. Therefore, validation results above 0.90 in these types of modelling are extremely hard to achieve. In our study, the AUC values of the final models ranged from 0.75 to 0.82 (Tab. 5). Phillips & Dudik (2008) suggested that models with AUC values above 0.75 should be considered as

useful for predictions (Ortiz et al. 2013). Therefore, the validation result of $AUC = 0.75$ can be considered as reasonable, and the susceptibility model can be considered as usable and reliable. Kärvmemo et al. (2014) used boosted regression trees for large-scale risk mapping of an eruptive bark beetle. The model was validated for multiple years and gained AUC results ranging from 0.729 to 0.818.

An interesting point is that our model was more accurate for the prediction of bark beetle attack in 2008 ($AUC = 0.82$) than in subsequent years. This result indicated that our model better predicted beetle infestations just after wind damage (from 2007, otherwise not included in study), which is likely explained by the spatial dependence of proximity to wind-thrown trees in subsequent years (Potter et al. 2019).

Validation of partial models

Our two types of validation (validation of models of all bark beetle infestations and separate validation of bark beetle infestations created by spot initiation and infestation caused by existing spots spreading) helped us more deeply understand underlying processes. The first (combined) type of validation did not show large differences between the partial models and the final predictive model (AUC for initiation: 0.71-0.82, AUC for spreading: 0.70-0.76, final predictive models AUC 0.75-0.82). One of the key variables in both models is the distance from the infestations of the previous year, therefore the partial models are, to a certain extent, interchangeable.

On the other hand, the second type of validation (separate) showed significantly better prediction of the processes of existing spot spreading (AUC 0.86-0.95) than prediction of new bark beetle infestations (AUC 0.56-0.64). This is caused by the significance of distance from the previous year's infestations and especially by hypothetical overlay between spot initiation and spot spreading at short distances from the last year's infestations. We can see that on ROC curves for spot initiation, the graph is even below the diagonal for small false positive rate (FPR) values, and vice versa; for spreading, the ROC curve grows rapidly. To summarise, it is easier to predict spot spreading than spot initiation (Økland et al. 2016).

The final effect of the described processes is the confirmation of relatively high accuracy of the final predictive model.

Implementation in forest management

The described and validated algorithm for prediction of bark beetle infestation is a core part of the TANABBO decision support system (Kissiyar et al. 2005, Jakuš et al. 2005, 2017). The one-year prediction can potentially be applied in the planning of forest protection measures in areas affected by bark beetle outbreaks or areas with extensive wind damage and expected

bark beetle outbreak. Important limitations of the practical use of our models are the difficulties in distinguishing between wind- and bark-beetle-caused tree mortality with the use of Landsat data. On the other hand, the interaction between wind- and bark-beetle-caused tree damage (Mezei et al. 2014a) could possibly improve those negative effects on model accuracy.

A limitation of Landsat data is the fact that in the mountainous conditions of central Europe, there is usually only one cloud-free Landsat scene available in the course of one year.

Conclusions

The beetle infestation model implemented within the TANABBO decision support system (Kissiyar et al. 2005, Jakuš et al. 2005, 2017) relatively accurately predicts tree mortality as a combination of spot initiation and spot spreading over the landscape ($AUC = 0.75-0.82$).

Acknowledgements

This work was supported by grants nos. 1/0300/19 and 1/0468/20 from the Grant Agency of Slovak Republic VEGA, the project of the Slovak Research and Development Agency under contracts APVV-15-0413, APVV-15-0761, and the National Agency for Agricultural Research within the framework of project no. QJ1220317 administered by the Ministry of Agriculture of the Czech Republic and by the grant "EXTEMIT - K", no. CZ.02.1.01/0.0/0.0/15_003/0000433 financed by OP RDE.

