Mapping dominant leaf type based on combined Sentinel-1/-2 data – Challenges for mountainous countries

Waser, Lars T; Rüetschi, Marius; Psomas, Achilleas; Small, David; Rehush, Nataliia

Abstract: Countrywide winter and summer Sentinel-1 (S1) backscatter data, cloud-free summer Sentinel-2 (S2) images, an Airborne Laser Scanning (ALS)-based Digital Terrain Model (DTM) and a forest mask were used to model and subsequently map Dominant Leaf Type (DLT) with the thematic classes broadleaved and coniferous trees for the whole of Switzerland. A novel workflow was developed that is robust, cost-efficient and highly automated using reference data from aerial image interpretation. Two machine learning approaches based on Random Forest (RF) and deep learning (UNET) for the whole country with three sets of predictor variables were applied. 24 subareas based on aspect and slope categories were applied to explore effects of the complex mountainous topography on model performances. The reference data split into training, validation and test data sets was spatially stratified using a 25 km regular grid. Model accuracies of both RF and UNET were generally highest with Kappa (K) around 0.95 when predictors were included from both S1/S2 and the topographic variables aspect, elevation and slope from the DTM. While only slightly lower accuracies were obtained when using S2 and DTM data, lowest accuracies were obtained when only predictors from S1 and DTM were included, with RF performing worse than UNET. While on countrywide level RF and UNET performed overall similarly, substantial differences in model performances, i.e. higher variances and lower accuracies, were found in subareas with northwest to northeast orientations. The combined use of S1/S2 and DTM predictors mitigated these problems related to topography and shadows and was therefore superior to the single use of S1 and DTM or S2 and DTM data. The comparison with independent National Forest Inventory (NFI) plot data demonstrated precisions of K around 0.6 in the predictions of DLT and indicated a trend of increasing deviations in mixed forests. A comparison with the Copernicus High Resolution Layer (HRL) DLT 2018 revealed overall higher map accuracies with the exception of pure broadleaved forest. Although, spatial patterns of DTL were overall similar, UNET performed better than RF in areas with a distinct DLT on forest stand level, with the largest differences occurring when only S1 and DTM data was used. In contrast, predictions obtained from RF were more accurate in mixed stands. This study goes beyond the case study level and meets the requirements of countrywide data sets, in particular regarding repeatability, updating, costs and characteristics of training data sets. The 10 m countrywide DLT maps add complementary and spatially explicit information to the existing NFI estimates and are thus highly relevant for forestry practice and other related fields.

DOI: https://doi.org/10.1016/j.isprsjprs.2021.08.017

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-209383
Journal Article
Published Version
The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:
Waser, Lars T; Rüetschi, Marius; Psomas, Achilleas; Small, David; Rehush, Natalia (2021). Mapping dominant leaf type based on combined Sentinel-1/-2 data – Challenges for mountainous countries. ISPRS Journal of Photogrammetry and Remote Sensing, 180:209-226.
DOI: https://doi.org/10.1016/j.isprsjprs.2021.08.017
Mapping dominant leaf type based on combined Sentinel-1/-2 data – Challenges for mountainous countries

Lars T. Waser a,*, Marius Rüetschi a, Achilleas Psomas a, David Small b, Nataliia Rehush c

a Department of Land Change Science, Swiss National Forest Inventory, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland
b Remote Sensing Laboratories (RSL), University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
c Department of Forest Resources and Management, Scientific Service NFI, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

ARTICLE INFO

Keywords: Broadleaved Coniferous Deep learning Forestry practice Random forest Wall-to-wall National forest inventory

ABSTRACT

Countrywide winter and summer Sentinel-1 (S1) backscatter data, cloud-free summer Sentinel-2 (S2) images, an Airborne Laser Scanning (ALS)-based Digital Terrain Model (DTM) and a forest mask were used to model and subsequently map Dominant Leaf Type (DLT) with the thematic classes broadleaved and coniferous trees for the whole of Switzerland. A novel workflow was developed that is robust, cost-efficient and highly automated using reference data from aerial image interpretation. Two machine learning approaches based on Random Forest (RF) and deep learning (UNET) for the whole country with three sets of predictor variables were applied. 24 subareas based on aspect and slope categories were applied to explore effects of the complex mountainous topography on model performances. The reference data split into training, validation and test data sets was spatially stratified using a 25 km regular grid. Model accuracies of both RF and UNET were generally highest with Kappa (K) around 0.95 when predictors were included from both S1/S2 and the topographic variables aspect, elevation and slope from the DTM. While only slightly lower accuracies were obtained when using S2 and DTM data, lowest accuracies were obtained when only predictors from S1 and DTM were included, with RF performing worse than UNET. While on countrywide level RF and UNET performed overall similarly, substantial differences in model performances, i.e. higher variances and lower accuracies, were found in subareas with northwest to northeast orientations. The combined use of S1/S2 and DTM predictors mitigated these problems related to topography and shadows and was therefore superior to the single use of S1 and DTM or S2 and DTM data. The comparison with independent National Forest Inventory (NFI) plot data demonstrated precisions of K around 0.6 in the predictions of DLT and indicated a trend of increasing deviations in mixed forests. A comparison with the Copernicus High Resolution Layer (HRL) DLT 2018 revealed overall higher map accuracies with the exception of pure broadleaved forest. Although, spatial patterns of DLT were overall similar, UNET performed better than RF in areas with a distinct DLT on forest stand level, with the largest differences occurring when only S1 and DTM data was used. In contrast, predictions obtained from RF were more accurate in mixed stands. This study goes beyond the case study level and meets the requirements of countrywide data sets, in particular regarding repeatability, updating, costs and characteristics of training data sets. The 10 m countrywide DLT maps add complementary and spatially explicit information to the existing NFI estimates and are thus highly relevant for forestry practice and other related fields.

1. Introduction

Large-area wall-to-wall maps on forest composition, as well as changes over time, are frequently required in the forestry sector and beyond. Such spatially explicit maps are indispensable for countrywide sustainable forest management strategies (MacDicken et al., 2015; Satir et al., 2017), biodiversity (Vihervaa et al., 2017) and biomass (Price et al., 2020) assessments, and appraisal of forest conditions – particularly in a changing climate (Keenan, 2015). Although National Forest Inventories (NFIs) provide a precise statistical estimation of the forest

* Corresponding author.
E-mail address: lars.waser@wsl.ch (L.T. Waser).

https://doi.org/10.1016/j.isprsjprs.2021.08.017
Received 16 February 2021; Received in revised form 26 July 2021; Accepted 16 August 2021
Available online 27 August 2021
0924-2716/© 2021 The Author(s). Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
composition for whole countries based on stratified samples, spatially explicit information beyond inventory domains is still lacking (McRoberts and Tomppo, 2007; Barrett et al., 2016; White et al., 2016; Ginzler et al., 2019). It is well established that several forest parameters used in NFIs, i.e. forest cover, forest type and composition, biomass etc., are described and quantified better with area-wide spatial data than with field plot data (Barrett et al., 2016; Ginzler et al., 2019). Thus, maps or remotely sensed products could help to improve estimates for forest inventories (McRoberts, 2010; Vidal et al., 2016).

Despite great advances in remote sensing data and technologies to an operational level over the last 40 years, spatially explicit maps of DLT and species composition with a high spatial resolution, regular update and robust verification on county-level, remain rare and were mainly conducted in individual case studies (Fassnacht et al., 2016; Waser et al., 2017). Only recently, with the Copernicus High Resolution Layers (HRL) (EEA, 2018; 2021) products a trend reversal was introduced in Europe. According to Ginzler et al. (2019), the lack of area-wide spatially explicit information was the main driver to develop countrywide products in the framework of the Swiss NFI over the last years. However, up to now, only few studies exist that have covered large geographical extents, such as forest attribute maps with a national extent, e.g. of Denmark (Nord-Larsen and Schumacher, 2012), Sweden (Nilsson et al., 2017), Switzerland (Waser et al., 2017) and Norway (Breidenbach et al., 2021). A relevant limitation of existing maps of the two classes broadleaved and coniferous based on remote sensing techniques and auxiliary data sets is that they are often not current due to the lack of input data and large realization time. For Switzerland, these maps include the European Forest Type (FTY) 2015 20-m-resolution map (EEA, 2018), the forest cover and type map from the European Commission’s Joint Research Centre (JRC) (Rempeneers et al., 2011) and the Copernicus HRL Forest and Copernicus DLT 2018 10-m-resolution maps (EEA, 2021). The latter is more promising since its update is planned to be on a three-year basis. However, in general, full suitability of all these maps is often limited because countrywide assessments are particularly challenging in areas with patchy forests and complex topography, such as mountainous terrain. For example, Waser et al. (2017) showed that the distinction between the classes broadleaved and coniferous at national level remains a challenge in mountainous countries due to shadows and frequent cloud coverage. Another limitation is the large heterogeneity of available training and validation data sources (e.g. NFIs, countrywide monitoring programs and individually collected data from case studies), their differences in quality and their often uneven distribution across classes (Japkowicz and Stephen, 2002; Mellor et al., 2015). Until now, only very few studies (Waser et al., 2017, Breidenbach et al. 2021) have successfully addressed these challenging conditions at the country level.

