Remote Sensing Applications: Society and Environment 36 (2024) 101349

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

journal homepage: www.elsevier.com/locate/rsase

A fully automated model for land use classification from historical maps using machine learning

Anneli M. Ågren *, Yiqi Lin

Department of Forest Ecology and Management, Swedish University of Agricultural Science, Skogsmarksgränd, SE-907 36, Umeå, Sweden

ARTICLE INFO

Keywords:
Automation
Economic map
Extreme gradient boosting
Historical maps
Land use
Machine learning

ABSTRACT

Digital land use data before the age of satellites is scarce. Here, we build a machine learning model, using Extreme Gradient Boosting, that can automatically detect land use classes from an orthophoto map of Sweden (economic maps, 1:10 000 and 1:20 000) constructed from 1942 to 1988. Overall, the machine learning model demonstrated robust performance, with Cohen's Kappa and Matthews Correlation Coefficient of 0.86. The F1 values of the individual classes were 0.98, 0.95, 0.84, and 0.87 for graphics, arable land, forest, and open land, respectively. While the model can be used to detect land use changes in arable land, higher uncertainties associated with forest and open land necessitate further investigation at regional scales or exploration of improved mapping techniques. The code is publicly available to enable easy adaptation for classifying other historical maps.

1. Introduction

Understanding the land use and land cover change (LULUC) is vital to tackling large-scale challenges such as biodiversity loss, climate change, and food production (Winkler et al., 2021). Habitat loss caused by land use changes is recognized as one of the most pressing threats to biodiversity worldwide (Newbold et al., 2015, 2016; Loucks et al., 2008). Studies showed the lasting effect of historical land use on plant diversity. For example, grassland plant diversity in agricultural landscapes was shown to be controlled by land use more than 200 years ago (Gustavsson et al.,2007). Similarly, persistent differences between the vegetation of primary forests and post-agricultural secondary forests indicate that distribution patterns for many plant species still reflect the open agricultural environment of the nineteenth century (Bellemare et al.,2002). This study focuses on Sweden, which has undergone significant land use changes in the last century. The amount of arable land peaked between 1910 and 1950 at over 3.7 million hectares but decreased to 2.5 million hectares by 2007 (Swedish Board of Agriculture 2011) southern Sweden's total open land cover declined by 17% during this period (Auffret et al.,2018), with afforestation being the major driver. However, some arable land has also been converted into open land uses (Auffret et al.,2018). Historical landscape predicted present-day species richness better than current landscape, suggesting the long-term effects of historical land use (Auffret et al.,2018). Furthermore, climate change can compound plant responses to habitat conversion, as both shape the re-distribution of plant species in space and over time (Auffret and Svenning 2022). While studies on land-use changes have been performed in southern Sweden (Auffret et al.,2018), research in northern Sweden is still lacking. Understanding the intricate interplay between climate change and land use change necessitates studies in these areas, as the northern latitudes are experiencing even more rapid climate change (Previdi et al.,2021; Laudon et al., 2021).

Land use and land-cover change contributed to 12.5% of anthropogenic global carbon emissions from 1990 to 2010 (Houghton

* Corresponding author.
E-mail address: anneli.agren@slu.se (A.M. Ågren).

https://doi.org/10.1016/j.rsase.2024.101349
Received 13 June 2024; Received in revised form 27 August 2024; Accepted 8 September 2024
Available online 12 September 2024

et al., 2012). Drainage of wet organic soils for agricultural purposes increases greenhouse gas (GHGs: CO_2, N_2O, and CH_4) emissions by roughly 1 t CO_2 equivalents ha^{-1} $year^{-1}$, compared to undrained peat (Kasimir-Klemedtsson et al.,1997). Globally, GHG emissions from drained organic soils make up 32% of the total emissions from the agricultural sector but for Sweden the corresponding number is 44%. Rewetting drained peatlands is therefore an important climate mitigation option in Sweden. It has been suggested that forested peatlands previously used for agriculture are more nutrient-rich and emit higher levels of N_2O (Klemedtsson et al.,2005). Hence, former arable soils, now often afforested, might be good objects for rewetting. While the nutrient status of mineral forest soils and peatlands generally increases towards the south (Högberg et al., 2021) peatlands display the opposite patterns (Laudon et al., unpublished manuscript) with more eutrophic peatlands in northern Sweden compared to the south. Whether these nutrient-rich peatlands in the north are naturally nutrient-rich sites or former arable land, is still unknown.

Sweden's dependence on food imports from countries facing a broad range of potential climate risks (Horn et al., 2022) highlights the importance of domestic food production. Since the end of the 2nd World War, there has been a decrease in the number of farms and the total area of arable land due to farm closure and consolidation, particularly in northern Sweden, where 55–60% of the arable land was lost from 1944 to 1988 (Riksantikvarieämbetet 1996). However, the 2017 Swedish new food strategy (2016/17:104) aiming to increase domestic food production suggests that parts of the previously abandoned arable land might be recultivated.

Understanding historical land use across large areas, especially in northern Sweden, is crucial for addressing ecological, climate change, and food security challenges. Though present-day geographical databases are available on both national and EU level (Abdi 2020; Naturvårdsverket 2023), land use data pre-dating the launch of satellite imagery programs like Landsat in the 1970s is scarce (Mäyrä et al., 2023). The semi-automated approach proposed by Auffret et al. (2017) that digitized land use from historical maps was

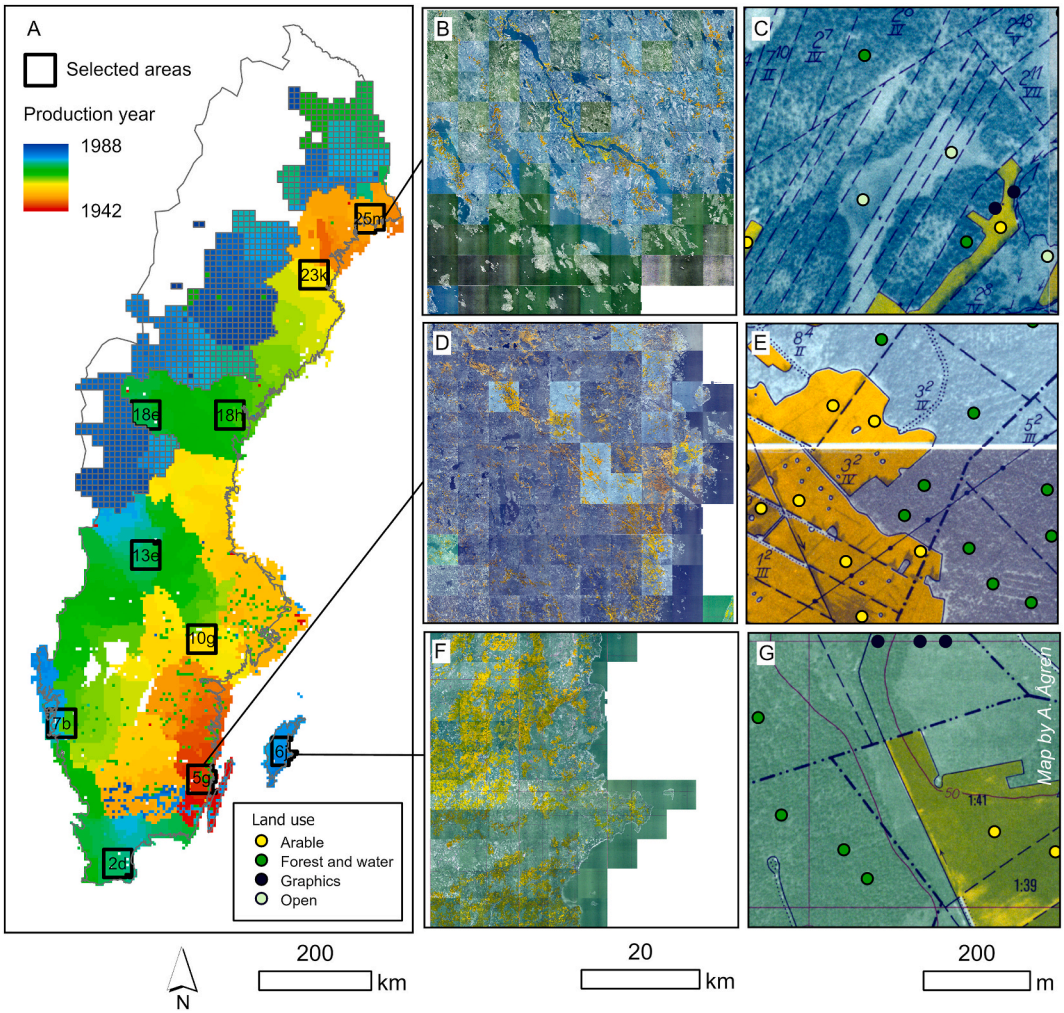

Fig. 1. A) Ten map regions for digitizing training and testing data were selected (A) to cover a range of production years and landscape types. B, D, F) Overview of three of the regions. Note the variability in color among the scanned maps. C, E, G) Example of digitized points, classified according to four classes, these are shown in the correct map scale (1:10 000). Note that the orthophotos are in the background of the map, with interpreted symbols drawn on top. The map tiles with a grey outline, in the northern inland of Sweden (A), are at the scale of 1:20 000, compared to 1:10 000 for the rest of Sweden. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

geographically limited to southern Sweden. Advancements in Machine Learning and Deep Learning algorithms, as well as the capability of supercomputers to handle high-resolution data for entire countries, are revolutionizing how maps are developed (Song et al., 2023). Recent successful applications in Sweden include automatic detection of ditch systems with high accuracy by deep learning (Laudon et al., 2022; Lidberg et al., 2023), machine learning generated soil moisture map at 2-m resolution (Ågren et al.,2021), which was further leveraged to create a novel peat map surpassing the accuracy of traditional methods (Ågren et al.,2022). These high-resolution maps provide invaluable new datasets for researchers and land-use planners working at large scales.

