Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach

Bas Kempen [a,b,*], Dick J. Brus [b], Gerard B.M. Heuvelink [a,b], Jetse J. Stoorvogel [a]

[a] Wageningen University, Land Dynamics Group, P.O. Box 47, 6700 AA Wageningen, The Netherlands
[b] Alterra, Soil Science Centre, Wageningen University and Research Centre, P.O. Box 47, 6700 AA Wageningen, The Netherlands

ARTICLE INFO

ABSTRACT

The 1:50,000 national soil survey of the Netherlands, completed in the early 1990s after more than three decades of mapping, is gradually becoming outdated. Large-scale changes in land and water management that took place after the field surveys have had a great impact on the soil. Especially oxidation of peat soils has resulted in a substantial decline of these soils. The aim of this research was to update the national soil map for the province of Drenthe (2680 km^2) without additional fieldwork through digital soil mapping using legacy soil data. Multinomial logistic regression was used to quantify the relationship between ancillary variables and soil group. Special attention was given to model-building as this is perhaps the most crucial step in digital soil mapping. A framework for building a logistic regression model was taken from the literature and adapted for the purpose of soil mapping. The model-building process was guided by pedological expert knowledge to ensure that the final regression model is not only statistically sound but also pedologically plausible. We built separate models for the ten major map units, representing the main soil groups, of the national soil map for the province of Drenthe. The calibrated models were used to estimate the probability of occurrence of soil groups on a 25 m grid. Shannon entropy was used to quantify the uncertainty of the updated soil map, and the updated soil map was validated by an independent probability sample. The theoretical purity of the updated map was 67%. The estimated actual purity of the updated map, as assessed by the validation sample, was 58%, which is 6% larger than the actual purity of the national soil map. The discrepancy between theoretical and actual purity might be explained by the spatial clustering of the soil profile observations used to calibrate the multinomial logistic regression models and by the age difference between calibration and validation observations.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The 1:50,000 soil map is the major source of soil information in the Netherlands. It is used for a wide variety of purposes such as agricultural and environmental policy making, nature and soil conservation and archeological prospection. However, the soil map, completed in the early 1990s after more than three decades of field surveys, is gradually becoming outdated (De Vries and Brouwer, 2006; Rosing et al., 2006). Large-scale changes in land and water management that took place after the field surveys have had a great impact on the soil. One of the most remarkable effects of human interventions in the Dutch landscape is the decline of large areas of peat soils through increased oxidation rates due to intensive tillage and lowering of groundwater levels. A quick scan on the status of the map units with thick peat soils (peat layer>40 cm) showed that 47% of the area mapped as thick peat soils during the 1:50,000 national survey changed into another soil type (Van Kekem et al., 2005). Finke et al.

(1996) mapped the thickness of the peat layer of peat soils for two map sheets of the national soil map for the province of Drenthe (800 km^2), and compared the results with peat thickness according to the soil map. They found that 82% of the thick peat soils had changed into thin peat soils, and that 63% of the thin peat soils had changed into mineral soils. Use of outdated soil information for environmental research or policy making may lead to erroneous decisions.

Although the need for updating the soil map has been recognized for a long time, the last update activities took place in the early nineties. Four map sheets of the national soil map were updated between 1988 and 1993. Brus et al. (1992) evaluated the merits of four update strategies for soil maps. Budgetary reasons currently hamper further map updates. As fieldwork is a major cost component in a project on map updating (Finke, 2000), methods that reduce the amount of fieldwork, such as used in digital soil mapping, are potentially interesting for future update activities.

In digital soil mapping soil observations are related to readily available, spatially exhaustive ancillary data. The relationships are then extended across a survey area to predict soil at unvisited locations (McBratney et al., 2003; Bui and Moran 2003). Such methods also may have potentials in the Netherlands considering its

---

* Corresponding author. Wageningen University, Land Dynamics Group, P.O. Box 47, 6700 AA Wageningen, The Netherlands. Tel.: +31 317 482416; fax: +31 317 419000.
E-mail address: bas.kempen@wur.nl (B. Kempen).

data rich environment. The Dutch soil information system (DSIS) contains soil profile descriptions and classifications at more than 260,000 locations. Most of these are located in areas where soil surveys at scale 1:10,000 have been carried out. Besides this, an extensive suite of spatially exhaustive environmental ancillary data is available at 25 m resolution. The ancillary data were combined with point observations to create maps of groundwater status for the sandy soils of the Netherlands, using time series analysis and geostatistical techniques (Finke et al., 2004). Brus et al. (2008) used over 8000 soil point observations stored in the DSIS to estimate the probabilities of occurrence of seven soil categories in the Netherlands. We hypothesize that existing, recent soil profile observations can be used to update the peat and other map units of the national soil map. Furthermore, we expect that the purity of the map units can be increased by using high-resolution ancillary data to delineate inclusions of soil classes other than the dominant soil class.

The objective of this paper is to update the national soil map for the province of Drenthe without additional fieldwork by using legacy soil data. A soil map update is urgent in Drenthe considering the large area of peat soils, the extensive areal decline of these soils (Van Kekem et al., 2005; De Vries and Brouwer, 2006) and the age of the existing soil map: the soil survey in Drenthe took place between 1965 and 1988. We explore the use of multinomial logistic regression (MLR) for digital soil mapping. MLR is widely used for spatial modeling in land use and ecology studies (Müller and Zeller, 2002; Rhemtulla et al., 2007; May et al., 2008; Suring et al., 2008). However, only a few studies applied MLR for digital soil mapping, see for instance Campling et al. (2002), Bailey et al. (2003), Hengl et al. (2007a) and Debella-Gilo and Etzelmüller (2009).

In this study, existing soil profile classifications in the DSIS are used to calibrate an MLR-model for each of the ten major map units of the national soil map of the province of Drenthe. With these models we re-map the soil group within each map unit. Careful attention is given to the model-building process, as this is perhaps the most crucial step in the digital soil mapping process. A framework for building logistic regression models is taken from the literature (Hosmer and Lemeshow, 1989) and adapted for soil mapping. We illustrate this framework for one of the peat map units. The purity of the updated soil map is assessed by an independent probability sample and compared to that of the existing soil map.

2. Methods

2.1. Study area

The province of Drenthe (2680 km^2) is situated in the northeastern part of the Netherlands between 52°12′ and 53°12′ northern latitude and 6°7′ and 7°5′ eastern longitude (Fig. 1). Altitude ranges between −1 and 30 m above sea level. The landscape of Drenthe is dominated by the gently west–east sloping till plateau and the Hunze valley that borders the plateau in the East. Glacial till was deposited under the continental ice sheet that covered the Northern Netherlands 160,000 years ago during the Saalian ice age. The most remarkable landscape feature of the till plateau is the Hondsrug, a straight, Northwest–Southeast oriented till ridge. Meltwater incised the till plateau during the early Weichselian (116–73 ka BP), forming large brook valley systems. The brook valleys were partly filled with fluvial deposits in the mid Weichselian (73-14.5 ka BP). A layer of coversand of up to 2 m thick was deposited on the till during the late Weichselian (14.5–11.5 ka BP) Sediments in the brook valleys were covered with fen peat during the early Holocene. At the same time, oligotrophic peat started to form in depressions on the plateau and in the Hunze valley and grew into highmoor bogs, which eventually covered one third of Drenthe (Spek, 2004). During the Middle Ages, drift-sand complexes formed on the plateau as a result of the open field farming system. Between the 17th and mid 20th century large-scale, systematic reclamation of the vast highmoor swamps took place resulting in a completely man-made landscape, the so-called peat-colonial landscape. Fig. 1 shows the major landforms of Drenthe.

All soils of Drenthe were formed during the Holocene. Podzols formed in poor coversand deposits on the plateau. In richer, more loamy parent material, brown forest soils formed. Plaggen soils, a result of the open field farming system (Pape, 1970; Spek, 2004), surround medieval settlements on the plateau. Peat soils dominate the centres of the brook valleys. Earth soils (soils with a humic topsoil overlying the C-horizon) are found in the brook valley–plateau transition zone. Vague soils (soils without pronounced signs of soil
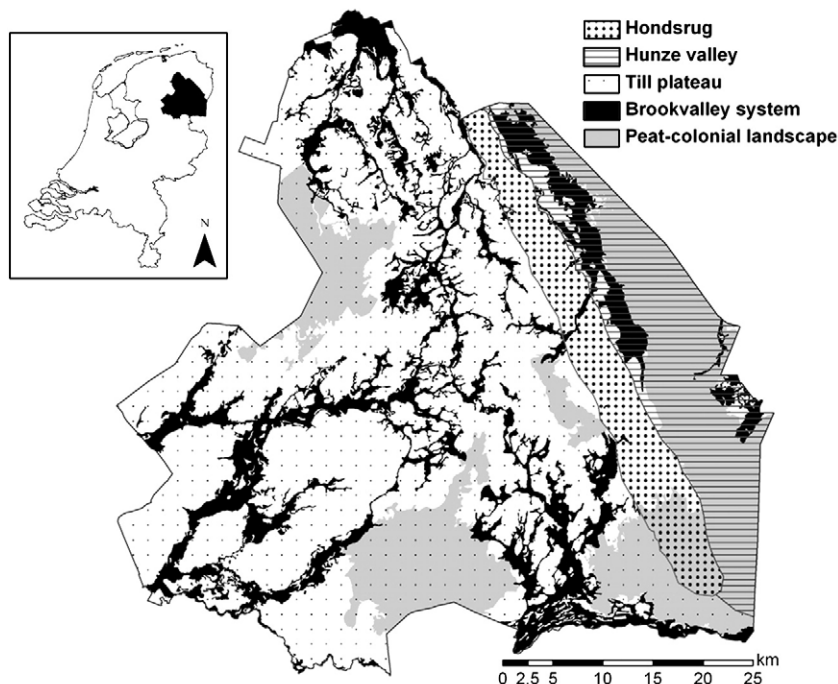

Fig. 1. The major landforms of the province of Drenthe. The insert top left shows the location of Drenthe in the Netherlands.

formation) are found in the drift-sand complexes and in the till of the Hondsrug. Soils of the peat-colonial landscape are distinguished from the peat soils of the brook valleys by their strong human disturbances to large depth due to deep cultivation and their man-made topsoil, the so-called peat-colonial topsoil.

2.2. Data sources

2.2.1. Soil data

The national soil map for Drenthe contains 96 map units describing 61 soil types at the subgroup level of the Dutch soil classification system (De Bakker and Schelling, 1989), and 35 associations. Because it is practically unfeasible to build an MLR-model for a categorical variable with 96 possible outcomes, we aggregated the 96 map units into ten map units, representing the major soil groups (Fig. 2):

1. *Thick peat soils* (*P*) (25,000 ha): soils with a peaty (organic matter content > 15%) surface horizon; at least 40 cm of peat within 80 cm from the surface;
2. *Thick peat soils with a mineral surface horizon* (*mP*) (24,800 ha): soils with at least 40 cm of peat within 80 cm from the surface, and a sandy, clayey, or peat-colonial surface horizon less than 40 cm thick;
3. *Thin peat soils* (*PY*) (13,400 ha): soils with a peaty surface horizon; at most 40 cm of peat within 80 cm from the surface;
4. *Thin peat soils with a mineral surface horizon* (*mPY*) (36,000 ha): soils with at most 40 cm of peat within 80 cm from the surface, and a sandy, clayey or peat-colonial surface horizon less than 40 cm thick;
5. *Brown forest soils* (*BF*) (900 ha): soils with a B-horizon formed by weathering of minerals and illuviation of moder humus;
6. *Podzol soils* (*PZ*) (93,000 ha): xeromorphic and hydromorphic podzols;
7. *Earth soils* (*E*) (13,000 ha): hydromorphic soils with a 15–50 cm thick humic A-horizon, overlying a sandy or loamy C-horizon with or without gleyic features;
8. *Plaggen soils* (*PS*) (17,000 ha): soils with an anthropogenic, humic A-horizon thicker than 30 cm overlying a podzol or brown forest soil; typical for the open fields on the Drenthe plateau;

9. *Till soils* (*T*) (3500 ha): soils with glacial till within 40 cm from the surface;
10. *Sandy vague soils* (*S*) (5800 ha): These are sandy soils with a humus-poor topsoil < 30 cm thick; subsoil only shows initial or no signs of soil formation.

