LUMA-GIS Thesis nr 25

# Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity?

**Yann Buhot**

2013
Department of
Physical Geography and Ecosystem Science
Centre for Geographical Information Systems
Lund University
Sölvegatan 1
S-223 62 Lund
Sweden

Yann Buhot (2013). Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape scenario and the impact on habitat diversity? Master degree thesis, 30/ credits in Master in Geographical Information Sciences Department of Physical Geography and Ecosystems Science, Lund University

# Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitats diversity?

---

Yann Buhot

Master thesis, 30 credits, in Physical Geography and Quaternary Geology, Stockholm University, Master Thesis Course GE9009.

Supervisor:

Sara Cousin
Professor in Landscape Ecology
Dept of Physical Geography & Quaternary Geology
Stockholm University

**Acknowledgments**

I am very grateful to my supervisor Sara Cousins for all appreciated advice and guidance in all aspects of this study. Thanks to the Ekoklim for providing all the needed data. Thanks to Lars-Ove Westerberg and Harry Lankreijer that made possible for me to carry out my master thesis with Stockholm University. Thanks to Lund University and the LUMA GIS programme for taking me into this incredible master programme. Also thanks to Helena Fridman-Konstantinidou who supported me not only for this thesis but also during my entire education.

**Abstract**

Biodiversity, and in particular plant species diversity decline is linked to landscape fragmentation and habitats loss due to land use transformations but also to decreasing dispersal vectors in the landscape. During the last 200 years the landscape has undergone dramatic changes in its structure and its habitats diversity. Among the different habitats that constitute rural landscape there are three key habitats that are important for plant species biodiversity: Semi-natural grasslands, deciduous forests and wetlands.

The aim of this study is first to create a model that predict the landscape as it was in 1800 using land cover data from 1900 and variables such as soil properties, distance to settlements, distance to water and slope percentage using logistic regression and GIS technology. Then land cover changes are also tested with soil properties, distance to settlement, distance to water and slope percentage. Thus it is possible to have a better understanding of the factor that influence changes from one land cover to another. The results show that between 1800 and 1900, a diversity of habitats has become lost, especially semi-natural grasslands (semi-natural grasslands on fine soil) and wetlands (fen and wooded mire). The small amount of data and the high diversity of habitats made difficult the construction of an accurate model for 1800. In 1800 it was likely too find any type of habitats on any type of soil or terrain conditions. The important finding of this study is that before 1850 there were a high diversity of semi-natural grasslands compare with today as well as a high diversity of wetlands. The model for 1900 has a very good accuracy and could be used for modelling the landscape in a larger extent than the study area.

Such a model can give more opportunities for studying landscape history and therefore get more data for studying the fragmentation of the landscape, which play an important role in the maintenance of a high plant species diversity. Conservationists could use this model for decision-making regarding the restoration of habitat.

## Table of content

## Introduction

Biodiversity, and in particular plant species diversity decline is linked to landscape fragmentation and habitat loss due to land use transformations (Sala, 2000; Foley et al., 2005; Lindborg, 2007) but also to decreasing dispersal vectors in the landscape (Cousins et al., 2007; Ozinga et al., 2009). Thus understanding the processes in the rural landscape that have an effect on biodiversity has become very important for geographers, ecologists and conservationists (Cousins, 2009). Unfortunately our knowledge of the land use distribution at the beginning of the 20$^{th}$ century and before is weak. Yet previous research stressed the fact that during the last 150 years the rural landscape has undergone many changes that have affected both its structure and its plant species diversity (Cousins, 2009). In this study the focus was on three different types of habitats: Semi-natural grasslands, deciduous forest and wetlands.

Semi-natural grasslands are grasslands that have never been treated with fertilizer or ploughed. These types of habitats are considered as one of the most important habitat in Europe regarding many aspect of the biodiversity (Berg et al., 2011; Wilson et al., 2012). Today semi-natural grassland degradation is a threat for plant species diversity with a direct impact on bird, insect and fungi populations (Critchley, 2004). In fact, 90% of the semi-natural grasslands have disappeared for the last 200 years in Europe (Stoate et al., 2001). Today in Sweden, semi-natural grasslands are located within a matrix of forests, arable lands and urban areas (Antonson and Jansson, 2011). As an example, in 1927 10% of the total landscape in Sweden was semi-natural grasslands but only 1.1% today. Moreover semi-natural grassland habitat can contain up to 40 to 60 species per square meter (Eriksson et al., 1995; Wilson et al., 2012). Different variables such as: soil, bedrock or distances to towns have influenced the change trajectories in the landscape during the last 200 years (Cousins, 2009; Antonson and Jansson, 2011).

Deciduous forests may harbour many semi-natural grassland plant species as well as many forest specie and is often the climax vegetation in many parts of the world. Thus abandoned non-forested land use will be spontaneously converted to forest. Therefore there is a mixture of forest patches of different ages and size in the landscape (Honnay et al., 2005). In Europe deforestation has begun from the Neolithic period. Then it was a succession of deforestation and afforestation periods (Hermy and Verheyen, 2007). For detailed forest cover data, only the data from the 18$^{th}$ century onwards are reliable. Thus it is difficult to map primary forest also called ancient forest. Ancient forests are characterised their continuity in time and space with a diversity in canopy strucuture and presence of dead trees (Hermy et al., 1999). Deciduous forests are important for regional economies providing fibres and wood. They offer ecological services such as: water supply, stream water quality, regulating soil erosion and biodiversity. In a context of climate change they play an important role in the carbon cycle. Deciduous forests are important for other human values such as: recreation, spirituality and aesthetic (Reich and Frelich, 2002). In this study deciduous forest were considered as wooded semi-natural grassland until 1900 because forest were grazed by animals at this time.

Wetlands are important habitat for many bird, insects, animals and plant species. They are usually classified in different categories on the basis of their hydrology, geography and flora. One of the most common classification is: *aquatic wetland* for wetland that are continuously under water, *marsh* for wetland that are dominated by herbaceous species but not located on peat soil, *swamp* for wooded wetland not located on peat soil, *fen* or *non wooded mire* for wetland dominated by herbaceous and shrub species located on peat soil and *wooded mire* that are located on peat soil. This classification show the diversity of wetlands and it is the classification used for modern paleo-wetlands (Keddy, 2010). In general all those habitats have many functions that make them essential in the landscape. They provide habitats and food for many species, support ground water recharge, play an important role in the dispersion of seeds (Ozinga et al., 2009), prevent water from pollution by regulating the nutrient cycle and they are important for water quality by filtering sediments and metals (Keddy, 2010). The specificity of wetlands allow the maintenance of a high water table and prevent flooding in contiguous (Greb and DiMichele, 2006). In the past farmers also used wetlands for haymaking and fishing. Until the mid twentieth century habitat loss was the main threat for wetland. Today the major threat is the loss of connectivity between the different habitats and the habitat degradation. These losses have a direct impact on the diversity of specialised wetland animals and plants (Gimmi et al., 2011).

There is a difference how species react to landscape or habitat change, short-lived plant species have a positive correlation with the present-day landscape connectivity and area while long-lived plant species are more correlated to past landscape patterns (Lindborg, 2007). The response of plant species when habitats are converted, destroyed or abandoned can be very long. For instance plant species can survived for decades in deciduous forest that were former grazed semi-natural grassland (Antonson and Jansson, 2011).

With this knowledge that many plant species response slowly to land use change it can be said that an extinction debt could occur which mean that in a future plant species will go extinct when the extinction threshold is reached (Kuussaari et al., 2009; Krauss et al., 2010). Thus it can be assumed that delayed responses to changes mean that current species patterns can be a legacy of past land use, therefore it is important not only to know where changes occurred but when and how they have changed (Cousins and Eriksson, 2002). Regarding the extinction threshold in Swedish grassland it has been proposed by Cousins (2009) that there is a decline of species number when there is around 10% grassland left.

Finally, as the understanding of past land use is important for understanding the present and future patterns in the landscape, two aspects have to be taken into account for conservation and restoration of habitats diversity. Indeed, we have to pay attention to remnant habitats and also the impact of cattle grazing and mowing on plant species richness since it was the main activity in the traditional Swedish landscape 200 years ago. Although, previous studies shows that traditional landscape is important to explain plant species patterns in rural landscapes today. These studies are based on small landscape sections of historical maps from the 17$^{th}$ and 18$^{th}$ century with an incomplete coverage of the whole landscape. The old maps do not cover the whole landscape, particularly not grazed forest, but only the villages that choose the land division. Thus, many of the important landscape changes occurred between 100 and 150 years ago where there is no coverage of historical landscape except as more or less point data. To understand the present processes it is important to create a model on how the landscape looked like 200 years ago, prior to the major changes.

Previous studies have used logistic regression to aid in understanding the driving mechanisms of land use change among them (Aspinall, 2004). Aspinall analysed the relationship between land use changes and a variety of environmental and socio-economic conditions. This relationship was tested at different time periods between 1860 and 2000 using logistic regression model and Geographical Information System (GIS). Millington et al., (2007) did one of the first studies with multinomial regression to analyse land use changes. This study examined the performance of multinomial regression models for considering land use changes caused by a combination of environmental and socio-economic conditions.

This study will analyse rural landscape change trajectories within a transect of 1700 km$^2$ that range from the Baltic coast to lake Mälaren from the 18$^{th}$ century to now. According to the data from 2003 the landscape today is mainly composed of arable fields and forest (figure 1).

