

**This is an electronic reprint of the original article.**

**This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Rasmus Erlandsson, Jarle W. Bjerke, Eirik A. Finne, Ranga B. Myneni, Shilong Piao, Xuhui Wang, Tarmo Virtanen, Aleksi Räsänen, Timo Kumpula, Tiina H.M. Kolari, Teemu Tahvanainen, Hans Tømmervik

**Title:** An artificial intelligence approach to remotely assess pale lichen biomass

**Year:** 2022

**Version:** Publisher's version

**Copyright:** The author(s) 2022

**Rights:** CC BY 4.0

**Rights url:** <https://creativecommons.org/licenses/by/4.0/>

**Please cite the original version:**

Rasmus Erlandsson, Jarle W. Bjerke, Eirik A. Finne, Ranga B. Myneni, Shilong Piao, Xuhui Wang, Tarmo Virtanen, Aleksi Räsänen, Timo Kumpula, Tiina H.M. Kolari, Teemu Tahvanainen, Hans Tømmervik, 2022. An artificial intelligence approach to remotely assess pale lichen biomass . 280, 113201.

<https://doi.org/10.1016/j.rse.2022.113201>

All material supplied via *Jukuri* is protected by copyright and other intellectual property rights. Duplication or sale, in electronic or print form, of any part of the repository collections is prohibited. Making electronic or print copies of the material is permitted only for your own personal use or for educational purposes. For other purposes, this article may be used in accordance with the publisher's terms. There may be differences between this version and the publisher's version. You are advised to cite the publisher's version.



Contents lists available at ScienceDirect

## Remote Sensing of Environment

journal homepage: [www.elsevier.com/locate/rse](http://www.elsevier.com/locate/rse)

## An artificial intelligence approach to remotely assess pale lichen biomass

Rasmus Erlandsson<sup>a,\*</sup>, Jarle W. Bjerke<sup>a</sup>, Eirik A. Finne<sup>a,b</sup>, Ranga B. Myneni<sup>c</sup>, Shilong Piao<sup>d</sup>, Xuhui Wang<sup>d</sup>, Tarmo Virtanen<sup>e</sup>, Alekski Räsänen<sup>e,f</sup>, Timo Kumpula<sup>g</sup>, Tiina H.M. Kolari<sup>h</sup>, Teemu Tahvanainen<sup>h</sup>, Hans Tømmervik<sup>a</sup>

<sup>a</sup> Norwegian Institute for Nature Research, FRAM – High North Research Centre for Climate and the Environment, NO-9296 Tromsø, Norway

<sup>b</sup> Department of Geosciences, University of Oslo, Postboks 1047 Blindern, NO-0316 Oslo, Norway

<sup>c</sup> Department of Earth and Environment, Boston University, 685 Commonwealth Avenue, Boston, MA US-02215, USA

<sup>d</sup> Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Shaw Building No.2, 5 Yiheyuan Road, Haidian District, CH-100871 Beijing, China

<sup>e</sup> Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, and Helsinki Institute of Sustainability Science (HELSUS), University of Helsinki, Yliopistonkatu 3, FI-00014, Finland

<sup>f</sup> Natural Resources Institute Finland (Luke), Paavo Havaksentie 3, FI-90570 Oulu, Finland

<sup>g</sup> Department of Geographical and Historical Studies, University of Eastern Finland, Joensuu Campus, PO Box 111, FI-80101 Joensuu, Finland

<sup>h</sup> Department of Environmental and Biological Sciences, University of Eastern Finland, Joensuu Campus, PO Box 111, FI-80101 Joensuu, Finland

## ARTICLE INFO

Edited by Dr. Marie Weiss

## Keywords:

Remote sensing  
Lichens  
Terricolous lichens  
Deep neural networks  
Artificial intelligence  
cladonia  
Reindeer lichen  
Light lichens  
Light coloured lichens  
Pale lichens  
Landsat

## ABSTRACT

Although generally given little attention in vegetation studies, ground-dwelling (terricolous) lichens are major contributors to overall carbon and nitrogen cycling, albedo, biodiversity and biomass in many high-latitude ecosystems. Changes in biomass of mat-forming pale lichens have the potential to affect vegetation, fauna, climate and human activities including reindeer husbandry. Lichens have a complex spectral signature and terricolous lichens have limited growth height, often growing in mixtures with taller vegetation. This has, so far, prevented the development of remote sensing techniques to accurately assess lichen biomass, which would be a powerful tool in ecosystem and ecological research and rangeland management. We present a Landsat based remote sensing model developed using deep neural networks, trained with 8914 field records of lichen volume collected for >20 years. In contrast to earlier proposed machine learning and regression methods for lichens, our model exploited the ability of neural networks to handle mixed spatial resolution input. We trained candidate models using input of  $1 \times 1$  ( $30 \times 30$  m) and  $3 \times 3$  Landsat pixels based on 7 reflective bands and 3 indices, combined with a 10 m spatial resolution digital elevation model. We normalised elevation data locally for each plot to remove the region-specific variation, while maintaining informative local variation in topography. The final model predicted lichen volume in an evaluation set ( $n = 159$ ) reaching an  $R^2$  of 0.57. NDVI and elevation were the most important predictors, followed by the green band. Even with moderate tree cover density, the model was efficient, offering a considerable improvement compared to earlier methods based on specific reflectance. The model was in principle trained on data from Scandinavia, but when applied to sites in North America and Russia, the predictions of the model corresponded well with our visual interpretations of lichen abundance. We also accurately quantified a recent historic (35 years) change in lichen abundance in northern Norway. This new method enables further spatial and temporal studies of variation and changes in lichen biomass related to multiple research questions as well as rangeland management and economic and cultural ecosystem services. Combined with information on changes in drivers such as climate, land use and management, and air pollution, our model can be used to provide accurate estimates of ecosystem changes and to improve vegetation-climate models by including pale lichens.

\* Corresponding author.

E-mail addresses: [rasmus.erlandsson@nina.no](mailto:rasmus.erlandsson@nina.no) (R. Erlandsson), [jarle.bjerke@nina.no](mailto:jarle.bjerke@nina.no) (J.W. Bjerke), [eirik.finne@nina.no](mailto:eirik.finne@nina.no) (E.A. Finne), [slpiao@pku.edu.cn](mailto:slpiao@pku.edu.cn) (S. Piao), [xuhui.wang@pku.edu.cn](mailto:xuhui.wang@pku.edu.cn) (X. Wang), [tarmo.virtanen@helsinki.fi](mailto:tarmo.virtanen@helsinki.fi) (T. Virtanen), [aleksi.rasanen@luke.fi](mailto:aleksi.rasanen@luke.fi) (A. Räsänen), [timo.kumpula@uef.fi](mailto:timo.kumpula@uef.fi) (T. Kumpula), [tiina.kolari@uef.fi](mailto:tiina.kolari@uef.fi) (T.H.M. Kolari), [teemu.tahvanainen@uef.fi](mailto:teemu.tahvanainen@uef.fi) (T. Tahvanainen), [hans.tommervik@nina.no](mailto:hans.tommervik@nina.no) (H. Tømmervik).

<https://doi.org/10.1016/j.rse.2022.113201>

Received 4 February 2022; Received in revised form 25 July 2022; Accepted 26 July 2022

Available online 9 August 2022

0034-4257/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. Background

While satellite-based methods for remote sensing of green vascular plants have been the most important source for monitoring global environmental change (e.g. Myneni et al., 1997; Gould, 2000; Piao et al., 2020; Ryu et al., 2019), reliable methods for lichens have been elusive (Falldorf et al., 2014; Nordberg and Allard, 2002). Lichens are symbiotic associations of fungi and algae and/or cyanobacteria characterized by low requirements of nutrient and low growth rates that enable them to pioneer in the colonization of exposed surfaces (Cutler, 2010). They can remain prominent in later succession stages, it is calculated that 6–8% of the Earth land surface is covered and dominated by lichens (Crittenden, 2000; Larson, 1987), and they may colonize nearly all terrestrial habitats while being functionally important in many ecosystems (Kappen, 1988). Especially in Arctic, alpine, boreal and dryland ecosystems lichens are major contributors to overall carbon and nitrogen cycling, biodiversity and biomass (Cutler, 2010; Elbert et al., 2012).

Terricolous (ground-dwelling) fruticose lichens, typically forming 10–20 cm thick mats in undisturbed regions, have the potential to be the dominating growth form in Arctic and alpine tundra, in open boreal forests, and on the drier parts of peatlands (Ahti and Hepburn, 1967; Joly et al., 2009; Rautiainen et al., 2007; Tømmervik et al., 2012). Their role as a vital winter food resource for reindeer/caribou (*Rangifer tarandus*, hereafter *Rangifer*) makes them economically and culturally important in areas with semi-domestic or wild populations of *Rangifer* (Joly et al., 2009; Riseth et al., 2016). Due to their wide distribution in the Arctic and boreal regions of Eurasia and North America, the area covered by reindeer husbandry alone amounts to some four million square kilometres (Jernsletten and Klokov, 2002; Tømmervik et al., 2012). *Rangifer* have impact on climate through their modification of vegetation structure by trampling and selective grazing (Cohen et al., 2013; Collins et al., 2011; Joly et al., 2010; Tømmervik et al., 2004).

While green plants have a relatively distinct reflective signal in the red and near-infrared bands of the light spectrum, lichens have considerably complex spectral signatures (Rees, 2004; Solheim et al., 2000). Notably, terricolous lichens consist of species both with melanised dark surfaces and pale surfaces with no or little melanin (Gauslaa, 1984), and these often grow in mixtures. Light reflectance properties are also affected by thallus humidity (Granlund et al., 2018) and growth form (crustose, foliose or fruticose). Another complicating factor is that terricolous lichens have limited growth height (maximum 25 cm), and they often grow in communities with a mixed cover of lichens, bryophytes and taller vascular plants, especially graminoids, shrubs and trees, that can partly shade lichens or hide them from optical sensors (Fraser et al., 2014; Tømmervik et al., 2004). Remote sensing assessment methods of lichen abundance should therefore ideally be able to detect lichens despite noise from varying vegetation structures. Simple arithmetic equations based on satellite reflectance values are typically sensitive to that type of noise (Nordberg and Allard, 2002), hence such methods should be used with caution, limiting their applicability to specific areas or habitats. This is for example the case with the Lichen Volume Estimator (LVE) model developed for pale lichens by Falldorf et al. (2014) and later applied by Rickbeil et al. (2017) and Macander et al. (2018).

Recently, methods to detect pale terricolous lichens based on machine learning have started to emerge (Fraser et al., 2021; He et al., 2021; Jozdani et al., 2021; Kennedy et al., 2020; Macander et al., 2020). Artificial intelligence-based approaches have several benefits since the combined information from all included bands are used without any a priori assumptions on their relationship to lichen reflectance. The blunt naivety of such an uninformed approach, compared to the more informed decisions when constructing a traditional equation, pays off in the ability to find context-dependent relationships (Brodrick et al., 2019). This allows detection algorithms of the lichen signal to vary

across the landscape, for example along gradients from open tundra to boreal taiga. In this way, the approach resembles the ability of the human brain to distinguish a relevant signal despite different types of noise (Fraser et al., 2021; Jozdani et al., 2021), however with the potential to process a large number of bands (Kennedy et al., 2020). In addition to reflectance, it is possible to include any additional rasterised data, such as elevation, emphasising the potential of artificial intelligence to efficiently combine different types of geographical information to solve complex remote sensing tasks (Fraser et al., 2021; Jozdani et al., 2021; Kennedy et al., 2020).