Although some detail is lost by aggregating map units, the ten soil groups still describe the major soil variation in Drenthe. We obtained 16,282 soil profile descriptions for Drenthe from the DSIS. Roughly 96% (15,580) of the soil profile observations are located in four areas where 1:10,000 soil surveys were carried out between 1996 and 2005. These areas cover 10% of the total area. The remaining 702 profile observations, collected during various research projects, are scattered across Drenthe (Fig. 2). Because of the variety of data sources, the profile observations were collected with different sampling designs. The sampling locations in the 1:10,000 survey areas were selected by purposive sampling. The other locations were selected by both purposive (*n* = 434) and probability sampling (*n* = 268).

The recorded soil types at the observation locations were reclassified to the ten soil groups. Cross-tabulation of the field-observed versus the mapped soil group showed that at 55% of the observation locations the recorded soil group corresponds to the soil group as depicted on the map. The podzol map unit is the most pure map unit (78%) and the thin peat soils are the least pure (PY: 19% and mPY: 30%). The latter shows the effect of oxidation of the peat.

2.2.2. Environmental ancillary data

Twelve spatially exhaustive primary datasets were available (Table 1). The polygon maps were converted to 25 m grids. The DEM was used to derive four relative elevation maps using the local mean elevation within search radii of 250, 500, 750 and 1000 m. Groundwater table classes were grouped into three classes. Historic land cover (HLC) was grouped into five classes and converted to a 25 m grid. The three recent land cover datasets (LC) were combined into a land cover history map with five classes for the period 1997–2003. The geomorphological units were grouped into 16 classes. The paleogeography grid contained 12 classes. After preprocessing the
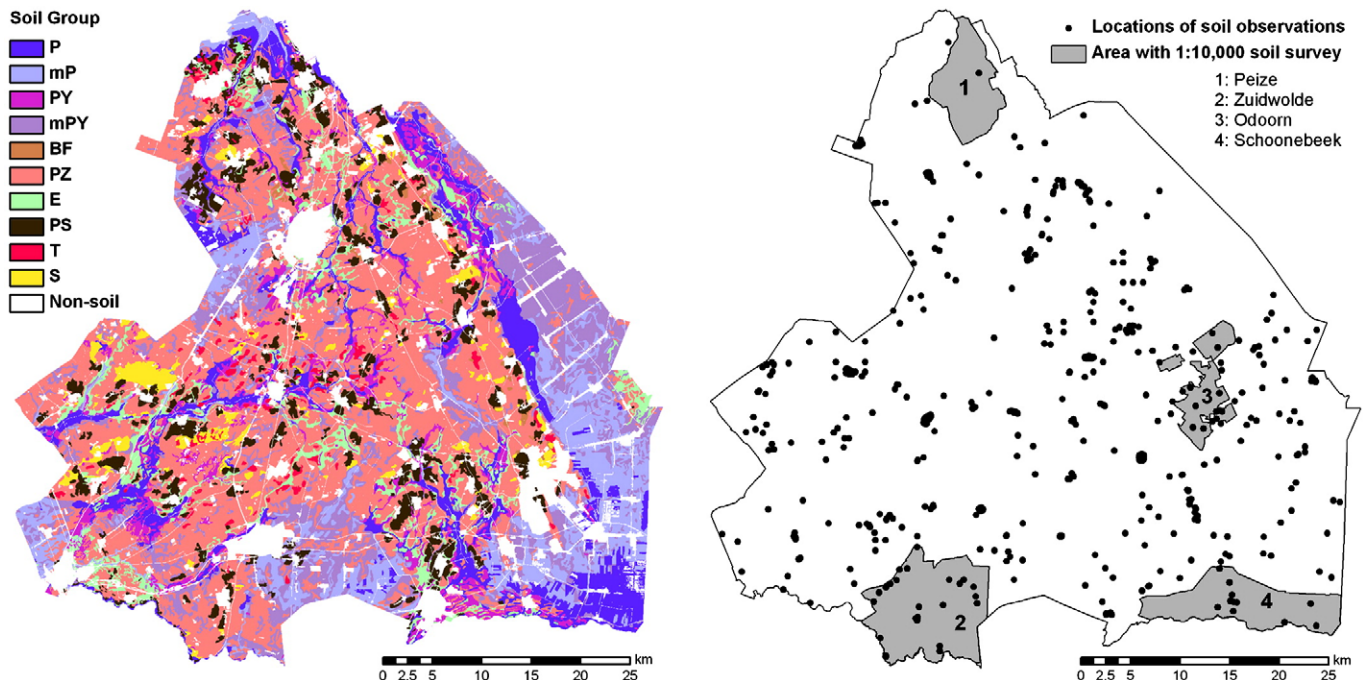

Fig. 2. Soil data: aggregated soil map at scale 1:50,000 of Drenthe (left) and locations of the four 1:10,000 soil survey areas and the 702 soil profile observations (right).

Table 1
Available primary environmental ancillary data.

Dataset	Description	Resolution	Reference
Digital elevation model	Absolute elevation	25 m	http://www.ahn.nl
Groundwater			
Groundwater table class (GT)	Seasonal fluctuation of phreatic water levels	1:50,000 (polygon)	Finke (2000)
Groundwater dynamics class (GD)	Updated GT map: quantitative set of parameters describing groundwater dynamics	25 m	Finke et al. (2004)
GD−mean highest water table (MHW)		25 m	
GD−mean lowest water table (MLW)		25 m	
Land cover			
HLC	Land cover in 1900	50 m	Knol et al. (2004)
LC1997	Land cover in 1997	25 m	http://www.lgn.nl
LC2000	Land cover in 2000	25 m	
LC2003	Land cover in 2003	25 m	
Paleogeography	Reconstruction of the landscape of Drenthe by the end of the early Middle Ages (ca. 1000 AD)	1:50,000 (polygon)	Spek (2004)
Geomorphology	Geomorphological units	1:50,000 (polygon)	Koomen and Maas (2004)
Soil map (scale 1:50,000)	Spatial distribution of soil classes	1:50,000 (polygon)	Steur and Heijink (1991)

primary datasets, ancillary data were grouped into eight groups: (1) elevation, (2) relative elevation, (3) groundwater, (4) historic land cover, (5) recent land cover, (6) paleogeography, (7) geomorphology, and (8) soil.

2.3. Multinomial logistic regression

2.3.1. The logistic model

The logistic model belongs to the family of generalized linear models and is used when the response variable is a categorical variable. Suppose that variable Y_i represents the observed soil group at a sampling location, with $i = 1,..., n$ and n is the number of soil groups in a survey area. In case n equals 2 and Y has outcomes Y_1 and Y_2. Both the counts of Y_1 and Y_2 follow a binomial distribution. The probability of occurrence of Y_1 is π_1 and that of Y_2 is π_2. Logistic regression relates probability π_1 to a set of predictors using the logit link function:

$$\text{logit}(\pi_1) = \ln\left(\frac{\pi_1}{\pi_2}\right) = \ln\left(\frac{\pi_1}{1-\pi_1}\right) = \mathbf{x}'\boldsymbol{\beta} \tag{1}$$

where $\mathbf{x}$ is a vector of predictors, and $\boldsymbol{\beta}$ is a vector of model coefficients that are typically estimated by maximum likelihood. Eq. (1) can be rewritten as:

$$\frac{\pi_1}{1-\pi_1} = \exp(\mathbf{x}'\boldsymbol{\beta}) = \exp(\eta). \tag{2}$$

The quotient in Eq. (2) is referred to as the *odds*. From Eq. (2) follows that:

$$\pi_1 = \frac{\exp(\eta)}{1 + \exp(\eta)}. \tag{3}$$

The binomial logistic regression model is easily generalized to the multinomial case. If there are n soil groups there are also n variables $Y_1,..., Y_n$ with corresponding probabilities of occurrence $\pi_1,..., \pi_n$. Analogous to binomial logistic regression the odds $\pi_1/\pi_n,..., \pi_{n-1}/\pi_n$ are modelled by means of $\exp(\eta_1),..., \exp(\eta_{n-1})$. From $\sum_{i=1}^{n} \pi_i = 1$ it follows that:

$$\pi_i = \frac{\exp(\eta_i)}{\exp(\eta_1) + \exp(\eta_2) + ... + \exp(\eta_n)} \tag{4}$$

where $\eta_n = 0$. This model ensures that all probabilities are in the interval [0,1] and that the probabilities sum to 1.

2.3.2. Assessing model significance and contribution of predictors

The significance of the logistic regression model is assessed with the *likelihood ratio test*. Central to this test is the *deviance* statistic, which is defined as (Hosmer and Lemeshow, 1989):

$$D = -2\ln\left(\frac{\text{likelihood fitted model}}{\text{likelihood saturated model}}\right) \tag{5}$$

where the quotient is the likelihood ratio. The larger the deviance D, the poorer the fit of the fitted model compared to the saturated model. The likelihood ratio test compares two logistic models by assessing the change in deviance due to inclusion of predictors (Hosmer and Lemeshow, 1989):

$$G = D(\text{model without the predictor}) - D(\text{model with the predictor}). \tag{6}$$

G is the log-likelihood ratio statistic, which is chi-square distributed under the null hypothesis that the model coefficients are zero, assuming independent and normally distributed residuals. The likelihood ratio test is used to assess the significance of the overall model by comparing the deviance of the intercept-only model with the full model, and that of the individual predictors. Note that the likelihood ratio test can only be used to compare nested models.

The significance of an individual model coefficient is assessed with the Wald statistic, which is obtained by comparing the estimated coefficient to an estimate of its standard error (Hosmer and Lemeshow, 1989):

$$W = \hat{\beta} / S\hat{E}\left(\hat{\beta}\right). \tag{7}$$

The Wald statistic follows the standard normal distribution under the null hypothesis that a model coefficient is zero. Important for the interpretation of the logistic regression is the value of $\exp(\beta)$, the odds ratio, which indicates the change in odds of an event resulting from a one-unit change in the predictor.

2.4. Model-building

2.4.1. Pedological knowledge for regression modeling

Regression modeling is a popular data-driven method for quantifying the relationship between soil and ancillary data (Thompson and Kolka, 2005; Meersmans et al., 2008; Schulp and Veldkamp, 2008). Usually a set of predictors is derived from ancillary data, coefficients are estimated for these predictors, followed by an evaluation of the selected model on the basis of some statistical performance

criterion such as R^2 or Mallows' Cp. The resulting regression model might be statistically sound but can be pedologically questionable if the selected predictors do not have a plausible relationship with the soil variable based on knowledge of the soil–landscape system.

An alternative to data-driven approaches to digital soil mapping are knowledge-driven approaches. It has become widely recognized that tacit knowledge of the soil–landscape system provides valuable information that should be integrated into the digital soil mapping process (Heuvelink and Webster, 2001; McKenzie and Gallant, 2007; Walter et al., 2007). Such knowledge can be used to build expert systems for mapping soils (Cook et al., 1996; Zhu et al., 2001) or to define a conceptual model of pedogenesis that forms the foundation of a quantitative (statistical) model for digital soil mapping (McKenzie and Ryan, 1999; McKenzie and Gallant, 2007). In case of regression modeling, use of knowledge of the soil–landscape system should be fully integrated throughout the process of model-building. Each step of the process must be critically reviewed from statistical as well as pedological perspectives. One should have confidence in the final regression model, not only statistically but also pedologically.

2.4.2. Model-building strategy

Hosmer and Lemeshow (1989) provide a methodological framework for building a binomial or multinomial logistic regression models. We adopted and extended their approach for digital mapping of multinomial, categorical soil variables. We describe now the eight steps in this approach. The statistical package SPSS (SPSS Inc., 2006) was used for model-building.

2.4.2.1. Definition of a conceptual model of pedogenesis. To ensure a sound pedological basis of the regression model, a conceptual model of pedogenesis is defined. This is an explicit, structured representation of knowledge of the soil–landscape system of the survey area, based on a review on soil development. The conceptual model identifies the driving factors and processes controlling pedogenesis and soil spatial distribution.

