A Benchmark Dataset for Sentinel-2 Based Forest Type Classification in the Siberian Summergreen-Evergreen Forest Transition Zone

Femke van Geffen, Ronny Hänsch, *Senior Member, IEEE*, Begüm Demir, *Senior Member, IEEE*, Stefan Kruse, Ulrike Herzschuh, and Birgit Heim

***Abstract*—Circumboreal forests covering about 30% of global forested areas are undergoing significant changes. In Siberia, global warming may reduce the dominance of summergreen larch forest inducing shifts towards evergreen forest types, specifically in the Eastern Siberian summergreen–evergreen forest transition zone. We create a remote sensing training dataset for summergreen and evergreen forest types from the SiDroForest Sentinel-2 image dataset. This new training dataset informed by expert field knowledge includes nearly two million Sentinel-2 pixels across the early summer, peak summer, and late summer phenophases. We create the equivalent seasonal SiDroTest dataset linked to in situ forest plots for benchmarking the seasonal training dataset. To optimize satellite-based monitoring, we train a random forest classifier on the train dataset to map summergreen and evergreen forest resulting in accuracies of 63% for early summer, 89% for peak summer, and 99% for late summer, with an average accuracy of 82% across all seasons. Feature importance analysis highlights the Sentinel-2 shortwave infrared as crucial for distinguishing forest types in all seasons. Additional key features include the normalized difference vegetation index (NDVI) and the red wavelength region for early summer, shortwave infrared and the visible wavelength region for peak summer, and shortwave infrared, near-infrared and NDVI for late summer. This study provides a benchmarked training dataset for mapping boreal forest types in the Siberian summergreen–evergreen transition zone. The random forest classifier performs best in late summer, leveraging distinct spectral differences between evergreen forests' greenness and the seasonal coloring of summergreen larch forests.**

***Index Terms*—Boreal forest, feature importance, multitemporal, random forest (RF) classification, Sentinel-2, summergreen–evergreen forest transition zone.**

Received 19 March 2025; revised 3 April 2025; accepted 6 April 2025. Date of publication 21 April 2025; date of current version 4 August 2025. The work of F. van Geffen was supported by the HEIBRiDS graduate school. This work was supported in part by Helmholtz Einstein International Berlin Research School in Data Science (HEIBRIDS), in part by Helmholtz Information & Data Science Academy (HIDA), in part by Methoden- und Anwendungsentwicklung AIvergreens KI-basierte Klassifizierung von sommer-/immergrünen Nadelbaumbeständen in borealen Wäldern in Sibirien (joint BMWi KI Project AWI, DLR, Technical University of Berlin), and in part by European Research Council ERC Consolidator under Grant 772852 'Glacial Legacy'. *(Corresponding author: Birgit Heim.)*

Femke van Geffen, Stefan Kruse, and Birgit Heim are with Polar Terrestrial Environmental Systems Research Section, Alfred-Wegener-Institute Helmholtz Centre for Polar and Marine Research (AWI), 14473 Potsdam, Germany (e-mail: femke.van.geffen@awi.de; stefan.kruse@awi.de; birgit.heim@awi.de).

Ronny Hänsch is with Microwaves and Radar Institute, German Aerospace Center (DLR), 82230 Weßling, Germany (e-mail: ronny.haensch@dlr.de).

Begüm Demir is with Berlin Institute for the Foundations of Learning and Data (BIFOLD), Technical University Berlin, 10587 Berlin, Germany (e-mail: demir@tu-berlin.de).

Ulrike Herzschuh is with the Institute of Environmental Science and Geography, University of Potsdam, 14476 Potsdam, Germany, and also with the Polar Terrestrial Environmental Systems Research Section, Alfred-Wegener-Institute Helmholtz Centre for Polar and Marine Research (AWI), 14473 Potsdam, Germany (e-mail: ulrike.herzschuh@awi.de).

Digital Object Identifier 10.1109/JSTARS.2025.3562912

I. Introduction

CIRCUMBOREAL forests located primarily in Alaska, Canada, and Northern Eurasia represent close to 30% of all forested land and are strongly changing in response to climate and increasingly frequent disturbances such as fires and drought [1], [2], [3], [4], [5], [6]. For Eastern Siberia, the dominant summergreen needleleaf *Larix* forests have been consolidated within a complex vegetation-fire-permafrost-equilibrium since the past Glacial [4], [5]. These summergreen needleleaf forests are now at risk of being slowly replaced by mixed to evergreen forests [2], [4], [5]. Understanding boreal forest type distribution, in relation to potential changes due to climate change and disturbance, is instrumental for assessing the impact of global warming and to better understand the consequences for the boreal permafrost ecosystems [6], [7]. Specifically, reliable data on the spatial distribution of forest types is needed to allow detailed assessments of the drivers and environmental niches of the current forest composition [8] and to enable monitoring methods of the potential changes over time.

However, for the Siberian summergreen–evergreen forest transition zone, determining a reliable distribution of forest types remains a challenge due to the scarcity of field observations and missing regional assessments of available Land Cover products for this region. Global land cover products that contain summergreen and evergreen forest categories are the European Space Agency's (ESA) Climate Change Initiative Globcover 300 m product [9] and the Copernicus PROBA-V 100 m product [10], however these products are not yet quality assessed for the Siberian boreal region [8]. For some of the regions in Yakutia and Chukotka (Eastern Siberia) where we have field knowledge [11], [12], [13], [14], [15], [16], [17] the forest composition in the global Land Cover products seems not to be reliable [8]. Other available global Land Cover products with higher spatial resolution, e.g., the ESA World Cover 2020 [18], derived from Sentinel-2 (S-2) and the newest ESA World Cover 2021 derived from Sentinel-1 and S-2 [19], both with 10 m resolution, only

have one class for 'forest', not distinguishing between summergreen needleleaf and evergreen needleleaf forest types. A reliable forest product at a high spatial resolution such as from S-2 satellite data would provide deeper insights into the current state of evergreen and summergreen forest in Siberia and could support the monitoring over time.

The accuracy of land cover maps derived from remote sensing data via machine learning strongly depends on the quality of the training and test datasets [20].

Therefore, in this study, we develop a benchmarked training dataset to map summergreen and evergreen needleleaf forest types. We select a random forest (RF) classifier, an ensemble learning based classification approach consisting of a large number of decision trees constructed during the training process [21], [22]. RFs have been successfully applied to Landsat and S-2 satellite data for Land Cover classification, as they are able to avoid overfitting, have low computation load, and are specifically robust with high accuracy for forest mapping across many continents and forest types [23], [24], [25], [26], [27]. Forest inventories are commonly used to build training datasets for forest related applications. For North America, Massey et al. [28] use extensive forest inventory data from Alaska (US) and Canada to train RF classifiers to map boreal forest composition change based on Landsat satellite long term series. In contrast, for Siberia, publicly available in situ forest inventories and forest inventories-related data publications remain rare and invaluable resources of information. With the exception of the data collection of Schepaschenko et al. [29] that contains plot-level data prepared for biomass estimation, these few publicly available forest inventory data for Siberia are not specifically prepared for machine learning and remote sensing applications.

In our previous work, the SiDroForest (Siberian Drone Mapped Forest Inventory) data collection [16], we produced a remote sensing training dataset containing forest type labeled S-2 image patches [30]. These 11 SiDroForest forest type labels linked to in situ forest plots from expeditions in Siberia [30] are highly dispersed in their thematic content and therefore contain only few members per class. A dataset that would better fit the machine learning requirements for forest remote sensing such as sample size and aligned and labeled S-2 image patches is the TreeSatAI Benchmark Archive [31]. TreeSatAI however represents central-European forest as it contains the labels of 20 European tree species derived from forest administration data of the federal state of Lower Saxony, Germany. While TreeSatAI represents a large and useful dataset, it does not cover forest types that are present in the Siberian summergreen-evergreen forest transition zone and it contains training data focused on one vegetation phenophase, the summer peak season.

However, we consider that the integration of the early to late summer forest phenophases might be useful to enhance spectral separability for forest-type classification in the Siberian summergreen-evergreen forest transition zone. Since the 1990s, research teams have explored the value of multitemporal analysis, in the form of seasonal and multiyear time series, primarily using long term data from the USGS/NASA Landsat satellite mission. Vegetation phenology is regarded as both a valuable source of information and a challenge due to the significant seasonal spectral variability it introduces [32], [33]. Using the different spectral characterization of phenophases of vegetation for an enhanced Land Cover classification is so far mainly applied for agricultural mapping [34], [35]. However, the use of multitemporal satellite data series also supports applications focusing on forests [36], [37], [38], [39], [40]. Therefore, in the present study, we investigate if seasonal information is useful to enable forest type mapping in the Siberian summergreen–evergreen transition zone.

The summergreen–evergreen Siberian forest transition zone is dominated by needleleaf summergreen *Larix* tree species *Larix gmelinii* and by needleleaf evergreen Pine (*Pinus sibirica, Pinus sylvestris*) and Spruce (*Picea obovata*) [41]. *Larix* undergoes a fundamental vegetation color change in the late summer season, from green to orange and yellow before finally shedding needles in early winter [42]. In contrast, the evergreen needleleaf tree taxa do not considerably change their spectral reflectance signal and remain "green" throughout the year [43] with only the spectral modulation due to seasonal snow cover.

In this study, we develop a training dataset optimized for summergreen needleleaf forest ("summergreen" class) and evergreen needleleaf forest ("evergreen" class), containing labeled S-2 pixel data for machine learning applications. We create the training datasets in the form of seasonal datasets corresponding to the phenological seasons of early summer, peak summer and late summer, using S-2 acquisitions from the SiDroForest S2-image dataset [30]. In addition, we construct the new S-2 seasonal test dataset SiDroTest linked to reference field data. The performance of the summergreen-evergreen training dataset provides insights into how seasonal factors influence the spectral signal for classification in each season.

The main contributions of this article are summarized as follows.

1) We develop a benchmarked optimized training dataset for S-2-based summergreen and evergreen forest classification, covering the three phenophases of early summer, peak summer, and late summer.
2) We assess whether combining multiple seasons improves prediction accuracy or if a single season is more effective for distinguishing summergreen from evergreen forests using a RF classifier.
3) We identify which S-2 spectral features are most significant in each season or combination of seasons for classifying evergreen and summergreen forests within the Siberian summergreen–evergreen forest transition zone.

II. Data and Methods

A. Study Region and Reference Data

The current study focuses on Eastern Siberia, where joint Russian-German fieldwork was conducted during several years until 2021. During the summer 2018 expedition, a wide range of different forest plots were assessed as part of a collaborative effort between the Alfred Wegener Institute (AWI) Helmholtz Centre for Polar and Marine Research in Germany and the North-Eastern Federal University of Yakutsk [13], [14] (see Fig. 1).

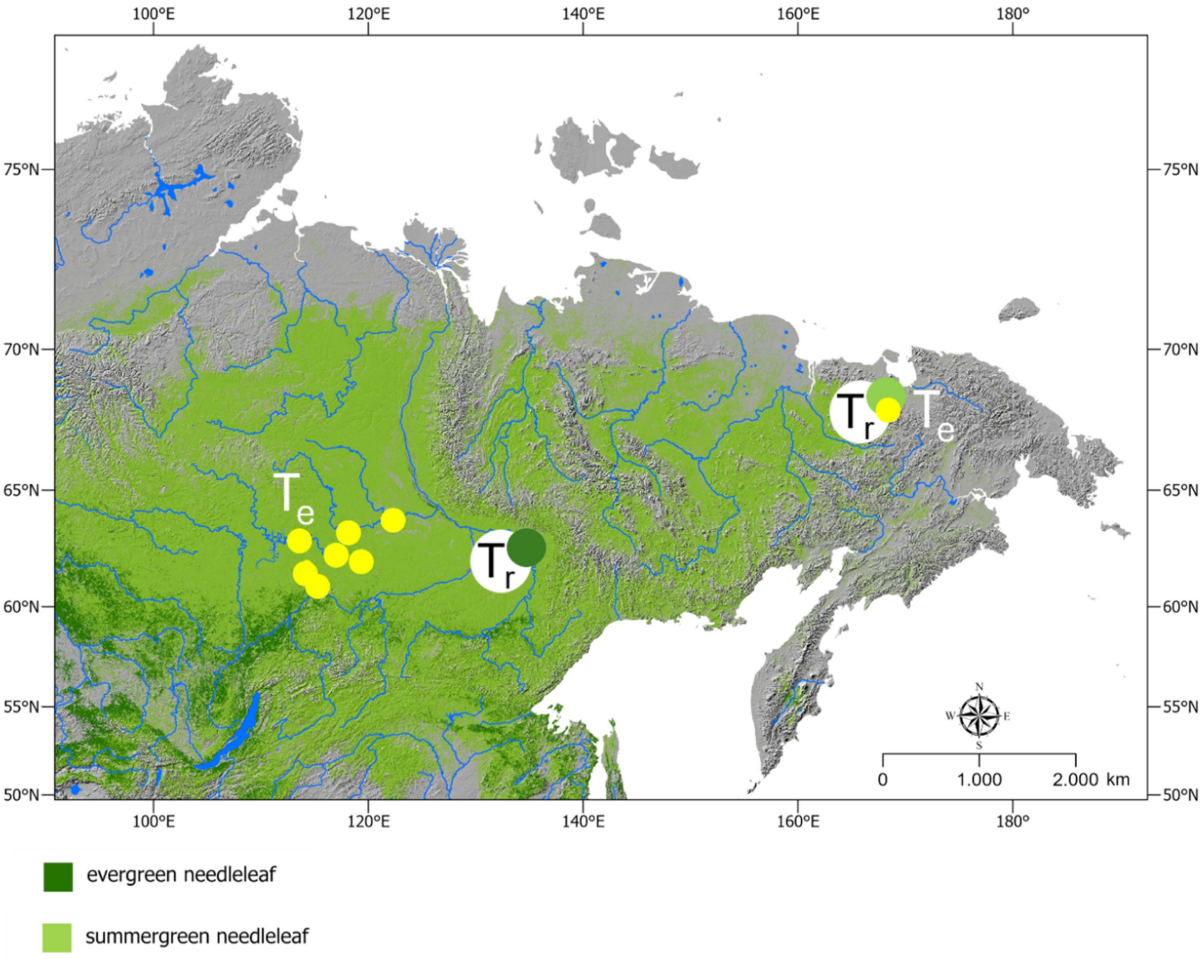

Fig. 1. Mapped evergreen and summergreen needleleaf forest from aggregated ESA CCI Land Cover (300 m spatial resolution) [9] with aggregation described in [4]) and the locations of the Training dataset (Tr) for summergreen (Tr bright green circle) in the Bilibino region and evergreen (Tr evergreen circle) in the Yakutsk region. The Test dataset (Te) created in this study is linked to fieldwork plots recorded in 2018 (yellow circles) [13], [16]).