Jozdani et al. (2021) and Fraser et al. (2021) investigated the ability to train computers on extremely high-resolution (< 3.5 cm) images taken with unmanned aerial vehicles (drones) and changed scale to high-resolution satellites (0.5–6 m). While the method of Jozdani et al. (2021) was limited to a binary classification of terricolous lichens in general; Fraser et al. (2021) assessed the ability of a random forest model to quantify cover of pale, fruticose lichens of the genus *Cladonia*, reaching  $R^2$  values from 0.07 to 0.49 for the *Planet* satellites, and 0.52 to 0.70 for the *World View* satellites, both at 6 m resolution. Although both studies showed the versatility and promising aspects of machine learning techniques to scale from local to regional scope, high resolution images are still often expensive, tend to have a limited availability, and low coverage or number of spectral bands. Not least, they have a very limited historical record, making such images unsuitable for long-term studies, and inapplicable with field records that are more than a few years old.

In contrast, Kennedy et al. (2020) presented a promising methodology using deep learning to train a neural network for assessment of lichen cover based on Landsat images, elevation data and climatic parameters. Kennedy et al. (2020) assessed pale lichen cover at three different high-latitude sites spanning the North American continent, achieving  $R^2$  values ranging from 0.27 to 0.55. This is a pioneering application of AI to remote sensing of terrestrial lichen cover, demonstrating the high potential of the technique. However, correlation coefficients were rather low in several areas, while lichen cover remains a rather crude measurement of actual lichen biomass. Although biomass is a function of cover and thickness, the correlation with cover is non-linear and highly sensitive to reindeer grazing pressure. In particular, when lichens are recovering after prolonged grazing, the cover can be substantial while the biomass is still very low due to rapid horizontal expansion from numerous lichen fragments while vertical growth is a slower process (Klein and Shulski, 2009). Thus volume is a more relevant ecological measure since it is closely correlated to biomass (Moen et al., 2007; Tømmervik et al., 2004; Tømmervik et al., 2012). Biomass is also socioeconomically relevant for reindeer herding and hunting-based communities as it quantifies the amount of food available for *Rangifer*, and lichen biomass can affect the ability of other vegetation to establish, since thick lichen mats can prevent vascular plant seeds from reaching the soil (Sedia and Ehrenfeld, 2003; Tømmervik et al., 2009). Furthermore, modellers of global ecosystem dynamics typically rely on biomass rather than cover. In addition, a thick layer of pale lichens would have a higher albedo compared to a thinner layer with the same cover, potentially affecting local climate (Stoy et al., 2012). However, while lichen cover can efficiently be obtained by using, for example, aerial photos or drone images, large datasets with lichen volume are tedious to collect since the thickness of the lichen cover must be measured in the field. As field data collection of representative lichen volumes is time-consuming and expensive, an accurate remote sensing model for estimation of lichen volumes would be a powerful tool in ecological research and rangeland management.

### 1.2. Objectives

We assembled a data set of 8914 field records collected over a period of 23 years from both open and forested lichen habitats covering a gradient from no or negligible grazing conditions to high-grazing

conditions. This dataset was used to address the following objectives: (1) To develop and evaluate a neural network-based method for assessment of pale lichen volume relying solely on reflectance and topography data and apply the model in selected circumpolar sites to evaluate the general validity of the model. (2) To test the ability of the model to detect a documented long-term change in lichen abundance.

## 2. Methods

We trained neural networks with Landsat images and ground-truth data collected over 20 years in northern Scandinavia and the Kola peninsula. In contrast to earlier proposed methods, we did not restrict the input data to single pixel values only, but exploited the ability of neural networks to handle mixed resolution input.

Although it is technically possible to feed virtually any type of geographic information to our neural network, we wanted to reduce complexity and limited the input data to Landsat reflectance and elevation data, since those datasets are easily retrievable and in high spatial resolution. To be able to utilise the historical records of Landsat images we only used bands that are similar across the Landsat 5, 7 and 8 satellites, allowing the method to be applied back to the 1980s.

### 2.1. Lichen abundance data

The ground-truth data we retrieved originated from different monitoring projects undertaken in Norway, Sweden, Finland and the Kola peninsula, Russia (Table 1, Fig. 1), and therefore, they are subject to some variation in methods and within-plot sampling regimes. The distribution of plots within sampling sites varied as well from random to systematic transects, depending on the original purpose of the sampling. All samples were based on 1 m<sup>2</sup> field plots except the data from Finland which was sampled from 0.25 m<sup>2</sup>. Values from fields plots that fell within the 30 m Landsat pixel size, were averaged (2–5 depending on area) and used as single values for training, evaluation and testing. A special case was made for the plots from west Finnmark that were located 50 m apart. Those plots were used as individual plots for training, to maintain a larger training set, but the mean was used for testing as the mean value, although sampled from adjacent Landsat pixels, more accurately reflected the general lichen biomass within the target pixel. All plots (n = 8914) were standardised to volume in litres (dm<sup>3</sup>) per m<sup>2</sup> of pale terricolous fruticose lichens mainly belonging to the two genera *Cladonia* and *Flavocetraria*. In 21 plots, lichen cover was

>100% reaching a maximum of 126% because of 2–3 layers of different lichen species, where the upper layers shade the lower ones. Vegetation cover measurements summing to over 100% are common in vegetation studies (Fehmi, 2010), and those records did not result in outlier volume measurements. Table S2 provides an overview of the main types of vegetation covered by the AI analysis in this study, following the concept developed by Fremstad (1997), which relies on ecological gradients in moisture, nutrient availability, soil type, topography, and climate. These types are dwarf shrub and lichen heaths, bilberry-lichen heath, alpine heather-lichen heath, lichen-birch woodland, lichen-pine woodland, and lichen-rich bog with heather and dwarf shrub.

Both a training and a validation set are needed to train neural networks. For an unbiased test of the model performance, an independent test set is also needed. We therefore randomly divided the dataset into 3 datasets: training (n = 3290, 73%), validation (n = 1054, 23%) and test (n = 159, 4%). Spatially overlapping records (within 300 m clusters) were allowed in the training and validation set, but plots that were closer than 300 m were removed from the test set in order not to bias the results (regarding the distance, see discussion below). For the test set, averaged values were used for multiple data points collected within a single or adjacent Landsat pixel.

We also rendered a ‘remote’ dataset, to investigate if boosting the training set with remotely assessed data could improve the models. Across an area of 280 km<sup>2</sup> in Central Norway, we classified a total area of 6.3 km<sup>2</sup> into 4 land-cover types based on lichen abundance through visual interpretation of aerial photography (0.5 m resolution). The categories were: a) vascular plants without lichens, b) bare ground with negligible amounts of lichens, c) intermediate lichen cover, d) lichen-dominated vegetation. Within the classified areas, we randomly placed 1734 points to which we assigned values for lichen cover and height corresponding to the average empirical values we measured in the field at comparable sites. To render variation within the assigned values, they were randomly picked from a class-specific distribution. The remotely assessed dataset was solely used as training data.

### 2.2. Image acquisition and pre-processing

The workflow described below is also graphically presented in a methodological flowchart (Fig. 2).

#### 2.2.1. Satellite images

Atmospherically corrected surface reflectance Landsat collection 1,

**Table 1**

Size of datasets used for training. Data points with high lichen volume were heavily underrepresented and therefore repeated in the training set to achieve a more even distribution (numbers given within parentheses).

Set	Finnmark	Hardanger	Røros	Nikel (RUS)	Teriberka	Väster-	Andøya	Lapland	Forolhogna	Sogn	Remote data
	(NO)	(NO)	(NO)	Pasvik (NO)	(RUS)	botten (SE)	(NO)	(FIN)	(NO)	(NO)	(NO)
	825 (+177)										
Core data		1382 (+903)	207 (+856)	9	24 (+66)	884 (+283)	49 (+4)				
Mixed sources	19	82	1	2	17			303 (+94)	65 (+153)	475 (+544)	
Remotely assessed											1734 (+6449)



**Fig. 1.** Map of the field plots used to train and test the deep neural network. Plots within a 1 km distance from each other was aggregated for readability. For sample sizes, see [Table 1](#).

Tier 1 images (Landsat 5, 7 or 8 depending on availability, U.S. Geological Survey) were acquired using Google Earth Engine ([Gorelick et al., 2017](#)). For each field site and year that data was recorded, a collection of all images taken between July 1st and August 31st was masked for clouds, cloud shadows and snow using the embedded quality assessment band (CFMASK-derived, see [Foga et al., 2017](#)). A composite was then formed using median values. In some years and locations, clouds completely prevented the compilation of a useful composite. In those cases, we instead used composite from the closest possible adjacent year with enough coverage ([Table 2](#)). An alternative method to handle atmospheric noise and other distortions is to use composites averaged from multiple years. However, this method is sensitive to vegetation changes between years. We decided not to use multi-year averages, since reindeer grazing impact in some of the field sites often vary considerably between years. Hence, multi-year averages would possibly add, and not reduce, noise.

In addition to the 7 bands that Landsat 5, 7 and 8 have in common, we included 3 normalised differentiated indices that have been used in previous efforts to sense lichens remotely: Normalised Differentiated Vegetation Index (NDVI, [Tucker, 1979](#), bands: NIR and red), ND Lichen Index (NDLI, [Nordberg, 1998](#), bands: SWIR1 and green), and ND Moisture Index (NDMI, [Wilson and Sader, 2002](#), bands: NIR and SWIR1). All pixel values were scaled to 8-bit values (0–255) in Google Earth Engine prior to export to facilitate an accurate conversion from Tag Image File Format (TIFF) to 8-bit Portable Network Graphics (PNG), which was the image file format supported by the artificial intelligence platform (Peltarion AB, Stockholm, Sweden) used in this study. For the scaling, the highest recorded reflectance value (0.4785) occurring in our training set was used as 255. Since the normalised indices range between  $-1$  and  $+1$ , those values were stretched to values between 0 and 255 separately.

