BALTIC FORESTRY

http://www.balticforestry.mi.lt
ISSN 1392-1355
eISSN 2029-9230

Baltic Forestry 2020 26(2): 244–249
Category: research article
https://doi.org/10.46490/BF467

Integration of a statistical forest reflectance model and Sentinel-2 MSI images into a continuous forest inventory system

ANDRES KUUSK [1]* AND MAIT LANG [1,2]

[1] *University of Tartu, Tartu Observatory, 61602 Tõravere, Estonia*
[2] *Estonian University of Life Sciences, Kreutzwaldi 1, 51014 Tartu, Estonia*

* *Corresponding author: andres@to.ee*

Kuusk, A. and Lang, M. 2020. Integration of a statistical forest reflectance model and Sentinel-2 MSI images into a continuous forest inventory system. *Baltic Forestry* 26(2): 244–249. https://doi.org/10.46490/BF467.

Received 20 February 2020 **Revised** 14 July 2020 **Accepted** 31 July 2020

Abstract

Spectral signatures of forest stands in Sentinel-2 MSI spectral bands are simulated with the statistical forest reflectance (SFRM) model and compared to the spectral signatures measured in spectral images at ten study sites in Estonia. As an overall measure of the agreement between simulated and measured spectral signatures we used the total error calculated as the sum of relative errors over spectral bands B2 to B11 of Sentinel-2. The distribution of the total error has strongly positive skewness at all study sites and all types of forests (broadleaf, pine and spruce forests). The right tail of the distribution is low. The stands of high value of the total error far right in the tail of the distribution may have some errors in their inventory data, or the inventory data are outdated. Pertinent stands should have priority in their *in situ* checking process. The SFRM model is a simple and reliable tool for the validity checking of forest inventory data, using routinely collected forest inventory data and operational satellite information of moderate spatial resolution. The model is simple and computationally efficient. Preparing input data for the model is a simple query in the forest inventory database. The suggested procedure can be incorporated into the automated systems of continuous forest inventory.

Keywords: Forest inventory, Sentinel-2 MSI images, Statistical forest reflectance model

Introduction

Forest monitoring provides information to understand, protect and manage forests. The main components of respective information are the extent of forest resources, the availability of wood and forest biomass, information about the carbon cycle, forest biodiversity, and forest health and vitality, and their response to air pollution and climate change. In Estonia, half of the land area is covered by forests. According to the National Forest Inventory (NFI), the area of forest land in Estonia is 23,308 km^2 (Raudsaar et al. 2019). Forest industry is a large component of Estonian economy and forests play an important role in Estonian environment and ecology. Collecting and updating of information for this large number of forests are labour intensive and costly. It is impossible to carry out field visits and *in situ* monitoring of all forest stands sufficiently frequently for observing changes and possible disturbances in forest growth and stand structure. If we assume only a small influence of disturbances or the forest owners have obligation to submit notices to confirm cuttings and afforestation, then the status of forests could be simulated with forest growth models (Kull and Kull 1989, Kiviste 1997) and updated in central inventory database. Usually *in situ* measurements of forest stands are carried out and respective corrections in forest inventory database are made in case of forest management events (thinning, clearcutting) are planned, or forest damages (fire, flooding etc.) have occurred. Supporting information for forest management inventories and cyclic database update is collected from aerial orthophotos, provided by the Estonian Land Board at two years interval (Estonian Land Board, 2019). However, the interpretation of orthophotos is done only visually with the aim to delineate forest stands.

Prediction of forest inventory variables based on feature variables obtained from remote sensing and sparse sampling network of National Forest Inventories was proposed in early 1990s (Poso et al. 1990, Fazakas et al. 1999). While a subsample of NFI plots is measured in each year it is possible to construct maps of forest inventory variables with the same interval for the whole country. The methods allow estimation of forest statistics at municipality level, but at the stand level the errors are large. Small area estimates can be done if the number of sample plots per unit area is increased. For example, Lang et al. (2014) used Landsat-8 OLI images and data from 444 sample plots for a 15 × 15 km forested area in Laeva forest district, Estonia, and predicted wood volume and tree species composition for forest stands.

