https://doi.org/10.17221/4/2022-SWR

Predictors for digital mapping of forest soil organic carbon stocks in different types of landscape

Luboš Borůvka[1]*, Radim Vašát[1], Vít Šrámek[2], Kateřina Neudertová Hellebrandová[2], Věra Fadrhonsová[2], Milan Sáňka[3], Lenka Pavlů[1], Ondřej Sáňka[3], Oldřich Vacek[1], Karel Němeček[1], Shahin Nozari[1], Vincent Yaw Oppong Sarkodie[1]

[1]Department of Soil Science and Soil Protection, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Prague, Czech Republic
[2]Forestry and Game Management Research Institute, Jíloviště-Strnady, Czech Republic
[3]RECETOX, Masaryk University, Brno, Czech Republic

*Corresponding author: boruvka@af.czu.cz

Citation: Borůvka L., Vašát R., Šrámek V., Neudertová Hellebrandová K., Fadrhonsová V., Sáňka M., Pavlů L., Sáňka O., Vacek O., Němeček K., Nozari S., Oppong Sarkodie V.Y. (2022): Predictors for digital mapping of forest soil organic carbon stocks in different types of landscape. Soil & Water Res., 17: 69–79.

Abstract: Forest soils have a high potential to store carbon and thus mitigate climate change. The information on spatial distribution of soil organic carbon (SOC) stocks is thus very important. This study aims to analyse the importance of environmental predictors for forest SOC stock prediction at the regional and national scale in the Czech Republic. A big database of forest soil data for more than 7 000 sites was compiled from several surveys. SOC stocks were calculated from SOC content and bulk density for the topsoil mineral layer 0–30 cm. Spatial prediction models were developed separately for individual natural forest areas and for four subsets with different altitude range, using random forest method. The importance of environmental predictors in the models strongly differs between regions and altitudes. At lower altitudes, forest edaphic series and soil classes are strong predictors, while at higher altitudes the predictors related to topography become more important. The importance of soil classes depends on the pedodiversity level and on the difference in SOC stock between the classes. The contribution of forest types as predictors is limited when one (mostly coniferous) type dominates. Better prediction results can be obtained in smaller, but consistent regions, like some natural forest areas.

Keywords: carbon stocks; digital soil mapping; environmental covariates; random forests; spatial distribution; terrain attributes

The total forest ecosystem carbon (C) stock is large and in dynamic equilibrium with its environment (Lal 2005). There is a high potential for C sequestration and forest soils can thus contribute significantly to climate change mitigation. The ratio of C storage between tree biomass and soil depends on climate. At colder climate, lower C amounts are incorporated in tree biomass, but the soil organic carbon (SOC) stocks in soil are increased due to slower decomposition (Wen & He 2016). As the built-up of organic matter is a long-term process, forest continuity is an important factor of the SOC stocks in forest soils (Nitsch et al. 2018), as well as forest age (Jonard et al. 2017). Recovery of SOC stocks after forest soil disturbance can take decades (Dobor et al. 2018).

Factors influencing SOC amount in forest soils include (Lal 2005): climatic factors, topography, soil characteristics, natural disturbance, and anthropogenic factors (forest management, afforestation, and deforestation). Chuman et al. (2021) concluded that

Supported by the Ministry of Agriculture of the Czech Republic, Project No. QK1920163.

https://doi.org/10.17221/4/2022-SWR

elevation (reflecting temperatures and precipitation levels) belongs to the most important factors controlling SOC pools in Podzols and Cambisols, together with legacy acid deposition of S and N compounds. The anthropogenic influence is particularly pronounced in forest floor and mineral topsoil.

The information on SOC stock spatial distribution and the influencing factors is important for the assessment of forest ecosystem functioning, soil ecosystem services, soil fertility, as well as a support for decision making in forest and environmental management. Digital soil mapping (DSM) provides a useful and efficient tool for the description and assessment of soil properties spatial distribution. General framework of DSM as the quantitative prediction of soil properties or classes using soil information and environmental covariates (scorpan model) was formalised by McBratney et al. (2003). Digital mapping of SOC contents or stocks is one of the most frequent applications of DSM (e.g. Lamichhane et al. 2019). Various prediction models are used, and various sets of covariates (predictors) are tested. Miller et al. (2015) tested a pool of 412 potential predictors and found that models with limited predictor pools can substitute other predictors to compensate for the missing variables.

Random forests (RF, Breiman 2001) is one of the most often used prediction methods in DSM (e.g. Calvo de Anta et al. 2020; Yamashita et al. 2022). However, Were et al. (2015) found that RF overestimated SOC stocks compared to models based on support vector regression and artificial neural networks. Martin et al. (2014) found that robust geostatistical modelling of residuals from tree-based models improved the prediction accuracy significantly when a limited number of predictors were included.

Many studies on digital mapping of SOC stocks focus on mineral topsoil 0–30 cm as there is usually the highest amount of SOC stored (e.g. Wiesmeier et al. 2012; Minasny et al. 2013; Yamashita et al. 2022). According to De Vos et al. (2015), the mineral layer 0–30 cm contains approximately 55–65% of the total SOC stock in forest soil profiles.

Various surveys of forest SOC content have been performed and various legacy data are available. However, different sampling designs, protocols and depth, different analytical methods, and data aging make the combination of data from different sources difficult and challenging (Borůvka et al. 2018; Bai & Fernandez 2020).

The aim of this study was to analyse the importance of environmental predictors for forest SOC stock prediction at the regional and national scale in the Czech Republic and to compare relative importance of the predictors in contrasting subsets of the national forest soil database compiled from several large-scale soil surveys.

MATERIAL AND METHODS

Study area and soil data. This study is done on the whole forested area of the Czech Republic, belonging to the temperate forest zone. The country has an elevation ranging from 115 to 1 602 m a.s.l. Mean annual temperatures are in the range from 1 to 10 °C, with mean annual precipitation ranging between 400 and 1 400 mm. Forests cover 26 551 km^2, forming 34.2% of the total country area. The Czech Republic is divided into 41 natural forest areas (NFA, http://www.uhul.cz/what-we-do/regional-plans-of-forest-development). These spatially compact areas are rather homogeneous territories defined on the basis of geological, climatic, orographic and phytogeographical conditions.