### 2.2.2. Elevation

Terricolous fruticose lichens at northern latitudes have the capability to form large mats from sea level to mountain summits, with the same species being abundant along the entire bioclimatic gradient ([Crittenden, 2000](#); [Hein et al., 2021](#); [Tømmervik et al., 2009](#)). Hence, altitude per se has a limited effect on lichen growth in the areas covered by this study and the study sites with the highest lichen abundance were found

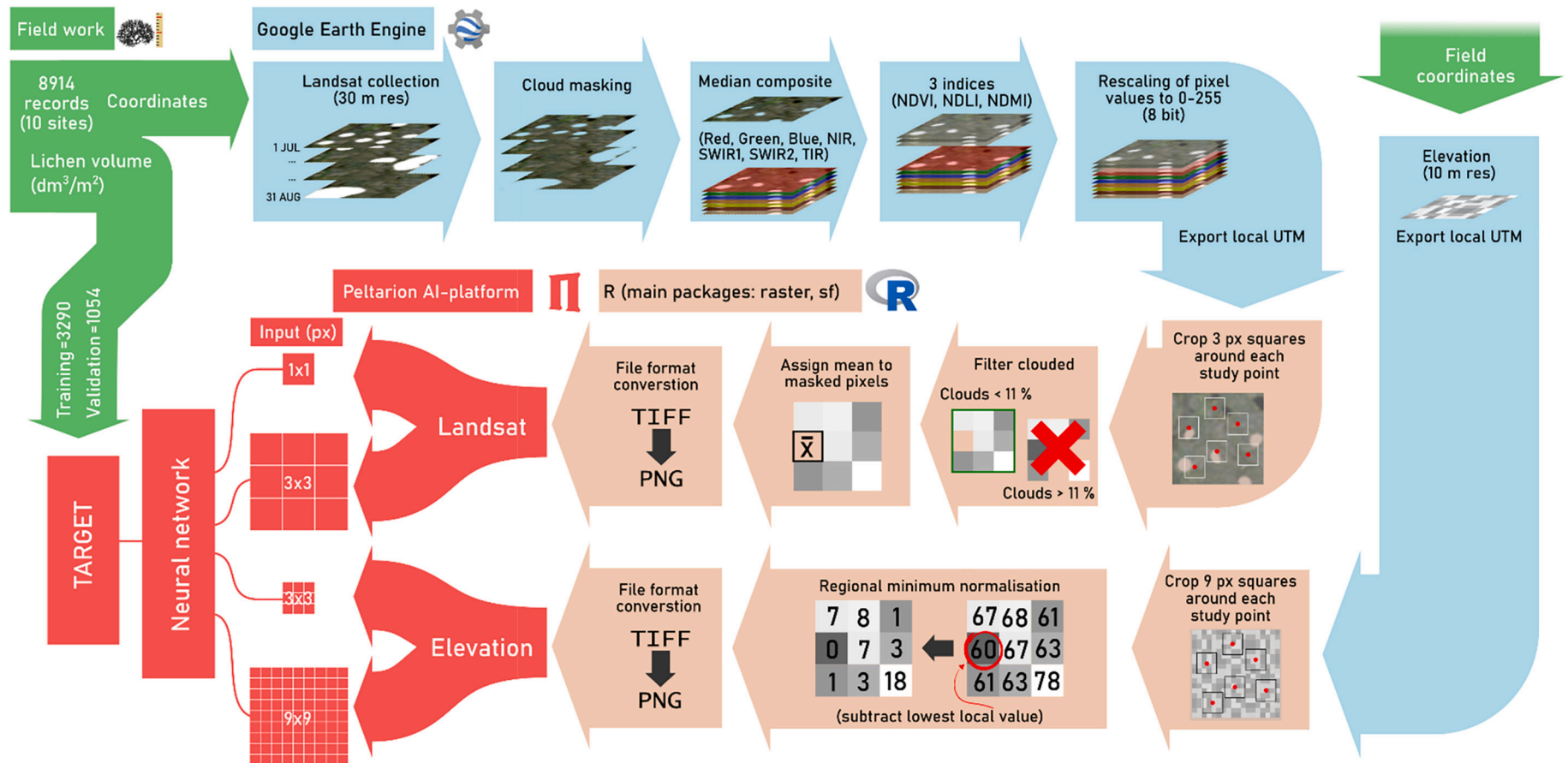
in upland areas with little reindeer grazing and low impacts of other land-use activities. Since this was largely an effect of variation in extent of reindeer husbandry, rather than in ecological processes relevant for lichen growth, there was a risk that absolute altitude could be exploited by the neural network to identify an artificial relationship between altitude and lichen abundance, limiting the generality of the model. However, local topography variation has strong impact on snow accumulation, which in turn can affect reindeer access to lichen forage resources ([Tømmervik et al., 2012](#)). By normalising the elevation data for each plot, we removed the region-specific variation, while maintaining informative local variation in topography. (The regional minimum value within a square of 3870 m (387 pixels) was subtracted from all pixels, resulting in elevation models ranging from 0 to the locally highest point.)

Elevation data was either acquired from the Arctic Digital Elevation Model (2 m resolution, [Porter et al., 2018](#)) or, for Norwegian sites, a 2 m resolution digital elevation model produced by the Norwegian Mapping Authority (“Kartverkets digitale terrenmodell”). The resolution was reduced to  $10 \text{ m} \times 10 \text{ m}$  using Google Earth Engine and default settings (NN: nearest neighbour). All images were exported as GeoTIFF raster files in the format of their local Universal Transversal Mercator (UTM) zone.

### 2.3. Raster handling

During the exploratory phase of the method development, we experimented with different cropping sizes of the input rasters before we settled on  $90 \times 90 \text{ m}$ . While for all applications (maps) of the model we used a  $90 \times 90 \text{ m}$  cropping straight away, the rasters for the training and test data were constructed using a larger crop size of 129 Landsat pixels.

For each Landsat-derived raster and data point, we cropped a square of  $3870 \times 3870 \text{ m}$  (Landsat:  $129 \times 129 \text{ px}$ , elevation data:  $387 \times 387 \text{ px}$ ) with the sampling coordinate in the centre. If 11% or more of the area was masked due to clouds, the data point was dropped entirely. Otherwise, any masked pixels were assigned with the mean value of the unmasked pixels as this could be assumed to render least noise for the artificial intelligence modelling. Rasters were handled in R ([R Core Team, 2020](#)) using packages *raster* ([Hijmans and van Etten, 2012](#)), *terra* ([Hijmans, 2021](#)), *sp* ([Bivand et al., 2013](#)) and *sf* ([Pebesma, 2018](#)).



**Fig. 2.** Methodological flow chart. Field work steps are shown with green background, while processing in Google Earth Engine, R, and the Peltarion AI-platform are shown in blue, beige, and red backgrounds, respectively. In Google Earth Engine, we compiled cloud-free median composites of all 1st tier surface reflectance Landsat images taken between 1st July and 31st August. We extracted the 7 bands that are in common for Landsat 5, 7 and 8 and calculated three spectral indices. Field records within pixel distance were averaged, resulting in the discrepancy between number of actual filed records and training/validation sets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 2**  
Table over the test data set.

Finnmark (NO)	Hardanger (NO)	Nikel (RUS) Pasvik (NO)	Røros (NO)	Västerbotten (SE)
13	108	1	6	31

#### 2.4. Structuring of data for training neural network models

We subdivided the datasets into 3 different categories (Table 1). The *Core dataset* consisted of data collected in the field by one or several of the authors or closely affiliated monitoring programs. The *Mixed sources* dataset was retrieved from external collaborators and collected following more diverse methods and with different levels of detail. The *Remote* dataset was compiled from aerial photos as described above.

We transferred 112 data points from the core dataset to the mixed sources dataset (Table 1), because these plots had comparatively high lichen volume-to-cover ratios (thick but patchy lichen cover, which occasionally occurs in forested systems with little reindeer grazing but high competition from shrubs) and could potentially add noise rather than contribute to an efficient training.

For each area where a dataset was collected, we randomly assigned 80% of the data for training and 20% for validation. Prior to this division, to avoid spatial overlap between training and evaluation sets, all data points that fell within a 300 m radius of any other point were grouped into clusters. In the assignment to the different set classes, training and evaluation points were never drawn from the same spatial clusters.

Before we trained the neural networks, we also extracted a random test-set from the Core training set. This set was completely removed prior to any training, including evaluation during training. To avoid spatial overlap between the test and the training data, the points in the test set were not allowed to come from the same spatial cluster as any data points in the training or validation set (similarly to the division of the training and test set). A recent debate has emerged regarding the implications of spatial autocorrelation when validating spatial predictions (Ploton et al., 2020; Wadoux et al., 2021). On the one hand, Ploton et al. (2020) argue that test data should be spatially independent from the training data. To test the spatial dependency, they propose to use variogram analysis to find out a threshold distance after which there is little spatial autocorrelation in the response variable. They suggest using such spatial cross-validation techniques in which test data points are not located within the spatial threshold distance from each training data point. On the other hand, Wadoux et al. (2021) argue that a random selection of points for evaluation should be used instead of spatial or traditional cross-validation, and that spatial cross-validation methods result in overly pessimistic assessment of model fit.

To assess the influence of spatial autocorrelation of our data set, we ran a spatial random forest analysis to analyse how well the lichen biomass in the test set could be predicted solely with the distance between data points (Hengl et al., 2018). To assess the scale of spatial autocorrelation, we constructed a variogram following the method described in Ploton et al. (2020). A distance sufficient to effectively handle the inherent spatial autocorrelation of the data set (at least 43 km, as indicated by a variogram analysis Supplementary Fig. S4) would render training impossible as it results in too few clusters per study site. The choice of clustering distance is a trade-off between spatial independence and having training sets in all study areas. Given that we were unable to compensate for the spatial autocorrelation, we settled for a clustering distance of 300 m to achieve a practical number of clusters while maintaining a minimum of spatial partitioning.

Data points with high lichen volumes were heavily under-represented. This can be problematic since each iteration of the training works with a subset of the complete training data, and a skewed representation of records is likely to result in few or no records of under-represented values in many training iterations. In our case, this would

result in biased models that were not sufficiently trained to predict high lichen volumes. A way to limit this effect is to repeat less represented values in the training set and artificially achieve a more even distribution. We binned the values into 19 volume bins of equal intervals and repeated values in the training set belonging to the 17 highest categories (repeating values from 25 to 140  $\text{dm}^3/\text{m}^2$ ) up to a maximum of 600 times (Fig. S1). Those values were repeated as is, without any additional noise.

#### 2.5. Training of neural network models

After exploratory testing, we chose a model architecture (Fig. 3) that was used in all models. Landsat pixels (30 m) and elevation pixels (10 m) were given separate inputs due to the difference in spatial resolution. In addition to input values of a single Landsat pixel ( $1 \times 1$  Landsat +  $3 \times 3$  elevation), we also included an input of  $3 \times 3$  Landsat pixels ( $9 \times 9$  elevation pixels, Fig. 3). The models were trained using a batch-size of 128, Adam optimizer, learning rate of 0.0005 and patience of 20. The final model discussed in detail below was trained for 33 epochs, reaching a plateau after 7 epochs. Epoch 13 was ranked best according to the training platform and thus used in the subsequent analyses.

#### 2.6. Model evaluation

To evaluate model fit, we fitted a linear model with the empirical lichen volume of the test set as explanatory variable and the AI-model output as the response variable to test the fit of each candidate model. The candidate models were compared based on intercept, coefficient of determination ( $R^2$ ),  $p$ -values and Root mean square error (RMSE). The  $R^2$  value we present is based on the comparison between the predicted lichen volume and the empirical field measurement. For the sake of comparison, however, we also present the  $R^2$  value derived from the linear regression, as this method might have been used in previous studies. To compare our model with the LVE (Falldorf et al., 2014) we applied the algorithm using Google Earth Engine on the same Landsat composites that we used for the AI analysis. We also calculated Bias and Mean Absolute Error (MAE) using the Metrics package (Hamner and Frasco, 2018) and  $R^2$  using the MLmetrics package (Yan, 2016).