In this study, we developed a fully automated machine learning model for land use classification from scanned economic maps

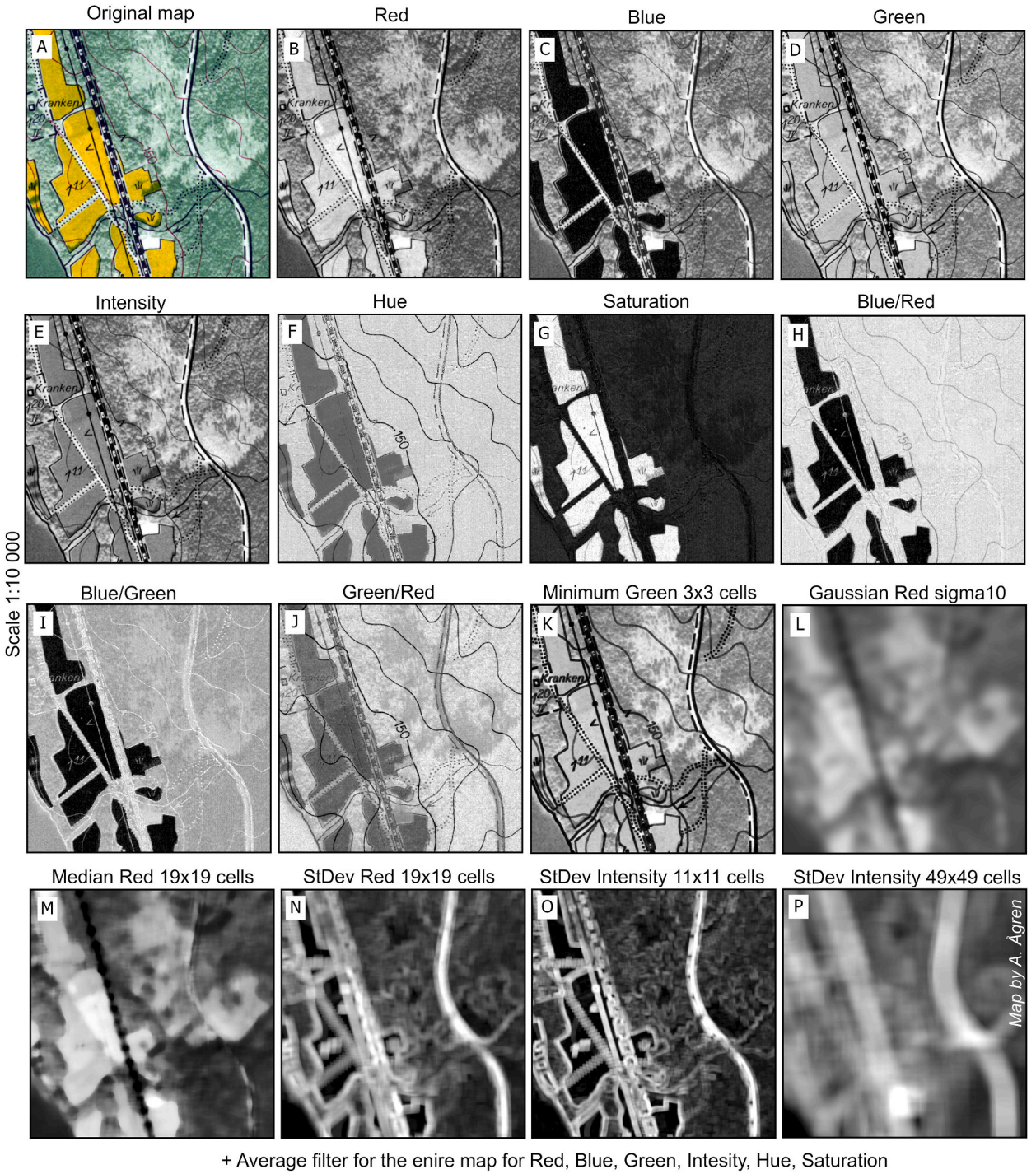

Fig. 2. The computed features from the scanned economic maps. 21 different features were derived. All maps show the same area, displayed at the scale of 1:10 000.

covering the entire country and evaluated the quality of the generated maps to determine their suitability for detecting land use changes over time.

2. Methods

The historical map series economic map (*Ekonomiska kartan in Swedish*) provides a detailed record of historical land use across most of the country (Fig. 1A), except the mountain range. In southern Sweden, each map sheet covers 25 km^2 (1:10 000). In northern inland Sweden, each sheet covers 100 km^2 (1:20 000) (Fig. 1A). The maps feature a monochrome aerial orthophoto in the background, with arable land colored in yellow and additional information such as roads, buildings, ownership boundaries, and place names in black (Fig. 1B–G). Maps from 1935 to 1947 have a black-and-white background photo, while for those produced after 1947, the photo was printed in green and white. More modern maps have topographical contour lines (Fig. 1G) and wetlands denoted with hatched areas. During the years 1992–1997, the maps were scanned as images by the Swedish Land Survey and released digitally as GeoTIFFs. The maps that were published online were produced from 1942 to 1988 (Fig. 1A).

2.1. Digitizing training/testing data for the machine learning model

The scanned economic maps from various production years and locations across Sweden were first explored visually. Due to variations in mapping methodologies, aging of the paper, and artifacts from printing and later scanning of the maps, land use categories can be expressed in a wide range of colors (Fig. 1B, D, F). To address this variability, 10 map regions were selected to cover a wide range of production years and landscape types (Fig. 1A). Within each region, up to 100 economic maps (Fig. 1B, D, F) were visually inspected to generate as diverse training/testing data as possible (Fig. 1C, E, G). Graphics were digitized at a scale of approximately 1:200–1:600 (i.e., zoomed in on pixel level), depending on the size of the object, while land-use classes were digitized at a scale of ca 1:2000–1:6000 (i.e., zoomed in on "landscape level"). Identifying land use classes relied primarily on color and texture within the orthophotos. Arable land was easily identifiable by its yellow color (though, in reality, it varied in shades from yellow to brown, orange, lime, and dark green with a hint of yellow). Generally speaking, 'forest and water' was darker (dark grey, blue, or green) than open land, with coniferous forest being darker and "spikier" than deciduous forest, which was lighter in color and "rounder" in appearance. Open land typically had the lightest colors. In the south, most open areas were fields, while in the north, most open areas were mostly peatlands or clear-cuts. However, land use was not always determined by color. The texture of the orthophoto on the map was another key determinant. Open areas often had a "smooth surface," i.e., little local variation in colors (Fig. 1C–G), while forests exhibited less uniform texture due to sunlit parts and shadows (Figs. 1C and 2A). Wet parts of peatlands, though often dark, could be visually distinguished from forests by their smoother homogeneous surface (Fig. 1C). Additionally, peatlands could also have bands with light and dark pixels, typical of string bogs. The minimum forest size, according to the definition from the Global Forest Resources Assessment (FRA), is 0.5 ha (ca 70 × 70 m) (Keenan et al., 2015). Open areas, however, can be smaller, such as a backyard with a lawn or the circular end of a cul-de-sac. Clear-cuts with a few seed trees were classified as open land (Fig. 1A), while a sparse natural forest was classified as forest land. Some of the classes were not easily identified even by human experts. We digitized numerous examples of such "tricky sites" with the aim of improving learning on them. Examples included dark, "patchy" green areas resembling forests but were situated in the mountains above the tree line where similar patterns were caused by field vegetation, or dark green areas on islands that were actually field vegetation and stoniness that had a forest-like appearance. A dense stand of even-aged forest could also be mistaken for an open field due to its homogeneous green impression. Based on map inspection, we expected arable land to be the easiest to classify and open land to be the most difficult. Therefore, we provided the model with additional examples of open land and forest. The final digitized points comprised 4312 for arable land, 2366 for 'forest and water', 4753 for open land, and 5223 for graphics. Local names provided additional information helpful for labelling, with descriptive names often ending with "-bog", "-moss", "-field", "-farm", "-forest", etc. In total we digitized 16 654 classified points from 938 maps, across the 10 study regions.

2.2. Extracting features to enhance the separation between land use classes

To enhance the separation between land use classes, several features were extracted from the original maps (Fig. 2A). The features were extracted with Whitebox tools v. 2.3.1 (Lindsay 2023), NumPy v. 1.24.4 (NumPy et al.,2023) and Rasterio v. 1.3.7 (Rasterio 2024). First, the composite images were separated into red, green, and blue bands (Fig. 2B–D), and converted to intensity, hue, and saturation (Fig. 2E–G). The ratios of the blue/red band, blue/green band, and green/red band were calculated (Fig. 2H–J). The graphics were most highlighted in the green band (Fig. 2D). To further enhance the graphics, a minimum filter on the green band with a kernel of 3 cells was applied (Fig. 2K). The arable areas appeared prominent in the blue band, saturation, blue/red, and blue/green features (Fig. 2C, G, I). 5 filters were calculated to improve the separation between forests and open land. In the clear-cut area in Fig. 2, individual trees/stones appeared as dark areas on the map. To prevent misclassification of these features as forest, we applied the Gaussian and median filters to the red band to smooth out color variations, as the background orthophoto was most visible in the red band (Fig. 2B). We used the 'fast almost Gaussian filter' (Kovesi 2010) algorithm with a Sigma of 10 cells (Fig. 2L), which provided a blurred effect visually similar to a Gaussian filter but was approximately 10 times faster. A median filter with a kernel of 19 cells was also applied to reduce color variations (Fig. 2M). 3 standard deviations were calculated to highlight texture. The standard deviations were calculated on the red band with a kernel of 19 cells and on the intensity band with kernels of 11 and 49 cells (Fig. 2N–P). To mitigate the impacts of color variations between maps (Fig. 1B, D, F) from different production years, the average red, green, blue, intensity, hue, and saturation for each individual map were derived. In summary, 21 different features were calculated for all 938 economic maps (Fig. 1). The values of these features were then extracted to the digitized point layer with classified land use.