A database of forest soil data from the years 2000 to 2020 was compiled from several resources: (*i*) National Forest Inventory (NFI) done by the Forest Management Institute (FMI, Forest Management Institute 2007); (*ii*) Data from permanent typological areas collected also by the FMI; (*iii*) Forest Soil Monitoring (FSM) done by the Central Institute for Supervising and Testing in Agriculture (Fiala et al. 2013); (*iv*) Data originating from the international projects ICP Forest and BioSoil (Lorenz & Becher 2012; Šrámek et al. 2013). As the surveys used different methodology and different sampling depths or horizons, the data were recalculated to the topsoil mineral layer 0–30 cm using weighted average. SOC content was mostly determined by oxidimetric method; comparability of other methods used in the surveys was tested. SOC stocks were calculated from the SOC content and bulk density (BD). Where the BD was not available, an estimate of BD was calculated using the model by Honeysett and Ratkowsky (1989):

$$BD = 1/(0.564 + 0.0556 \times OM) \ (g/cm^3)$$

where:
OM (organic matter) = 1.724 × SOC (%).

Rock fragments were not taken into account as this information was not available on all sites and, moreover, the accuracy of rock fragment content is generally low. In total, SOC stock values at the

https://doi.org/10.17221/4/2022-SWR

0–30 cm depth were collected from 7 338 forest stands all over the country, though the spatial distribution is not even and there are some gaps (see Figure 1D).

Potential covariates. Terrain data were extracted from the digital elevation model (DEM) ArcČR®500 with resolution 200 m (ARCDATA PRAHA, ZÚ, ČSÚ, 2016; Figure 1A). Secondary terrain characteristics were calculated using Terrain Analysis Toolbox in SAGA GIS 2.1.4 (Conrad et al. 2015). The following terrain attributes were determined: elevation (m a.s.l.), slope, aspect (cos and sin), planar and profile curvatures, convergence index, catchment area, valley depth, relative slope position (RSP), channel network base level (CNBL), channel network distance (CND), topographic wetness index (TWI), LS factor (LSF), and analytical hillshade.

Soil classes were obtained from the Czech soil information system PUGIS at the resolution 1 : 250 000 (Kozák et al. 1996). The individual classes were grouped into 13 groups (see Table 3). While some soil classes were grouped to larger sets as they are less represented in forests (like Chernozems, Phaeozems and Vertisols), or have similar properties (like Luvisols and Retisols), the most abundant Cambisols forming in total more than 50% of the country were divided into 3 subclasses (mostly Eutric, Dystric and Arenic Cambisols; Figure 1C). Mean annual precipitation and temperatures were obtained from the database WorldClim.org at resolution 1 km (Fick & Hijmans 2017). Land cover/land use categories, particularly forest types (deciduous/mixed/coniferous) were obtained from the database CORINE Land Cover 2018 (EEA 2018) at resolution of 100 m (Figure 1B). Forest typology (Viewegh et al. 2003) information on stands (forest vegetation zones – FVZ, and edaphic series) were obtained from the map of forest typology at scale 1 : 10 000 (ÚHÚL 2019).

Model selection, calibration and validation. Several model types were tested for SOC stock prediction, namely artificial neural networks, boosted regression trees, random forests (RF), and multivariate adaptive regression splines. Based on the results,

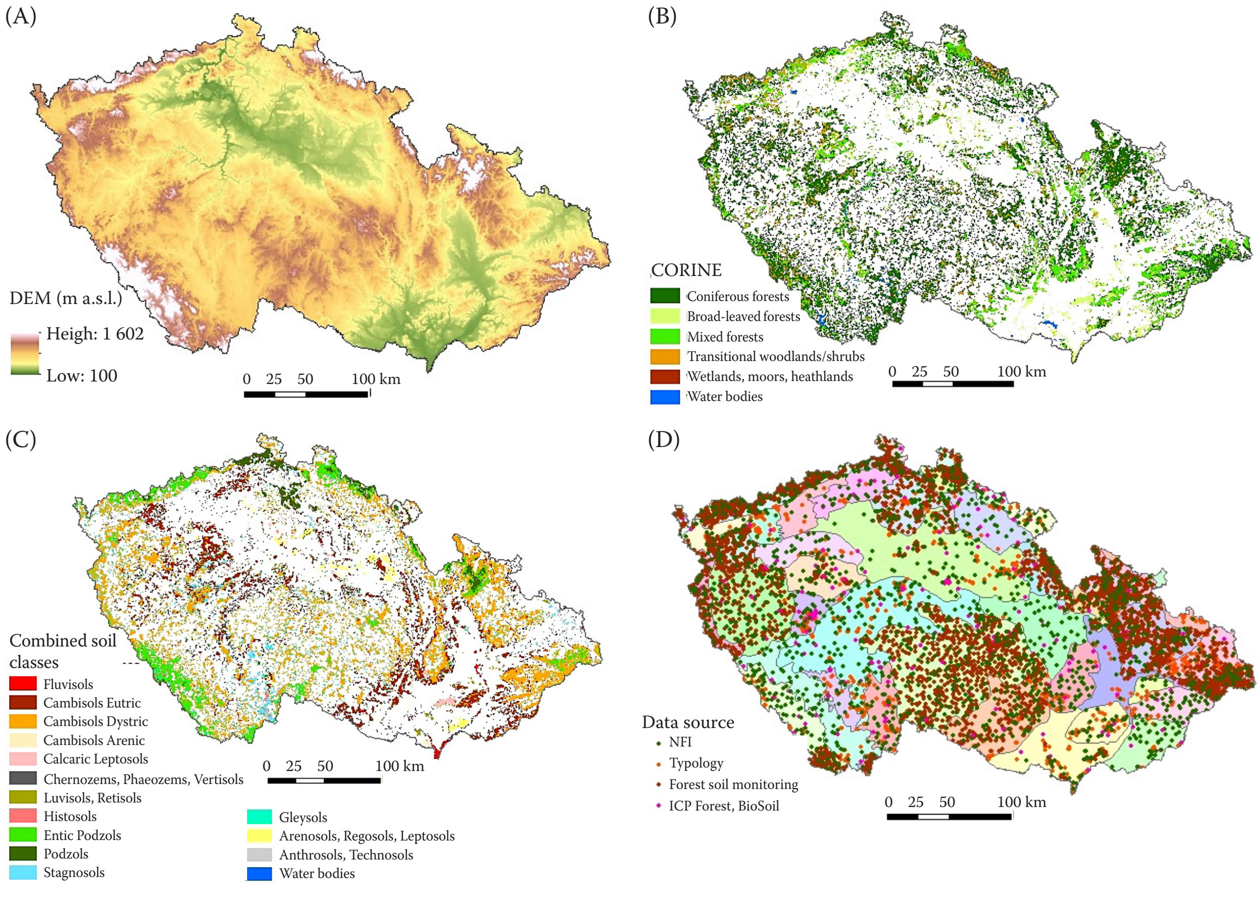

Figure 1. Map of the Czech Republic with digital elevation model (DEM) (A), forest types (B), combined soil classes (C) and sampling points in the natural forest areas (D)

https://doi.org/10.17221/4/2022-SWR

and taking into account its common utilization, robustness to model overfitting and intercorrelation of predictors, and its ability to quantify relative predictor importance, the method of random forests (Breiman 2001) was chosen. 70% of data were used for model calibration, 30% for model validation. Index of determination (R^2) and root mean square error (RMSE) of validation were used for model performance evaluation.