A detailed forest inventory was performed for each plot using 15 m radius circular plots to estimate the projected cover of trees and tall shrubs [11], nested within 30 m × 30 m rectangular plots to assess the ground projective cover of vegetation taxa [16].

B. Sentinel-2 Data Preprocessing

This study uses pre-Collection-1 S-2 Level-2A (orthorectified surface reflectance, Universal Transverse Mercator / World Geodetic System 1984 (UTM/WGS84)) (v3.0) image subscenes and labeled image patches from our SiDroForest data collection [16], [30] to further generate the training and the test datasets. Specific characteristics of the S-2 SiDroForest dataset are as follows.

1) All S-2 spectral bands in SiDroForest are resampled to the same 10 m spatial resolution with the exception of the 60 m bands that are removed as they target atmospheric-related applications and are not optimal for land surface classification. The normalized difference vegetation index (NDVI) band is added to enhance vegetation signals. An overview of the features included are in Table II.
2) The SiDroForest S-2 dataset contains three phenological time stamps: early summer, ES (May to early June, depending on latitude), peak summer, PS (July to August), and late summer, LS (late August to early October, depending on latitude).

The training dataset is derived from S-2 k-means based clustering and expert knowledge and the test dataset from S-2 pixels linked to the in-situ forest plots [14], [16], [17].

We also acquire Collection-1 Copernicus S-2 Level-2A (orthorectified surface reflectance) data (v4.0) in Standard Archive Format for Europe (SAFE, named SA in this study) as well as a subset in the Copernicus Data Space Ecosystem gridded background format (named GR in this study) for the Lake Khamra region to investigate the transferability of our method. We preprocess the Collection-1 S-2 files similarly to the SiDroForest S-2 images by removing the 60 m spectral bands used for atmospheric correction and using all remaining ten spectral bands and the produced NDVI feature in upsampled 10 m pixel resolution. An overview of the S-2 images used in this study are in Table III.

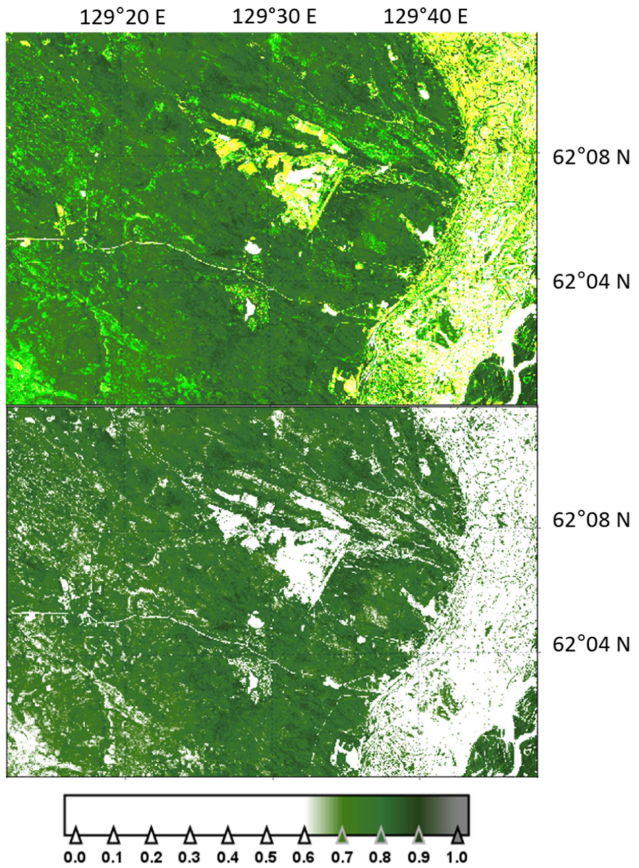

Fig. 2. The two figures show a subscene of the SiDroForest NDVI S-2 image [30] of the Yakutsk (YA) region (showing the Yakutsk city region in the Southeastern part) both displayed as color-coded NDVI maps representing the peak summer in the year 2020. Upper image: NDVI subscene with masked surface water. Lower image: NDVI subscene with masked nonforested areas NDVI $<$ 0.7. Color code is from white (NDVI $<$ 0.7) to green colors for NDVI $>$ 0.7 with nonforest areas (NDVI $<$ 0.05) masked out.

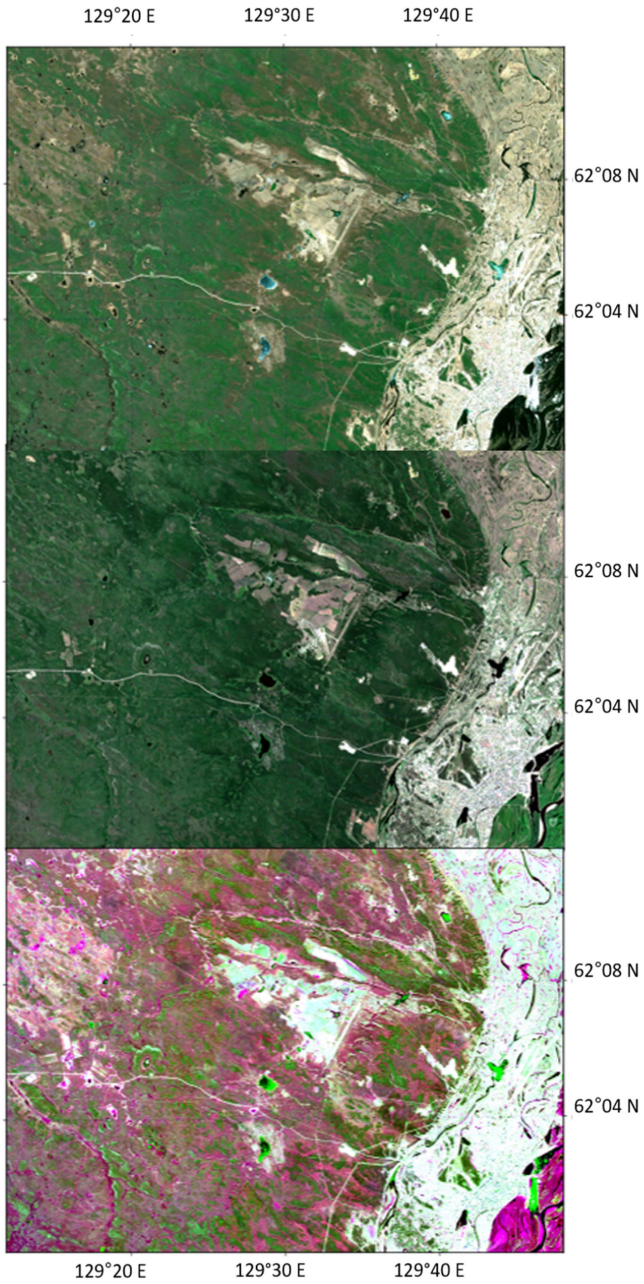

Fig. 3. The same S-2 subscene of the region Yakutsk (YA) (showing the Yakutsk city region in the Eastern part) displayed as quasi-true Red Green Blue (RGB) composites with red = B4, green = B3, and blue = B2 surface reflectance. Upper image: early summer 2020-05-16, center image: peak summer 2020-07-23, bottom image: late summer 2020-09-18. The quasi-true RGB composites are stretched to visualize the landscape contrast. The color-stretched late summer quasi-true RGB image already displays well the contrast between summergreen and evergreen showing purple colors in areas represented by summergreen vegetation in contrast to the evergreen forest patches appearing in bright green colors.

To optimize the available S-2 image subscenes in this study we mask out all non-forest classes. The masked-out nonforest Land Cover includes grasslands, shrublands, agriculture, and barren areas (i.e., rock, floodplains, infrastructure and urban, mining, open soil). These non-forest classes are all characterized by low NDVI during peak summer. A 0.7 NDVI threshold is iteratively found and justified by examining different NDVI threshold values with the expert knowledge of the area from fieldwork. The nonforest mask is created when the NDVI values are at their maximum, in midsummer, using only the peak summer S-2 images. Early summer and late summer S-2 images are also masked as non-forest based on the pixel locations from the peak summer nonforest mask.

Fig. 2 shows as an example the SiDroForest S-2 image subset of the Yakutsk region [30] with surface water masked (NDVI $<$ 0.05) and the final masking of all nonforest classes using the NDVI threshold $<$ 0.7. Fig. 19 (see the Appendix) shows the early spring and late autumn situation with snowmelt in spring transgressing from homogenous snow cover to heterogenous snow cover, often depending on topography and wind pattern, and the increasing snow coverage in autumn. Therefore, the final multitemporal S-2 image data collection covers the three important snowfree forest phenophases: the early summer greening, the maximum peak summer greenness, and the late summer coloring of summergreen forest. Fig. 3 shows an example of the SiDroForest S-2 image [30] in the Yakutsk region for the three seasons. An overview of the S-2 images used in this study is collected in Table III.

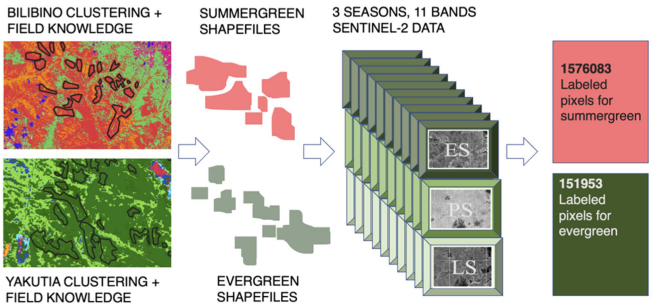

Fig. 4. Creation of the training dataset based on k-means clustering, fieldwork information, and expert knowledge. The final extracted labeled number of pixels per class can be observed on the right.

C. Training Dataset

We select two distinct regions to extract representative pixels from pure forest stands for each forest class for our training dataset. For the summergreen class, we use the Bilibino region SiDroForest S-2 image [30] located in Chukotka (N 68.46, E 163.35 center coordinate), characterized by dense larch forests with high crown coverage. For the evergreen class, we extract pixels from the Yakutsk region SiDroForest S-2 image [30], including the Yakutian capital Yakutsk (N 62.08, E 129.62 center coordinate). In the Yakutsk region, the wider area is dominated by summergreen forests, however, close to the Lena River, dense evergreen forests can grow on the sandy soils due to deep thaw depths in summer. These patches of dense evergreen forests are chosen to create the evergreen training dataset.

We apply the k-means algorithm for pixel-wise clustering of the S-2 images and evaluate the results visually using expert knowledge; this process is shown in Fig. 4. Through iterative visual inspection, comparing with ESRI© high-resolution images we find that eight clusters are sufficient for both regions when using the S-2 stacked spectral bands plus the NDVI. For the Bilibino region, peak summer clusters differentiate well between forest and shrubland. For the Yakutsk region, late summer clusters differentiate well between evergreen and summergreen forests.

After iterative inspection, we create polygonal shapefiles on the classified maps to best represent the summergreen and evergreen forest classes. The manual delineation of these polygonal shapefiles is guided by regional expert knowledge and ESRI high-resolution images. For the Bilibino region, polygonal shapefiles are centered on *Larix* forests without contributions from the *Pinus pumila* evergreen shrub species. For the Yakutsk region S-2, we avoid evergreen forest patches in the urban areas and only delineate evergreen forest on the upper sandy Lena River terrace.

Using the labeled polygonal shapefiles we assign the labels "summergreen" or "evergreen" to the extracted pixels. For each of the three seasons, we extract pixels for each class from all the prepared ten S-2 spectral bands and the NDVI (see Table II in the Appendix). The final training set contains 151 953 pixels for the evergreen class and 1 576 083 pixels for the summergreen class.