There have been attempts to use satellite information for estimating forest inventory parameters at the stand level. Jakubauskas and Price (1997) tried to estimate stand structure parameters of lodgepole pine forests as the multiple regression of Landsat-5 TM reflectance values in optical bands. While some stand parameters were predictable from remotely sensed data, factors relating specifically to understory condition were poorly predicted by spectral data, even with the inclusion of data transformations or indices. Data transformations (e.g. NDVI, Tasseled Cap) provided some measure of data reduction, but did not substantially increase the strength of the statistical relationship between spectral and biotic variables.

Kuusk et al. (2019) suggested a method how to support forest inventory by satellite information. A statistical forest reflectance model SFRM allows simulating reflectance spectrum of a forest stand in the wavelength range 400–1700 nm, having the regular forest inventory data from the forest management database as its input. The comparison of the stand reflectance simulated with the SFRM model to the stand reflectance in spectral images of Sentinel-2 MSI reveals which stands in the forest management database have erroneous or outdated data. Pertinent stands should have priority in their *in situ* checking process. The method was developed using data of airborne spectral measurements of forest reflectance in South-Eastern Estonia and regular forest management inventory data. In the present work the method is applied on several regions over the whole Estonia, thus covering regions of rather different climatic and forest growth conditions from an island in the Baltic Sea to inland forests in Eastern Estonia. The result of this study is the suggestion to the Estonian State Forest Management Centre and other forest owners which forest stands should have priority in their *in situ* checking process.

Materials and methods

Forest management inventory database which follows the concepts by Ferretti and Fisher (2013) and Burkhart and Tomé (2012) is available for the state forests in Estonia. Some private forests are covered by the database as well. The stand-wise inventory is following the Estonian forest inventory regulations (Government of Estonia 2015). Forest stands are delineated using aerial photos, data from previous inventory, and field visits, to construct a 1:10,000 map and update database records. A forest stand is a patch of forest homogeneous in species composition, age, tree height, tree density, and site type. The main forest inventory variables measured in the field for forest stand elements (a combination of tree species, dominance, and age) are height, stand basal area, and diameter of stems at breast height. Tree species composition in each social layer is calculated according to the wood volume of trees. Stand relative density is the ratio of stand basal area to the standard value according to forest height. The primary forest parameters in the database, which are the input parameters of the SFRM model, are listed in Table 1.

Table 1. The primary forest parameters

Parameter	Range
Stand age	6–286 years
Percent of main species	30–100%
Stand height	1–40 m
Relative density of the upper layer	0–328%
Relative density of the lower layer	0–161%
Basal area of the upper layer	0–80 m²/ha
Basal area of the lower layer	0–23 m²/ha
Stem volume of the upper layer	0–960 m³/ha
Stem volume of the lower layer	0–176 m³/ha
Percent of birch	1–100%
Percent of aspen	1–100%
Percent of common alder	1–100%
Percent of gray alder	1–100%
Percent of pine	1–100%
Percent of spruce	1–100%

Although some additional tree species are listed in the forest inventory database, their share in the forests under study is negligible.

Under the forest inventory rules, the minimum area of a stand is 0.1 ha. Growth conditions at the study sites range from poor, where the site index H_{100} is less than 10 m, to very good, where H_{100} can be over 35 m. The site index H_{100} is the stand height at the stand age of 100 years.

The database is updated continuously regarding state forests, but updates for private forests are only mandatory when the owner is planning harvest operations. In this work the forest inventory database of October 2018 is used. In the study are considered stands for which database contains sufficient information to run SRFM and which are large enough regarding spatial resolution of Sentinel-2 images. The selected stands were older than 5 years with the relative density, basal area, and the stem volume of the upper layer greater than zero. The second condition was that at least 10 Sentinel-2 pixels must be located inside the stand perimeter that is buffered inward for 8 meters.