2.3. Tuning, training, feature reduction, and application of the final XGBoost model

The dataset of 16 654 classified land use sites with extracted feature values was first divided into a training dataset (80%) and a separate test data set (20%) based on the digitized map tiles. All points from 750 randomly selected map tiles were used to train the model, and all points from the remaining 181 map tiles were set aside for evaluation. This approach allowed us to evaluate the model performance on "unseen maps" with different colors due to different production years. The machine learning algorithm Extreme Gradient Boosting (XGBoost) was chosen because it is designed to optimize both computational efficiency and model performance (Chen and Guestrin 2016). It incorporates parallelization and tree-pruning techniques and utilizes a more optimized algorithm that reduces memory usage and computational overhead.

The hyperparameter tuning of the model was performed with Bayesian optimization using the Optuna framework version 3.5.0 (Akiba et al.,2019). Specifically, the Tree-structured Parzen Estimator (TPE) sampler was applied on the training set (Watanabe 2023). The objective was to maximize the mean Matthews Correlation Coefficient (MCC) obtained from all 5 cross-validation folds over 100 trials. MCC gives the best measure of overall model performance for imbalanced multi-class datasets (Delgado and Tibau 2019). The overall quality of the model performance was high for our initial model, which included all 21 features and explored a wide hyperparameter search space using TPE. However, clear indications of overfitting were observed as the loss curves of training and testing datasets diverged (Fig. 3A). Various techniques exist to reduce model overfitting and improve the model's generalizability on unseen data. The model complexity was reduced by adjusting the bounding values of the hyperparameter search space. We decreased the number of features, maximum depth, and learning rate, and increased L1 and L2 regularization parameters. We implemented subsampling and early stopping during model fitting. Training features were iteratively removed based on their importance and collinearity. Hyperparameters were fine-tuned iteratively, with adjustments made in steps of 20%, while monitoring improvements in overfitting without compromising overall model performance. The total decrease in MCC after all iterations was <5‰. The process was stopped when an additional iteration resulted in a decrease of MCC by 1.5%. Finally, the best model hyperparameters from the tuning process were used to train the XGBoost model with 100 boosting rounds (Fig. 3B). The trained model with the 12 remaining features was applied to predict land use on unseen maps. For more details, we refer to the Python scripts on GitHub (https://github.com/anneliagren/Land-Use-Mapping-from-Historical-Maps). All processing was performed on two local Linux servers running Ubuntu 22.04.4. Server 1 is equipped with two Intel Xeon Platinum 8362 CPUs (2.8 GHz, together 64 cores/128 threads) and 1024 GB RAM. Server 2 is equipped with one AMD Ryzen ThreadRipper 3990X CPU (2.9 GHz, 64 cores/128 threads) and 256 GB RAM.

2.4. Evaluation metrics and SHAP-values

The model performance was evaluated on the separate test dataset of 181 maps (20%). We evaluated model performance based on these measures: Recall (user's accuracy), Precision (producer's accuracy), F1-value (Powers 2011), Matthews Correlation Coefficient (MCC) (Matthews 1975), and Cohen's Kappa (κ) (Cohen 1960). Recall answers the question; out of all the points for a certain class, how many were retrieved? Precision answers the question: out of all the points that the model retrieved for a certain class, how many were classified correctly? F1 is the harmonic mean between precision and recall (Powers 2011). Cohen's Kappa and Matthews Correlation Coefficient can often be used interchangeably, except for certain conditions, where Matthews Correlation Coefficient behaves more robustly (Delgado and Tibau 2019).

To gain insights into how the XGB model classifies land use, we calculated the Shapley Additive exPlanations (SHAP) values. Derived from game theory, SHAP values provide a unified approach to explaining the output of any machine learning model (Winter, 2002; Li 2022). Originally it was developed to calculate how much an individual player contributed to a game (Lundberg and Lee 2017). We calculated a global measure of feature importance, defined as the mean absolute value of each feature over all instances of the dataset using the Python library "Shap" version 0.44.1, on model predictions during testing. The absolute SHAP-values were calculated to compute the contribution of each feature (Fig. 2) to the detection of each land-use class and not if the feature contribute positively or negatively to the prediction (Li and Managi 2024).

2.5. Adding water features to the classified raster maps

We adopted the method used by Auffret et al. (2017) by adding water features using a modern vector layer, as water and forest could not be separated effectively. We opted for the more detailed water polygons from the Swedish Property map (1:12 500)

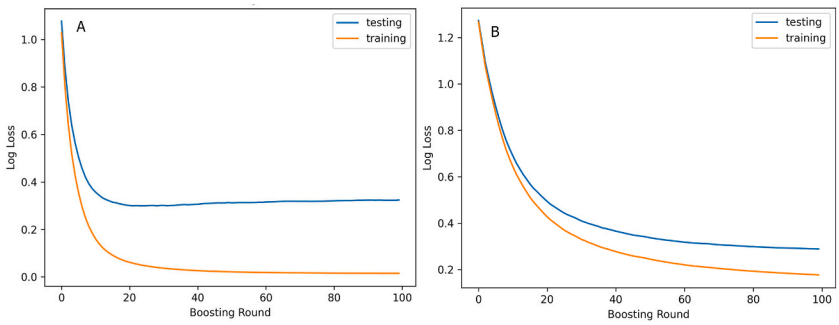

Fig. 3. Loss curves for A) Initial model with all 21 features. B) Final model after feature reduction and further tuning of the XGBoost model to reduce overfitting.

(downloaded 2021-08-09). The package gdalUtils was used to burn the modern water vector layer onto the final classified raster map (GDAL /OGR 2024). Given this approach, users of the maps should not use this map for detecting changes in water feature boundaries.

2.6. Post-processing

In some of the predicted maps, we observed the occasional irregularities in land use classification affecting one or a few cells (Fig. 4B). A post-processing step using the 'generalized classified raster' algorithm in Whitebox Tools v 2.3.0 was undertaken to remove the noise. This algorithm was preferred over a general smoothing filter (for example, Gaussian or majority), due to its capability to selectively modify without affecting feature boundaries within the classified raster (Lindsay 2023). The algorithm first identifies each contiguous group of cells in the input (i.e. a clumping operation) and then defines the subset of features that are smaller than the user-specified minimum feature size. In our case, the threshold was set to 9 cells to avoid removing small features representing footpaths. We used the method "longest" for class feature generalization, which reassigns small clumps with the same class value as the neighboring land use category with the longest shared border (Fig. 4C).

2.7. Mapping the northern inland of Sweden

While the 1:20 000 maps for the northern inland had been scanned, they had not yet been georeferenced and the borders around the maps remained (Fig. 5A). To address this issue, we implemented the following steps: 1) converted the original image to grayscale, 2) applied Otsu's thresholding (Otsu 1979) to create a binary image (Figs. 5B), 3) moved a bounding box of 5000 by 5000 pixels (red dashed box, Fig. 5) iteratively over the binary image, progressively decreasing the step size from 200, 100, 20, to 5 pixels. This process aimed to find the region of the target size that contains the most white pixels, corresponding to the map content area in the original map (green area), 4) cropped the original map with the location of the bounding box, and 5) the cropped image (Fig. 5C) was georeferenced using the shape file with the location of each file, in Fig. 1A. Apart from these extra steps, the maps were processed identically (calculating features and applying the classification model) as the rest of Sweden.

2.8. Comparison with Auffret et al. (2017) for the south of Sweden

Five of the digitized study areas overlapped with the area previously classified by Auffret et al. (2017) in their work on semi-automated land use digitization in south Sweden using R (https://github.com/AGAuffret/HistMapR/). We compared our manually digitized points with the predictions by Auffret et al. (2017), focusing only on points classified as arable land, open land, and forest. All points could be used for evaluation purposes since our digitized points were independent of the training data used in Auffret et al. (2017). Some key differences existed between the two methods. While both studies incorporated water features from modern-day maps after the final predictions, Auffret et al. (2017) included a separate class for pixels that could not be classified, labeled as "borders". Due to these class discrepancies, direct comparisons of the total maps were impossible. However, we could perform comparison for the individual classes of arable land, forest, and open land, which were present in both studies.

2.9. Optional removal of graphics

Finally, we provided the option to remove graphic elements (e.g. contour lines, ownership borders, place names) from the maps. Classifying graphics as a separate class provided us the option to use it as a mask to remove them while preserving as much of the original land use data as possible. First, all graphics were converted to a value indicating missing data, commonly referred to as 'Nodata'. Then we employed the GDAL (GDAL /OGR 2024) function 'gdal.FillNodata' to fill the 'NoData' values. It started by searching within a neighborhood of 3 x 3 pixels and exponentially increasing neighborhood sizes (3, 6, 12, 24, 48, etc. pixels) until all 'NoData' cells were filled. Progressively increasing the search area allowed for more efficient computations while preserving local variations in the data.