RESULTS AND DISCUSSION

General data description and predictor selection. The calculated SOC values in the depth 0–30 cm ranging from 0.07 to 38.59 kg/m^2, with a mean of 10.30 kg/m^2 (Table 1) correspond to values compiled by Lal (2005) for temperate forests, as well as those reported for Germany and other Central European countries (Wiesmeier et al. 2012), Slovakia (Priwitzer et al. 2009), Austria (Baumgarten et al. 2021), Spain (Calvo de Anta et al. 2020), or EU (De Vos et al. 2015). Prietzel and Christophel (2014) found slightly lower values in mineral topsoils in German Alps, which may be caused by higher elevations and consequently higher proportion of SOC in forest floor, and by the rock fragments that were not taken into account in our study. Lower SOC stock values were found also in Russian forests (Osipov et al. 2021) or in Hesse, Germany (Heitkamp et al. 2021).

Table 1. Basic statistical parameters of soil organic carbon (SOC) stock dataset (in kg/m^2, layer 0–30 cm)

Parameter	SOC stock
Count	7 338
Mean	10.30
Median	10.05
Geometric mean	9.06
Variance	22.04
SD	4.69
CV (%)	45.59
Standard error	0.05
Minimum	0.07
Maximum	38.59
Range	38.52
Lower quartile	6.53
Upper quartile	13.65
Skewness	0.33
Kurtosis	–0.33

SD – standard deviation; *CV* – coefficient of variation

Correlation analysis showed that SOC stocks are positively correlated with altitude (r = 0.438 ; Table 2), forest vegetation zones (0.413) and mean annual precipitation (0.347), and negatively correlated with average annual temperature (–0.425). An increase of SOC stocks with increasing altitudes was reported also by Bojko and Kabala (2017), but only to the altitude of 1 000 m a.s.l. Above this level, the SOC stocks started to drop again. Decreasing SOC stocks with increasing altitudes above 900 m a.s.l. were found also by Tungalag et al. (2020) in Mongolia. Weak correlation of the other predictors with SOC stock does not necessarily mean that there are no relationships; there can be some, but not linear.

The correlation analysis showed also mutual relationships between the predictors. Thanks to the large dataset, even weak relationships are significant. Though RF model is not too sensitive to interrelations of predictors, we removed from further model calibration the predictors strongly correlated with other predictors to avoid redundant information in the model input. Finally, only seven continuous auxiliary variables were retained: annual precipitation, analytical hillshade, LS factor, catchment area, profile curvature, convergence index, and channel network distance. Three categorical ones were added: combined soil classes, edaphic series indicating trophic conditions and thus indirectly reflecting soil and geological conditions, and forest type. These ten predictors were used in all further models and their relative importance was evaluated.

SOC stocks prediction in natural forest areas. Separate models for SOC stock prediction were developed for individual NFA if the number of sampling points was sufficient, or for groups of two or a few neighbouring NFA that were similar. The NFA can correspond to the soil-landscape systems described by Mulder et al. (2015) who concluded that these systems have homogeneous conditions with respect to the combination of SOC controlling factors. This may explain why the prediction in some of these NFA was more successful than the groups defined by altitude ranges as shown further, or than the whole national model; the highest R^2 was 0.564, the lowest RMSE 2.31 kg/m^2. However, prediction accuracy for some other NFA was rather poor (minimum R^2 0.001, highest RMSE 4.53 kg/m^2). Similar results were reported by Hounkpatin et al. (2021) after comparison of national model with local (regional) models. Though the prediction accuracy generally improves (R^2 increases and RMSE decreases) with increasing

https://doi.org/10.17221/4/2022-SWR

Table 2. Correlation matrix of soil organic carbon (SOC) stocks and covariates for the whole dataset (layer 0–30 cm); the predictors finally used for model development are in bold in the first column