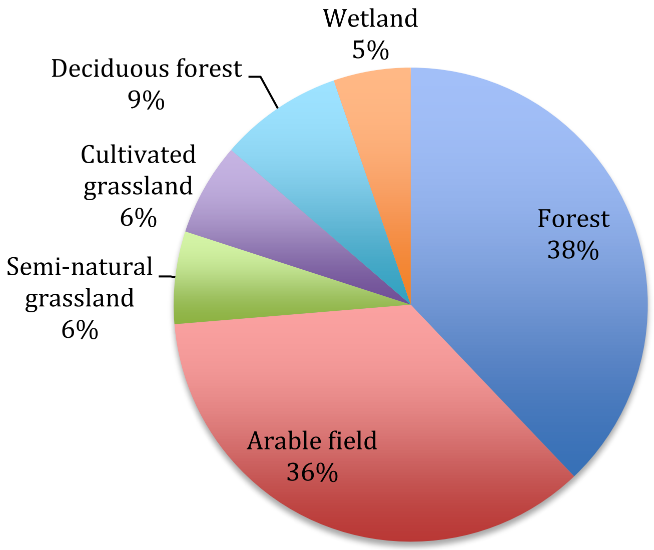


Figure 1: The average repartition of land cover in 2003 based on 11 small landscape sections within the study area in South-Eastern Sweden. Semi-natural grasslands are grasslands that have never been ploughed or fertilised. Cultivated grasslands are grassland that are fertilised and sometimes ploughed.

Cousins (2009) carried out a similar study using old cadastral maps however she only used the sample of 12 small landscapes in the analysis, which makes the generality weak. The 12 landscapes can be considered as cookie cutters of the historical landscape. Nevertheless, in the study an interesting correlation between soil properties and grassland change was found. The aims of this study are to: (i) create a model that can predict the land cover (predict for semi-natural grasslands, arable field and wetland) for 1800 in a larger extent than the cookie cutters by using old historical maps (cookie cutters), soil properties, and geographic features; then (ii) to test different parameters that explain land cover changes in the past. The parameters tested are: soil type, distance to the closest settlement, distance to the closest water source and the slope percentage.

## Methodology

### *Study area*

The study was carried out in the boreonemoral zone in a transect of 1700 km$^2$ (midpoint 59°00'N, 17°11'E) in the county of Södermanland in South-Eastern Sweden (figure 2). Transect land cover vector layer data for the period 1900 was provided by the EkoKlim project. The mean temperature for January is -3°C, and 18°C for July, with a mean annual precipitation of 600 mm (Cousins, 2009). The topography ranges from the Baltic Sea level to higher land with an altitude of maximum 85m. Forest, arable land and lakes constitute the main part of the landscape. The Södermanland's landscape has a long tradition of livestock grazing and haymaking. The landscape is constituted of fine soil such as clay in basins and valleys and coarser soils on slopes and hills. Today the land use is mainly forestry, agriculture and cultivated grasslands with some patches of semi-natural grassland, wetland and deciduous forests.

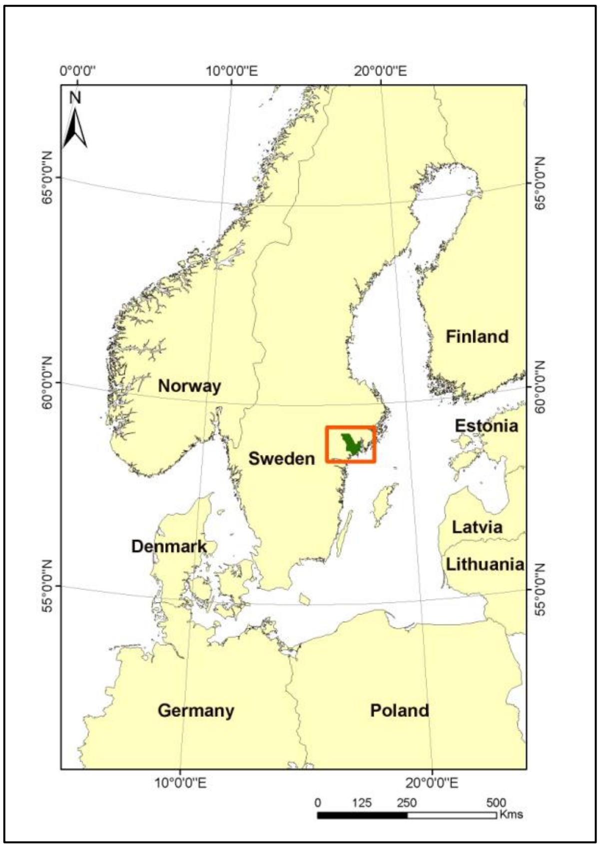


Figure 2: The study area is located in Sweden in the county of Södermanland (midpoint 59°00'N, 17°11'E). The red square shows where the study area is located while the green inside shows the extent of the study area.

Within the study area I have 11 small landscape sections, hereafter called samples, in a GIS from a previous study (Cousins, 2009). The average area of the 11 samples is 1.24 $km^2$ with a maximum of 1.85 $km^2$ and a minimum of 0.44 $km^2$. The 11 samples represent the landscape at different time periods: 1850, 1900, 1950 and 2003.

*Preparation of data*

All the different layers were prepared within a Geographical Information System (GIS) using the ESRI software ArcGIS. Land use in 1800, 1900, 1950 and 2003 as well as changes between 1800-1900 and 1800-1950 were tested with 4 different parameters (table 2).

In the GIS, each landscape vector layer for each period was rasterised with a 10 m resolution. For each time period the landscape was divided in different land use classes with a code for each different class (table 1). Even if forest did exist in the 1800 and 1900 period, forest was classified as wooded semi-natural grasslands, as farmers used the forest as grazing area for livestock up until the 1930's (Antonson and Jansson, 2011).

Table 1: Summary of the different land cover classes used in this study for each time period. Each class has it own land use code. Forest stands for all the forestry industry and arable field is all the agriculture fields used for harvesting. Open SG means Open semi-natural grassland and included all semi-natural grassland with less than 25% trees. Wooded SG means Open semi-natural grassland and included all semi-natural grassland with more than 25% trees. Culti grassland is cultivated grasslands. The symbol -- means the class did not exist in the data layer.

| Land use code | 1800 | 1900 | 1950 | 2003 |
|---|---|---|---|---|
| 1 | -- | -- | Forest | Forest |
| 2 | Arable field | Arable field | Arable field | Arable field |
| 3 | Open SG | Open SG | Open SG | Open SG |
| 4 | Wooded-SG | Wooded SG | Wooded SG | Wooded SG |
| 5 | -- | Wetland | Wetland | Wetland |
| 6 | -- | -- | Deciduous forest | Deciduous forest |
| 7 | -- | -- | Culti grassland | Culti grassland |

Four different 10 m resolution raster were created for testing the four parameters: soil properties, slope, distance to water and distance to settlement.

The soil properties raster was created from the soil vector layer. The soil raster was composed of 6 categories with a value for each categories ranging from 1 to 6: (1) bedrock, (2) rock and stone, (3) till and sand, (4) fine sand, (5) clay and (6) peat. It was important for the classification to be aware of that in Sweden the soil is measured at a 50 cm depth, therefore a 48 cm of clay on top of bedrock will be classified as bedrock.

The slope raster was created using a Digital Elevation Model (DEM) available from Lantmäteriet. Initially, the DEM had a resolution of 2 m but due to the space it took in computer's memory it was transformed to a coarser 10 m resolution. The slope percentage was calculated using spatial analyst package in ArcGIS.

The distance to settlement raster layers was created for each time period. By extracting the settlement polygons from the land use vector layer, it was possible to use them as a source to carry out a Euclidian distance analysis using the spatial analyst package in ArcGis. For the 1800's period only cookie cutters were available, hence 1900's settlement polygons were used as a source making the assumption that the settlement did not change too much between the two periods. The distance to water raster layers went through the same process as the distance to settlement. However only the 1800's and 1900's period were covered. Because only cookie cutters were available for 1950 and 2003 it was not possible to use the same 1900 layers for those period. In Sweden 95% of rivers and wetland have been drained during the last century. The assumption that the water areas have not changed much between 1800 and 1900 was made.

Table 2: Data layers used in the Geographical Information System for this study. (1) Soil type classification of the study area with the following classes: Bedrock, stone, till and sand, fine sand, clay, peat. (2) Each cell of the raster gives the distance to the nearest settlement. (3) Each cell of the raster gives the slope percentage. (4) Each cell of the raster gives the distance to the nearest source of water.

| **Variable** | **Unit of measurement** | **Year of data** | **Source** |
|---|---|---|---|
| Land use 1800 | Land cover | $17^{th}$ and $18^{th}$ century | Derived from historical maps |
| Land use 1900 | Land cover | 1900 | Derived from historical maps |
| Land use 1950 | Land cover | 1950 | Derived from Aerial Photo |
| Land use 2003 | Land cover | 2003 | Derived from infrared aerial photos |
| Soil Type (1) | Soil cover | | Geological Survey of Sweden |
| House distance 1800 (2) | Meters | | Derived from 1800 land use map |
| House distance 1900 | Meters | | Derived from 1900 land use map |
| House distance 1950 | Meters | | Derived from 1950 land use map |
| House distance 2003 | Meters | | Derived from 2003 land use map |
| Slope (3) | Slope percentage | | Derived from 2m resolution DEM from Lantmäteriet |
| Water distance (4) | Meters | | Derived from the 1900 land use map |

To test the relationship between land cover and the parameters shown in table 2 it was necessary to create a layer with sample points. A random raster was created over the area available from 1800. The raster was divided in between 15 different random values and one value was taken and used as sample points. The sample point raster contained 8381 cell points and it was used for all the periods. The values for each parameter were extracted from the different layer at each cell point using the sample extraction tool in ArcGIS.

*Statistical analysis*

The aim of this study was to create a model that can predict the land cover for 1800 in a larger extent than the cookie cutters. The relationship between land use type and soil, slope percentage, distance to water and distance to settlement was tested.