Nowadays, existing countrywide campaigns devoted to acquiring remote sensing data (mostly aerial images and ALS) are not always carried out during the vegetation period because their goal is to update topographic maps or terrain models rather than obtain information on canopy cover. Unprecedented opportunities for countrywide mapping approaches arise by combining Copernicus Sentinel-1/-2 data (hereafter S1 and S2) from the European Union and European Space Agency (ESA) (ESA, 2021). They are freely available and operating in twin mode enables to acquire high temporal (~5 days or better) and high spatial (10–20 m) resolution data sets across a variety of wavelengths. A number of studies in Central and Northern Europe used S2 time series to classify broadleaved and coniferous trees (e.g. Wessel et al., 2018; Immitzer et al., 2019) or tree species (e.g. Immitzer et al., 2016; Karasiak...
et al., 2017; Persson et al., 2018; Puletteri et al. 2018; Grabksa et al., 2019; Grabksa et al., 2020, Breidenbach et al., 2021). With the exception of Breidenbach et al. (2021) all these studies investigated relatively small areas and achieved high overall accuracies.

While the use of S2 images is often limited for mountainous areas owing to restrictions related to topography, i.e. shadows and atmospheric conditions – in particular convective cloud cover, Rüetschi et al. (2018) showed that S1 Synthetic Aperture Radar (SAR) C-band data is very promising in this regard. They used opposing seasonal backscatter behavior to distinguish between the classes broadleaved and coniferous in Switzerland. SAR data with its independence of weather conditions and daytime has been frequently used to assess forest attributes in all parts the world (e.g. Reiche et al., 2016; Dostalova et al., 2021). The well recognized synergetic use of optical and radar data for assessing forest attributes (Puhl et al., 2020) and more specific for classifying tree types and species (Passmaclt et al., 2016) is recently also found in the combined use of S1 and S2 data. For example, Erinjery et al. (2018) combined S1 and S2 data to classify vegetation types in tropical rain forests, Mngadi et al. (2019) to examine its effectiveness for commercial tree species mapping in South Africa, and Campos-Taberner et al. (2019) to classify shrub and forest types in Spain. In all these studies, it was concluded that S1 complements the capabilities of optical S2 data by contributing data acquired in the long-wavelength microwave domain.

A large body of research has been devoted to a variety of machine learning methods in environmental remote sensing for mapping as well as modeling purposes. RF, as introduced by Breiman (2001), is a popular and often preferred classifier because of its high performance, its robustness to overfitting and nonlinearities of predictors, its ability to handle high-dimensional data, and its computational efficiency (Rush et al., 2018). While RF requires the hand-engineered predictor extraction (e.g. vegetation indices, textural variables) (Belgium and Driga, 2016) Deep learning (DL) approaches “learn” directly from the input data to the target prediction (LeCun et al., 2015; Schmidhuber, 2015). During the last decade, DL techniques have been increasingly applied for image segmentation and classification, as well as object detection in various research areas. Since being proposed by Hinton et al. (2006), there has been increasing interest in using deep learning in earth observations (Zhu et al., 2017; Rammer and Seidl, 2019; Ma et al., 2019; Hoeser and Kuenzer, 2020), and lately also for classifying tree species (e.g. Fricker et al., 2019; Hartling et al., 2019; Gao and Zhang, 2020; Nezami et al., 2020). Nevertheless, some expertise is still required to set up a deep network architecture and tune the optimization hyperparameters (Hoeser and Kuenzer, 2020).

The main goal of the present study was to map the Dominant Leaf Type (DLT) with the thematic classes broadleaved and coniferous trees on a national extent based on freely available S1 and S2 data and a countrywide DTM. The following main points were addressed: (i) to investigate the effectiveness of using freely available remote sensing data and software, along with existing reference data, to map DLT at 10 m spatial resolution across Switzerland, (ii) to evaluate the potential of single use of S1 and DTM or S2 and DTM data to map DLT, (iii) to evaluate if the combination of S1/S2 and DTM improves accuracies and overcomes problems of complex topography and atmospheric constraints in a mountainous country, (iv) to develop a workflow that is highly automated, as simple as possible, and reproducible, (v) to evaluate the performance of RF and a DL approach (UNET) to map DLT, and (vi) to produce output that is relevant and applicable by practitioners of different fields. The present study builds on the countrywide tree type mapping approach of Waser et al. (2017) and was motivated by the increasing need for spatially explicit, wall-to-wall, accurate and frequently updated information on forest composition in the forestry sector and beyond.

2. Materials and methods

2.1. Overview

To map DLT at a national level, two supervised classification approaches, i.e. RF (Breiman, 2001) and UNET (Ronneberger et al., 2015) were applied. The workflow of the countrywide DLT mapping approach is shown in Fig. 1. We used three sets of predictor variables from S1/S2 data and from a DTM, reference data from orthoimage interpretation and a forest mask. All data sets were resampled to 10 m ground sample distance (GSD). The reference data was split as follows: for model calibration into training (90%) (grey boxes in Fig. 1) and validation (10%) (red boxes), and for model testing (“blind”, not used for model calibration) (blue boxes), which were spatially stratified using a 25 km regular grid. Model calibration was repeated five times, each time the training, validation and test data sets were newly selected. Additionally, the predicted final maps were compared with independent NFI plot data. All modeling and validation were performed in R 4.0.0 (R Core Team, 2021) and Python 3.7 (VanRossum and Drake, 2001). RF was run using the R package caret, and UNET was applied in the Python package PyTorch.

2.2. Study area

Switzerland is a central European country on the Alpine arc, located between 5°57’ and 10°29’ E and 45°49’ and 47°48’ N, covering an area of 41,285 km² and an elevation range between 197 and 4634 m a.s.l. (see Fig. 1). Mountainous areas cover two-thirds of the country, and the topography in these regions is complex, ranging from flat valley bottoms to steep slopes and narrow gorges. Switzerland belongs to the temperate climatic zone but ranges from polar conditions at high elevations in winter to Mediterranean warmth in the southern region during summer. From low to high elevation, vegetation belts present in Switzerland include colline, montane, sub-alpine and alpine. Forest area has continuously increased over recent decades, largely related to processes of agricultural land abandonment and associated afforestation, while tree species richness decreases with increasing elevation. Forests currently cover approximately 31% of Switzerland’s area and consist of 47.1% broadleaved and 52.9% coniferous trees (Abegg et al., 2014; Brändli et al., 2020). According to the fourth Swiss NFI report (Brändli et al., 2020), the forests up to approx. 1200 m a.s.l. are managed to a high degree and dominated by European beech (Fagus sylvatica L.), European aspen (Praxinus excelsior L.), oak (Quercus sp.), European white fir (Abies alba Mill.), and Norway spruce (Picea abies (L.) H. Karst), with typical broadleaved and softwood species growing along rivers and streams. High-elevation forests near the upper treeline, which varies between 1600 m a.s.l. in the northern Prealps to an average of 2300 m a.s.l. in the Central Alps, consist of the dominant tree species Scots pine (Pinus sylvestris L.), European larch (Larix decidua Mill.) and Norway spruce (Picea abies (L.) H. Karst), with a minor presence of dwarf mountain pine (Pinus mugo Turra) and green alder (Alnus viridis (CHAIX) DC.).

2.3. Remote sensing data and pre-processing

2.3.1. Digital terrain model

The DTM swissALTI3D from Swisstopo is based on ALS data (<2000 m a.s.l.) and stereo-image-based point clouds (>2000 m a.s.l.). Both the ALS and image data were acquired between 2010 and 2015, with 0.5–2 points per m². The original swissALTI3D DTM has a spatial resolution of 2 m. To match the sample interval of the S1/S2 data, it was bilinearly resampled to 10 m (Ginzler and Hobi, 2015).

2.3.2. Forest mask

The forest mask from the Swiss NFI (Waser et al., 2015), which is based on Digital Surface Models (DSMs) from aerial stereo-images, was
used to mask the final predictions. The original forest mask is a countrywide data sets which is regularly updated and has a spatial resolution of 1 m. To match the sample interval of the S1/S2 data, it was bilinearly resampled to 10 m.