Variable	SOC stock	FVZ	Altitude	Anal. hillshade	Slope	Aspect (sin)	Aspect (cos)	Planar curv.	Prof. curv.	Conv. index	Catch. area	TWI	LSF	CNBL	CND	Valley depth	RSP	Aver. temp.	Ann. prec.
SOC stock	**1.000**	**0.413**	**0.438**	**0.241**	**0.199**	−0.021	−0.004	−0.012	−0.007	**0.037**	−0.001	**−0.061**	**0.221**	**0.210**	**0.162**	**0.171**	**0.138**	**−0.425**	**0.347**
FVZ	**0.413**	**1.000**	**0.900**	**0.202**	**0.229**	**0.047**	−0.015	0.012	**0.026**	**0.054**	**−0.061**	**−0.235**	**0.215**	**0.659**	**−0.027**	**−0.035**	**0.413**	**−0.824**	**0.633**
Altitude	**0.438**	**0.900**	**1.000**	**0.186**	**0.295**	**0.017**	**−0.045**	**0.040**	**0.068**	**0.132**	**−0.066**	**−0.299**	**0.259**	**0.700**	**−0.088**	**−0.108**	**0.508**	**−0.889**	**0.718**
Analytical hillshade	**0.241**	**0.202**	**0.186**	**1.000**	**−0.026**	**0.302**	**−0.317**	**−0.067**	**−0.062**	**0.045**	0.022	**0.210**	**0.067**	**−0.052**	**0.288**	**0.329**	0.000	**−0.227**	**0.158**
Slope	**0.199**	**0.229**	**0.295**	**−0.026**	**1.000**	0.017	**0.027**	**0.070**	0.003	−0.009	**−0.047**	**−0.575**	**0.923**	**0.178**	**0.168**	**0.151**	**0.147**	**−0.353**	**0.310**
Aspect (sin)	−0.021	**0.047**	0.017	**0.302**	0.017	**1.000**	0.011	−0.012	−0.001	0.004	0.013	**−0.053**	0.011	0.000	−0.005	−0.013	**0.043**	**−0.053**	**0.060**
Aspect (cos)	−0.004	−0.015	**−0.045**	**−0.317**	**0.027**	0.011	**1.000**	0.017	−0.017	−0.007	**−0.038**	−0.012	0.011	**−0.045**	**−0.028**	**−0.029**	−0.013	**0.040**	−0.005
Planar curv.	−0.012	0.012	**0.040**	**−0.067**	**0.070**	−0.012	0.017	**1.000**	**0.521**	**0.339**	−0.010	**−0.280**	**−0.063**	−0.002	**−0.066**	**−0.138**	**0.218**	**−0.035**	0.022
Profile curv.	−0.007	**0.026**	**0.068**	**−0.062**	0.003	−0.001	−0.017	**0.521**	**1.000**	**0.178**	−0.012	**−0.190**	**−0.080**	0.017	**−0.064**	**−0.187**	**0.312**	**−0.063**	**0.042**
Conv. index	**0.037**	**0.054**	**0.132**	**0.045**	−0.009	0.004	−0.007	**0.339**	**0.178**	**1.000**	**−0.051**	**−0.332**	**−0.095**	−0.019	**−0.232**	**−0.244**	**0.349**	**−0.111**	**0.075**
Catch. area	−0.001	**−0.061**	**−0.066**	0.022	**−0.047**	0.013	**−0.038**	−0.010	−0.012	**−0.051**	**1.000**	**0.227**	0.013	**−0.056**	**0.033**	**0.034**	**−0.084**	**0.070**	**−0.042**
TWI	**−0.061**	**−0.235**	**−0.299**	**0.210**	**−0.575**	−0.053	−0.012	**−0.280**	**−0.190**	**−0.332**	**0.227**	**1.000**	**−0.357**	**−0.201**	**0.179**	**0.226**	**−0.470**	**0.308**	**−0.251**
LSF	**0.221**	**0.215**	**0.259**	**0.067**	**0.923**	0.011	0.011	**−0.063**	**−0.080**	**−0.095**	0.013	**−0.357**	**1.000**	**0.154**	**0.270**	**0.272**	**0.031**	**−0.323**	**0.282**
CNBL	**0.210**	**0.659**	**0.700**	**−0.052**	**0.178**	0.000	**−0.045**	−0.002	0.017	−0.019	**−0.056**	**−0.201**	**0.154**	**1.000**	**−0.025**	**−0.057**	**0.221**	**−0.753**	**0.682**
CND	**0.162**	**−0.027**	**−0.088**	**0.288**	**0.168**	−0.005	**−0.028**	**−0.066**	**−0.064**	**−0.232**	**0.033**	**0.179**	**0.270**	**−0.025**	**1.000**	**0.943**	**−0.478**	−0.017	**0.073**
Valley depth	**0.171**	**−0.035**	**−0.108**	**0.329**	**0.151**	−0.013	**−0.029**	**−0.138**	**−0.187**	**−0.244**	**0.034**	**0.226**	**0.272**	**−0.057**	**0.943**	**1.000**	**−0.601**	0.022	**0.036**
RSP	**0.138**	**0.413**	**0.508**	0.000	**0.147**	**0.043**	−0.013	**0.218**	**0.312**	**0.349**	**−0.084**	**−0.470**	**0.031**	**0.221**	**−0.478**	**−0.601**	**1.000**	**−0.489**	**0.352**
Aver. temp.	**−0.425**	**−0.824**	**−0.889**	**−0.227**	**−0.353**	**−0.053**	**0.040**	**−0.035**	**−0.063**	**−0.111**	**0.070**	**0.308**	**−0.323**	**−0.753**	−0.017	0.022	**−0.489**	**1.000**	**−0.744**
Annual prec.	**0.347**	**0.633**	**0.718**	**0.158**	**0.310**	**0.060**	−0.005	0.022	**0.042**	**0.075**	**−0.042**	**−0.251**	**0.282**	**0.682**	**0.073**	**0.036**	**0.352**	**−0.744**	1.000

FVZ – forest vegetation zones; TWI – topographic wetness index; LSF – LS factor; CNBL – channel network base level; CND – channel network distance; RSP – relative slope position; correlations at $P \leq 0.05$ are in bold

https://doi.org/10.17221/4/2022-SWR

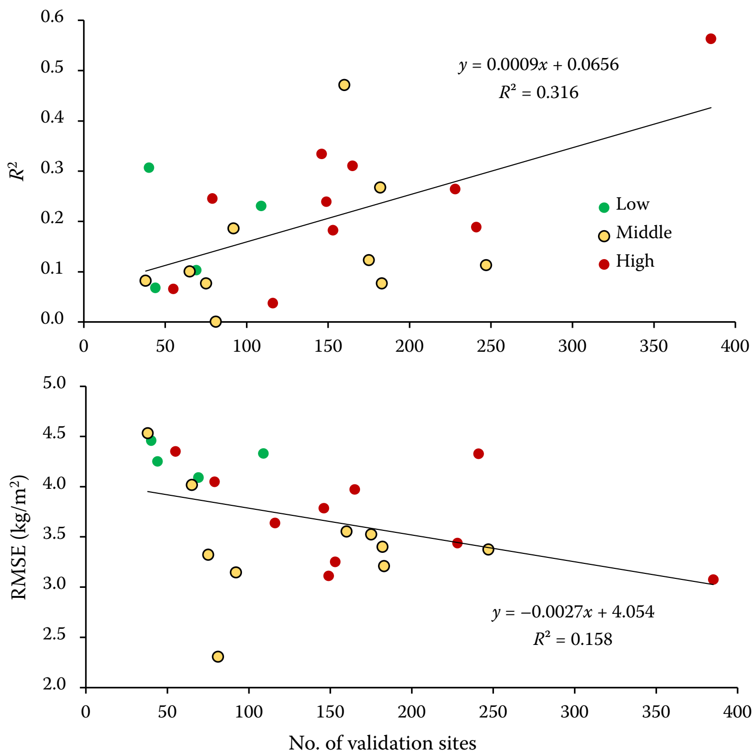

Figure 2. Relationship between model performance measures (R^2 and RMSE of validation) and the number of validation sites for regional models on individual NFA or groups of several neighbouring NFA

The sets are divided to groups according to the prevailing altitude range (lower/middle/higher); RMSE – root mean square error; NFA – natural forest areas

size of the dataset (Figure 2), there are large datasets with poor models, and, in contrast, small datasets with good prediction accuracy. Moreover, though there are different combinations of important predictors for lower NFA and higher NFA, there is not a consistent trend of a better model performance in any altitude group of NFA. To analyse the different combination of important predictors at different altitudes, and to avoid criticism for different size of the groups, we divided the whole national dataset to four equal groups according to altitudes.