Logistic regression analysis was used because it was the equation needed to analyse the relationship between dependant categorical values and both independent categorical and continuous values. The result of the logistic regression equation (1) was the probability that the dependent variable occurs given the independent variable tested. In this case the dependent variable was land use and the independent variable are soil, slope percentage, distance to settlement and distance to settlement.

$$\pi(x) = \frac{e^{\alpha+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_i x_i}}{1+e^{\alpha+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_i x_i}} \quad (1)$$

Where $\pi(x)$ was the probability that the dependent variable equals 1, $\alpha$ was the equation constant also called the intercept and $\beta_i$ was the coefficient of the independent variable $x_i$.

Before starting with the analysis, the continuous data such as distance to settlement water and slope percentage were log transformed because they did not a normal distribution.

The analysis of the relationship between land cover and each parameter gave the coefficient of probability, the Wald test (W) as well as its significance. Thus, all the independent variables that are not statistically significant are discarded. Once the significant variables are found they were tested all together to get the best model possible as in equation (1).

In the logistic regression summary table the goodness of fit ($R^2$) is missing. The goodness of fit was used in the summary table of a One Way ANOVA analysis. It gave the proportion of the deviance in the model that has been explained by the independent variable tested. However it was possible to calculate a goodness of fit called $D^2$ (Rossiter and Loza, 2004) that was equivalent to $R^2$ (equation (2)).

$$D^2 = 1 - \frac{Model\ Deviance}{Null\ Deviance} \quad (2)$$

The best model's intercept and coefficient are used in raster calculator to produce the predicted models. In the raster calculator the following equation must be written:

$$\alpha + \beta \times [soil] + \beta \times [dist\_to\_sett] + \beta \times [dist\_to\_water] + \beta \times [slope] \quad (3)$$

Where α was the constant of the equation also called the intercept and β was the coefficient of the independent variable. The analysis gave a coefficient of each soil type. Therefore only the significant soil types are used to create the land cover model.

After running all the equations in raster calculator, it gave one probability model layer for each type of land cover tested: arable field, open semi-natural grassland and wooded semi-natural grassland. However the target in this study was to get a land cover model for the whole landscape. This was possible taking out the highest probability between the three probability layers. This operation can be carried out in ArcMap with the *highest position* tool available in the Arctoolbox.

After getting the predicted models, the accuracy had to be tested to know whether the results are accurate enough to be used. To do so the result model was compared with an accurate map. In this case the accurate models were the cookie cutters for 1800 and the whole landscape for 1900. To avoid circularity the sample points used for deriving the coefficient were taken away from the accuracy analyses. Another set of 8263 random sample points was created. The coefficient of agreement also called Kappa coefficient (Millington et al., 2007) was used to assess the accuracy of the output model and was calculated as:

$$\hat{x} = \frac{nd - q}{n^2 - q} \quad (4)$$

Where $\hat{x}$ was the Kappa coefficient, $n$ was the total number of tested cells, $d$ the sum of correctly mapped cells and $q$ was the sum of the product of the number of cell mapped for each land cover class in the accurate model and the predicted model. For example if we have 2 classes A and B in the accurate model and 2 classes A' and B' in the predicted model with x the number of cell in A, y the number of cell in B, x' the number of cell in A' and y' the number of cell in B' we have q = xx' + qq'. The Kappa coefficient gets rid of the parameter of chance and assess if a model was better or worse than chance. The result of the kappa coefficient was a percentage of accuracy.

The accuracy for each class is also tested. The equation for this coefficient was:

$$\hat{x} = \frac{nd_i - q_i}{nb_i - q_i} \quad (5)$$

Where $\hat{x}$ was the individual kappa coefficient, n was the total number of tested cells d the sum of correctly mapped cells for the class of interest, q was the product of the number of cells mapped in this class for the accurate model and the tested model and b was the number of cell mapped for one class in the accurate model.

The accuracy for the 1800's model has been tested comparing the results with the 1900's model. Thus the predicted model for 1900 could be compared with the map from 1900 that cover the whole study area using the Kappa coefficient.

Land cover change trajectories between 1800 and 1900 could be estimated by using data on soil, distance to settlement, distance to water or slope. The different changes were tested with the same for parameters. Semi-natural grasslands were divided into 3 different classes. Open semi-natural grassland, wooded semi-natural grassland and semi-natural grassland. The class semi-natural grassland includes both open and wooded semi-natural grassland. Thus when further in the thesis semi-natural grassland was mentioned it meant both open and wooded semi-natural grassland. The different changes tested were explained in figure 3.

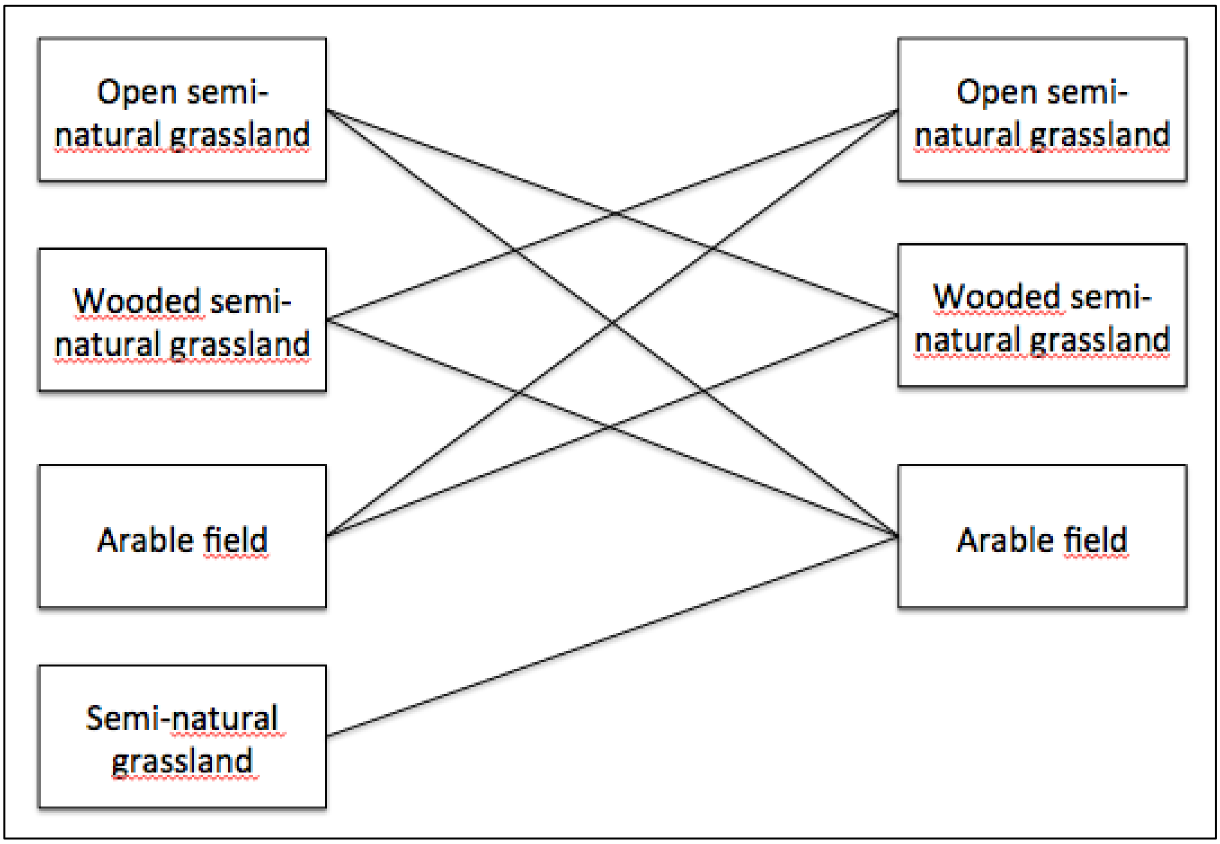


Figure 3: This figure explains the different changes from one habitat to another that have been tested in the analysis. The case semi-natural grassland is gathering together open semi-natural grassland and wooded grassland.

Then the process was the same as the one for the other predicted model. However, both changes between the predicted model and the accurate 1900 model and changes between the accurate 1800 model and the accurate model for 1900 were tested to see if there was a significant difference in the results.

In the last part of the study I tested the 1950 and 2003's land cover to see how they were driven by the variables tested before.

The preparation of the data was done within the software produced by ESRI called ArcMap 9.3 and the tools available in the ArcTools box. All statistical indicators were calculated using the open source programme R 2.15.1. The regression coefficients were used of each parameter within the raster calculator tool of ArcMap.

## Results

In the results part I will highlight the most important findings. The most relevant results are for the periods 1800 and 1900 that is the reason why no results for 1950 and 2003 figure in the result part. However the results for these two periods are available in the appendix 1.

*Arable field*

In 1800 the best model for arable field is based on the addition of soil properties, water distance and settlement distance. For this period the model explained 18.6% of the occurrence of arable fields while for the 1900 period it explained 46.3% of the occurrence of arable fields. Thus in 1800 it was most likely to find arable fields on any type of soil closed to both water area and settlement while in 1900 it was more likely to find arable fields on fine soil, far from water, close to settlement and flat terrain.

In 1800 the most important variable for explaining the arable field were the distance to settlement that explained 11% of the occurrence of arable fields and soil type that explained 8% of the occurrence of arable field. In 1900 the most important variable for explaining the arable fields were the soil type that explained 33.4% of the occurrence of arable fields and slope that explained 17.3% of the occurrence of arable fields (table 3).

*Semi-natural grassland*

The best model for 1800 based on all the parameters explained 21.6% of the occurrence of wooded semi-natural grassland while the best model for 1900 explained 38.3% of the occurrence of wooded semi-natural grasslands. In 1800, wooded semi-natural grasslands were more likely to be found on bedrock with steep slope or peat soil, and far from water and settlement and it was still the same for 1900. The relationship between wooded semi-natural grassland and peat soil ($W= -8.7; p < 0.001$) was negative in 1900, which can be explained by the fact that in 1900 the land cover class wetland is used. In 2003, according the 11 landscape sections 67% of the wooded semi-natural grasslands that were considered as forest in 1800 are now classified as forest. However it is not obvious that broad leaves trees still remain in this type of forest. Today only 8.2% are classified as deciduous forest (table 3).