Hengl et al. (2018) suggest that random forest can be used as a generic framework for spatial interpolation and to test to what degree the response variable is spatially autocorrelated. Following their method, we constructed a random forest (150 trees) with measured lichen volume as the dependent variable and buffer distance from other plots as independent variable (Hengl et al., 2018).

#### 2.7. Importance of variables

One downside of deep neural networks is that the training process essentially is a black box offering little insight compared to more traditional methods, and the model gives no straight forward coefficients for the different predictors. Therefore, we did a permutation test to evaluate the explanatory contribution of each predictor variable. The idea is to compare how the model efficiency is affected when one of the input variables is permuted (randomly shuffled among the data points). A small decrease in model efficiency indicates that the permuted predictor contributes little to the predictive power of the model, and vice versa. We made 11 copies of the original data set, one for each predictor (7 bands, 3 indices and elevation), and permuted the input rasters for a single predictor in each copy. The resulting 11 runs (one for each predictor variable) were then compared to the output of the original test set, using the Nash–Sutcliffe model efficiency coefficient (Nash and Sutcliffe, 1970; Zambrano-Bigiarini, 2020).

#### 2.8. Applicability

The Finnmark region in northern Norway has experienced a

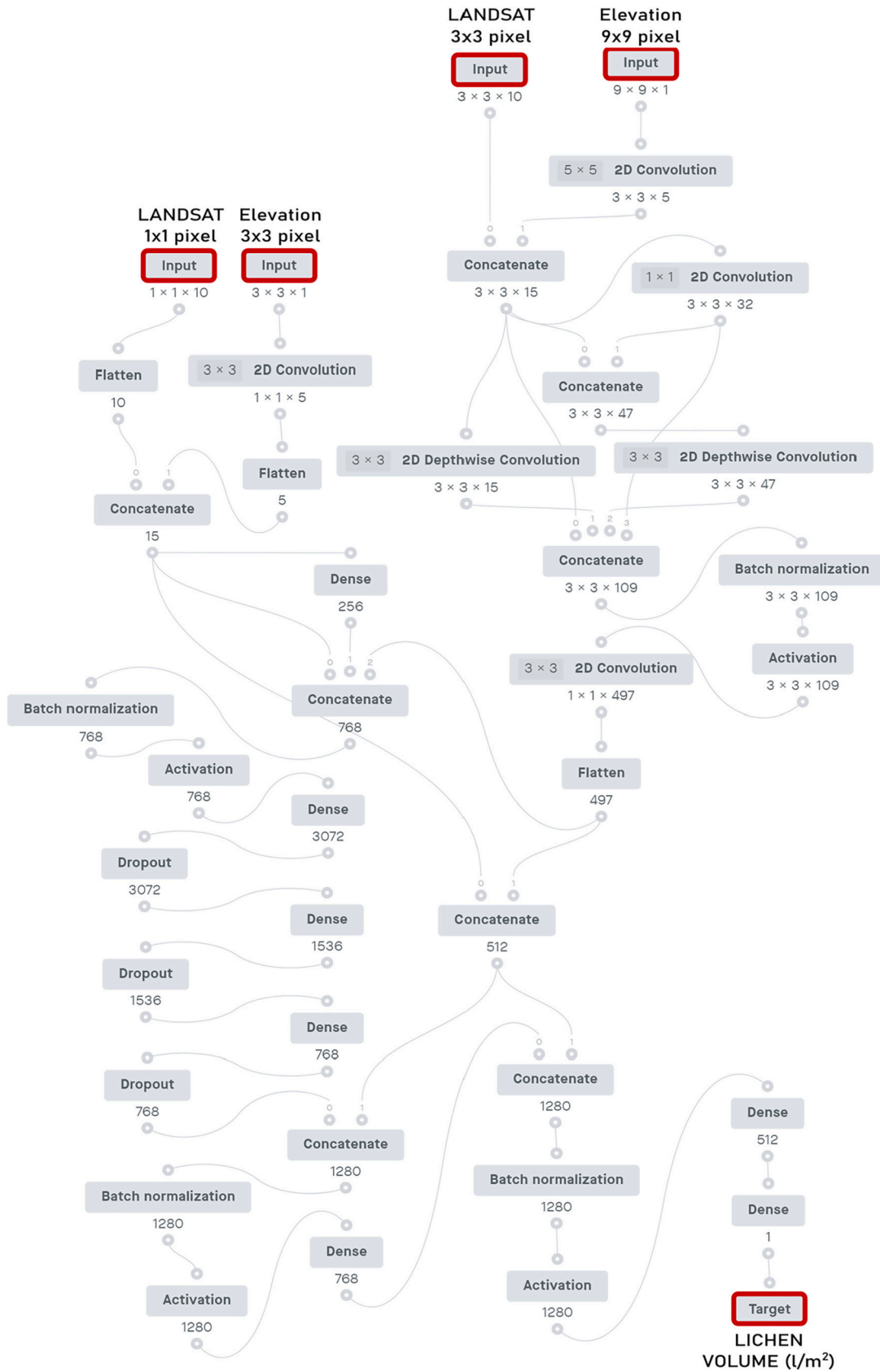


Fig. 3. The architecture of the deep neural network. Upper red boxes are input nodes. The lower red box is the output. Figures below boxes indicate the size of output tensors. Figures within boxes are kernel sizes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



considerable reduction in lichen abundance since the 1980s (Tømmervik et al., 2009). To test if the model could accurately detect changes in lichen abundance over time, we created a grid of points (240 m between point) covering the Kautokeino reindeer winter grazing area (5400 km<sup>2</sup>) and applied the model based on 2 Landsat mosaics from 1984 to 1985 and 2020–2021 respectively. We interpolated the values between the gridded points and masked water bodies based on NIR pixel values (threshold = 90) in the original Landsat mosaics.

To evaluate if the model produced reasonable results in areas that were not used for training, we also applied the model to two sites in Canada and Russia, with a grid resolution of 120 m, and interpolating the values between. (For the Canadian site we used the Canadian Digital Elevation Model from Natural Resources Canada). We also applied the model in native Landsat resolution to two smaller areas in Norway to illustrate how the model responded to bright rocks and sand.

### 3. Results

When applied to the test set ( $n = 159$ ), all models described a positive relationship between predicted values and empirical field measurements. The models were capable of detecting volumes as low as 1 dm<sup>3</sup>/m<sup>2</sup> and as high as 100 dm<sup>3</sup>/m<sup>2</sup> in heaths, tundra and open pine and birch forests. Applying the Lichen Volume Estimate algorithm (Falldorf et al., 2014) to our test set rendered an R<sup>2</sup> of 0.17 (RMSE = 12.31 dm<sup>3</sup>/m<sup>2</sup>,  $n = 158$ ,  $df = 1$ ,  $p < 0.01$ ). In the evaluation of spatial autocorrelation, the random forest model explaining measured lichen volume based on distances between plots (to evaluate the influence of spatial autocorrelation) reached a R<sup>2</sup> = 0.33 (RMSE = 15.1 dm<sup>3</sup>/m<sup>2</sup>).

#### 3.1. Candidate model training

The use of different training sets (Table 1) affected the goodness of fit and variation explained by the candidate AI models with R<sup>2</sup> ranging from 0.53 to 0.57 and RMSE from 9.95 to 11.24 dm<sup>3</sup>/m<sup>2</sup> (Table 3, Fig. 4). The model based on all available training data (Core + Mixed-sources) performed similarly as compared to the model trained with the core subset alone (Table 3). However, the model based on the Core and Mixed-sources had an intercept value of 4.05 compared to 5.64 in

the Core set, ( $p_{intercept} > 0.16$ ). Including the “remote” dataset (derived from aerial photos) in the training resulted in lower RMSE, but the highest intercept value (9.95). We considered the model trained on the Core + Mixed-sources as the best, based on the lowest intercept, highest regression score and lowest mean absolute error (See Table 4). This is the model subject for all following analyses, mapped estimations and the discussion.

#### 3.2. Permutation test

The full model had a Model Efficiency (ME) of 0.55. We ran 70 complete permutation iterations and calculated the mean ME. There was a sequential decrease in mean explanatory value (absolute change in ME) from the most important predictor variable, NDVI ( $\Delta ME = 0.13$ ), to the least important, TIR ( $\Delta ME = 0$ ), showing no drastic shifts across the range (Fig. 5). Elevation and the Green band was the second and third most important variables, whereas NDMI and the Red band was the 2nd and 3rd least important. For some predictors, the effect of the permutations varied considerably between iterations as indicated by the comparably large standard deviation ( $SD_{Green}: 0.07$ ,  $SD_{Elevation}: 0.05$ ,  $SD_{NIR}: 0.05$ ). The TIR, however, was constantly uninformative ( $SD = 0.01$ ).

#### 3.3. Applying the best candidate model to circumpolar sites

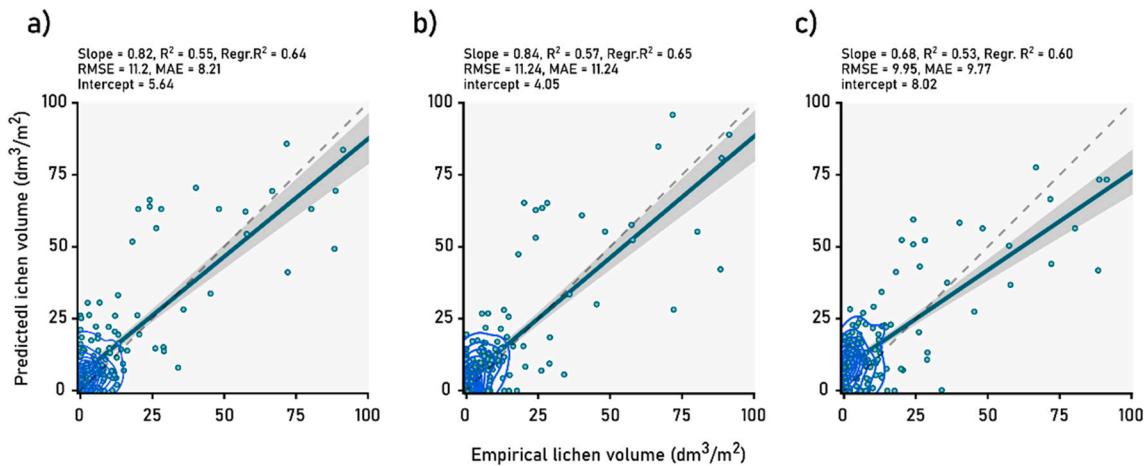
The model was applied to test sites in Quebec, Canada and the eastern Kola Peninsula, Russia (Fig. 6). Although we did not have ground-truth data from the sites outside of the Nordic countries and the western Kola Peninsula to evaluate the accuracy of the model, the predictions of the model corresponded well with our visual interpretations of lichen abundance using images on Google Earth, and literature confirm that these regions are rich in lichens. Visual comparisons against high-resolution aerial photos in Norway confirmed that the model performed well, detecting high lichen abundance in sparse Scots pine (*Pinus sylvestris*) tree forest stands (Supplementary Fig. S2).