SOC stocks prediction in different altitude ranges. Figure 3 shows that the relative importance of predictors differs between different altitudes. At the first group with the lowest altitudes, there is the strongest effect of edaphic series, followed by combined soil classes, catchment area, annual precipitation, and analytical hillshade. Edaphic series indicate the trophic state of the stands, which definitely has a strong effect on SOC accumulation. The effect of soil classes is important because there is a strong variation of soil types in group 1, as it is shown by higher level of pedodiversity (Vacek et al. 2020), and also there are significant differences in SOC stock between soil types as confirmed by analysis of variance (ANOVA, Table 3). The highest stocks are in Fluvisols, which corresponds to the general features of this class, and in Calcareous Leptosols (mainly Rendzinas), where soil organic matter is stabilized by carbonates. However, Ostrowska et al. (2010) stated that the SOC accumulation in the profile is to a greater extent affected by the site type and stand age than by the soil type. In contrast, rather low importance

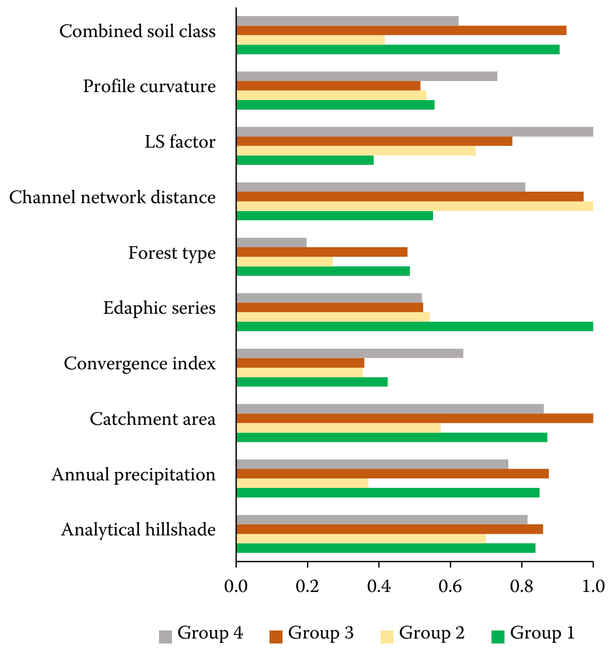

Figure 3. Relative importance of predictors for four altitude groups of equal size according to the altitude

https://doi.org/10.17221/4/2022-SWR

was found at low altitudes for most relief-related predictors. Tziachris et al. (2019) also reported that terrain-based covariates have the least importance in flatness area. Only analytical hillshade, which is a terrain parameter, but with a strong relationship to the extent of solar radiation reaching the stand, has similar importance in all the altitude groups, as the sunlight undoubtedly influences organic matter production, decomposition and accumulation.

In group 2, the biggest importance was achieved for channel network distance, followed by other terrain characteristics like LS factor, catchment area or profile curvature. The importance of forest type is rather low, as this group is dominated by coniferous forests and the difference between broadleaved and mixed forests is not significant. The low importance of soil classes is probably caused by the dominance of Cambisols in this group (reflected by the lowest pedodiversity), and even if there are three sets of Cambisols distinguished, they do not differ much in SOC stocks.

The higher altitudes, groups 3 and 4, have generally even more heterogenous relief, and therefore the importance of relief-related predictors is rather

Table 3. Basic characteristics of four altitude classes and mean soil organic carbon (SOC) stocks in the layer 0–30 cm (in kg/m^2) in separate soil class and forest type subsets; number of sampling points in each subset is given in parentheses

	Group 1	Group 2	Group 3	Group 4
Altitude range (m a.s.l.)	145–421	421–550	550–748	748–1479
Mean annual temperature range (°C)	2.7–9.5	2.6–9.5	2.9–8.9	1.3–8.5
Annual precipitation range (mm)	470–1157	494–1233	494–1175	519–1318
Mean C stock	8.48 (1 836)	8.54 (1 836)	10.44 (1 836)	13.73 (1 836)
Combined soil classes				
Chernozems, Phaeozems, Vertisols	9.16 (44)^{bc}	–	–	–
Fluvisols	11.77 (91)^{e}	–	–	–
Cambisols Eutric	8.40 (349)^{abc}	9.59 (300)^{d}	11.87 (284)^{cd}	13.80 (146)^{bc}
Cambisols Dystric	7.87 (557)^{a}	8.42 (1 163)^{bc}	9.78 (1 100)^{b}	12.93 (281)^{a}
Cambisols Arenic	8.33 (312)^{ac}	9.04 (85)^{cd}	10.19 (21)^{abc}	–
Calcaric Leptosols	11.27 (37)^{de}	12.87 (8)^{e}	–	–
Luvisols, Retisols	8.02 (272)^{ac}	7.39 (56)^{ab}	8.89 (1)^{abcde}	15.20 (2)^{abcd}
Histosols	5.66 (3)^{abc}	7.25 (9)^{abcd}	10.78 (13)^{abcd}	12.40 (149)^{a}
Entic Podzols	–	10.80 (11)^{cde}	12.48 (218)^{de}	13.84 (664)^{b}
Podzols	–	–	13.55 (26)^{e}	14.31 (575)^{c}
Stagnosols	9.15 (151)^{b}	7.31 (189)^{a}	9.03 (129)^{a}	18.26 (7)^{d}
Gleysols	9.01 (1)^{abcde}	9.35 (12)^{abcde}	9.88 (44)^{ab}	11.77 (12)^{ab}
Technosols	9.54 (19)^{abcd}	12.42 (3)^{cde}	–	–
F ratio	9.65	6.65	19.22	7.28
P	< 0.001	< 0.001	< 0.001	< 0.001
Shannon index of pedodiversity (relative)	1.846 (0.780)	1.192 (0.518)	1.275 (0.580)	1.485 (0.714)
Forest types				
Coniferous	7.61 (668)^{a}	8.10 (1 172)^{a}	10.04 (1 318)^{a}	13.69 (1 476)^{a}
Mixed	8.20 (303)^{b}	8.99 (269)^{b}	11.04 (264)^{b}	13.91 (230)^{a}
Deciduous	9.27 (859)^{c}	9.57 (391)^{b}	11.85 (254)^{c}	13.93 (130)^{a}
F ratio	29.96	20.61	23.54	0.44
P	< 0.001	< 0.001	< 0.001	0.645
Prediction results (validation subset)				
R^2	0.140	0.207	0.240	0.093
RMSE	3.96	3.62	3.71	3.76