The best model for 1800 for open semi-natural grassland used only a combination of two parameters: soil type and distance to settlement. The model explained 7.8% of the occurrence of semi-natural grassland. It shows that semi-natural grasslands were more likely to be found on any type of soil, far from settlements. In 1900 the best model used all the parameters and explained 7.2% of the occurrence of semi-natural grasslands. The model shows that open semi-natural grasslands were more likely to be found on steep slope, and far for settlement.

For semi-natural grassland as a whole, which means that the model included both open semi-natural grassland and wooded grassland, in 1800 15.6% of the occurrence of semi-natural grassland is explained. In contrast, for the 1900 period the model explained 45.1% of the occurrence of semi-natural grasslands. The two models showed that it was more likely to find semi-natural grasslands on any type of soil in 1800 while in 1900 it appeared clearly that semi-natural grasslands were more located on coarse soil (table 3).

*Wetland*

There was no result for the 1800's period because wetland was not in the classification. However it can be assumed that it was classified as semi-natural grassland or wooded semi-natural grasslands. Thus the location of wetlands in 1800 is probably the same as in 1900.

The best model for wetland in 1900 used the combination of soil type and slope. The model shows that wetland were more likely to be found on peat soil. The model explained 48.9% of the occurrence of wetland. In 1900 the relationship between soil alone and wetland was significant with bedrock ($W= -13.9; p < 0.001$), till ($W= 3; p < 0.001$) and peat ($W= 11.1; p < 0.001$). Soil type alone explained 40.5% of the occurrence of wetland in 1900 (table 3).

Table 3: $D^2$ value for the relationship between the different habitats (arable field, open semi-natural grassland, wooded semi-natural grassland, semi-natural grassland and wetland) and the four variables tested: soil type, slope percentage, distance to water and distance to settlements. The column "All" for soil type gives the $D^2$ value for soil type as a whole and for each soil type it gives the wald test results as well as its significance. For the other variables, the first line is the $D^2$ value and the second line the wald test and its significance. Finally the Full model column gives the $D^2$ value for the full model that includes all the variables. To make it easier the $D^2$ is in bold style.

| | Soil type | | | | | | | Slope | Water | Settlement | Full model |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Arable** | Bedrock | Stone | Fine sand | Till and sand | Clay | Peat | All | | | | |
| 1800 | w= -31.4; p<0.001 | w= 3.3; p<0.001 | w= 5.3; p<0.001 | w= -3.4; p<0.001 | w= 16.3; p<0.001 | -- | **8** | -- | **0.19**<br>w= 3.5; p<0.001 | **11**<br>w= -36.7; p<0.001 | **18.6** |
| 1900 | w= -32.6; p<0.001 | -- | w= 4.2; p<0.001 | w= -1.8; p<0.001 | w= 43.5; p<0.001 | -- | **33.4** | **17.3**<br>w= -36.7; p<0.001 | **2.7**<br>w=16.6; p<0.001 | **6**<br>w= -24.2; p<0.001 | **46.3** |
| **Open semi-natural grassland** | | | | | | | | | | | |
| 1800 | w= -26.2; p<0.001 | w= 3.5; p<0.001 | w= 2.2; p<0.05 | w= 4.3; p<0.001 | w= 20.3; p<0.001 | -- | **5.2** | **4.6**<br>w= -19.7; p<0.001 | **0.3**<br>w= -6; p<0.05 | **0.25**<br>w= -4.9; p<0.001 | **7.8** |
| 1900 | w= -32.9; p<0.001 | w= 3.9; p<0.001 | w= 3; p<0.01 | w= 5.7; p<0.001 | w= -7.2; p<0.001 | -- | **3.8** | -- | **1.4**<br>w= -8.4; p<0.001 | **1.4**<br>w= -8.6; p<0.001 | **7.2** |
| **Wooded semi-natural grassland** | | | | | | | | | | | |
| 1800 | w= 19.3; p<0.001 | w= -4.6; p<0.001 | w= -4.1; p<0.001 | -- | w= -31.2; p<0.001 | w= 2; p<0.05 | **12.9** | **6.6**<br>w= 24.2; p<0.001 | **0.2**<br>w= 4.1; p<0.001 | **6.1**<br>w=23.3; p<0.001 | **21.6** |
| 1900 | w= 26.8; p<0.001 | w= -3.9; p<0.001 | w=-5.2; p<0.001 | w= -3.6; p<0.001 | w= -43.5; p<0.001 | w= -8.7; p<0.001 | **25.1** | 16.36<br>w= 36.8; p<0.001 | **1.3**<br>w= -12.1; p<0.001 | **8.6**<br>w= 28; p<0.001 | **38.3** |
| **Semi-natural grassland** | | | | | | | | | | | |
| 1800 | w= 31.6; p<0.001 | w= -2.6; p<0.01 | w= -4.8; p<0.001 | w= 4.2; p<0.001 | w= -14.8; p<0.001 | -- | **6.6** | -- | **0.07**<br>w= -2.2; p<0.001 | **9.6**<br>w= 23.9; p<0.001 | **15.6** |
| 1900 | w=32.5; p<0.001 | -- | w= -4.2; p<0.001 | -- | w= -43.5; p<0.001 | w= -9.5; p<0.001 | **32.2** | **17.8**<br>w= 37.3; p<0.001 | **2.6**<br>w= -16.7; p<0.001 | **5.8**<br>w= 23.9; p<0.001 | **45.1** |
| **Wetland** | | | | | | | | | | | |
| 1900 | w= -11.6; p<0.001 | -- | -- | w= 3.3; p<0.001 | -- | w= 9.7; p<0.001 | **46.4** | -- | -- | **0.8**<br>w= 2.3; p<0.05 | **46.3** |

*Changes between 1800 and 1900*

The best model for changes from open semi-natural grassland to arable field used all the parameters. The model showed that those changes mostly occurred on area classified as bedrock and clay with flat terrain far from water and close to settlements. This model explained 39.1% of the changes from open semi-natural grasslands to arable fields. It is the same results for changes from wooded semi-natural grassland to arable field. This model explained 45.1% of changes from wooded semi-natural grasslands to arable fields. By comparison open semi-natural grasslands change to wooded semi-natural grasslands on coarse hilly terrain close to water and far from settlement. This model explained 30.4% of changes from open semi-natural grasslands to wooded semi-natural grasslands (table 4) .

The relationship between land use changes and soil properties is the most important in all cases. Both open and wooded semi-natural grassland had a high probability to be converted into arable field when they were located on clay, peat, or fine sand for wooded grassland (i.e. semi-natural grasslands on fine soils are more likely to be converted to arable field than semi-natural grasslands on coarser soils). Slope has also a significant relationship with changes in landscape. All classes of semi-natural grassland located on flat terrain had a high probability to be converted to arable field. On the contrary, arable field located on steep terrain had a high probability to be converted into grassland, which can be explained by the fact that it is more difficult to cultivate. The accuracy for the different change models is ranging from 11.3% to 19.5%.

Table 4: $D^2$ value for the relationship between land cover changes and the four variables tested: soil type, slope, distance to water and distance to settlement. The last column gives the $D^2$ value for the full model that combine all the variable. The $D^2$ value is the goodness of fit of the model, which is the proportion of the deviance in the model that has been explained by the independent variable tested. Abbreviations used in the table: OSNG = Open semi-natural grassland, WSNG = Wooded semi-natural grassland, SNG = Semi-natural grassland, Ara = Arable field, "—" not used in the full model.

| Land cover changes between 1800 and 1900 | | | | | |
|---|---|---|---|---|---|
| | Soil type | Slope | Water | Settlement | Full model |
| OSNG -> Ara | 27.7 | 15 | 5.5 | 0.2 | 39.1 |
| OSNG -> Ara | 21.1 | 14.3 | 1.6 | 1.4 | 30.3 |
| WSNG -> Ara | 29.1 | 15.5 | 9 | 4.1 | 45.1 |
| SNG -> Ara | 32.6 | 11.3 | 3.4 | 3.11 | 44.7 |
| Ara -> OSNG | 12.7 | 19.5 | -- | 7.8 | 30.4 |
| Ara -> WSNG | 16.4 | 17.2 | -- | 2.8 | 31.8 |

*Visual models of landscape structure*

<u>1800's model</u>

Two different models based on the variable were derived from the regression coefficients. The first model included three land cover classes (figure 4): arable field, open semi-natural grassland and wooded semi-natural grassland. The second model included 2 land cover classes: arable field and semi-natural grassland. The overall kappa for the three classes model is 25.8% better than a random model, which is considered as moderate accuracy while the value for the 2 classes model is 30.2% better than a random model, which is also considered moderate accuracy. In both cases arable fields are the land cover that is better explain by the model with a individual coefficient of 33% compared to 27% and 22% for open semi-natural grassland and wooded semi-natural grassland respectively.

<u>1900's model derived with the land cover coefficients</u>

Two models based on the variable were derived from the regression coefficient. The first model had four classes and the second model had three classes. Both models had the same classes as the 1800's model but for this period the class wetland is added (figure 5). The coefficient of agreement for both 4 and 3 classes model are 53% and 54% respectively better than a random model. This is considered has a good accuracy. Looking at the individual coefficient the best one is for arable fields with a value of 76% considered as very good and wooded semi-natural grassland and wetland had a good coefficient with 49% and 56% respectively. By contrast the coefficient for open semi-natural grassland is not significant at all with a value of 0.03% (table 5).

<u>1900's model derived with the land cover changes coefficient (between 1800 and 1900)</u>

Two models based on the variables were derived using the land cover changes coefficient. The three land-cover classes 1800's model was used for deriving the 1900's model. However the wetland class was derived using the land cover coefficient since wetlands were not classified in the 1800' model (figure 5). The overall coefficient for the four classes model is 55% considered as good. The coefficient for 3 classes is slightly bigger 55.3%. Looking at the individual coefficient the accuracy is still considered as very good for arable fields with values of 75% for four classes and 74% for 3 classes. The coefficient for wooded semi-natural grassland is good with a value of 52% as well as for wetland with 45% (table 5).