TABLE I
PLOT CODES, LOCATIONS AND CLASS LABELS OF THE S-2 IMAGE PATCHES FROM THE SIDROFOREST DATA COLLECTION [14], [16], [30] INCLUDED IN THE SIDROTEST DATASET

PLOT	S2-SUBSET	CLASS LABEL
EN18068	NYURBA	EVERGREEN
EN18070PROXY	NYURBA	EVERGREEN
EN18074	SUNTAR	EVERGREEN
EN18075PROXY	MIRNY	EVERGREEN
EN18077PROXY	MIRNY-LENSK	EVERGREEN
EN18030	BILIBINO	SUMMERGREEN
EN18035	BILIBINO	SUMMERGREEN
EN18063	VILNUYI	SUMMERGREEN
EN18073PROXY	SUNTAR	SUMMERGREEN
EN18076	MIRNY	SUMMERGREEN
EN18078	MIRNY-LENSK	SUMMERGREEN

We undertake a principal component analysis (PCA) on the training dataset with both classes, summergreen and evergreen, together to explore patterns and the spectral separability of the applied S-2 features.

D. SiDroTest Dataset

The in-situ forest plot-linked SiDroTest dataset is made by extracting the pixels from summergreen or evergreen labeled 30 m by 30 m SiDroForest plot polygonal shapefiles [30]. The original 30 m by 30 m SiDroForest labeled image patches [30] represent 11 forest types containing subcategories and mixed forest closely following the in-situ forest inventories [17]. To reduce the class diversity, we assign one of the two forest classes "summergreen" and "evergreen" to each 30 m by 30 m in situ plot shapefile based on the forest inventories. For this study, we select only those in situ plots of dominant or pure stands of evergreen or summergreen from medium dense to closed forest canopies. To achieve a pure summergreen or evergreen spectral signal, all plots that represent early successional stages of forest after fire only, i.e., woodland plots are excluded from the SiDroTest set.

An overview of the tree species distribution in the SiDroForest forest plot inventories that are used in this case study can be seen in Fig. 16 (see the Appendix). We additionally check the in situ and drone image data of each plot as part of the selection process. In four cases, we need to shift the evergreen or summergreen forest plot shapefile to get summergreen or evergreen with denser canopies for a good spectral characterization (see Tables II and III). We undertake shifting to other locations very carefully, in close proximity to the original plots, and guided by field knowledge and ESRI© high spatial resolution background imagery.

The 30 m by 30 m SiDroForest S-2 image patches selected to construct the SiDroTest dataset are from different regions: Bilibino, Vilnuyi, Suntar, Nyurba, Mirny, and Mirny-Lensk [30] (see Table I). The final independent SiDroTest dataset contains 1470 pixels for the evergreen and 1489 pixels for the summergreen class.

E. Summergreen–Evergreen Forest Classification

We train four classifiers, a support vector machine (SVM), a K-nearest neighbors (KNN), a Gaussian Naive Bayes (GNB) and a RF on the training datasets for summergreen and evergreen for all seasons combined (ES, PS, and LS). We train these classifiers to get an initial idea of the separability of the dataset with all seasons included. We observe that testing on a 20% test dataset and training on the remaining 80% results in a perfect or near perfect average accuracy (99% for SVM, 100% for KNN, 98% for GNB and 100% for RF). The high accuracy is likely due to the clear class assignment of summergreen and evergreen in our training dataset. We want to exclude the possibility of overfitting on this specific dataset and therefore for our results test on another, unrelated dataset.

We create the SiDroTest dataset to evaluate the performance on a separate dataset that is linked to fieldwork knowledge. The average accuracy on the SiDroTest dataset for all seasons is 79% for SVM, 80% for KNN, 60% for GNB, and 80% for RF. Based on the outcomes of these trained algorithms and other previous work [22], [23], [24], [25], [26] we select the RF to optimize for our training dataset and we will test on the independent SiDroTest set.

Hyperparameter tuning of the RF on a training dataset adjusts the parameters to the best optimal settings for the specific dataset that we use. With "balanced" class weights, it addresses class imbalance by adjusting the weights inversely proportional to class frequencies. The ensemble comprises ten decision trees (n_estimators = 10), each with a minimum of ten samples required to split an internal node (min_samples_split = 10) and to be at a leaf node (min_samples_leaf = 10). The "max_features" parameter, set to the square root of the input dimensions, controls the number of features to consider when looking for the best split, promoting diversity among trees. The maximum depth of each tree is capped at 70 (max_depth = 70) to prevent overfitting, while bootstrap sampling is disabled (bootstrap = False), ensuring each tree is trained on the entire dataset (151 953 "evergreen" and 1 576 083 "summergreen"). We evaluate the RF with four standard evaluation metrics: precision, recall, the F-1 score, and the average accuracy. In our case, we have an imbalanced dataset that can lead to bias when overall accuracy is used. We therefore choose the average accuracy. The precision shows how often a machine learning model is correct when predicting the target class. Recall shows whether a machine learning model can find all objects of the target class. By using all the metrics, we can evaluate the performance of each class in detail.

Finally, we inspect the feature importance of the RF on the training datasets per season to evaluate the importance of the ten S-2 spectral bands and the NDVI on the classification results. We additionally inspected what the optimal selection of features is by using the recursive feature elimination method [44]. We then use the season with the trained RF with the highest average accuracy to perform a pixel-wise classification of the Bilibino region and Yakutsk region SiDroForest S-2 images and the Lake Khamra region SiDroForest S-2 and Copernicus S-2 images. The results are 10 m resolution classified maps of the two forest classes evergreen and summergreen.

F. Code Availability

The trained random forests per season and an example of how to use them to predict on a Sentinel-2 image are published on GitHub.[1]

III. Results

A. Properties of the Training and SiDroTest Datasets

We compare the surface reflectance of the S-2 bands for the test and training datasets [see Fig. 16(a)–(f) (Appendix), Tables V (4-1), (4-2), (4-3), VI] and observe the strongest separability in the late summer (LS) where the evergreen class shows vegetation vitality in terms of "greenness" in contrast to the summergreen class showing reduced vegetation vitality. This high potential of spectral separability between the summergreen and evergreen class for the late summer season is also shown in the PC1–PC2 biplot visualization optimally separating the two classes of the training dataset along the PC1 axes that is also visualized by strong loadings in the PC1 [see Figs. 5(c), 18(c) (Appendix)]. The early summer season PC1–PC2 biplot visualization also shows a spectral separability potential of the two classes summergreen and evergreen [see Fig. 5(b), 18(a) (Appendix)], with strong loadings in the PC1 and PC2, however not as distinct as in the late summer case for the PC1 axis.

In contrast, the peak summer PC1–PC2 biplot visualization [see Fig. 5(a)] shows that both classes, evergreen and summergreen, overlay each other in the PC1-P2 space confirming a low spectral separability in the first PCs.

The LS S-2 surface reflectance of the evergreen class shows high green reflectance building a green reflectance peak and low blue and red reflectance due to high chlorophyll and additional pigments absorption, together with high NIR reflectance due to vegetation-inherent multiple NIR scattering [see Fig. 16(c), (f) (Appendix)]. In contrast, the LS S-2 surface reflectance of the summergreen class does not contain a green reflectance peak like it shows in peak summer [see Fig. 16(b), (e) (Appendix)], and is further characterized by high red reflectance, low NIR, and a generally flat reflectance curve [see Fig. 16(c), (f) (Appendix)].

Similarly, high NDVI for evergreen stands out versus low NDVI for summergreen in late summer [see Tables V(4-3), VI, *Appendix*] when the summergreen class is characterized by needle coloring and has already lost its vitality. The LS reflectance and LS NDVI show a clear separability between summergreen and evergreen in both data sets, the SiDroTest dataset and the training dataset. The summergreen class shows the most similar spectral shapes between the training dataset and the SiDroTest dataset throughout all three seasons, whereas the evergreen class of the SiDroTest dataset shows lower reflectance in the NIR than the training dataset through all seasons [see Fig. 16(a)–(f) (Appendix)].

The spectral reflectance shape and the NDVI magnitude is much more similar for summergreen and evergreen during peak summer [see Fig. 5(b), (e), Tables IV(3-2), V], in contrast to the more divergent pattern for early summer and late summer

[1][Online]. Available: https://github.com/fvangef/Code-supplement-for-Summer-and-Evergreen.git

Fig. 5. (a)–(c) PC1–PC2 biplots visualize the potential of the separability of the S-2 features of the training dataset for (a) early summer, (b) peak summer, (c) late summer, in descending order, from top to the bottom.

[see Fig. 5(a), (c), (d), (f) Tables IV(3-1), (3-3), V]. In early summer also the contrast between the "greenness" of the evergreen class to the flat spectral curve of the summergreen class comes out [see Fig. 5(a) and (d)]. However, the NDVI range is wide for summergreen and the NDVI difference between evergreen and summergreen is lower than in late summer [see Tables IV(3-3), V].

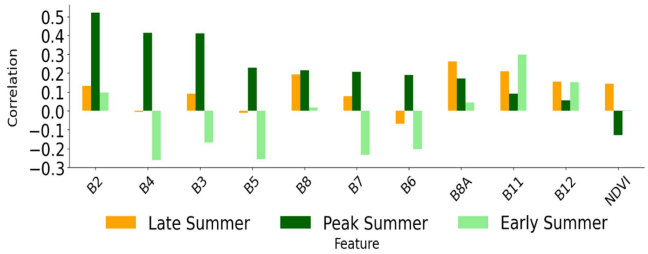

Fig. 6. The correlation coefficient between the training dataset and SiDroTest dataset for each S-2 feature and each season in descending order, early summer, peak summer, and late summer, in descending order, related to the peak summer.

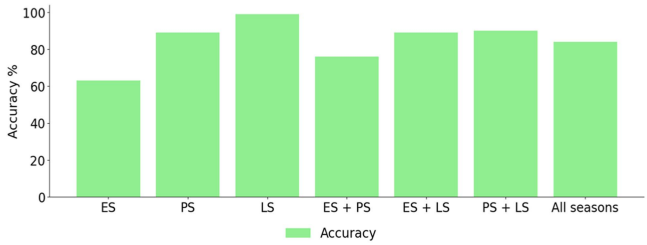

Fig. 7. Accuracy scores in percentages for the early summer, ES, peak summer, PS, late summer, LS, and the combination of the seasons on the SiDroTest dataset.

We then calculate the correlation between the SiDroTest and training datasets to further compare the two datasets' compatibility (see Fig. 6). We observe that there is negative correlation between the early summer training dataset and for the green, the red and the red-edge bands (B3–B7). The highest positive correlation between the early summer training and SiDroTest dataset is for the shortwave infrared (B11).

For the peak summer season, we see positive correlation between the two datasets for all features except the NDVI, with the highest positive correlation for the S-2 bands in the visible wavelength range, the highest for the blue (B2), then the red (B4) and the green wavelength region (B3). For the late summer, we see the highest positive correlation with the NIR (B8, B8A), and in general, a moderate overall positive correlation for all features, except for the mid red-edge band (B6).

The highest average accuracy score is obtained with late summer at 99% and we also observe a high accuracy score for the combination of the peak summer and late summer at 90% (see Fig. 6). The lowest average accuracy score is observed for early summer with 63% for a single season and the early summer and late summer combination at 76% (see Fig. 7).

Upon further inspection of the precision, recall and F-1 score per class per season (see Fig. 8) we see that the results for late summer show 99% average accuracy for both of the summergreen and evergreen classes. We see 100% precision for the evergreen class for peak summer, and 100% recall for the summergreen class, the F-1 score or harmonic mean is similar with 86% for evergreen and 89% for summergreen (see Fig. 8). For early summer we see that the evergreen class outperforms the summergreen class with 65% over 56% . Next, we take a closer look at the precision, recall, and F-1 score for the combination of seasons. Here we observe the highest scores for the combination

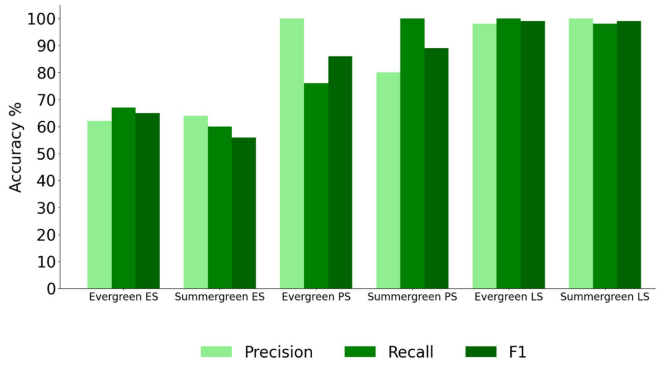

Fig. 8. Percentage scores of precision, recall and the F-1 score for each class for the early summer, ES (left), peak summer, PS (center) and late summer, LS (right) on the SiDroTest dataset.

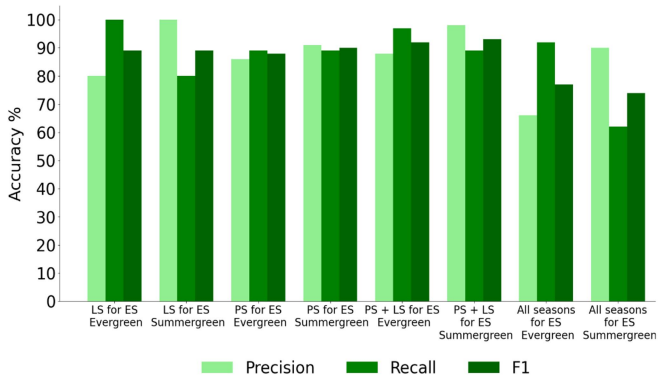

Fig. 9. Percentage scores of precision, recall and the F-1 score for the combinations of seasons, early summer, ES, peak summer, PS, late summer, LS, on the SiDroTest dataset.

of the peak summer and late summer seasons (see Fig. 9). We observe high precision, recall, and F-1 scores for both, evergreen and summergreen. We also inspect the combination of all seasons and observe that the summergreen and evergreen classes both obtain around 80% for all evaluation metrics. The combination of seasons does not outperform the performance of the RF trained with the late summer, alone.