2.4.2.2. Collection of predictors from available environmental ancillary data. In quantitative prediction models, the drivers of pedogenesis are represented or proxied by predictors. The predictors are identified and collected from available environmental ancillary data. The result is a set of predictors, all of pedological importance, that are candidates for the MLR-model.

2.4.2.3. Univariate analysis and selection of candidate predictors. Selection of predictors for an MLR-model from the set of candidates starts with a univariate analysis of each predictor. For categorical predictors this involves cross-tabulation of the response variable versus each predictor followed by the chi-square test of independence. Attention must be paid to contingency tables with zero frequency cells as these may cause numerical instability during parameter estimation, which is marked by extreme model coefficients and associated standard errors (Hosmer and Lemeshow, 1989). The analysis of the contingency tables is followed by the fit of a univariate MLR-model for each predictor that showed at least a moderate level of association with the response variable. Univariate MLR-models are also fit for continuous predictors.

The estimated coefficient and odds ratio of each logit function of the univariate MLR-models should be checked for pedological consistency. Predictors that are significant in the univariate analysis are selected for the next step. Hosmer and Lemeshow (1989) suggest to retain predictors with p-value<0.25. The large p-value used is based on the work of Bendel and Afifi (1977) and Mickey and Greenland (1989) who showed that the 0.05 level often fails to identify predictors known to be important. Predictors that are only weakly correlated with the response variable may become strong predictors when taken together in the multivariate model. When

univariate analysis resulted in a very large set of candidate predictors we selected from each variable group (Table 1) only the predictors with the strongest association to the response variable, as predictors within each group are expected to be strongly associated.

2.4.2.4. Multivariate analysis of selected candidate predictors. A multicollinearity assessment is carried out to identify associated predictors. Next, multivariate MLR-models are fitted, with the aim of selecting one or more competing preliminary models. We used the stepwise forward method for model selection with entry probability 0.20 and removal probability 0.25, as recommended by Lee and Koval (1997). Selected MLR-models must be checked for numerical stability and multicollinearity. Numerical problems can be solved by replacing the predictor with another (associated) predictor that describes the same soil forming process; by grouping the levels of the predictor; by omitting the predictor from the model; or by omitting the outcome class of the response variable that shows numerical instability (this will induce bias in the predictions). Multivariate analysis of candidate predictors might result in several competing MLR-models, especially when some of the selected predictors are associated, as for each of these predictors a separate model can be fitted.

2.4.2.5. Evaluation of adequacy of the multivariate model(s). The fit of the MLR-model(s) is followed by verification of the importance of each included predictor using the Wald statistic (Eq. (7)). When there are competing MLR-models, then verification is done with the best model. Competing MLR-models that are not nested cannot be compared with the likelihood ratio test but are compared with goodness-of-fit measures. Assessing goodness-of-fit of logistic regression models is not as straightforward as for linear regression models, and the appropriateness of the various goodness-of-fit measures for logistic regression models is a subject of debate in the literature (Mittlböck and Schemper, 1996; Hosmer et al., 1997; Menard, 2000). We used three goodness-of-fit measures: Pearson chi-square statistic, classification tables and the McFadden-R^2. The Pearson chi-square statistic indicates how well the model fits the data. Hosmer and Lemeshow (1989) advise caution when using this statistic for models containing continuous predictors. The chi-square distribution then becomes an inadequate approximation of the true distribution of the statistic. Therefore the p-value for this statistic becomes meaningless, although the statistic itself is a good measure of model adequacy (Hosmer et al., 1997): the lower the statistic, the better the model fit. Classification tables were used to derive the calibration purity. The McFadden-R^2 (Menard, 2000) measures the reduction in maximized log-likelihood. It is conceptually and mathematically close to the ordinary least squares R^2.

Once an MLR-model is chosen from the alternatives, the included predictors can be verified. Predictors that are not significant should be deleted from the model one by one, starting with the least significant. A new model is fitted each time a predictor is deleted and compared to the old model with the log-likelihood ratio test. Careful attention should be paid to predictors whose coefficient has changed markedly after another predictor is removed, indicating that the deleted predictor is a confounder of other predictors (Hosmer and Lemeshow, 1989). A strong confounder should be kept in the model, even when the predictor is not significant. Next the odds ratios of the predictors are checked for pedological consistency.

2.4.2.6. Checking the assumption of linearity in the logit. Logistic regression assumes a linear relationship between continuous predictors and the logit. We used the Box–Tidwell approach and logit graphs to test this assumption (Hosmer and Lemeshow, 1989). The Box–Tidwell approach adds the transformed predictor $x \ln(x)$ to the model, where x is the value of the predictor. Statistical significance of this predictor suggests non-linearity in the logit. We also used the logit graph approach, which replaces the continuous predictor with a categorical predictor with four levels using the quartiles as cut-points. The

estimated coefficients of this predictor are plotted against the midpoints of the quartiles. Non-linear plots indicate non-linearity in the logit. The relationship shown by the graph should be pedologically plausible, as before.

2.4.2.7. Checking for interactions between predictors. To check whether interactions between predictors should be included in the MLR-model, pairwise interactions are created for each possible combination of predictors or only for those predictors that the model-builder expects to interact. We used the stepwise forward method to select interactions from all possible combinations of predictors. Interactions are tested for significance with the likelihood ratio test. Significant interactions are included unless these are not pedologically plausible or cause numerical instability during parameter estimation. Goodness-of-fit statistics are used to check if the model fit improved.

2.4.2.8. Statistical and visual assessment of the final model. Statistical assessment of the final MLR-model is based on the goodness-of-fit measures as described in step 5. If the model is judged statistically acceptable then the model is applied to create a preliminary soil map. If unrealistic soil patterns are found the model should be adjusted. This means a return to step 4 of the model-building framework.

2.5. Model application

Ten calibrated MLR-models, one for each map unit, were used to estimate the probabilities of occurrence of the ten major soil groups on a 25 m grid. The soil group with the largest probability was used to construct a prediction map. The theoretical purity was computed as the mean of the maximum probability at each grid cell of the prediction grid (Brus et al., 2008). Prediction uncertainty was quantified by Shannon entropy:

$$H_z = -\sum_{i=1}^{n_y} \hat{\pi}(z_i, \mathbf{s}) \log_{n_z} \hat{\pi}(z_i, \mathbf{s}) \tag{8}$$

where $\hat{\pi}(y_i, \mathbf{s})$ is the estimated probability that random variable Z at location $\mathbf{s}$ takes the value z_i, and n_z is the number of outcomes (Brus et al., 2008). By using the logarithm with base n_z the maximum entropy is 1, which occurs when all outcomes have equal probability. The minimum value for the entropy is 0, which occurs when there is no uncertainty and one of the outcomes has probability 1. It should be noted that the entropy indicates whether the predicted soil group has a large probability, it does not indicate that the prediction itself is correct. The accuracy of the predicted soil groups was validated with an independent data set (Section 2.6).

2.6. Model validation

2.6.1. Sampling strategy

The results were validated with an independent, stratified simple random sample (De Gruijter et al., 2006). Strata were obtained by overlaying the aggregated national soil map, henceforth referred to as the *reference map*, with a map depicting three regions that roughly coincide with the major drainage basins and the areas with 1:10,000 soil maps. The latter map improves the spreading of the sample locations over the study area and facilitates the separate estimation of purity for the subareas with high and low density of calibration data. This resulted in 34 strata. A total of 150 locations were allocated to the strata in proportion to their area, with a minimum of two per stratum to allow estimation of the sampling variance for each stratum. Locations where permission was denied or proved otherwise impossible to sample were replaced with locations from a reserve list.

2.6.2. Statistical inference

Validation resulted in an indicator variable taking value 1 if the mapped soil group equals the observed soil group and 0 else. The

estimated statistical parameter was the spatial mean of the indicator, which corresponds to the fraction of the survey area that is correctly mapped and is known as the *actual purity* (or user's accuracy). The actual purity was also estimated for the ten 'soil strata' (map units of reference map) separately, and for the two 'mapping-scale strata' (1:50,000 and 1:10,000). The actual purity was estimated by (De Gruijter et al., 2006):

$$\hat{f} = \sum_{h=1}^{l} w_h \hat{f_h} \tag{9}$$

where w_h is the weight (relative area) of stratum h, $\hat{f_h}$ is the estimated areal fraction of stratum h correctly classified, and l is the number of strata. The stratum fractions were estimated by the fraction correctly predicted locations in each stratum since the locations in each stratum were selected by simple random sampling:

$$\hat{f_h} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_i \tag{10}$$

where n_h is the number of sampling locations in stratum h, and y_i is the indicator variable at sampling location i. Eqs. (9) and (10) were also used to compare the predictive capabilities of the updated soil map and reference map by substituting y_i for $d_i = y_i^{(u)} - y_i^{(r)}$, the difference between the indicators for the updated $(y_i^{(u)})$ and for the reference map $(y_i^{(r)})$. This variable can have values -1, 0, and 1 and is used to estimate $\hat{d}$, which is the mean difference in actual purity of the updated and reference maps. Under the null hypothesis that the expected value of the estimated mean difference is zero, $\hat{d}$ follows approximately a normal distribution with zero mean and variance $Var(\hat{d})$.

We used group ratio-estimators (De Gruijter et al., 2006) to estimate the actual purity and sensitivity (or producer's accuracy) of the map units on the updated soil map because the strata used in random selection of the validation locations did not coincide with the map units on the updated map. The actual purity of map unit k of the updated soil map was estimated by:

$$\hat{p}^{(k)} = \frac{\sum_{h=1}^{l} A_h \overline{y}_h^{(k)}}{\sum_{h=1}^{l} A_h \overline{x}_h^{(k)}} \tag{11}$$

where A_h is the area of stratum h, $\overline{y}_h^{(k)}$ is the sample mean of indicator $y_{i,h}^{(k)}$ taking value 1 if the mapped and observed soil group at sampling location i equal soil group k and 0 else, and $\overline{x}_h^{(k)}$ is the sample mean of indicator $x_{i,h}^{(k)}$ taking value 1 if the mapped soil group equals soil group k and 0 else. The sensitivity is defined as the fraction of the true area of soil group k that is mapped as soil group k. The sensitivity of map unit k of the updated soil map was estimated by:

$$\hat{s}^{(k)} = \frac{\sum_{h=1}^{l} A_h \overline{y}_h^{(k)}}{\sum_{h=1}^{l} A_h \overline{z}_h^{(k)}} \tag{12}$$

where and $\overline{z}_h^{(k)}$ is the sample average of the indicator $z_{i,h}^{(k)}$ taking value 1 if the observed soil group equals soil group k and 0 else.

3. Results

3.1. Model-building

We describe the model-building process for map unit "thin peat soils with a mineral surface horizon (mPY)" by applying the eight steps described in Section 2.4.2. For the other nine map units we followed a similar approach.

3.1.1. Definition of a conceptual model of pedogenesis

The map unit mPY (36,000 ha) is the second largest of Drenthe. The national soil map subdivides this unit in *iW*, *zW* and *kW*, which have different topsoils due to different soil forming processes. Map unit *kW*, covering 200 ha in the northern tip of Drenthe, has a clayey topsoil that is formed by deposition of marine clay on peat in the brook valleys. The topsoils of units *iW* (22,000 ha) and *zW* (14,000 ha) are formed by anthropogenic processes. The spatial extent of *iW* is limited to the peat-colonial landscape. The topsoil is formed by repetitive mixing of the sand cover, applied after peat excavation, with peat remnants in the subsoil. This reclamation method resulted in topsoils that are spatially highly variable in thickness (15–40 cm) and organic matter content (10–25%). Map unit *zW* is found in small areas within the peat colonies, along the edges of brook valleys or in depressions (pingo remnants) on the Drenthe plateau. The sandy topsoil can be formed by (1) cultivation by application of sand-rich manure, (2) leveling of the irregular surface of the Drenthe plateau during agricultural reclamation, or (3) sand application on the peaty surface to improve trafficability. The *zW* topsoil is spatially less heterogeneous than the *iW* topsoil and its organic matter content is on average 5–15%. All soils in map unit mPY have a sandy subsoil that may contain a podzol, depending on the position in the landscape. A podzol-B horizon in the subsoil is generally found on higher positions.

