Geoderma 352 (2019) 314–329

Contents lists available at ScienceDirect

Geoderma

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

Prediction of soil drainage classes in Denmark by means of decision tree classification

Anders Bjørn Møller*, Bo V. Iversen, Amélie Beucher, Mogens H. Greve

Department of Agroecology, Aarhus University, Blichers Allé 20, 8830 Tjele, Denmark

ARTICLE INFO

Keywords:
Soil
Drainage
Mapping
Water management
Models
Denmark

ABSTRACT

Soil drainage conditions are highly important to farmers and the environment. To map drainage classes efficiently, several analytical approaches, such as decision tree classification, can be used. Decision tree classification can be improved by combining the predictions of several trees with boosting and bagging techniques. This study tested the relative performance of boosting and bagging for the prediction of drainage classes. Furthermore, as drainage classes form an ordered series rather than unrelated classes, differential costs for misclassification were tested in combination with each technique. Decision tree models were trained from 1135 observations of soil drainage classes and validated using leave-one-out cross validation and a hold-out validation sample with 567 observations. The best model was achieved using bagging combined with differential costs for misclassification (overall accuracy = 52.0%). On the other hand, differential costs for misclassification reduced the overall accuracy of boosted decision trees from 50.8% to 49.2%. The best models obtained with boosting and bagging were used to produce maps of drainage classes on a national extent. The maps predicted the same drainage class in 81% of the study area. Finally, with boosting as well as bagging, the models had a high usage of the predictor variables wetlands, slope to channel network, clay content, land use and geology.

1. Introduction

Soil drainage conditions, defined as the degree and frequency with which the soil matrix is free of water saturation, are highly important to farmers, environmentalists and policy makers. They affect plant growth, soil biota, the leaching of nutrients, the release of greenhouse gases, and the preservation of archaeological artefacts (Gambrell et al., 1975; Levine et al., 1994; Nuutinen et al., 2001; Smith et al., 2003; Van de Noort et al., 2002; Watson et al., 1976). The poorest natural drainage conditions are found in wetlands, where the depth to the groundwater is shallow. In other areas, the natural drainage condition is usually the result of the topography and the soil texture, as a fine texture can hinder the infiltration of water. As insufficient soil drainage can severely reduce crop growth, it is often ameliorated by artificial means such as ditching and subsurface drain pipes in agricultural fields (Brady and Weil, 1996). The artificial drainage systems can in turn affect the hydrologic cycle and all of the processes above, as they alter the natural drainage conditions of the soil (Ernstsen et al., 2015; Schelde et al., 2006). In order to ensure optimal agricultural practices and protect the environment, knowledge about the soil drainage conditions is needed. However, the knowledge is often found lacking, and assessments will usually have to rely on the farmer's experience or time-consuming field work.

As an alternative, soil drainage can be digitally mapped by various analytical approaches, of which a large number have been applied over the last decades. Bell et al. (1992, 1994), Kravchenko et al. (2002) and Liu et al. (2008) mapped soil drainage classes using discriminant analysis. Bell et al. (1992, 1994) mapped three generalized drainage classes for a 144 km^2 area in Pennsylvania, USA based on soil observations and maps of geological and topographic variables. The study found the achieved accuracy of 74% was higher than the accuracy achieved with a conventional soil survey of the same area (69%).

Kravchenko et al. (2002) mapped three generalized drainage classes for a 20 ha area in Illinois, USA based on soil observations and maps of topographic variables and soil electrical conductivity. The study used discriminant analysis, indicator kriging, and cokriging, finding that discriminant analysis and cokriging provided the best results. Liu et al. (2008) mapped three drainage classes for two adjacent fields in Ontario, Canada based on soil observations. The variables used included topographic variables, apparent soil electrical conductivity and remote sensing imagery obtained during bare soil conditions. Niang et al. (2012) mapped five drainage classes for a 167 km^2 area in Quebec, Canada, using soil profiles and data derived from remote sensing images as input data. The study tested the results obtained with discriminant analysis versus the results using decision tree classification,

[*] Corresponding author.
E-mail address: anbm@agro.au.dk (A.B. Møller).

http://dx.doi.org/10.1016/j.geoderma.2017.10.015
Received 6 February 2017; Received in revised form 8 June 2017; Accepted 12 October 2017
Available online 21 October 2017

and found that decision tree classification provided the best results. Decision tree classification was also used by Cialella et al. (1997) and Lemercier et al. (2012). Cialella et al. (1997) mapped five drainage classes for a 24 km^2 boreal forest site in Maine, USA. The training and validation data used in the study were sampled from a map of drainage classes derived from a conventional soil map, and the predictor variables included topographic variables and NDVI from satellite imagery. Lemercier et al. (2012) mapped four drainage classes for a 4645 km^2 area in Brittany, France. The model was trained on points extracted from detailed soil maps of smaller areas, combined with layers of topographic variables, geology, land use, data from gamma-ray spectrometry and a map of soil parent material. Zhao et al. (2008) mapped seven drainage classes for a 14.5 km^2 area in New Brunswick, Canada using an artificial neural network (ANN) with DEM derived topographic variables and soil parameters derived from a coarse resolution conventional soil map as input. The study used drainage classes from a high resolution conventional soil map for training and validation. Zhao et al. (2013) proceeded to map seven drainage classes for Nova Scotia, Canada (55,000 km^2) using the ANN from the previous study, an ANN trained on observations from the new study area, and a linear transformation model with models for 12 different landforms. The study found that the ANN from the previous study predicted the drainage classes poorly, while the linear transformation model had the best prediction. Other methods used include logistic modelling used by Campling et al. (2002) and unsupervised classification used by Peng et al. (2003). Campling et al. (2002) mapped six drainage classes for a 589 km^2 area of South Eastern Nigeria based on soil observations. The study found that the best predictions were achieved by combining terrain and vegetation information. Peng et al. (2003) mapped three drainage classes for a 57 ha area with bare soil in Maine, USA using topographic variables and spectral images from several sources. The study found that the best predictions were achieved with high-resolution imagery.

Decision tree classification is a machine learning technique, which works by recursive partitioning of a dataset in order to arrive at a homogenous classification of a target variable. At each split the algorithm aims to reduce the entropy of the target variable in the resulting datasets by choosing the optimal split from of a number of independent variables. The technique has several advantages, as it is computationally inexpensive, makes no assumptions about the distribution of the predictor variables and is robust towards missing data and redundant predictor variables (Mitchell, 1997; Quinlan, 1996; Rokach and Maimon, 2005; Tan et al., 2014). As single decision trees are usually weak classifiers, it is common to combine the predictions of several decisions trees to obtain a better predictive performance by means of ensemble techniques such as boosting or bagging (Bauer and Kohavi, 1999; Dietterich, 2000). Of the previous studies using decision tree classification for mapping drainage classes, Cialella et al. (1997) and Niang et al. (2012) used single decision trees, while Lemercier et al. (2012) used stochastic gradient boosting, which combines boosting and bagging.

Apart from the prediction of soil drainage classes, decision tree classification has been successfully applied to map other categorical soil variables in a large number of studies (Adhikari et al., 2014; Chaney et al., 2016; Giasson et al., 2011; Kheir et al., 2010a; Kheir et al., 2010b; Lagacherie and Holmes, 1997; Lemercier et al., 2012; Moran and Bui, 2002; Odgers et al., 2014; Scull et al., 2003; Taghizadeh-Mehrjardi et al., 2014). The technique has most frequently been applied to map soil types, but alternative uses have included Cu-content classes and fertility classes (Chen and Ma, 2010; Zhang et al., 2008).

By default, decision tree classification does not take similarities between the predicted classes into account, and treat the cases as either correctly or incorrectly classified. However, in the case of drainage classes, some misclassifications can be more severe than others. For example the misclassification of a poorly drained soil as a moderately well-drained soil is closer to reality than misclassifying it as a well-drained soil, and would have practical implications for the soil management. Some decision tree algorithms can be altered by implementing differential costs for misclassification (Drummond and Holte, 2000; Kuhn et al., 2015; Ting, 1998), and it is conceivable that this method may improve the prediction of drainage classes.

In Denmark, soil drainage conditions are divided into five classes: very well-drained soils (DC1), well-drained soils (DC2), moderately well-drained soils (DC3), poorly drained soils (DC4), and very poorly drained soils (DC5). These classes are mainly defined from morphological characteristics such as the presence and depth of pseudogley, reduced horizons and histic epipedons (Madsen and Jensen, 1988). The classes have been used in the description of soil profiles but have not previously been mapped in any areas of Denmark.

The study aims at mapping the drainage classes of Danish soils on a national extent by means of decision tree classification. It will test the relative effects on predictive performance of boosting vs. bagging and investigate how the performances of each of the two techniques are affected by the implementation of differential costs for misclassification.

2. Materials and methods

2.1. Study area

Denmark is located in northern Europe with a total area of 42,895 km^2. The landscape is generally flat with a mean elevation of 31 m above sea level. The climate is temperate with a mean annual temperature of 8.7 °C in the years 2001–2010, ranging from 1.5 °C in the winter to 16.3 °C in the summer. The mean annual precipitation was approximately 770 mm, ranging from about 650 mm in the eastern part of the country to about 875 mm in the western part of the country (Wang, 2013).

The geology consists mostly of loamy Weichselian moraine in the eastern part of the country and sandy glacial outwash plains and Saalian moraine in the western part. The dominant soil types are Luvisols and Podzols in the eastern and western parts of the country respectively (Jacobsen, 1984). Historically, topographic maps have shown wetlands covering > 20% of the country. The wetlands comprised meadows in river valleys and along the coast, salt marshes in the south-western part of the country, raised sea beds in the northern part of the country, and sinks in kettled moraine landscapes. However, due to drainage activities in the 19th and 20th centuries, the extent of wetlands has been much reduced. Historically, They have also been present in moorlands, where placic horizons could hinder the infiltration of water (Madsen et al., 1992).

The main land use is agriculture, which accounts for 66% of the land area, followed by natural vegetation (16%) and urban areas (10%) (Statistics Denmark, n.d.).

2.2. Training and validation data

1702 soil profiles were used as input data. 860 profiles were located along a major gas pipeline, and 841 profiles were located on a 7 km grid (Fig. 1). The soil profiles along the gas pipeline were described in the years 1981–1985, while the 7 km grid was investigated in the years 1987–1989 (Madsen et al., 1992). The soil profiles were dug to a depth of 1.5–1.8 m, or less if a water table or bedrock was encountered < 1.8 m from the surface. The profiles were photographed and described according to “A key To Soil Profile Description” (Madsen and Jensen, 1988), which is an adaptation of FAO's” Guidelines For Soil Profile Description” (FAO, 1977) for Danish conditions. The descriptions included horizon sequence, depth, colour, texture and other characteristics, and samples were collected from the horizons in order to measure the physical and chemical properties of the soil.