C. Feature Importance

We use feature importance to inspect which features are the most informative per season [see Fig. 10(a)–(c)] by observing the change in mean accuracy when computed with and without each feature. In addition, we use the PCA to understand if a dimensionality reduction toward the most important features will reduce complexity. The S-2 shortwave infrared (B12) ranks as the most important feature for early summer and late summer and as the second most important feature for peak summer. In addition, the NDVI ranks as the second most important feature for early summer and late summer.

We observe for early summer the most important feature to be the shortwave infrared (B12), the second most important the NDVI followed by the red (B4) covering the strong chlorophyll absorption and the green covering the green reflectance peak [see Fig. 9(a)]. The NDVI and the red band's significance highlights

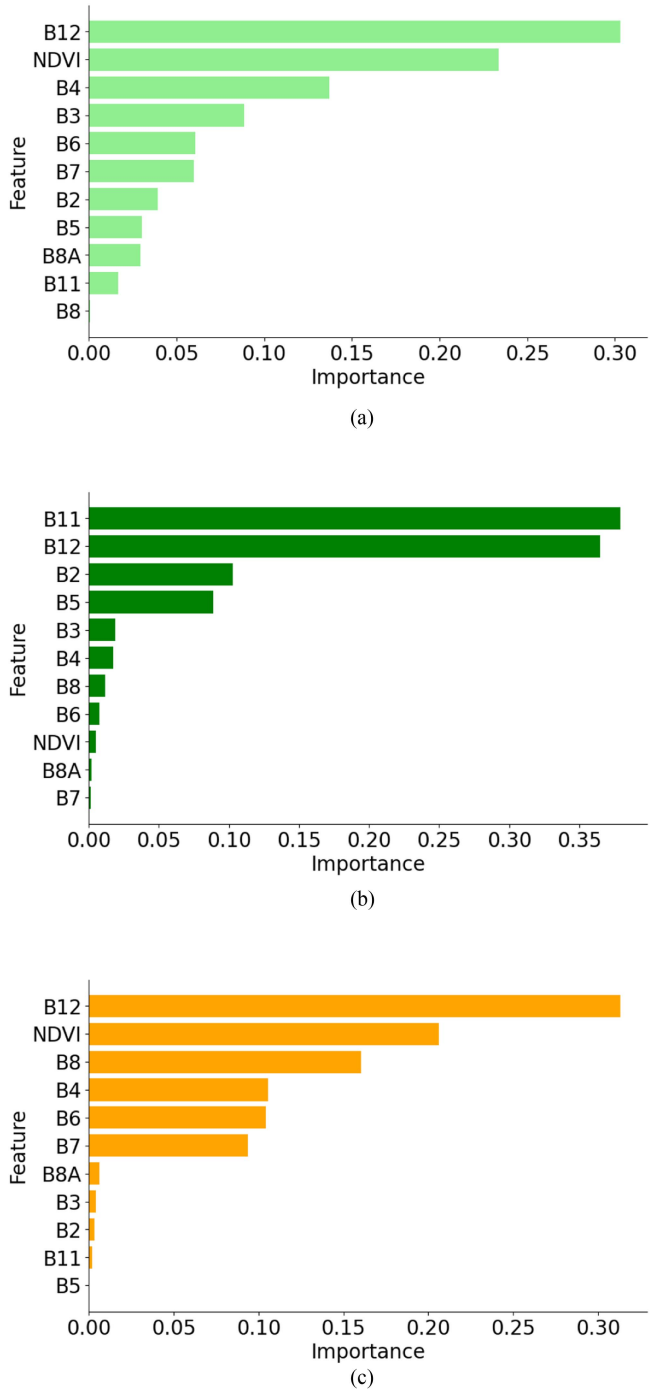

Fig. 10. (a)–(c). Feature importance for NDVI and the S-2 spectral bands for (a) early summer, (b) peak summer, and (c) late summer, in descending order from top to bottom.

the role of vegetation vitality in distinguishing between the "greener" evergreen class and the summergreen class not yet characterized by vegetation vitality, i.e., showing lower NDVI, low chlorophyll absorption and therefore high red reflectance and no green reflectance peak in early summer [see Fig. 16(a) and (d), (Appendix), Tables V (4-1), VI]. In the PCA, an additional but smaller potential of the feature separability for the two classes is indicated in the spread of the evergreen class

along the PC2 versus the summergreen class [see Fig. 5(a)], however with the two classes overlapping. The PC2 loadings also show considerable contributions of NDVI and both SWIR bands, and to a lesser extent, the red pigment absorption band [see Appendix, Fig. 17(a)]. The highest ranked features in peak summer are the shortwave infrared bands (B11, B12) followed by the blue (B2) and the first red-edge band (B5) [see Fig. 10(b)].

Shortwave infrared is specifically sensitive to water content of the vegetation canopy during the summer season when the differences in water content and leaf structure between summergreen and evergreen vegetation seem to be pronounced [see Fig. 16(b), (e) (Appendix), Table V](4-2).

The potential of feature separability seems to be low for peak summer, as both classes, summergreen and evergreen, are similarly distributed in the PC1–PC2 space [see Fig. 5(c)], with all features equally contributing to PC1 [see Appendix, Fig. 17(b)]. During the season of the summergreen vegetation senescence in late summer, the most important features are the shortwave infrared (B12), the NIR (B8) and the NDVI [see Fig. 10(c)] that clearly show the low vegetation vitality and low water content of the senescing summergreen vegetation with low NDVI values, no green reflectance peak and relatively high shortwave infrared (B12) reflectance indicating a low water canopy content [see Fig. 16(c), (f) (Appendix), Tables V(4-3), VI]. Vegetation vitality is also linked to the red pigment absorption band (B4) and the red-edge bands B6 to B7 that follow in the ranking of the feature importance [see Appendix, Fig. 17(c), (f), Tables V(4-3), VI]. In the PCA, the high potential of the feature separability for the two classes shows up along the PC1 axis [see Fig. 5(c)], with the PC1 loadings showing the contributions of NDVI, SWIR and red edge and NIR bands [see Appendix, Fig. 18(c)].

The largest variance of the features is explained in the PCA similarly in each season for the first three PCs [see Appendix, Fig. 17(a)–(c)]. Already the 1st and 2nd PCs in each season contain the majority of loadings of all features, similarly as the feature importance indicating that all features could be kept across the seasons and a reduction of the dimensionality, e.g., to reduce noise is not needed [see Appendix, Fig. 18(a)–(c)]. The heatmap of the early summer and late summer loadings [see Fig. 18(a) and (c)] indicate that all features considerably contribute to the first two PCs, whereas for the peak summer season, the PC1 already contains the major loadings of all the features [see Fig. 18(b)].

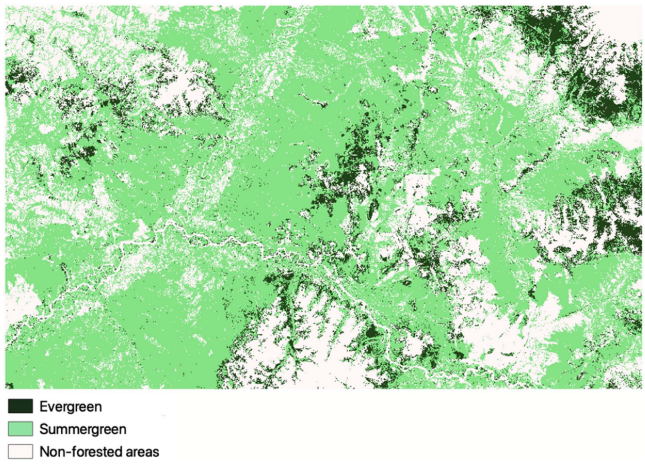

Fig. 11. 2020 SiDroForest S-2 LS image for the Bilibino region [30] classified into the two classes summergreen and evergreen with bright green color assigned to the summergreen forest type class and dark green color assigned to the evergreen forest type class. The white color-coded pixels represent the nonforest classes of peak summer NDVI < 0.7.

Fig. 12 2020 SiDroForest S-2 LS image for the Yakutsk region [30] classified into the two classes summergreen and evergreen with bright green color assigned to the summergreen forest type class and dark green color assigned to the evergreen forest type class. The white color-coded pixels represent the nonforest classes of peak summer NDVI < 0.7.

D. Classified Summergreen–Evergreen Forest Maps

We use the RF trained on the late summer season to map the summergreen and evergreen pixels using the late summer SiDroForest S-2 subscenes. For the Bilibino region, the resulting forest map shows the dominance of summergreen forest in Chukotka (see Fig. 11), agreeing with the Land Cover spatial pattern displayed by visual inspection of the high-spatial resolution ESRI background image and field knowledge, [13], [14], [17]. Evergreen is classified in the higher elevations of mountain areas above the *Larix* treeline representing shrublands composed of *Pinus pumila*, a dense and high growing evergreen shrub with high biomass but not representing evergreen forest. The polygons where we assigned the summergreen labels to extract the clustered dataset are robustly classified as summergreen.

The RF-mapped Yakutsk region shows summergreen forest over a large area as dominating forest type (see Fig. 12), agreeing with thematic and field knowledge [13], [14], [17]. The mapping reveals that the evergreen forest class in the Yakutsk region occurs in mostly linearly elongated features agreeing well with the visual inspection of a high-spatial resolution ESRI background image showing evergreen forest following topographic features. The polygons where we assigned the evergreen labels to extract the clustered dataset are robustly classified as evergreen (see Fig. 13).

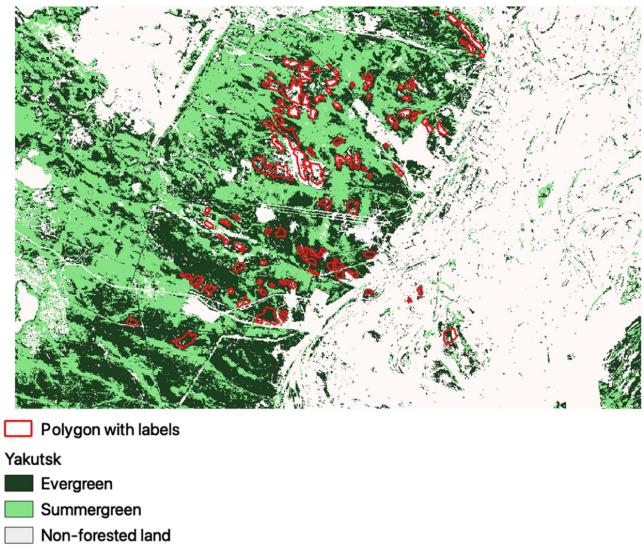

Fig. 13. Mapped evergreen and summergreen needleleaf forest derived from the prepared 10 m S-2 data for the Yakutsk region [30] and the shapefiles for the evergreen-labeled pixel extraction location visualization. The white-colored pixels represent the strict mask (NDVI < 0.7) applied for nonforest abundance. The shapefiles drawn with expert knowledge also included sparser and more diffuse forest patches that were also masked out by the peak summer nonforest threshold mask.

To further examine how well the trained RF works on unseen S-2 data, a part of the Lake Khamra region representing the summergreen-evergreen forest transition zone is classified for the late summer season. We use the pre-Collection-1 SidroForest S-2 surface reflectance image subset from late summer 2020 as well as the Collection-1 S-2 images from late summer 2018 (see Table III *in Appendix*).

Fig. 14 shows the mapped summergreen and evergreen forest types for the three S-2 classifications covering the Lake Khamra region.

IV. Discussion

A. Classification and Importance of Seasons

A wide range of studies show highly diverse performances of the multiseasonal approach across a range of forest types in different climates with different seasonal drivers. However, multitemporal phenological studies for forests in Central and Eastern Siberia have not been carried out up to date. In our case study, the evaluation metrics (accuracy, precision recall, and the F-1 score) of the summergreen–evergreen RF land cover classification on S-2 surface reflectance data show that the late summer season outperforms the other seasons of early summer and peak summer. The combination of peak summer and late summer also outperforms the other combinations of seasons. We find that using multiple seasons together in the same RF classification (92% accuracy) does not improve the classification of the forest types when compared to the late summer dataset (99% accuracy).