The frequency distribution of observed soil groups in map unit mPY shows that at only 30% of the locations a thin peat soil with a mineral topsoil was found. The podzol is the most common observed soil group (42%), which is the result of oxidation of the peat layer. Oxidation rate depends on several factors. These include *land use*, because oxidation rate is faster under arable land than under grassland or nature; *groundwater level*, because oxidation rate increases as the groundwater level decreases; and *peat type*, because mesotrophic peat is less resistant to oxidation than oligotrophic peat. Where peat has disappeared, either a "podzol" (PZ) or an "earth soil" (E) is present now. Soils of the peat colonies are better drained and under more intensive agricultural use than peat soils in the brook valleys. We therefore expect soils in the peat colonies to be more strongly affected by oxidation than soils in the brook valleys.

Not all impurities in map unit mPY can be explained by peat oxidation. Part of the inclusions were present from the beginning, due to generalization errors. Confusion of soil groups close to boundaries of map delineations is expected to be larger than in the centre of the delineations due to the positional accuracy of the delineations. We therefore assumed that the probability of occurrence of soil groups within an impure map delineation is also governed by the soil groups of adjacent map delineations and by soil groups that dominate the direct neighbourhood of a location. Generalization errors are also caused by the large short-scale variability of the soils in the peat colonies, which cannot be adequately expressed at the 1:50,000 map scale. Because of this variability "thin peat soils" (PY), "thin peat soil with a mineral topsoil" (mPY), thick peat soils (P) or "thick peat soils with a mineral topsoil" (mP) can all occur in areas smaller than the minimum delineation size. Inclusion of podzols or earth soils can be found at higher and drier positions in the peat colonies and brook valleys, such as coversand ridges. These geomorphological features are in general too small to be mapped at the 1:50,000 scale.

3.1.2. Collection of predictors from available environmental ancillary data

The pedogenic processes and factors that cause inclusions of soil groups other than the mPY soil group were represented by the set predictor variables described below. This set comprises 46 predictors (Table 2).

1. *Land cover*. The effect of land cover on peat oxidation is represented by datasets "recent land cover" and "historic land cover". Five indicator predictors were derived from both datasets.

2. *Groundwater*. The effect of groundwater level on peat oxidation is represented by datasets "GD", "GT", "GD_MHW", and "GD_MLW". Three indicator predictors were derived from each dataset. An ordinal categorical predictor with three levels was derived from datasets GD and GT.

3. *Peat type*. Peat type is proxied by subsoil type as described by the soil map (Finke et al., 1996). If a podzol-B horizon is present in the subsoil then it was assumed that the peat is of oligotrophic origin otherwise it was assumed that the peat is of mesotrophic origin.

4. *Oxidation risk*. Finke et al. (1996) mapped peat oxidation risk (high-low) for two map sheets of the soil map of Drenthe by combining data on groundwater and peat type. We created two such risk predictors, one using groundwater data from the GT data, and one using the groundwater data from the GD data.

5. *Topsoil lithology*. One indicator predictor was derived from the soil map to represent the topsoil type.

6. *Landscape*. Information from the soil and geomorphology maps was combined to delineate the peat-colonial landscape. The paleogeography map was used to delineate the former highmoor landscape and the brook valley system.

7. *Elevation*. Elevation was used to map out inclusions of soil groups PZ and PY.

8. *Relative elevation*. Four relative elevation grids captured local height variation to identify for example local depressions or coversand ridges.

9. *Proximity to boundary of map delineations*. Two indicator maps were generated from the soil map indicating whether a grid cell within a mPY delineation fell into boundary zones with widths 125 and 250 m.

10. *Neighbouring soil group*. The soil group of the nearest neighbouring delineation was determined for each grid cell within map unit mPY. The resulting map was recoded into two categorical maps: one with two and one with three levels.

11. *Dominant soil group*. The dominant soil group within search radii 125, 250, and 500 m was determined for each location within map unit mPY, resulting in three maps. Each map was reclassified into two categorical maps similar to the neighbouring soil group predictors.

3.1.3. Univariate analysis of candidate predictors

Our data set contained 2894 soil profile observations within map unit mPY. Each of the ten soil groups is observed at least once in the map unit. The soil groups "*Brown Forest soils*" and "*Till soils*" are observed only three and two times, respectively. These two soil groups were eliminated as outcome level because there were not enough observations to fit the logit functions. This implies that the probability of occurrence of these soil groups in map unit mPY was set to zero.

Each cross-tabulation of a categorical predictor with the response variable resulted in a significant Pearson chi-square statistic. Furthermore, cross-tabulations showed that response outcome "*Plaggen soil*" (17 observations) had zero cell frequencies for several predictors (Table 2). To reduce the number of candidate predictors we selected those with the strongest association to the response variable from variable groups "groundwater", "recent land cover", "historic land cover" and "soil map". This selection resulted in 19 categorical and 5 continuous predictors (Table 2). A univariate MLR-model was fitted for each of the selected predictors, with soil group mPY as reference level. The likelihood ratio test was significant for each univariate model, indicating that all predictors are candidates for the multivariate model. The odds ratios were generally in accordance with our knowledge on the soil–landscape system.

3.1.4. Multivariate analysis of selected candidate predictors

A multicollinearity assessment confirmed the assumption that predictors within variable groups are associated. Furthermore, moderate and strong associations were found between predictors from different groups (Table 2).

Table 2
Candidate predictors per variable group for MLR modeling for soil map unit mPY.

Variable group	Description	Codes/Values	Predictor name	Associated predictors of other groups
1	Elevation			
	Absolute elevation*	cm a.s.l.	ELEV	
2	Relative elevation			
	Search radius 250 m*	cm	RELELEV250	
	Search radius 500 m*	cm	RELELEV500	
	Search radius 750 m*	cm	RELELEV750	
	Search radius 1000 m*	cm	RELELEV1000	
3	Groundwater			
	GD*[a]	1 = Wet/2 = Moist/3 = Dry	GD	PEATOX_GD
				PEATOX_GT
	GD wet*[a]	1 = Yes/0 = No	GD_W	PEATOX_GD
	GD moist	1 = Yes/0 = No	GD_M	
	GD dry	1 = Yes/0 = No	GD_D	
	GD_MHW wet	1 = Yes/0 = No	MHW_W	
	GD_MHW moist	1 = Yes/0 = No	MHW_M	
	GD_MHW dry	1 = Yes/0 = No	MHW_D	
	GD_MLG wet	1 = Yes/0 = No	MLG_W	
	GD_MLG moist	1 = Yes/0 = No	MLG_M	
	GD_MLG dry	1 = Yes/0 = No	MLG_D	
	GT*	1 = Wet/2 = Moist/3 = Dry	GT	PEATOX_GT
	GT wet*	1 = Yes/0 = No	GT_W	PEATOX_GT
	GT moist	1 = Yes/0 = No	GT_M	
	GT dry	1 = Yes/0 = No	GT_D	
4	Recent land cover, 1997–2003			
	Permanent grassland*	1 = Yes/0 = No	RLC_GR	
	Permanent cropland*	1 = Yes/0 = No	RLC_CR	
	Gras–crop rotation	1 = Yes/0 = No	RLC_GRCR	
	Gras–crop rotation or cropland*	1 = Yes/0 = No	RLC_ROTCR	
	Nature	1 = Yes/0 = No	RLC_NAT	
5	Historic land cover, 1900			
	Grassland*	1 = Yes/0 = No	HLC_GR	
	Cropland	1 = Yes/0 = No	HLC_CR	
	Heath*	1 = Yes/0 = No	HLC_HEATH	
	Forest	1 = Yes/0 = No	HLC_FOR	
	Nature*	1 = Yes/0 = No	HLC_NAT	
6	Paleogeography			
	Brook valley system*[a]	1 = Yes/0 = No	BROOKVAL	PEATTYPE
	Former highmoor areas*	1 = Yes/0 = No	HIGHMOOR	PEATTYPE
7	Geomorphology–soil map			
	Peat-colonial landscape*	1 = Yes/0 = No		SOILCOV
			PEATCOL	GT, GT_W
8	Soil map			
	Peat type*	1 = Oligotrophic	PEATTYPE	BROOKVAL
		0 = Mesotrophic		HIGHMOOR
	Topsoil lithology*	1 = Peat-colonial	SOILCOV	PEATCOL
		0 = Sandy/clayey		
	Distance to boundary delineation			
	<125 m*[a]	1 = Yes/0 = No	DIST125 m	
	<250 m	1 = Yes/0 = No	DIST250 m	
	Nearest neighbouring soil group			
	2 levels	1 = Peat soil	NEIGHB_2L	
		0 = Mineral soil		
	3 levels*[a]	1 = Thick peat soil	NEIGHB_3L	
		2 = Thin peat soil		
		3 = Mineral soil		
	Dominant soil group, 125 m radius			
	2 levels	See nearest neighb. soil group	DOMSOIL125_2L	
	3 levels	See nearest neighb. soil group	DOMSOIL125_3L	
	Dominant soil group, 250 m radius			
	2 levels	See nearest neighb. soil group	DOMSOIL250_2L	
	3 levels	See nearest neighb. soil group	DOMSOIL250_3L	
	Dominant soil group, 500 m radius			
	2 levels	See nearest neighb. soil group	DOMSOIL500_2L	
	3 levels*	See nearest neighb. soil group	DOMSOIL500_3L	
9	Groundwater–soil map			
	Oxidation risk, using GD*[a]	1 = High/0 = Low	PEATOX_GD	GD, GD_W
	Oxidation risk, using GT	1 = High/0 = Low	PEATOX_GT	GD, GT, GT_W

*Predictors selected after univariate analysis.
[a] Outcome soil group PS has a zero cell frequency for one of the levels of the variable.

The first MLR-model estimation started with all 24 predictors. Soil group mPY was used as reference level. All predictors except GD, GT_W and HLC_GR were selected resulting in a model that showed strong multicollinearity effects, evidenced by highly inflated coefficients and standard errors for several predictors. To remove the multicollinearity effects we started with omitting the least significant predictor of variable groups *elevation*, *recent land cover* and *historic land cover* until the two strongest predictors within these groups

Table 3
Competing MLR-models with their goodness-of-fit measures.

	Model 1	Model 2	Model 3	Model 4	Model 4*
Variable group					
1	ELEV	ELEV	ELEV	ELEV	ELEV
2	RELELEV250	**RELELEV1000**	RELELEV250	RELELEV250	RELELEV250
3	GT	GT	GT	GT	GT
4	RLC_ROTCR	RLC_ROTCR	**RLC_GR**	RLC_ROTCR	RLC_ROTCR
5	HLC_HEATH	HLC_HEATH	HLC_HEATH	**HLC_NAT**	**HLC_NAT**
6	BROOKVAL	BROOKVAL	BROOKVAL	BROOKVAL	
7	PEATCOL	PEATCOL	PEATCOL	PEATCOL	PEATCOL
8	PEATTYPE	PEATTYPE	PEATTYPE	PEATTYPE	PEATTYPE
	SOILCOV	SOILCOV	SOILCOV	SOILCOV	SOILCOV
	DIST125 m	DIST125 m	DIST125 m	DIST125 m	
	NEIGHB_3L	NEIGHB_3L	NEIGHB_3L	NEIGHB_3L	
	DOMSOIL500_3L	DOMSOIL500_3L	DOMSOIL500_3L	DOMSOIL500_3L	DOMSOIL500_3L
9	PEATOX_GD	PEATOX_GD	PEATOX_GD	PEATOX_GD	
Goodness-of-fit measure					
Pearson−χ^2 (df)	19,592 (20,034)	19,824 (20,041)	19,636 (20,034)	18,716 (20,062)	24,119 (20,027)
McFadden−R^2	0.13	0.12	0.13	0.13	0.10
Calibration purity	0.484	0.480	0.481	0.484	0.476

Competing predictors are indicated in bold type.

remained. With these predictors we fitted four competing MLR-models (Models 1–4, Table 3). Model 1 is the model after selection of the most significant predictor from each variable group. Models 2 to 4 are competing models in which one of the competing predictors is substituted for the other competing predictor of the same variable group. Because none of these models showed effects of multi-collinearity, we decided to keep the weakly and moderately associated predictors that belong to different variable groups in the model. Four predictors caused numerical instability during the fit of the logit function of outcome level PS. Because eliminating an outcome level is at first less preferable than omitting predictors, we fitted an MLR-model without the predictors that caused numerical instability during the fit of the logit function of map unit PS (Model 4*, Table 3).