The drainage classes assigned during the original surveys were used for 1697 profiles. Five profiles which did not a have a drainage class

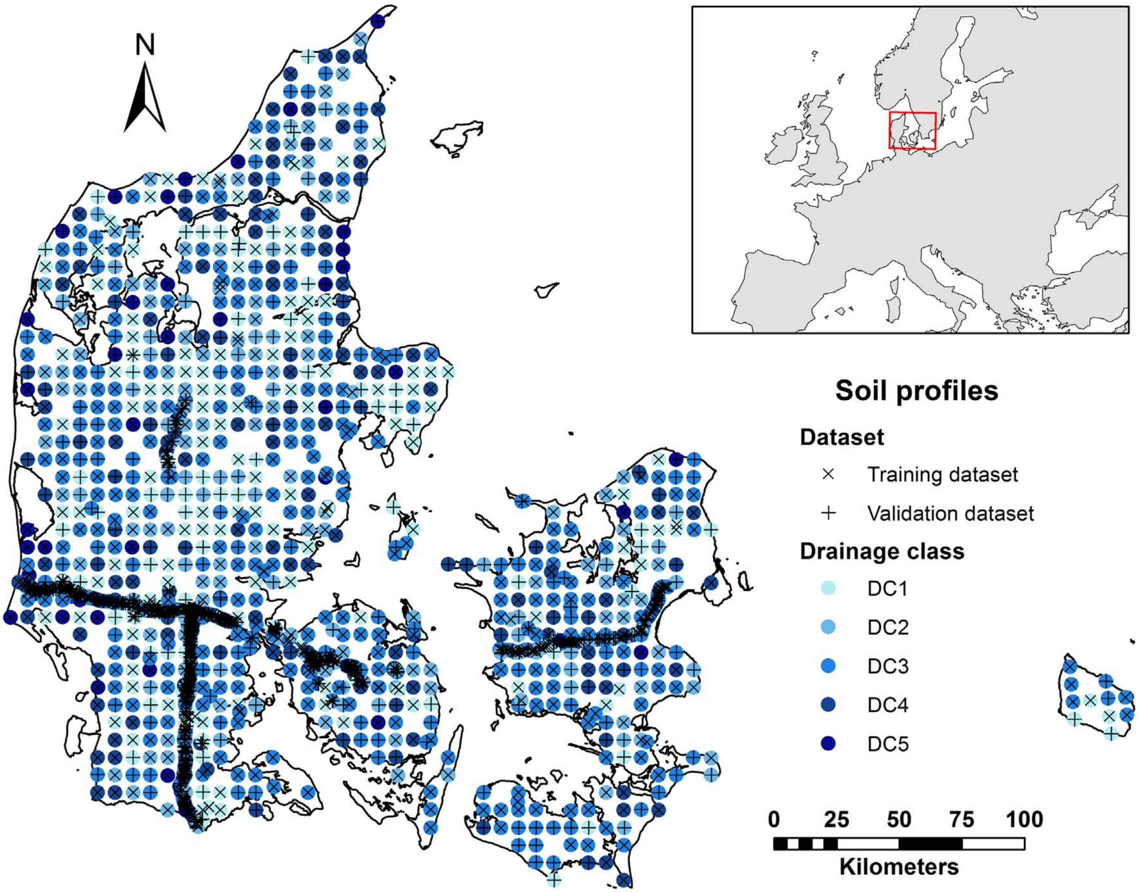

Fig. 1. Training and validation datasets used in the study and the drainage classes of the soil profiles. The location of the study area is shown in the upper left corner.

assigned to them during the surveys had their drainage classes determined from the profile description and pictures. The drainage classes are defines as (Madsen and Jensen, 1988):

- DC1 (Very well-drained soils): Soils with no hydromorphic features in the upper 120 cm.
- DC2 (Well-drained soils): Soils with pseudogley at 80–120 cm or weak signs of temporary water stagnation in the upper 120 cm.
- DC3 (Moderately well-drained soils): Soils with gley at 80–120 cm, pseudogley at 0–80 cm or clear signs of temporary water stagnation at 40–120 cm.
- DC4 (Poorly drained soils): Soils with a reduced horizon at 80–120 cm, gley at 0–80 cm, an organic-rich A horizon, a histic top-horizon, a water table at 50–100 cm during the summer in most years or clear signs of temporary water stagnation at 0–40 cm.
- DC5 (Very poorly drained soils): Soils with a reduced horizon at 0–80 cm or a constant water table at 0–50 cm.

Two thirds of the data were used for training, and one third for validation. The data were split using a stratified random split in order to ensure equal proportions of drainage classes in the training data and the validation data (Table 1). The most common drainage class in the input data is DC3 with 37.5% of the profiles, while DC5 is the rarest drainage class with 4.3% of the profiles.

Table 1
Number of soil profiles per drainage class for the whole dataset and the training and validation datasets.

DC	Name	Whole dataset	Training data	Validation data
1	Very well-drained soils	331	221	110
2	Well-drained soils	286	191	95
3	Moderately well-drained soils	639	426	213
4	Poorly drained soils	373	248	125
5	Very poorly drained soils	73	49	24
Total		1702	1135	567

2.3. Predictor variables

In the study, 31 predictor variables were used (Table 2). A LiDAR based DEM in 1.6 m resolution was aggregated to 30.4 m resolution by calculating the mean value of the 19 × 19 cells covered by each of the cells in the new layer. 15 topographic parameters were derived from the 30.4 m DEM. Horizontal distance to channel network was calculated as the two-dimensional Euclidean distance to vector layers of watercourses, lakes and the sea. Slope to channel network calculates the slope angle to the hydrologically nearest waterbody as arctan (*Vertical distance/Horizontal distance*), with surface flow direction taken into account. Depth to groundwater was calculated from a groundwater table modelled at a 500 m resolution (Henriksen et al., 2012). The groundwater table was resampled to a 30.4 m resolution by means of bilinear interpolation and subtracted from the DEM. Multiresolution valley bottom flatness was calculated as described by Gallant and Dowling (2003).

Layers of soil clay contents were calculated in four depth intervals by aggregating clay contents predicted by Adhikari et al. (2013), using weighed averages of the finer intervals used in the original study. Cropping history was extracted from digital field maps with farmers' declarations for area payments for the period 2011–2014 (The Danish Agrifish Agency, 2014). The crops were classified as either drainage-dependent, possibly drainage-dependent or drainage-independent, and

Table 2
Predictor variables used in the study.

Predictor variable	Description	Mean (range)/number of classes	Reference
Topography			
Blue spot analysis	Depth of sinks	0.1 (0.0–92.5)	
Depth to groundwater	Depth to upper groundwater table	5.8 (0.0–126.0)	(Henriksen et al., 2012)
Detrended elevation model	Elevation minus the mean elevation within a 4 km radius	1.0 (− 57.9–105.4)	
Direct insolation	Potential incoming solar radiation calculated for a single year (kWh)	1269 (122–1707)	
Elevation	LiDAR produced elevation of the land surface	30.9 (− 39.5–170.5)	(National Survey and Cadastre, 2012)
Flow accumulation	Number of upslope cells	60 (1–110,908)	
Horizontal distance to channel network	Calculates the horizontal distance to the nearest waterbody	279 (0–4401)	
Mid-slope position	Covers the warmer zones of slopes	0.27 (0–1)	
Multiresolution valley bottom flatness	Calculates depositional areas	4.3 (0.0–10.9)	(Gallant and Dowling, 2003)
Reclassified slope aspect	Slope aspect reclassified into 8 general directions	8 classes	
SAGA Wetness Index	Topographic wetness index with modified catchment area	14.5 (6.9–19.1)	(Böhner et al., 2002)
Slope aspect	Direction of the steepest slope from the North	181.1 (0–360)	
Slope gradient	Local slope gradient (degrees)	1.6 (0.0–90)	
Slope to channel network	Slope (degrees) to the hydrologically nearest waterbody	1.0 (0.0–78.9)	
Topographic wetness index	Calculated as TWI = ln(a/tan b): where a is flow accumulation, and b is local slope gradient	5.87 (− 15.83–63.30)	
Valley depth	Extent of the valley depth	7.5 (0.0–89.9)	
Vertical distance to channel network	Calculates vertical distance to waterbodies	1.4 (0.0–45.9)	
Landsat 8 derived indices			
NDMI	Normalized difference moisture index	0.08 (− 1–1)	(NASA Landsat Program, 2014)
NDVI	Normalized difference vegetation index	0.52 (− 1–1)	(NASA Landsat Program, 2014)
NDWI	Normalized difference water index	− 0.54 (− 0.99–1)	(NASA Landsat Program, 2014)
SAVI	Soil adjusted vegetation index	0.29 (− 0.29–0.72)	(NASA Landsat Program, 2014)
Choropleth maps			
Geology	Scanned and registered geological map (Scale 1:25,000)	10 classes	(Jakobsen et al., 2015)
Geo-regions	Scanned geographical regions map (Scale 1:100,000)	7 classes	(Adhikari et al., 2013)
Landscape elements	Landform types (Scale 1:100,000)	12 classes	(Madsen et al., 1992)
Wetlands	Shows the presence of wetlands, central wetlands and peat (Scale 1:20,000)	4 classes	(Greve et al., 2014)
Soil texture			
Clay content (0–30 cm)	Clay content (%) for 0–30 cm soil depth	8.1 (0.0–51.2)	(Adhikari et al., 2013)
Clay content (30–60 cm)	Clay content (%) for 30–60 cm soil depth	10.1 (0.0–62.7)	(Adhikari et al., 2013)
Clay content (60–100 cm)	Clay content (%) for 60–100 cm soil depth	11.2 (0.0–59.1)	(Adhikari et al., 2013)
Clay content (100–200 cm)	Clay content (%) for 100–200 cm soil depth	10.9 (0.0–57.1)	(Adhikari et al., 2013)
Land cover			
Cropping history	Years with drainage dependent crops minus years with drainage independent crops	1.6 (− 4–4)	(The Danish Agrifish Agency, 2014)
Land use	CORINE land cover data adopted in Denmark (Scale 1:100,000)	5 classes	(European Environment Agency, 2014)

the number of years with each crop category was counted. The number of years with drainage-independent crops was subtracted from the number of years with drainage-dependent crops, while the number of years with possibly drainage-dependent crops was not used in the final layer. Choropleth maps of geology at 1 m depth (Jakobsen et al., 2015), geo-regions (Adhikari et al., 2013), landscape elements (Madsen et al., 1992), Corine Land Cover 2012 land use (European Environment Agency, 2014) and wetlands (Greve et al., 2014) were used. The geological map was reclassified to 11 classes and the land use map was reclassified to 5 classes.

The spectral indices NDVI (Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index), NDMI (Normalized Difference Moisture Index), and SAVI (Soil Adjusted Vegetation Index) were derived from a mosaic of Landsat 8 images in 30 m resolution from 10 scenes recorded during March 2014 (NASA Landsat Program, 2014), which was the only month during which cloud free images could be obtained for all of Denmark. The indices were calculated as:

$$NDMI = \frac{NIR - IR}{NIR + IR} \quad (1)$$

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (2)$$

$$NDWI = \frac{NIR - G}{NIR + G} \quad (3)$$

$$SAVI = \frac{NIR - RED}{NIR + RED + L}^{*}(1 + L) \quad (4)$$

NIR: Near infrared (Band 5); *IR*: Infrared (Band 6); *RED*: Red (Band 4); *G*: Green (Band 3). *L*: Factor set to 0.5.

The maps of the spectral indices were resampled to the same resolution as the DEM derived variables using bilinear interpolation, and the vector maps were converted to raster layers with the same resolution. The clay contents, geology and land use of the soil profiles used as training and validation data were derived from observations from the soil profiles, while the remaining predictor variables were extracted from the map layers.

2.4. Models

Two predictive models were trained using the C5.0 decision tree algorithm (Quinlan, 1993). The first model combined trees by means of boosting and is henceforth referred to as the *boosted model*, while the second model combined trees by means of bagging and is henceforth referred to as the *bagged model*. In the creation of both models the effects of differential costs for misclassification were tested.

Boosting assigns weights to the instances in the training set. Instances which are misclassified are assigned higher weights, and a new classifier is learned. Thereby, instances which are consistently misclassified are assigned higher weights with each iteration, causing the algorithm to focus on the instances that are difficult to classify. The final prediction is a weighted vote between the classifiers of the ensemble (Freund and Schapire, 1996).

Bagging draws a number of bootstrap samples from the training data and builds a classifier from each bootstrap sample. The predictions of the classifiers are then combined by simple voting (Breiman, 1996). Probabilistic bagging is a variant of the method wherein the class probabilities predicted by the individual classifiers are averaged. The method has been shown to provide better performance than ordinary bagging in most cases (Bauer and Kohavi, 1999).

In the original C5.0 algorithm, the splits of the decision tree are chosen in order to reduce entropy (Quinlan, 1993). Ting (1998) proposed a cost sensitive version of the C5.0 algorithm's predecessor C4.5. In this approach the instances in the training data are assigned weights, which are proportional to the cost of misclassifying the class to which the instance belongs. The algorithm then chooses the splits which minimize the costs associated with the resulting classification. In a classification task with m classes, the costs for misclassification are given in a matrix of size $m*m$, wherein the rows indicate the predicted classes and the columns indicate the reference classes. The off-diagonal elements give the costs for misclassification and the diagonal elements are zero as they represent the cost for correct classification. The method for differential costs for misclassification is included in the C5.0 R package (Kuhn et al., 2015). In order to take into account the fact that drainage classes form an ordered series, costs equal to the absolute differences between the predicted and the reference drainage classes were used (Table 3).

In the optimization of the boosted model, decision trees were grown with and without differential costs for misclassification. In each experiment, the number of trees grown was varied from one (no boosting) to 100 in order to find the optimal number of boosting trials within this range. The final model was chosen as the model with the best performance on the validation data. The performance was measured as overall accuracy (OA) and kappa (K), treating the drainage classes as a categorical variable:

$$OA = \frac{\sum_{i=1}^{m} E_{ii}}{N} \tag{5}$$

m: number of drainage classes; E_{ii}: sum of diagonal elements; N: total number of observations.

$$K = 1 - \frac{1 - OA}{1 - p_e} \tag{6}$$

p_e: hypothetical probability of chance agreement.