The spectral signatures of the forest stands were collected from Sentinel-2 MSI spectral images acquired in midsummer of 2019. Satellite images are provided by the satellite data centre of the Estonian Land Board (ESTHub 2019). The study sites, dates of satellite images, and atmospheric conditions during acquisition are listed in Table 2. Figure 1 shows the study area. The study sites are presented in Figure 1 with Sentinel-2 MSI spectral imag-

Table 2. Study sites and Sentinel-2 images

Site	Center coordinates	Image date	SZA degree	AOD	H2O, cm	Method	N
Aegviidu	59°17'N 25°33'E	02.09.2019	51.4	0.07	1.5	Sen2Cor	11272
Alutaguse	59°06'N 27°22'E	28.08.2019	49.5	0.17	2.2	LUT	8864
Häädemeeste	58°05'N 24°41'E	28.08.2019	48.9	0.15	2.2	LUT	6461
Järvselja	58°19'N 27°16'E	18.08.2019	45.3	0.07	1.9	SkySpec	3027
Käru	58°45'N 25°07'E	28.08.2019	49.8	0.15	2.2	LUT	9168
Laeva	58°32'N 26°26'E	18.08.2019	45.3	0.07	1.9	SkySpec	5374
Nõva	59°01'N 23°52'E	29.07.2019	41.0	0.06	1.4	LUT	7835
Saaremaa	58°21'N 22°10'E	25.07.2019	39.0	0.14	2.8	LUT	8919
Tõstamaa	58°27'N 23°51'E	28.08.2019	48.9	0.10	1.8	LUT	9662
Võru	57°49'N 26°49'E	29.07.2019	40.0	0.06	1.4	LUT	7524

SZA stands for sun zenith angle, AOD stands for aerosol optical density at 550 nm, H2O stands for column water content, N stands for number of forest stands, LUT stands for atmospheric correction using look-up-table (Kuusk, 1998), SkySpec stands for optical parameters of the atmosphere estimated from SkySpec data at Järvselja (Kuusk and Kuusk, 2018).

es B7 (NIR, 780 nm), the background is the false colour with near infrared orthophoto by the Estonian Land Board (2019).

Simulated and Sentinel-2 MSI spectral reflectance of forest stands were compared to each other in 11 spectral bands of Sentinel-2 MSI: B2–B8, B8A, and B11. As in Kuusk et al. (2019), the mean relative difference,

$$d\rho(\lambda) = \frac{\rho_s(\lambda) - \langle\rho_m(\lambda)\rangle}{\langle\rho_m(\lambda)\rangle}, \quad (1)$$

is analyzed for detecting inconsistencies in forest inventory records. Here, $\rho_s(\lambda)$ is the simulated spectral reflectance of a stand in the spectral bands of Sentinel-2 MSI, and $\langle\rho_m(\lambda)\rangle$ is the average spectral reflectance of this stand measured from the Sentinel-2 MSI spectral image.

For the conversion of satellite-level spectral radiances to the top-of-canopy reflectance factor three different procedures of atmospheric correction were applied to satellite images. For the Järvselja and Laeva sites (the Sentinel-2 MSI acquisition T35VME of 18.08.2019) the data of SkySpec spectroradiometer that measures solar irradiance spectra at Järvselja were available and the optical parameters of the atmosphere were estimated using the procedure by Kuusk and Kuusk (2018), and the method of look-up-table (LUT) as described by Kuusk (1998) was applied. At the Aegviidu site the Sentinel-2 tool Sen2Cor was used (ESA 2019). At other sites the Sen2Cor procedure overestimated aerosol optical density (AOD), and thus, Sentinel-2 spectral images B2 (blue), B3 (green) and B4 (red) were overcorrected. That resulted in systematically underestimated stand reflectance in visible bands B2–B4, and consequently systematically high values in the relative error, Eq. (1). Therefore, the estimated AOD was adjusted so that the simulated with SFRM and corrected with LUT method spectral reflectance of broadleaf stands in blue and red bands coincided in the average.