Table 5: A summary of the different coefficient of agreement for the different model created. SNG means Semi-natural grassland. 1900* is the model that is derived using the 1800's model and the land use change regression coefficient. Abbreviation used in the table: Kappa O = Kappa overall, SNG = Semi-natural grassland.

| | | Kappa individual | | | | |
|---|---|---|---|---|---|---|
| Model | Kappa O | Arable | Open SNG | Wooded SNG | SNG | Wetland |
| 1800 3 classes | 25.7 | 33 | 27.6 | 22.3 | -- | -- |
| 1800 2 classes | 30.2 | 33.1 | -- | -- | 27.8 | -- |
| 1900 4 classes | 53.6 | 76.1 | 0.03 | 49.2 | -- | 56.6 |
| 1900 3 classes | 54.1 | 76.1 | -- | -- | 42.3 | 56.6 |
| 1900* 4 classes | 55 | 75 | 0.8 | 52.4 | -- | 46 |
| 1900* 3 classes | 55.29 | 74.6 | -- | -- | 44.3 | 56.6 |

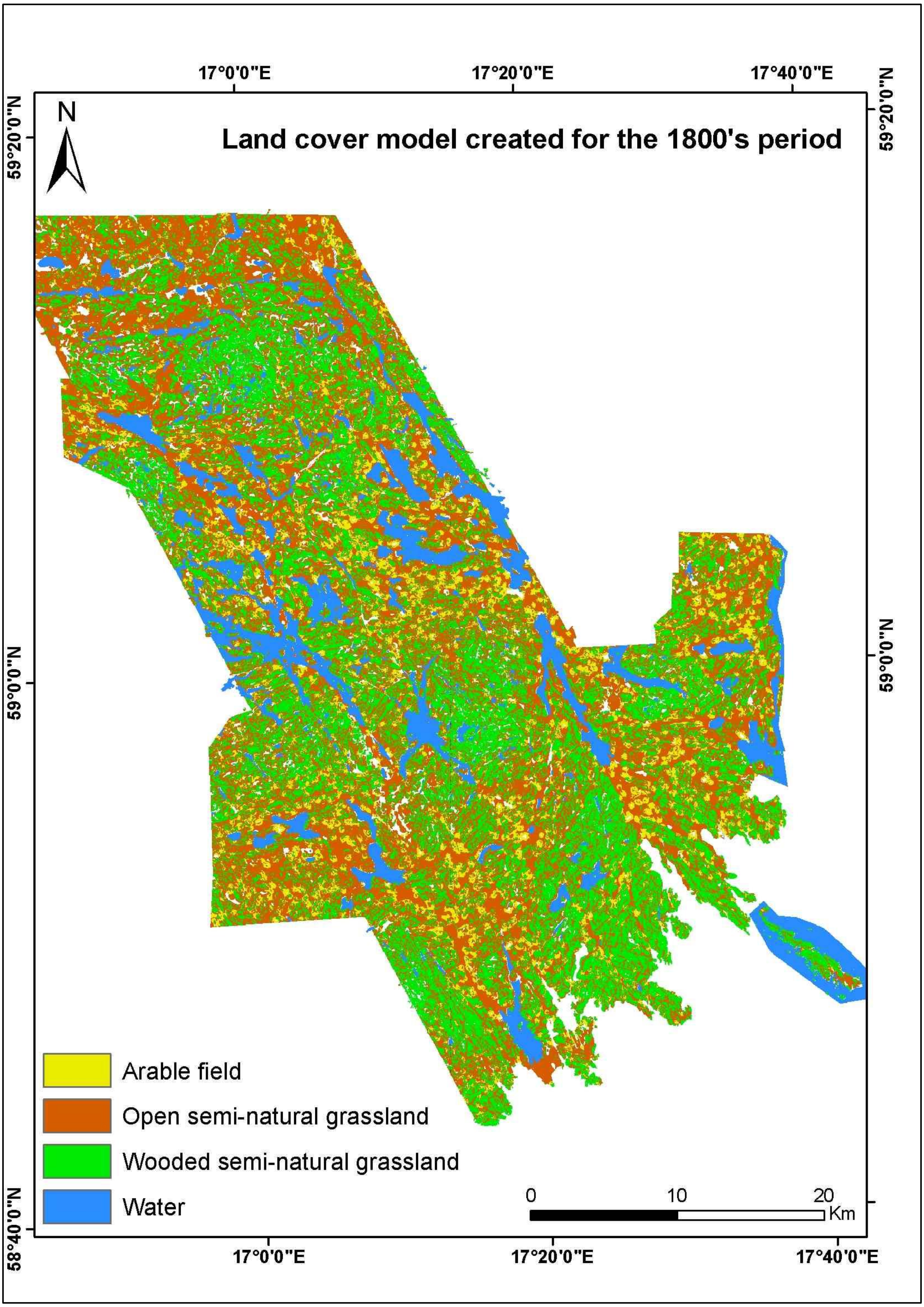


Figure 4: Visual representation of the 1800's best model. Arable field, open and wooded semi-natural grassland are the output classes of the model while the class water is only here for visualization purpose.

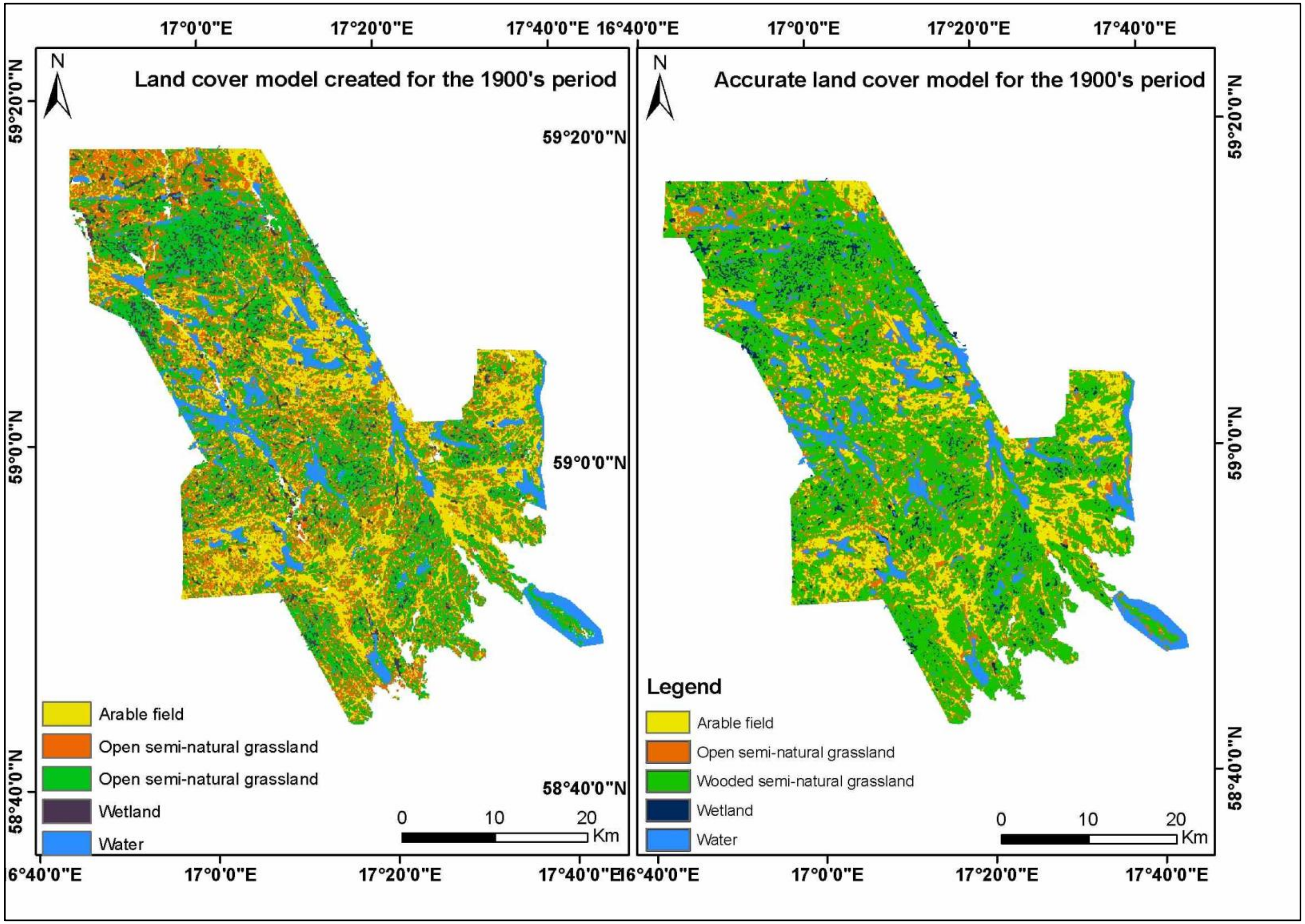


Figure 5: Visual representation of the best model for 1900's derived from the 1800's model and the changes coefficient.

## Discussion

*Landscape change trajectories*

This study aimed to create land cover models based on soil type, distance to settlement, distance to water, and how steep the terrain is. Furthermore to explain the parameters that are related to the land cover at the four different time periods: 1800, 1900, 1950 and today. This study shows clearly that between 1800 and 1900 the landscape undergone dramatic changes in its land cover.