3.1.5. Evaluation of adequacy of the multivariate model(s)

Summary measures of goodness-of-fit were calculated for each of the four competing multivariate models (Models 1–4, Table 3). All goodness-of-fit measures are very similar for the four competing models. As Model 4 performed slightly better for Pearson Chi-squared and calibration purity, we chose this model for the next steps in the model-building process. Model 4* performed worse than Model 4 (Table 3). We decided to eliminate PS as outcome level because soil group PS was observed at only 17 locations within map unit mPY. A pedological justification is that soils belonging to the PS soil group are unlikely to occur in map unit mPY as they are characteristic for the open field farming system found on coversand ridges and not for peat reclamation areas. The refitted MLR-model contained the same predictors as Model 4.

The number of significant predictors differed between the six logit functions and showed a clear relationship with the number of observations of each soil group. Only three predictors were significant (at the 0.15 level) in the logit of soil group S (19 observations) whereas ten predictors were significant in the logit of soil group PZ (1234 observations). The number of significant predictors in the logits of the other soil groups varied from seven (PY, 423 observations) to four (P, 60 observations). Predictors SOILCOV, HLC_NAT and NEIGHB_3CL were not significant for five logit functions. Since the likelihood ratio test is significant for SOILCOV and SOILCOV is a pedologically important predictor, we decided to retain this predictor in the model. The likelihood ratio test for HLC_NAT is not significant. Furthermore, HLC_NAT is only a moderately strong confounder of one coefficient in the logit of soil group PZ. We therefore omitted HLC_NAT from the model. NEIGHB_3L contributes significantly to the model and is a strong confounder of other predictors, in spite of five

non-significant model coefficients. Collapsing NEIGHB_3L to two levels improves the Wald statistics: the coefficient of the binary predictor is significant in three logits. However, the likelihood ratio test suggests that the model with the three-level predictor performs better than the model with the binary predictor so we decided to retain NEIGHB_3L in the model. Predictors DOMSOIL500_3L, RLC_ROTCR and BROOKVAL were not significant for four logit functions. Collapsing DOMSOIL_3L into a binary predictor or omitting the predictor did not improve the model. BROOKVAL and RLC_ROTCR were kept in the model for their pedological significance although the likelihood ratio test of BROOKVAL did not confirm its importance. Like the odd ratios of the univariate model, the odd ratios of the multivariate MLR-model generally agree with the conceptual model of pedogenesis.
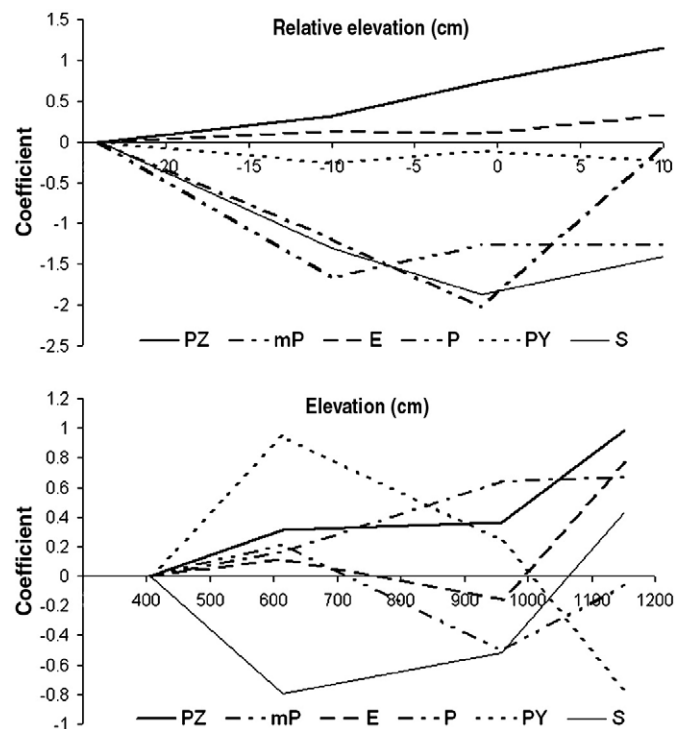

Fig. 3. Logit graphs for predictors RELELEV250 (top) and ELEV (bottom).

Table 4
The estimated model coefficients (Coeff) and odds ratios (OR) of the final MLR-model for map unit mPY.

Predictor	Logit function					
	PZ		mP		E	
	Coeff	OR	Coeff	OR	Coeff	OR
Intercept	−0.63		−3.61*		−4.26*	
NEIGHB_3L=1	−0.22*	0.81	1.12*	3.08	0.08	1.08
NEIGHB_3L=2	−2.04*	0.13	0.12	1.13	0.33	1.38
DOMSOIL500_3L=1	−0.83*	0.44	0.81*	2.24	−1.05*	0.35
DOMSOIL500_3L=2	−0.36*	0.70	0.26	1.30	−0.52*	0.60
DIST_125M=0	−0.24*	0.79	−0.37*	0.69	0.31	1.37
SOILCOV=0	0.12	1.13	−0.40*	0.67	−0.47	0.62
GT=1	−0.02	0.98	−0.88*	0.42	0.87*	2.38
GT=2	0.10	1.11	−0.658	0.52	0.04	1.04
RLC_ROTCR=0	−0.23*	0.80	−0.07	0.93	−0.48*	0.62
PEATOX_GD=0	−0.49*	0.61	0.46*	1.58	−0.15	0.86
PEATTYPE=0	0.12	1.12	0.28	1.32	1.50*	4.49
PEATCOL=0	0.82*	2.27	0.75*	2.12	1.85*	6.39
BROOKVAL=0	0.42*	1.52	0.52	1.69	−0.12	0.88
ELEV	0.001*	1.001			0.002*	1.002
BROOKVAL=0×RELELEV250	0.02*	1.02	−0.01*	0.99	0.02*	1.02
BROOKVAL=1×RELELEV250	0.05*	1.05	−0.03*	0.97	0.02*	1.02
RLC_ROTCR=0×RELELEV250	0.01	1.01	−0.02*	0.98	−0.01	0.99

Predictor	Logit function					
	P		PY		S	
	Coeff	OR	Coeff	OR	Coeff	OR
Intercept	−6.91*		−3.37*		−5.70*	
NEIGHB_3L=1	0.24	1.27	−0.17	0.84	0.25	1.29
NEIGHB_3L=2	−1.02	0.36	−0.53*	0.59	0.68	1.97
DOMSOIL500_3L=1	1.44*	4.20	0.10	1.11	−1.60	0.20
DOMSOIL500_3L=2	0.64	1.89	0.49*	1.63	−0.64	0.53
DIST_125M=0	0.23	1.26	0.14	1.16	−1.88*	0.15
SOILCOV=0	−0.58	0.56	−1.41*	0.24	−1.21	0.30
GT=1	0.93*	2.54	1.80*	6.04	1.25	3.49
GT=2	−0.07	0.93	0.90*	2.45	0.54	1.71
RLC_ROTCR=0	0.60*	1.83	0.00	1.00	0.73	2.08
PEATOX_GD=0	1.02*	2.77	0.36*	1.44	0.48	1.62
PEATTYPE=0	1.85*	6.33	0.58*	1.78	0.32	1.38
PEATCOL=0	−0.10	0.91	0.91*	2.47	1.87*	6.47
BROOKVAL=0	1.23*	3.43	0.81*	2.24	−0.98	0.38
RELELEV250	0.001	1.001	0.001*	1.001	0.002*	1.002
BROOKVAL=0×RELELEV250	−0.03*	0.97	0.01	1.01	0.01	1.01
BROOKVAL=1×RELELEV250	−0.07*	0.93	−0.02	0.98	0.02	1.02
RLC_ROTCR=0×RELELEV250	0.02	1.02	−0.01	0.99	0.01	1.01

*Wald statistic is significant at the 0.15 level.

3.1.6. Checking the assumption of linearity in the logit

So far the continuous predictors ELEV and RELELEV250 were treated as linear in the logit. The coefficients of both Box–Tidwell transformed predictors are not significant for four of six logit functions. The logit graphs for RELELEV250 (Fig. 3) show that this predictor is linear in the logits of soil groups PZ, E and S, is somewhat linear in the logits of mP and PY, and is non-linear in the logit of P. Both logit graph and Box–Tidwell transformation for RELELEV250 suggest that this predictor can be treated as linear in the logit. The logit graph for ELEV

Table 5
Statistical assessment of the ten final MLR-models.

Model	R^2_{MF}	Calibration purity
mPY	0.13	0.49
P	0.31	0.61
mP	0.28	0.62
PY	0.21	0.51
BF	0.19	0.63
PZ	0.21	0.79
E	0.21	0.57
PS	0.30	0.63
T	0.31	0.75
S	0.49	0.83

R^2_{MF} is the McFadden-R^2.

(Fig. 3) shows that this predictor is linear in the logits of outcome levels PZ and P, and non-linear in the logits of mP and E. The logit graphs of PY and S show linearity between the second, third and fourth quartiles. The results of the logit graphs and Box–Tidwell transformation for ELEV do not convincingly support linearity in the logit, nor do they rule it out. As the fit of the model with ELEV as continuous predictor was much better than the fit with ELEV as categorical predictor we decided to keep ELEV in the model as continuous.

3.1.7. Checking for interactions between predictors

Three interactions remained after exclusion of interactions that were not pedologically plausible or that caused numerical instability during coefficient estimation. BROOKVAL*RELELEV250 and RLC_ROTCR*RELELEV250 were statistically significant and pedologically plausible and were added to the main effects model. BROOK-VAL*PEATOX_GD was significant but did not improve the model fit, and was therefore not included. The final MLR-model is presented in Table 4.

3.1.8. Statistical and visual assessment of the final model

The results of the statistical assessment of the final MLR-model for map unit mPY are presented in Table 5. The deviance of the fitted model is 13% (McFadden-R^2) smaller than the intercept-only model. This model predicts the most frequently observed outcome, in this case soil group PZ, at each calibration location. Overall calibration purity shows that the model correctly predicts the soil group at 49% of the calibration locations, which is a 19% increase compared to the reference map. Statistical assessment of the other nine MLR-models shows that the models explain a substantial part of the variation within the soil data set (Table 5). Global calibration purity is 66%, which is an 11% increase compared to soil classification at the calibration locations using the reference soil map. The gain is on average about 20% for the peat map units and 3% for the mineral map units.