Results

The simulated and Sentinel-2 MSI spectral reflectances of 78,100 forest stands at ten study sites were compared to each other. The distributions of the relative error

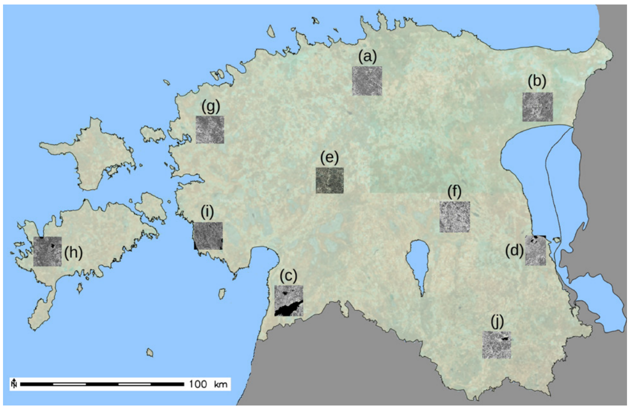

Figure 1. Map of the study sites
(a) – Aegviidu, (b) – Alutaguse, (c) – Häädemeeste, (d) – Järvselja, (e) – Käru, (f) – Laeva, (g) – Nõva, (h) – Saaremaa, (i) – Tõstamaa, (j) – Võru. The study sites are presented with the Sentinel-2 MSI images B7 (NIR). The background is the forestry orthophoto by the Estonian Land Board (2019).

$d\rho(\lambda)$ are plotted separately for every study site in Figures 2 and 3.

The main part of all distributions is between –0.5 and +0.5. Most of distributions are unimodal, and the mean value is close to zero. As in Kuusk et al. (2019), systematic differences vary from band to band. The dark bands in red and blue spectral region, where forests absorb incident radiation for the photosynthesis, are the most sensitive to errors in atmospheric correction of Sentinel images. The undercorrection of satellite images in the shortwave bands results in the negative relative error $d\rho(\lambda)$. An overprediction of the reflectance of broadleaf stands in NIR bands by SFRM as in Kuusk et al. (2019) results in the positive relative error $d\rho(\lambda)$.

In case of pine stands, different systematic errors in different spectral bands lead to the bimodal distribution of the relative error at some study sites when combined into a single error indicator variable. As commented in Kuusk et al. (2019), using spectral signatures at the Järvselja site for developing the regression model, the selection of pine stands was less representative compared to that of broadleaf and spruce stands. Obviously, there are some system-

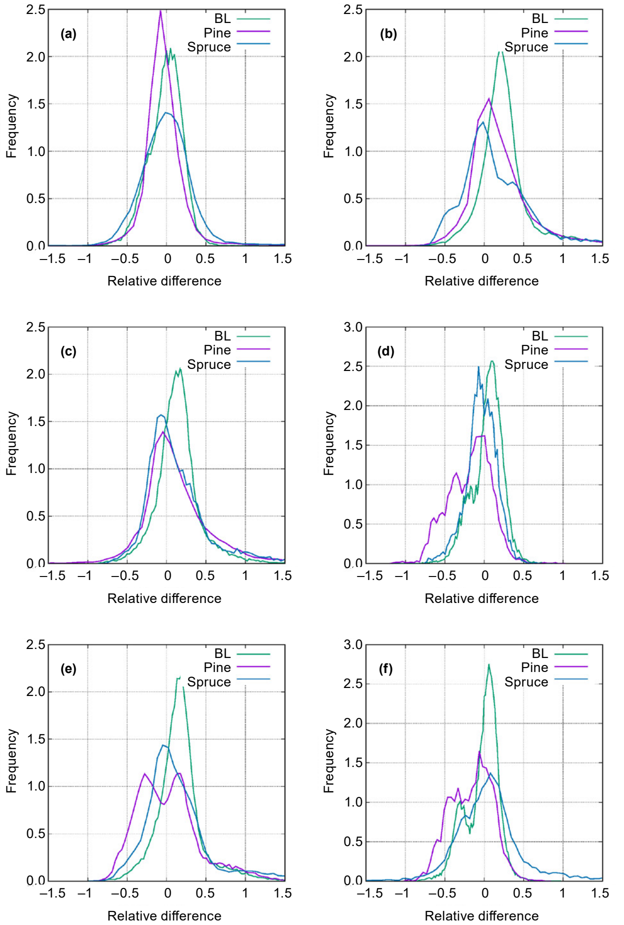

Figure 2. Distribution of the relative error Eq. 1 at study sites (a) – Aegviidu, (b) – Alutaguse, (c) – Häädemeeste, (d) – Järvselja, (e) – Käru, (f) – Laeva; BL – broadleaf stands.