Before the 1900's period it was likely to find semi-natural grasslands on every type of soil. The diversity of semi-natural grassland was high with habitats such as: moist-wet meadows and grassland on coarser and drier soil. The open semi-natural grasslands located in the infield system were mostly used for fodder and haymaking while the other type of semi-natural grasslands, located in the outland system, were grazed during summer (Cousins, 2009). Thus at this period the terrain slope percentage as well as the distance to water were not important parameters for the land cover. Because of this diversity of semi-natural grasslands the accuracy of creating a model based on soil type, slope, distance to settlement and terrain steepness was not particularly good for the period 1800.

From the 1900's the soil became an important parameter for the repartition of arable fields together with the terrain slope percentage. Arable fields were more likely to be found on fine soil, such as clay, and on flat terrain. Analysing changes between 1800 and 1900 show that grassland located on clay soil were more likely to be converted to arable fields. Thus from this period semi-natural grassland were mostly located on coarser and drier soil. The fact that the diversity of habitats decreased, it is likely to have an effect also on the biodiversity. Particularly on specialist plant species that are found on moist-wet semi-natural grasslands (Kuussaari et al., 2009; Krauss et al., 2010).

Even though the distance to the closest settlement was important, this parameter was less important for the landscape changes between 1800 and 1900, and especially for the location of arable fields. This can be the consequence of the land consolidation (*laga skifte*) that occurred until 1870 (Jerneck, 2005). The landscape structure after 1900 was clearly divided with arable fields on flat terrain on fine and rich soils while wooded grassland or grazed forest was located on the steepest terrain. It is interesting how the soil type peat was important for wetland's location in the 1900's time layer while it became not significant in the 1950's. In the 1950's slope became the most important parameter together with the distance to water. This can perhaps be explained by the fact that river systems were regulated at the beginning of the 20th century. From the 1950's is clear that most of the wooded semi-natural grassland disappeared in favour of the forest industry.

Since the source of historical maps in Sweden is unique it was hard to find other studies from other countries that have analysed parameters that can explain landscape repartition as well as changes between different periods. Nevertheless, two studies have been performed with different type of physical and socioeconomic data at different periods (Aspinall, 2004; Millington et al., 2007). Both studies used logistic regression models . Even though the techniques are equivalent the parameters that are tested cannot be the same because of the lack of data for some socioeconomic factors in the 18th century. It would be interesting to investigate more socioeconomic data from the 1800 in order to add more parameters in the model. This can be achieved together with historians that may have more knowledge of the socioeconomic features at this time for example: land ownership, number of animal present in each area, number of people working in the village or land abandonment rate. For instance the university of Gothenburg is working on the creation of a database that aims to collect agriculture statistics for a period between 1570 and 1810, where the aim is to gather information on population, agriculture, land ownership or cattle present in the villages.

*Limitations of the model*

The degree of precision between the cookie cutters and the whole landscape from 1900 differs significantly. The cookie cutters were much more detailed in both thematic and spatial resolution while the landscape description for 1900 is coarser. Because the cookie cutters were both thematically and spatially detailed, it was difficult to get an output model that could retrieve the landscape structure as it was in 1800. This certainly explains a lower accuracy when comparing the result model with the cookie cutters. When comparing the accuracy of the 1900's model it is much more accurate and that might be explain by the fact that the accurate model had a coarser representation.

When performing the equation (3) in the raster calculator, it was difficult to avoid the fact that the soil type influenced the results because it was the only parameters that was not a continuous value. In further studies it would interesting to convert the categorical soil data into continuous soil data. This could be done by ranking the soil texture from one to ten. The value one can be given to coarser soils and the value 10 given to finer soils.

Regarding the analyses of slope percentage it could have been interesting to test the slope with finer resolution. With a resolution of 10 m it is more difficult to model the small changes in steepness of the terrain within the landscape. However it was impossible in this student project, as it required that all the layers have the same resolution, which required much more hardware power to execute the calculations and store data.

*Encouraging results*

The results are encouraging when looking at the results for the 1900's model derived from the 1800 model with the changes coefficient. Since the model for 1800 seems to be working it can be interesting to carry out further studies on this model with more cookie cutters in order to get a stronger statistical power. It is possible that such information will give an even more accurate model for the 1800 period. It also important to note that the settlement used for the whole landscape model for 1800 was the one from 1900. Thus it is possible that there are some differences.

With such a model it will be possible to analyse the history of landscape and to be able to analyse the impact of landscape changes on species diversity. It would also be possible to pinpoint areas for restoration where in the past semi-natural grassland were present. It is important not only to preserve actual grassland but also to restore authentic semi-natural grassland in order to increase the diversity both in plant species but also in landscape, which is one of the most important factors for the diversity of plant species. Most of the grasslands in the landscape today are not semi-natural grassland but cultivated grassland located on former arable fields.

Finally such a model like the one created for 1800 is important to increase our knowledge about the role played by landscape history in the diversity of plant species. This model shows that there was such a diversity of grassland habitats on different type of soil and different terrain that have become lost since the 1900's. The collection of more socioeconomic parameters for the 1800's periods as well as for the latest periods is one of the priorities for the construction of more accurate model.

## Conclusion

This study aimed to analyse rural landscape change trajectories using historical maps soil type and variable such as slope distance to settlement and distance to water. The analysis was carried out using GIS and General Linear Model. The results showed that between 1800 and 1900 the diversity of habitats among semi-natural grasslands decreased significantly. Also the model created for 1800 has a good accuracy even though it could be more precise if more variable were added to the model equation. With such a model it will be possible to analyse the history of rural landscape and understand more thoroughly the fragmentation processes. Therefore it will be possible to pinpoint areas for restoration of old habitats that were important for plant species diversity. Finally this showed that GIS is a good tool for modelling past landscape patterns. It can be used for when deciding which areas should be restored or conserved.

References


Antonson, H., Jansson, U., 2011. Agriculture and forestry in Sweden since 1900. Geographical and historical studies [WWW Document]. Bokus.com. URL http://www.bokus.com/bok/9789186573119/agriculture-and-forestry-in-sweden-since-1900-geographical-and-historical-studies/ (accessed 4.15.13).

Aspinall, R., 2004. Modelling land use change with generalized linear models—a multi-model analysis of change between 1860 and 2000 in Gallatin Valley, Montana. Journal of Environmental Management 72, 91–103.

Berg, Å., Ahrné, K., Öckinger, E., Svensson, R., Söderström, B., 2011. Butterfly distribution and abundance is affected by variation in the Swedish forest-farmland landscape. Biological Conservation 144, 2819–2831.

Cousins, S., Ohlson, H., Eriksson, O., 2007. Effects of historical and present fragmentation on plant species diversity in semi-natural grasslands in Swedish rural landscapes. Landscape Ecology 22, 723–730.

Cousins, S.A.O., 2009. Landscape history and soil properties affect grassland decline and plant species richness in rural landscapes. Biological Conservation 142, 2752–2758.

Cousins, S.A.O., Eriksson, O., 2002. The influence of management history and habitat on plant species richness in a rural hemiboreal landscape, Sweden. LANDSC ECOL 17, 517–529.

Cousins, S.A.O., Lavorel, S., Davies, I., 2003. Modelling the effects of landscape pattern and grazing regimes on the persistence of plant species with high conservation value in grasslands in south-eastern Sweden. LANDSC ECOL 18, 315–332.

Critchley, C., 2004. Conservation of lowland semi-natural grasslands in the UK: a review of botanical monitoring results from agri-environment schemes. Biological Conservation 115, 263–278.

Eriksson, O., Eriksson, A., Berglund, H., 1995. Species abundance patterns of plants in Swedish semi-natural pastures. Ecography 18, 310–317.

Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N., Snyder, P.K., 2005. Global Consequences of Land Use. Science 309, 570–574.

Gimmi, U., Lachat, T., Bürgi, M., 2011. Reconstructing the collapse of wetland networks in the Swiss lowlands 1850–2000. Landscape Ecol 26, 1071–1083.

Greb, S.F., DiMichele, W.A., 2006. Wetlands Through Time. Geological Society of America.

Hermy, M., Honnay, O., Firbank, L., Grashof-Bokdam, C., Lawesson, J.E., 1999. An ecological comparison between ancient and other forest plant species of Europe, and the implications for forest conservation. Biological Conservation 91, 9–22.

Hermy, M., Verheyen, K., 2007. Legacies of the past in the present-day forest biodiversity: a review of past land-use effects on forest plant species composition and diversity, in: Nakashizuka, T. (Ed.), Sustainability and Diversity of Forest Ecosystems. Springer Japan, pp. 361–371.

Honnay, O., Jacquemyn, H., Bossuyt, B., Hermy, M., 2005. Forest fragmentation effects on patch occupancy and population viability of herbaceous plant species. New Phytol. 166, 723–736.

Keddy, P.A., 2010. Wetland ecology : principles and conservation. Cambridge University Press, Cambridge.

Krauss, J., Bommarco, R., Guardiola, M., Heikkinen, R.K., Helm, A., Kuussaari, M., Lindborg, R., Öckinger, E., Pärtel, M., Pino, J., Pöyry, J., Raatikainen, K.M., Sang, A., Stefanescu, C., Teder, T., Zobel, M., Steffan-Dewenter, I., 2010. Habitat fragmentation causes immediate and time-delayed biodiversity loss at different trophic levels. Ecology Letters 13, 597–605.

Kuussaari, M., Bommarco, R., Heikkinen, R.K., Helm, A., Krauss, J., Lindborg, R., Ockinger, E., Pärtel, M., Pino, J., Rodà, F., Stefanescu, C., Teder, T., Zobel, M., Steffan-Dewenter, I., 2009. Extinction debt: a challenge for biodiversity conservation. Trends Ecol. Evol. (Amst.) 24, 564–571.

Lindborg, R., 2007. Evaluating the distribution of plant life-history traits in relation to current and historical landscape configurations. Journal of Ecology 95, 555–564.

Millington, J.D.A., Perry, G.L.W., Romero-Calcerrada, R., 2007. Regression Techniques for

Examining Land Use/Cover Change: A Case Study of a Mediterranean Landscape. Ecosystems 10, 562–578.