However, when applied to sandy areas or human infrastructure, the model was unable to distinguish some bright surfaces from lichens and produced incorrect high estimations of lichens (Supplementary Fig. S3a,

**Table 3**

Table over the years field data was collected and what year the corresponding satellite image used for remote sensing was acquired. Satellite images originating from a different year than the field data was used when satellite coverage for the correct year was too patchy. Those instances are marked with \* and the year of the used image.

Sample year	Finnmark (NO)	Hardanger (NO)	Røros (NO)	Nikel (RUS) Pasvik (NO)	Teriberka (RUS)	Västerbotten (SE)	Andøya (NO)	Lapland (FIN)	Forolhogna (NO)	Sogn (NO)
1998	1998									
1999	1999									
2000		2003*		2000		2000			2000	
2001		2003*				2001				
2002						2002				
2003		2003								
2004		2003*								
2005	2005	2003*								
2006										
2007										
2008					2008					
2009							2009			
2010	2010									
2011										
2012										
2013	2013									
2014										
2015							2015			
2016										
2017										2018*
2018	2018						2018			2018
2019										2019
2020			2020					2020, 2019*		2020, 2019*

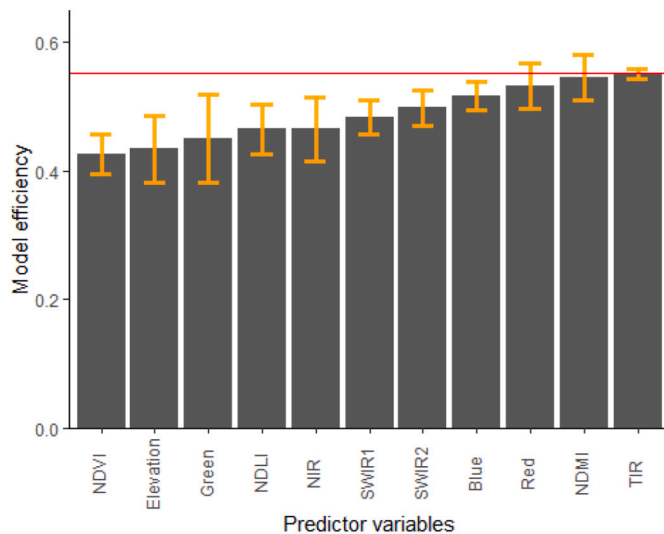


**Fig. 4.** The performance of the three candidate models predicting empirical lichen volume in plots that were new to the deep neural network. The three models were trained with datasets of increasing size: a) solely the core data set, b) The core set as well as a data set from mixed sources, c) The core, mixed and remotely assessed data set.

**Table 4**  
Performance of lichen volume candidate models.

Model	Intercept	Slope	p	R <sup>2</sup>	Regr. R <sup>2</sup>	Bias	MAE	RMSE
Trained with Core data	5.64	0.82	<0.001	0.55	0.64	-3.66	8.21	11.20
Trained with Core and Mixed sources	4.05	0.84	<0.001	0.57	0.65	-2.33	7.50	11.24
Trained with Core and Remotely assessed	8.02	0.68	<0.001	0.53	0.60	-4.48	9.77	9.95

RMSE = Root Mean Square Error, MAE = Mean Absolute Error, Regr. R<sup>2</sup> = Regression R<sup>2</sup>, see text for details.



**Fig. 5.** Model efficiency as a function of each predictor being permuted (deprived of its explanatory value through random shuffling). Bars indicate mean change in efficiency (error bars are standard deviation). The red line marks the model efficiency of the un-permuted model. The bigger the decrease compared to the red bar, the more the predictor variable contributes to the model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

an industrial site and sandy river sides in North Norway).

### 3.4. Time series analysis

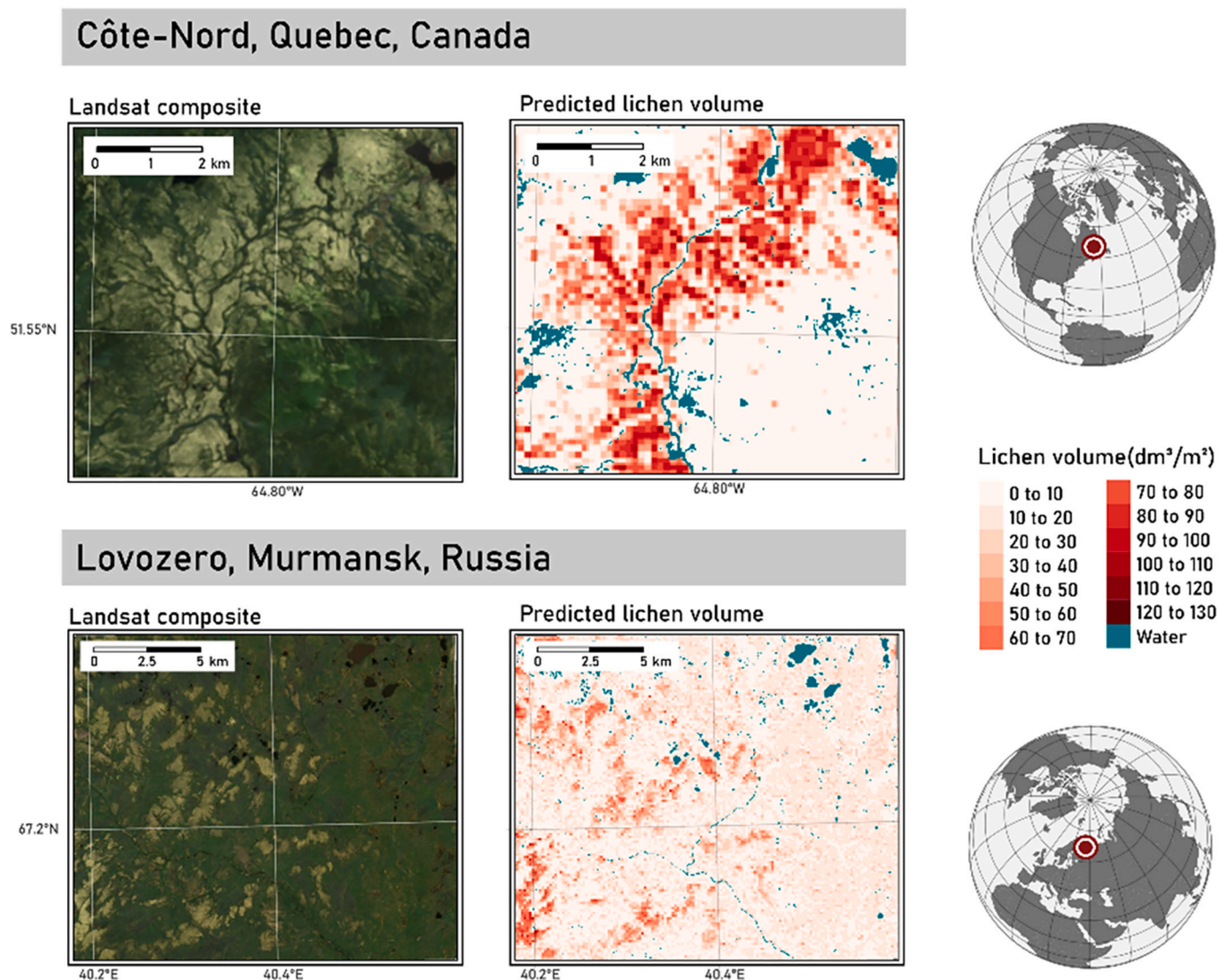
The model detected that lichen abundance halved between 1984 and 2020, both in mean lichen volume per m<sup>2</sup> (1984: 23.1 ± 14.45 dm<sup>3</sup>/m<sup>2</sup>, 2020: 11.9 ± 7.14 dm<sup>3</sup>/m<sup>2</sup>) and in total volume for the whole

Kautokeino area (1984: 116 million cubic meters, 2020: 60 million cubic meters, Fig. 7). The decrease was particularly visible near the reindeer fence along the Finnish-Norwegian border where the striking contrast between the thicker lichen layers on the Norwegian side and the scarcer lichen vegetation on the Finnish side decreased considerably during the period (Fig. 8).

## 4. Discussion

We assessed pale lichen volume using neural networks trained on Landsat satellite images and elevation data. By the use of artificial intelligence, we have here developed models that rendered considerably better fit than the method developed by Falldorf et al. (2014) did against field-based measurements.

In general, all candidate models managed to predict lichen volume up to 100 dm<sup>3</sup>/m<sup>2</sup> with high statistical significance ( $p < 0.001$ ) and R<sup>2</sup> values over 0.53. The predictive power of the models was high, given the results of Räsänen and Virtanen (2019) and Virtanen and Ek (2014) who concluded that at least 5 m spatial resolution is required for accurate vegetation mapping in treeless northern ecosystems. However, all of the models tended to overestimate lichen volume at low values, as indicated by the intercept values. Model outputs for areas with low lichen abundance are therefore likely to result in some overestimations. The model did also underestimate highest volumes (indicated by a slope lower than 1), and the general accuracy for higher lichen volumes was lower. This is probably largely an effect of the training dataset having comparably few data points with high lichen volumes and that the repetition of high values could not entirely compensate for this. However, it could also be related to a decreasing difference between plots with varying thickness, but consistently high cover over a certain level. Although our experience from fieldwork is that the thickness of lichen cover has a major impact on the reflected light, satellites, in a strict sense, are only able to register surface properties. If the difference between plots with varying lichen thickness is becoming less pronounced with increasing cover, there is likely a saturation threshold where remote sensing methods become less



**Fig. 6.** The model applied to an area in Canada (Côte-Nord, Quebec) and Russia (Eastern Kola Peninsula, Lovozero, Murmansk). The pale yellow and grey areas in the Landsat composite are areas covered with lichens, while green areas are dominated by vascular plants, primarily shrubs and graminoids. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

accurate or unable to efficiently quantify differences in high lichen biomass. This kind of problem is not new to remote sensing, a parallel example is the issues with accurate remote sensing of biomass in forests (Lu et al., 2016). More high cover datapoints would be needed to assess this effect and improve the model.

The inclusion of additional training data sets did not considerably affect the model performance (Table 3), but the intercept value improved (i.e. got closer to zero) by including the mixed sources dataset. The opposite happened when including the remotely assessed lichen volumes, which increased the intercept value. Altogether, this suggests that these additional datasets did not provide any information of fundamental value to the training process. While the mixed-sources dataset potentially resulted in more noise due to the variable quality of input data, the remotely assessed data points represented clear categories that likely did not introduce any new information to the training. This method, as we implemented it, did not provide an alternative to the time-consuming collection of empirical data in the field. Given the nature of neural networks, the current model can continuously be trained with new field data.

There was spatial autocorrelation in the field measured lichen data, and the test set was not spatially independent from the training set.