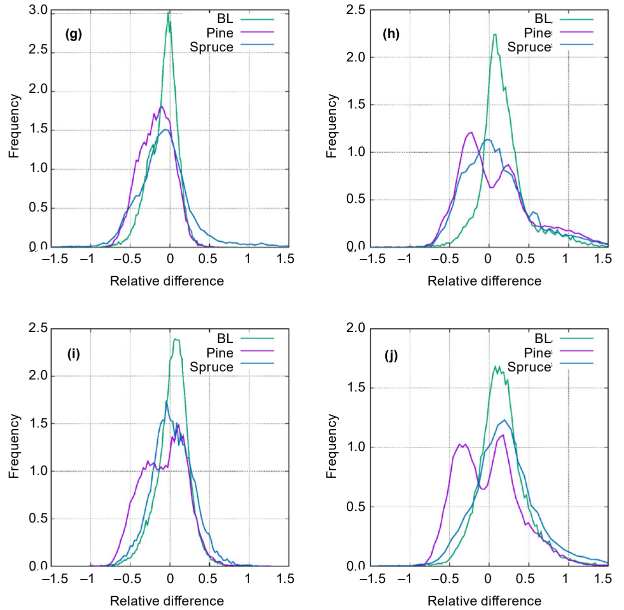

Figure 3. Distribution of the relative error Eq. 1 at study sites (Figure 2 continued)
(g) – Nõva, (h) – Saaremaa, (i) – Tõstamaa, (j) – Võru; BL – broadleaf stands.

atic differences between structure, age distribution, and growth conditions of pine stands at different study sites.

As an overall measure of the agreement between simulated and measured spectral reflectance can be used the total error S as defined by Kuusk et al. (2019)

$$S = \sum_{j=B_2}^{B1_1} \frac{\left|\rho_s(\lambda_j) - \langle \rho_m(\lambda_j) \rangle\right|}{\langle \rho_m(\lambda_j) \rangle}. \qquad (2)$$

The distribution of the total error S is plotted in Figures 4 and 5 separately for every study site. All these distributions have strongly positive skewness, and the right tail is low. The stands of high value of the total error S far right in the tail of the distribution may have some errors in their inventory data, or the inventory data is outdated.

A test case

For the validation of the method we analyzed why there is a large total error S at the Käru site, Figure 4(f), where the maximum value of total error was $S = 32.4$. The range of primary forest parameters at the Käru site is reported in Table 3.