The reason why multiseasons performed better than only one season in other studies seems to be linked to regionally specific forest types and their seasonal spectral characterizations. Immitzer et al. [38] find in their multitemporal forest classification

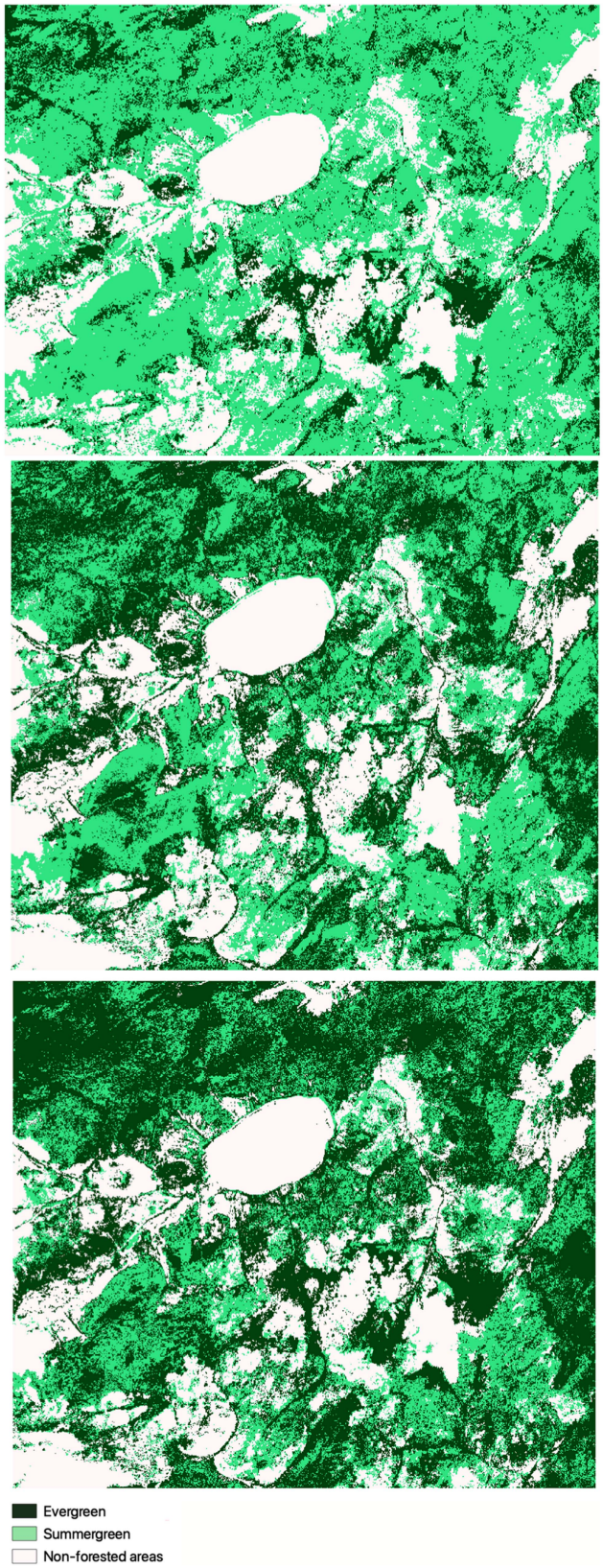

Fig. 14. (a)–(c). Top: late summer 2018 Collection-1 SA S-2 RF-classification, Center: late summer 2018 Collection-1 GR S-2 RF-classification, Bottom: late summer 2020 pre-Collection-1 SidroForest S-2 [30] RF-classification. The white color-coded pixels represent the non-forest classes of peak summer NDVI < 0.7.

that it is difficult to depict the importance of a season, as several effects are counterbalancing, e.g., satellite image quality or the vegetation's specific phenophase. Also a challenge remains in regions and seasons with high cloud coverage retrieving all phenophases, as discussed in Persson et al. [45]. The authors [45] report a successful RF S-2 classification in the case study of a mixed mature broadleaf and needleleaf forest in Sweden with five different tree species with the late spring image as the most informative season to separate tree species at 80.5% accuracy. The authors in [45] discuss that the high performance of the late spring season in their case study is due to the outstanding spectral characteristics of the intense greenish and dense broadleaf forest taxa development versus the "less green" needleleaf forest taxa. This is in contrast to our study case region where we encounter in early summer summergreen needleleaf forest not yet showing high vegetation vitality and greenness. In [28], RF models for the North-American boreal forest trained with three-season Landsat satellite data from all summer phenophases had a better performance than trained with the peak summer only. We also encounter in our study that the peak summer season has a low spectral separability. We did a first experiment in [8], using late summer S-2 images for forest type classification in Siberia that was successful [46]. Our work in [8] also shows that to apply gradual forest classes [46] was challenging in terms of accuracy and validation. Similarly, in our earlier classification works in [41] producing land cover maps for estimation of pollen productivity in Siberia we did not encounter as robust results as with the classification in our study using the aggregated two forest classes summergreen needleleaf and evergreen needleleaf.

In their work, Schriever and Congalton [39] and Mickelson et al. [40] noted that satellite images from the start and the end of the growing season are important, since in the early and late summer season the phenological separability between tree species seems to be the highest.

We, however, see that including the early summer, the best performing time window in the Persson et al. [45] study, in our study lowers the accuracy of the RF classification when the early summer is used on its own or in combination with the other seasons. For our case study and forest types, the early summer season also shows high spectral separability but we see the lowest classification results for the summergreen and evergreen pixels and negative correlation between test and training datasets. This could be related to the early season ground vegetation cover providing high greenness and high spectral variability across the SiDroTest dataset. In addition, in terms of challenges for a robust application, the early season start in the high latitude boreal biome changes between years due to the timing of snowmelt and spring temperatures. Integrating the peak summer season in our case study still keeps a high performance. However, the high cloud coverage in the case study region also limits high-quality optical satellite imagery in the summer months. In the late summer in our investigated region, the *Larix* late summer coloring occurs in a wide time window from late August to early October before the shedding of needles. Despite lower sun zenith angles, the spectral quality of S-2 images stays high.

This indicates that a summergreen–evergreen forest classification using the late summer season is a robust approach for this specific Siberian region rather than working fully multi-temporal that will be limited by the availability of cloud-free, smoke-free, and snow-free optical satellite imagery throughout all seasons.

B. Classification and Feature Importance

There are key differences in feature importance in remote sensing-based land cover classification studies, related to different sensors, land cover types, the season, the amount of training data and the geographic location which limits the comparison of such studies. In our study, the order of the feature's importance is different in different seasons. We observe, similar to Massey et al. [28], that in all three seasons the shortwave infrared is an important feature. This high feature importance indicates that differences in water content of the vegetation canopy and structural properties of the vegetation are key differentiators between the summergreen and the evergreen forest classes.

In the early summer, the shortwave infrared, the red covering the chlorophyll absorption and the NDVI rank important for the RF to distinguish between evergreen and summergreen forest classes. In the peak summer, the shortwave infrared as well as the visible wavelength region are shown as most important. For late summer, we see a large impact of the NDVI and the NIR in addition to the long short-wave infrared. Evergreen forests generally maintain higher water content in the canopy throughout the year compared to deciduous forests with coloring and leaf fall processes in late summer. In contrast, during peak summer, the spectral separability for both forest classes in their full vegetation vitality is low with the exception in the shortwave infrared. The NDVI that builds on using the red and the NIR supports distinguishing between evergreen and summergreen during the seasons when summergreen shows less "greenness" and vegetation vitality, i.e., specifically during the early and late summer.

The reflectance of evergreen forests in early and late summer is characterized by high chlorophyll absorption in the red and high multiple NIR scattering, resulting in low red and high NIR reflectance. Both processes combined result in higher NDVI values for the evergreen tree canopy in comparison to summergreen tree canopy not yet in full foliage in early summer or affected by fall vegetation coloring in late summer.

In late summer, the spectral difference between summergreen and evergreen is the highest and across the majority of features. The shortwave infrared, NDVI and NIR are crucial because they capture the spectral differences due to different water content, leaf structure, and chlorophyll absorption. Our case study reveals that high latitude boreal forest phenology in the summergreen–evergreen Siberian transition zone results in differences in pigment absorption and spectral properties from early to peak to late summer and is a reason for the difference in feature importance between the seasons.

In [28], for the Alaska and western Canada regions, the peak summer Landsat NIR is the most important feature in the spectral band dataset, for both peak summer and three-season RF models, with visible, NIR, as well as shortwave infrared. In the eastern Canadian region, this study [28] shows Landsat SWIR-1 as the most important feature in peak summer, and Landsat SWIR-2 for the three-season model agreeing well with the high feature performance of the long S-2 shortwave infrared in our study throughout all seasons. There are S-2 or Landsat-based Land

Cover classifications where the inclusion of NDVI does not improve the vegetation classification. However, this seems to be due to the choice of the season used in the classification studies. For example, Massay et al. [28] in their forest type classification of the North-American boreal domain use the peak summer where the forest vegetation is fully developed but spectrally not so highly differentiated. Similarly, in our study, the NDVI performance for the peak summer only is low. In the case of our study, the NDIV is of higher importance for the late summer when NDVI differentiation is high between summergreen and evergreen.

As we see in land cover classifications across forest types, and seasons and in our study, that the feature importance is widely spread across the visible, NIR, and shortwave infrared of the commonly used Landsat and Sentinel-2 satellite data for forest classification. We therefore suggest using all S-2 spectral features (without the 60 m resolution bands that support the atmospheric correction because they would introduce more noise in the data [16]) and also include the NDVI due to its power to separate between summergreen and evergreen in late summer.

Other studies that were successful in forest classification by including topography in addition to the spectral features, seem to be performed in regions that are characterized by high topographical differences. For example, Waśniewski et al. [23] assessed the potential of S-2 images and the RF classifier for mapping forest types in northwest Gabon, Africa, also including topography metrics derived from a digital elevation model (DEM). The study [23] investigated the impact of S-2 spectral bands, the NDVI and DEM on the accuracy of the classification of five subtropical forest classes, one being evergreen broadleaf forest.

The DEM was successfully included in [23] as the region covered lowland as well as mountain forest regions with an overall RF accuracy of forest type classification between 83.4 to 97.4% and the DEM derived features showing the biggest impact. Similarly, in the RF developed in Massay et al. [28] the North-American boreal domain is covering high contrasting topography with forests in lowland as well as mountainous regions high up to elevational treelines, therefore the DEM-derived features in [28] show high feature importance.

These conditions in other case studies involving high topographic variability are in contrast to our study in the summergreen-evergreen forest transition zone mainly representing a lowland region. Therefore, the addition of DEM-derived features in our case study seems to be not meaningful for mapping forest types and could introduce more signal noise. Waśniewski et al. [23] additionally show that the RF model for their forest classification can be successfully transferred from one primary satellite image to other satellite images. This is important for our study, because we aim to classify all S-2 images for the region in future studies with the method created in this study.

C. Performance of the Training and Test Datasets

The short snow-free time span of three to four months only in the high latitude Siberian region poses a challenge for optical vegetation-related remote sensing. This is one reason why [47] include SAR imagery in their circum-Arctic land cover classification of the tundra region North of the treeline. During the snow-free time window, cloud-free and smoke-free optical satellite image acquisitions are rare. Therefore, the previously published SiDroForest S-2 labeled image patch dataset [30] contains the best images of the summers from a wider time range of several summers from 2016 to 2020. As the forests do not change rapidly between years, with the exception of wildfire events, we can construct the S-2 training and test datasets for the different phenophases also from the neighboring years, as we consistently check for fires and other disturbances that would cause obvious and drastic changes in the land cover. For this study we need to specifically construct a large and not as diverse training dataset than the SiDroForest S-2 labeled image patch dataset [30], with labeled pixels for summergreen and evergreen respectively. The Bilibino region where we extract the summergreen training data is located in Chukotka in Siberia further to the east. Because the forest in Chukotka well represents "summergreen" with high biomass and crown coverage metrics [17] it is also representative for *Larix* stands in Central Yakutia. The in-situ forest plots [12], [13], [14] that are covered by the SiDroForest S-2 image patch dataset [30] are selected to have the most diverse samples of the region's forest vegetation, therefore including all successional forest states, also with low tree biomass and low crown coverage. For the creation of the SiDroTest dataset, we aggregate the SiDroForest S-2 image patch dataset [30] into summergreen and evergreen and further exclude forest plots of the open woodland type by the application of a minimum NDVI threshold anchored to the peak summer and re-assuring low NDVI for the late summer summergreen forest plots.

For the optimization of the SiDroTest dataset, we even slightly shift the 30 m by 30 m coverage for four out of the 11 forest plots guided by our expert field and thematic knowledge constructing an optimized SiDroTest dataset resulting in five evergreen and six summergreen plots. The SiDroTest dataset is however small with 1470 pixels for the evergreen and 1489 pixels for the summergreen classes but serves as a robust, independent test dataset.

Investigating the spectral surface reflectance and NDVI characteristics between the training and the test datasets, we see that the SiDroTest summergreen matches the summergreen training dataset very well, whereas the SiDroTest evergreen seems to represent a more diverse evergreen forest class with a lower NIR reflectance than the evergreen training dataset. When we inspect the correlation between the three seasons of the training dataset and the SiDroTest dataset, we see that some features are negatively correlated in the early summer and peak summer. Our contributions are prepared S-2 training datasets and test datasets designed with regional and field knowledge for the binary classification into summergreen and evergreen needleleaf forest categories, related in a way that they are representative for the occurring forest in the case study region and meaningful for the application.

D. Application in the Summergreen–Evergreen Forest Transition Zone

The Lake Khamra region lies nowadays within a part of the summergreen–evergreen forest transition zone where evergreen

is mixed with summergreen forest [48]. According to paleoclimate lake sediment records, evergreen needleleaf tree taxa expanded in this region into the summergreen dominated forests already since the Mid-Holocene along a south-west to north-east transect [48]. We see in our case study that for the current state of the summergreen-evergreen forest transition zone, S-2 satellite-based mapping can provide forest type distribution at a fine landscape scale. The classification of late summer S-2 surface reflectance data for the Lake Khamra region reveals distinct evergreen forest patches in a matrix of a summergreen to mixed evergreen-summergreen forest landscape. The three different summergreen-evergreen S-2 classifications of the Lake Khamra region show that important landscape characteristics such as the percentage of each mapped class, the heterogeneity, the degree of fragmentation, or the patch density of evergreen forest that provide the metrics to observe the structural change in time series (for example, [49]) varies. The S-2 SidroForest and the S-2 Collection-1 SA maps both show the evergreen forest patches south of Lake Khamra in a similar sharp outline and the S-2 SidroForest and the S-2 Collection-1 GR maps both show dominant evergreen in the northern part of the Lake Khamra region. In general, the S-2 mapped forest distribution displays high levels of landscape fragmentation, heterogeneity and complexity and a current state of a mixed summergreen and evergreen forest landscape.

