Identifying European Old-Growth Forests using Remote Sensing: A Study in the Ukrainian Carpathians

Benedict D. Spracklen and Dominick V. Spracklen

Forests **2019**, *10*(2), 127; https://doi.org/10.3390/f10020127

Keywords: old-growth forest; multispectral satellite imagery; random forest; forest classification

Background

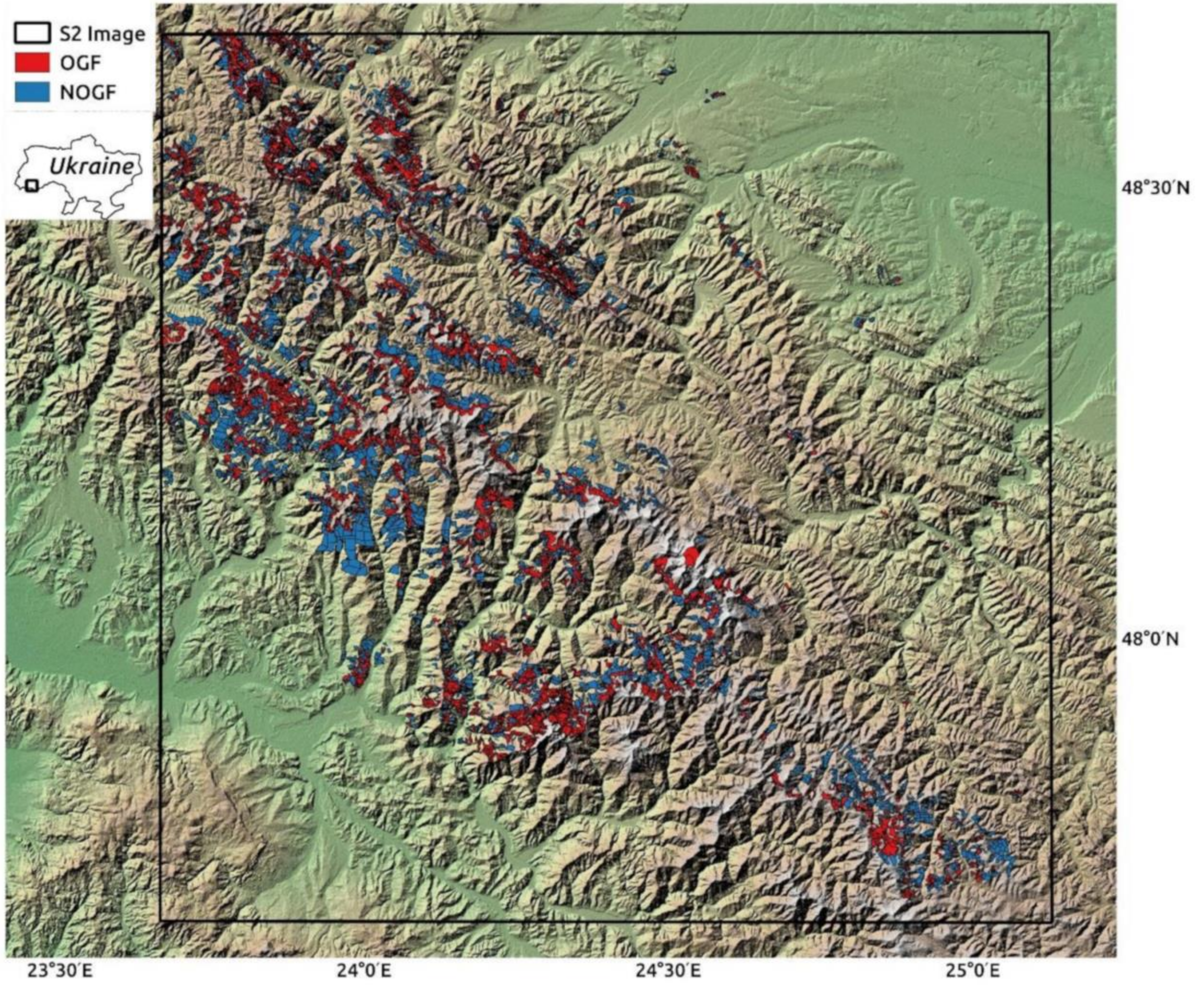
Purpose:
This research aims to use remote sensing methods to identify the old-growth forests, an important, rare, and endangered habitat in Europe, and the dominant tree species within old-growth forests, which would be helpful for both conservation and forest management.

Data:

1. Data on Beech, Norway spruce, and mountain pine old-growth forests in the Ukrainian Carpathians.
2. Summer and autumn 2017 Sentinel-2 satellite images comprising 10m and 20m resolution bands.