2.3.3. Sentinel-1 data sets

To take advantage of the opposing seasonal backscatter signal between broadleaved and coniferous trees in SAR C-band data, we obtained all S1 Ground Range Detected High-resolution (GRDH) products covering Switzerland taken within 24-days leaf-off and a leaf-on periods in three different years. In the Interferometric Wide (IW) swath mode, S1 acquires data at incident angles between 31° and 46°, with a swath width of 250 km (Torres et al., 2012). The leaf-off period was defined as 2–25 March and the leaf-on period as 29 August to 21 September. For the leaf-off period, March was chosen to reduce potential issues with meteorological influences on the backscatter (e.g. wet snow, temperatures < 0 °C) while still ensuring leafless tree crowns over most of the country (Rüetschi et al., 2018). To produce countrywide products, the S1 data was geometrically and radiometrically corrected and composited in two steps. In the first step, the GRDH images with a pixel spacing of 10 m were radiometrically calibrated and terrain-corrected, and then terrain-geocoded using the Radiometric Terrain Correction (RTC) methodology of Small (2011) and the above-mentioned DTM. In the second step, the Local Resolution Weighting (LRW) methodology was applied, merging all intermediate images from both leaf-off and leaf-on periods for each year into one composite backscatter image (Small, 2012, Small et al., 2021). The LRW methodology further reduces terrain-induced artefacts and enhances spatial resolution by combining multiple acquisitions from multiple geometries (esp. ascending and descending) at the cost of temporal blurring. These two pre-processing steps were conducted for S1 data of both polarizations, VV and VH, resulting in a total of five VV and five VH composites (leaf-off 2015, leaf-off 2016, leaf-on 2016, leaf-off 2017, leaf-on 2017) with a sample interval of 10 m.

2.3.4. Sentinel-2 data sets

Multispectral S2 Level-2A products, which provide bottom-of-atmosphere reflectance values, acquired between 2016 and 2018 were used. To cover the extent of Switzerland, 11 S2 tiles were required. To minimize shadows and snow-covered forest areas and to capture with flushed leaves, all available S2 tiles (a total of 1814) in the period between the beginning of June to the end of September were considered. Since the cloud mask provided by ESA (Zhu et al., 2015) did not sufficiently remove all cloud pixels, particularly in mountainous areas we continued the analyses including only S2 tiles that had a cloud coverage ≤ 30%. This resulted in a total of 870 tiles available for the subsequent analyses. For each S2 pixel, all cloud-free observations were used to calculate its median value for each S2 spectral band. For the present study, only bands that are informative for vegetation analysis were used, i.e. those with a 10 m (B2–B4, B8) or 20 m (B5–B7, B8A, B11, B12) spatial resolution. The 20 m bands were resampled to 10 m so that pixel spacing was the same for all bands.

2.4. Predictor variables

The predictor variables consisted of S1, S2 and DTM data. First, both VV and VH backscatter information from S1 was used. Second, S2 original bands with an initial spatial resolution of 10 m and 20 m, respectively, were used. In addition, we used three Vegetation Indices (VI), i.e. Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974), Modified Soil-Adjusted Vegetation Index (MSAVI2) (Qi et al., 1994) and Global Environment Monitoring Index (GEMI) (Pinty and Verstraete, 1992). Third, the three topographic variables elevation (m a. s.l.) (E), aspect (A) and slope (S) were derived from the DTM to deal with problems in the data sets regarding illumination and shadows in mountainous regions (Franklin et al., 1986; Wilson and Gallant, 2000; Grabhska et al. 2020; Waser et al., 2017). Table 1 gives an overview of all calculated predictor variables.

In total, the three sets of predictor variables S1 + EAS, S2 + EAS and S1S2 + EAS were used to train the RF and the UNET classification algorithms.

2.5. Reference data

2.5.1. Data acquisition

The reference data consisted of pixels that were assigned to the classes broadleaved and coniferous trees within the area of the forest mask. The original data set was produced in the approach of Waser et al. (2017) and consists of digitized polygons of large tree crowns or tree groups were obtained from orthophotos with a 25–50 cm spatial resolution. In the present study, to handle the 10 m spatial resolution of the Sentinel data, only pixels that lay 100% within the digitized polygons were selected (Fig. 2). This resulted in a total of 195,222 pixels (112,940 polygons) and 230,520 pixels (72,300 polygons) for the classes broadleaved and coniferous trees, respectively.

2.5.2. Data splitting

The reference data was stratified into a) training and validation data for model calibration and b) data for model testing using a 25 km regular grid. For this, the whole of Switzerland was first divided into 92 equally sized and non-overlapping blocks. Then, 10% of these blocks were randomly selected for testing the classification models (so called “blind
data”). The reference data in the remaining 90% of the blocks were used for model calibration (training 90% and validation 10%) (Fig. 1). For model calibration, the training and validation data sets were built as follows: for RF (see Section 2.6.1.), a 10-fold spatial cross-validation was applied, where the blocks were also used to generate training and validation data sets; for UNET (see Section 2.6.2.), the remaining 90% of the blocks were again randomly subdivided into training and validation data sets (90% and 10%, respectively). This procedure was run five times (in Fig. 1 the blue parts represent the regions used for model testing in one run). Each time, the same test data sets were used for both RF and UNET. Such a reference data stratification strategy enables to reduce bias in training, validation and testing and, thus, leads to a better evaluation of model performances (Roberts et al., 2017; Schratz et al., 2019).

2.6. Classification algorithms

The two machine learning algorithms RF and UNET were applied to assess their utility for countrywide classification of DLT using the three sets of predictor variables.

2.6.1. Random forest

RF (Breiman, 2001) is a widely used ensemble learning classifier that uses a randomly selected subset of training samples and predictor variables to produce multiple decision trees. The final classification output is created by aggregating the predictions of all individual decision trees using majority voting. The two hyperparameters, which have to be tuned while training RF, were handled as follows: ntree (number of trees) was set to its default (500) and mtry (number of predictor variables randomly sampled for each node) was tuned using a search grid by testing values between two and the total number of predictor variables with a step of five. To evaluate the performances of the RF models while training, spatial 10-fold cross-validation was used. The final models were then evaluated using the test data set (“blind data”). Variable importance was assessed using the Mean Decrease Gini (MDG) that measures the reduction in the Gini impurity metric by a variable for a particular class (Breiman, 2001) (Fig. A1).

2.6.2. UNET

UNET (Ronneberger et al., 2015) is a fully convolutional network for fast and precise segmentation of images. UNET has a typical encoder-decoder architecture and consists of a contracting path and a symmetric expanding path. The contracting path enables the algorithm to learn global features through a series of convolution blocks, which are followed by pooling layers (Fig. 3). Thus, the size of the feature maps is substantially reduced, while their number permanently increases. In the expanding path, pooling layers are replaced with up-sampling layers, the output of which is concatenated with information from earlier layers and followed by convolutional blocks. The number of feature maps is decreased continually. The fusion with information from earlier blocks forces the network to learn local features as well, which improves the accuracy of the segmentation.

In our study, each convolutional block consisted of two convolutional layers. The output of each convolutional layer was normalized and passed through the Rectified Linear Unit (ReLU). To prevent overfitting of the network and increase its generalization ability, a dropout layer with a rate of 0.2 was introduced before every second convolutional layer (Fig. 3). The dropout rate indicated the probability of a neuron being dropped out randomly. UNET was trained over 60 epochs using a batch size of 16. The initial learning rate was set to 0.01 and was divided by 10 every 10 epochs. Stochastic Gradient Decent (SGD) with a momentum of 0.9 and weight decay of 0.0005 was used to optimize the network parameters. As a loss function, a linear combination of binary cross-entropy loss (L_BCE) and dice loss (L_dice) was applied:

\[
L(T, P) = L_{BCE}(T, P) + \lambda L_{dice}(T, P)
\]

(1)

where \(T\) indicates true values and \(P\) indicates predicted values. Learning curves of the three UNET models are shown in Fig. A2.

To generate input data for the UNET, the S2 mosaic and S1 composite of the whole of Switzerland were divided into overlapping image patches (overlay = 50%) with a window size of 128 × 128 pixels. UNET received a batch of the image patches with an input shape of 128 × 128 × N × n, where \(N\) indicates the number of image bands (i.e. predictor variables), and \(n\) the batch size. Correspondingly, only the internal non-overlapping parts of the image patches were used for optimizing network parameters, as well as producing the final probability map (see Fig. 3 for details). To compare the classification algorithms, the same three sets of predictor variables (S1 + EAS, S2 + EAS and S1S2 + EAS) were used as in the RF classification. The pixel values in the training, validation and test data sets were normalized using minimal and maximal values derived from the training data set.