While this case study demonstrates the effectiveness of the strategy to use a late-summer trained RF S-2-based classifier for summergreen-evergreen forest classification, limitations remain.

A crucial point in the transferability of mapping applications, specifically for change detection in multiannual satellite imagery is the radiometric consistency of the satellite reflectance data. Consistent reflectance time series are required for land surface monitoring [49]. Pixel-based temporal aggregation is a method to generate cloud-free, radiometrically and phenologically consistent, spatially contiguous image composites [50], [51]. It is beyond the scope of this case study to construct temporally aggregated, phenologically consistent annual Collection-1 S-2 time series for the late summer season to produce consistent forest maps over large areas. We consider this study as a case study for a benchmarked training dataset for the summergreen–evergreen forest transition zone. The RF trained with evergreen and summergreen training datasets robustly performs specifically for the late summer season with the strongest spectral contrast between summergreen and evergreen.

V. Conclusion

Our case study presents a benchmarked S-2 training dataset for mapping summergreen and evergreen needleleaf forests in the Siberian summergreen–evergreen forest transition zone using Random Forest classifiers on S-2 optical satellite data. This region is experiencing significant land cover changes due to climate change and increasing forest disturbances. Existing global land cover products, which distinguish between summergreen and evergreen forest types, are not quality controlled for this region and lack high spatial resolution. Although there are forest classification case studies using Landsat and S-2 data for other regions and continents, an in-depth analysis of seasonal variations and spectral features' importance for the summergreen–evergreen transition zone in Siberia has been lacking.

Our multitemporal S-2 training dataset representing summergreen and evergreen needleleaf forest classes contains nearly two million labeled pixels covering the early summer, peak summer, and late summer forest phenophases. The evaluation, using the SiDroTest dataset linked to reference forest plot data, demonstrated that applying our trained Random Forest classifier provides the highest classification accuracy at 99% for the late summer season. Feature importance analysis reveals that the long shortwave infrared is crucial across all seasons, likely due to water content differences in the plant canopies between the two forest types. The long shortwave infrared, the NDVI and the red band feature that is the major chlorophyll absorption band are important in the early summer datasets, the shortwave infrared and the blue band feature that represents also chlorophyll absorption are important in the peak summer, and the long shortwave infrared, NDVI and NIR are important in the late summer to distinguish between summergreen and evergreen.

The distinct spectral separability during the late summer phenophase between the two forest types underscores the value of focusing on late summer for optimal classification performance. This is particularly advantageous for the Siberian boreal forest region, where the late summer season offers a time window with less frequent wildfire activities, and cloud cover.

We plan in future work to apply this strategy to classify temporally aggregated late summer S-2 satellite data across the summergreen–evergreen forest transition zone, contributing to monitoring the changes in boreal forests under climate pressures. We are also contributing tools such as the multiseasonal S-2 training dataset for evergreen–summergreen needleleaf forest patch detection in the boreal forest domain.

Appendix

The Appendix presents details on the data characteristics of the input data: Tables II and III give an overview on the S-2 features and satellite acquisitions prepared for this study, Table IV shows the details on the shifted forest plots and Fig. 15 on characteristics of tree species that is the base for the classes "summergreen" and "evergreen" of the SiDroTest dataset. Tables V-12,3 and VI are showing the NDVI characteristics of the SiDroTest, Fig. 16(a)–(f) the surface reflectance of the training datasets and SiDroTest dataset. Fig. 17 shows the training dataset variances explained in the principal components, Fig. 18(a) heatmap visualization of the principal component loadings of the training dataset and Fig. 19 exemplarily shows seasonal S-2 subscenes of the spring and the autumn season with snow cover.

TABLE II
S-2 SPECTRAL FEATURES USED IN THIS STUDY (LEVEL-2A SURFACE REFLECTANCE AND NDVI AND ITS RADIOMETRIC AND SPATIAL CHARACTERISTICS

S-2 Spectral Feature	Central Wavelength (nm)	Original Pixel Length (m)
Band 2 - blue	490	10
Band 3 - green	560	10
Band 4 - red	665	10
Band 5 - red edge	705	20
Band 6 - red edge	740	20
Band 7 - red edge	783	20
Band 8 - near-infrared	842	10
Band 8A - near-infrared	865	20
Band 11 - shortwave infrared	1610	20
Band 12 - shortwave infrared	2190	20
NDVI (B8 - B4)/ (B8 + B4)	-	10

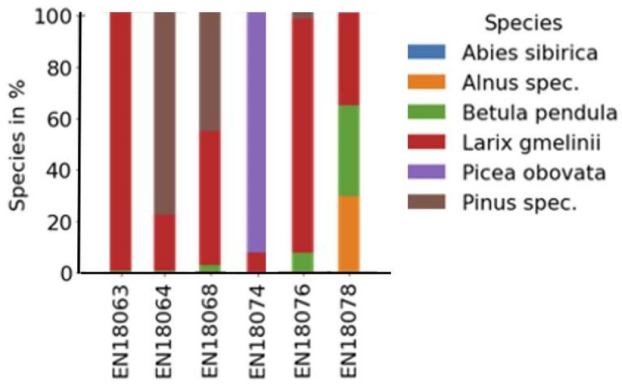

Fig. 15. Tree species coverage in percentage per plot of the vegetation plots in Yakutia (SiDroForest data collection) [14], [16] for labeling the S-2 image patches of the SiDroTest dataset used in this study with aggregation of the tree species into the two class assignments "summergreen" and "evergreen".

TABLE III
OVERVIEW ON S-2 ACQUISITION DATES AND SOURCES USED IN THIS STUDY

Site	S-2 Source	ES Date	PS Date	LS Date
S-2 AQUISITIONS -TRAINING DATASET				
Bilibino	SiDroForest	2020-05-21	2020-07-21	2020-09-13
Yakutsk	SiDroForest	2020-05-16	2020-07-23	2020-09-18
S-2 ACQUISITIONS- TEST DATASET				
Bilibino	SiDroForest	2020-05-21	2020-07-21	2020-09-13
Nyurba	SiDroForest	2020-05-09	2020-08-04	2020-10-11
Suntar	SiDroForest	2020-05-09	2020-07-23	2020-10-11
Mirny	SiDroForest	2020-05-22	2020-07-31	2020-09-19
Mirny-Lensk	SiDroForest	2020-06-01	2020-08-25	2020-09-19
Vilnyui	SiDroForest	2020-05-21	2020-08-04	2020-09-18
S-2 ACQUISITIONS- CLASSIFICATION SET				
Bilibino	SiDroForest	-	-	2020-09-13
Yakutsk	SiDroForest	-	-	2020-09-18
Lake Khamra	SiDroForest	-	-	2020-09-24
Lake Khamra	Copernicus GRID	-	-	2018-09-25
Lake Khamra	Collection-1 SAFE	-	-	2018-09-25

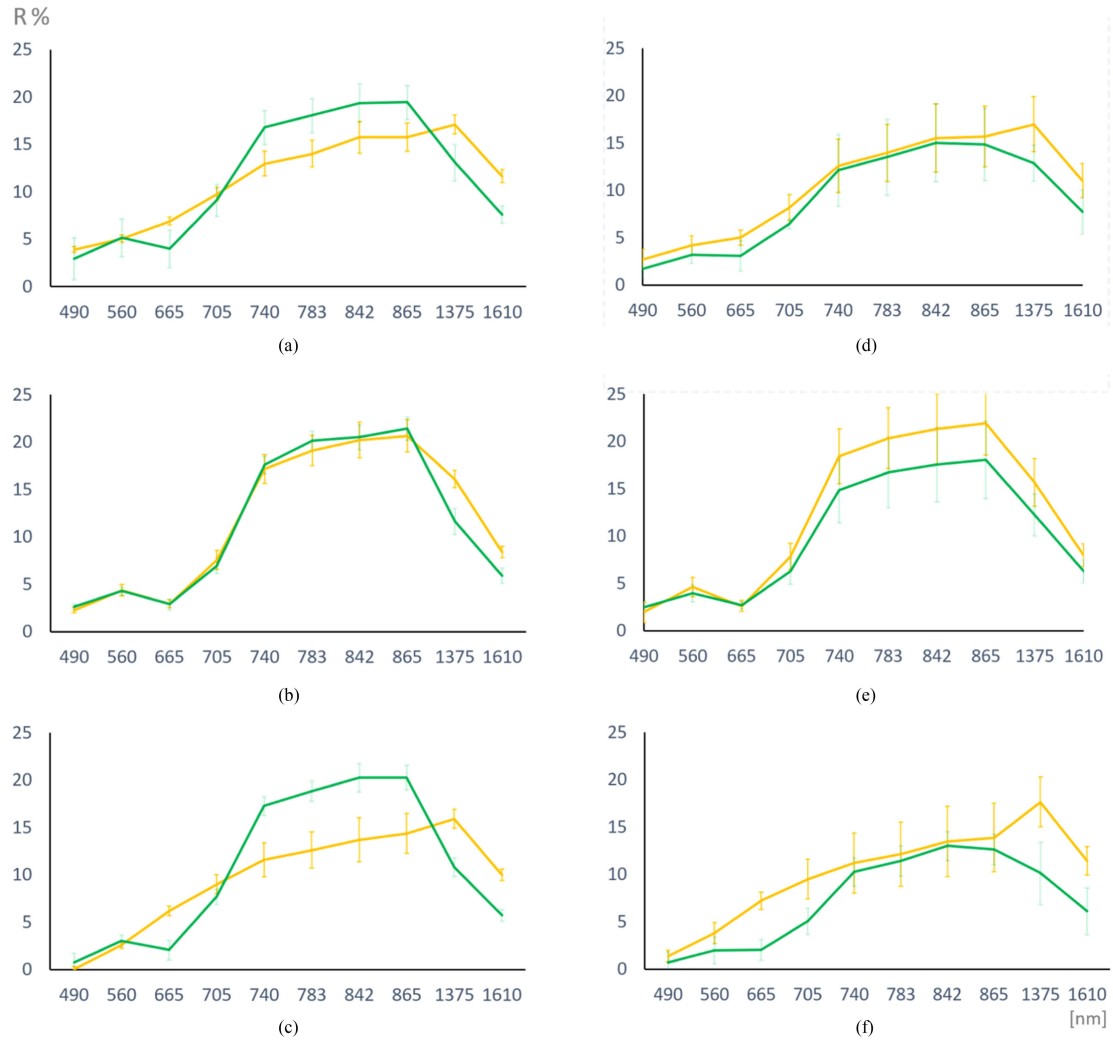

Fig. 16. (a)–(f). S-2 surface reflectance training dataset and SiDroTest dataset (with the standard deviation indicated as bars) at the center wavelengths of the 10 used S-2 bands (B 2, 3, 4, 5, 6, 7, 8, 8A, 11, 12) class *Summergreen* (orange), class *Evergreen* (dark green). (a)–(c) Training dataset: early summer, peak summer, late summer. (d)–(f) SiDroTest dataset: early summer, peak summer, late summer.