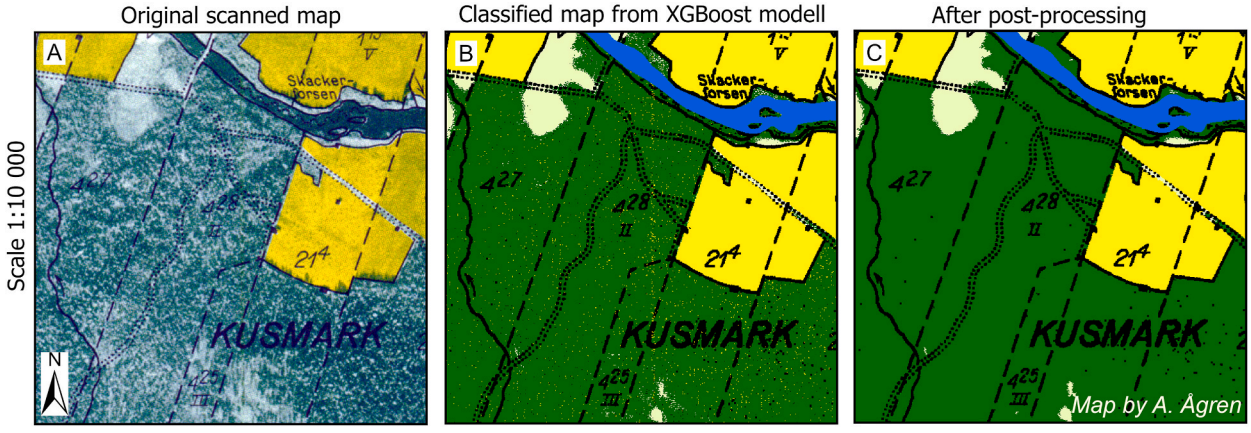

Fig. 4. A) The figure shows an example of the original scanned map, B) the classified land use map from the XGBoost model, and C) the map after the post-processing step after the removal of individual misclassified small "clumps".

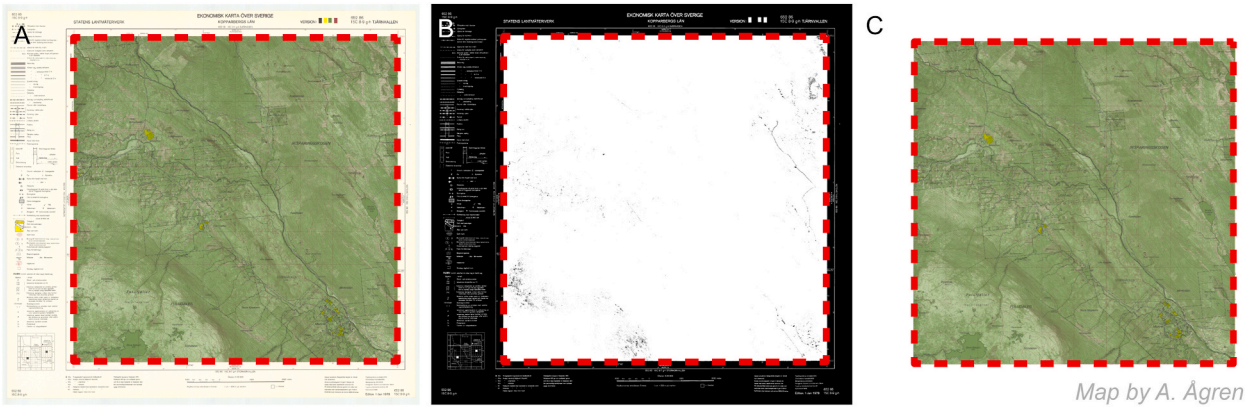

Fig. 5. Example of automated cropping of images. A) Original map. B) Binary image from Otsu's thresholding. C) Cropped map.

3. Results

3.1. Quality measures of the maps

The overall quality of the classified land use maps was high, as indicated by the statistical evaluation using the test dataset consisting of 181 maps (Table 1). While we presented accuracy for comparisons with other studies, we considered κ and MCC to be more robust metrics for the overall quality of the map, given the slightly unbalanced nature of our dataset. The accuracy for our final model was 0.90, κ and MCC both achieved a score of 0.86. Graphics and arable land had the highest performance for the individual classes, with F1 scores above 0.90. The most difficult land use to distinguish/identify was the forest and open land, with F1-scores of 0.84 and 0.87, respectively.

3.2. Feature importance with SHAP values

The SHAP values (Fig. 6) quantify the contribution of each feature to XGBoost's ability to accurately classify each land-use class. The length of each bar within a feature represents the feature's relative importance for individual classes. The cumulative length of all bars for a given feature reflects its overall impact on the model across all classes. For graphics, the minimum filter on the green band, green/red, and average green were the most important features. Blue/red and green/red contributed the most to classifying arable

Table 1
Precision, Recall, F1-score on the testing dataset for each class, and the overall performance measures of the model; Accuracy, Cohen's kappa (κ), Matthews Correlation Coefficient (MCC), and number of observations (n). The table presents the measures from the final XGBoost model after feature reduction and further tuning of the model to reduce overfitting (see section 2.3).

	Precision	Recall	F1-score	Accuracy	κ	MCC	n
Graphics	0.97	0.98	0.98				716
Arable	0.97	0.92	0.95				468
Forest	0.83	0.86	0.84				930
Open	0.88	0.86	0.87				976
Total			**0.91**	**0.90**	**0.86**	**0.86**	

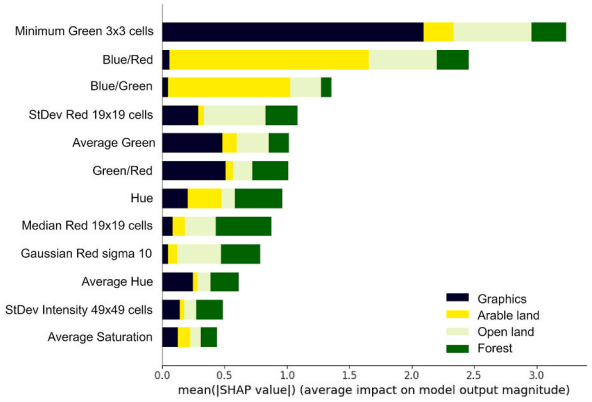

Fig. 6. The absolute SHAP values for the XGBoost model.

land. Classifying open land relied on a wider range of features, with the most important ones being the minimum filter on the green band, the standard deviation of the red band with a kernel of 19 cells, blue/red, and a Gaussian filter with a sigma of 10. Forest classification primarily involved the median filter on the red band with a 19 cells kernel, hue, and a Gaussian filter with a sigma of 10.

3.3. Visualization of the predicted maps

Visual inspection of the predicted maps and the original maps revealed a high degree of accuracy (Fig. 7A and B). There was also a high level of agreement between our new fully automated land use classification with the classification by Auffret et al. (Fig. 7B and C). The primary difference between our methods lay in the handling of graphics, which we maintained as a separate class, while Auffret et al. masked them out during preprocessing. This was especially evident in urban areas with high density of graphics (Fig. 7D–F). Both models delineated open and forest land based on the brightness or darkness of the orthophoto, which led to some limitations. Light-colored deciduous forests were misclassified as open areas (Fig. 7G–I), and some dark open areas (often wet soils) were misclassified as forests (Fig. 7J–L). The percentage of graphics ranged 0–27% on the evaluation maps. We observed an increase in graphics over time, from an average of ca 5% on the maps produced in the early 40s to ca 12% in the late 70s (Fig. 8).

3.4. Statistical comparison with Auffret et al., (2017) for the south of Sweden

In addition to the visual comparison with Auffret et al. (2017), we also compared Auffret et al.'s maps statistically with our manually digitized points. We compared the performance of the individual classes' arable land, forests, and open land, excluding classes that were not represented in both maps. Both models achieved high recall, or user's accuracy, for arable land, specifically, 0.92 in our study and 0.93 in Auffret et al. (2017). This indicates that 92% and 93% of the actual arable land were correctly identified by the models. Our model demonstrated higher precision, or producer's accuracy, for arable land 0.97 (Table 1) compared to Auffret et al.'s 0.58 (Table 2). This metrics measures the proportion of correctly identified arable land out of all pixels identified as arable land. For forest and water, our model had a lower recall 0.86 but higher precision 0.83 in comparison to Auffret et al.'s higher recall 0.91, and lower precision 0.58. For open areas, our model outperformed that of Auffret et al.'s by ca 10%, with recall and precision values of 0.86 and 0.88, respectively (Table 1), compared to 0.70 and 0.75, for Auffret et al. (Table 2). Based on F1 scores, which is the harmonic mean for precision and recall, our model generally performed better than the previous model by Auffret et al., with F1 scores 0.24 higher on arable land, 0.19 higher for forest, and 0.14 higher for open areas (Table 2).

3.5. Masking away graphics

By using the classified graphic as a mask (Fig. 9B), we could remove the graphics without changing any other cells in the maps and, therefore, maintain the integrity of the land use classes (Fig. 9C). For the majority of the map tiles three iterations (filter size 3x3, 6x6, and 12x12) was enough to fill all NoData cells, and mask away all graphics. Another common way to remove graphics is to use a general smoothing filter (Fig. 9D) as applied by Auffret et al., (2017). Which is similar to our post-processing step (Section 2.6, Fig. 4), but, with a larger kernel. However, this affects more pixels across the landscape and introduce more uncertainty in the maps (Fig. 9D).

4. Discussion

4.1. Land use classification model

4.1.1. Objective and significance of land use classification

The purpose of classifying the land use in these historical maps is to generate a dataset that can be used to detect changes in land use (Auffret et al.,2018) and learn more about the interactions on species diversity in a changing world where both climate and land use are changing (Auffret and Svenning 2022). We also aim for these maps to serve as planning tools for rewetting nutrient-rich peatlands, as historical land use may have long-lasting effects on soil carbon (Schulp and Verburg 2009), microbial diversity (Li et al., 2021), and N_2O emissions (Klemedtsson et al.,2005). This article focuses on the method development of detecting land use in historical maps and assessing the machine learning model classified maps for their intended use.

4.1.2. Modern land use data quality and XGBoost model assessment

As background to our classification, we examined the quality of modern-day land use data for Sweden (NMD), at 10 m resolution, derived from a combination of 41 different data layers of satellite imagery, laser data, digital terrain indices, current vector maps, etc. (Naturvårdsverket 2023). The NMD maps identify forests, wetlands, arable land, other open land, urban areas, and water. The quality of the NMD maps was assessed based on regional precision for Sweden, divided into 4 regions based on field data (Nilsson et al., 2020). The precision was (the range among the 4 regions is given in the in brackets): forest (0.93–0.97), arable land (0.96–0.97), wetland (0.91–0.98), other open land (0.38–0.75), urban land (0.46–0.58), water (0.99–100). Nilsson et al. (2020) further concluded that a precision above 0.8 is very good, 0.7–0.8 is good, 0.6–0.7 is acceptable, but below 0.6 there is a poor agreement between the map and reality.