Performances of the UNET models during training and validation were evaluated using the loss function. Similarly to the RF models, performance of the final UNET models was evaluated using OA and K.
2.7. Accuracy assessment

2.7.1. Model accuracies

Performances of the RF and UNET models (based on the three sets of predictor variables) were evaluated using Overall Accuracy (OA), Cohen’s Kappa coefficient ($K$) (Cohen, 1960), Precision ($P$) and Recall ($R$). For further details and use see e.g., Powers (2011) and Richter et al. (2012). In order to better reflect the highly complex topography, Switzerland was additionally divided into 24 subareas comprising eight aspect categories ($N = 337.5^\circ - 22.5^\circ, NE = 22.5^\circ - 67.5^\circ, E = 67.5^\circ - 112.5^\circ, SE = 112.5^\circ - 157.5^\circ, S = 157.5^\circ - 202.5^\circ, SW = 202.5^\circ - 247.5^\circ, W = 247.5^\circ - 292.5^\circ, NW = 292.5^\circ - 337.5^\circ$) and three slope categories.

Fig. 4. Distribution and number (in pixels) of broadleaved (orange) and coniferous (green) samples across the 24 subareas, as used for model accuracy assessment. The 24 subareas are a combination of the three categories of slope (grey scale) and eight categories of aspect (red scale). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 5. Example of the aerial stereo-image interpretation of the 25 sub-sampling points (Right) of the Swiss NFI 1.4 km regular raster (Left) for coniferous trees (green) and broadleaved trees (light green) decisions. For the interpretation color-infrared (CIR) stereo-images were used. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Fig. 6. Boxplots are showing A) overall statistics (OA, K) and B) per-class statistics (P and R) obtained from RF (grey) and UNET (yellow) for the three sets of predictors S1 + EAS, S2 + EAS, S1S2 + EAS, where E = elevation, A = aspect and S = slope. The central lines within each box are the medians. The boxes’ edges are the upper and lower quartiles. Outliers are depicted as grey dots. Testing performance of the two modeling approaches is based on the 5 runs of model training and testing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 7. Boxplots are showing the testing performance of the RF and UNET for each of the 24 subareas based on the 5 runs of model training and testing. The central lines within each box are the medians. The boxes’ edges are the upper and lower quartiles. Outliers are depicted as points. Weighted K values are based on the RF model (grey) and the UNET (yellow). The following three sets of predictor variables were used: (a) S1 + EAS, (b) S2 + EAS, and (c) S1S2 + EAS, where E = elevation, A = aspect and S = slope. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
deviation (≥10°, 10° to < 30°, ≥30°). Model accuracies of RF and UNET were separately evaluated for each of the 24 subareas (see Fig. 4).

2.7.2. Map accuracies

Swiss NFI data were used as an independent data set to assess the accuracy of the total of 6 final predictions (maps) from the RF and UNET models. The Swiss NFI is a continuous survey with 6,500 plots on a 1.4 km grid that are visited over a period of nine years. It follows a two-phase sampling approach that incorporates terrestrial surveys and aerial stereo-image interpretation. The aerial stereo-image interpretation is conducted yearly in 25 sub-sampling plots within a 50 m square (2,500 m²) around the terrestrial plot centers for one-ninth of the NFI sample plots using stereo-imagery from 2009 to 2018. In addition to other land cover classes, broadleaved and coniferous trees are interpreted (see Fig. 5). For more details on the NFI interpretation see e.g. Mathys et al. (2006) and Waser et al. (2017). According to the Swiss NFI forest definition, only interpreted trees with a minimum height of 3 m were chosen.

The pixel value of the predictions was extracted and compared at each sub-sampling point with its land cover information. The means of the statistics OA, K, P and R were calculated for the 3173 plots that included a minimum of nine trees. In addition to this, the Standard deviation (SD) and Standard error (SE) of the OA was also derived, to get an idea of the variability between the NFI plots. Five classes of proportion of broadleaved, i.e. > 80%, 60–80%, 40–60%, 20–40% and < 20% enabled to evaluate the prediction in mixed forest conditions.

### 3. Results

#### 3.1. Model performances

##### 3.1.1. Countrywide model performances

The statistics obtained for RF and UNET based on the three sets of predictor variables were evaluated and are shown in Fig. 6. The corresponding cumulative confusion matrices from model predictions on the test data sets are given in Table A1. Furthermore, the Gini importance for the RF model and the learning curves of the UNET are given in the Appendix (Figs. A1 and A2).

The classification performance of UNET was generally higher when compared to RF. Lowest OA and K values (Fig. 6 A) were obtained when S1 + EAS was used, for RF (0.90, 0.79) and for UNET (0.94, 0.88) with the latter performing better. Models based on S2 + EAS or combined S1S2 + EAS predictors revealed comparable OA and K values for RF and UNET and were only slightly lower for RF. Moreover, RF produced similar accuracies (0.96, 0.92) independent of the predictors S2 + EAS or S1S2 + EAS. UNET produced OA and K values (0.97, 0.94) when
using S2 + EAS and (0.97, 0.95) when using all predictors (S1S2 + EAS).
The variances were higher for RF than for UNET. Fig. 6 B shows the per-class accuracies. The P statistic measures how accurate the broadleaved and coniferous classes are, i.e. the proportion of the predicted pixels of the two classes really belong to these classes, per set of predictor variables for both RF and UNET. The R statistic shows the proportion of correctly classified broadleaved and coniferous pixels per class. P and R were almost similar for both classes when using UNET and generally higher than for RF. This difference increased when only S1 + EAS were used, i.e. UNET produced higher P and R values (both around 0.93) for both broadleaved and coniferous classes compared with the RF-based predictions (P = 0.88–0.91, R = 0.89–0.90). For both RF and UNET models, P and R for both classes were similar when either using S2 + EAS or S1S2 + EAS.

3.1.2. Subarea-wide model performances

To cope with the complex topography of Switzerland, the whole area of Switzerland was divided into 24 subareas based on the combination of the eight aspect and three slope categories, see Fig. 4). The testing performances of RF and UNET on each of the 24 subareas are illustrated in Fig. 7 for K per set of predictor variables and Fig. A3 per classifier.

Fig. 7 shows that for all 24 subareas highest weighted K values were obtained when using the predictors S1S2 + EAS, ranging between 0.85 and 0.98 for UNET and between 0.8 and 0.95 for RF, respectively. When using RF, differences, between S2 + EAS and S1S2 + EAS were highest in flat subareas of any exposition. In contrast, when using UNET, such differences were obtained in subareas with steeper slopes. A more detailed illustration of the subsets for which the predictors S1S2 + EAS were superior to S2 + EAS is given in Fig. A4. While lowest K values were obtained when using the predictors S1 + EAS, ranging between 0.7 and 0.9 for UNET and between 0.5 and 0.85 for RF, they were 5–20% higher when S2 + EAS was used. Largest differences were found in subareas 8, 9, 16, 18, 23 and 24 with lowest K (around 0.6 and 0.5) for subareas 23 and 24. Similarly, variances of K were overall highest when using S1 + EAS and higher when RF was used. They were reduced when using S1S2 + EAS, but remained high for subareas 17, 18, 23 and 24.

Fig. 8. Predictions of DLT with the classes broadleaved and coniferous for an area of interest (2.5 km × 2.5 km) in the Central Plateau of Switzerland with relatively flat terrain. The forest is mixed and typically patchy and managed. The RF and UNET predictions were masked out to the extent of the forest mask and based on the three sets of variables S1 + EAS, S2 + EAS and S1S2 + EAS, where E = elevation, A = aspect and S = slope. Polygons of the training and test data are depicted in yellow. © Includes modified Copernicus-Sentinel-Data (2015–18), swisstopo and Swiss NFI. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
3.2. Map accuracies

3.2.1. Comparison with independent NFI data

The accuracy of the final maps (see Fig. A5 for an example of the countrywide DLT map) was evaluated by comparing (per NFI sample plot) the predicted DLT pixels and the observed (NFI aerial stereo-image interpretation) information for the corresponding lattice point. The statistics (overall and per class of broadleaved fraction) of the comparisons for the analysis is shown in Table 2.

The results of the countrywide analysis show that the obtained accuracies were generally very similar between RF and UNET predictions when either S2 + EAS or S1S2 + EAS were used, with OA around 0.8 and K around 0.6, and slightly lower when S1 + EAS were used (OA around 0.77 and K around 0.54). Per-class statistics revealed high accuracies for both RF and UNET for relatively pure forest stands (OA around 0.94, K around 0.88) for lowest proportion of broadleaved (<20%) followed by highest proportion of broadleaved (OA around 0.83, K around 0.6).