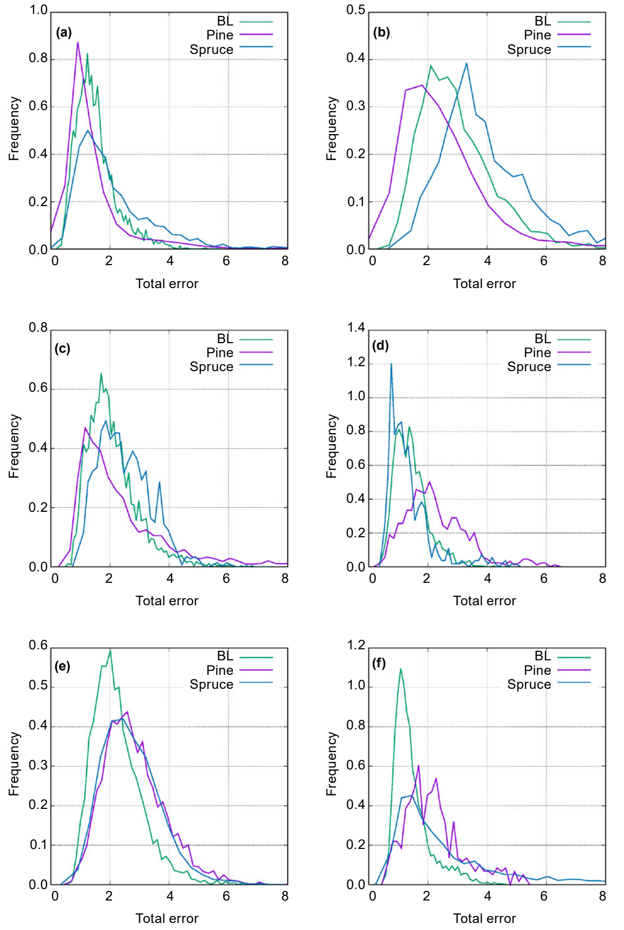

Figure 4. Distribution of the total error Eq. 2 at study sites (a) – Aegviidu, (b) – Alutaguse, (c) – Häädemeeste, (d) – Järvselja, (e) – Käru, (f) – Laeva; BL – broadleaf stands.

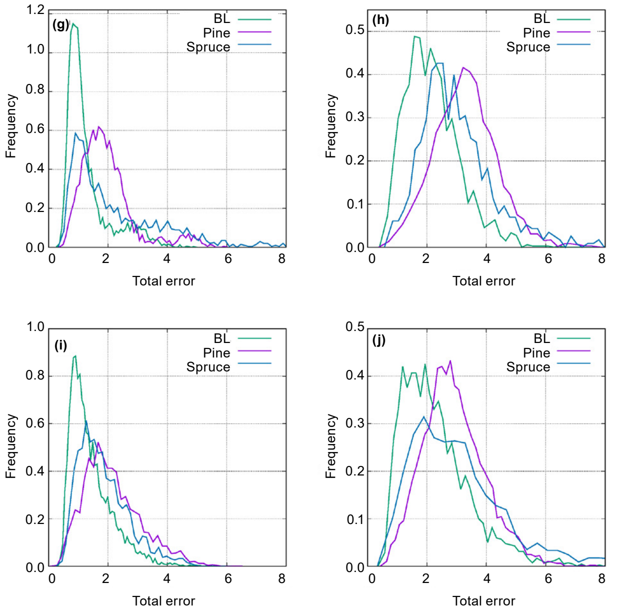

Figure 5. Distribution of the total error Eq. 2 at study sites (Figure 4 continued)
(g) – Nõva, (h) – Saaremaa, (i) – Tõstamaa, (j) – Võru; BL – broadleaf stands.

Table 3. The range of primary forest parameters at the Käru site

Parameter	Range
Stand age	6–217 years
Percent of main species	30–100%
Stand height	1–36 m
Relative density of the upper layer	2–328%
Relative density of the lower layer	0–86%
Basal area of the upper layer	0.2–47 m²/ha
Basal area of the lower layer	0–16 m²/ha
Stem volume of the upper layer	0–587 m³/ha
Stem volume of the lower layer	0–158 m³/ha

The stands with extreme values of forest parameters (see Table 3) do not have very high value of the total error *S*. A subset of 200 stands with the highest value of the total error *S* was selected for the detailed analysis. The distribution of the age of inventory data in the forestry database from year 2018 for these 200 stands is plotted in Figure 6. These 2.2% of stands from 9168 stands at the Käru site have the total error *S* values in the range from 4.93 to 32.4. In the sample, 8 stands have inventory data older than 15 years, however these stands have no exceptionally high value of the total error *S*.