Our machine learning model could not be directly assessed with field data, due to the lack of accurate positioning of the available field data during the map production years. Hence, we partitioned our digitized point layer into a separate evaluation dataset. Our XGBoost model had a precision of 0.97, 0.83, and 0.88 for arable land, forest and open areas, respectively. The inherent uncertainty of the original paper maps (García-Alvarez et al.,2019) is acknowledged but remains nonquantifiable, as the only information given by the National Land Survey on this, is that newer maps are more accurate than old maps.

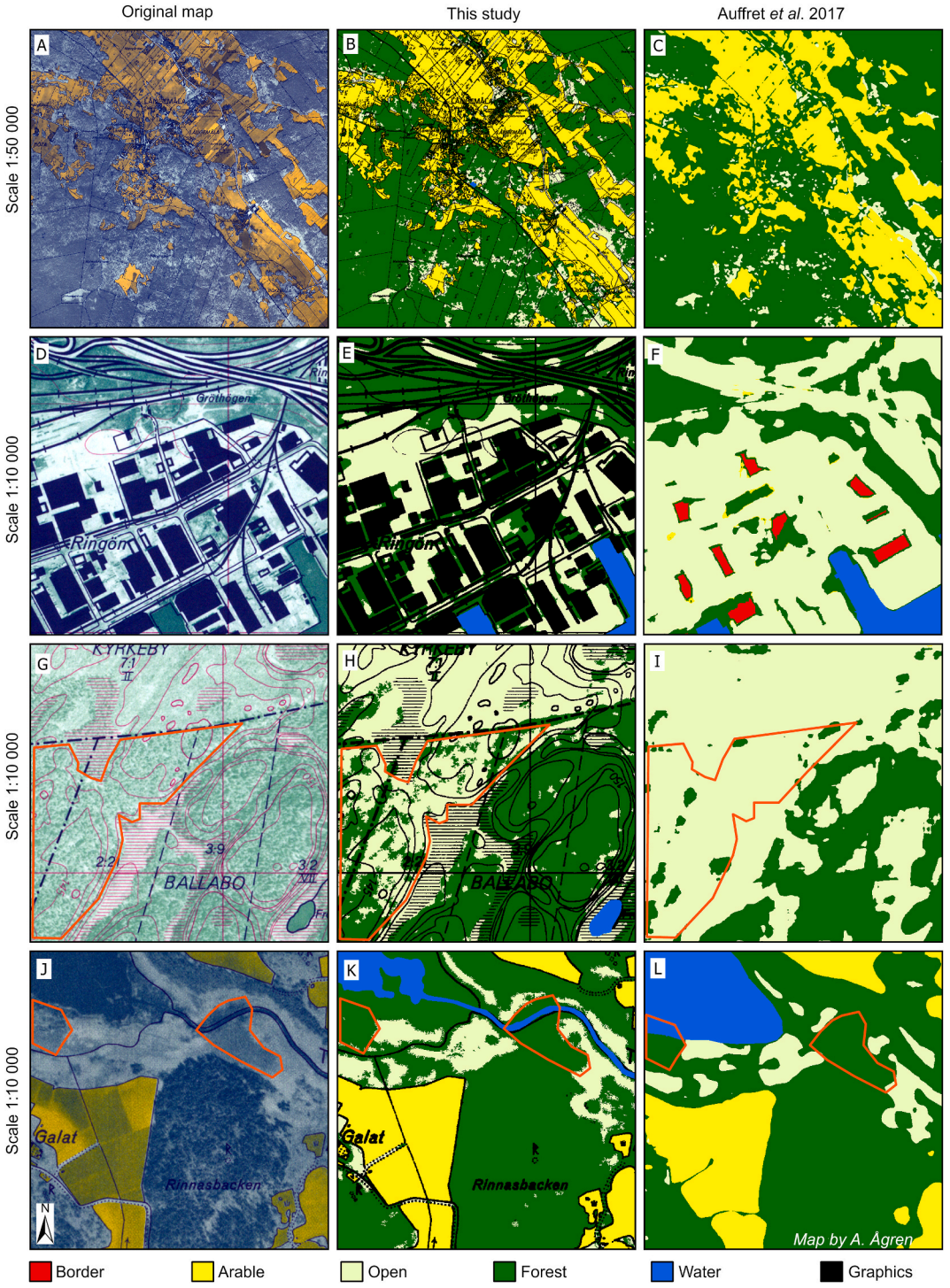

Fig. 7. Comparison of land use classification for four different areas in southern Sweden. The left panels show the original scanned maps, the center panels our land use classification, and the right panels display the classification by Auffret et al., (2017). Panels A–C show an entire map tile compressed to 1:50 000 while, panels D–L show zoomed-in sections of maps originally produced at a scale of 1:10 000. Orange polygons exemplify misclassified areas. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

4.1.3. Challenges and solutions in historical orthomap classification

Both our XGBoost method and Auffret's model encountered difficulties in correctly mapping dark open areas (Fig. 7K and L), and light forests (Fig. 7H and I). In a separate model run (not shown) we digitized a further 5500 points representing dark open areas and

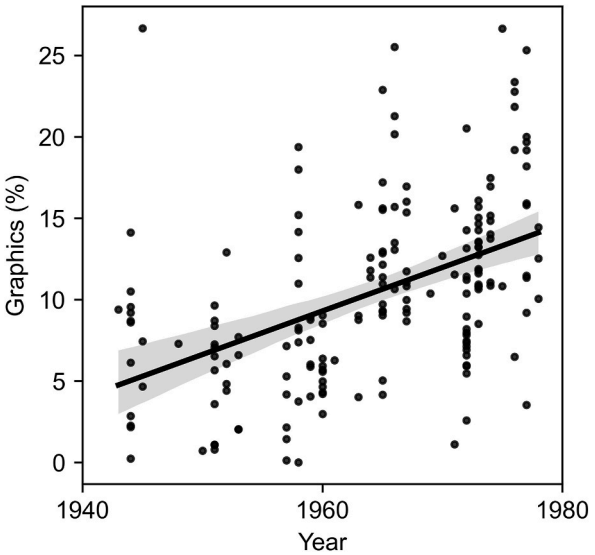

Fig. 8. Linear regression reveals a significant increase in graphics content of historical maps over time ($p < 0.001$, $R^2 = 0.22$). The grey shaded area depicts the 95% confidence interval for the regression line.

Table 2
Precision, Recall, F1-score, and number of observations (n) for the map by Auffret et al. (2017). The map was statistically evaluated against all of our training and testing points where the study areas overlapped. Furthermore, we show the difference between our map and Auffret et al.'s. A positive number indicates that our map has a higher quality, while a negative number indicates that Auffret's map has a higher quality.

	Precision	Recall	F1 score	n	Difference in precision	Difference in recall	Difference in F1
Arable	0.58	0.93	0.71	1740	0.39	−0.01	0.24
Forest	0.51	0.91	0.65	3022	0.32	−0.05	0.19
Open	0.75	0.70	0.73	3246	0.13	0.16	0.14

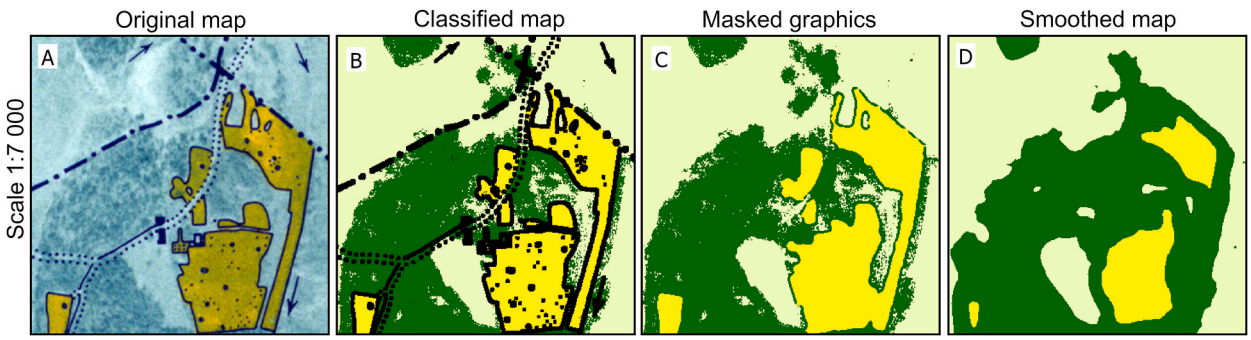

Fig. 9. All four maps show the same area. A) Original map. B) Our classified map. C) The map after masking graphics in our study. D) Classified map by Auffret et al., (2017), who applied a general smoothing filter in the first step of land use classification.

light forests. The intention was to improve the classification with additional training data, allowing it to learn to differentiate land use based on both color (light vs. dark) and texture (smooth vs. "textured"). However, this attempt to improve performance instead introduced noise in the model, leading to the deterioration of predictive performance. Hence, we opted to revert back to the original, more well-balanced dataset. Our analysis suggested that classification based on light and dark areas of the orthophoto would yield more accurate predictions overall. We observed illumination non-uniformity in some of the scanned maps (visible in Fig. 1B), characterized by lighter and darker stripes across the maps (Baltsavias 1994). These artifacts affected the land use classification in some of the images. Underexposed dark bands were over-classified as forest, while overexposed light bands were over-classified as open land. This banding was most pronounced for the northern inland maps, but, the same patterns were also present in the maps by Auffret et al., (2017) (not shown). To potentially improve the classification of forest and open land, we recommend evaluated a novel deep learning model Correcting Uneven Illumination Network (CUI-Net) for correcting light and dark stripes in the original images (Chao et al., 2023).