In addition, the performance was measured as mean absolute error (MAE) and relative error (RE), treating the drainage classes as a numeric variable:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |o_i - y_i| \tag{7}$$

Table 3
Matrix with differential costs for misclassification of drainage classes.

		Reference DC				
		1	2	3	4	5
Predicted DC	1	0	1	2	3	4
	2	1	0	1	2	3
	3	2	1	0	1	2
	4	3	2	1	0	1
	5	4	3	2	1	0

o_i: observed value of observation i; y_i: predicted value of observation i; n: number of observations.

Expected error (EE):

$$EE = \frac{\sum_{j=1}^{m}\sum_{k=1}^{m} |j - k|\ n_j n_k}{N^2} \tag{8}$$

j: predicted drainage class; k: observed drainage class; n_j: number of predictions of drainage class j; n_k: number of observations of drainage class k.

$$RE = MAE/EE \tag{9}$$

In the optimization of the bagged model, decision trees were also grown with and without differential costs for misclassification. Furthermore, the effects of ordinary bagging and probabilistic bagging were tested. When probabilistic bagging was used, tree pruning was turned off as recommended by Bauer and Kohavi (1999). Probabilistic bagging was not used in conjunction with differential costs for misclassification as the implementation of differential costs prevents the calculation of class probabilities (Kuhn et al., 2015). As a result, a total of three experiments were carried out. In each experiment, bagging was repeated 30 times with an ensemble of 100 trees grown in each repetition. In order to ensure comparability, the same 30 sets of 100 bootstrap samples were used in each experiment. Each model had its performance evaluated from the validation data based on Eqs. (5) to (9). The final model was chosen as the model with the best performance on the validation data.

The two final models obtained using boosting and bagging were furthermore validated by leave-one-out cross validation. The two models had their ability to predict the individual drainage classes measured as map unit purity (MUP), which is the probability that the predicted class matches the observed class, and class representation (CR), which measures how well the observed class was predicted by the model:

$$MUP_j = \frac{X_{jj}}{\sum_{j=1}^{m} X_{jk}} \tag{10}$$

$$CR_k = \frac{X_{kk}}{\sum_{k=1}^{m} X_{jk}} \tag{11}$$

X_{ii}: diagonal value for each class in one row; X_{kk}: diagonal value for each class in one column; X_{jk}: sum of values in one row or column.

The two metrics are identical to the user's accuracy and the producer's accuracy, respectively, but the former terminology is adopted as recommended by Lark (1995). Eqs. (10) and (11) were calculated on the validation data and the results of the cross validation.

The model uncertainties associated with the predictions given by the two final models were calculated from the class probabilities. When ordinary bagging was used, class probabilities were calculated as the proportion of trees in the model predicting a given drainage class. The model uncertainty was calculated as the expected mean absolute deviation from the predicted drainage class for a given point:

$$Model\ uncertainty = \sum_{k=1}^{m} |q - r|\ p_q \tag{12}$$

q: drainage class with the predicted probability p_q; r: drainage class with the highest probability (the predicted drainage class for the point).

The model uncertainty was calculated for all instances in the cross validation. The boosted and the bagged model were then used to produce maps of drainage classes and model uncertainties for the nonurban land area of Denmark.

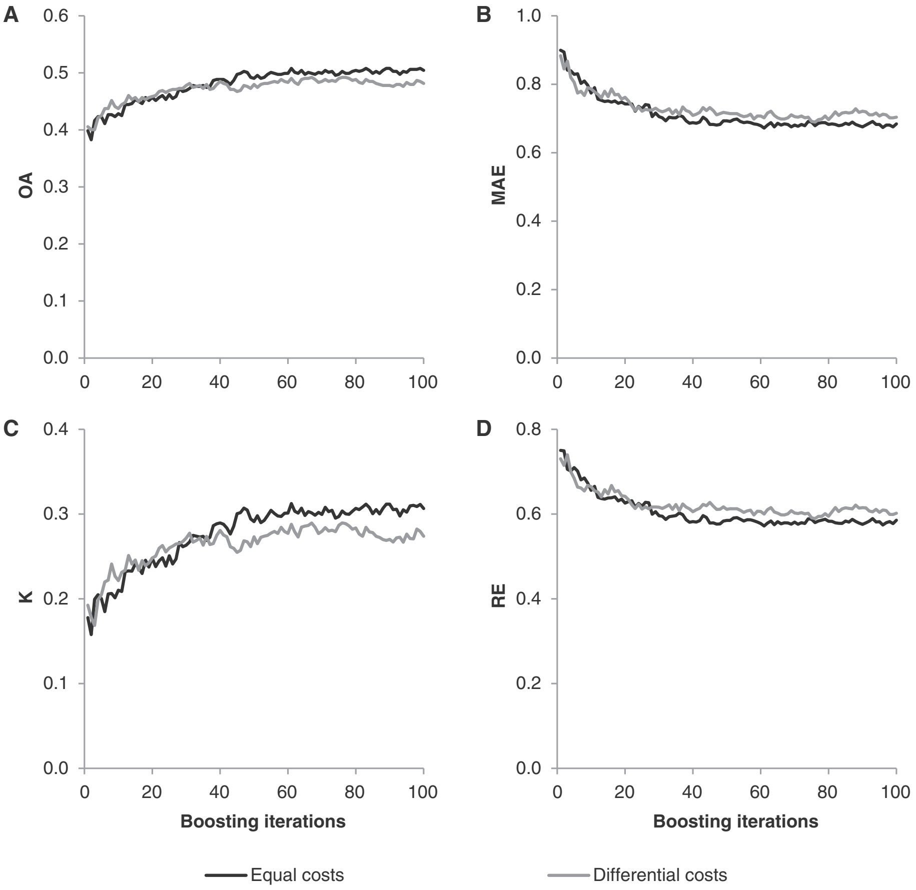

Fig. 2. Performance of each method relative to the number of boosting iterations as evaluated by OA (A), MAE (B), K (C) and RE (D).

3. Results

3.1. Model optimization

Boosting improved the performance for all performance metrics, as evaluated by Eqs. (5) to (9) (Fig. 2). The four performance metrics showed similar patterns of improvement. The largest gains in performance were observed in the first iterations after which they decreased until performance reached a plateau. The model using differential costs for misclassification performed better than the model with equal costs when the number of boosting iterations was low. However, the model using differential costs for misclassification lost its advantage as the number of boosting iterations increased. From a number of boosting iterations of 37 and onwards, the model with equal costs for misclassification was the best model as evaluated by OA, K, MAE and RE. It reached its best performance at 61 boosting iterations as indicated by all four performance metrics. Therefore, the model with equal costs for misclassification and 61 boosting iterations was selected as the final boosted model.

Bagging produced variable performances for each method with large overlaps between most of the methods (Fig. 3; Table 4). The mean OA was similar for the three methods. The highest mean K was achieved with differential costs for misclassification with the ensembles using equal costs for misclassification and ordinary bagging achieving the second-highest mean K. The lowest mean MAE was achieved using probabilistic bagging with the ensembles using equal costs for misclassification in combination with ordinary bagging producing the second-lowest mean MAE. Lowest RE was achieved with differential costs for misclassification with the ensembles using equal costs for misclassification in combination with ordinary bagging producing the second-lowest mean RE. The method using differential costs for misclassification produced the most variable performances. This method produced the best as well as the worst performances for each of the four performance metrics.

The single ensemble with the best performance was found using differential costs for misclassification. This ensemble had the best of all performances measured both as OA, K, MAE and RE. This ensemble was therefore selected as the final bagged model.

3.2. Model performance

Once the two models had been optimized, the final bagged model had a better performance than the final boosted model. This was observed in the validation sample as well as the cross validation as evaluated by all four performance metrics (Table 5). Both models had a poorer performance in the cross validation than in the validation sample.

For each drainage class, the cases which were misclassified were mostly classified as ± 1 drainage class of the reference drainage class, with the only exception being DC1, which was mostly misclassified as DC3 (Table 6). For the boosted model 85.5% of the cases in the validation sample and 81.4% of the cases in the cross validation were predicted within ± 1 drainage class of the reference DC. For the bagged model, 86.4% of the cases in the validation sample and 81.3% of the cases in the cross validation were predicted within ± 1 drainage class of the reference drainage class.

DC1, DC3, and DC4 were generally predicted well, while DC2 and DC5 were poorly predicted (Table 7). DC3 had the highest CR and MUP in both the validation sample and cross validation of both models. DC5 had the lowest CR and MUP in the validation sample of both models, and DC2 has the lowest CR and MUP in the cross validation of both models.

In the validation sample, the bagged model had a higher CR and MUP than the boosted model for all drainage classes with the exception

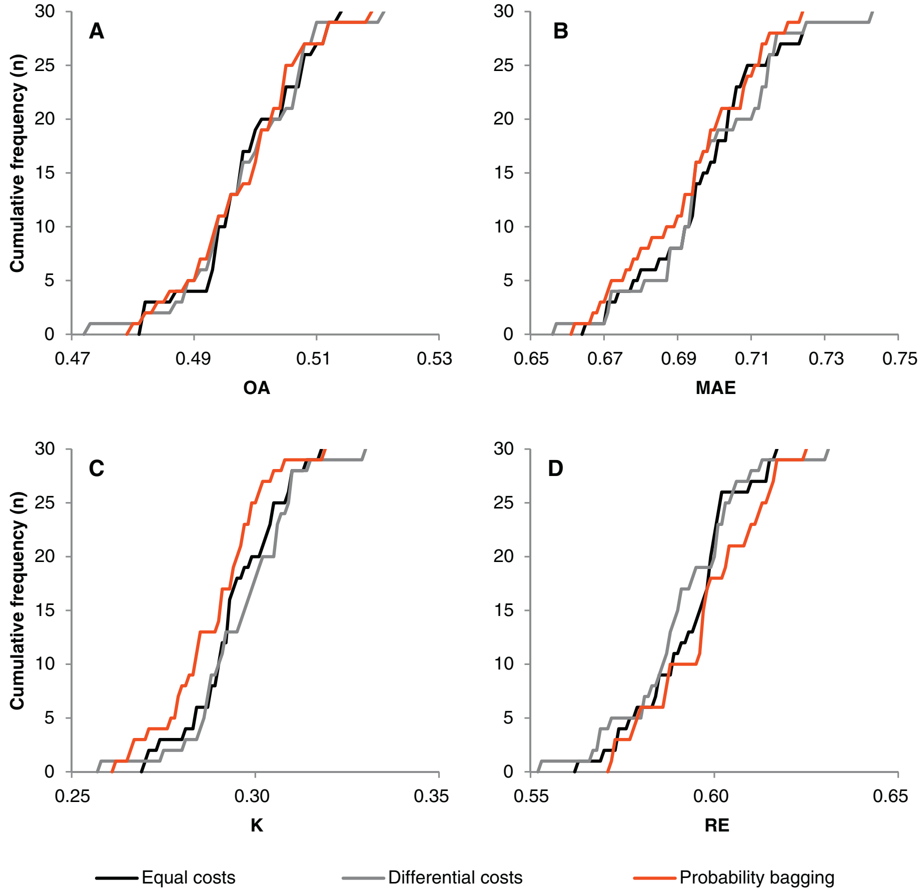

Fig. 3. Cumulative frequency of bagged ensembles obtained with each method relative to their performance as evaluated by OA (A), MAE (B), K (C) and RE (D).

Table 4
Mean value and range of the performance achieved with each bagging method.

Method	OA; mean (range)	K; mean (range)	MAE; mean (range)	RE; mean (range)
Equal costs	0.498 (0.481–0.513)	0.294 (0.269–0.317)	0.697 (0.665–0.725)	0.593 (0.562–0.614)
Differential costs	0.498 (0.473–0.520)	0.296 (0.258–0.330)	0.698 (0.656–0.725)	0.591 (0.553–0.613)
Probabilistic bagging	0.498 (0.480–0.519)	0.289 (0.262–0.319)	0.694 (0.661–0.720)	0.597 (0.572–0.616)

Table 5
Performance of the boosted and the bagged model based on the validation sample and the cross validation.