Generally lower accuracies were obtained for classes with intermediate proportions of broadleaved (mixed forest stands), with lowest for the class 40–60% broadleaved (OA around 0.6, K around 0.17, with the exception of S1 (K 0.06)). This trend was similarly independent on used set of predictors. One exception was found for the proportion of broadleaved (<20%), where UNET only based on predictors S1 + EAS outperformed RF based on S1S2 + EAS. The differences for P and R between the classes broadleaved and coniferous indicate a slight over-estimation of coniferous for all approaches. Both SD and SDE varied and were highest for the class > 80% of broadleaved, and lowest for the intermediate class (40–60%).

3.2.2. Visual inspection

For visual inspection of the maps, we selected two representative areas with different topography and tree species composition (Figs. 8 and 9).

Fig. 8 illustrates that, the spatial patterns of the classes broadleaved and coniferous are overall quite similar in the entire area when using either the predictors S2 + EAS or S1S2 + EAS independent of using RF or UNET. The largest accuracies were obtained when S1 + EAS was used alone, in particular for UNET compared with RF. Fig. 8 also shows that UNET produced a more generalized prediction compared to RF with the
largest differences when S1 + EAS predictors were used alone. In this case, the RF prediction also seemed to be noisier and UNET seemed to underestimate the class coniferous. However, a closer look at the CIR-orthophoto shows that the prediction of UNET is more realistic. Obviously, the class coniferous tends to be slightly overestimated when S2 + EAS predictors were used alone. In this case, the CIR-orthophoto includes the forests

Fig. 9 illustrates that the predictions from RF and UNET models differ in terms of spatial patterns and level of detail. Obviously, UNET again produced a more generalized prediction compared with the RF model, with minor exceptions in the center of the area of interest. Compared with the example from the Central Plateau (Fig. 8), the overall similar patterns of the spatial distribution of the two classes broadleaved and coniferous were less pronounced and differences occurred at pixel-level rather than per area. RF predictions based on the single use of S1 + EAS were noisy, particularly in mixed stands, and were difficult to interpret. The largest differences in the predictions were obtained between models based on S1 + EAS and S2 + EAS or S1S2 + EAS, respectively, and were most

Table 3
Comparison of the Copernicus DLT 2018 data set with the RF and UNET based DLT maps. The three sets of predictor variables S1 + EAS, S2 + EAS and S1S2 + EAS were used, where E = elevation, A = aspect and S = slope. The comparison is based on independent NFI plot data, that incorporates 3,173 sample plots. The means of K were obtained per class of proportion of broadleaved, i.e. > 80%, 60–80%, 40–60%, 20–40% and < 20% and in total. Bold numbers indicate for which class of proportion of broadleaved the Copernicus DLT 2018 performed better than our approaches with RF and UNET.

| Broadleaved in % | S1 + EAS | S2 + EAS | S1S2 + EAS |
|-----------------|---------|---------|-----------|
|                | >80     | 60-80   | 40-60     | 20-40    | 0-20   | Total  |
| No of NFI plots| 1056    | 389     | 333       | 296      | 1099   | 3173   |
| K               | 0.70    | 0.17    | 0.10      | 0.28     | 0.71   | 0.54   |
| Copernicus DLT  | 0.56    | 0.24    | 0.16      | 0.34     | 0.86   | 0.60   |
| RF S1 + EAS     | 0.67    | 0.25    | 0.15      | 0.31     | 0.86   | 0.59   |
| UNET S1 + EAS   | 0.60    | 0.14    | 0.08      | 0.28     | 0.89   | 0.56   |
| UNET S2 + EAS   | 0.65    | 0.24    | 0.16      | 0.33     | 0.98   | 0.60   |

Table 4
List of criteria for user-specific applications and interpretation of the DLT maps for the single use of S1 + EAS, S2 + EAS and combined S1S2 + EAS data for both RF and UNET, where E = elevation, A = aspect and S = slope. Note: The recommendations are to be interpreted as follows: - (not recommended), +/- (partly recommended) and + (recommended). *Average computation time using a conventional desktop PC with a high-performance graphic card. **Per area, e. g. on forest stand level. Note, that a forest stand might be individually defined by each country.

| Criteria                    | Set of predictor variables |
|-----------------------------|---------------------------|
| Data acquisition time       | S1 + EAS                  |
| 1–3 years                   | 2–3 years during leaf-off/|
|                            | during leaf-on            |
| Computation time*           | (modelling,               |
|                            | prediction)               |
| RF (2 h, 3 h)               | RF (2 h, 3 h)             |
| UNET (10 h, 3 h)            | UNET (10 h, 3 h)          |
| Classification              | RF                         |
| RF +                        | RF +                       |
| UNET +/-                    | UNET +                     |
| Product update              | 2–3 years                  |
|                            | 3 years                    |
| Map accuracy                | Pure stands +/-           |
|                            | Mixed stands +/           |
|                            | Mixed stands +            |
| Application scale           | Any location +/           |
|                            | Any location +            |
|                            | +/                         |
|                            | Any location +/           |
|                            | **Per area +/             |
|                            | Pixel level -             |
|                            | **Per area +/             |
|                            | Pixel level -             |
| Users                       | Federal agencies,         |
|                            | Mapping agencies,         |
|                            | NFI, Reporting,           |
|                            | Researchers,              |
|                            | Foresters                 |
pronounced by RF. Independent on the use of sets of predictors, RF tended to overestimate the class coniferous in mixed forest stands (e.g. in the center) and underestimate the class coniferous in broadleaved-dominated stands (right lower corner). In contrast, this underestimation by UNET was only found when S1 + EAS predictors were used alone.

4. Discussion

4.1. Input data

In the following, traits of the used input data are discussed. The generation of predictor variables was based on the overall target of the approach to guarantee high classification accuracies and relevance for practice, i.e. effectiveness in considering the required input data, reproducibility and reasonable computation times.

The inclusion of DTM predictors such as elevation, aspect and slope was based on the idea of overcoming classification problems in complex topography. Since the pioneer work of Strahler et al. (1978) and Franklin et al. (1986), the use of terrain data has been examined in other recent studies (e.g. Waser et al. 2017 based on ALS data; Droin et al., 2018; Hoszilo and Lewandowska, 2019 based on SRTM data). Thus, the three sets of predictors from S1, S2 and combined S1/S2 data always included elevation, aspect and slope. Although the high-resolution DTM as used in the present study is not available for every country, existing data sets, e.g. on European level (Hengl et al., 2020) in combination with upcoming data sets such as e.g. GEDI will be an appropriate alternative.

The use of S1 SAR C-band backscatter data was based on the idea of compensating for problems with optical S2 data and testing its capability to distinguish the classes broadleaved and coniferous, building on results reported by Rütschi et al. (2018). This is in contrast to assumptions made in other studies, where radar data was principally not applied to mountainous areas due to topographic shadows (e.g. Immitzer et al., 2019). According to our knowledge, there is no other study in which the utility of S1 data in mountainous regions has been examined for the distinction of the two classes broadleaved and coniferous on a country level. The main reason might be the non-trivial radiometric terrain correction and compositing of S1 data, as well as meteorological influences (e.g. wet snow and temperatures below 0 °C), in particular in higher elevations. These influences were minimized by choosing March data for the leaf-off period. The SAR data pre-processing enabled the retrieval of relatively “flat” backscatter signatures even over regions with severe terrain undulations. Without such pre-processing, it can become difficult to guarantee a consistent treatment over broad regions, i.e. local terrain-induced distortions can intrude. The successful application of LRW to generate wide-area composite backscatter products as input to analysis of forest signatures was reported in (Rüetschi et al., 2018) for forest type classification, and in (Rüetschi et al., 2019) for rapid detection of windthrow events over a forested region. As expected, VH predictors from S1 proved to be more important than VV predictors because the opposing seasonal backscatter behavior between the classes broadleaved and coniferous was more distinct in VH (Fig. A1). In general, the March predictors were expected to be more important because the backscatter differences between the classes broadleaved and coniferous are higher during the leaf-off than the leaf-on period in both polarizations (Rüetschi et al., 2018). The low importance of the March predictors could be attributed to the above-mentioned meteorological influences. Future studies should address this issue by defining the leaf-off period for the backscatter composite production for different areas of the country based on their elevation.