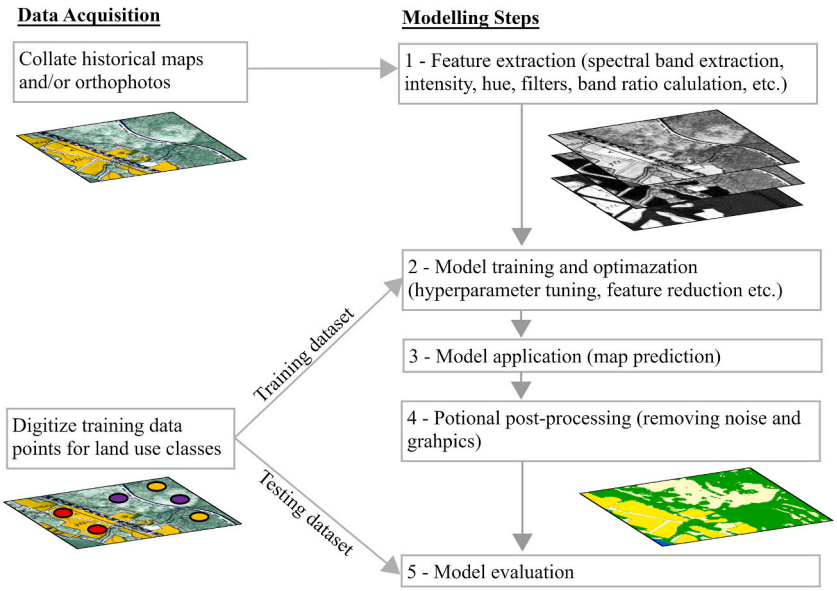

Fig. 10. Modeling Framework for Classifying Land Use from Pre-Satellite Historical Data. Python scripts for all calculations and instructions are found on https://github.com/anneliagren/Land-Use-Mapping-from-Historical-Maps.

The extent of graphic coverage on the maps depended on the production year, with more graphics added over time (Fig. 8). Among our 181 evaluation maps, graphics covered between 0 and 26%. Since graphics obscured the underlying photo, it was challenging to determine the land use hidden underneath. However, by assigning graphics to a separate class, we could quantify the uncertainty introduced by the overall coverage of the graphics (Fig. 8). Having a separate class allows users to easily mask the graphics away and interpolate with the surrounding land use if needed (Fig. 9). While the masking and interpolating method works well for rural areas where most of the graphics are wetlands, topographic lines, and ownership boundaries, it will introduce a larger bias for urban areas (Fig. 7D–F). Hence, we opted to retain the graphic as a separate class, allowing potential future map users the flexibility depending on their research questions. Additionally, we have included a Python script to facilitate efficient removal of graphics when desired. We observed a further complication in the analysis. Recent maps, characterized by a higher graphic content (Fig. 8), exhibited a reduction in background orthophoto contrast to enhance graphic visibility. This contrast reduction appeared to have negatively impacted land-use classification performance, particularly for forest and open land categories, as evidenced by the test map evaluations.

The automated clipping of the maps of northern inland generally performed well. However, some maps experienced distortion from scanning (often a minor rotation). The misalignment with the 5000x5000 cell window has, in some cases, resulted in partial map loss or map border inclusion in the clipping area. The location uncertainty introduced by scanning distortion amounted to ca 1%. To make the clipping more efficient, we enforced early stopping when the bounding box was within 5 m of the optimal placement. Except for the distorted map areas, we argue that this method still surpasses modern land use classification, which has a 10 m resolution.

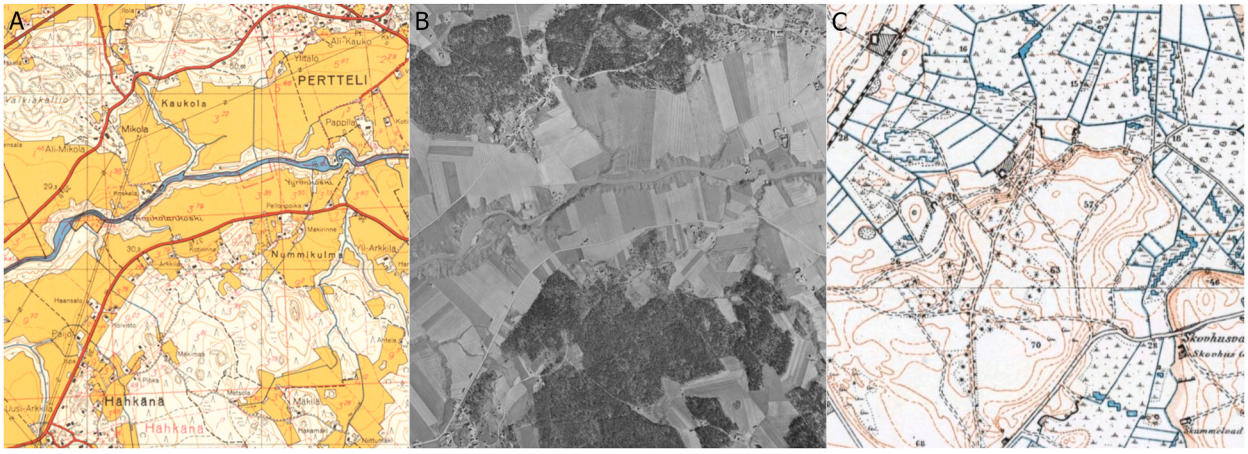

Fig. 11. Example of a Finnish basic map 1:20 000 (1964) (A) and a historical orthophoto (1959) (B) of the same area, from National Land Survey of Finland. The Danish topographical map Lave målebordsblade (C) 1:20 000, 1901–1971, Generalstaben.

4.2. Can the maps be used for detecting land-use changes over time?

4.2.1. Arable land

Ensuring the trustworthiness of land-use maps is important as they are crucial tools for understanding some of the world's most pressing challenges like biodiversity loss and climate change (Griffin 2020). Accurate assessment of land-use change from maps necessitates the estimation of map uncertainty (Prestby 2023). To properly evaluate the uncertainty, we calculated several measures of uncertainty both for the entire map and for the individual classes (Table 1). The uncertainties of the maps also have to be considered in the context of land-use change over time. Statistical records indicate a 21% decrease in arable land in Sweden from the period 1910–1950 to the year 2023 (Swedish Board of Agriculture 2011; Swedish Official Statistics 2023). While this data comes from the farmers' self-reporting to the board, which lacked georeferencing (before 1998), most of the individual fields can be identified on our model-produced maps. Our model had an F1-value of 0.95 for arable land (Table 1), and modern-day land use maps had a precision of 0.96–0.97 (Nilsson et al., 2020). These metrics suggest that a few local arable sites might be mislabeled by the model, and one should not trust the maps implicitly on the local level. However, when examining arable land at regional scales, the low uncertainty associated with both our model-produced maps and modern land use maps suggests minimal impact on the overall results.

4.2.2. Forest and open land

For forest land and open land, which were harder to detect, we can estimate that our map has an uncertainty of around 13–16% based on the F1 values (Table 1). In comparison, the modern-day land use maps had 3–7% uncertainty based on the producer's accuracy for forest land, while the uncertainty for wetland was 2–9% and other open land 25–68% (Nilsson et al., 2020). The higher uncertainties for forest and open land in both our and existing maps, limit the effective detection of small-scale land-use changes. Based on field measurements from the National Forest Inventory the productive forest land has increased by 5% from 1926 to 2020, while mires decreased by 13% over the same period (Swedish Official Statistics). While Auffret et al. (2018) reported a 17% decline in total open land cover in southern Sweden, our evaluation indicated an uncertainty of 27% for open land in their maps (Table 2). Furthermore, the broad range of uncertainty (25–68%) associated with open land in modern maps (Nilsson et al., 2020) undermines the certainty of this reported land-use change. Despite having a higher overall quality (14–24% more accurate (Table 2)), our map still exhibits an uncertainty of 13% for open land (Table 2). We therefore advise using the map with caution. Given the higher uncertainty of the classes forest and open land for both modern and historical maps, the suitability of the maps for detecting land use changes for those land use classes is questionable. Therefore, we recommend users conduct further investigations at more regional and local scales to assess the uncertainty further.

4.3. Applying the model to other countries

Long-term records of land use change require alternative data sources beyond satellites, as satellites dedicated to land-cover observation were not launched until 1972 (Belward and Skoien 2015). Historical maps and orthophotos provide valuable tools for detecting land use changes extending back more than 250 years. This approach is demonstrated by a Czech case study where historical land use was manually interpreted and digitized as vector data (Skalos et al.,2011). Manual digitization remains a viable option for small study areas (Skalos et al.,2011) and smaller countries, such as Latvia (65 300 km^2) (Piskinaite and Veteikis 2023), or even for larger countries when using low resolutions, as shown in a study of India at 5 arc-minute resolution (Tian et al., 2014). However, detailed manual interpretation of historical maps and orthophotos at scales of 1:10 000 to 1:20 000 is prohibitively time-consuming and expensive to perform across entire countries. Semi-automated methods using R and GRASS have been employed to detect land use in southern Sweden and Italy (Auffret et al.,2017; Gobbi et al., 2019). In this study, we developed a fully automated method for land use classification that could be applied across an entire country (470 000 km^2) at high resolution (1 m) by leveraging efficient calculations with parallel processing using Python scripts running on Linux servers. This approach allowed us to map the entire country within ca. two months of full time work. The same methodology can be adapted to other countries with similar datasets (Fig. 10).