Metric	Validation sample		Cross validation	
	Boosting	Bagging	Boosting	Bagging
OA	0.508	0.520	0.471	0.489
K	0.313	0.330	0.259	0.280
MAE	0.672	0.656	0.757	0.744
RE	0.572	0.553	0.642	0.631

of DC2. In the cross validation, the boosted model had the highest CR for DC2 and DC5 and the highest MUP for DC1 and DC5. In all other cases the bagged model had the highest CR and MUP in the cross validation.

The MUP and CR were lower in the cross validation than in the validation sample for all drainage classes with only two exceptions: For DC5, the cross validation generally indicated a better prediction than the validation sample, and DC3, as predicted by the bagged model, had a higher CR in the cross validation than in the validation sample. MUP was generally lower than CR for DC1 and DC3. For the other drainage classes MUP was higher than CR.

Both in the boosted and the bagged model, the frequency of DC3 was overpredicted, especially in the cross validation (Fig. 4). The predicted frequencies of DC1 and DC4 were more or less equal to the frequencies in the reference data (whole dataset, see Table 1), while the frequencies of DC2 and DC5 were strongly underpredicted.

The cases which were correctly classified in the cross validation generally associated with low model uncertainties in both the boosted and the bagged model (Fig. 5). The model uncertainties generally increased with the absolute prediction error up to an absolute prediction error of three. Cases with a prediction error of four had lower model uncertainties associated with them than cases with a prediction error of three. Only three cases in the cross validation of each model had a prediction error of four.

3.3. Predictor variables

The usage of the predictor variables in the two models was similar (Spearman's rank correlation = 0.91, n = 31, $p < 0.05$). The five most used predictor variables were the same in the boosted and the bagged model: wetlands, slope to channel network, clay content (100–200 cm), land use, and geology. These five predictor variables were on average used for > 50% of the training cases in the individual trees of the two models (Table 8). Furthermore, geo-regions and the vertical and horizontal distances to the channel network were all amongst the ten most used predictor variables in both models.

Cropping history was the least used predictor variable in both models, and was the only predictor variable with a total usage of < 95%. Also its mean usage of 1.7% was much lower than the usage of the second least used predictor variable (mid-slope position).

Table 6
Confusion matrices for the boosted model and the bagged model, tested on the validation sample and with cross validation.

	Boosting: Validation sample					
	Reference					
Prediction	DC1	DC2	DC3	DC4	DC5	Total
DC1	60	23	18	12	1	114
DC2	15	20	14	6	0	55
DC3	29	46	151	46	4	276
DC4	6	6	30	57	19	118
DC5	0	0	0	4	0	4
Total	110	95	213	125	24	567

	Bagging: Validation sample					
	Reference					
Prediction	DC1	DC2	DC3	DC4	DC5	Total
DC1	65	24	19	13	3	124
DC2	13	16	13	4	0	46
DC3	28	49	154	45	0	276
DC4	4	6	27	59	20	116
DC5	0	0	0	4	1	5
Total	110	95	213	125	24	567

	Boosting: Cross validation					
	Reference					
Prediction	DC1	DC2	DC3	DC4	DC5	Total
DC1	111	45	41	22	3	222
DC2	21	15	29	21	0	86
DC3	69	103	303	98	7	580
DC4	20	28	53	95	28	224
DC5	0	0	0	12	11	23
Total	221	191	426	248	49	1135

	Bagging: Cross validation					
	Reference					
Prediction	DC1	DC2	DC3	DC4	DC5	Total
DC1	115	51	45	27	3	241
DC2	16	12	16	13	1	58
DC3	72	102	319	94	7	594
DC4	18	26	46	104	33	227
DC5	0	0	0	10	5	15
Total	221	191	426	248	49	1135

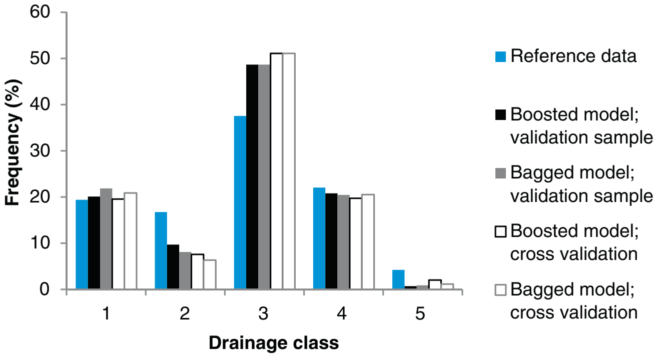

Fig. 4. Frequency of each drainage class in the reference data (whole dataset) and in the predictions of the boosted model and the bagged model on the validation sample and in the cross validation.

Furthermore, the spectral indices (NDMI, NDVI, NDWI and SAVI), detrended elevation model, slope aspect (raw and reclassified), topographic wetness index and valley depth were amongst the least used predictor variables in both models.

Landscape elements, elevation, and slope gradient had a higher relative usage in the boosted model than in the bagged model with a rank difference of six or more, while clay contents at 0–30 cm and 60–100 cm and direct insolation had a higher relative usage in the bagged model than in the boosted model.

The clay content at a depth of 100–200 cm generally increased from DC1 to DC3 and decreased from DC3 to DC5. The slope to the channel network generally decreased with the drainage class (Fig. 6).

Glacial and aeolian sands were generally the geologies most closely associated with well-drained soils, while freshwater peat and marine clay were mostly associated with poorly drained soils. During the original investigations, two profiles with buried peat horizons at a depth of approximately 1 m were classified as DC1 and DC2, respectively. The two profiles had drainage classes assigned to them based on the colluvial deposits covering the peat. However, their geological class was freshwater peat, as it was defined at a depth of 1 m.

The land use class natural vegetation was mostly associated with either well-drained (DC2) or very well-drained soils (DC1), while wetland vegetation was almost exclusively associated with poorly drained (DC4) or very poorly drained soils (DC5). Most soils in agricultural areas were moderately well-drained (DC3). The soil drainage classes generally increased from non-wetland areas through the classes wetlands and central wetlands to peat areas, with increasing frequencies of DC4 and DC5 and decreasing frequencies of DC1 and DC2 (Table 9).

3.4. Mapping

The boosted and the bagged model produced similar maps (Fig. 7). The agreement between the drainage classes mapped by the two models

Table 7
CR and MUP for each DC calculated from the validation sample and by cross validation for the boosted model and the bagged model.

	Boosting				Bagging			
	Validation sample		Cross validation		Validation sample		Cross validation	
DC	CR	MUP	CR	MUP	CR	MUP	CR	MUP
1	0.55	0.53	0.50	0.50	0.59	0.52	0.52	0.48
2	0.21	0.36	0.08	0.17	0.17	0.35	0.06	0.21
3	0.71	0.55	0.71	0.52	0.72	0.56	0.75	0.54
4	0.46	0.48	0.38	0.42	0.47	0.51	0.42	0.46
5	0.00	0.00	0.22	0.48	0.04	0.20	0.10	0.33

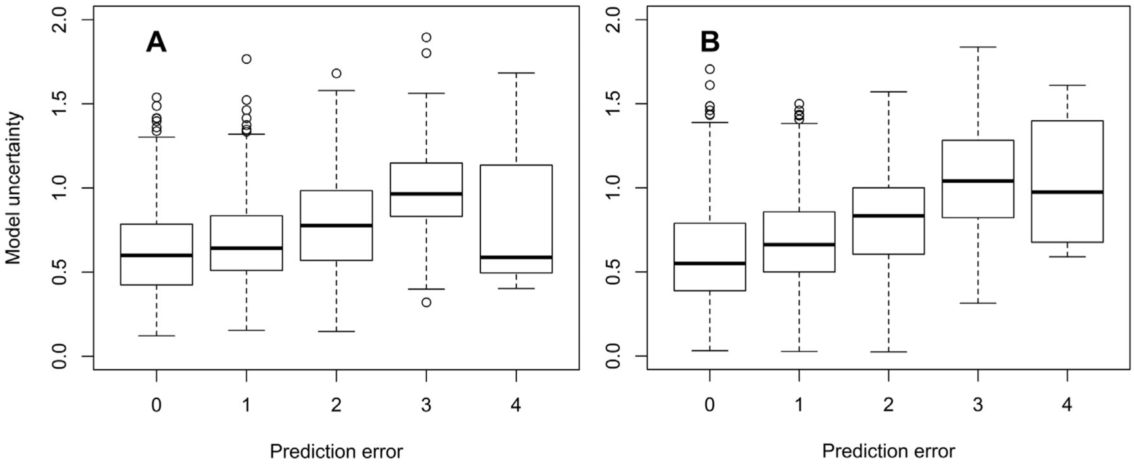

Fig. 5. Box plot showing model uncertainty versus the absolute prediction error (both using the numeric value of the drainage classes) of the leave-one-out predictions of the boosted model (A) and the bagged model (B). Whiskers mark 1.5 IQR (interquartile range) distances from the boxes.

was 81%, and the mean absolute difference between the mapped drainage classes was 0.27. In both maps, DC3 was the dominant drainage class on the loamy tills of eastern Denmark, and DC4 was the dominant drainage class in wetlands areas. In western Denmark, the maps showed a mixture of drainage classes with DC1 and DC2 in high-lying sandy areas and DC3 and DC4 in clayey areas and low-lying areas. DC5 was rarely found in the maps and exclusively in wetland areas.

The proportions of the drainage classes in the maps produced by the two models were similar. The bagged model predicted higher proportions of DC1, DC3, and DC4 and lower proportions of DC2 and DC5. DC3 was the most common drainage class with > 45% of the mapped area in both maps, followed by DC4, which covered approximately 25% of the mapped area in both maps. DC5 was only present in approximately 2% of the mapped area in both maps (Fig. 8).

Maps of the five most used predictor variables are shown for an example area in northern Denmark in Fig. 9, and the predicted drainage classes for the same area are shown in Fig. 10. For both models, DC4 was the dominant drainage class in wetland areas. On the glacial till, which is mostly sandy in this area, DC1 and DC2 were the dominant drainage classes, with DC1 predicted for the areas with the lowest clay contents and the largest slope to the channel network. A large portion of these areas are covered with natural vegetation. On the marine deposits, both models predicted a mixture of DC2 and DC4. Although it was the most frequently predicted drainage class on a national scale, DC3 was rarely predicted for the example area with only a small presence in areas with clayey till and intermixed with DC2 and DC4 on the marine deposits. The edges between the drainage classes were in many cases abrupt. For example, DC2 was found directly next to DC4 in the marine deposits of the example area, and in other areas DC1 was observed next to DC3 or DC4. The predictions of the two models were similar, but the boosted model generally predicted a larger presence of DC4 in the marine deposits not classified as wetlands, while the bagged model predicted a larger presence of DC1 on the sandy till.

The maps of the model uncertainties for the boosted and bagged models generally showed the same patterns. The mapped model uncertainties of the boosted and bagged models are significantly positively correlated ($R^2 = 0.88$, $n = 4.2*10^7$, $p < 0.05$). The mean model uncertainties were 0.71 ± 0.29 for the boosted model and 0.49 ± 0.15 for the bagged model (± 1 standard deviation). In general, model uncertainties were larger in western Denmark than in eastern Denmark. High model uncertainties were mostly found in areas with ambiguity in the predictor variables such as low-lying sandy areas not classified as wetlands, along the edges of river valleys, and in areas with intermediate clay contents (Fig. 10). On the other hand, low model uncertainties were found in wetland areas, in sandy areas high above the nearest waterbody, and in areas with high clay contents.

4. Discussion

4.1. Model optimization and performance

When boosting was applied, the best model was achieved with equal weights for misclassification while the best model produced by means of bagging was achieved with differential costs for misclassification.

It could be argued that the best bagged ensemble was achieved with differential costs for misclassification only because the ensembles using differential costs for misclassification had a larger variation in performance. However, the method produced the highest mean K and lowest mean RE. Overall, the ensembles using differential costs for misclassification were only outperformed by the other methods, when performance was evaluated by MAE. It is noteworthy that the inclusion of differential costs for misclassification did not reduce OA and K. Ting (1998) found that differential costs would reduce the model's ability to predict the correct class since the model would instead seek to minimize the costs. However, the costs implemented in this study are not arbitrary but instead represent real differences between the drainage classes. For this reason, the costs may guide the model towards the correct predictions.

The reasons for the detrimental effect of differential costs for misclassification on the performance of boosted decision trees are not clear. However, both boosting and differential costs for misclassification work by means of case weights (Freund and Schapire, 1996; Ting, 1998), and it is possible that the enforced dual purposes of the case weights decreases the effectiveness of the boosting process.