Earlier, Ploton et al. (2020) argued that spatial dependency between training and test set might lead to overestimates of model fit. However, the example of Ploton et al. (2020) represented an extreme case, since their model predicting aboveground forest biomass in Africa, in addition to 9 satellite bands, included 27 explanatory factors with strong inherent spatial autocorrelation (e.g. precipitation, temperature, absolute elevation, solar radiation). On the contrary, our model did not include any climatic or similar explanatory variables with strong patterns across pixels, which likely makes our modelling less sensitive to bias from spatial autocorrelation. This is further supported by the fact that the random forest model that used distances from other plots as the explanatory variables to explain lichen biomass (following Hengl et al., 2018) had notably lower explanatory capacity ( $R^2 = 0.33$ ) than our neural network models ( $R^2$  0.53–0.57). Wadoux et al. (2021) showed that a random K-fold cross-validation is appropriate for systematic random and simple random sample designs. However, our randomly selected test set is likely not able to compensate for the spatial distribution and autocorrelation of our data, and the test statistics can be suspected to be somewhat inflated. A tendency that is likely amplified by the fact that datapoints from Hardangervidda form a dominating part of the dataset. This is a consequence of the data not being collected with

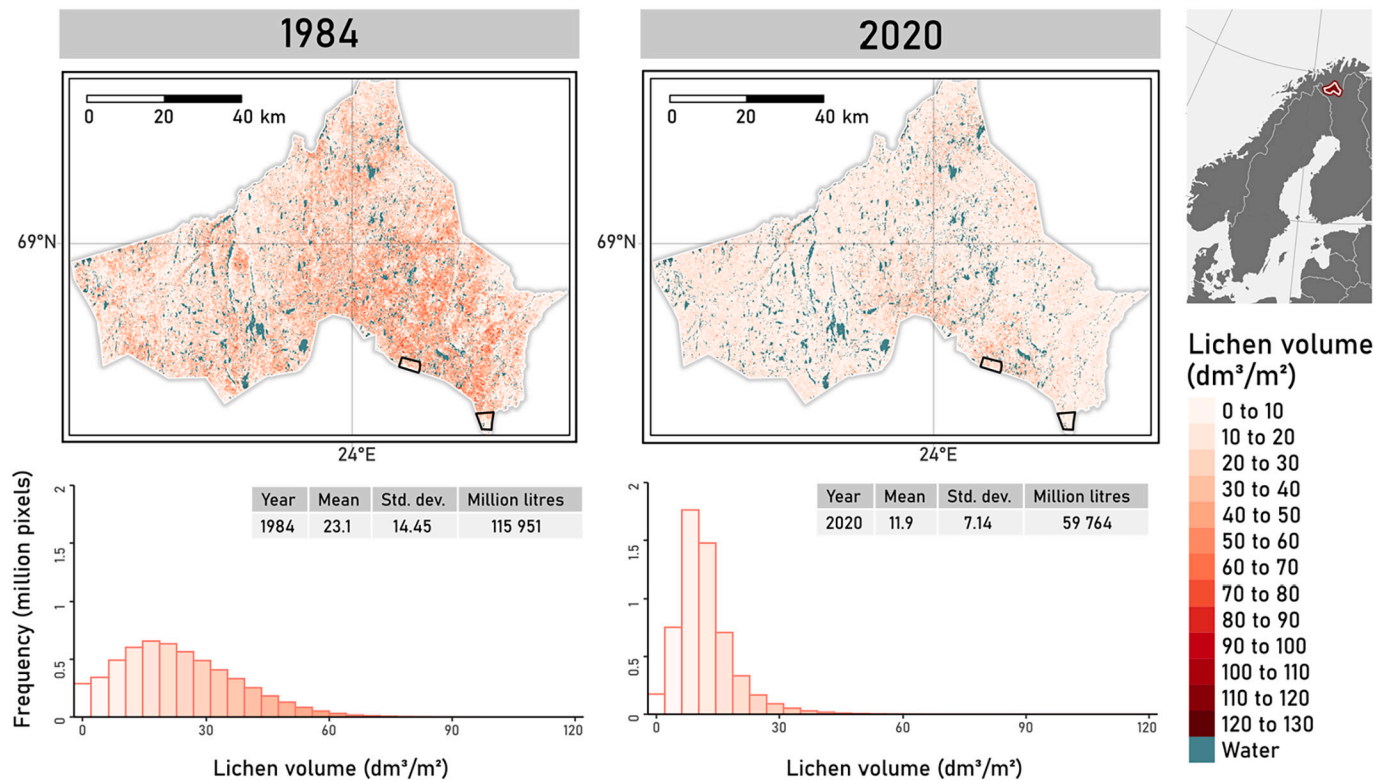


Fig. 7. The model applied to Landsat images from 1984 and 2020, respectively, over the winter reindeer grazing area of Kautokeino, north Norway. The model detected that lichen abundance halved during the period, both in mean lichen volume per  $m^2$ , and total lichen volume. The results are in line with the results of previous estimations based on field records. The outlined areas in the south-western corner are shown in greater detail in Fig. 8.

the current application in mind, but for different local studies and monitoring programs. Although desirable, a simple random or systematic random field sampling would demand a large investment in terms of money and time which was beyond our financial capacity.

#### 4.1. Predictive power of the explanatory factors

The permutation test showed that NDVI was the predictor with the strongest explanatory capacity, closely followed by regional elevation (Fig. 5). The importance of NDVI differed strongly from the results of Kennedy et al. (2020), where NDVI had a very low explanatory value, while elevation turned out as the by far most important predictor. However, there are some differences in the elevation input that can be critical when comparing the role of topography in the two studies. Kennedy et al. (2020) used absolute altitude while our model applies locally normalised elevation. While topography is likely to be important, lichens are not limited by elevation per se (Crittenden, 2000). However, if altitude and lichen abundance differed considerably between study sites, a neural network could infer an elevation signal that represents regional differences (for example in *Rangifer* grazing regime) rather than an actual effect of topography. This can result in an artificially boosted predictive power in the training regions at the cost of general applicability. Locally normalised elevation values, on the other hand, remove altitudinal differences between study sites, while keeping the effect of local topography structure. In this regard, our elevation input was likely more comparable to the aspect and slope predictors used by Kennedy et al. (2020) which were the 5th and 8th most important predictors in their model. An important difference, however, is that Kennedy et al. (2020) used elevation data with the same spatial resolution as the satellite imagery, while we used a higher spatial resolution (10 m) providing 9 elevation pixels per reflectance pixel (30 m). This way we did not need to calculate separate values for aspect or slope as the neural network could derive such information directly from the data. Our single

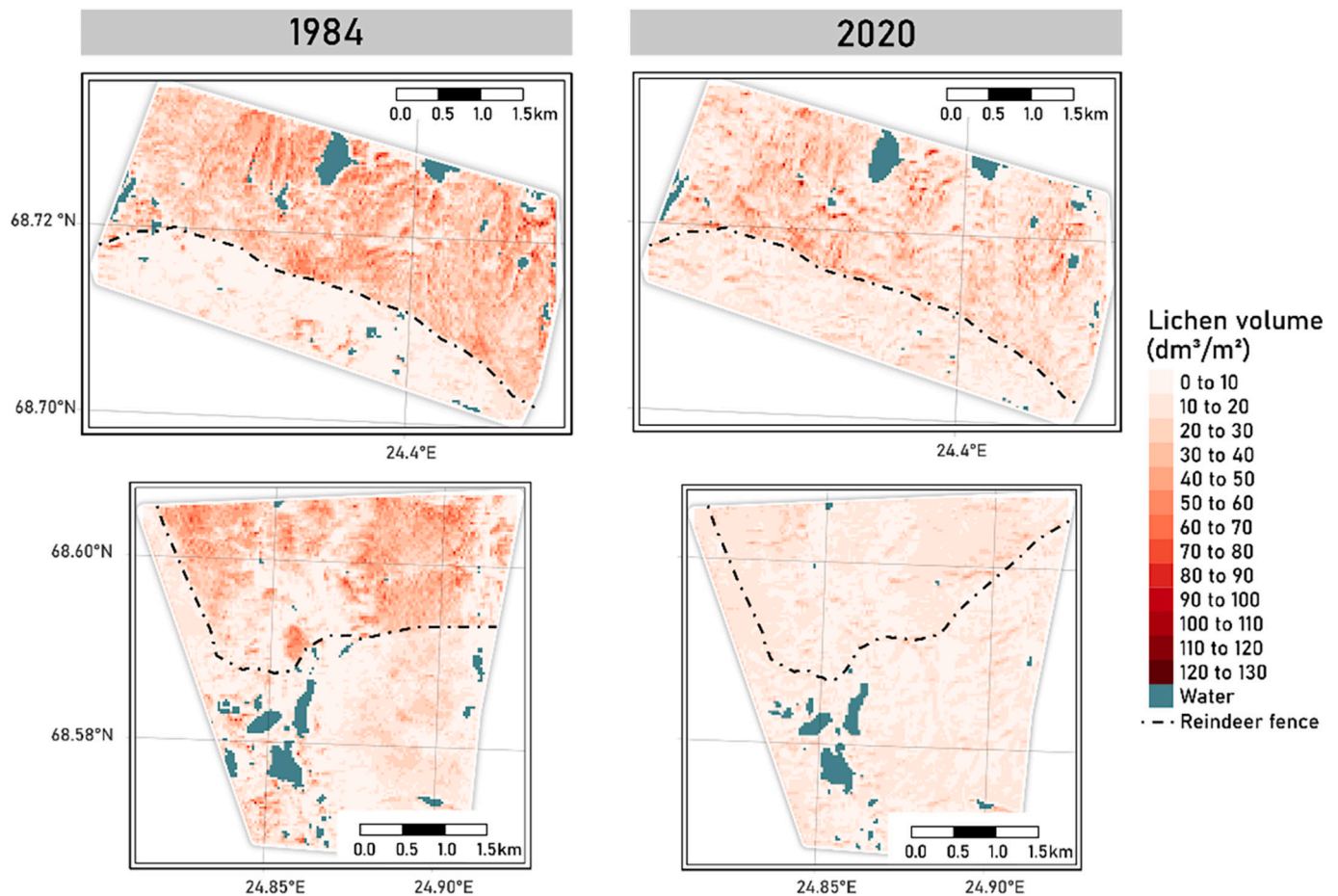
elevation input, hence, comprises the elevation, aspect and slope inputs of the model of Kennedy et al. (2020). Yet, despite a higher explanatory potential (more pixels provided), elevation had a lower explanatory capacity in our model compared to Kennedy et al. (2020). This, combined with the relatively lower predictive power of aspect and slope in their study, could be an indication that the model of Kennedy et al. (2020) may rely too much on regional differences in absolute altitude.

The predictive power of the green band (3rd most important) varied the most between iterations. The high variation suggests that certain combinations of the green band and other explanatory variables have a disproportional influence on the predictions. This would then also be the case, to a decreasing extent, for Elevation, NIR and NDLI.

The bands red, green, NIR and SWIR1 are all used in the calculation of the normalised indices. This could cloud the interpretation of their importance, since the model could, to some extent, compensate for the permutation of one of them by using the information in the corresponding index, and vice versa. The three indices diverge in this aspect. The most important variable, NDVI, is based on NIR and the red band, none of which turned out particularly informative. NDMI on the other hand, was less informative than its components NIR and SWIR1, while NDLI, and its components green and the SWIR1, ended up in the middle. Since the permutation test did not include correlation between variables, it is likely that the importance of less correlated predictors (e.g. elevation) are somewhat exaggerated.

The thermal band, TIR, could be expected to hold little to no information, and the results confirm this. The band could hence be excluded in further development of the model, but fills the function of a reference variable for the others to be compared against.