The soil map for map unit mPY did not show unexpected patterns of soil groups. The area of soil group mPY was, as expected, greatly reduced. Podzols were predicted at 62% of the map unit area. The MLR-
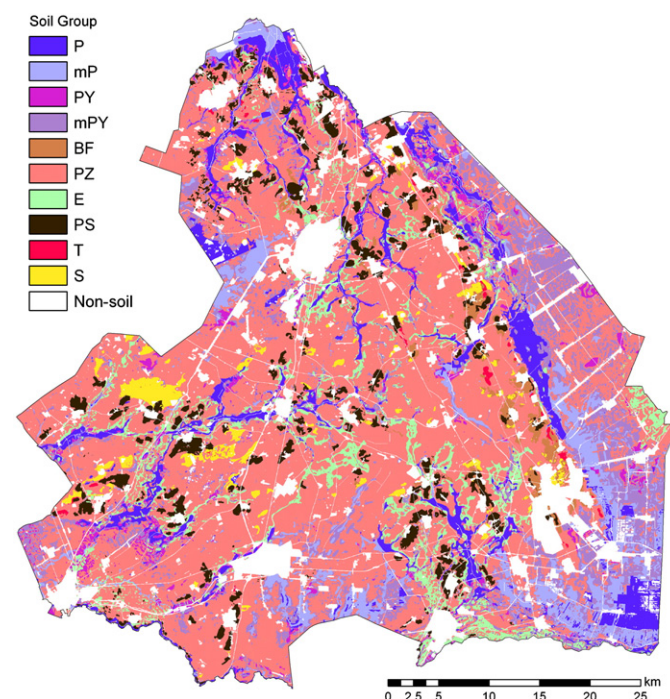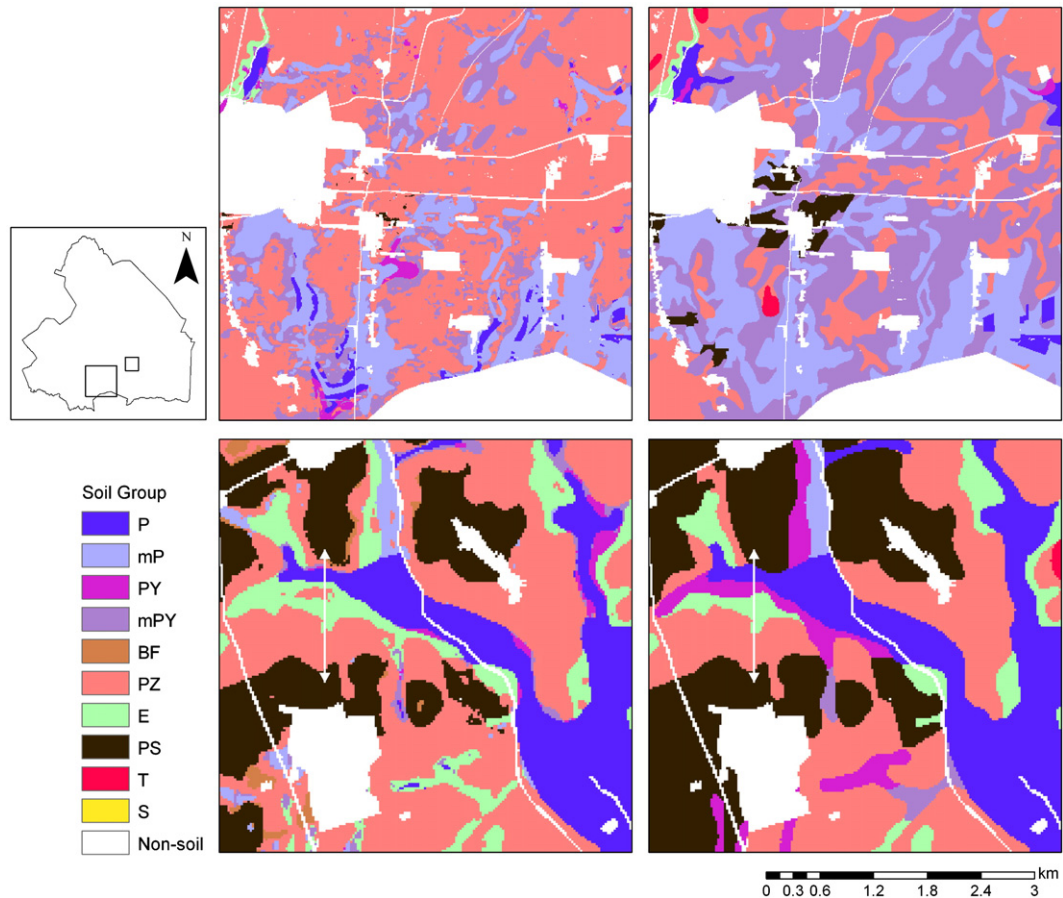

Fig. 4. Updated soil map as predicted with the ten MLR-models.

Fig. 5. Details of the updated and reference soil maps: peat-colony updated map (top left), peat-colony reference map (top-right), brook valley updated map (bottom-left), and brook valley reference map (bottom-right). The arrows in the bottom figures indicate the location of the north–south oriented catena shown in Fig. 7.

models for map units P, E and T were adjusted after inspection of the pattern of the predicted soil groups.

3.2. Model application

Ten MLR-models were used to re-map soil distribution within the ten map units of the reference map (Fig. 4). The general pattern of soil groups on the updated map resembles that of the reference map, although the updated soil group differs from the reference soil group at 31.5% of the area. Changes are most dramatic, as expected, for the peat map units. The area with peat soils declined with 34% (33,525 ha) compared to the reference map. Only 45%, 20% and 30% of the soils mapped as mP, PY and mPY, respectively are predicted as such. Roughly 60% of the soils mapped as thin peat soils are predicted to be transformed to mineral soils: the extent of the podzol soil group increased with almost 40,000 ha. Thirty-six percent of the thick peat soils with a mineral topsoil (mP), typical for the peat-colonial landscape, are predicted to be transformed to thin peat soils with mineral topsoil. Fig. 5 clearly shows these changes. Changes within map unit P are less severe: only 22% is predicted to be transformed to thin peat soils. The reason for this is that soil group P primarily occurs in the brook valleys where peat layers are thicker and where conditions for oxidation are less favorable compared to the peat colonies. The strong decline of soil group T can be explained by the fact that most till soils occur in association with podzol soils on the reference map. The majority of the observations used to calibrate the model for map unit T are classified as PZ, which results in PZ as the dominant predicted soil group in map unit T. The area with plaggen soils, PS, is reduced with 32% compared to the reference soil map. Affected areas are the edges of PS map delineations (Fig. 5) and the

plaggen soils on the Hondsrug. The former is explained by the decrease in thickness of the plaggen A-horizon from the centre of the open fields towards the edges: if the thickness does not exceed 30 cm, then the soil is not classified as plaggen soil. The latter is a direct result of the observations on the Hondsrug, which were all located in a relative small area (intensively surveyed area 3, Fig. 2). Many profile observations within map unit PS were classified as brown forest soils (BF). This is also the reason for the strong increase in area of map unit BF on the updated soil map compared to the reference map.

Table 6
Theoretical purity and entropy of the updated soil map for the areas corresponding to the map units of the reference map, i.e. the areas for which the MLR-models were calibrated and applied, and for pooled strata peat (P–mP–PY–mPY) and mineral (BF–PZ–E–PS–T–S).

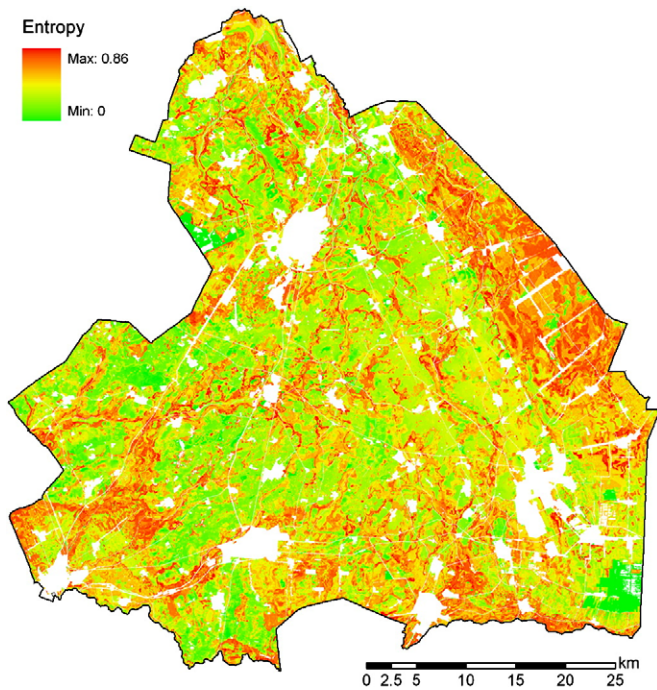|	Theoretical purity	Entropy
Global	0.67	0.40
Stratum		
P	0.64	0.42
mP	0.63	0.43
PY	0.50	0.54
mPY	0.50	0.54
BF	0.83	0.21
PZ	0.79	0.34
E	0.58	0.47
PS	0.66	0.38
T	0.79	0.24
S	0.77	0.24
Peat	0.57	0.48
Mineral	0.75	0.35

Fig. 6. The prediction uncertainty, quantified with Shannon entropy.

The overall theoretical purity of the updated soil map is 67% (Table 6). In general the theoretical purity is smaller and the uncertainty is larger for the areas mapped as peat soils on the reference map than for the mineral soils. The areas with the smallest theoretical purity are the areas that were originally mapped as thin peat soils (PY, mPY) and earth soils (E). Predictions for these areas are also the most uncertain (Table 6). Map unit mPY is characteristic for the peat-colonial landscape, whereas PY and E are mainly found along brook valley sides. Fig. 6 shows high entropy values for these two parts of the

landscape. The soil group pattern in the peat-colonial areas is very heterogeneous by itself and is further complicated by peat oxidation. This makes soil spatial prediction challenging in this area, which is evidenced by highly uncertain predictions. The brook valley sides are topographical transition zones where gradual changes in soil weaken relationships between soil groups and predictors. Prediction uncertainty will be larger in such areas than in areas with stronger relationships such as in the centre of the brook valleys and the high parts of the plateau.

Fig. 7 depicts a catena of predicted and mapped soil groups along a 1500 m long transect from plateau through a brook valley to plateau, as well as the change of estimated probabilities and entropy of the main predicted soil groups. The location of the catena is indicated by the arrow in Fig. 5. The typical catena in Drenthe has plaggen soils on top of the plateau, bordered by podzols or brown forest soils depending on parent material. When going from plateau to brook valley one would typically encounter a gradual transition from podzols to earth soils to thin peat soils to thick peat soils. The first difference between predicted and mapped soil groups is the soil sequence from the northern plateau towards the brook valley. In the reference map plaggen soils border thick peat soils whereas the updated map shows a pedologically more realistic transition from plaggen soils to podzols to thin peat soils to thick peat soils. The second difference is the prediction of podzols on a coversand undulation in the earth soil map unit at the southern side of the brook valley centre. These undulations are better drained than the surrounding, lower terrain, creating more favourable conditions for podzol formation. Earth soils are predicted at the sides of the undulation and podzols at the top, which is pedologically plausible. The thin peat soil mapped at the southern side of the brook valley is predicted to be oxidized. Earth soils are predicted at these locations. We did not validate the predicted soil sequence along the transect but based on our knowledge of the soil–landscape system the updated map shows a more realistic soil sequence along the transect than the reference map. The probability graph shows that the MLR-models do not have much difficulty in differentiating soil groups at the
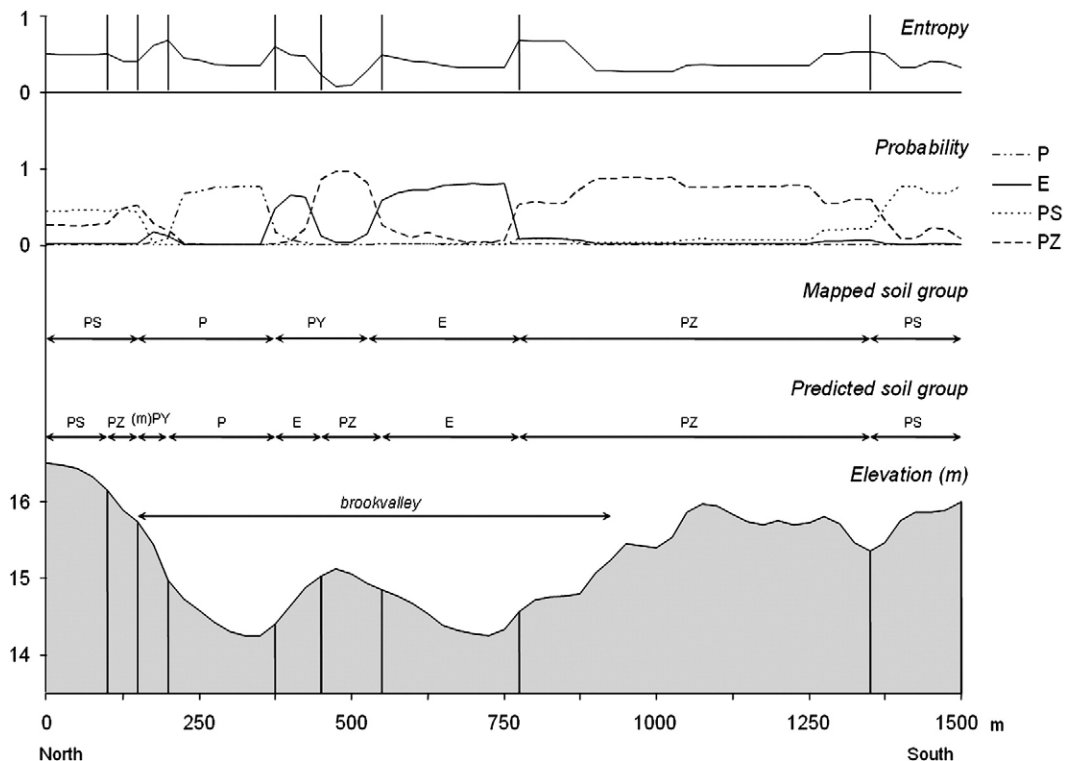

Fig. 7. Change of the entropy, maximum probability, mapped and predicted soil groups along a typical catena in the Drenthe landscape.

topographical extremes: the lowest parts of the brook valleys, the plateaus and the top of the coversand undulation. The difference between largest and second largest estimated probability is relatively large. Small differences in estimated probabilities are found at topographical transition zones. At these locations the models easily confuse between soil groups, which is evidenced by an increase in entropy at these zones.