The results of the cross validation were not used for model optimization and selection, and therefore they probably represent a less biased estimate of model performance. The cross validation confirms the better performance of the bagged model. Bauer and Kohavi (1999) found that boosting generally performed better than bagging over a long range of applications, but not uniformly so. This finding was confirmed by Dietterich (2000), who also found that boosting performed best, when the level of noise in the dataset was low. On the other hand, if a large degree of classification noise was present, bagging generally outperformed boosting. It is possible that the better performance of the bagged model in this study is due to classification noise in the dataset, arising from subjective judgements and temporal variation. The profiles in the dataset were described over a period of eight years by a large number of investigators, and the drainage classification requires a degree of subjective judgement. For example, the distinction between DC2 and DC3 requires discrimination between “weak” and “clear” signs of water stagnation in the soil. It is also possible that the profiles were more likely to be classified as DC4 and DC5 during wet seasons, as the groundwater level would be higher.

Table 8
Predictor variables ordered by their mean usage (% of training cases) in the boosted model and the bagged model. The means are based on the individual trees in the models.

Boosted model

Variable	Usage
Land use	93.6
Geology	76.0
Slope to channel network	72.7
Wetlands	71.2
Clay content (100–200 cm)	69.5
Geo-regions	46.0
Landscape elements	38.2
Blue spot analysis	34.9
Horizontal distance to channel network	34.6
Vertical distance to channel network	28.4
Depth to groundwater	27.5
SAGA Wetness Index	26.8
Slope gradient	26.5
Clay content (60–100 cm)	25.3
Clay content (30–60 cm)	23.0
Elevation	21.0
Clay content (0–30 cm)	19.7
Multi-resolution valley bottom flatness	18.9
Direct insolation	18.4
Valley depth	17.8
Reclassified slope aspect	16.7
Flow accumulation	16.2
NDMI	14.2
Topographic wetness index	13.9
Slope aspect	10.1
Detrended elevation model	10.1
NDWI	10.0
NDVI	9.5
SAVI	9.4
Mid-slope position	9.1
Cropping history	1.7

Bagged model

Variable	Usage
Wetlands	68.9
Slope to channel network	65.8
Clay content (100–200 cm)	65.7
Land use	63.2
Geology	63.1
Horizontal distance to channel network	46.6
Geo-regions	45.3
Clay content (60–100 cm)	45.0
Vertical distance to channel network	39.8
Clay content (0–30 cm)	35.4
Blue spot analysis	30.7
SAGA Wetness Index	30.3
Direct insolation	29.8
Depth to groundwater	29.0
Clay content (30–60 cm)	28.9
Landscape elements	28.5
Flow accumulation	28.3
Multi-resolution valley bottom flatness	28.1
Slope gradient	26.9
NDMI	26.0
Valley depth	24.0
Elevation	22.3
Topographic wetness index	21.5
SAVI	19.0
Slope aspect	16.6
Reclassified slope aspect	16.6
Detrended elevation model	15.2
NDWI	14.9
NDVI	11.8
Mid-slope position	11.7
Cropping history	1.7

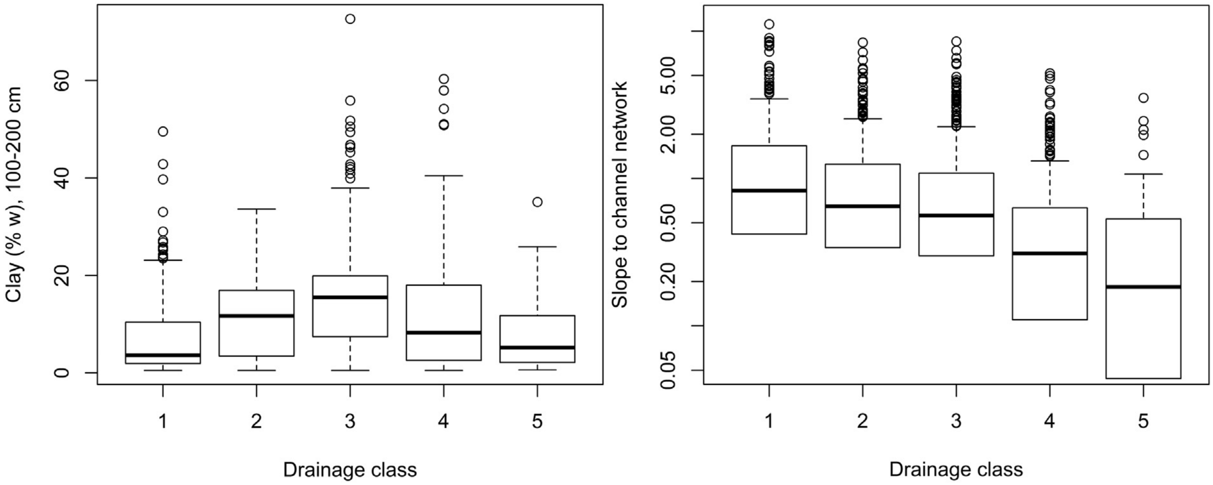

Fig. 6. Box plots of the clay content at a depth of 100–200 cm and the slope to the channel network relative to the drainage classes of the profiles used in the study. Whiskers mark 1.5 IQR (interquartile range) distances from the boxes.

Table 9
Drainage class distribution in percentages of the geological classes, land use classes and wetland classes of the soil profiles used in the study.

Drainage class (%)	1	2	3	4	5	n
Geological class						
Aeolian sand	38.0	14.1	16.9	23.9	7.0	71
Freshwater clay	3.6	17.9	25.0	32.1	21.4	28
Freshwater sand	19.2	4.1	16.4	47.9	12.3	73
Freshwater peat	2.6	2.6	0.0	60.5	34.2	38
Marine clay	0.0	0.0	30.0	50.0	20.0	20
Marine sand	5.9	14.7	11.8	50.0	17.6	34
Glacial clay	10.7	17.4	54.7	16.0	1.2	758
Glacial sand	34.3	19.9	28.7	15.0	2.1	341
Meltwater clay	12.5	29.2	29.2	25.0	4.2	24
Meltwater sand	27.0	17.5	24.8	26.7	4.1	315
Land use class						
Agriculture	15.5	17.0	41.1	22.6	3.7	1364
Natural vegetation	39.5	17.6	25.9	14.6	2.3	301
Wetland vegetation	0.0	2.7	0.0	56.8	40.5	37
Wetland class						
Non-wetlands	21.3	18.4	41.4	17.3	1.6	1425
Wetlands	13.2	11.8	19.1	41.4	14.5	152
Central wetlands	9.6	7.2	22.9	43.4	16.9	83
Peat	0.0	0.0	2.4	64.3	33.3	42
All	19.4	16.8	37.5	21.9	4.3	1702

Overall, the performance achieved in this study is poor when compared to other studies mapping drainage classes (Table 10). The overall performance, as evaluated by all four performance metrics, only compares favourably with Zhao et al. (2013), while the achieved K was also higher than the values achieved by Zhao et al. (2008) and Lemercier et al. (2012). A reason for the poor performance may be the size of the study area, as it is the largest to date with the exception of Zhao et al. (2013). The number of drainage classes may also have had an influence, as some of the best performances were achieved by studies mapping only three drainage classes. However, Niang et al. (2012), Campling et al. (2002) and Niang et al. (2012) achieved better performances with a higher or similar number for drainage classes. It is furthermore possible that classification noise may have affected the performance, as explain above. The overall performance suggests that the produced maps are not reliable, if knowledge of the exact drainage class is required. However, they do display the general patterns well, and, for practical purposes, the associated maps of prediction uncertainties give an indication of the reliability of the maps in specific areas.

4.2. Mapping

The maps produced are generally consistent with a priori knowledge about soil drainage. For example, poorly drained soils (DC4 and DC5) are found in wetland areas while well-drained soils are found in sandy upland areas, and moderately well-drained soils (DC3) are found in relatively clayey areas. However, a clear bias is present in the predicted drainage classes as the most frequent drainage class DC3 is overpredicted while the least frequent drainage classes DC2 and DC5 are underpredicted.

The underprediction of rare classes is often seen when decision tree classification is used (Chawla, 2003; Cieslak and Chawla, 2008). Decision trees have a bias towards more general splitting rules, meaning that rare classes, which are often very specific, are predicted poorly (Holte et al., 1989). Furthermore, the presence of noise in the dataset will make cases with different classes appear similar, leading them to be classified as the majority class (Weiss, 1995). Common solutions include balancing the dataset by “downsampling” (dropping cases from the majority classes), “upsampling” (repeating cases from the rare classes), and the generation of new synthetic cases (Alhammady and Ramamohanarao, 2004; Chawla, 2003; Haibo and Garcia, 2009). Other studies have proposed changes to the induction of the decision trees by changes to the pruning method, splitting criteria, and the calculation of class probabilities (Chawla, 2003, Cieslak and Chawla, 2008). Lastly, the implementation of differential misclassification costs can be used for imbalanced problems, as they can make it more costly to misclassify rare classes (Haibo and Garcia, 2009; Ling and Sheng, 2010). The last solution would most likely be complicated, if the differential costs for misclassification were also employed to take into account similarities between the classes, as in the present study.

The underprediction of DC5 is most likely related only to its rarity in the training dataset, as it is the least frequent drainage class. On the other hand, the underprediction of DC2 is probably related to the overprediction of the majority class, DC3, as about half of the cases classified as DC2 are predicted as DC3 in the validation dataset and the cross validation (Table 6). For the most important predictor variables, DC2 occupied an intermediate position between the more frequent classes DC1 and DC3 (Table 9, Fig. 6), leading it to be easily misclassified.

The abrupt borders between some of the mapped drainage classes are unexpected, as the logical pattern would be a gradual progression of drainage classes. For example, an occurrence of moderately drained soils would be expected between well-drained soils and poorly drained soils. This may partly be explained by the use of categorical variables, which cannot reproduce this pattern, as three of the five most used

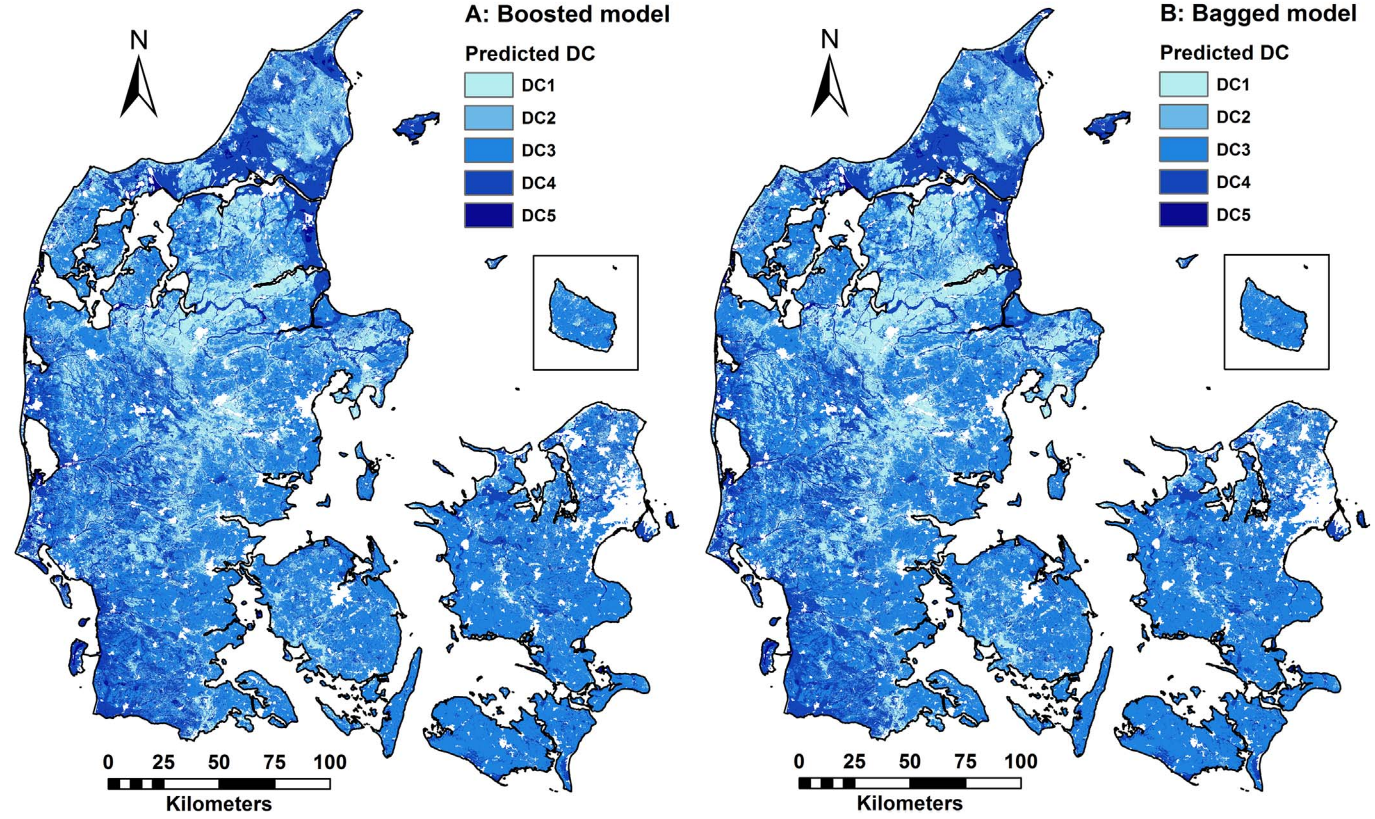

Fig. 7. Maps of the drainage classes predicted by the boosted model and the bagged model.