References

- Barka I, Lukeš P, Bucha T, Hlásny T, Strejček R, Mlčoušek M, Krístek S (2018). Remote sensing-based forest health monitoring systems - case studies from Czechia and Slovakia. *Central European Forestry Journal* 64: 259-275. - doi: [10.1515/forj-2017-005](https://doi.org/10.1515/forj-2017-005)
- Bucha T, Barka I (2010). Satellite-based regional system for observation of forest response to global environmental changes. In: "Advances in Geoinformation Technologies" (Horák J, Halounová L, Hlásny T, Kusendová D, Vozenílek V eds). Technical University of Ostrava, Ostrava, Czech Republic, pp. 1-14.
- Caha J, Nevtířilová V, Dvorsky J (2014). Constraint and preference modelling for spatial decision making with use of possibility theory. In: *Proceedings of the "Hybrid Artificial Intelligence Systems"* (Polycarpou M, De Carvalho ACPLF, Pan JS, Wozniak M, Quintian H, Corchado E eds). Salamanca (Spain) 11-13 June 2014. Springer International Publishing, Switzerland, pp. 145-155. - doi: [10.1007/978-3-319-07617-1_13](https://doi.org/10.1007/978-3-319-07617-1_13)
- Culek M (1996). Biogeografické členění České republiky [Biogeographical division of the Czech Republic]. Enigma, Praha, Czech Republic. [in Czech]
- Eastman JR (1999). Guide to GIS and image processing - Volume 2. IDRISI production Clark University, Worcester, UK, pp. 170.
- Egan J (1975). Signal detection theory and ROC analysis. Academic Press, New York, USA, pp.