Monochrome orthophotos became widespread in the 1930s–1960s and changed how maps were constructed. The production of photomaps began after the First World War, initially intended for military purposes but eventually expanding to civil applications as well (Baltsavias 1996). Maps based on orthophotos gained widespread popularity because orthophotomaps provided a wealth of ground details and information that could not be captured by standard vector maps (McKenzie 1973). The Swedish economic map is an orthophotomap with the orthophoto in the map background and interpreted vectorgraphics on top. Most European countries has similar maps that are in the scale of ca 1:10 000–1: 50 000, and were drawn based on aerial photography. However, in some European countries, these datasets can sometimes exist separately. In the Finnish basic map (1:20 000, 1949–1998, National Land Survey of Finland) (Fig. 11A), the land uses water, arable land, and graphics can be classified from the maps, while open land and forest land can be classified from the orthophotos (Fig. 11B) and then used to update the white background areas in the map (Fig. 11A). Unlike many modern maps, early geographical information was not standardized across countries. As a result, our tuned model cannot be directly applied to maps from other countries. However, the methods (https://github.com/anneliagren/Land-Use-Mapping-from-Historical-Maps) can be easily transferred to any country with historical maps and/or orthophotos from the same period (Fig. 10). For maps that use a greater number of symbols to indicate different land uses within polygons, such as the Danish topographical map (Lave målebordsblade, 1:20 000, 1901–1971, Generalstaben) (Fig. 11C), we suggest that a deep learning approach might be more effective. The complexity and diversity of symbols in such maps could be better interpreted by advanced deep learning techniques, which classify land use based on image segmentation techniques, possibly with a U-Net model (Mäyrä et al.,2023).

5. Conclusions

Our method demonstrates the potential for automated, high-quality arable land mapping. The maps can thus facilitate the identification of afforested former arable sites, which are likely to be nutrient-rich sites. Such nutrient rich sites often displays a high growth and yield, high biodiversity, and can be good candidates for rewetting, which can lead to tradeoffs for best management practice. Very little is known about nutrient-rich drained peatlands in northern regions; this map is, therefore, an important contribution to further research. However, with the developed methods we could not get a good estimation of open and forest land use change, despite F1 values above the threshold for "very good" suggested by Nilsson (>80%), as the magnitude of change for these land uses fall within the range of uncertainty estimates. To promote further research and broader application, the code for this method is publicly available on GitHub (https://github.com/anneliagren/Land-Use-Mapping-from-Historical-Maps).

Code availability

All code is openly available and published on GitHub under The Creative Commons CC0 Public Domain Dedication (https://github.com/anneliagren/Land-Use-Mapping-from-Historical-Maps)

Data sources

The original maps can be downloaded from: https://www.lantmateriet.se/sv/geodata/vara-produkter/produktlista/ekonomiska-kartan/

Declaration of generative AI

During the preparation of this work the authors used GitHub Copilot to assist in writing the Python scripts. Following this, the authors thoroughly reviewed and edited the content and take full responsibility for the content of the publication.

Ethical statement

No other related manuscripts are under consideration or in press elsewhere by the authors. We agree and comply with the general terms and general obligations of the authors.

None of the authors of this article have coauthored an article with any of the suggested reviewers.

CRediT authorship contribution statement

Anneli M. Ågren: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Yiqi Lin:** Writing – review & editing, Software.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Anneli Agren reports financial support was provided by Swedish Reserch Council Formas. Anneli Agren reports financial support was provided by Knut and Alice Wallenberg Foundation. Anneli Agren reports a relationship with Knut and Alice Wallenberg Foundation that includes: funding grants. Anneli Agren reports a relationship with Swedish Research Council Formas that includes: funding grants. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All code is openly available and published on GitHub under The Creative Commons CC0 Public Domain Dedication (https://github.com/anneliagren/Land-Use-Mapping-from-Historical-Maps)

Acknowledgements

This work was funded by The Swedish Research Council Formas (proj no. 2021-00713, 2021-00115) and Knut and Alice Wallenberg Foundation (2018.0259 Future Silviculture). The funding sources had no involvement in study design, collection, analysis and interpretation of data, nor in the writing of the report.

References

2016/17:104 A National Food Strategy for Sweden – more jobs and sustainable growth throughout the country. Short Version of Government Bill 2016/17:104. Government Offices of Sweden, Ministry of Enterprise and Innovation. Elanders. 24 p. Stockholm.

Abdi, A.M., 2020. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. GIScience Remote Sens. 57, 1–20. https://doi.org/10.1080/15481603.2019.1650447.

Ågren, A.M., Larson, J., Paul, S.S., Laudon, H., Lidberg, W., 2021. Use of multiple LIDAR-derived digital terrain indices and machine learning for high-resolution national-scale soil moisture mapping of the Swedish forest landscape. Geoderma 404, 115280. https://doi.org/10.1016/j.geoderma.2021.115280.

Ågren, A.M., Hasselquist, E.M., Stendahl, J., Nilsson, M.B., Paul, S.S., 2022. Delineating the distribution of mineral and peat soils a the landscape scale in northern boreal regions. Soil 8, 733–749. https://doi.org/10.5194/soll-8-733-2022.

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: a next-generation hyperparameter optimization framework. Kdd'19: Proceedings of the 25th Acm Sigkdd International Conferencce on Knowledge Discovery and Data Mining, pp. 2623–2631. https://doi.org/10.1145/3292500.3330701.

Auffret, A.G., Svenning, J.C., 2022. Climate warming has compounded plant responses to habitat conversion in northern Europe. Nat. Commun. 13 https://doi.org/10.1038/s41467-022-35516-7. ARTN 7818.

Auffret, A.G., Kimberley, A., Plue, J., Skånes, H., Jakobsson, S., Waldén, E., Wennbom, M., Wood, H., Bullock, J.M., Cousins, S.A.O., Gartz, M., Hooftman, D.A.P., Tränk, L., 2017. HistMapR: rapid digitization of historical land-use maps in R. Methods Ecol. Evol. 8, 1453–1457. https://doi.org/10.1111/2041-210x.12788.

Auffret, A.G., Kimberley, A., Plue, J., Waldén, E., 2018. Super-regional land-use change and effects on the grassland specialist flora. Nat. Commun. 9 https://doi.org/10.1038/s41467-018-05991-y. ARTN 3464.

Baltsavias, E.P., 1994. Test and calibration procedures for image scanners. In: ISPRS Commission I Symposium, pp. 163–170. https://doi.org/10.3929/ethz-a-004334530. Como, Italy, September 12-16: IAPRS.

Baltsavias, E.P., 1996. Digital ortho-images - a powerful tool for the extraction of spatial- and geo-information. ISPRS J. Photogrammetry Remote Sens. 51, 63–77. https://doi.org/10.1016/0924-2716(95)00014-3.

Bellemare, J., Motzkin, G., Foster, D.R., 2002. Legacies of the agricultural past in the forested present: an assessment of historical land-use effects on rich mesic forests. J. Biogeogr. 29, 1401–1420. https://doi.org/10.1046/j.1365-2699.2002.00762.x.

Belward, A.S., Skoien, J.O., 2015. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. ISPRS J. Photogrammetry Remote Sens. 103, 115–128. https://doi.org/10.1016/j.isprsjprs.2014.03.009.

Chao, K., Song, W., Shao, S., Liu, D., Liu, X.C., Zhao, X.B., 2023. CUI-Net: a correcting uneven illumination net for low-light image enhancement. Sci. Rep. 13 https://doi.org/10.1038/s41598-023-39524-5. ARTN 12894.

Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, San Francisco, California, USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educational and Psychologial Measurment 20, 37–46. https://doi.org/10.1177/001316446002000104.

Delgado, R., Tibau, X.A., 2019. Why Cohen's Kappa should be avoided as performance measure in classification. PLoS One 14. https://doi.org/10.1371/journal.pone.0222916.

García-Alvarez, D., Van Delden, H., Olmedo, M.T.C., Paegelow, M., 2019. Uncertainty challenge in geospatial analysis: an approximation from the land use cover change modelling perspective. Geospatial Challenges in the 21st Century 289–314. https://doi.org/10.1007/978-3-030-04750-4_15.

GDAL/OGR, 2024. Geospatial Data Abstraction software. Library, Open Source Geospatial Foundation. https://doi.org/10.5281/zenodo.5884351.

Gobbi, S., Ciolli, M., La Porta, N., Rocchini, D., Tattoni, C., Zatelli, P., 2019. New tools for the classification and filtering of historical maps. ISPRS Int. J. Geo-Inf. 8 https://doi.org/10.3390/ijgi8100455. ARTN 455.

Griffin, A.L., 2020. Trustworthy maps. Journal of Spatial Information Science 5–19. https://doi.org/10.5311/Josis.2020.20.654.

Gustavsson, E., Lennartsson, T., Emanuelsson, M., 2007. Land use more than 200 years ago explains current grassland plant diversity in a Swedish agricultural landscape. Biol. Conserv. 138, 47–59. https://doi.org/10.1016/j.biocon.2007.04.004.

Högberg, P., Wellbrock, N., Högberg, M.N., Mikaelsson, H., Stendahl, J., 2021. Large differences in plant nitrogen supply in German and Swedish forests - implications for management. For. Ecol. Manag. 482 https://doi.org/10.1016/j.foreco.2020.118899. ARTN 118899.

Horn, B., Ferreira, C., Kalantari, Z., 2022. Links between food trade, climate change and food security in developed countries: a case study of Sweden. Ambio 51, 943–954. https://doi.org/10.1007/s13280-021-01623-w.

Houghton, R.A., House, J.I., Pongratz, J., van der Werf, G.R., DeFries, R.S., Hansen, M.C., Le Quéré, C., Ramankutty, N., 2012. Carbon emissions from land use and land-cover change. Biogeosciences 9, 5125–5142. https://doi.org/10.5194/bg-9-5125-2012.

Kasimir-Klemedtsson, A., Klemedtsson, L., Berglund, K., Martikainen, P., Silvola, J., Oenema, O., 1997. Greenhouse gas emissions from farmed organic soils: a review. Soil Use Manag. 13, 245–250. https://doi.org/10.1111/j.1475-2743.1997.tb00595.x.