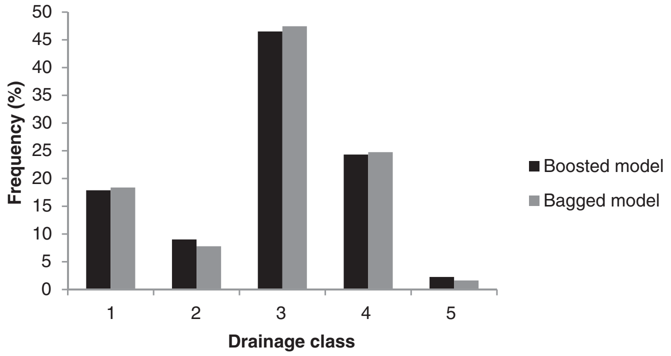

Fig. 8. Frequency of each drainage class in the maps produced from the boosted model and the bagged model.

predictor variables (geology, land use and wetlands) are categorical. It may also be related to the bias towards majority classes mentioned above. For example, DC2 might not appear between DC1 and DC3, because the model does not predict the class as often as it appears in the training data. In the example area, the abrupt change from DC2 to DC4 found in areas with marine sand may be due to the fact that DC3 is rare for this geological class (Table 9).

It is noteworthy that DC5 is better predicted in the cross validation than in the validation sample in opposition to the other drainage classes. As the pattern is repeated in the boosted model as well as the bagged model, it may be due to differences between the training sample and the validation sample. For example, in the training dataset, DC4 and DC5 had slope to channel values of 0.50 ± 0.62 and 0.30 ± 0.45 (± 1 standard deviation), respectively, while in the validation dataset, they had slope to channel values of 0.62 ± 0.93 and 0.60 ± 0.86 (± 1 standard deviation), respectively. A Wilcoxon rank sum test showed that the difference was statistically significant in the training dataset ($p < 0.05$, $n = 297$), but not in the validation dataset ($p > 0.05$, $n = 149$). In effect, the distinction between DC4 and DC5 was more easily determined in the training dataset than in the validation dataset, leading to a better prediction of DC5 in the cross validation. The difference between the two datasets is accidental, as the split was random. However, for rare classes such as DC5, differences in the predictor variables can occur.

4.3. Predictor variables

The high importance of the wetland layer as a predictor variable was not surprising as wetlands are defined by poor drainage conditions. As seen from the results, the layer effectively separated DC4 and DC5 from the other drainage classes. It is noteworthy that no other mapping studies of drainage classes have used similar layers (Bell et al., 1992, 1994; Campling et al., 2002; Cialella et al., 1997; Kravchenko et al., 2002; Lemercier et al., 2012; Liu et al., 2008; Niang et al., 2012; Peng et al., 2003; Zhao et al., 2013; Zhao et al., 2008). The reason may be that similar layers did not exist for the study areas in question as intensive field work would be needed to produce them.

The most important topographic variables were related to the relative position of water bodies (slope to channel network, vertical distance to channel network, and horizontal distance to channel network). A clear relationship was observed between slope to channel network and the drainage classes. Similar predictor variables were found to be important when mapping drainage classes in other studies (Bell et al., 1992, 1994; Kravchenko et al., 2002; Lemercier et al., 2012; Zhao et al., 2013; Zhao et al., 2008).

Variables containing information about the parent material were generally important. Geology was one of the top predictor variables, geo-regions were important in both models, and landscape elements were somewhat important. The soil drainage classes were clearly related to geology as poorly drained soils were mostly found on marine and freshwater deposits while well-drained soils were more often found

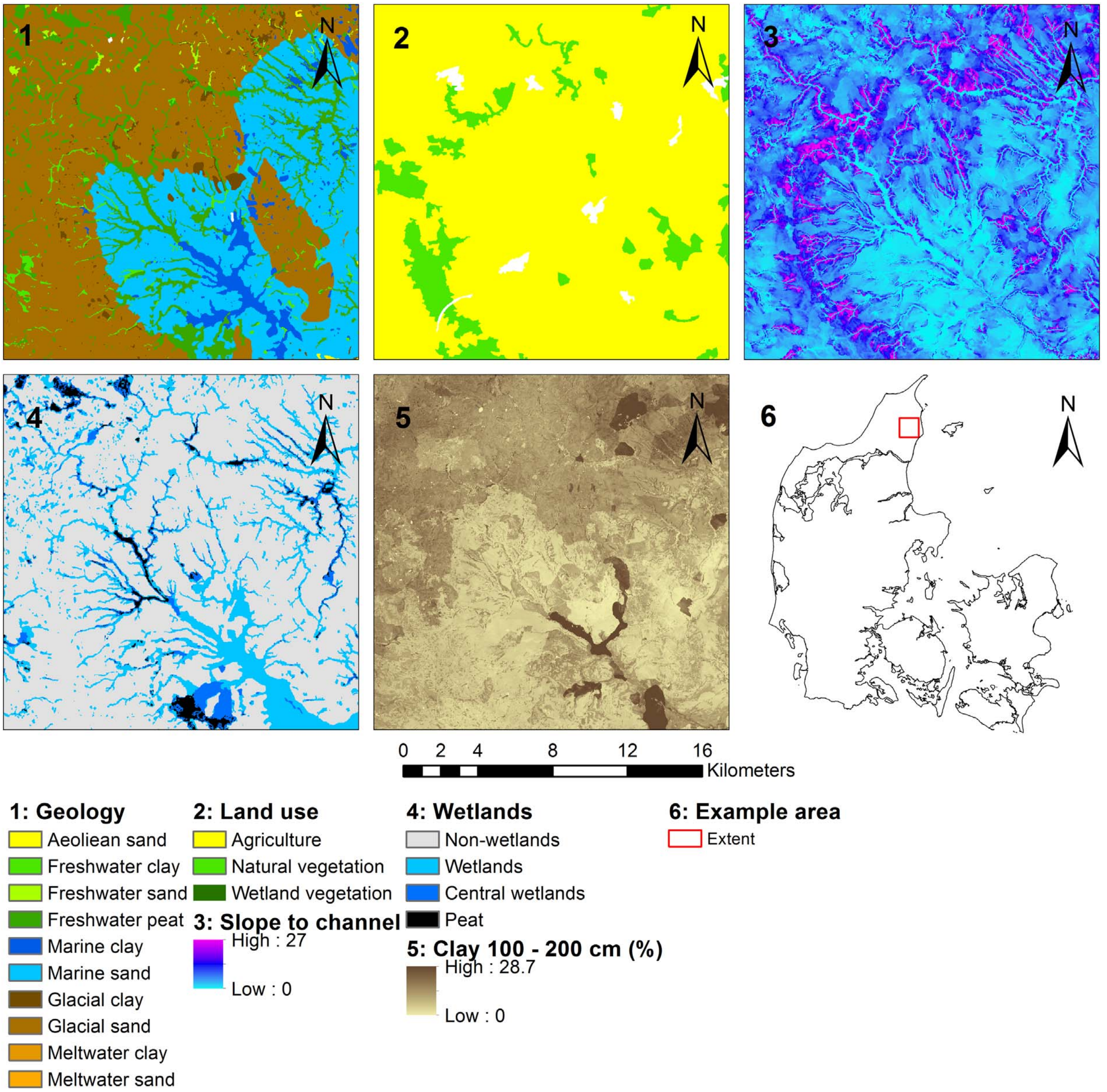

Fig. 9. Maps of the five most used predictor variables in a 19 × 19 km example area in northern Denmark.

on sandy deposits. Despite their usefulness, variables relating to the parent material have rarely been used in other mapping studies of drainage classes. Bell et al. (1992) used a geological map, Zhao et al. (2013) used a map of landforms, which was conceptually similar to the landscape element layer in the present study, and Lemercier et al. (2012) used maps of geology and soil parent material in the prediction. In all three studies, maps related to the parent material were found to be important.

Land use was an important predictor variable while cropping history was the least used variable. This indicates that the land use is related to the soil drainage conditions while crops grown on the fields are largely unrelated to the soil drainage conditions. The latter conclusion is surprising as some crops are highly sensitive to poor drainage conditions (Collaku and Harrison, 2002; Ren et al., 2014; Watson et al., 1976). The layer of cropping history only contained information about agricultural areas. This may have lead the algorithm to choose other variables for the splits. The C5.0 algorithm can handle missing values, but has a preference for informative variables. Furthermore, the drainage classes are defined from soil morphological characteristics, which may remain unchanged when the poor drainage conditions are ameliorated, for example through the installation of a subsurface drainage system. Therefore, drainage dependent crops could theoretically be grown on a soil with a drainage class which indicates that the soil is poorly drained. Despite its usefulness, land use has only been used to map drainage classes in two studies (Lemercier et al., 2012; Niang et al., 2012).

Clay content was found to be important to the soil drainage classes especially at lower depth intervals (60–100 cm; 100–200 cm). The increase in the clay content from DC1 to DC3 is probably related to the low hydraulic conductivity of the clayey soils causing stagnation of water in the soil matrix. On the other hand, the lower clay contents of DC4 and DC5 relative to DC3 may be due to the prevalence of DC4 and DC5 in wetland areas, where the soil drainage conditions are mostly controlled by the water table, while DC3 is generally found outside wetlands, in areas where the drainage conditions are mostly controlled by the soil texture. Predictor variables relating to the soil texture were only used to map drainage classes in the studies of Zhao et al. (2008) and Zhao et al. (2013). In both cases the soil texture was derived from conventional soil maps.