Product:
6 vegetation indices and 9 textural features

Method:
Random Forest classification model



Background

What is old-growth forest:
Old-growth forest (OGF), also referred to as primary, virgin or ancient forest, are forests that have developed for a long period of time without significant human intervention and are characterised by the presence of old and large trees, multi-layered vertical structure and abundant standing and lying deadwood in different stages of decay.

Function of OGF:
Supporting significant biodiversity, storing and sequestering large amounts of carbon, and buffering microclimate.

Why use remote sensing to identify OGF:
Classic identification requires time-intensive field surveys that generally involve surveying indicators such as dead wood quantity and quality, forest structure, and the degree of anthropogenic influence. Enabling the identification of such stands by remote sensing or even establishing the sites of potential OGF stands that could later be verified by field teams could help save time and expense.

Method:
Random Forest classification model

OGF and NOGF Data

Tree Species	Number of Polygons	Area (km^2)	Mean Elev (m)	Min Elev (m)	Max. Elev (m)	Mean Slope (º)
Beech	1281	139.2	1055	394	1565	24.2
Oak	21	3.2	507	334	871	13.2
Mountain Pine	219	37.4	1477	1061	1982	22
Norway Spruce	1784	182.4	1343	519	1688	22.1
Silver Fir	20	1.3	598	481	946	12.5
Beech BCMix	189	16.0	1052	425	1443	24.5
Norway Spruce CBMix	226	19.2	1136	514	1620	29.2
Silver Fir CBMix	60	5.3	933	515	1286	24.8
Beech BMix	59	5.1	1039	454	1497	26.1
Other Bmix	15	1.5	618	342	1131	14
Other BCMix	6	0.7	1410	1030	1719	21.2
Norway Spruce CMix	98	10.6	1266	703	1568	22.3
Other CMix	48	6.0	1209	591	1953	23.1
Other CBMix	2	0.1	1598	1374	1722	24
Other B	2	0.07	1522	1422	1633	31.2
Other C	4	0.15	929	733	1373	24.5
Total Conifer	2173	237.6	1341	481	1982	22
Total Broadleaf	1378	149	1042	334	1633	24
Total Mixed	486	41.4	1084	342	1689	24.3
Total	4037	428	1208	334	1982	23

Forest Type	Number of Polygons	Area (km^2)	Mean Elev. (m)	Min Elev. (m)	Max Elev. (m)	Mean Slope (º)
Conifer	2563	299.6	1238	457	1792	20.2
Broadleaved	1343	206.1	888	357	1456	23.6
Mixed	543	57.5	1045	438	1566	22.9
Total	4449	560.5	1108	357	1792	21.5

OGF samples (left image):
This data was provided by WWF Ukraine and covered the survey years 2010–2017 inclusive. This survey includes information on the location and spatial extent of OGF (shapefile polygons of identified OGF stands) as well as detailed information on tree species composition and age.

NOGF samples (top image):
Created 4000 polygons randomly located within a buffer of 2 km of the OGF. This distance was chosen as it enabled the requisite number of appropriately sized NOGF polygons to fit in.

Sentinel-2 Data

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)
B2–Blue	0.490	10
B3–Green	0.560	10
B4-Red	0.665	10
B5–Red edge	0.705	20
B6–Red edge	0.740	20
B7–Red edge	0.783	20
B8–Near IR	0.842	10
B8A–Near IR	0.865	20
B11–SWIR	1.610	20
B12–SWIR	2.190	20

Image used:
Use the 10 and 20m bands of the Sentinel-2 (S2) features.

Two S2 images were downloaded as Level-1C Top-of-Atmosphere reflectance products: one for summer (2 August 2017) and one for autumn (16 October 2017)

Image evaluation:
Used object-based classification because the WWF data included mixed forest polygons which suited an object-based approach. The mean and standard deviation of the pixel spectra and the mean of the associated vegetation indices and textural features within a forest polygon were used for the analysis.