TABLE IV
PLOT CODES, SITES, DECIMAL GEOGRAPHIC COORDINATES (WGS84, NORTH; EAST) AND CLASS LABELS OF THE S-230 M X 30 M IMAGE PATCHES FROM THE SIDROFOREST DATA COLLECTION [14], [16] THAT WERE SHIFTED BASED ON EXPERT KNOWLEDGE

Forest Plot	Site	Adjusted Longitude	Adjusted Latitude	Label
EN18070 proxy	Nyurba	63.08291	117.98491	Evergreen
EN18073 proxy	Suntar	62.18776	117.40928	Summer green
EN18075 proxy	Mirny	62.68398	113.67319	Evergreen
EN18077 proxy	Mirny-Lensk	61.89186	114.28564	Evergreen

TABLE V
(4-1), (4-2), (4-3) NDVI MEAN, MEDIAN (MD) AND STANDARD DEVIATION (SD)) OF THE SIDROTEST DATASET: A4-1 EARLY SUMMER, A4-2 PEAK SUMMER, A4-3 LATE SUMMER

Forest Plot	Site	Label	Mean	MD	SD
EN18068	Nyurba	Evergreen	0.72	0.72	0.013
EN18070	Nyurba	Evergreen	0.69	0.72	0.074
EN18074	Suntar	Evergreen	0.61	0.61	0.018
EN18075	Mirny	Evergreen	0.58	0.58	0.011
EN18077	Mirny-Lensk	Evergreen	0.68	0.69	0.016
EN18030	Bilibino	Summer green	0.43	0.43	0.020
EN18035	Bilibino	Summer green	0.40	0.39	0.013
EN18063	Vilnuyi	Summer green	0.54	0.54	0.010
EN18073	Suntar	Summer green	0.30	0.30	0.006
EN18076	Mirny	Summer green	0.40	0.39	0.013
EN18078	Mirny-Lensk	Summer green	0.82	0.82	0.007

TABLE VI
TRAINING DATASET NDVI MEAN, MEDIAN (MD), AND STANDARD DEVIATION (SD) FOR EARLY SUMMER, PEAK SUMMER, AND LATE SUMMER

Forest Plot	Site	Label	Mean	MD	SD
EN18068	Nyurba	Evergreen	0.73	0.73	0.011
EN18070	Nyurba	Evergreen	0.69	0.69	0.020
EN18074	Suntar	Evergreen	0.74	0.76	0.048
EN18075	Mirny	Evergreen	0.76	0.76	0.025
EN18077	Mirny-Lensk	Evergreen	0.71	0.71	0.013
EN18030	Bilibino	Summer green	0.74	0.74	0.011
EN18035	Bilibino	Summer green	0.73	0.73	0.009
EN18063	Vilnuyi	Summer green	0.87	0.87	0.005
EN18073	Suntar	Summer green	0.79	0.79	0.009
EN18076	Mirny	Summer green	0.80	0.80	0.012
EN18078	Mirny-Lensk	Summer green	0.81	0.81	0.008

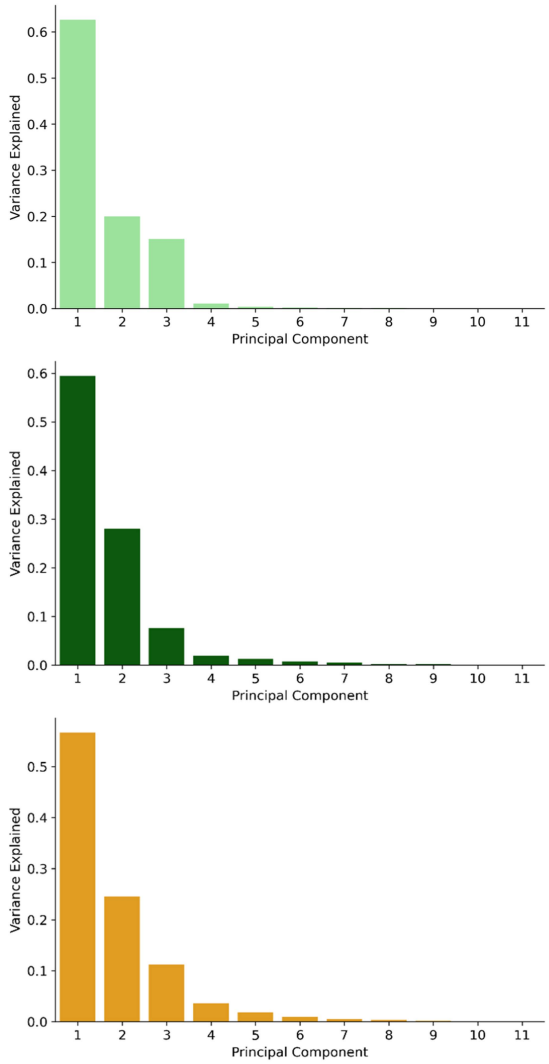

Fig. 17. (a)–(c) Training dataset variances explained in the principal components for early summer, peak summer, and late summer, in descending order, from top to the bottom. a) variances explained in early summer, b) peak summer and c) late summer.

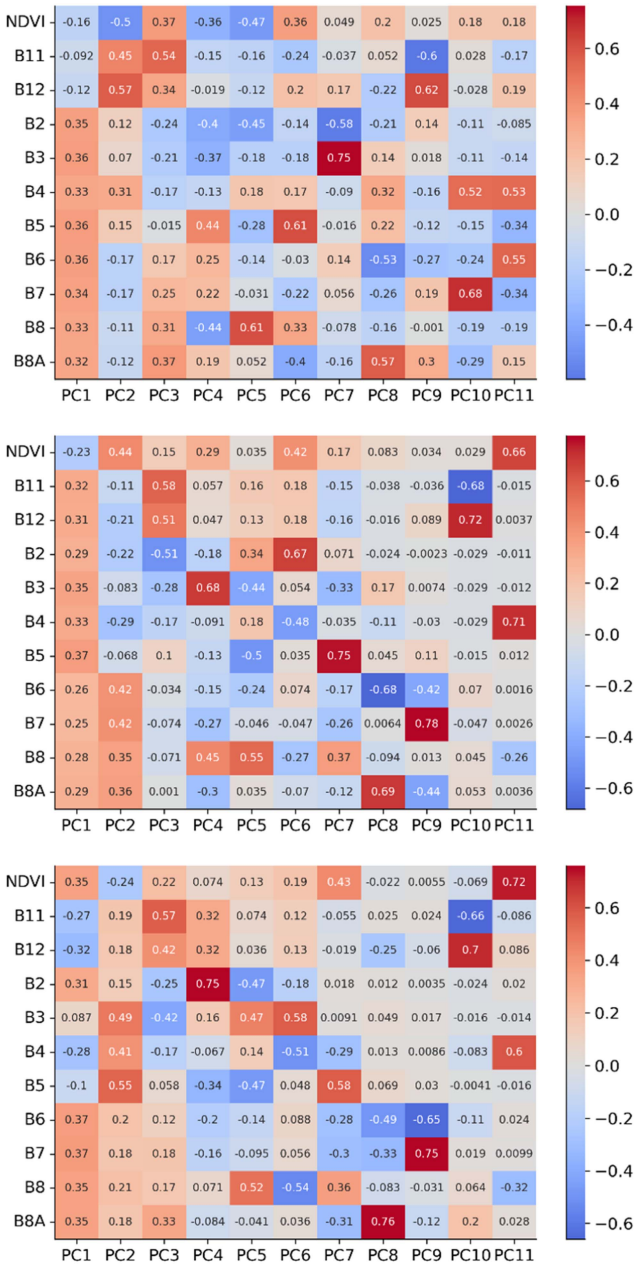

Fig. 18. (a)–(c) Heatmap visualization of the principal component loadings of the training dataset for early summer, peak summer, and late summer, in descending order, from top to the bottom.

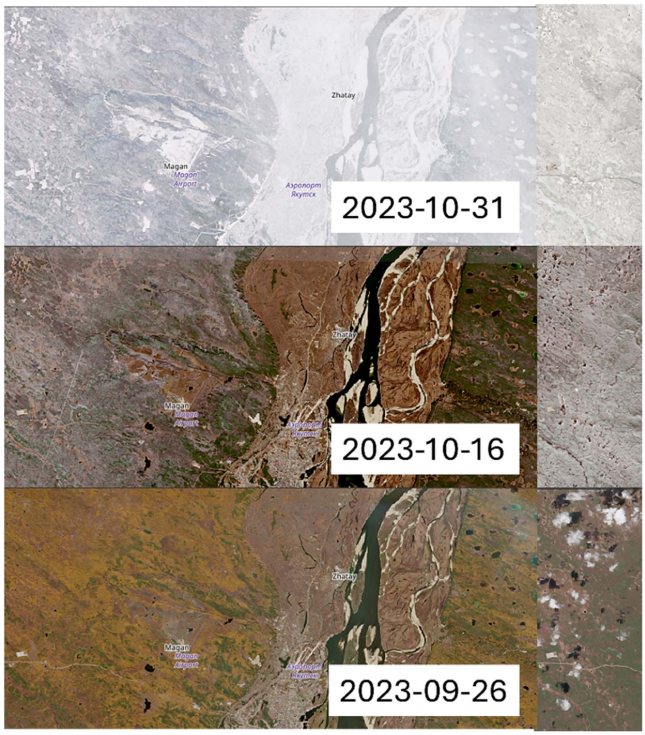

Fig. 19. Seasonal S-2 subscenes of the region Yakutsk (YA) (showing the Yakutsk city region in the Eastern part) displayed as quasi-true Red Green Blue (RGB) composites with red = B4, green = B3 and blue = B2 surface reflectance showing spring acquisitions with heterogeneous snow coverage in early May due to snow melting and a snow-free land surface in late May; and the late autumn situation with first heterogenous snow cover occurring during the month of October. Note how the late October S-2 subscene already shows low radiometric contrast due to the low sun zenith and low-incoming solar energy.

ACKNOWLEDGMENT

The authors would like to thank their Russian and German colleagues from the joint German–Russian expedition Chukotka 2018 for support in the field. The authors would also like to thank the time spent at the DLR in Munich with the HIDA exchange program and all the support.

REFERENCES

[1] G. B. Bonan, "Forests and climate change: Forcings, feedbacks, and the climate benefits of forests," *Science*, vol. 320, no. 5882, pp. 1444–1449, Jun. 2008, doi: 10.1126/science.1155121.

[2] M. M. Loranty et al., "Changing ecosystem influences on soil thermal regimes in northern high-latitude permafrost regions," *Biogeosciences*, vol. 15, no. 17, pp. 5287–5313, Sep. 2018, doi: 10.5194/bgd-15-5287-2018.

[3] S. D. Mamet et al., "Shifting global larix distributions: Northern expansion and southern retraction as species respond to changing climate," *J. Biogeography*, vol. 46, no. 1, pp. 134–145, Jan. 2019, doi: 10.1111/jbi.13465.

[4] U. Herzschuh, "Legacy of the last glacial on the present-day distribution of deciduous versus evergreen boreal forest," *Glob. Ecol. Biogeography*, vol. 29, no. 2, pp. 198–206, Feb. 2020, doi: 10.1111/geb.13018.

[5] U. Herzschuh et al., "Glacial legacies on interglacial vegetation at the pliocene-pleistocene transition in NE Asia," *Nat. Commun.*, vol. 7, Jul. 2016, Art. no. 11967, doi: 10.1038/ncomms11967.

[6] S. Kruse et al., "Novel coupled permafrost–forest model (LAVESI–CryoGrid v1.0) revealing the interplay between permafrost, vegetation, and climate across eastern Siberia," *Geoscientific Model Develop.*, vol. 15, no. 6, pp. 2395–2422, Jun. 2022, doi: 10.5194/gmd-15-2395-2022.

[7] S. N. Kirpotin et al., "Impacts of environmental change on biodiversity and vegetation dynamics in Siberia," *Ambio*, vol. 50, no. 11, pp. 1926–1952, Nov. 2021, doi: 10.1007/s13280-021-01619-3.

[8] L. Enguehard et al., "Unraveling boreal forest composition and drivers across scales in eastern Siberia," *Environ. Res. Lett.*, vol. 19, no. 7, Jun. 2024, Art. no. 074050, doi: 10.1088/1748-9326/ad5742.

[9] ESA Land Cover CCI Project Team and P. Defourny, "ESA land cover climate change initiative (Land_Cover_cci): Global land cover maps, version 2.0.7," Centre for Environmental Data Analysis, 2019, Accessed: Feb. 2024. [Online]. Available: https://climate.esa.int/en/projects/land-cover/

[10] M. Buchhorn et al., "Copernicus global land service: Land cover 100 m: Collection 3: Epoch 2019: Globe 2020," Zenodo, Geneva, Switzerland, 2020, doi: 10.5281/zenodo.3939050.

[11] T. Miesner et al., "Forest structure and individual tree inventories of north-eastern Siberia along climatic gradients," *Earth Syst. Sci. Data*, vol. 14, no. 12, pp. 5695–5716, Dec. 2022, doi: 10.5194/essd-14-5695-2022.

[12] T. Miesner et al., "Tree dataset from forest inventories in North-Eastern Siberia," *PANGAEA*, 2022, doi: 10.1594/PANGAEA.943547.

[13] S. Kruse et al., "Sampling mixed species boreal forests affected by disturbances and mountain lake and alas lake coring in Central Yakutia," in *Russian-German Cooperation: Expeditions to Siberia in 2018*, S. Kruse et al., Eds., Bremerhaven, Germany: Alfred Wegener Inst. Polar Mar. Res., 2019, pp. 148–153.

[14] S. Kruse et al., "Forest inventories on circular plots on the expedition Chukotka 2018, NE Russia," *PANGAEA*, 2020, doi: 10.1594/PANGAEA.923638.

[15] I. Shevtsova et al., "Foliage projective cover of 57 vegetation sites of central Chukotka from 2016," *PANGAEA*, 2019, doi: 10.1594/PANGAEA.908570.

[16] F. van Geffen et al., "SiDroForest: A comprehensive forest inventory of Siberian boreal forest investigations including drone-based point clouds, individually labeled trees, synthetically generated tree crowns, and Sentinel-2 labeled image patches," *Earth Syst. Sci. Data*, vol. 14, no. 11, pp. 4967–4994, 2022, doi: 10.5194/essd-14-4967-2022.

[17] F. van Geffen et al., "SiDroForest: Individual-labelled trees acquired during the fieldwork expeditions that took place in 2018 in Central Yakutia and Chukotka, Siberia," *PANGAEA*, 2021, doi: 10.1594/PANGAEA.932821.

[18] D. Zanaga et al., "ESA WorldCover 10 m 2020 v100," Zenodo, Geneva, Switzerland, 2022, doi: 10.5281/zenodo.5571936.

[19] D. Zanaga et al., "ESA WorldCover 10 m 2021 v200," Zenodo, Geneva, Switzerland, 2021, doi: 10.5281/zenodo.7254221.

[20] A. Elmes et al., "Accounting for training data error in machine learning applied to earth observations," *Remote Sens.*, vol. 12, Mar. 2020, Art. no. 1034, doi: 10.3390/rs12061034.