- 277.
- Fawcett T (2006). An introduction to ROC analysis. *Pattern Recognition Letters* 27: 861-871. - doi: [10.1016/j.patrec.2005.10.010](https://doi.org/10.1016/j.patrec.2005.10.010)
- Fink AH, Brücher T, Ermert V, Krüger A, Pinto JG (2009). The European storm Kyrill in January 2007: synoptic evolution, meteorological impacts and some considerations with respect to climate change. *Natural Hazards and Earth System Sciences* 9: 405-423. - doi: [10.5194/nhess-9-405-2009](https://doi.org/10.5194/nhess-9-405-2009)
- Hair JF, Black WC, Babin BJ, Anderson RE (2010). *Multivariate data analysis* (7th edn). Pearson Education Limited, London, UK, pp. 734.
- Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO, Townshend JRG (2014). High-resolution global maps of 21st-century forest cover change. *Science* 324: 850-853. - doi: [10.1126/science.1244693](https://doi.org/10.1126/science.1244693)
- Gini C (1912). Variabilità e mutabilità [Variability and mutability]. Reprinted in: "Memorie di metodologica statistica", 1955 (Pizetti E, Salvemini T eds). Libreria Eredi Virgilio Veschi, Rome, Italy, pp. 156. [in Italian]
- Gomarasca MA (2009). *Basics of geomatics*. Springer-Verlag, New York, USA, pp. 697. [online] URL: <http://books.google.com/books?id=B AQ3FJiXDGsc>
- Havašová M, Bucha T, Ferenčík J, Jakuš R (2015). Applicability of a vegetation indices-based method to map bark beetle outbreaks in the High Tatra Mountains. *Annals of Forest Research* 58: 295-310. - doi: [10.15287/afr.2015.388](https://doi.org/10.15287/afr.2015.388)
- Havašová M, Ferenčík J, Jakuš R (2017). Interactions between windthrow, bark beetles and forest management in the Tatra national parks. *Forest Ecology and Management* 391: 349-361. - doi: [10.1016/j.foreco.2017.01.009](https://doi.org/10.1016/j.foreco.2017.01.009)
- Hofierka J, Šúri M (2002). The solar radiation model for open source GIS: implementation and applications. In: *Proceedings of the "Open source GIS - GRASS user's conference 2002"* (Ciulli M, Zatelli P eds). Trento (Italy) 11-13 Sept 2002. University of Trento, Trento, Italy, pp. 19.
- Jakuš R, Grodzki W, Jezik M, Jachym M (2003). Definition of spatial patterns of bark beetle *Ips typographus* (L.) outbreak spreading in Tatra Mountains (Central Europe), using GIS. In: *Proceedings of the Conference "GTR NE-311 Survey and Management of Forest Insects"* (Mc Manus M, Liebhold A eds). Kraków (Poland) 1-5 Sept 2002. USDA Forest Service, Newtown Square, PA, USA, pp. 25-32. [online] URL: <http://www.nrs.fs.fed.us/pubs/9117>
- Jakuš R, Jezik M, Blazenec M (2005). Prognosis of bark beetle attacks in TANABBO model. In: "GIS and Databases in the Forest Protection in Central Europe" (Grodzki W eds). Centre of Excellence PROFEST at the Forest Research Institute, Warsaw, Poland, pp. 35-43.
- Jakuš R, Blazenec M, Koren M, Barka I, Lukášová K, Lubojacký J, Holuša J (2017). TANABBO II model pro hodnocení rizika napadení lesních porostů lýkožroutem smrkovým *Ips typographus* (L.) [TANABBO II model for risk of Ips typographus (L.) attack assessment]. *Lesnický průvodce* 1/2017. Výzkumný ústav lesního hospodářství a myslivosti, v. v. i., Jilovište, Czech Republic, pp. 71. [in Czech]
- Karell L, Munko M, Duračiová R (2017). Applicability of support vector machines in landslide susceptibility mapping. In: "The Rise of Big Spatial Data, Lecture Notes in Geoinformation and Cartography" (Ivan I, Singleton A, Horák J, In-spektor T eds). Springer International Publishing AG, Cham, Switzerland, pp. 373-386. - doi: [10.1007/978-3-319-45123-7_27](https://doi.org/10.1007/978-3-319-45123-7_27)
- Kautz M, Dworschak K, Gruppe A, Schopf R (2011). Quantifying spatio-temporal dispersion of bark beetle infestations in epidemic and non-epidemic conditions. *Forest Ecology and Management* 262: 598-608. - doi: [10.1016/j.foreco.2011.04.023](https://doi.org/10.1016/j.foreco.2011.04.023)
- Kärveemo S, Van Boeckel TP, Gilbert M, Grégoire JC, Schroeder M (2014). Large-scale risk mapping of an eruptive bark beetle - Importance of forest susceptibility and beetle pressure. *Forest Ecology and Management* 318: 158-166. - doi: [10.1016/j.foreco.2014.01.025](https://doi.org/10.1016/j.foreco.2014.01.025)
- Kissiyar O, Blazenec M, Jakuš R, Willekens A, Jezik M, Baláz P, Valckenborg JV, Celer S, Fleischer P (2005). TANABBO model: a remote sensing based early warning system for forest decline and bark beetle outbreaks in Tatra Mts. - overview. In: "GIS and Databases in the Forest Protection in Central Europe" (Grodzki W eds). Centre of Excellence PROFEST at the Forest Research Institute, Warsaw, Poland, pp. 15-34.
- Kuhn M, Johnson K (2013). *Applied predictive modeling*. Springer Science + Business, New York, USA, pp. 600.
- Lieskovsky T, Duračiová R, Karell L (2013). Selected mathematical principles of archaeological predictive models creation and validation in the GIS environment. *Interdisciplinaria Archaeologica - Natural Sciences in Archaeology* IV (2): 33-46. [online] URL: <http://www.researchgate.net/publication/261223595>
- Meddens AJH, Hicke JA (2014). Spatial and temporal patterns of Landsat-based detection of tree mortality caused by a mountain pine beetle outbreak in Colorado, USA. *Forest Ecology and Management* 322: 78-88. - doi: [10.1016/j.foreco.2014.02.037](https://doi.org/10.1016/j.foreco.2014.02.037)
- Metz CE (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine* 8: 283-298. - doi: [10.1016/S0001-2998\(78\)80014-2](https://doi.org/10.1016/S0001-2998(78)80014-2)