In the present study, obtaining countrywide cloud-free S2 level 2A imagery from summer was only feasible with data from three years, and 870 tiles were required to cover an area of 41,000 km². Beyond the seasonal frequency of cloud cover, high solar radiation during the summer half-year in mountains is responsible for the development of convective clouds. Thus, chances of obtaining cloud-free imagery from close acquisition dates are low and a multi-season set of predictors, as suggested in other studies (e.g. Wessel et al., 2018; Grabksa et al., 2019; Immitzer et al., 2019), is not feasible for countrywide approaches within reasonable time. Our study further demonstrates that time series are not a prerequisite to distinguish between the classes broadleaved and coniferous. Based on a previous countrywide mapping approach (Waser et al., 2017) and tree species classification studies (Waser et al., 2014; Fassnacht et al., 2016), we applied the three vegetation indices GEMI, NDVI and MSAVI2, which substantially contributed (>50% besides the NIR and Red-Edge bands) to the distinction between the classes broadleaved and coniferous. Least contribution was obtained from the DTM and S1 predictors (Fig. A1). However, since several of the S2 bands and indices might be correlated, the absolute contributions may vary depending on the RF model and have to be interpreted with care.

For the present study, reference data from Waser et al. (2017) based on aerial image interpretation was used. This was advantageous in regard to both quality and time resources since the collection of sufficient and representative reference data remains a demanding task for countrywide approaches. We assume that the large number of samples in our approach (as it is suggested for deep learning approaches by e.g. Hoeser and Fuenzler, 2020) was sufficient for the UNET model. Although existing reference data from other sources (available at different scales) could potentially be used, problems are related to their heterogeneity and acquisition time span. On national level, automated generation of reference data by using deep learning methods, as suggested by Zhu et al. (2017) could be a promising alternative and will be tested in future.

4.2. Model performances and map accuracies

4.2.1. Model performances

In our study, the highest model accuracies were obtained when all predictors (S1S2 + EAS) were used. While model accuracies were only slightly lower when the predictors S2 + EAS were used, they were substantially lower when only the predictors S1 + EAS were used. Adding more than five 24-day S1 seasonal snapshots might be considered in the future. Not only that UNET slightly outperformed RF, vari- ances of K and OA were also lower. These two main differences substantially increased for models that were only based on predictors S1 + EAS. We assume that the high performance of UNET using S1 data is related to the architecture of the algorithm, which creates textural features based on neighborhood pixels. UNET models solely based on S2 + EAS slightly overestimated the class broadleaved, in contrast to RF models for which a slight overestimation of the class coniferous was observed. The use of S1S2 + EAS predictors reduced this trend, resulting in more balanced predictions when UNET was used, but obviously not entirely for RF.

When using the predictors S1S2 + EAS, the obtained accuracies (K and OA) in the present study, with values (0.95 and 0.97) for UNET and 0.92 and 0.96 for RF, respectively, are slightly higher than those reported in other studies that distinguished between the two classes broadleaved and coniferous. For example, Hoszilo and Lewandowska (2019) obtained a K of 0.89, and overall accuracies of 93% (Immitzer et al., 2019), around 90% (Wessel et al., 2018; Cheng and Wang, 2019) and around 75% (Droin et al. 2018) have been reported. However, direct comparisons with other studies have to be interpreted with care, as most previous investigations have been carried out on a case study level, with a relatively small extent and optimized training data (large sample) and S2 data (cloud-free multi-temporal images). In such case studies, classification of DLT was the first step in the classification of up to six tree species.

4.2.2. Comparing map accuracies with NFI data

Overall, the high level of precision obtained by RF and UNET was not entirely found in the DLT predictions. The comparison with independent NFI data (Table 2) indicated largest differences in areas with high degree of mixture of the classes broadleaved and coniferous and a slight trend of overestimation of coniferous as dominant leaf type. However, the absolute deviations must be carefully interpreted since the NFI data is an
independent data source for comparisons. The comparisons were affected by the time gap, which ranged from a few months to nine years and the different spatial levels of detail between the NFI data and the predictions. A comparison with terrestrial surveys was not considered, due to the difference in data collection (ground-based measurements versus canopy cover from Sentinel data).

4.2.3. Comparing map accuracies with other DLT maps

The increased quality of our DLT maps is clearly visible in Fig. 10, which is representative for large parts in Switzerland, with a comparison of aerial image-based DLT (Waser et al. 2017) and the Copernicus DLT 2018 (EEA, 2021).

Large differences between the two DLT predictions based on Sentinel data and the one based on aerial images are apparent in broadleaved forests, in particular in the eastern part of the map (see the mark X in Fig. 10). A general overestimation of the class coniferous is seen when aerial images are used. Even in the photograph, it is difficult to distinguish between broadleaved and coniferous trees located in the shadows. The whole eastern side of the valley is dominated by broadleaved trees (mainly beech, birch and chestnut). In the remaining parts of the valley, the dominance of the class coniferous based on the predictors S1S2+EAS is clear. The class broadleaved (mainly beech, birch and chestnut) is visible on the western part of the valley, where broadleaved stands are not clearly visible.

In the eastern part of the map (see the mark X in Fig. 10), a general overestimation of the class coniferous is seen when aerial images are used. Even in the photograph, it is difficult to distinguish between broadleaved and coniferous trees located in the shadows. The whole eastern side of the valley is dominated by broadleaved trees (mainly beech, birch and chestnut). In the remaining parts of the valley, the dominance of the class coniferous based on the predictors S1S2+EAS is clear. The class broadleaved (mainly beech, birch and chestnut) is visible on the western part of the valley, where broadleaved stands are not clearly visible.

In the eastern part of the map (see the mark X in Fig. 10), a general overestimation of the class coniferous is seen when aerial images are used. Even in the photograph, it is difficult to distinguish between broadleaved and coniferous trees located in the shadows. The whole eastern side of the valley is dominated by broadleaved trees (mainly beech, birch and chestnut). In the remaining parts of the valley, the dominance of the class coniferous based on the predictors S1S2+EAS is clear. The class broadleaved (mainly beech, birch and chestnut) is visible on the western part of the valley, where broadleaved stands are not clearly visible.

4.3. Operational use

Effective operational use is ensured to a high degree because the focus was laid on relevance for practice (user friendly) by using open source software and freely available, frequently acquired remote sensing data with a high product consistency. While the acquisition and preprocessing of S2 data has become standard and user-friendly (e.g. Google Earth Engine), the required robust geometric and radiometric terrain corrections applied to the S1 data (Small et al., 2021), are more time intensive. Since the latter is currently being adapted to full automatically, operationality in this regard will be guaranteed as well. Table 4 serves as basis for the user-specific applications and interpretation of the DLT maps.

We offer three main recommendations:

1. To avoid misinterpretations, in the sense of spatial explicitly at 10 m level, the use of DLT information per pixel is not recommended. Instead of an area-based use of DLT, e.g. per forest stand is preferable.

2. Overall reliable predictions from countrywide to forest stand level were achieved by using S1S2+EAS - regardless of the use of RF or UNET. Both predictions show similar patterns of the spatial distribution of the two classes broadleaved and coniferous. DLT obtained from UNET is more homogenous and the dominant class per area, e.g. forest stand, is more precisely predicted than by RF - which is helpful for the forest practice, e.g. for updating stand maps. If the dominant class is less pronounced for an area, RF predictions vary more. If the focus is laid on detecting pure broadleaved or coniferous stands, UNET is preferable as well.

3. While the single use of S2+EAS is recommended for countrywide assessments and principally also for a more precise assessment on local scale, the single use of S1+EAS predictors is not recommended when cloud-free optical composites are available.

The predictions of DLT can be used to optimize forest management and planning and (inter)national reporting, and they might be important for reliable assessments of forest resources. The countrywide DLT maps of Switzerland presented here already serve as an important information source for many other regional or national uses, e.g. for modelling biomass (Price et al., 2020); handling protection issues, e.g. concerning rockfall and avalanches; managing natural hazards, i.e. drought, insects and diseases; and providing new perspectives for biodiversity studies. Based on first feedback from the forestry sector, the DTL maps complement the sample-based NFI data by giving spatially explicit information on distribution of the two classes broadleaved and coniferous, enabling densification of the existing 1.4 km NFI sample plot grid for smaller areas of interest. The DLT (https://doi.org/10.16904/1000001.6), provided by the Environmental Research Data of Switzerland (EnviDat) (https://www.envidat.ch/#/metadata/forest-type-nfi) and the mapping platform of the Swiss Confederation (https://map.geo.admin.ch) can be freely used and downloaded.

We assume that the presented DLT mapping approach can be applied to other countries with similar topography and forest ecosystems provided that a DTM, similarly processed S1 data and a representative training data set will be available and an independent validation will be performed. The training data set should be representative and spatially distributed over the whole area of interest. In that case, we also assume that other sources of training data could be used as well.