3.3. Model validation

Table 7 summarizes the validation results of the updated and reference soil map. The estimated actual purity of the updated map is 58%, which is 6% larger ($P = 0.039$) than the purity of the reference map. The purity is 9% smaller than the theoretical purity, possibly because the calibration locations are concentrated in the four areas with a detailed soil map (Fig. 2). Apparently, the calibrated relationships were unable to explain a similar amount of variation outside the four detailed survey area than within the four areas.

At the level of the soil-strata soil spatial distribution is better represented by the MLR-model than by the reference map for soil-strata mP, PY, mPY, PS and T. The largest increase is for stratum PY (35%, $P = 0.000$). Purity gains for strata mP (5.4%, $P = 0.318$) and mPY (14%, $P = 0.224$) are not significant due to the small numbers of validation locations, but they are pedologically relevant, especially the 14% purity gain for stratum mPY. The cause of the large purity increase of soil-stratum T is outlined in Section 3.2. Purity gain of stratum PS is 10% ($P = 0.174$). There is no difference in actual purity between the updated and reference maps for soil-strata P, PZ and S. The reference map better represents soil spatial distribution within strata E and BF. The actual purity of both strata is 13% smaller ($P = 0.105$ for E; $P = 0.159$ for BF) for the updated soil map than for the reference soil map. These figures indicate that the global increase in map purity of the updated map compared to the reference map is largely attributed to the increase in purity in the peat strata. The pooled purity increase for these strata is a pedologically relevant 11% ($P = 0.069$) whereas the pooled purity increase for the mineral strata is 2.4% ($P = 0.086$).

Table 8 shows the actual purity of the ten map units of the updated soil map. The MLR-models predict the spatial distribution of soil groups P, BF, PZ, PS and S fairly well, while map units PY and T have purities close to 0. The large variation in purities can have several reasons. Firstly, the effect of peat oxidation is underestimated: mineral soils were observed at three validation locations in map unit PY and at five validation locations in map unit mPY; thin peat soils were

Table 7
Estimated actual purities of the two soil maps: global purity, soil-strata purity, grouped soil-strata (peat–mineral) purity and purity for the two mapping-scale strata.

	n	Updated map	Reference map
Global	150	0.58 (0.04)	0.52 (0.04)
Soil stratum			
P	15	0.45 (0.14)	0.45 (0.14)
mP	15	0.31 (0.11)	0.26 (0.08)
PY	9	0.40 (0.17)	0.05 (0.05)
mPY	22	0.50 (0.11)	0.36 (0.11)
BF	4	0.13 (0.13)	0.26 (0.00)
PZ	55	0.73 (0.06)	0.73 (0.06)
E	10	0.42 (0.14)	0.55 (0.17)
PS	11	0.75 (0.15)	0.65 (0.10)
T	4	0.98 (0.03)	0.00 (0.00)
S	5	0.64 (0.32)	0.64 (0.32)
Peat	61	0.43 (0.07)	0.32 (0.06)
Mineral	89	0.70 (0.05)	0.67 (0.05)
Mapping-scale stratum			
Area with 1:10,000 map	26	0.72 (0.12)	
Area without 1:10,0000 map	124	0.56 (0.04)	

The number in brackets is the estimated standard error and n is the number of validation samples used to estimate the actual purity.

Table 8
The estimated actual purity and sensitivity of the ten map units of the updated soil map.

Map unit	n	Actual purity	Sensitivity
P	9	0.77	0.52
mP	10	0.28	0.56
PY	8	0.05	0.06
mPY	13	0.37	0.25
BF	5	0.71	0.40
PZ	82	0.67	0.90
E	9	0.46	0.21
PS	8	0.80	0.47
T[a]	2	0.00	–
S	4	0.94	0.52

[a] The sensitivity of map unit T could not be computed because soil group T was not observed in the validation sample.

observed at four validation locations in map unit mP. Secondly, the models have difficulty in predicting topsoil lithology of peat soils, which is to a large extent influenced by human activities and is not easily associated to environmental predictors. In map unit PY, soil group mPY was observed at three validation locations and soil group P was observed at three validation locations in map unit mP. Thirdly, soil groups in topographical transition zones are easily confounded as indicated by Fig. 7. In map unit E, typical for such transition zones, earth soils, podzols and peat soils were observed. Fourthly, sample size. The small number of validation locations in several map units results in highly uncertain purity estimates. An example is the purity of map unit T that might be attributed to chance as only two validation locations were located in this map unit.

It is interesting to note that if we would aggregate the updated soil map to two map units (peat and mineral) and then validate the map, the purity of the peat map unit would increase to 80% while the pooled purity of the four separate peat map units is 42%. This indicates that the main confusion between predicted soil groups within the predicted peat map units is between the four peat soil groups. Thus at the locations where the MLR-models predict peat soils, it is very likely that a peat layer is present in the soil profile at these locations. However, the model is uncertain about the thickness of the peat layer and the topsoil lithology (peaty or mineral). The purity of the mineral map unit is 88%.

4. Discussion

4.1. Multinomial logistic regression for soil mapping

MLR is computationally a simple method compared to more demanding methods such as indicator kriging (IK) (Bierkens and Burrough, 1993) and Bayesian Maximum Entropy (BME) (Brus et al., 2008). It does not suffer from shortcomings of IK like probabilities that are outside the interval [0,1] or probabilities that do not sum to 1. Nor is it as computationally demanding as BME. However, building a statistically and pedologically sound MLR-model requires careful attention as many choices have to be made, and interpretation (both statistical and pedological) of the MLR-model is not as straightforward as that of linear regression models. The framework based on Hosmer and Lemeshow (1989) proved a valuable guideline for building MLR-models, although the steps should be meticulously applied.

The main drawback of applying MLR to spatial data is that spatial autocorrelation during coefficient estimation and prediction is ignored. This may bias estimated effects of the predictors on the response variable. Hengl et al. (2007a) showed that MLR did not perform as well as methods that incorporate spatial autocorrelation in soil spatial prediction. Autologistic regression can account for spatial autocorrelation in the response variable and is a popular method in spatial ecology (Smith, 1994; Augustin et al., 1996). Unfortunately, the autologistic regression model can only handle binomially distributed data. There are no examples known to us that extend the autologistic

model to the multinomial case. An alternative to account for spatial autocorrelation in MLR is to integrate MLR with BME. The estimated probability distributions at the nodes of the prediction grid can be used in the BME framework as one-point bivariate probabilities that constrain the computation of the multi-point probability density function instead of using spatially invariant one-point bivariate probabilities. This might result in predictions with independent errors but is mathematically less elegant because the coefficients of the MLR-model are still estimated as if the observations were independent, like in regression-kriging (Hengl et al., 2007b).

Certain structures in the calibration data, such as predictors that completely separate outcome levels or predictors that lack observations for one or more of its levels, cause non-existent maximum likelihood estimates or infinite odds ratio estimates (Hosmer and Lemeshow, 1989). This is a serious drawback of MLR as it limits the number of outcome levels and the number of predictor levels. Even with only ten soil groups we were often forced to omit outcome levels of the response variable during model-building. Furthermore, predictors that cause complete separation or that have zero cell counts are in theory strong predictors as they suggest that certain soil groups do not occur under certain conditions. Several candidate predictors that were highly correlated with the response variable could not be used as predictors because of complete separation or zero cell counts.

4.2. Soil spatial prediction

Like Hengl et al. (2007a) we find that the most frequently observed soil groups within a map unit are overrepresented on the predicted soil maps (e.g. soil groups mPY and PZ in map unit mPY). Hengl et al. (2007a) argues that this is caused by weak association of the predictors with some of the soil groups. However, the odds ratios of the predictors of the logits of PY and E show that the several predictors are strongly related to soil groups PY and E (Table 4). There is no evidence that weak associations with less frequently observed soil groups cause overrepresentation of the most frequently observed soil groups. The predictors strongly influence the estimated probabilities of less frequently observed soil groups, but apparently this is not enough to exceed the estimated probabilities of the most frequently observed soil groups. This is caused by strong influence of the marginal probability distribution on the conditional probability distribution.

We can only hint at the reason why the clustered distribution of the calibration locations causes the 9% discrepancy between actual and theoretical purities. The four survey areas can be regarded as reference areas used to obtain predictive relationships. These relationships are then extrapolated across the entire survey area. This resembles the mapping approaches presented by Lagacherie et al. (1995), Bui and Moran (2003) and Grinand et al. (2008). Lagacherie et al. (1995) states that the reference area approach works when similar soil forming processes act in the two areas, creating similar soil patterns. If soil forming processes were similar for the areas with and without a detailed soil map, as we would expect, then we would expect similar purities for these areas. But validation indicates that the actual purities of the two mapping-scale strata differ (Table 7). The estimated actual purity of the updated map for the areas with detailed soil maps is 16% larger than for the area without the detailed maps. Note, however, that this estimate is very uncertain given the large standard errors of the purities. This might indicate that the modelled relationships are not so easily extrapolated across the province. If we assume that the natural soil forming processes within and outside the detailed survey areas are similar, then the difference between global theoretical and actual purities and the difference between actual purities within and outside the detailed soil survey areas might be attributed to human influence on soil formation. Since peat soils are much more sensitive to human interventions in the landscape than mineral soils, we would expect the discrepancy between theoretical

and actual purity to be larger for peat soils than for mineral soils. This is supported by the figures in Tables 6 and 7. For the pooled peat soils this discrepancy is 14% and for the mineral soils 5%.

The influence of human activities on soil distribution might also explain the 27% difference in actual purity of the peat soil strata compared to mineral soil strata (Table 7). Human influence can sometimes be proxied by predictors, e.g. human influence on peat oxidation can be proxied by land cover and groundwater. However, in many cases such influence is hard to proxy with biophysical data. For example, the decision to apply a sand cover on a peaty topsoil is made on the scale of individual agricultural fields. This leads to occurrence of soil group mP in map unit P or mPY in PY. Furthermore, discerning between soil groups P and mP or PY and mPY within the peat-colonial landscape is almost impossible to relate to environmental predictors as the organic matter content of the topsoil exhibits large short-distance variation as a result of the reclamation method used. We likely captured some of the local soil variation resulting from human activities with our environmental predictors and extrapolated this across the province where the relationships might not be valid. The actual purity of the peat soil stratum of the updated soil map would increase to 56.6% and global actual purity would increase to 64.1% if we would pool map units P–mP and PY–mPY, meaning that we would only discriminate between thick and thin peat soils and not by topsoil lithology.