Temperature and precipitation were among the 4 most important predictors in the model of Kennedy et al. (2020). We deliberately excluded these variables to keep our model a strict remote sensing product, and we believe that the general applicability of our model is higher because of this choice. In addition to the increased simplicity, the



**Fig. 8.** Modelled lichen volume in two sections along the fenced Finnish-Norwegian border in 1984 and 2020 (for location, see Fig. 7). The fence (dotted line) does not strictly follow the border, and was partly moved during the period. (Both detailed sections have the model applied to all pixels, hence no interpolation.)

fact that our model relied more on reflectance features is likely an advantage when applied to areas with homogeneous topography, for example lichen tundra or lichen-dominated shrub tundra at lower altitudes.

#### 4.2. Applicability

Our results showed that our model was efficient in assessing pale lichen volume in ecosystems with moderate tree cover density, thereby offering a considerable improvement compared to earlier methods based on specific reflectance, for example the LVE equation by Falldorf et al. (2014). Furthermore, we observed that LVE, which is supposed to be applicable to open lichen heaths and lichen tundra, worked poorly on our evaluation set. Further, LVE cannot measure volumes higher than  $60 \text{ dm}^3/\text{m}^2$ , while our method has a theoretical limit of  $130 \text{ dm}^3/\text{m}^2$  (maximum value in the training set), although no plots over  $90 \text{ dm}^3$  appeared in our random evaluation set. Our method detected even small volumes of lichens in lichen tundra, dry heaths, and even in sparsely wooded forests. However, areas with bright sand, such as river sides, was falsely assessed as lichen rich. This means that model, although producing realistic results in our test applications, should be used with caution in areas with bare sand or artificially bright features and such larger areas should be masked out using a landcover map or visual classification. In future iterations of the model, specific training to improve the ability to distinguish sand and buildings from lichens should be a priority. Although the model correctly assessed low lichen abundance in a mosaic of dwarf shrub-lichen heath and bare bright bed rock with a scarce cover of fruticose lichens (Supplementary Fig. S3b),

we cannot rule out that bright bedrocks could result in overestimated lichen volumes as well. Furthermore it remains unexplored how the model would perform in estimation of lichen volume in forests with a denser canopy cover. From a reindeer herder perspective, dense forests are not the most important source for winter forage, as lichen volumes tend to decline with increasing tree crown density (Forbes et al., 2019). Thus, in economic and socio-cultural perspective, it is not so critical that the model has not yet been tested on datasets from dense forests.

Given that the training data was sampled in areas where lichens do occur to some extent (Table S2), we have not formally tested the applicability of the model in other vegetation types. Our results, when applying the model to larger areas, indicates that the model is fairly robust as it did not predict high lichen volume in areas where lichens clearly were absent. Although assessment of potentially lichen rich areas are of primary interest, the applicability of the model in other vegetation types is also important. It is possible that the model returns less reliable assessments in vegetation types that we have not investigated, and model output should therefore be manually investigated, especially if applied to vegetation types that were not covered by our study.

The results from the comparison between 1984 and 2020 in the Kautokeino winter grazing area in North Norway fitted well with the results reported by Tømmervik et al. (2009). Our results confirm the possible use of the Landsat catalogue of historical images (back to the launch of Landsat 5). The clear difference between years is an indication of the importance of the spectral bands for the model output, since the same topography input was used in 1984 and 2020 and hence remained constant across the years. The detected change over time confirms that although topography was the second most important explanatory

variable, the model is not overinterpreting topography, but using topography in combination with the spectral information. The ability of the model to accurately assess historical changes, also confirms that the model is generally applicable using older images than the collections that were used for training.

The historical comparison also puts the problem with overestimation of lichen abundance in lichen free areas in perspective. Although the model had a tendency to assign moderately high lichen abundance in some lichen free areas, it was still sensitive to the drastic decrease in Kautokeino winter grazing area where large areas became virtually lichen free. This shows that the model, despite some inaccuracies in lichen free areas, is able to detect relative changes. However, when used for absolute assessments in low-lichen areas, it should be combined with control measurements, preferably in the field.

The consistency between the model output and our visual interpretation, when we applied the model on landscape-scale across different areas in North America and Russia, suggests that the model performs consistently across regions and continents. Although we lack ground-truth data to verify that the accuracy is consistent outside our field sites, we find those results promising, opening up for circumpolar investigation and monitoring of lichen volume. Together with the models ability to detect lichen in sparsely forested areas (Fig. S2), potentially also in areas with mosaics of shrubs and trees. This is a major improvement compared to the LVE, which was originally only intended for use in exposed lichen heaths and lichen tundra (Falldorf et al., 2014) and therefore having a more limited regional applicability.

Our results confirm that neural networks are useful in complex remote sensing tasks. But contrary to the results of Kennedy et al. (2020) and Jozdani et al. (2021) our best model achieved higher accuracy with fewer explanatory variables; hence delimiting the input data requirements to elevation and the bands that Landsat 5, 7 and 8 have in common, i.e. datasets that are easily retrievable and do not rely on any environmental assumptions or weather data. In addition, our method appears to be robust for direct volume estimations of lichens over vast areas. The cropping of satellite images is however a time-consuming process for larger areas, and hence we applied the model on a lower spatial resolution (120–140 m) than the native 30 m Landsat resolution, an limited more detailed assessments to the specific areas along the Finnish-Norwegian border and the examples in the supplementary data. Preliminary analyses of Tveraa et al. (unpublished data) showed that change from 120 and 240 resolution had a limited effect on total lichen volume assessment, but high computational load could still be a limitation for mapping over larger areas if higher resolution is needed. More efficient raster cropping, or implementation of the model directly in Google Earth Engine, have the potential to improve high spatial resolution processing in large scale studies.

To summarize, volume is arguably a more ecologically and socially relevant metric of lichen abundance than cover, as several field-based studies have confirmed a close relationship between volume and biomass of terricolous fruticose lichens. The ability to remotely assess lichen volume, combined with the long historical catalogue of Landsat images, enables spatial and temporal studies of variation and changes in lichen biomass related to multiple research questions. The approach we have developed can be used to predict the quality of lichen pastures for *Rangifer* and, ultimately, facilitate and improve *Rangifer* habitat and population management (Rickbeil et al., 2017; Tømmervik et al., 2012). Combined with information on changes in drivers, for example climate, land use and management, and air pollution, the model can be used to provide estimates of environmental changes, as well as ecosystem services (Riseth et al., 2016; Stoy et al., 2012). Finally, as current vegetation-climate models typically disregard lichen abundance (Ellis, 2019), our lichen volume model may help to improve vegetation-climate models by facilitating continental-scale assessments of ground-lichen volumes.

## Funding

This study was supported by the Research Council of Norway (287402, 294948) and the National Natural Science Foundation of China (41861134036). Additional support was provided by FRAM - High North Research Centre for Climate & the Environment (369910, 369911) and CHARTER - EU Horizon 2020 (869471).

## CRedit authorship contribution statement

**Rasmus Erlandsson:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Jarle W. Bjerke:** Conceptualization, Investigation, Resources, Data curation, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Eirik A. Finne:** Investigation, Data curation, Writing – review & editing. **Ranga B. Myneni:** Writing – review & editing, Project administration, Funding acquisition. **Shilong Piao:** Writing – review & editing, Project administration, Funding acquisition. **Xuhui Wang:** Writing – review & editing, Project administration, Funding acquisition. **Tarmo Virtanen:** Investigation, Resources, Data curation, Writing – review & editing, Project administration, Funding acquisition. **Aleksi Räsänen:** Investigation, Resources, Data curation, Writing – review & editing. **Timo Kumpula:** Investigation, Resources, Data curation, Writing – review & editing, Project administration, Funding acquisition. **Tiina H.M. Kolari:** Investigation, Data curation, Writing – review & editing. **Teemu Tahvanainen:** Investigation, Data curation, Writing – review & editing. **Hans Tømmervik:** Conceptualization, Investigation, Resources, Data curation, Writing – review & editing, Supervision, Project administration, Funding acquisition.

## Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: co-author is a close relative of a co-founder of Peltarion AB, R.E.

## Data availability

Data will be made available on request.

## Acknowledgements

The authors would like to thank Marit K. Arneberg, Zander Venter, Geir R. Rauset, Olav Strand, Emelie Winqvist and Tobias Falldorf for their support during the study. Lena M. Tallaksen and Frode Stordal contributed financially to the participation of EAF in this study.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rse.2022.113201>.

## References

- Ahti, T., Hepburn, R., 1967. Preliminary Studies on Woodland Caribou Range, Especially on Lichen Stands, 74. Ontario Department of Lands and Forests, Technical Service Research Report, Ontario, pp. 1–134.
- Bivand, R.S., Pebesma, E., Gomez-Rubio, V., 2013. Applied Spatial Data Analysis with R, Second edition. Springer, NY.
- Brodrick, P.G., Davies, A.B., Asner, G.P., 2019. Uncovering ecological patterns with convolutional neural networks. Trends Ecol. Evol. 34, 734–745. <https://doi.org/10.1016/j.tree.2019.03.006>.
- Cohen, J., Pulliainen, J., Ménard, C.B., Johansen, B., Oksanen, L., Luojus, K., Ikonen, J., 2013. Effect of reindeer grazing on snowmelt, albedo and energy balance based on satellite data analyses. Remote Sens. Environ. 135, 107–117.
- Collins, W.B., Dale, B.W., Adams, L.G., McElwain, D.E., Joly, K., 2011. Fire, grazing history, lichen abundance, and winter distribution of caribou in Alaska's taiga. J. Wildl. Manag. 75, 369–377.