There are 4 small Cumulus clouds in the Sentinel scene of 28.08.2019, and 14 stands were entirely or partly in the shadow of clouds. The value of total error *S* for these stands exceeded 7.0, and the highest value of total error occurred in the stands which are entirely in shade. We excluded these stands from the further analysis. In the most recent forest registry database from January 2020 the inventory parameters of 46 stands from the 186 stands of

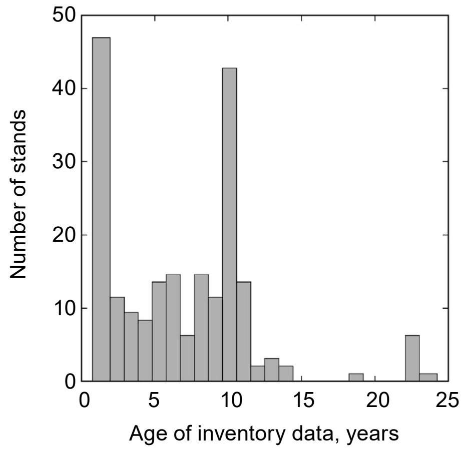

Figure 6. Age distribution of inventory data in 2018

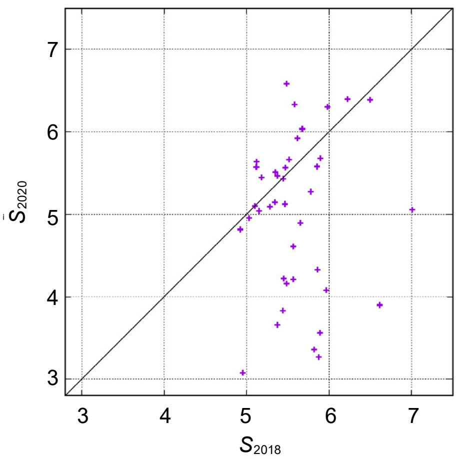

Figure 7. Comparison of the total error of simulated spectral signatures using outdated (2018) and updated (2020) forest inventory data

the highest values of the total error *S* while using inventory data of 2018 have been updated. The border contour of 6 stands was changed on the map. These 6 stands were also excluded from the further analysis.

Total error of simulated spectral signatures using updated forest inventory data is compared to that of using the 2018 inventory data in Figure 7. We see both increase and decrease in them, while decrease dominates. The aspen stand which was cut had the largest increase of total error, while according to the forestry database of 2018 there was a mature stand. The increase exceeds 0.3 for old pine stands (over 100 years) where other parameters but stand age have changed a little or have not changed at all. This indicates that the influence of stand age on spectral signatures of pine stands is overestimated in the SFRM model. For half of stands with the updated inventory data the total error of the simulated signatures decreased substantially, i.e. more than 0.3, while the maximal decrease was 2.7. This confirms that high values of the total error *S* indicate that stand parameters may have changed since the last inventory and the stands should be inspected in the field.

Conclusions

The statistical forest reflectance model SFRM is a simple and reliable tool for checking of forest data accuracy, using routinely collected but aging forest inventory data and operational satellite information of moderate spatial resolution. The comparison of the Sentinel-2 MSI spectral signatures to the simulated ones reveals the forest stands, inventory data of which may be inaccurate for some reason: outdated due to forest management events, or some damages, or illegal cutting. The stands with large discrepancies between measured and predicted reflectance should be visited *in situ* to check of inventory data. The model was applied for the simulation of forest spectral signatures at different sites in Estonia. Growth conditions vary at these study sites in a wide range from mild mid-latitude at Saaremaa to cold mid-latitude in the eastern regions of Estonia. All main weather parameters (temperature, precipitations, sunshine duration) vary in rather wide range at the study sites. Nevertheless, there were no big differences in the distribution of neither relative differences between simulated and measured spectral signatures nor of the total error S (Eq. 2). The model is simple and computationally efficient. Preparing input data for the model is a simple query in the forest inventory database. The suggested procedure can be incorporated into the automated systems of continuous forest inventory. The most problematic step in the routine application of this procedure is the preparation of satellite data – selection of satellite scene and preparing input data for the atmospheric correction of satellite images.

Acknowledgements

This study was financially supported by the European Regional Development Fund within the National Programme for Addressing Socio-Economic Challenges through R&D (RITA). Copernicus Sentinel data of 2019 provided by the Estonian Land Board. The forest inventory database was provided by the Estonian Environment Agency. The authors thank two anonymous reviewers for commenting of the manuscript. The authors thank Dr. Tõnu Viik for the language proofreading of the manuscript.