Identical letters in each column indicate homogeneous groups according to ANOVA at $P \leq 0.05$; RMSE – root mean square error

https://doi.org/10.17221/4/2022-SWR

high. Similarly, Ellili et al. (2019) found that slope and elevation are the most important covariates for predicting SOC. Soil classes are very important predictor in group 3 as there are Cambisols and Podzols that differ in SOC stocks. Group 4 is dominated by Podzols and therefore the importance of soil classes as predictor is again smaller. The highest SOC stocks at higher altitudes are in Podzols which corresponds to results reported by Bojko and Kabala (2017), and in Stagnosols where water saturation reduces mineralization process. The importance of forest type is still rather small, as the forests are dominated by conifers, and moreover, in group 4 the SOC stock in mineral topsoil under broadleaved and mixed forests does not differ significantly from coniferous forests.

General discussion. The validation results of the models were mostly weak, with quite low R^2 values. Similarly, Yamashita et al. (2022) obtained R^2 value of 0.38 in spatial prediction of SOC stocks in forested areas of Japan. Even lower R^2 values were obtained by Ottoy et al. (2017), Hounkpatin et al. (2021), Nussbaum et al. (2014), Baltensweiler et al. (2021) and Hoffmann et al. (2014). The predictions overestimated low values and underestimated high values, creating thus much narrower range of values. Similar result was obtained for Swedish forest soils by Hounkpatin et al. (2021). Much better prediction accuracy was obtained by Li et al. (2021) when using remote sensing indices as additional predictors. Another potential source of auxiliary information for SOC prediction can be found in soil spectroscopy (Gholizadeh et al. 2021). Using some covariates in a more detailed resolution can possibly improve the prediction. However, more detailed environmental covariates do not need necessarily lead to more accurate soil maps (Samuel-Rosa et al. 2015). An important part of the uncertainty in the models could have been introduced by combination of data from different surveys using different sampling designs, methods and approaches, by recalculation of the data to unified depth, and by uncertainty in bulk density estimation. Potential sources of errors and uncertainties in the assessment of forest SOC stocks from sample to continental scale are clearly reviewed and summarized by Vanguelova et al. (2016).

The importance of soil classes depends on the heterogeneity of soil cover (described for example by Shannon's index of pedodiversity), and also on the significance of difference between soil classes in SOC stocks. Surprisingly, the SOC stocks in Histosols were among the lowest. However, there are just a few sites with Histosols particularly in the first three altitude groups, so that it cannot be considered significant, either. It indicates rather some inconsistencies or errors in the database, in spite of numerous checks applied.

SOC stock in the depth of 0–30 cm is lower under coniferous (mainly spruce) forests than under broadleaved and mixed forests (Table 3); at lower altitudes this difference is significant. However, as the coniferous forests have usually thicker O horizons and larger SOC amounts are retained in the surface organic horizons (Kjønaas et al. 2021), the total SOC stock in the whole profile is generally bigger under coniferous forests than under broadleaved ones (Bojko & Kabala 2017; Nitsch et al. 2018). Nevertheless, Cremer et al. (2016) reported higher SOC stocks under coniferous forests even in the mineral topsoil. The dominance of coniferous forests, particularly at higher altitudes, and very similar SOC stock values in all forest types make forest type a less important predictor. A more detailed description of forest species composition might improve the prediction. The effect of climate on building SOC stocks was shown e.g. by Rial et al. (2017), Černý et al. (2020), or Calvo de Anta et al. (2020).

CONCLUSION

The study showed that the importance of environmental predictors in the models for SOC stock prediction can strongly differ between regions and altitudes. At lower altitudes, edaphic series and soil classes are strong predictors, while at higher altitudes the predictors related to topography become more important. The importance of soil classes depends on the pedodiversity level and on the difference in SOC stock between the soil classes distinguished. The contribution of forest types as predictor is limited when one type dominates. Collection and selection of influential covariates is a very important part of digital mapping of soil properties. It was found that better prediction results can be obtained in smaller, but consistent regions, like in some natural forest areas; however, in some NFA the models failed. It was also shown than even very exhaustive datasets used for modelling do not ensure highly accurate prediction. Data harmonization, transformation, standardization and recalculation bring additional uncertainty and error that are projected in developed prediction models and model estimates. Nevertheless, in spite of the uncertainties of the models, the

https://doi.org/10.17221/4/2022-SWR

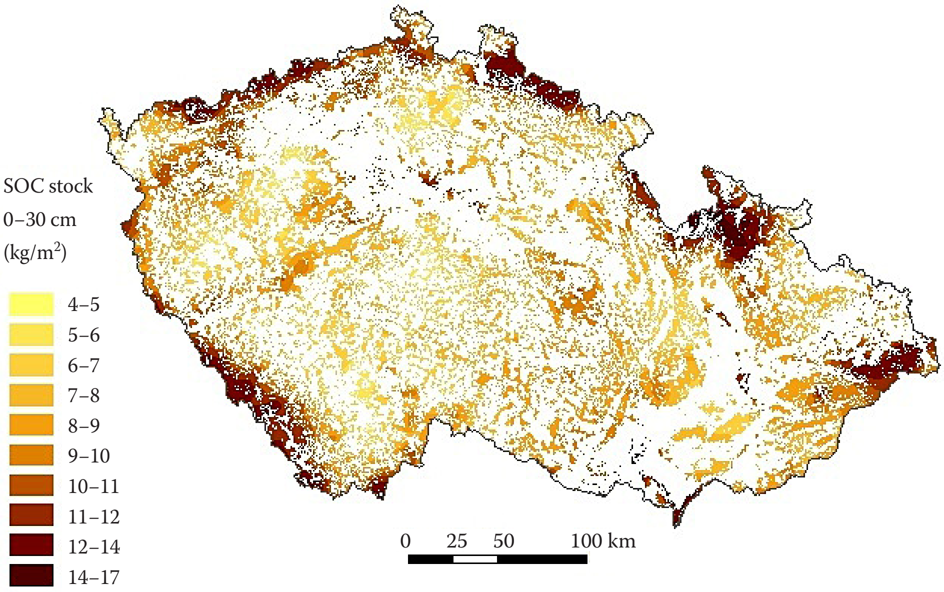

Figure 4. Predicted soil organic carbon (SOC) stock values for the mineral topsoil (0–30 cm) of forest soils using random forest model (R^2 = 0.32, RMSE = 3.91 kg/m^2)
The agricultural and other non-forest soils are masked by white colour; RMSE – root mean square error

prediction shows well the general trends and factors of SOC stock distribution, at least at the national scale (Figure 4).