Modeling

6 vegetation indices:
2 forest classification indices:
the Normalized Vegetation Difference Index (NDVI)
the Enhanced Vegetation Index (EVI)

2 forest structure indices (OGF has a more heterogeneous structure):
Advanced Vegetation Index (AVI)
the Shadow Index (SI)

Distinguishing mature and OGF through the difference between SWIR and NIR bands:
Normalized Difference Infrared Index (NDII)

Exploit information in the red-edge bands
The red edge Normalized Difference Vegetation Index (RENDVI)

9 texture features (spatial variation):
Texture measurements quantitively describe relationships of spectral values with neighboring pixels, which information has been used to improve forest stand classification accuracy.

the Grey Level Co-occurrence Matrix (GLCM):
OGF has more contrast in brightness

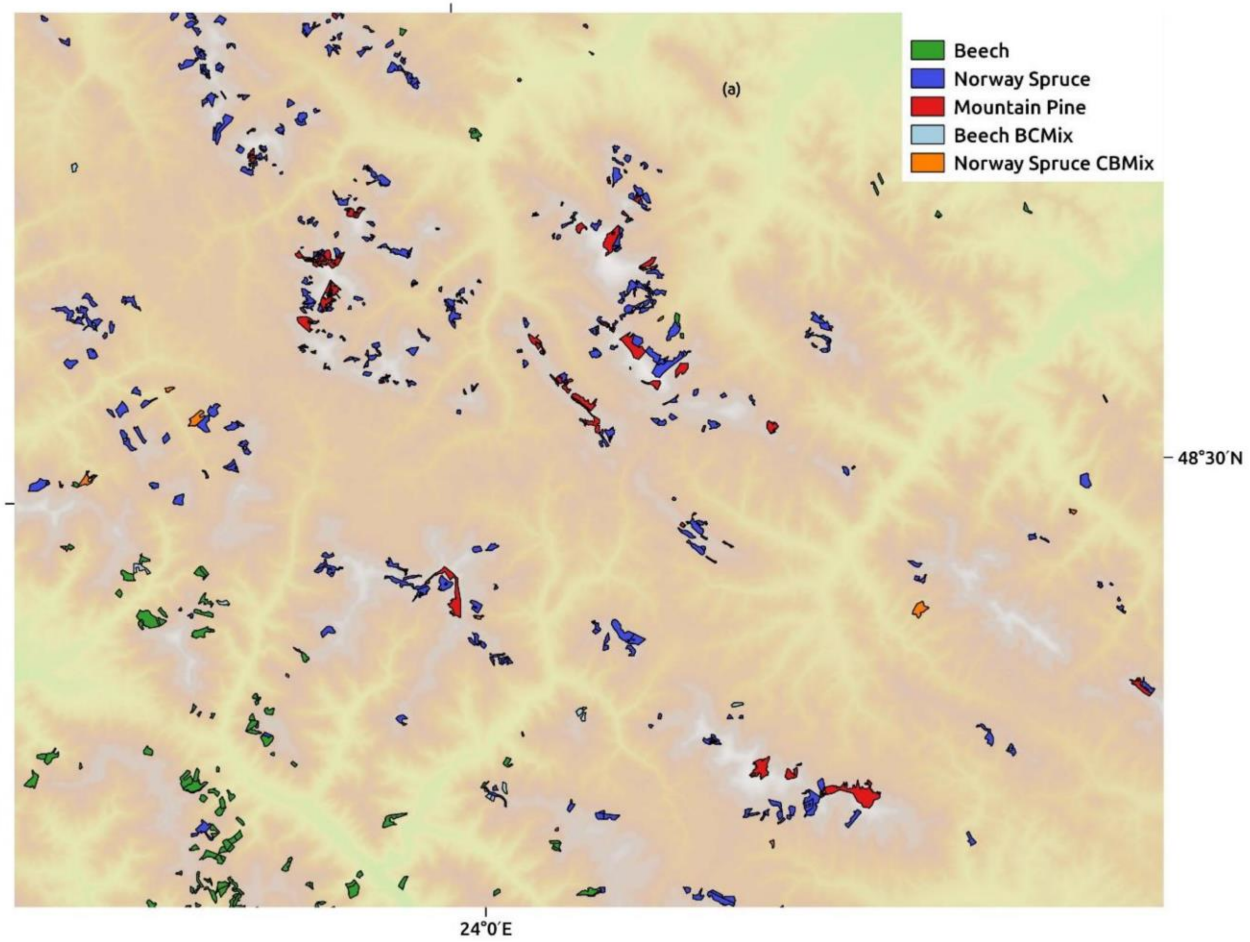
9 texture features are contrast, entropy, and GLCM mean for visual, near IR, and shortwave IR band (B3, B8, and B12)

Random Forest

Scikit-learn Python library:
The polygons were randomly divided into training and validation sets in a ratio of 75% and 25% respectively.