Ozinga, W.A., Römermann, C., Bekker, R.M., Prinzing, A., Tamis, W.L.M., Schaminée, J.H.J., Hennekens, S.M., Thompson, K., Poschlod, P., Kleyer, M., Bakker, J.P., van Groenendael, J.M., 2009. Dispersal failure contributes to plant losses in NW Europe. Ecology Letters 12, 66–74.

Reich, P.B., Frelich, L., 2002. Temperate Deciduous Forests, in: Encyclopedia of Global Environmental Change. pp. 565–569.

Rossiter, D.G., Loza, A., 2004. Technical Note: Analyzing land cover change with R, in: Enschede (NL): International Institute for Geo-information Science & Earth Observation (ITC.

Sala, O.E., 2000. Global Biodiversity Scenarios for the Year 2100  Science 287, 1770–1774.

Stoate, C., Boatman, N.., Borralho, R.., Carvalho, C.R., Snoo, G.R. d., Eden, P., 2001. Ecological impacts of arable intensification in Europe. Journal of Environmental Management 63, 337–365.

Wilson, J.B., Peet, R.K., Dengler, J., Pärtel, M., 2012. Plant species richness: the world records. Journal of Vegetation Science 23, 796–802.

## Appendix 1

### Regression coefficients for each land cover class as well as for each land cover change class.

Table 6: Regression coefficient for arable fields at the different time-periods with each variable as an individual parameter. For each parameter there is also its accuracy with the D2 value. Only the significant coefficients are written in the table. The p value are written as follow: ‘***’ p < 0.001, ‘**’ p < 0.01, ‘*’ p < 0.05, ‘.’ p < 0.1 and “—“ no significant.

| arable | | | Soil properties | | | | | | Geographical variables | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **$D_2$** | **Slope** | **$D_2$** | **Water** | **$D_2$** | **Settlement** | **$D_2$** |
| 1800 | -2.607 *** | 1.155 *** | -0.67 *** | 2.502 *** | 1.474 *** | -- | 8.01 | -0.866 *** | 1.48 | 0.117 *** | 0.19 | -0.906 *** | 10.95 |
| 1900 | -2.074 *** | -- | -0.217 . | 1.417 *** | 3.168 *** | -- | 33.44 | -3.276 *** | 17.27 | 0.432 *** | 2.65 | -0.867 *** | 5.95 |
| 1950 | -2.485 *** | -- | -0.326 * | 1.934 *** | 3.542 *** | -- | 38.37 | -3.671 *** | 19.89 | 0.410 *** | 2.63 | -0.549 *** | 3.88 |
| 2003 | -2.646 *** | -- | -0.470 ** | -- | 3.335 *** | -- | 34.04 | -3.336 *** | 16.76 | 0.601 *** | 4.94 | -0.556 | 3.78 |

Table 7: Best model regression coefficient for arable fields at each period. The regression coefficients are given for each parameter that are included in the model. The best model accuracy is given with the $D^2$ value. Only the significant coefficients are written in the table. The p value are written as follow: ‘***’ p < 0.001, ‘**’ p < 0.01, ‘*’ p < 0.05, ‘.’ p < 0.1 and “—“ no significant.

| Arable | | | Soil properties | | | | Geographical variables | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **Slope** | **Water** | **Settlement** | **$D_2$** |
| 1800 | 3.649 *** | 0.74 . | -0.699 *** | 1.38 ** | 1.549 *** | -- | -- | -0.145 *** | -0.964*** | 18.58 |
| 1900 | 8.028 *** | -- | -0.346 ** | 1.014 ** | 2.809 *** | -0.642 * | -2.556 *** | 0.640 *** | -1.209 *** | 46.31 |
| 1950 | 7.228 *** | -0.992 . | -0.491 ** | 1.414 *** | 3.147 *** | -0.987 ** | -2.914 *** | 0.496 *** | -0.754 *** | 49.89 |
| 2003 | 3.238 *** | -- | -0.673 *** | -- | 2.888 *** | -1.562 *** | -2.485 *** | 0.822 | -0.665 | 45.80 |

Table 8: Regression coefficient for forest at the different time-periods with each variable as an individual parameter. For each parameter there is also its accuracy with the D2 value. Only the significant coefficients are written in the table. The p value are written as follow: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.1 and "—" no significant.

| forest | Soil properties | | | | | | | Geographical variables | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **$D_2$** | **Slope** | **$D_2$** | **Water** | **$D_2$** | **Settlement** | **$D_2$** |
| 1800 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1900 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1950 | 0.386 *** | -1.286 *** | 0.199 ** | -0.533 . | -2.685 *** | 0.794 *** | 25.57 | 2.114 *** | 10.23 | -0.145 *** | 0.38 | 0.634 | 4.63 |
| 2003 | 0.756 *** | -1.692 *** | -- | -0.609 . | -2.723 | 1.051 *** | 27.07 | 2.014 *** | 9.2 | -- | -- | 0.856 | 7.38 |

Table 9: Best model regression coefficient for forest at each period. The regression coefficients are given for each parameter that are included in the model. The best model accuracy is given with the $D^2$ value. Only the significant coefficients are written in the table. The p value are written as follow: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.1 and "—" no significant.

| forest | Soil properties | | | | | Geographical variables | | | |
|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **Slope** | **Water** | **Settlement** | **$D_2$** |
| 1800 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1900 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1950 | -7.407 *** | -1.243 *** | 0.342 *** | -- | -2.352 *** | 1.474 *** | 1.778 *** | -- | 0.718 *** | 31.67 |
| 2003 | -7.585 *** | -1.561 *** | 0.235 ** | -- | -2.450 *** | 1.805 *** | 0.978 *** | -- | 0.941 *** | 34.44 |

Table 10: Regression coefficient for semi-natural grasslands at the different time-periods with each variable as an individual parameter. The regression coefficients are given for each parameter as an individual parameter. For each parameter there is also its accuracy with the D2 value. Only the significant coefficients are written in the table. The p value are written as follow: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.1 and "—" no significant.

| | Soil properties | | | | | | | Geographical variables | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **$D_2$** | **Slope** | **$D_2$** | **Water** | **$D_2$** | **Settlement** | **$D_2$** |
| **1800** | | | | | | | | | | | | | |
| Open grass | -1.371 *** | 0.951 *** | 0.402 *** | 1.053 * | 1.234 *** | -- | 5.02 | -1.455 *** | 4.62 | -0.148 *** | 0.37 | -0.124 *** | 0.25 |
| Cover grass | 0.895 *** | -1.243 *** | -- | -3.035 *** | -1.805 *** | 0.823 * | 12.86 | 1.67 *** | 6.6 | 0.101 *** | 0.17 | 0.654 *** | 6.1 |
| Grass | 2.343 *** | -0.891 ** | 0.755 *** | -2.238 *** | -1.229 *** | -- | 6.58 | 0.679 *** | 0.96 | -0.070 * | 0.07 | 0.842 *** | 9.66 |
| **1900** | | | | | | | | | | | | | |
| Open grass | -2.224 *** | 1.059 *** | 0.588 *** | 1.093 ** | -0.703 *** | -- | 3.84 | -- | -- | -0.324 | 1.41 | -0.43 | 1.44 |
| Cover grass | 1.317 *** | -0.912 *** | -0.296 *** | -1.662 *** | -2.709 *** | -1.492 *** | 24.47 | 2.995 *** | 16.36 | -0.295 *** | 1.33 | 1.122 *** | 8.64 |
| grass | 2.053 *** | -- | -- | -1.397 *** | -3.153 *** | -1.687 *** | 32.18 | 3.345 *** | 17.89 | -0.434 *** | 2.68 | 0.855 *** | 5.78 |
| **1950** | | | | | | | | | | | | | |
| open grass | -2.393 *** | 1.109 *** | -- | -- | -0.202 * | -- | 0.52 | 0.270 | 0.15 | -0.164 *** | 0.39 | -0.259 *** | 0.71 |
| cover grass | -1.923 *** | 0.566 * | -- | -- | -1.536 *** | -0.822 * | 6.29 | 0.959 *** | 2.04 | -- | -- | -- | -- |
| cugrass | -3.855 *** | -- | 0.726 *** | 1.316 * | -- | -- | 1.18 | -- | -- | -- | -- | -0.786 *** | 6.04 |
| grass | -1.623 *** | -1.321 * | -0.409 *** | -- | -- | -0.521 . | 0.6 | -- | -- | 0.512 *** | 2.88 | 0.088 ** | 0.1 |
| **2003** | | | | | | | | | | | | | |
| ogra | -3.779 *** | -- | 0.599 ** | -- | -- | -- | 1.42 | -- | -- | -0.284 | 0.99 | -0.278 *** | 0.62 |
| cover grass | -2.706 *** | 1.302 *** | -0.613 *** | -- | -0.891 *** | -- | 3.21 | 1.105 | 2.51 | 0.167 ** | 0.28 | -0.533 | 2.76 |
| culti grass | -3.400 *** | 1.336 *** | -- | 1.819 *** | 1.111 *** | -- | 4.12 | -0.885 *** | 1.19 | -0.207 *** | 0.6 | -.605 *** | 3.88 |
| grass | -2.019 *** | 1.218 *** | -- | 0.887 * | 0.165 * | -1.866 | 0.88 | -- | -- | -0.129 | 0.25 | -0.590 | 4.04 |

Table 11: Best model regression coefficient for grassland with the different classes at each period. The regression coefficients are given for each parameter included in the model. The best model accuracy is given with the $D^2$ value. Only the significant coefficients are written in the table. The p value are written as follow: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$ and "—" no significant.