Acknowledgement: We thank to the Forest Management Institute and the Central Institute for Supervising and Testing in Agriculture for providing the forest soil data.

REFERENCES

ARCDATA PRAHA, ZÚ, ČSÚ (2016): ArcČR 500 – Digital Geogrphic Database, Version 3.3. Available at www.arcdata.cz (accessed Oct 23, 2020). (in Czech)

Bai X., Fernandez I.J. (2020): Comparing publicly available databases to evaluate soil organic carbon in Maine, USA. Soil Science Society of America Journal, 84: 1722–1736.

Baltensweiler A., Walthert L., Hanewinkel M., Zimmermann S., Nussbaum M. (2021): Machine learning based soil maps for a wide range of soil properties for the forested area of Switzerland. Geoderma Regional, 27: e00437.

Baumgarten A., Haslmayr H.-P., Schwarz M., Huber S., Weiss P., Obersteiner E., Aust G., Englisch M., Horvath D., Leitgeb E., Foldal C., Rodlauer C., Bohner A., Spiegel H., Jandl R. (2021): Organic soil carbon in Austria – Status quo and foreseeable trends. Geoderma, 402: 115214.

Bojko O., Kabala C. (2017): Organic carbon pools in mountain soils – Sources of variability and predicted changes in relation to climate and land use changes. Catena, 149: 209–220.

Borůvka L., Penížek V., Kozák J., Marhoul A.M., Vacek O., Šrámek V., Žížala D., Poláková Š., Smatanová M., Fiala P. (2018): Combining and harmonizing soil data from different sources: Problems and approaches. In: Arrouays D., Savin I., Leenaars J., McBratney A.B. (eds.): GlobalSoilMap. Digital Soil Mapping from Country to Globe. Leiden, CRC Press/Balkema: 29–34.

Breiman L. (2001): Random forests. Machine Learning, 45: 5–32.

Calvo de Anta R., Luís E., Febrero-Bande M., Galiñanes J., Macías F., Ortíz R., Casás F. (2020): Soil organic carbon in peninsular Spain: Influence of environmental factors and spatial distribution. Geoderma, 370: 114365.

Černý J., Pokorný R., Vejpustková M., Šrámek V., Bednář P. (2020): Air temperature is the main driving factor of radiation use efficiency and carbon storage of mature Norway spruce stands under global climate change. International Journal of Biometeorology, 64: 1599–1611.

Conrad O., Bechtel B., Bock M., Dietrich H., Fischer E., Gerlitz L., Wehberg J., Wichmann V., Böhner J. (2015): System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geoscientific Model Development, 8: 1991–2007.

Cremer M., Kern N.V., Prietzel J. (2016): Soil organic carbon and nitrogen stocks under pure and mixed stands of European beech, Douglas fir and Norway spruce. Forest Ecology and Management, 367: 30–40.

https://doi.org/10.17221/4/2022-SWR

Chuman T., Oulehle F., Zajícová K., Hruška J. (2021): The legacy of acidic deposition controls soil organic carbon pools in temperate forests across the Czech Republic. European Journal of Soil Science, 72: 1780–1801.

De Vos B., Cools N., Ilvesniemi H., Vesterdal L., Vanguelova E., Carnicelli S. (2015): Benchmark values for forest soil carbon stocks in Europe: Results from a large scale forest soil survey. Geoderma, 251–252: 33–46.

Dobor L., Hlásny T., Rammer W., Barka I., Trombik J., Pavlenda P., Šebeň V., Štěpánek P., Seidl R. (2018): Post-disturbance recovery of forest carbon in a temperate forest landscape under climate change. Agricultural and Forest Meteorology, 263: 308–322.

EEA (2018): CORINE Land Cover (CLC) 2018, Version 20, European Environment Agency, Kopenhagen, Denmark. Available at https://www.eea.europa.eu/data-and-maps/ data/corine-land-cover-accounting-layers (accessed Oct 23, 2020).

Ellili Y., Walter C., Michot D., Pichelin P., Lemercier B. (2019): Mapping soil organic carbon stock change by soil monitoring and digital soil mapping at the landscape scale. Geoderma, 351: 1–8.

Fiala P., Reininger D., Samek T., Němec P., Sušil A. (2013): Forest Nutrition Survey at the Territory of the Czech Republic 1996–2011. Brno, Central Institute for Supervising and Testing in Agriculture. (in Czech)

Fick S.E., Hijmans R.J. (2017): WorldClim 2: New 1 km spatial resolution climate surfaces for global land areas. International Journal of Climatology, 37: 4302–4315.

Forest Management Institute (2007): National Forest Inventory in the Czech Republic 2001–2004. Introduction, Methods, Results. Brandýs and Labem, Forest Management Institute.

Gholizadeh A., Viscarra Rossel R.A., Saberioon M., Borůvka L., Kratina J., Pavlů L. (2021): National-scale spectroscopic assessment of soil organic carbon in forests of the Czech Republic. Geoderma, 385: 114832.

Heitkamp F., Ahrends B., Evers J., Meesenburg H. (2021): Spatial 3D mapping of forest soil carbon stocks in Hesse, Germany. Journal of Plant Nutrition and Soil Science, 184: 635–656.

Hoffmann U., Hoffmann T., Johnson E.A., Kuhn N.J. (2014): Assessment of variability and uncertainty of soil organic carbon in a mountainous boreal forest (Canadian Rocky Mountains, Alberta). Catena, 113: 107–121.

Honeysett J., Ratkowsky D. (1989): The use of ignition loss to estimate bulk density of forest soils. European Journal of Soil Science, 40: 299–308.

Hounkpatin K.O.L., Stendahl J., Lundblad M., Karltun E. (2021): Predicting the spatial distribution of soil organic carbon stock in Swedish forests using a group of covariates and site-specific data. Soil, 7: 377–398.

Jonard M., Nicolas M., Coomes D.A., Caignet I., Saenger A., Ponette Q. (2017): Forest soils in France are sequestering substantial amounts of carbon. Science of the Total Environment, 574: 616–628.

Kjønaas O.J., Bárcena T.G., Hylen G., Nordbakken J.-F., Økland T. (2021): Boreal tree species change as a climate mitigation strategy: Impact on ecosystem C and N stocks and soil nutrient levels. Ecosphere, 12: e03826.