- Crittenden, P., 2000. Aspects of the ecology of mat-forming lichens. *Rangifer* 20, 127–139.
- Cutler, N., 2010. Long-term primary succession: a comparison of non-spatial and spatially explicit inferential techniques. *Plant Ecol.* 208, 123–136.
- Elbert, W., Weber, B., Burrows, S., Steinkamp, J., Büdel, B., Andreae, M.O., Pöschl, U., 2012. Contribution of cryptogamic covers to the global cycles of carbon and nitrogen. *Nat. Geosci.* 5, 459–462.
- Ellis, C.J., 2019. Climate change, bioclimatic models and the risk to lichen diversity. *Diversity* 11, 54.
- Falldorf, T., Strand, O., Panzacchi, M., Tømmervik, H., 2014. Estimating lichen volume and reindeer winter pasture quality from Landsat imagery. *Remote Sens. Environ.* 140, 573–579. <https://doi.org/10.1016/j.rse.2013.09.027>.
- Fehmi, J.S., 2010. Confusion among three common plant cover definitions may result in data unsuited for comparison. *J. Veg. Sci.* 21, 273–279.
- Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Joseph Hughes, M., Laue, B., 2017. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* 194, 379–390. <https://doi.org/10.1016/j.rse.2017.03.026>.
- Forbes, B.C., Turunen, M.T., Soppela, P., Rasmus, S., Vuojala-Magga, T., Kitti, H., 2019. Changes in mountain birch forests and reindeer management: comparing different knowledge systems in Sápmi, northern Fennoscandia. *Polar Record* 55, 507–521. <https://doi.org/10.1017/S0032247419000834>.
- Fraser, R.H., Lantz, T.C., Olthof, I., Kokelj, S.V., Sims, R.A., 2014. Warming-induced shrub expansion and lichen decline in the Western Canadian Arctic. *Ecosystems* 17, 1151–1168.
- Fraser, R.H., Pouliot, D., van der Sluijs, J., 2021. UAV and high resolution satellite mapping of forage lichen (*Cladonia* spp.) in a rocky Canadian shield landscape. *Can. J. Remote Sens.* 0, 1–14. <https://doi.org/10.1080/07038992.2021.1908118>.
- Fremstad, Eli, 1997. Vegetasjonstyper i Norge. Norsk Institutt for Naturforskning, Trondheim.
- Gauslaa, Y., 1984. Heat resistance and energy budget in different Scandinavian plants. *Ecography* 7, 5–6.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- Gould, W., 2000. Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots. *Ecol. Appl.* 10, 1861–1870.
- Granlund, L., Keski-Saari, S., Kumpula, T., Oksanen, E., Keinänen, M., 2018. Imaging lichen water content with visible to mid-wave infrared (400–5500 nm) spectroscopy. *Remote Sens. Environ.* 216, 301–310. <https://doi.org/10.1016/j.rse.2018.06.041>.
- Hammer, B., Frasco, M., 2018. Metrics: Evaluation Metrics for Machine Learning.
- He, L., Chen, W., Leblanc, S.G., Lovitt, J., Arseneault, A., Schmelzer, I., Fraser, R.H., Latifovic, R., Sun, L., Prévost, C., White, H.P., 2021. Integration of multi-scale remote sensing data for reindeer lichen fractional cover mapping in Eastern Canada. *Remote Sens. Environ.* 267, 112731.
- Hein, N., Merkelbach, J., Zech, K., Weijers, S., 2021. Drought sensitivity of *Empetrum nigrum* shrub growth at the species' southern lowland distribution range margin. *Plant Ecol.* 222, 305–321. <https://doi.org/10.1007/s11258-020-01107-z>.
- Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B.M., Gräler, B., 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ* 6, e5518. <https://doi.org/10.7717/peerj.5518>.
- Hijmans, R.J., 2021. Terra: Spatial data analysis.
- Hijmans, R.J., van Etten, J., 2012. Raster: Geographic Analysis and Modeling with Raster Data.
- Jernsletten, J.-L.L., Klokov, K., 2002. Sustainable reindeer husbandry. University of Tromsø, Tromsø, Norway, Centre for Saami Studies.
- Joly, K., Jandt, R.R., Klein, D.R., 2009. Decrease of lichens in Arctic ecosystems: the role of wildfire, caribou, reindeer, competition and climate in North-Western Alaska. *Polar Res.* 28, 433–442.
- Joly, K., Chapin, F.S., Klein, D.R., 2010. Winter habitat selection by caribou in relation to lichen abundance, wildfires, grazing, and landscape characteristics in Northwest Alaska. *Ecoscience* 17, 321–333.
- Jozdani, S., Chen, D., Chen, W., Leblanc, S.G., Prévost, C., Lovitt, J., He, L., Johnson, B. A., 2021. Leveraging deep neural networks to map caribou lichen in high-resolution satellite images based on a small-scale, noisy UAV-derived map. *Remote Sens.* 13, 2658.
- Kappen, L., 1988. Handbook of lichenology. In: Galun, Margalith (Ed.), *Ecophysiological Relationships in Different Climatic Regions*, vol. 2. CRC Press, Boca Raton, Florida.
- Kennedy, B., Pouliot, D., Manseau, M., Fraser, R., Duffe, J., Pasher, J., Chen, W., Olthof, I., 2020. Assessment of landsat-based terricolous macrolichen cover retrieval and change analysis over caribou ranges in northern Canada and Alaska. *Remote Sens. Environ.* 240, 111694. <https://doi.org/10.1016/j.rse.2020.111694>.
- Klein, D.R., Shulski, M., 2009. Lichen recovery following heavy grazing by reindeer delayed by climate warming. *Ambio* 38 (1), 11–16.
- Larson, D., 1987. The absorption and release of water by lichens. *Bibliotheca Lichenologica* 25, 351–360.
- Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., Moran, E., 2016. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth* 9, 63–105. <https://doi.org/10.1080/17538947.2014.990526>.
- Macander, M.J., Frost, G.V., Palm, E., 2018. Lichen cover mapping for the range of the Fortymile Caribou herd, Alaska and the Yukon Territory. In: Bureau of Land Management, Fairbanks, Alaska, USA. Department of Environment, Government of Yukon, Whitehorse, Yukon, Canada.
- Macander, M.J., Palm, E.C., Frost, G.V., Herriges, J.D., Nelson, P.R., Roland, C., Russell, K.L., Suito, M.J., Bentzen, T.W., Joly, K., Goetz, S.J., 2020. Lichen cover mapping for caribou ranges in interior Alaska and Yukon. *Environ. Res. Lett.* 15 (5), 055001.
- Moen, J., Danell, Ö., Holt, R., 2007. Non-destructive estimation of lichen biomass. *Rangifer* 27. <https://doi.org/10.7557/2.27.1.172>.
- Myneni, R.B., Keeling, C., Tucker, C.J., Asrar, G., Nemani, R.R., 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* 386, 698–702.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — a discussion of principles. *J. Hydrol.* 10, 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Nordberg, M.-L., 1998. Vegetation and biomass changes in mountainous areas in Sweden using satellite and airborne imaging scanner data. *Information for sustainability* 431–435.
- Nordberg, M.-L., Allard, A., 2002. A remote sensing methodology for monitoring lichen cover. *Can. J. Remote Sens.* 28, 13.
- Pebesma, E., 2018. Simple features for r: standardized support for spatial vector data. *The R Journal* 10, 439–446. <https://doi.org/10.32614/RJ-2018-009>.
- Piao, S., Wang, X., Park, T., Chen, C., Lian, X., He, Y., Bjerke, J.W., Chen, A., Ciais, P., Tømmervik, H., Nemani, R.R., Myneni, R.B., 2020. Characteristics, drivers and feedbacks of global greening. *Nat. Rev. Earth Environ.* 1, 14–27. <https://doi.org/10.1038/s43017-019-0001-x>.
- Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S., Péliissier, R., 2020. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nat. Commun.* 11, 4540. <https://doi.org/10.1038/s41467-020-18321-y>.
- Porter, C., Morin, P., Howat, I., Noh, M.-J., Bates, B., Peterman, K., Keesey, S., Schlenk, M., Gardiner, J., Tomko, K., Willis, M., Kelleher, C., Cloutier, M., Husby, E., Foga, S., Nakamura, H., Platson, M., Wethington, M., Williamson, C., Bauer, G., Enos, J., Arnold, G., Kramer, W., Becker, P., Doshi, A., D'Souza, C., Cummins, P., Laurier, F., Bojesen, M., 2018. ArcticDEM. <https://doi.org/10.7910/DVN/OHHUKH>.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing (Version 4.0.3). R Foundation for Statistical Computing, Vienna, Austria.
- Räsänen, A., Virtanen, T., 2019. Data and resolution requirements in mapping vegetation in spatially heterogeneous landscapes. *Remote Sens. Environ.* 230, 111207. <https://doi.org/10.1016/j.rse.2019.05.026>.
- Rautiainen, M., Suomalainen, J., Möttöus, M., Stenberg, P., Voipio, P., Peltoniemi, J., Manninen, T., 2007. Coupling forest canopy and understory reflectance in the Arctic latitudes of Finland. *Remote Sens. Environ.* 110, 332–343.
- Rees, W., 2004. Reflectance spectra of subarctic lichens between 400 and 2400 nm. *Remote Sens. Environ.* 90, 281–292. <https://doi.org/10.1016/j.rse.2003.12.009>.
- Rickbeil, G.J.M., Hermosilla, T., Coops, N.C., White, J.C., Wulder, M.A., 2017. Estimating changes in lichen mat volume through time and related effects on barren ground caribou (*Rangifer tarandus groenlandicus*) movement. *PLoS One* 12, e0172669. <https://doi.org/10.1371/journal.pone.0172669>.
- Riseth, J.Å., Tømmervik, H., Bjerke, J.W., 2016. 175 years of adaptation: North Scandinavian Sámi reindeer herding between government policies and winter climate variability (1835–2010). *J. For. Econ.* 24, 186–204.
- Ryu, Y., Berry, J.A., Baldocchi, D.D., 2019. What is global photosynthesis? History, uncertainties and opportunities. *Remote Sens. Environ.* 223, 95–114.
- Sedia, E.G., Ehrenfeld, J.G., 2003. Lichens and mosses promote alternate stable plant communities in the New Jersey pinelands. *Oikos* 100, 447–458.
- Solheim, I., Engelsen, O., Hosgood, B., Andreoli, G., 2000. Measurement and modeling of the spectral and directional reflection properties of lichen and moss canopies. *Remote Sens. Environ.* 72, 78–94.
- Stoy, P.C., Street, L.E., Johnson, A.V., Prieto-Blanco, A., Ewing, S.A., 2012. Temperature, heat flux, and reflectance of common subarctic mosses and lichens under field conditions: might changes to community composition impact climate-relevant surface fluxes? *Arct. Antarct. Alp. Res.* 44, 500–508.
- Tømmervik, H., Johansen, B., Tombre, I., Thannheiser, D., Høgda, K., Gaare, E., Wielgolaski, F., 2004. Vegetation changes in the Nordic mountain birch forest: the influence of grazing and climate change. *Arct. Antarct. Alp. Res.* 36, 323–332.
- Tømmervik, H., Johansen, B., Riseth, J., Karlsen, S., Solberg, B., Høgda, K., 2009. Above ground biomass changes in the mountain birch forests and mountain heaths of Finnmarksvidda, northern Norway, in the period 1957–2006. *For. Ecol. Manag.* 257, 244–257.
- Tømmervik, H., Bjerke, J.W., Gaare, E., Johansen, B., Thannheiser, D., 2012. Rapid recovery of recently overexploited winter grazing pastures for reindeer in northern Norway. *Fungal Ecol.* 5, 3–15. <https://doi.org/10.1016/j.funeco.2011.08.002>.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150.
- Virtanen, T., Ek, M., 2014. The fragmented nature of tundra landscape. *International Journal of Applied Earth Observation and Geoinformation, Special Issue on Polar Remote Sensing 2013* (27), 4–12. <https://doi.org/10.1016/j.jag.2013.05.010>.
- Wadoux, A.M.J.-C., Heuvelink, G.B.M., de Bruin, S., Brus, D.J., 2021. Spatial cross-validation is not the right way to evaluate map accuracy. *Ecol. Model.* 457, 109692. <https://doi.org/10.1016/j.ecolmodel.2021.109692>.
- Wilson, E.H., Sader, S.A., 2002. Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sens. Environ.* 80, 385–396. [https://doi.org/10.1016/S0034-4257\(01\)00318-2](https://doi.org/10.1016/S0034-4257(01)00318-2).
- Yan, Y., 2016. *MLmetrics: Machine Learning Evaluation Metrics*.
- Zambrano-Bigiarini, 2020. hydroGOF: Goodness-of-Fit Functions for Comparison of Simulated and Observed Hydrological Time Series. <https://doi.org/10.5281/zenodo.839854>.