5. Conclusions

In the present study, high-quality DLT maps with the two classes broadleaved and coniferous were produced for the whole of Switzerland (41,000 km²). The large set of reference data was stratified using 92 equally sized blocks in order to reduce bias between training, validation and test data sets, thus enabling a better evaluation of the classification models. RF and UNET overall performed very similarly when using all predictors from S1, S2 and DTM data (K around 0.95). Differences were found in the single use of S1 and S2 and were less pronounced by UNET than by RF. Lowest accuracies were obtained when only using S1 and DTM predictors. The same effects were also found in the predictions. While UNET tended to predict the dominant class more accurately than RF, predictions from RF varied more and the dominant class was less pronounced. From a practical point of view, the use of DLT is recommended per area, e.g. forest stand rather than per pixel. A comparison of our approach and the Copernicus DLT 2018 with independent NFI data revealed overall slightly higher accuracies for our approach with the exception of pure broadleaved stands.

With this study, we have demonstrated that the combined use of S1/S2 and DTM mitigates problems of complex topography and atmospheric constraints in a mountainous country. We explored the requirements for mapping countrywide DLT, in particular regarding the use of freely and openly available S1/S2 data, as well as open source software, repeatability, updating and characteristics of training data.
sets. The introduced workflow is robust, cost-efficient, highly-automated and helps strengthen the bridge between practice and research and further promote RS-based products as spatially explicit and thus complementary information to existing sample-based NFI data. It could therefore help stimulate activities in the forestry sector that rely on accurate spatial descriptions of forest resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Sentinel-1 and Sentinel-2 data are courtesy of ESA and the DTM of Swiss Federal Office of Topography swisstopo. This study was carried out in the framework of the Swiss National Forest Inventory (NFI), a cooperative effort between the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) and the Swiss Federal Office for the Environment (FOEN). We thank Adrian Lanz, Christian Ginzler and Dominque Weber for fruitful discussions on data validation, Susette Haegi for digitizing training data and Dr. Melissa Dawes for professional language editing.

Appendix

Fig. A1. Gini importance from the RF model (based on the five runs of model training and testing) using all predictor variables (S1S2 + EAS). The most important variables were GEMI index (99%) and NIR (B8) (80%), Red-Edge 4 (B8A) (70%), NDVI (60%), MSAVI (55%) and RedEdge3 (B7) (48%) were the less important were the variables aspect and elevation from the DTM and least important were the variables from S1.

Fig. A2. Learning curves of the UNET models based on the three sets of predictor variables S1 + EAS, S2 + EAS and S1S2 + EAS, where E = elevation, A = aspect and S = slope. The solid lines indicate mean loss values, whereas the bounds of the shaded area indicate minimum and maximum of loss values per epoch. Each UNET model was trained five times, and each time training and validation data were newly resampled on a spatial basis (same blocks as for the RF models were used).
Fig. A3. The boxplots are showing the testing performance of the RF and UNET for each of the 24 subareas based on the 5 runs of model training and testing. The central lines within each box are the medians. The boxes’ edges are the upper and lower quartiles. Outliers are depicted as points. OA values are based on the RF model (grey) and the UNET (yellow). The following three sets of predictor variables S1 + EAS, S2 + EAS, and S1S2 + EAS were used, where E = elevation, A = aspect and S = slope. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. A4. Boxplots are showing the K of the RF and UNET for each of the 24 subareas based on the 5 runs of model training and testing. The central lines within each box are the medians. The boxes’ edges are the upper and lower quartiles. Outliers are depicted as points. Weighted K values are based on the three sets of predictor variables with S1 + EAS (grey), S2 + EAS (yellow) and S1S2 + EAS (red), where E = elevation, A = aspect and S = slope. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Fig. A5. DLT map of Switzerland with the two classes broadleaved (orange) and coniferous (green) obtained from RF using the predictor variables S1S2 + EAS. As background the relief of the terrain model swissAlti3D was used. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table A1
Cumulative confusion matrices (based on 5 model runs) of the RF and UNET models using the three sets of predictor variables S1 + EAS, S2 + EAS, S1S2 + EAS for the two classes broadleaved (Brod.) and coniferous (Con.). Given are Reference (Ref.), Prediction (Pred.), Precision (P), Recall (R), Overall accuracy (OA) and Kappa (K).

|        | RF S1 + EAS | RF S2 + EAS | RF S1S2 + EAS |
|--------|-------------|-------------|---------------|
| Pred.  | Ref.        | Pred.       | Ref.          | Pred.       | Ref.        |
| Broad. | 90.3%       | 10.8%       | 0.88          | Broad.      | 95.1%       | 3.4%        | 0.96         | Broad.      | 95.5%       | 3.1%        | 0.97         |
| Con.   | 9.7%        | 89.2%       | 0.91          | Con.        | 4.9%        | 96.6%       | 0.95         | Con.        | 4.5%        | 96.9%       | 0.96         |
| R      | 0.90        | 0.89        |               | R           | 0.97        | 0.97        |              | R           | 0.95        | 0.97        |
| OA     | 0.90        |             | 0.96          | OA          |             | Kappa       | 0.92         |              | Kappa       |             |
| Kappa  | 0.79        |             |               | Kappa       |             |             |              | Kappa       |             |

|        | UNET S1 + EAS | UNET S2 + EAS | UNET S1S2 + EAS |
|--------|---------------|---------------|-----------------|
| Pred.  | Ref.          | Pred.         | Ref.            | Pred.       | Ref.        |
| Broad. | 93.4%         | 5.2%          | 0.94            | Broad.      | 97.9%       | 3.7%        | 0.96         | Broad.      | 97.6%       | 2.9%        | 0.97         |
| Con.   | 6.6%          | 94.8%         | 0.94            | Con.        | 2.1%        | 96.3%       | 0.98         | Con.        | 2.4%        | 97.1%       | 0.98         |
| R      | 0.93          | 0.95          |                | R           | 0.98        | 0.96        |              | R           | 0.98        | 0.97        |
| OA     | 0.94          | OA            | 0.97            | OA          |             | OA           |              | Kappa       |             |
| Kappa  | 0.88          | Kappa         | 0.94            | Kappa       |             |             |

Table A2
OA of the Copernicus DLT 2018 data set compared to the RF and UNET based DLT maps with the three sets of predictor variables S1 + EAS, S2 + EAS and S1S2 + EAS, where E = elevation, A = aspect and S = slope. The comparison is based on independent NFI plot data, that incorporates 3,173 sample plots. The means of OA were obtained per class of proportion of broadleaved, i.e. >80%, 60–80%, 40–60%, 20–40% and < 20% and in total. Bold numbers indicate for which class of proportion of broadleaved the Copernicus DLT 2018 performed better than our approaches with RF and UNET.

| Broadleaved in % | >80 | 60–80 | 40–60 | 20–40 | 0–20 | Total |
|-----------------|-----|-------|-------|-------|------|-------|
| Nr. of NFI plots | 1056 | 389 | 333 | 296 | 1099 | 3173 |
| OA              | 0.70 | 0.17 | 0.10 | 0.28 | 0.71 | 0.54 |
| HRL             | 0.62 | 0.16 | 0.08 | 0.26 | 0.06 | 0.53 |
| RF S1 + EAS     | 0.68 | 0.21 | 0.13 | 0.31 | 0.79 | 0.57 |
| RF S2 + EAS     | 0.68 | 0.21 | 0.13 | 0.31 | 0.78 | 0.57 |
| RF S1S2 + EAS   | 0.65 | 0.21 | 0.11 | 0.28 | 0.79 | 0.55 |
| UNET S1 + EAS   | 0.68 | 0.21 | 0.13 | 0.29 | 0.78 | 0.57 |
| UNET S2 + EAS   | 0.68 | 0.13 | 0.13 | 0.31 | 0.79 | 0.57 |
| UNET S1S2 + EAS | 0.68 | 0.13 | 0.13 | 0.31 | 0.79 | 0.57 |
References

Abegg, M., Brandl, U.-B., Ciolfi, F., Fischer, C., Herold-Bonardit, A., Huber, M., Keller, M., Kelle, R., Röder, A., et al., 2014. Fourth National Forest Inventory—Result Tables and Maps on the Internet for the NFI 2009-2013 (NF14b). Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland.

Barrett, F., McRoberts, R.E., Tomppo, E., Cienciala, E., Waser, L.T., 2016. A questionnaire-based review of the operational use of remotely sensed data by national forest inventories. Remote Sens. Environ. 174, 279–289.

Belgium, Dragut, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31.