The low importance of the topographic wetness index is surprising,

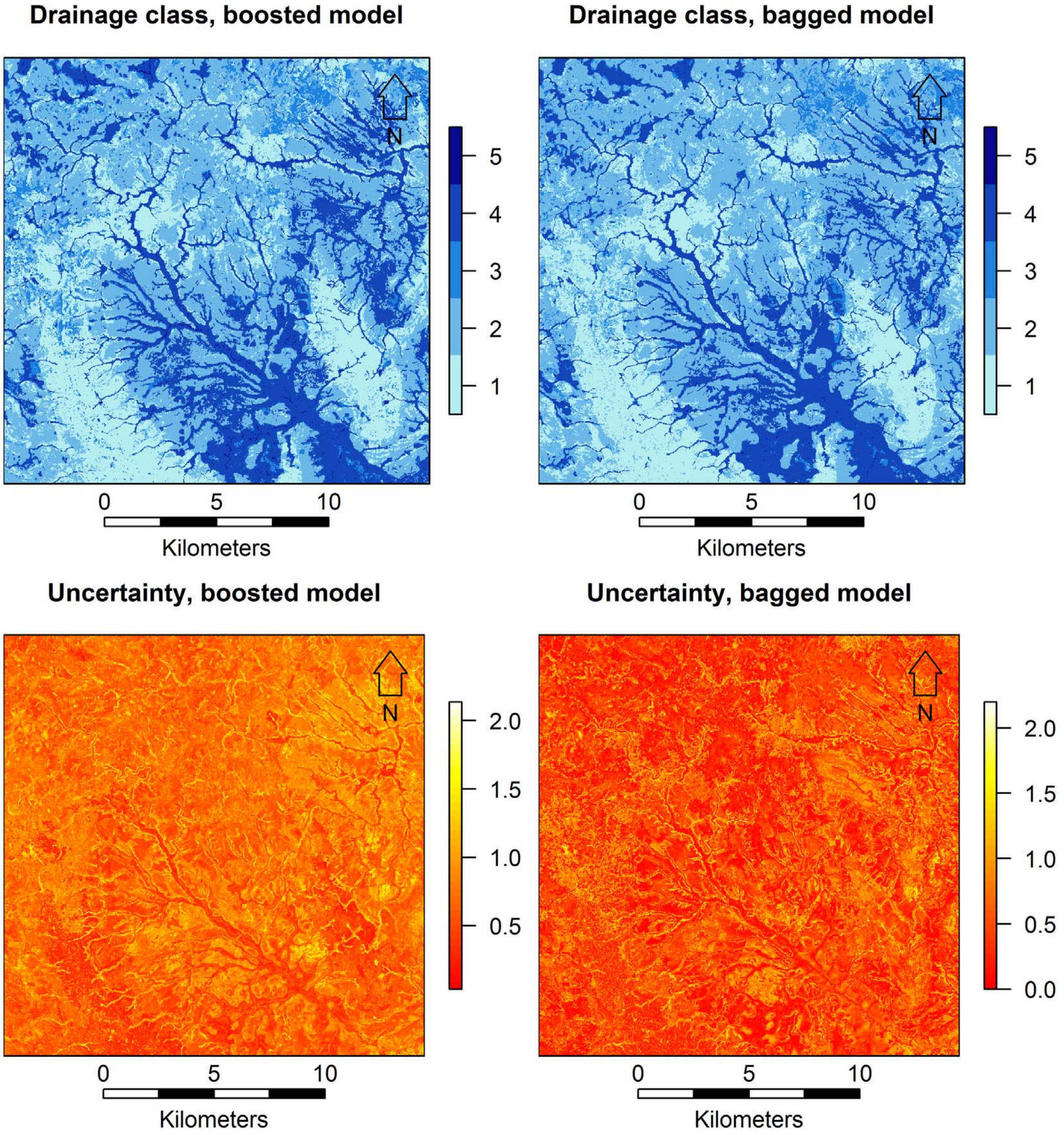
Drainage class, boosted model
Drainage class, bagged model
N
5
4
3
2
1
0
5
10
Kilometers
Uncertainty, boosted model
Uncertainty, bagged model
2.0
1.5
1.0
0.5
0.0

Fig. 10. Drainage classes and the associated uncertainties predicted by the boosted model and the bagged model for the example area shown in Fig. 9.

Table 10
Performance achieved in the present study (cross validation) compared to the performance achieved in other studies mapping drainage classes. The values are taken as presented by the authors or calculated from the data presented, when available, based on Eqs. (5) to (9).

Study	OA	K	MAE	RE	Classes
Bell et al. (1992, 1994)	0.74				3
Cialella et al. (1997)	0.81	0.72	0.33	0.32	5
Campling et al. (2002)	0.71	0.71			6
Kravchenko et al. (2002)	0.65	0.42	0.41	0.51	3
Peng et al. (2003)	0.76	0.62			3
Liu et al. (2008)	0.87	0.70	0.13	0.26	3
Zhao et al. (2008)	0.51	0.25	0.55	0.62	7
Lemercier et al. (2012)	0.52	0.27			4
Niang et al. (2012)	0.60	0.40			5
Zhao et al. (2013)	0.34	0.10	0.83	0.79	7
This study	0.49	0.28	0.74	0.63	5

as it is known to correlate with soil drainage (Ågren et al., 2014). Other studies found the topographic wetness index, or modified versions of it, to be either somewhat important (Campling et al., 2002; Liu et al., 2008; Zhao et al., 2008) or highly important (Lemercier et al., 2012) for the prediction of drainage classes. The reason for the low importance of the topographic wetness index may be that the models had a preference for the similar SAGA wetness index, which is an important variable in both of the final models. However, in this study the SAGA wetness index is the 12th most important variable in both models, while in the study of Lemercier et al. (2012), it was the second most important variable. A reason for this contrast may be the relative flatness of the study area in the present study, as the study area of Lemercier et al. (2012) had a larger range in variation and was characterized by steep slopes. It is therefore likely that topography plays a less obvious role in the present study. Secondly, the present study included a number of variables, which were not used by Lemercier et al. (2012), such as wetlands, slope to channel network, clay contents and blue spot analysis, which were found to be more important than the SAGA wetness index.

The spectral indices (NDMI, NDVI, NDWI, SAVI) had a low importance in this study. This is surprising, as spectral imagery and other remote sensing products were found to be highly important to the mapping of drainage classes in other studies (Campling et al., 2002; Cialella et al., 1997; Lemercier et al., 2012; Liu et al., 2008; Niang et al., 2012; Peng et al., 2003). In some cases the success of remote sensing products may be related to the fact that the studies were carried out on bare soil at a field scale (Liu et al., 2008; Peng et al., 2003). The study area of Cialella et al. (1997) was dominated by natural vegetation, which was found to correlate with the soil drainage conditions. In contrast, this study was carried out on a national scale with a thick vegetation cover and varying land uses, which may explain the insufficiency of spectral indices as a predictor variable. Perhaps a large bare soil image could be obtained by combining images from several points in time, as a large part of the agricultural area of Denmark is tilled. The high revisit time of many newer satellites could make this possible. For example, the Sentinel-2 mission will have a revisit time of 2–3 days at mid-latitudes once both satellites are launched (European Space Agency, n.d.). With a large bare soil image at hand, soil drainage conditions could probably be mapped more accurately using spectral imagery. Furthermore, Lemercier et al. (2012) used data from gamma-ray spectrometry, which is not as strongly affected by the vegetation cover as electromagnetic radiation in the visible to mid-infrared range (Schetselaar and Rencz, 1997; Wilford et al., 1997). It is likely that the inclusion of gamma-ray spectrometry would have improved the prediction of soil drainage conditions in the present study.

It is noteworthy that the studies which used remote sensing products generally did not include the variables, which were found to be most important in the present study, namely geology, land use, slope and distance to channel network, wetlands and clay content. The only exceptions to this rule are the studies of Campling et al. (2002), who used the horizontal distance to the channel network, Niang et al. (2012), who included a land use map, and Lemercier et al. (2012), who used layers with distances to the channel network, geological variables and land use.

5. Conclusions

The effect of implementing differential costs for misclassification depended on the ensemble technique in use. When combined with bagging, the differential costs for misclassification provided a slight increase in the performance. On the other hand, the performance of boosted decision trees deteriorated, when differential costs for misclassification were in use. The best model was achieved using differential costs for misclassification combined with bagging.

Despite the advantage of the bagged model, the results of the two final models were very similar both in terms of the predicted drainage classes and the associated uncertainties. For both models, the mapped uncertainties were related to the prediction error thereby providing an indication of the reliability of the maps. Additionally, both models had a high usage of the predictor variables wetlands, slope to channel network, clay content (100–200 cm), land use, and geology. Despite their high importance in other studies, remote sensing products had a low importance in this study.

Soil drainage classes can be mapped with many different techniques, and it cannot be ruled out that other techniques could predict drainage classes more accurately than decision tree classification. For this reason it would be relevant to test other methods as well, such as ANNs, for the prediction of drainage classes.

Overall, decision tree classification with differential costs for misclassification is an interesting technique for digital soil mapping. It is likely that the method also could be used for soil class mapping as it would be able to include similarities between soil classes. However, it should generally be tested against decision trees with equal costs for misclassification, as the results may vary.

Acknowledgements

The help of Kabindra Adhikari in producing the predictor variables is acknowledged. Furthermore, the help of the reviewers in improving the paper is acknowledged and appreciated.

References

Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P., Minasny, B., McBratney, A.B., Greve, M.H., 2013. High-resolution 3-D mapping of soil texture in Denmark. Soil Sci. Soc. Am. J. 77 (3), 860–876. http://dx.doi.org/10.2136/sssaj2012.0275.

Adhikari, K., Minasny, B., Greve, M.B., Greve, M.H., 2014. Constructing a soil class map of Denmark based on the FAO legend using digital techniques. Geoderma 214-215, 101–113. http://dx.doi.org/10.1016/j.geoderma.2013.09.023.

Ågren, A., Lidberg, W., Strömgren, M., Ogilvie, J., Arp, P., 2014. Evaluating digital terrain indices for soil wetness mapping–a Swedish case study. Hydrol. Earth Syst. Sci. 18 (9), 3623–3634.

Alhammady, H., Ramamohanarao, K., 2004. Using emerging patterns and decision trees in rare-class classification. In: Fourth IEEE International Conference on Data Mining, Proceedings, pp. 315–318. http://dx.doi.org/10.1109/Icdm.2004.10058.

Bauer, E., Kohavi, R., 1999. An empirical comparison of voting classification algorithms: bagging, boosting, and variants. Mach. Learn. 36 (1–2), 105–139. http://dx.doi.org/10.1023/A:1007515423169.

Bell, J.C., Cunningham, R.L., Havens, M.W., 1992. Calibration and validation of a soil-landscape model for predicting soil drainage class. Soil Sci. Soc. Am. J. 56, 1860–1866. http://dx.doi.org/10.2136/sssaj1992.03615995005600060035x.

Bell, J.C., Cunningham, R.L., Havens, M.W., 1994. Soil drainage class probability mapping using a soil-landscape model. Soil Sci. Soc. Am. J. 58, 464–470. http://dx.doi.org/10.2136/sssaj1994.03615995005800020031x.

Böhner, J., Köthe, R., Conrad, O., Gross, J., Ringeler, A., Selige, T., 2002. Soil regionalisation by means of terrain analysis and process parameterisation. In: Micheli, E., Nachtergaele, F., Montanarella, L. (Eds.), Soil Classification 2001 - Research Report No. 7, EUR 20398 EN. European Soil Bureau, Luxembourg, pp. 213–222.

Brady, N.C., Weil, R.R., 1996. The Nature and Properties of Soils. Prentice-Hall Inc.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140. http://dx.doi.org/10.1023/A:1018054314350.

Campling, P., Gobin, A., Feyen, J., 2002. Logistic modeling to spatially predict the probability of soil drainage classes. Soil Sci. Soc. Am. J. 66 (4), 1390–1401. http://dx.doi.org/10.2136/sssaj2002.1390.

Chaney, N.W., Wood, E.F., McBratney, A.B., Hempel, J.W., Nauman, T.W., Brungard, C.W., Odgers, N.P., 2016. POLARIS: a 30-meter probabilistic soil series map of the contiguous United States. Geoderma 274, 54–67. http://dx.doi.org/10.1016/j.geoderma.2016.03.025.

Chawla, N.V., 2003. C4.5 and imbalanced data sets: investigating the effect of sampling method, probabilistic estimate, and decision tree structure. Proc. ICML 3.

Chen, G., Ma, L., 2010. Research on Rough Set and Decision Tree Method Application in Evaluation of Soil Fertility Level. Computer and Computing Technologies in Agriculture, Nanchang, China, October 22–25, 2010, Selected Papers, Part II. pp. 408–414. http://dx.doi.org/10.1007/978-3-642-18336-2_50.

Cialella, A.T., Dubayah, R., Lawrence, W., Levine, E., 1997. Predicting soil drainage class using remotely sensed and digital elevation data. Photogramm. Eng. Remote. Sens. 63 (2), 171–178.

Cieslak, D.A., Chawla, N.V., 2008. Learning Decision Trees for Unbalanced Data. Machine Learning and Knowledge Discovery in Databases, Part I, Proceedings. 5211. pp. 241–256. http://dx.doi.org/10.1007/978-3-540-87479-9_34.

Collaku, A., Harrison, S.A., 2002. Losses in wheat due to waterlogging. Crop Sci. 42 (2), 444–450.

Dietterich, T.G., 2000. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. Mach. Learn. 40 (2), 139–157. http://dx.doi.org/10.1023/A:1007607513941.

Drummond, C., Holte, R.C., 2000. Exploiting the Cost (In) Sensitivity of Decision Tree Splitting Criteria. Proceedings of the Seventeenth International Conference on Machine Learning. Morgan Kaufmann Publishers Inc., pp. 239–246.

Ernstsen, V., Olsen, P., Rosenbom, A.E., 2015. Long-term monitoring of nitrate transport to drainage from three agricultural clayey till fields. Hydrol. Earth Syst. Sci. 19 (8), 3475–3488. http://dx.doi.org/10.5194/hess-19-3475-2015.