| | Soil properties | | | | | | Geographical variables | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **Slope** | **Water** | **Settlement** | **$D_2$** |
| **1800** | | | | | | | | | | |
| Open grass | 4.12 *** | 0.709 * | 0.431 *** | -- | 0.920 *** | -- | -0.982 *** | -0.268 *** | -0.156 *** | 7.75 |
| cov grass | -9.52 *** | -0.783 *** | -0.225 * | -- | -1.702 *** | 0.819 . | 0.796 *** | 0.504 *** | 0.853 *** | 21.62 |
| grassland | -2.489 *** | -- | 0.818 *** | -1.342** | -1.231 *** | -- | -- | -- | 0.856 *** | 15.61 |
| **1900** | | | | | | | | | | |
| Open grass | 3.914 *** | 0.582 * | 0.604 *** | -- | -0.938 *** | | -0.521 *** | -0.303 *** | -0.504 *** | 7.15 |
| cov grass | -10.581 *** | -- | -0.276 *** | -1.123 ** | -2.260 *** | -1.023 *** | 2.286 *** | -0.349 *** | 1.317 *** | 38.26 |
| grassland | -7.806 | -- | 0.238 . | -1.024 ** | -2.782 *** | -1.279 *** | 2.557 *** | -0.653 *** | 1.177 *** | 45.13 |
| **1950** | | | | | | | | | | |
| Open grass | -- | 0.993 *** | -- | -1.741 . | -0.211 * | -- | -- | -0.195 *** | -0.293 *** | 1.78 |
| cov grass | -7.168 *** | -1.248 * | -0.441 *** | -- | -0.888 *** | -1.401 * | 0.699 *** | -- | 0.517 *** | 7.74 |
| culti grass | -- | -- | 0.767*** | -- | -- | -- | -- | -- | -0.811 *** | 7.41 |
| grassland | -5.363 *** | -1.121 * | -0.469 *** | -- | -0.125 . | -- | -- | 0.507 *** | 0.114 ** | 3.52 |
| **2003** | | | | | | | | | | |
| Open grass | -- | -- | 0.608 ** | | -0.311 . | -- | -- | -0.301 *** | -0.358 *** | 3.27 |
| cov grass | -2.818 *** | 0.574 * | -- | -- | -1.425 *** | -0.697 * | 0.291 * | -- | -0.57 | 6.44 |
| culti grass | 2.768 *** | 0.928 * | -- | 1.011 * | 0.947 *** | -- | -0.33 * | -0.327 *** | -0.588 *** | 8.55 |
| grassland | 2.445 *** | 0.863 ** | -- | -- | -- | -2.034 *** | -- | -0.189 *** | -0.597 *** | 5.14 |

Table 12: Regression coefficient for wetlands at the different time-periods with each variable as an individual parameter. For each parameter there is also its accuracy with the D2 value. Only the significant coefficients are written in the table. The p value are written as follow: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.1 and "—" no significant.

| wetland | Soil properties | | | | | | | Geographical variables | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **$D_2$** | **Slope** | **$D_2$** | **Water** | **$D_2$** | **Settlement** | **$D_2$** |
| 1800 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1900 | -6.720 *** | -- | 2.128 *** | -- | -- | 5.850 *** | 46.43 | -2.870 *** | 5.21 | -- | -- | 0.502 * | 0.86 |
| 1950 | -7.827 *** | -- | -- | -- | 3.287 ** | -- | 7.38 | -5.282 *** | 10.79 | -1.404 | 28.82 | 0.793 ** | 3.13 |
| 2003 | -6.033 *** | -- | -- | -- | 2.341 *** | 1.735 * | 7.05 | -5.902 *** | 14.07 | -1.101 | 19.42 | 0.277 | 0.47 |

Table 13: Best model regression coefficient wetlands at each period. The regression coefficients are given for each parameter included in the model. The best model accuracy is given with the $D^2$ value. Only the significant coefficients are written in the table. The p value are written as follow: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.1 and "—" no significant.

| wetland | Soil properties | | | | | | Geographical variables | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **Slope** | **Water** | **Settlement** | **$D_2$** |
| 1800 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1900 | -- | -- | -- | -- | -- | 5.152 *** | -2.203 *** | -- | -- | 48.88 |
| 1950 | 12.151 *** | -- | -- | -- | 2.888 ** | -- | -4.282 *** | -1.697 *** | -- | 47.10 |
| 2003 | 15.351 *** | -- | -- | | 1.842 *** | -- | -5.131 *** | -1.443 *** | -- | 39.97 |

Table 14: Regression coefficient for arable fields at the different time-periods with each variable as an individual parameter. For each parameter there is also its accuracy with the D2 value. Only the significant coefficients are written in the table. The p value are written as follow: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.1 and "—" no significant.

| deciduous | Soil properties | | | | | | | Geographical variables | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **$D_2$** | **Slope** | **$D_2$** | **Water** | **$D_2$** | **Settlement** | **$D_2$** |
| 1800 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1900 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1950 | -2.388 *** | 0.876 ** | -0.306 * | -- | -0.987 *** | -2.609 ** | 3.29 | 1.233 *** | 3.32 | -0.396 *** | 2.42 | 0.949 *** | 6.41 |
| 2003 | -2.102 *** | 1.301 *** | 0.332 ** | -- | -0.851 *** | -- | 3.7 | 1.096 *** | 2.77 | -0.550 *** | 5.19 | 0.190 ** | 0.32 |

Table 15: Best model regression coefficient deciduous forest at each period. The regression coefficients are given for each parameter included in the model. The best model accuracy is given with the $D^2$ value. Only the significant coefficients are written in the table. The p value are written as follow: '***' p < 0.001, '**' p < 0.01, '*' p < 0..05, '.' p < 0.1 and "—" no significant.

| deciduous | Soil properties | | | | | | Geographical variables | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **Slope** | **Water** | **Settlement** | **$D_2$** |
| 1800 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1900 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1950 | -8.573 *** | 0.999 ** | -- | -- | -0.509 *** | -1.971 . | 0.794 *** | -0.216 *** | 0.806 *** | 10.95 |
| 2003 | -1.579 *** | 1.253 *** | 0.469 *** | -- | -0.423 *** | -- | 0.756 *** | -0.509 *** | -- | 9.09 |

Table 16: Regression coefficient for each type of changes at each period. The regression coefficients are given for each parameter included in the model. The best model accuracy is given with the $D^2$ value. Only the significant coefficients are written in the table. The p value are written as follow: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$ and "—" no significant.

| | Soil properties | | | | | | | Geographical variables | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **$D_2$** | **Slope** | **$D_2$** | **Water** | **$D_2$** | **Settlement** | **$D_2$** |
| ograss-> ara | -1.334 *** | -- | -0.557 ** | -- | 2.815 *** | 0.864 * | 27.69 | -2.884 *** | 14.99 | 0.602 *** | 5.45 | -0.142 * | 0.15 |
| ograss -> cgrass | 0.414 *** | -- | 0.336 * | -- | -2.331 *** | -1.618 *** | 21.11 | 2.642 *** | 14.28 | -0.322 *** | 1.64 | 0.479 *** | 1.41 |
| cgrass -> ara | -3.145 *** | 1.164 * | -- | 1.248 * | 2.318 *** | 3.173 *** | 29.08 | -3.361 *** | 15.47 | 1.046 *** | 9.01 | -0.715 *** | 4.1 |
| grass -> | -2.474 *** | 0.811 * | -- | 0.825 . | 3.295 *** | 2.08 *** | 32.6 | -2.597 | 11.27 | 0.492 *** | 3.36 | -0.632*** | 3.11 |
| arable-> ogra | -2.000 *** | 0.391 ** | -- | -- | -1.800 *** | -- | 12.68 | 4.584 *** | 19.48 | -- | -- | 1.324 *** | 7.82 |
| arable -> cgra | -0.763 *** | 2.268 ** | -- | -- | -1.908 *** | -- | 16.37 | 4.251 *** | 17.21 | -- | -- | 0.673 *** | 2.76 |

Table 17: Best model regression coefficient for each type of change at each period. The regression coefficients are given for each parameter included in the model. The best model accuracy is given with the $D^2$ value. Only the significant coefficients are written in the table. The p value are written as follow: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$ and "—" not significant.

| | Soil properties | | | | | | Geographical variables | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **model** | **Bedrock** | **Rockstone** | **Till & sand** | **Finesand** | **Clay** | **Peat** | **Slope** | **Water** | **Settlement** | **$D_2$** |
| ograss-> ara | 3.759 *** | -- | -0.544 * | -- | 2.522 *** | -- | -2.213 *** | 0.878 *** | -0.735 *** | 39.01 |
| ograss -> cgrass | -8.591 *** | -- | 0.307 . | -- | -1.906 *** | -0.965 . | 2.035 *** | -0.417 *** | 1.009 *** | 30.39 |
| cgrass -> ara | 3.090 *** | 0.100 . | -- | 2.138 ** | 2.777 *** | -- | -2.505 *** | 1.111 *** | -1.079 *** | 45.10 |
| grass -> | 6.15 *** | -- | -- | -- | 2.908 *** | 1.623 *** | -2.592 *** | 0.708 *** | -0.991 *** | 44.74 |
| arable -> ogra | -20.843 | -- | -- | -- | -0.924 ** | -- | 4.077 *** | -- | 1.352 *** | 30.41 |
| arable -> cgra | -17.206 *** | 3.761 *** | 0.862 * | -- | -1.307 *** | -- | 4.069 *** | -- | 0.846 *** | 31.81 |

Series from Lund University

Department of Physical Geography and Ecosystem Science

**Master Thesis in Geographical Information Science (LUMA-GIS)**

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland. (2008).
2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research. (2008).
3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing. (2009).
4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia. (2009).
5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).
6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).
7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010)
8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010)
9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010)
10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation

and malaria transmission intensities in selected parts of Africa (2011)

11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011)
12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).
13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).
14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011)
15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011)
16. *David Kanyari*: Nairobi City Journey Planner An online and a Mobile Application (2011)

17 *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012)

18 *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012)

19 *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012)

20 *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013)

21 *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine*' (2013)

22 *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013)

23 *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013)

24 *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013)