[21] I. Reis, D. Baron, and S. Shahaf, "Probabilistic random forest: A machine learning algorithm for noisy data sets," *Astron. J.*, vol. 157, 2018, Art. no. 16.

[22] M. Pal, "Random forest classifier for remote sensing classification," *Int. J. Remote Sens.*, vol. 26, pp. 217–222, Jan. 2005.

[23] A. Waśniewski et al., "Assessment of Sentinel-2 satellite images and random forest classifier for rainforest mapping in Gabon," *Forests*, vol. 11, no. 9, Sep. 2020, Art. no. 941.

[24] M. Ma et al., "Tree species classification based on Sentinel-2 imagery and random forest classifier in the eastern regions of the Qilian Mountains," *Forests*, vol. 12, no. 12, Dec. 2021, Art. no. 1736.

[25] P. Phan Noi and M. Kappas, "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery," *Sensors*, vol. 18, no. 1, Jan. 2017, Art. no. 18.

[26] C. Bolyn et al., "Forest mapping and species composition using supervised per pixel classification of Sentinel-2 imagery," *Biotechnologie, Agronomie, Société Environnement*, vol. 22, pp. 172–187, Mar. 2018.

[27] M. Wessel, M. Brandmeier, and D. Tiede, "Evaluation of different machine learning algorithms for scalable classification of tree types and tree species based on Sentinel-2 data," *Remote Sens.*, vol. 10, Jul. 2018, Art. no. 1419, doi: 10.3390/rs10091419.

[28] R. Massey et al., "Forest composition change and biophysical climate feedbacks across boreal North America," *Nat. Climate Change*, vol. 13, no. 12, pp. 1368–1375, Dec. 2023, doi: 10.1038/s41558-023-01851-w.

[29] D. Schepaschenko et al., "The forest observation system: Building a global reference dataset for remote sensing of forest biomass," *Sci. Data*, vol. 6, Jan. 2019, Art. no. 198, doi: 10.1038/s41597-019-0196-1.

[30] F. van Geffen et al., "SiDroForest: Sentinel-2 level-2 bottom of atmosphere labeled image patches with seasonal information for Central Yakutia and Chukotka vegetation plots (Siberia, Russia)," *PANGAEA*, 2021, doi: 10.1594/PANGAEA.933268.
[31] S. Ahlswede et al., "TreeSatAI benchmark archive: A multi-sensor, multi-label dataset for tree species classification in remote sensing," *Earth Syst. Sci. Data*, vol. 15, pp. 681–695, Mar. 2023, doi: 10.5194/essd-15-681-2023.
[32] C. E. Woodcock et al., "Transitioning from change detection to monitoring with remote sensing: A paradigm shift," *Remote Sens. Environ.*, vol. 238, Mar. 2020, Art. no. 111558, doi: 10.1016/j.rse.2019.111558.
[33] S. E. Martiniano de Oliveira et al., "Reducing the effects of vegetation phenology on change detection in tropical seasonal biomes," *GIScience Remote Sens.*, vol. 56, no. 5, pp. 699–717, Jul. 2018, doi: 10.1080/15481603.2018.1550245.
[34] L. Pan et al., "Mapping winter crops using a phenology algorithm time-series Sentinel-2 and Landsat-7/8 images and Google Earth Engine," *Remote Sens.*, vol. 13, no. 13, Jun. 2021, Art. no. 2510, doi: 10.3390/rs13132510.
[35] Y. Chabalala et al., "Mapping fruit tree dynamics using phenological metrics from optimal Sentinel-2 data and deep neural network," *CABI Agriculture Biosci.*, vol. 4, Nov. 2023, Art. no. 51, doi: 10.1186/s43170-023-00193-z.
[36] M. Li et al., "Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington wildlife forest," *GIScience Remote Sens.*, vol. 50, no. 4, pp. 361–384, 2013, doi: 10.1080/15481603.2013.819161.
[37] E. Grabska et al., "Forest stand species mapping using the Sentinel-2 time series," *Remote Sens.*, vol. 11, Nov. 2019, Art. no. 1197.
[38] M. Immitzer et al., "First experience with Sentinel-2 data for crop and tree species classifications in Central Europe," *Remote Sens.*, vol. 8,, 2016, Art. no. 166.
[39] J. R. Schriever and R. G. Congalton, "Evaluating seasonal variability as an aid to cover-type mapping from Landsat Thematic Mapper data in the northeast," *Photogrammetry Eng. Remote Sens.*, vol. 3, pp. 321–327, 2005.
[40] J. G. Mickelson et al., "Delineating forest canopy species in the Northeastern United States using multi-temporal TM imagery," *Amer. Soc. Photogrammetry Remote Sens.*, vol. 64, pp. 891–904, 1998.
[41] R. Geng et al., "Modern pollen assemblages from lake sediments and soil in East Siberia and relative pollen productivity estimates for major taxa," *Front. Ecol. Evol.*, vol. 10, 2022, Art. no. 837857, doi: 10.3389/fevo.2022.837857.
[42] S. Lev-Yadun, "The phenomenon of red and yellow autumn leaves: Hypotheses, agreements and disagreements," *J. Evol. Biol.*, vol. 35, no. 10, pp. 1245–1128, Oct. 2022, doi: 10.1111/jeb.14069.
[43] Y. Luo et al., "Evergreen greenness and its relationship with leaf flushing, aging, and water fluxes," *Agricultural Forest Meteorol.*, vol. 323, 2022, Art. no. 109060, doi: 10.1016/j.agrformet.2022.109060.
[44] P. F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *JMLR*, vol. 12, pp. 2825–2830, 2011, doi: 10.48550/.arXiv.1201.0490.
[45] M. Persson, E. Lindberg, and H. Reese, "Tree species classification with multi-temporal Sentinel-2 data," *Remote Sens.*, vol. 10, 2018, Art. no. 1794, doi: 10.3390/rs10111794.
[46] L. Enguehard et al., "Classified maps of forest types in Eastern Siberia based on field surveys and Sentinel-2 imagery," *PANGAEA*, 2024, doi: 10.1594/PANGAEA.967133 </Dataset>.
[47] A. Bartsch et al., "Circumarctic land cover diversity considering wetness gradients," *Hydrol. Earth Syst. Sci.*, vol. 28, no. 11, pp. 2421–2481, 2024, doi: 10.5194/hess-28-2421-2024.
[48] I. Baisheva et al., "Late glacial and holocene vegetation and lake changes in SW Yakutia, Siberia, inferred from sedaDNA, pollen, and XRF data," *Front. Earth Sci.*, vol. 12, 2024, Art. no. 1354284, doi: 10.3389/feart.2024.1354284.
[49] M. A. Syariz et al., "Spectral-consistent relative radiometric normalization for multitemporal Landsat 8 imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 147, pp. 56–64, 2019.
[50] T. Phan, M. Kappas, and T. Koellner, "Assessment of temporal aggregation of Sentinel-2 images on seasonal land cover classification using machine learning," *Environ. Monit. Assessment*, vol. 192, p. 142, 2020, doi: 10.1007/s10661-024-13596-w.
[51] H. J. White et al., "Ecosystem stability at the landscape scale is primarily associated with climatic history," *Funct. Ecol.*, vol. 36, no. 3, pp. 622–634, 2022, doi: 10.1111/1365-2435.13957.

Femke van Geffen received the M.Sc. degree in data science from Tilburg University, Tilburg, the Netherlands, in 2018. She is currently working towards the Ph.D. degree with the University of Potsdam, Potsdam, Germany, and the Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI) as a part of the Helmholtz Einstein International Berlin Research School in Data Science (HEIBRiDS: https://www.heibrids.berlin/).

The HEIBRiDS research project is a collaboration between the AWI Research Unit Potsdam, and the Microwave and Radar Institute of the German Aerospace Center (DLR), Munich, Germany, facilitated by the Helmholtz Information & Data Science Academy (HIDA) grant. She currently works with the Hague University of Applied Sciences, Hague, the Netherlands, as a Researcher in image analysis and artificial intelligence.

Ronny Hänsch (Senior Member, IEEE) received the Diploma degree in computer science and the Ph.D. degree in computer science from TU Berlin, Berlin, Germany, in 2007 and 2014, respectively.

He is a Scientist with the Microwave and Radar Institute, German Aerospace Center (DLR), Germany, where he leads the Machine Learning Team in the Signal Processing Group of the SAR Technology Department. He served as the Chair of the GRSS Image Analysis and Data Fusion (IADF) technical committee and editor of the GRSS eNewsletter.

Dr. Hänsch is currently the Co-Chair of the ISPRS working group on Image Orientation and Sensor Fusion, an Editor in Chief of the *Geoscience and Remote Sensing Letters*, an Associate Editor of the *ISPRS Journal of Photogrammetry and Remote Sensing*, a Co-Organizer of the CVPR Workshop EarthVision (2017–2025), and the IGARSS Tutorial on Machine Learning in Remote Sensing (2017–2024). He has extensive experience in organizing remote sensing community competitions, serves as the GRSS Representative within SpaceNet and was the Technical Lead of the SpaceNet 8 and 9 Challenges.

Begüm Demir (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electronic and telecommunication engineering from Kocaeli University, Kocaeli, Türkiye, in 2005, 2007, and 2010, respectively.

She is currently a Full Professor and the Founder Head of the Remote Sensing Image Analysis (RSiM) group, Faculty of Electrical Engineering and Computer Science, TU Berlin, Berlin, Germany, and the Head of the Big Data Analytics for Earth Observation research group at the Berlin Institute for the Foundations of Learning and Data (BIFOLD), Berlin. Her research interests include the intersection of machine learning, remote sensing and signal processing. Specifically, she performs research in the field of processing and analysis of large-scale Earth observation data acquired by airborne and satellite-borne systems.

Dr. Demir was the recipient of the prestigious "2018 Early Career Award" by IEEE Geoscience and Remote Sensing Society for her research contributions in machine learning for information retrieval in remote sensing. In 2018, she received a Starting Grant from the European Research Council (ERC) for her project "BigEarth: Accurate and Scalable Processing of Big Data in Earth Observation". She is a Fellow of European Lab for Learning and Intelligent Systems (ELLIS). She is a Scientific Committee member of several international conferences and workshops. She is a referee for several journals such as the PROCEEDINGS OF IEEE, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE TRANSACTIONS ON IMAGE PROCESSING, *Pattern Recognition*, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, the *International Journal of Remote Sensing*, and several international conferences. She is currently an Associate Editor for IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE.

Stefan Kruse received the M.Sc. and Ph.D. degrees in geoecology from Potsdam University, Potsdam, Germany, in 2013 and 2017, respectively.

Dr. Kruse has been the Group Lead of the Working group High-latitudinal vegetation changes in the Polar Terrestrial Environmental Systems Research Section, AWI Potsdam, Potsdam, since 2017. He has long-term experiences in fieldwork at high-latitudes and extensive knowledge in the boreal forest system of Siberia and North America. His aim is reducing uncertainties arising from nonlinear relationships and time lags in boreal systems, with a particular emphasis on their implications in the context of global warming. To this end, he employs a multidisciplinary approach, also optimizing fieldwork at all spatial and temporal scales. He collects forest structure across all scales using remote sensing from UAV remote sensing up to the spatial scale of operational satellite remote sensing (optical and SAR). His research interests include latitudinal treeline dynamics, combining knowledge from population genetics and calibrating the individual-based model LAVESI that he had developed himself using multiscale remote sensing.

Ulrike Herzschuh received the Diploma degree in biology and the Ph.D. degree in biology from Potsdam University, Potsdam, Germany, in 1999 and 2005, respectively.

She is a Full Professor on "Statistical Integration of Palaeoenvironmental Data" with the Institute for Environmental Science and Geography and with the Institute for Biochemistry and Biology, University Potsdam, Potsdam, and the Head of the Polar Terrestrial Environmental Systems Research Section, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI), Research Section Potsdam, Germany. She is also the Head of the AWI high-latitude Biodiversity working group with a research focus on environmental data sciences and data science applications. She is serving as long-term Member of the International Arctic Science Committee (IASC) Terrestrial Working Group (TWG). She has long-term experiences in fieldwork at high latitudes and extensive knowledge in the boreal forest system of Siberia. In 2018, she received a Starting Grant from the European Research Council (ERC) for her project "Glacial Legacy on the establishment of evergreen vs. summergreen boreal forests' in Siberia." Her main research interests include ecosystem changes, ancient DNA and pollen analysis, and Arctic and boreal vegetation dynamics.

Dr. Herzschuh was awarded the Gottfried Wilhelm Leibniz-Preis in 2024.

Birgit Heim received the Diploma degree in geology and paleontology from the Free University of Berlin, Berlin, Germany, in 2001, and the Ph.D. degree in geosciences from Potsdam University, Potsdam, Germany, in 2006.

She visited a range of research opportunities in the field of terrestrial and aquatic spectral remote sensing at German research institutions and is a Senior Scientist with the Polar Terrestrial Environmental Systems Research Section, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI), providing remote sensing and data science support. She has long-term experiences in fieldwork at the high-latitudes and extensive knowledge in the boreal forest system of Siberia and North America. She has extensive experience in spectral optical remote sensing, spectral characteristics of high-latitude landscapes, and multiscale working on interdisciplinary level. She was involved in integrating satellite remote sensing products in the national Helmholtz Regional Climate Change and Humans (REKLIM) program. She is serving as Topical Editor of *Earth System Science Data* (ESSD), an international, interdisciplinary journal for the publication of articles on original research data.