- Mezei P, Grodzki W, Blazenec M, Jakuš R (2014a). Factors influencing the wind-bark beetles' disturbance system in the course of an *Ips typographus* outbreak in the Tatra Mountains. *Forest Ecology and Management* 312: 67-77. - doi: [10.1016/j.foreco.2013.10.020](https://doi.org/10.1016/j.foreco.2013.10.020)
- Mezei P, Grodzki W, Blazenec M, Skvarenina J, Brandysova V, Jakuš R (2014b). Host and site factors affecting tree mortality caused by the spruce bark beetle (*Ips typographus*) in mountainous conditions. *Forest Ecology and Management* 331: 196-207. - doi: [10.1016/j.foreco.2014.07.031](https://doi.org/10.1016/j.foreco.2014.07.031)
- Mezei P, Jakuš R, Pennerstorfer J, Havašová M, Skvarenina J, Ferencik J, Slivinsky J, Bicarova S, Bilcik D, Blazenec M, Netherer S (2017). Storms, temperature maxima and the Eurasian spruce bark beetle *Ips typographus* - An infernal trio in Norway spruce forests of the Central European High Tatra Mountains. *Agricultural and Forest Meteorology* 242: 85-95. - doi: [10.1016/j.agrformet.2017.04.004](https://doi.org/10.1016/j.agrformet.2017.04.004)
- Ortiz SM, Breidenbach J, Kändler G (2013). Early detection of bark beetle green attack using TerraSAR-X and RapidEye data. *Remote Sensing* 5: 1912-1931. - doi: [10.3390/rs5041912](https://doi.org/10.3390/rs5041912)
- Økland B, Nikolov C, Krokene P, Vakula J (2016). Transition from windfall- to patch-driven outbreak dynamics of the spruce bark beetle *Ips typographus*. *Forest Ecology and Management* 363: 63-73. - doi: [10.1016/j.foreco.2015.12.007](https://doi.org/10.1016/j.foreco.2015.12.007)
- Pontius RG, Schneider LC (2001). Land-cover change model validation by an ROC method for the Ipswich watershed. *Agriculture, Ecosystems and Environment* 85: 239-248. - doi: [10.1016/S0167-8809\(01\)00187-6](https://doi.org/10.1016/S0167-8809(01)00187-6)
- Potterf M, Nikolov C, Kočícká E, Ferenčík J, Mezei P, Jakuš R (2019). Landscape-level spread of beetle infestations from windthrown- and beetle-killed trees in the non-intervention zone of the Tatra National Park, Slovakia (Central Europe). *Forest Ecology and Management* 432: 489-500. - doi: [10.1016/j.foreco.2018.09.050](https://doi.org/10.1016/j.foreco.2018.09.050)
- Phillips S, Dudík M (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31: 161-175. - doi: [10.1111/j.0906-7590.2008.5203.x](https://doi.org/10.1111/j.0906-7590.2008.5203.x)
- Raffa KF, Gregoire J, Lindgren BS, Gre J (2015). Natural history and ecology of bark beetles. In: "Bark Beetles: Biology and Ecology of Native and Invasive Species". Academic Press, Elsevier, pp. 1-40. - doi: [10.1016/B978-0-12-417156-5.00001-0](https://doi.org/10.1016/B978-0-12-417156-5.00001-0)
- Robertson C, Nelson T, Boots B (2007). Mountain pine beetle dispersal: the spatial-temporal interaction of infestations. *Forest Science* 53: 395-405. [online] URL: <http://academic.oup.com/forests/article/53/3/395/4604057>
- Rossi F, Breidenbach J, Puliti S, Astrup R, Talbot B (2019). Assessing harvested sites in a forested boreal mountain catchment through global forest watch. *Remote Sensing* 11: 543. - doi: [10.3390/rs11050543](https://doi.org/10.3390/rs11050543)
- Rouault G, Candau JN, Lieutier F, Nageleisen LM, Warzee N, Martin JC, Warzee N (2006). Effects of drought and heat on forest insect populations in relation to the 2003 drought in Western Europe. *Annals of Forest Science* 63: 613-624. - doi: [10.1051/forest:2006044](https://doi.org/10.1051/forest:2006044)
- Senf C, Pflugmacher D, Wulder MA, Hostert P (2015). Characterizing spectral-temporal patterns of defoliating and bark beetle disturbances using Landsat time series. *Remote Sensing of Environment* 170: 166-177. - doi: [10.1016/j.rse.2015.09.019](https://doi.org/10.1016/j.rse.2015.09.019)
- Simard M, Powell EN, Raffa KF, Turner MG (2012). What explains landscape patterns of tree mortality caused by bark beetle outbreaks in Greater Yellowstone? *Global Ecology and Biogeography* 21: 556-567. - doi: [10.1111/j.1466-8238.2011.00710.x](https://doi.org/10.1111/j.1466-8238.2011.00710.x)
- Stoklasa M (2003). Monitoring zdravotního stavu lesu z družicových snímků [Forest health monitoring with the use of satellite images]. *Ochrana přírody* 58: 228-232. [in Czech]
- Thatcher RC, Searcy JL, Coster JE, Hertel GD (1980). The southern pine beetle. Technical Bulletin 1631, Expanded Southern Pine Beetle Research and Application Program, USDA Forest Service, Science and Education Administration, Pineville, LA, USA, pp. 265.
- Wackernagel H (2003). *Multivariate geostatistics. An introduction with applications*. Springer-Verlag, Berlin, Heidelberg, Germany, pp.

257.

Wulder MA, Dymond CC, White JC, Leckie DG, Carroll AL (2006). Surveying mountain pine beetle damage of forests: a review of remote sensing opportunities. *Forest Ecology and Management* 221: 27-41. - doi: [10.1016/j.foreco.2005.09.021](https://doi.org/10.1016/j.foreco.2005.09.021)

Supplementary Material

Fig. S1 - The predictive model of the susceptibility of the study area to bark beetle attack in 2009.

Fig. S2 - The predictive model of the susceptibility of the study area to bark beetle attack in 2010.

Fig. S3 - The results of the first (combined) type of validation of predictive models: the ROC curves and the calculated AUC values of the predictive models from 2008 to 2010.

Link: [Duriaciova_3271@suppl001.pdf](#)