Brandl, U., Abegg, M., Allgaier Leuch, B. (Eds.), 2020. Swiss National Forest Inventory: Results of the fourth assessment. Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland, 2010 2009-2017, 341 p.

Breidbach, J., Waser, L.T., Ginzler, C., Hobi, M., Hauglin, M., Putili, S., 2021. 2021. forest type mapping and estimation of forest area by dominant tree species using Sentinel-2 data. Can. J. For. Res. 51 (3), 365-379.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Campos-Taberner, M., García-Haro, F.J., Martínez, B., Sander, C., 2018. Countrywide stereo-image matching for updating digital terrain model at 30 m resolution based on GEDI and background layers (Version 2015). https://land.copernicus.eu/pan-european/high-resolution-layers/forests/leaf-type-2018. Zenodo. http://doi.org/10.5281/zenodo.4057883 (accessed on 5 May 2021).

Cao, K., Zhang, X., 2020. An improved Res-UNet model for tree species classification using airborne high-resolution images. Remote Sens. Environ. 236, 111600.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Measur. 20, 41–45.

Dost, C., Halim, A., Drigger, R., 2019. Urban forest mapping through fusion of different spatial resolution remote sensing images applied to forest-type mapping. IEEE Transaction Geoscience Remote Sens. 49, 4977–4986.

Ewel, J.K., Verduzco, S., Jones, D., 2007. Remote sensing and GIS for forest resource assessment in the five forested regions of Costa Rica. ISPRS J. Photogramm. Remote Sens. 62, 166–177.

MacDicken, K.G., Sola, P., Hall, J., Sabogal, C., Tadoun, M., de Wasseige, C., 2015. Global progress toward sustainable forest management. Forest Ecol. Manag. 352, 47-56.

Mathis, J., Ginzler, C., Zimmermann, N.M., Brasel, P., Woldi, O., 2006. Sensitivity assessment on continuous landscape variables to classify a discrete forest area. Forest Ecol. Manag. 229, 111–119.

McRoberts, R.E., Tomppo, E.O., 2007. Remote sensing support for national forest inventories. Remote Sens. Environ. 110 (4), 416–425.

McRoberts, R.E., 2010. Probability-and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. Remote Sens. Environ. 114 (5), 1017–1025.

Melis, S., Veroult, S., Haywood, A., Jones, S., 2015. Exploring issues of training data imbalance and mislabeling on random forest performance for large area land cover classification using the ensemble margin. ISPRS J. Photogramm. Remote Sens. 105, 155-168.

Minta, M., Olindi, J., Peerbhay, K., Mutanga, O., 2019. Examining the effectiveness of Sentinel-1 and 2 imagery for commercial forest species mapping. Geocarto Int. (12, 4).

Nezami, S., Khorasanchi, E., Nevalainen, O., Polinon, I., Hankavaara, E., 2020. E. Tree species classification of airborne hyperspectral and RGB imagery with Deep Learning. Convolutional Neural Networks. Remote Sens. 12, 1070.

Puhm, M., Deutscher, J., Hirschmugl, M., Wimmer, A., Schmitt, U., Schardt, M., 2020. Puletti, N., Chianucci, F., Castaldi, C., 2018. Use of Sentinel-2 for forest classification in urban and suburban areas in Northern Italy. ISPRS Journal of Photogrammetry and Remote Sensing 142, 185-198.

R Core Team, 2021. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria; Available online: https://www.r-project.org/ (accessed on 14 May 2021).

Rehush, N., Abegg, M., Waser, L.T., Brandl, U.-B., 2018. Identifying tree-related bark beetle outbreaks. Front. Plant Sci. 10, 1327.

Reichert, K., Atzberger, C., Hank, T.B., Mauer, W., 2012. Derivation of biophysical variables from earth observation data: Validation and statistical measures. J. Appl. Earth Obs. Geo. 90, 102116.

Richter, K., Atzberger, C., Hank, T.B., Mauser, W., 2012. Derivation of biophysical variables from earth observation data: Validation and statistical measures. J. Appl. Earth Obs. Geo. 90, 102116.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Kéry, M., 2016. Spatial prediction dynamics at the data’s true extent from digital aerial photographic inventories. Int. J. Appl. Earth Obs. Geoinf. 4986–444.

R Core Team, 2021. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria; Available online: https://www.r-project.org/ (accessed on 14 May 2021).

R Core Team, 2021. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria; Available online: https://www.r-project.org/ (accessed on 14 May 2021).

R Core Team, 2021. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria; Available online: https://www.r-project.org/ (accessed on 14 May 2021).

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Kéry, M., 2016. Spatial prediction dynamics at the data’s true extent from digital aerial photographic inventories. Int. J. Appl. Earth Obs. Geoinf. 4986–444.
Rüetschi, M., Schaepman, M.E., Small, D., 2018. Using Multitemporal Sentinel-1 C-band Backscatter to Monitor Phenology and Classify Deciduous and Coniferous Forests in Northern Switzerland. Remote Sens. 10, 55.

Rüetschi, M.; Small, D.; Waser, L.T., 2019. Rapid Detection of Windthrows Using Sentinel-1 C-Band SAR Data. Remote Sens. 11, 115.

Satir, O., Berberoglu, S., Akca, E., Yeler, O., 2017. Mapping the dominant tree distribution using a combined image classification approach in a complex Eastern Mediterranean basin. J. Spatial Sci. 62 (1), 157–171.

Schmidtuber, J., 2015. Deep learning in neural networks: An overview. Neural Networks 61, 85–117.

Small, D., 2011. Flattening Gamma: Radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Rem. Sens. 49, 3081–3093.

Small, D., 2012. SAR backscatter multitemporal compositing via local resolution weighting. In International Geoscience and Remote Sensing Symposium (IGARSS); IEEE: Munich, Germany, 2012, pp. 4521–4524.

Small, D., Rohner, C., Miranda, N., Rüetschi, M., Schaepman, M.E., 2021. Wide-area analysis ready radar backscatter composites. IEEE Trans. Geosci. Rem. Sens. 59 https://doi.org/10.1109/TGRS.2021.3055562.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., Brenning, A., 2019. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. Ecol. Model. 406, 109–120.

Strahler, A.H., Logan, T.L., Bryant, N.A., 1978. Improving forest cover classification accuracy from Landsat by incorporating topographic information, in: Proceedings of the 12th International Symposium on Remote Sensing of Environment. Ann Arbor, Michigan, pp. 927-942.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B.O., Floury, N., Brown, M., Traver, I.N., Deghaye, P., Doesmann, B., Rossich, B., Miranda, N., Bruno, C., L’Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24.

Van Rossum, G., Drake, F.L., 2001 (eds.). Python Reference Manual, PythonLabs, Virginia, USA. https://www.python.org (accessed on 5 May 2021).

Vidal, C., Albérdi, J.A., Hernández Mateo, L., Redmond, J.J., 2016 (eds.). National Forest Inventories - Assessment of Wood Availability and Use. Cham, Springer, 783-805.

Vihervaara, P., Auvinen, A.P., Mononen, L., Torma, M., Ahlroth, P., Anttila, S., Böttcher, K., Forsius, M., Heino, J., Helioä, J., et al., 2017. How essential biodiversity variables and remote sensing can help national biodiversity monitoring. Glob. Ecol. Conserv. 10, 43–59.

Waser, L.T., Küchler, M., Jütte, K., Stampfer, T., 2014. Evaluating the Potential of WorldView-2 Data to Classify Tree Species and Different Levels of Ash Mortality. Remote Sens. 6, 4515-4545.

Waser, L.T., Fischer, C., Wang, Z., Ginzler, C., 2015. Wall-to-wall forest mapping based on digital surface models from image-based point clouds and a forest definition. Remote Sens. 7, 766.

Wessel, M., Brandmeier, M., Tiede, D., 2018. Evaluation of Different Machine Learning Algorithms for Scalable Classification of Tree Types and Tree Species Based on Sentinel-2 Data. Remote Sens. 10, 1419.

White, J.C., Coops, N.C., Wulder, M.A., Vastaranta, M., Hilker, T., Tomalski, P., 2016. Remote sensing technologies for enhancing forest inventories: A review. Can. J. Remote Sens. 42, 619–641.

Wilson, J.P., Gallant, J.C., 2000. Terrain Analysis: Principles and Applications; John Wiley and Sons: Hoboken, NJ, USA.

Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the FMask algorithm: Cloud, cloud shadow, and snow detection for LANDSATs 4–7, 8, and Sentinel-2 images. Remote Sens. Environ. 159, 269–277.

Zhu, X., Tuia, D., Mou, L., Xia, G-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci. Rem. Sens. M. 5 (4), 8–36,