4.3. Legacy soil data

The utility of legacy soil data for updating soil maps in landscapes with highly dynamic soils such as peat soils, greatly depends on the age of legacy data. Roughly 60% of the calibration locations located in the four detailed survey areas were already 12 years old at the time this study was carried out. The other 40% was between 3 and 6 years old. We are aware that 12 year-old observations on peat soils are also becoming outdated as peat oxidation continued since the time the soil profile was recorded and classified. But omitting these data from the calibration dataset would have greatly reduced the number of locations for calibrating the models for the peat map units. Furthermore, the soil survey for the reference soil map was conducted at least 15 years before the detailed soil surveys, which contained almost all calibration observations. This means that these observations are still useful for updating. However, using decade-old observations on peat soils for map updating might overestimate the area of (thin) peat soils on the predicted soil map, although this is not strongly supported by the validation. This can also contribute to the difference between theoretical and actual purity.

4.4. Validation of soil maps

Accuracy assessment of soil maps is imperative for any soil mapping study; traditional or digital. For thematic (soil) maps, commonly used statistics to quantify map accuracy are the purity and the kappa index (Hengl et al., 2007a; Li and Zhang, 2007; Grinand et al., 2008). Another statistic for accuracy assessment that provides valuable information is the sensitivity, or producer's accuracy, of the map. This statistic is often used in image classification studies (Foody, 2002) but is hardly reported for (digital) soil maps. Sensitivity values of the map units of the updated soil map are presented in Table 8. An example of the merit of the sensitivity statistic is given by map unit P. This map unit has a high purity (77%), which tells the user that soil group P is found at 77% of the area predicted as soil group P. However, the sensitivity of map unit P is 52%, meaning that only 52% of the true area of soil group P is mapped as P. One can question the usefulness of a soil map, for example for a carbon stock inventory, if the area with peat soils is heavily underestimated, which is not indicated by the purity. If we aggregate the updated soil map to two map units (peat and mineral), the sensitivity of the peat map unit is

72% and that of the mineral map unit 92%. This indicates that the updated soil map distinguishes peat soils from mineral soils fairly well.

We used pedological knowledge during the model-building phase. Such knowledge could also be exploited for error analysis or validation of the model. One could argue if predicting a site which is P as mP (two different thick peat soils) is equally wrong as predicting it as PZ (a mineral soil) and vice versa. Whether all errors are equal or different, clearly depends on the application for which a soil type map is used. Predicting a site which is P as mP would not have a large impact on estimates of carbon stock but it would have for the organic matter content of the topsoil; while predicting a site which is mP as PZ would have a large effect on carbon stock estimates but not on the topsoil organic matter content. We did not use pedological knowledge for computation of the validation statistics or for the evaluation of the classification tables, but it would be worthwhile to investigate this in future mapping studies.

5. Conclusions

Legacy soil data in combination with high-resolution environmental ancillary data can be used to update a soil map. Although an independent validation showed that updating proved to have more effect for the peat map units than for the mineral map units.

The presented framework provides a systematic approach for building MLR-models but it can also be easily used for linear regression models. Furthermore it allows integration of expert knowledge during model selection and evaluation.

When calibration locations are clustered in small areas within the survey area such as in this study, the transferability of the calibrated relationships might be limited, especially when there is strong human influence on soil development. This leads to overestimation of the theoretical map purity, and emphasizes the importance of validation of soil maps, preferably with an independent probability sample, which gives an unbiased estimate of the map purity (De Gruijter et al., 2006). Sensitivity is a useful accuracy measure complementary to the purity, for thematic soil maps and is easily computed from a validation sample.

References

Augustin, N.H., Mugglestone, M.A., Buckland, S.T., 1996. An autologistic model for the spatial distribution of wildlife. Journal of Applied Ecology 33, 339–347.

Bailey, N., Clements, T., Lee, J.T., Thompson, S., 2003. Modelling soil series data to facilitate targeted habitat restoration: a polytomous logistic regression approach. Journal of Environmental Management 67, 395–407.

Bendel, R.B., Afifi, A.A., 1977. Comparison of stopping rules in forward stepwise regression. Journal of the American Statistical Association 72, 46–53.

Bierkens, M.F.P., Burrough, P.A., 1993. The indicator approach to categorical soil data. I. Theory. European Journal of Soil Science 44, 361–368.

Brus, D.J., De Gruijter, J.J., Breeuwsma, 1992. Strategies for updating soil survey information: a case study to estimate phosphate sorption characteristics. European Journal of Soil Science 43, 567–581.

Brus, D.J., Bogaert, P., Heuvelink, G.B.M., 2008. Bayesian Maximum Entropy prediction of soil categories using a traditional soil map as soft information. European Journal of Soil Science 59, 166–177.

Bui, E.N., Moran, C.J., 2003. A strategy to fill gaps in soil survey over large spatial extents: an example from the Murray–Darling Basin of Australia. Geoderma 111, 21–44.

Campling, P., Gobin, A., Feyen, J., 2002. Logistic modeling to spatially predict the probability of soil drainage classes. Soil Science Society of America Journal 66, 1390–1401.

Cook, S.E., Corner, R.J., Grealish, G., Gessler, P.E., Chartres, C.J., 1996. A rule-based system to map soil properties. Soil Science Society of America Journal 60, 1893–1900.

De Bakker, H., Schelling, J., 1989. In: Brus, D.J., van Wallenburg, C. (Eds.), System of Soil Classification for the Netherlands: The Higher Levels. Second, Adjusted Edition. Winand Staring Centre, Wageningen.

Debella-Gilo, M., Etzelmüller, B., 2009. Spatial prediction of soil classes using digital terrain analysis and multinomial logistic regression modeling integrated in GIS: examples from Vestfold County, Norway. CATENA 77, 8–18.

De Gruijter, J.J., Brus, D.J., Bierkens, M.F.P., Knotters, M., 2006. Sampling for Natural Resource Monitoring. Springer, Berlin.

De Vries, F., Brouwer, F., 2006. De bodem van Drenthe in beeld. Alterra-rapport 1381. Alterra, Wageningen.

Finke, P.A., 2000. Updating the (1:50,000) Dutch groundwater table class map by statistical methods: an analysis of quality versus cost. Geoderma 97 (3–4), 329–350.

Finke, P.A., Groot Obbink, D.J., Rosing, H., De Vries, F., 1996. Actualisatie Gt-kaarten 1: 50 000 Drents deel kaartbladen 16 Oost (Steenwijk) en 17 West (Emmen). Rapport 439. DLO-Staring Centrum, Wageningen.

Finke, P.A., Brus, D.J., Bierkens, M.F.P., Hoogland, T., Knotters, M., De Vries, F., 2004. Mapping groundwater dynamics using multiple sources of exhaustive high resolution data. Geoderma 123, 23–39.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment 80, 185–201.

Grinand, C., Arrouays, D., Laroche, B., Martin, M.P., 2008. Extrapolating regional soil landscapes from an existing soil map: sampling intensity, validation procedures, and integration of spatial context. Geoderma 143, 180–190.

Hengl, T., Toomanian, N., Reuter, H.I., Malakouti, M.J., 2007a. Methods to interpolate soil categorical variables from profile observations: lessons from Iran. Geoderma 140, 417–427.

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007b. About regression-kriging: from equations to case studies. Computers & Geosciences 33, 1301–1315.

Heuvelink, G.B.M., Webster, R., 2001. Modelling soil variation: past, present, and future. Geoderma 100, 269–301.

Hosmer, D.W., Lemeshow, S., 1989. Applied Logistic Regression. Wiley, New York.

Hosmer, D.W., Hosmer, T., Le Cessie, S., Lemeshow, S., 1997. A comparison of goodness-of-fit tests for the logistic regression model. Statistics in Medicine 16, 965–980.

Knol, W.C., Kramer, H., Gijsbertse, H., 2004. Historisch Grondgebruik Nederland: een landelijke reconstructie van het grondgebruik rond 1900. Alterra-rapport 573. Alterra, Wageningen.

Koomen, A.J.M., Maas, G.J., 2004. Geomorfologische Kaart Nederland (GKN); Achtergronddocument bij het landsdekkende digitale bestand. Alterra-rapport 1039. Alterra, Wageningen.

Lagacherie, P., Legros, J.P., Burfough, P.A., 1995. A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. Geoderma 65, 283–301.

Lee, K.I., Koval, J.J., 1997. Determination of the best significance level in forward stepwise logistic regression. Communications in Statistics Part B: Simulation and Computation 26, 559–575.

Li, W., Zhang, C., 2007. A random-path Markov chain algorithm for simulating categorical soil variables from random point samples. Soil Science Society of America Journal 71, 656–668.

May, R., Van Dijk, J., Wabakken, P., Swenson, J.E., Linnell, J.D.C., Zimmermann, B., Odden, J., Pedersen, H.C., Andersen, R., Landa, A., 2008. Habitat differentiation within the large-carnivore community of Norway's multiple-use landscapes. Journal of Applied Ecology 45, 1382–1391.

McBratney, A.B., Mendonca Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52.

McKenzie, N.J., Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma 89, 67–94.

McKenzie, N.J., Gallant, J., 2007. Digital soil mapping with improved environmental predictors and models of pedogenesis. In: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), Digital Soil Mapping: An introductory perspective. Elsevier, Amsterdam, pp. 327–349.

Meersmans, J., De Ridder, F., Canters, F., De Baets, S., Van Molle, M., 2008. A multiple regression approach to assess the spatial distribution of Soil Organic Carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 143, 1–13.

Menard, S., 2000. Coefficients of determination for multiple logistic regression analysis. The American Statistician 54, 17–24.

Mickey, R.M., Greenland, S., 1989. The impact of confounder selection criteria on effect estimation. American Journal of Epidemiology 129, 125–137.

Mittlböck, M., Schemper, M., 1996. Explained variation for logistic regression. Statistics in Medicine 15, 1987–1997.

Müller, D., Zeller, M., 2002. Land use dynamics in the central highlands of Vietnam: a spatial model combining village survey data with satellite imagery interpretation. Agricultural Economics 27, 333–354.

Pape, J.C., 1970. Plaggen soils in the Netherlands. Geoderma 4, 229–255.

Rhemtulla, J.M., Mladenoff, D.J., Clayton, M.K., 2007. Regional land-cover conversion in the U.S. upper Midwest: magnitude of change and limited recovery (1850–1935–1993). Landscape Ecology 22, 57–75.

Rosing, H., Thijssen, G.L. and Brouwer, F., 2006. Actualisatie en modernisering van de Bodemkaart van Nederland, schaal 1:50 000; Een test in de omgeving van Helmond. Alterra-rapport 1057, Alterra, Wageningen.

Schulp, C.J.E., Veldkamp, A., 2008. Long-term landscape–land use interactions as explaining factor for soil organic matter variability in Dutch agricultural landscapes. Geoderma 146, 457–465.

Smith, P.A., 1994. Autocorrelation in logistic regression modelling of species' distributions. Global Ecology & Biogeography Letters 4, 47–61.

Spek, T., 2004. Het Drentse esdorpenlandschap: een historisch-geografische studie. PhD thesis, Wageningen University, Wageningen, The Netherlands.

SPSS Inc., 2006. SPSS 15.0 for Windows, Chicago, USA.

Steur, G.G.L. and Heijink, W. (Editors), 1991. Bodemkaart van Nederland, schaal 1:50.000; Algemene begrippen en indelingen (4e editie). Staring Centrum, Wageningen.

Suring, L.H., Goldstein, M.I., Howell, S.M., Nations, C.S., 2008. Response of the cover of berry-producing species to ecological factors on the Kenai Peninsula, Alaska, USA. Canadian Journal of Forest Research 38, 1244–1259.

Thompson, J.A., Kolka, R.K., 2005. Soil carbon storage estimation in a forested watershed using quantitative soil–landscape modeling. Soil Science Society of America Journal 69, 1086–1093.

Van Kekem, A.J., Hoogland, T., Van der Horst, J.B.F., 2005. Uitspoelingsgevoelige gronden op de kaart; werkwijze en resultaten. Alterra-rapport 1080. Alterra, Wageningen.

Walter, C, Lagacherie, P., Follain, S., 2007. Integrating pedological knowledge into soil digital mapping. In: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), Digital Soil Mapping: An introductory perspective. Elsevier, Amsterdam, pp. 281–300.

Zhu, A.X., Hudson, B., Burt, J., Lubich, K., Simonson, D., 2001. Soil mapping using GIS, expert knowledge, and fuzzy logic. Soil Science Society of America Journal 65, 1463–1472.