References

Burkhart, H. and Tomé, M. 2012. Modeling forest trees and stands. Springer, Dordrecht. 457 pp. https://doi.org/10.1007/978-90-481-3170-9.

ESA. 2019. Sen2cor v2.8. Accessible at: http://step.esa.int/main/third-party-plugins-2/sen2cor/sen2cor_v2-8.

ESTHub. 2019. Esthub services. Accessible at: https://geoportaal.maaamet.ee/eng/Services/ESTHub-Services-p654.html.

Estonian Land Board. 2019. Spatial data. Orthophotos. Accessible at: https://geoportaal.maaamet.ee/eng/Spatial-Data/Orthophotos-p309.html.

Fazakas, Z., Nilsson, M. and Olsson, H. 1999. Regional forest biomass and wood volume estimation using satellite data and ancillary data. *Agricultural and Forest Meteorology* 98/99: 417–425. https://doi.org/10.1016/S0168-1923(99)00112-4.

Ferretti, M. and Fisher, R. (Eds.) 2013. Forest Monitoring. Vol. 12. 1st ed. Methods for terrestrial investigations in Europe with an overview of North America and Asia. Elsevier, Oxford, UK. 536 pp.

Government of Estonia. 2015. Metsa korraldamise juhend [Forest inventory act]. In: Riigi Teataja. RT I, 24.11.2015, 6 (in Estonian). Available online at: https://www.riigiteataja.ee/akt/13124148.

Jakubauskas, M. and Price, K. 1997. Empirical relationships between structural and spectral factors of Yellowstone lodgepole pine forests. *Photogrammetric Engineering and Remote Sensing* 63: 1375–1381.

Kiviste, A. 1997. Eesti riigimetsa puistute kõrguse, diameetri ja tagavara vanuseridade diferentsmudel 1984.–1993. a. metsakorralduse takseerkirjelduste andmeil [A differential equation model of the height, diameter and volume depending on age series and site factors for Estonian state forest stands over the period of 1984–1993 according to forest management assessment descriptions]. In: Eesti Põllumajandusülikooli teadustööde kogumik [Transactions of the Estonian Agricultural University], 189. EPMÜ, Tartu, p. 63–75 (in Estonian).

Kull, K. and Kull, O. 1989. Dinamicheskoe modelirovanie rosta dereviev [Dynamic modelling of tree growth]. Valgus, Tallinn, Estonia, 232 pp. (in Russian with English summary and contents).

Kuusk, A. 1998. Monitoring of vegetation parameters on large areas by the inversion of a canopy reflectance model. *International Journal of Remote Sensing* 19: 2893–2905. https://doi.org/10.1080/014311698214334.

Kuusk, A. and Kuusk, J. 2018. Estimation of variable atmospheric parameters for the atmospheric correction of satellite images. *International Journal of Remote Sensing* 39: 565–576. https://doi.org/10.1080/01431161.2017.1388937.

Kuusk, A., Kuusk, J. and Lang, M. 2019. A statistical forest reflectance model. *Remote Sensing* 11: 2749. https://doi.org/10.3390/rs11232749.

Lang, M., Arumäe, T., Lükk, T. and Sims, A. 2014. Estimation of standing wood volume and species composition in managed nemoral multi-layer mixed forests by using nearest neighbour classifier, multispectral satellite images and airborne lidar data. *Forestry Studies | Metsanduslikud Uurimused* 61: 47–68. https://doi.org/10.2478/fsmu-2014-0010.

Poso, S., Karlsson, M., Pekkonen, T. and Härmä, P. 1990. A system for combining data from remote sensing, maps and field measurements for forest planning purposes. Research Notes No. 23. University of Helsinki, Department of Forest Mensuration and Management, 40 pp.

Raudsaar, M., Sims, A., Timmusk, T., Pärt, E., Nikopensius, M. and Matson, T. 2019. Mets. 1. Metsavarud [Forest. 1. Forest resources]. In: Raudsaar, M., Siimon, K.L. and Valgepea, M. (Eds.) Aastaraamat Mets 2018 [Yearbook Forest 2018]. Keskkonnaagentuur, Tallinn, 75 pp.