Kozák J., Němeček J., Jetmar M. (1996): The database of soil information system – PUGIS. Rostlinná Výroba, 42: 529–534.

Lal R. (2005): Forest soils and carbon sequestration. Forest Ecology and Management, 220: 242–258.

Lamichhane S., Kumar L., Wilson B. (2019): Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. Geoderma, 352: 395–413.

Li Y., Zhao Z., Wei S., Sun D., Yang Q., Ding X. (2021): Prediction of regional forest soil nutrients based on Gaofen-1 remote sensing data. Forests, 12: 1430.

Lorenz M., Becher G. (eds.) (2012): Forest Condition in Europe, 2012. Technical Report of ICP Forests. Work Report of the Thünen Institute for World Forestry 2012/1. Hamburg, ICP Forests.

Martin M.P., Orton T.G., Lacarce E., Meersmans J., Saby N.P.A., Paroissien J.B., Jolivet C., Boulonne L., Arrouays D. (2014): Evaluation of modelling approaches for predicting the spatial distribution of soil organic carbon stocks at the national scale. Geoderma, 223–225: 97–107.

McBratney A.B., Mendonca Santos M.L., Minasny B. (2003): On digital soil mapping. Geoderma, 117: 3–52.

Miller B.A., Koszinski S., Wehrhan M., Sommer M. (2015): Impact of multi-scale predictor selection for modeling soil properties. Geoderma, 239–240: 97–106.

Minasny B., McBratney A.B., Malone B.P., Wheeler I. (2013): Digital mapping of soil carbon. Advances of Agronomy, 118: 1–47.

Mulder V.L., Lacoste M., Martin M.P., Richer-de-Forges A., Arrouays D. (2015): Understanding large-extent controls of soil organic carbon storage in relation to soil depth and soil-landscape systems. Global Biogeochemical Cycles, 29: 1210–1229.

Nitsch P., Kaupenjohann M., Wulf M. (2018): Forest continuity, soil depth and tree species are important parameters for SOC stocks in an old forest (Templiner Buchheide, northeast Germany). Geoderma, 310: 65–76.

Nussbaum M., Papritz A., Baltensweiler A., Waltheart L. (2014): Estimating soil organic carbon stocks of Swiss forest soils by robust external-drift kriging. Geoscientific Model Development, 7: 1197–1210.

https://doi.org/10.17221/4/2022-SWR

Osipov A.F., Bobkova K.S., Dymov A.A. (2021): Carbon stocks of soils under forest in the Komi Republic of Russia. Geoderma Regional, 27: e00427.

Ostrowska A., Porebska G., Kanafa M. (2010): Carbon accumulation and distribution in profiles of forest soils. Polish Journal of Environmental Studies, 19: 1307–1315.

Ottoy S., De Vos B., Sindayihebura A., Hermy M., Van Orshoven J. (2017): Assessing soil organic carbon stocks under current and potential forest cover using digital soil mapping and spatial generalisation. Ecological Indicators, 77: 139–150.

Prietzel J., Christophel D. (2014): Organic carbon stocks in forest soils of the German Alps. Geoderma, 221–222: 28–39.

Priwitzer T., Pajtík J., Konôpka B., Ištoňa J., Pavlenda P. (2009): Carbon allocation in forest ecosystem – Case studies from young pine stands. Lesnícky Časopis – Forestry Journal, 55: 239–250.

Rial M., Martínez Cortizas M., Rodríguz-Lado L. (2017): Understanding the spatial distribution of factors controlling topsoil organic carbon content in European soils. Science of the Total Environment, 609: 1411–1422.

Samuel-Rosa A., Heuvelink G.B.M., Vesques G.M., Anjos L.H.C. (2015): Do more detailed environmental covariates deliver more accurate soil maps? Geoderma, 243–244: 214–227.

Šrámek V., Jurkovská L., Fadrhonsová V., Hellebrandová-Neudertová K. (2013): Forest soil chemistry in relation to the forest site classification categories used in the Czech Republic – Results of the EU "BioSoil" forest soil monitoring project. Zprávy Lesnického Výzkumu, 58: 314–323.

Tungalag M., Gerelbaatar S., Lobanov A.I. (2020): Organic carbon stocks in the forest soils of Northern Mongolia. Lesnoy Zhurnal, 2020: 169–176.

Tziachris P., Aschonitis V., Chatzistathis T. (2019): Assessment of spatial hybrid methods for predicting soil organic matter using DEM derivatives and soil parameters. Catena, 174: 206–216.

ÚHÚL (2019): Typological map 2019. Ústav pro hospodářskou úpravu lesů. Available at http://geoportal.uhul.cz/ (in Czech)

Vacek O., Vašát R., Borůvka L. (2020): Quantifying the pedodiversity-elevation relations. Geoderma, 373: 114441.

Vanguelova E.I., Bonifacio E., De Vos B., Hoosbeek M.R., Berger T.W., Vesterdal L., Armolaitis K., Celi L., Dinca L., Kjønaas O.J., Pavlenda P., Pumpanen J., Püttsepp Ü., Reidy B., Simončič P., Tobin B., Zhiyanski M. (2016): Sources of errors and uncertainties in the assessment of forest soil carbon stocks at different scales – Review and recommendations. Environmental Monitoring and Assessment, 188: 630.

Viewegh J., Kusbach A., Mikeska M. (2003): Czech forest ecosystem classification. Journal of Forest Science, 49: 85–93.

Wen D., He N. (2016): Forest carbon storage along the north-south transect of eastern China: Spatial patterns, allocation, and influencing factors. Ecological Indicators, 61: 960–967.

Were K., Bui D.T., Dick Ø.B., Singh B.R. (2015): A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. Ecological Indicators, 52: 394–403.

Wiesmeier M., Spörlein P., Geuß U., Hangen E., Haug S., Reischl A., Schilling B., von Lützow M., Kögel-Knabner I. (2012): Soil organic carbon stocks in southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. Global Change Biology, 18: 2233–2245.

Yamashita N., Ishizuka S., Hashimoto S., Ugawa S., Nanko K., Osone Y., Iwahashi J., Sakai Y., Inatomi M., Kawanishi A., Morisada K., Tanaka N., Aizawa S., Imaya A., Takahashi M., Kaneko S., Miura S., Hirai K. (2022): National-scale 3D mapping of soil organic carbon in a Japanese forest considering microtopography and tephra deposition. Geoderma, 406: 115534.

Received: January 7, 2022
Accepted: January 24, 2022
Published online: February 9, 2022