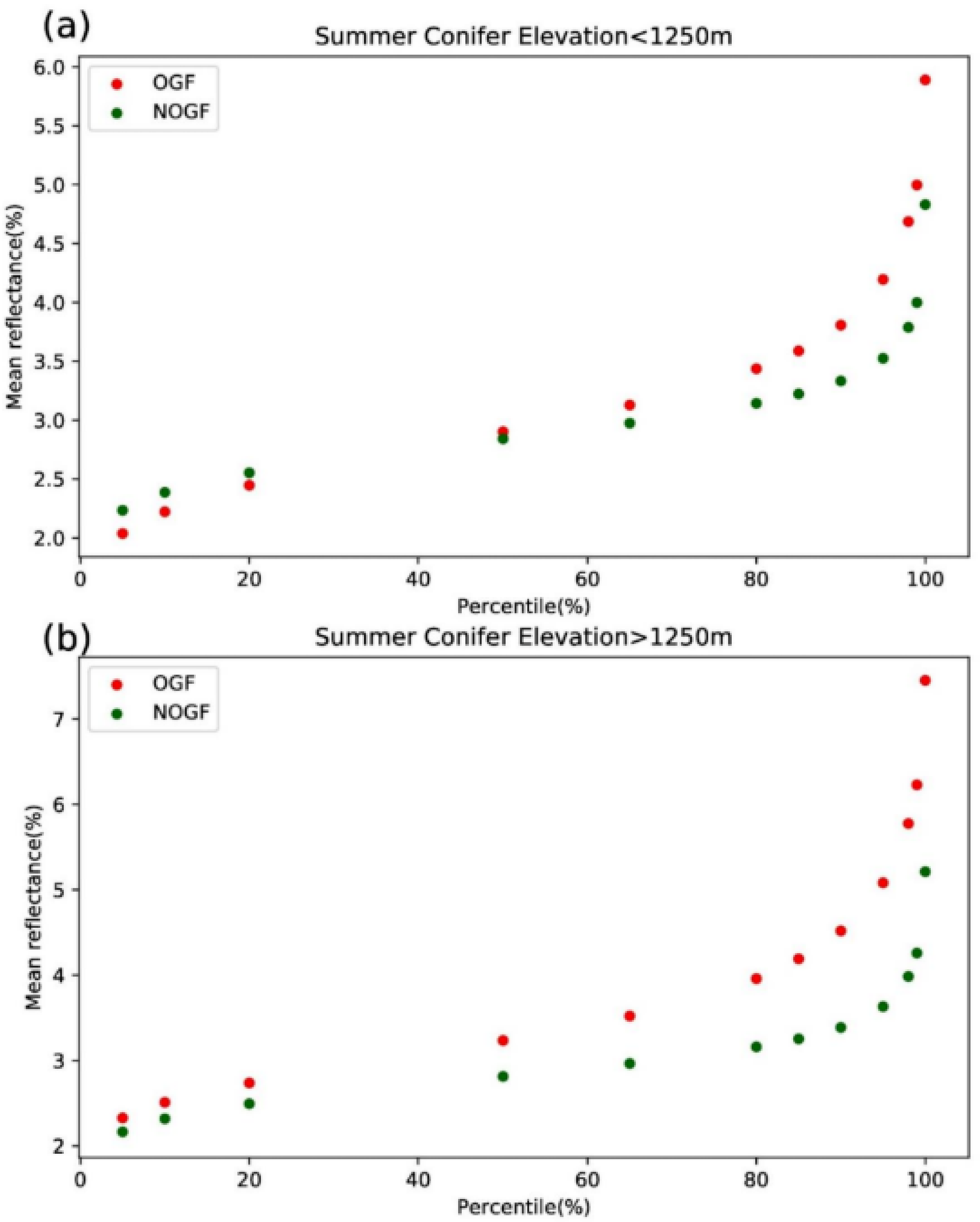
Maximum number of features Random Forest was allowed to try:
The square root of the total number of features.

Number of trees: 500



Key Takeaways

1. An overall accuracy of about 85% was achieved in separating OGF from the surrounding forest, with classification accuracies higher for conifer and broadleaved than mixed forest.
2. The addition of band standard deviations, combining summer and autumn images and adding elevation data improved overall accuracy.
3. Vegetation indices gave only a minimal performance improvement.



Question and Discussion

1. The control NOGF polygons were usually forests lower in height, which makes the samples not truly random.
2. ground identification will generally include criteria such as deadwood quantity and quality, presence of non-native tree species, and human impact such as livestock grazing that cannot be surveyed remotely.
3. Potential improvements could involve exploring the use Support vector Machines (SVM), which has been found to be more accurate than Random Forest in tree species classification studies.
4. OGF could cover more tree species

How will this study deal with future climate change when finding the location of OGF?