European Environment Agency, 2014. Corine land cover (CLC) 2012 - Denmark, version 1, Oct. 2014. http://download.kortforsyningen.dk/content/corine-land-cover.

European Space Agency Missions: sentinel-2. https://sentinel.esa.int/web/sentinel/missions/sentinel-2 (n.d., accessed 14-12-16).

FAO, 1977. Guidelines for Soil Profile Description. FAO, Rome, Italy.

Freund, Y., Schapire, R.E., 1996. Experiments with a new boosting algorithm. ICML 96, 148–156.

Gallant, J.C., Dowling, T.I., 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. Water Resour. Res. 39 (12), 1347. http://dx.doi.org/10.1029/2002wr001426.

Gambrell, R.P., Gilliam, J.W., Weed, S.B., 1975. Denitrification in subsoils of the North Carolina Coastal Plain as affected by soil drainage. J. Environ. Qual. 4 (3), 311–316. http://dx.doi.org/10.2134/jeq1975.00472425000400030005x.

Giasson, E., Sarmento, E.C., Weber, E., Flores, C.A., Hasenack, H., 2011. Decision trees for digital soil mapping on subtropical basaltic steeplands. Sci. Agric. 68 (2), 167–174. http://dx.doi.org/10.1590/S0103-90162011000200006.

Greve, M.H., Christensen, O.F., Greve, M.B., Kheir, R.B., 2014. Change in peat coverage in Danish cultivated soils during the past 35 years. Soil Sci. 179 (5), 250–257. http://dx.doi.org/10.1097/ss.0000000000000066.

Haibo, H., Garcia, E.A., 2009. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. 21 (9), 1263–1284. http://dx.doi.org/10.1109/tkde.2008.239.

Henriksen, H.J., Højberg, A.L., Olsen, M., Seaby, L.P., van der Keur, P., Stisen, S., Troldborg, L., Sonnenborg, T.O., Refsgaard, J.C., 2012. Klimaeffekter på hydrologi og grundvand - Klimagrundvandskort. Aarhus University.

Holte, R.C., Acker, L., Porter, B.W., 1989. Concept learning and the problem of small disjuncts. IJCAI 89, 813–818.

Jacobsen, N.K., 1984. Soil map of Denmark according to the FAO-UNESCO legend. Dan. J. Geogr. 84, 93–98. http://dx.doi.org/10.1080/00167223.1984.10649206.

Jakobsen, P.R., Hermansen, B., Tougaard, L., 2015. Danmarks digitale jordartskort 1:25000 version 4.0. In: GEUS.

Kheir, R.B., Bøcher, P.K., Greve, M.B., Greve, M.H., 2010a. The application of GIS based decision-tree models for generating the spatial distribution of hydromorphic organic landscapes in relation to digital terrain data. Hydrol. Earth Syst. Sci. 14 (6), 847–857. http://dx.doi.org/10.5194/hess-14-847-2010.

Kheir, R.B., Greve, M.H., Bocher, P.K., Greve, M.B., Larsen, R., McCloy, K., 2010b. Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: the case study of Denmark. J. Environ. Manag. 91 (5), 1150–1160. http://dx.doi.org/10.1016/j.jenvman.2010.01.001.

Kravchenko, A.N., Bollero, G.A., Omonode, R.A., Bullock, D.G., 2002. Quantitative mapping of soil drainage classes using topographical data and soil electrical conductivity. Soil Sci. Soc. Am. J. 66 (1), 235–243. http://dx.doi.org/10.2136/sssaj2002.0235.

Kuhn, M., Weston, S., Coulter, N., Culp, M., Quinlan, J.R., 2015. Package 'C50'. https://cran.r-project.org/web/packages/C50/C50.pdf (accessed 20-07-16).

Lagacherie, P., Holmes, S., 1997. Addressing geographical data errors in a classification tree for soil unit prediction. Int. J. Geogr. Inf. Sci. 11 (2), 183–198. http://dx.doi.org/10.1080/136588197242455.

Lark, R.M., 1995. Components of accuracy of maps with special reference to discriminant-analysis on remote sensor data. Int. J. Remote Sens. 16 (8), 1461–1480. http://dx.doi.org/10.1080/01431169508954488.

Lemercier, B., Lacoste, M., Loum, M., Walter, C., 2012. Extrapolation at regional scale of local soil knowledge using boosted classification trees: a two-step approach. Geoderma 171-172, 75–84. http://dx.doi.org/10.1016/j.geoderma.2011.03.010.

Levine, E., Knox, R., Lawrence, W., 1994. Relationships between soil properties and vegetation at the Northern Experimental Forest, Howland, Maine. Remote Sens. Environ. 47 (2), 231–241. http://dx.doi.org/10.1016/0034-4257(94)90158-9.

Ling, C.X., Sheng, V.S., 2010. Cost-sensitive learning and the class imbalance problem. In: Sammut, C., Webb, G.I. (Eds.), Encyclopedia of Machine Learning. Springer.

Liu, J., Pattey, E., Nolin, M.C., Miller, J.R., Ka, O., 2008. Mapping within-field soil drainage using remote sensing, DEM and apparent soil electrical conductivity. Geoderma 143 (3–4), 261–272. http://dx.doi.org/10.1016/j.geoderma.2007.11.011.

Madsen, H.B., Jensen, N.H., 1988. Vejledning til beskrivelse af jordbundsprofiler. Arealdatakontoret, Landbrugsministeriet.

Madsen, H.B., Nørr, A.H., Holst, K.A., 1992. The Danish Soil Classification. The Royal Danish Geographical Society, Copenhagen, Denmark.

Mitchell, T., 1997. Decision tree learning. In: Machine Learning. McGraw Hill, New York, pp. 52–80.

Moran, C.J., Bui, E.N., 2002. Spatial data mining for enhanced soil map modelling. Int. J. Geogr. Inf. Sci. 16 (6), 533–549. http://dx.doi.org/10.1080/13658810210138715.

NASA Landsat Program, 2014. Landsat OLI/TIRS scenes LC81920222014090LGN00, LC81940212014072LGN00, LC81940222014072LGN00, LC81950212014079LGN00, LC81950222014079LGN00, LC81960202014070LGN00, LC81960212014070LGN00, LC81960222014070LGN00, LC81980202014068LGN00, LC81980212014068LGN00, L1T. USGS, Sioux Falls.

National Survey and Cadastre, 2012. Danmarks Højdemodel 2007, DHM-2007/Terræn. National Survey and Cadastre.

Niang, M.A., Nolin, M., Bernier, M., Perron, I., 2012. Digital mapping of soil drainage classes using multitemporal RADARSAT-1 and ASTER images and soil survey data. Appl. Environ. Soil Sci. 2012, 1–17. http://dx.doi.org/10.1155/2012/430347.

Nuutinen, V., Pöyhönen, S., Ketoja, E., Pitkänen, J., 2001. Abundance of the earthworm *Lumbricus terrestris* in relation to subsurface drainage pattern on a sandy clay field. Eur. J. Soil Biol. 37 (4), 301–304. http://dx.doi.org/10.1016/S1164-5563(01)01105-0.

Odgers, N.P., Sun, W., McBratney, A.B., Minasny, B., Clifford, D., 2014. Disaggregating and harmonising soil map units through resampled classification trees. Geoderma 214-215, 91–100. http://dx.doi.org/10.1016/j.geoderma.2013.09.024.

Peng, W., Wheeler, D.B., Bell, J.C., Krusemark, M.G., 2003. Delineating patterns of soil drainage class on bare soils using remote sensing analyses. Geoderma 115 (3–4), 261–279. http://dx.doi.org/10.1016/s0016-7061(03)00066-1.

Quinlan, J.R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann.

Quinlan, J.R., 1996. Learning decision tree classifiers. ACM Comput. Surv. 28 (1), 71–72. http://dx.doi.org/10.1145/234313.234346.

Ren, B., Zhang, J., Li, X., Fan, X., Dong, S., Liu, P., Zhao, B., 2014. Effects of waterlogging on the yield and growth of summer maize under field conditions. Can. J. Plant Sci. 94 (1), 23–31. http://dx.doi.org/10.4141/cjps2013-175.

Rokach, L., Maimon, O., 2005. Decision trees. In: Data Mining and Knowledge Discovery Handbook. Springer, pp. 165–192.

Schelde, K., de Jonge, L.W., Kjaergaard, C., Laegdsmand, M., Rubæk, G.H., 2006. Effects of manure application and plowing on transport of colloids and phosphorus to tile drains. Vadose Zone J. 5 (1), 445. http://dx.doi.org/10.2136/vzj2005.0051.

Schetselaar, E.M., Rencz, A.N., 1997. Reducing the effects of vegetation cover on airborne radiometric data using Landsat TM data. Int. J. Remote Sens. 18 (7), 1503–1515.

Scull, P., Franklin, J., Chadwick, O., McArthur, D., 2003. Predictive soil mapping: a review. Prog. Phys. Geogr. 27 (2), 171–197. http://dx.doi.org/10.1191/0309133303pp366ra.

Smith, K.A., Ball, T., Conen, F., Dobbie, K.E., Massheder, J., Rey, A., 2003. Exchange of greenhouse gases between soil and atmosphere: interactions of soil physical factors and biological processes. Eur. J. Soil Sci. 54 (4), 779–791. https://doi.org/10.1046/j.1351-0754.2003.0567.x.

Statistics Denmark Arealanvendelse. https://www.dst.dk/da/Statistik/emner/areal/arealanvendelse (n.d., accessed 09-12-16).

Taghizadeh-Mehrjardi, R., Sarmadian, F., Minasny, B., Triantafilis, J., Omid, M., 2014. Digital mapping of soil classes using decision tree and auxiliary data in the Ardakan Region, Iran. Arid Land Res. Manag. 28 (2), 147–168. http://dx.doi.org/10.1080/15324982.2013.828801.

Tan, P.-N., Steinbach, M., Kumar, V., 2014. Classification: Basic Concepts, Decision Trees, and Model Evaluation. In: Introduction to Data Mining. 2014. Pearson Education, Limited, pp. 145–205.

The Danish Agrifish Agency, 2014. Markkort. http://www.geodata-info.dk/Portal/ShowMetadata.aspx?id=6e3bc77f-c193-4508-80d4-836e1668db91 (accessed 29-05-2017).

Ting, K.M., 1998. Inducing cost-sensitive trees via instance weighting. In: European Symposium on Principles of Data Mining and Knowledge Discovery. Springer, pp. 139–147.

Van de Noort, R., Fletcher, W., Thomas, G., Carstairs, I., Patrick, D., 2002. Monuments at risk in England's wetlands. University of Exeter.

Wang, P.R., 2013. Referenceværdier: Døgn-, måneds- og årsværdier for regioner og hele landet 2001–2010, Danmark for temperatur, relativ luftfugtighed, vindhastighed, globalstråling og nedbør. Teknisk Rapport 12–24. Danish Meteorological Institute.

Watson, E.R., Lapins, P., Barron, R.J.W., 1976. Effect of waterlogging on the growth, grain and straw yield of wheat, barley and oats. Anim. Prod. Sci. 16 (78), 114–122. http://dx.doi.org/10.1071/EA9760114.

Weiss, G.M., 1995. Learning with rare cases and small disjuncts. ICML 558–565.

Wilford, J., Bierwirth, P., Craig, M., 1997. Application of airborne gamma-ray spectrometry in soil/regolith mapping and applied geomorphology. AGSO J. Aust. Geol. Geophys 17 (2), 201–216.

Zhang, X., Lin, F., Jiang, Y., Wang, K., Wong, M.T., 2008. Assessing soil Cu content and anthropogenic influences using decision tree analysis. Environ. Pollut. 156 (3), 1260–1267. http://dx.doi.org/10.1016/j.envpol.2008.03.009.

Zhao, Z.Y., Chow, T.L., Yang, Q., Rees, H.W., Benoy, G., Xing, Z.S., Meng, F.R., 2008. Model prediction of soil drainage classes based on digital elevation model parameters and soil attributes from coarse resolution soil maps. Can. J. Soil Sci. 88 (5), 787–799. http://dx.doi.org/10.4141/CJSS08012.

Zhao, Z.Y., Ashraf, M.I., Meng, F.-R., 2013. Model prediction of soil drainage classes over a large area using a limited number of field samples: a case study in the province of Nova Scotia, Canada. Can. J. Soil Sci. 93 (1), 73–83. http://dx.doi.org/10.4141/cjss2011-095.