Keenan, R.J., Reams, G.A., Achard, F., de Freitas, J.V., Grainger, A., Lindquist, E., 2015. Dynamics of global forest area: results from the FAO global forest Resources assessment 2015. For. Ecol. Manag. 352, 9–20. https://doi.org/10.1016/j.foreco.2015.06.014.

Klemedtsson, L., von Arnold, K., Weslien, P., Gundersen, P., 2005. Soil CN ratio as a scalar parameter to predict nitrous oxide emissions. Global Change Biol. 11, 1142–1147. https://doi.org/10.1111/j.1365-2486.2005.00973.x.

Kovesi, P., 2010. Fast almost-Gaussian filtering. In: 2010 International Conference on Digital Image Computing: Techniques and Applications, pp. 121–125. https://doi.org/10.1109/DICTA.2010.30.

Laudon, H., Hasselquist, E.M., Peichl, M., Lindgren, K., Sponseller, R., Lidman, F., Kuglerová, L., Hasselquist, N.J., Bishop, K., Nilsson, M.B., Ågren, A.M., 2021. Northern landscapes in transition: evidence, approach and ways forward using the Krycklan Catchment Study. Hydrol. Process. 35, e14170 https://doi.org/10.1002/hyp.14170.

Laudon, H., Lidberg, W., Sponseller, R.A., Hasselquist, E.M., Westphal, F., Ostlund, L., Sandstrom, C., Jarveoja, J., Peichl, M., Agren, A.M., 2022. Emerging technology can guide ecosystem restoration for future water security. Hydrol. Process. 36 https://doi.org/10.1002/hyp.14729. ARTN e14729.

Li, Z.Q., 2022. Extracting spatial effects from machine learning model using local interpretation method: an example of SHAP and XGBoost. Comput. Environ. Urban Syst. 96 https://doi.org/10.1016/j.compenvurbsys.2022.101845. ARTN 101845.

Li, C., Managi, S., 2024. Mental health and natural land cover: a global analysis based on random forest with geographical consideration. Sci. Rep. 14 https://doi.org/10.1038/s41598-024-53279-7. ARTN 2894.

Li, J.Y., Hua, Z.S., Liu, T., Wang, C.W., Li, J., Bai, G., Lucker, S., Jetten, M.S.M., Zheng, M., Guo, J.H., 2021. Historical land use has long-term effects on microbial community assembly processes in forest soils. Isme Communications 1. https://doi.org/10.1038/s43705-021-00051-x. ARTN 48.

Lidberg, W., Paul, S.S., Westphal, F., Richter, K.F., Lavesson, N., Melniks, R., Ivanovs, J., Ciesielski, M., Leinonen, A., Agren, A.M., 2023. Mapping drainage ditches in forested landscapes using deep learning and aerial laser scanning. J. Irrigat. Drain. Eng. 149 https://doi.org/10.1061/Jidedh.Ireng-9796. ARTN 04022051.

Lindsay, J.B., 2023. WhiteboxTools User Manual. Geomorphometry and Hydrogeomatics Research Group, University of Guelph, Guelph, Canada, version 2.3.0.

Loucks, C., Ricketts, T.H., Naidoo, R., Lamoreux, J., Hoekstra, J., 2008. Explaining the global pattern of protected area coverage: relative importance of vertebrate biodiversity, human activities and agricultural suitability. J. Biogeogr. 35, 1337–1348. https://doi.org/10.1111/j.1365-2699.2008.01899.x.

Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Curran Associates Inc, Long Beach, California, USA, pp. 4768–4777. https://doi.org/10.48550/arXiv.1705.0787.

Matthews, B.W., 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim. Biophys. Acta Protein Struct. 405, 442–451. https://doi.org/10.1016/0005-2795(75)90109-9.

Mäyrä, J., Kivinen, S., Keski-Saari, S., Poikolainen, L., Kumpula, T., 2023. Utilizing historical maps in identification of long-term land use and land cover changes. Ambio 52, 1777–1792. https://doi.org/10.1007/s13280-023-01838-z.

McKenzie, M.L., 1973. Photoimages for map bases. J. Res. U. S. Geol. Surv. 1, 327–339.

Naturvårdsverket, 2023. Nationella Marktäckedata 2018 Basskikt, Produktbeskrivning, p. 59. Stockholm, Version 2.3.

Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Lysenko, I., Senior, R.A., Börger, L., Bennett, D.J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M.J., Feldman, A., Garon, M., Harrison, M.L.K., Alhusseini, T., Ingram, D.J., Itescu, Y., Kattge, J., Kemp, V., Kirkpatrick, L., Kleyer, M., Correia, D.L.P., Martin, C.D., Meiri, S., Novosolov, M., Pan, Y., Phillips, H.R.P., Purves, D.W., Robinson, A., Simpson, J., Tuck, S.L., Weiher, E., White, H.J., Ewers, R.M., Mace, G.M., Scharlemann, J.P.W., Purvis, A., 2015. Global effects of land use on local terrestrial biodiversity. Nature 520, 45–50. https://doi.org/10.1038/nature14324.

Newbold, T., Hudson, L.N., Arnell, A.P., Contu, S., De Palma, A., Ferrier, S., Hill, S.L.L., Hoskins, A.J., Lysenko, I., Phillips, H.R.P., Burton, V.J., Chng, C.W.T., Emerson, S., Gao, D., Pask-Hale, G., Hutton, J., Jung, M., Sanchez-Ortiz, K., Simmons, B.I., Whitmee, S., Zhang, H.B., Scharlemann, J.P.W., Purvis, A., 2016. Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. Science 353, 288–291. https://doi.org/10.1126/science.aaf2201.

Nilsson, M., Ahlkrona, E., Jönsson, C., Allard, A., 2020. Regionala Jämförelser Mellan Nationella Marktäckedata Och Fältdata Från Riksskogstaxeringen Och NILS. Sveriges lantbruksuniversitet & Naturvårdsverket, Umeå & Stockholm, p. 34.

NumPy, D., Steppi, A., van Beek, B., Harris, C., Laíns, F., Vanderplas, J., Yan, L., van Kerkwijk, M., Picus, M., Weber Mendonça, M., Shanker, N., Goldbaum, N., Gommers, R., Goswami, R., Adel, S., Berg, S., van der Walt, S., Reddy, T., Weckesser, W., 2023. NumPy user guide, version 1.26. https://numpy.org/doc/stable/index.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9, 62–66. https://doi.org/10.1109/TSMC.1979.4310076.

Piskinaite, E., Veteikis, D., 2023. The results of digitizing historical maps: comparison of Lithuanian land-use structure in the 19th and 21st centuries. Land 12. https://doi.org/10.3390/land12050946. ARTN 946.

Powers, D.M.W., 2011. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. J. Mach. Learn. Technol. 2, 37–63. https://doi.org/10.48550/arXiv.2010.16061.

Prestby, T.J., 2023. Trust in maps: what we know and what we need to know. Cartography and Geographic Information Science 1–18. https://doi.org/10.1080/15230406.2023.2281306.

Previdi, M., Smith, K.L., Polvani, L.M., 2021. Arctic amplification of climate change: a review of underlying mechanisms. Environ. Res. Lett. 16, 093003 https://doi.org/10.1088/1748-9326/ac1c29.

Rasterio, 2024. https://rasterio.readthedocs.io/en/stable/index.html.

Riksantikvarieämbetet, 1996. ODLINGSLANDSKAPET - en lång markanvändnings historia *Informationsavdelningen*. Kulturmiljöavdelningen 58. Stockholm.

Schulp, C.J.E., Verburg, P.H., 2009. Effect of land use history and site factors on spatial variation of soil organic carbon across a physiographic region. Agric. Ecosyst. Environ. 133, 86–97. https://doi.org/10.1016/j.agee.2009.05.005.

Skalos, J., Weber, M., Lipsky, Z., Trpáková, I., Santrucková, M., Uhlírová, L., Kukla, P., 2011. Using old military survey maps and orthophotograph maps to analyse long-term land cover changes - case study (Czech Republic). Appl. Geogr. 31, 426–438. https://doi.org/10.1016/j.apgeog.2010.10.004.

Song, Y.Z., Kalacska, M., Gasparovic, M., Yao, J., Najibi, N., 2023. Advances in geocomputation and geospatial artificial intelligence (GeoAI) for mapping. Int. J. Appl. Earth Obs. Geoinf. 120 https://doi.org/10.1016/j.jag.2023.103300. ARTN 103300.

Swedish Board of Agriculture, 2011. Agriculture in Figures Years 1866-2007. SCB-tryck, p. 209. Örebro.

Swedish Official Statistics, 2023. Jordbruksmarkens användning 2023. Slutlig Statistik. https://jordbruksverket.se/om-jordbruksverket/jordbruksverkets-officiella-statistik/jordbruksverkets-statistikrapporter/statistik/2024-02-07-jordbruksmarkens-anvandning-2023.-slutlig-statistik.

Tian, H.Q., Banger, K., Bo, T., Dadhwal, V.K., 2014. History of land use in India during 1880-2010: large-scale land transformations reconstructed from satellite data and historical archives. Global Planet. Change 121, 78–88. https://doi.org/10.1016/j.gloplacha.2014.07.005.

Watanabe, S., 2023. Tree-structured parzen estimator: understanding its algorithm components and their roles for better empirical performance. https://doi.org/10.48550/arXiv.2304.11127.

Winkler, K., Fuchs, R., Rounsevell, M., Herold, M., 2021. Global land use changes are four times greater than previously estimated. Nat. Commun. 12 https://doi.org/10.1038/s41467-021-22702-2. ARTN 2501.

Winter, E., 2002. Chapter 53 the shapley value. In: *Handbook Of Game Theory With Economic Applications*, 2025-2054. Elsevier. https://doi.org/10.1016/S1574-0005(02)03016-3.