Land use modelling of Austria with the EuClueScanner

Calibration and validation of explanatory factors

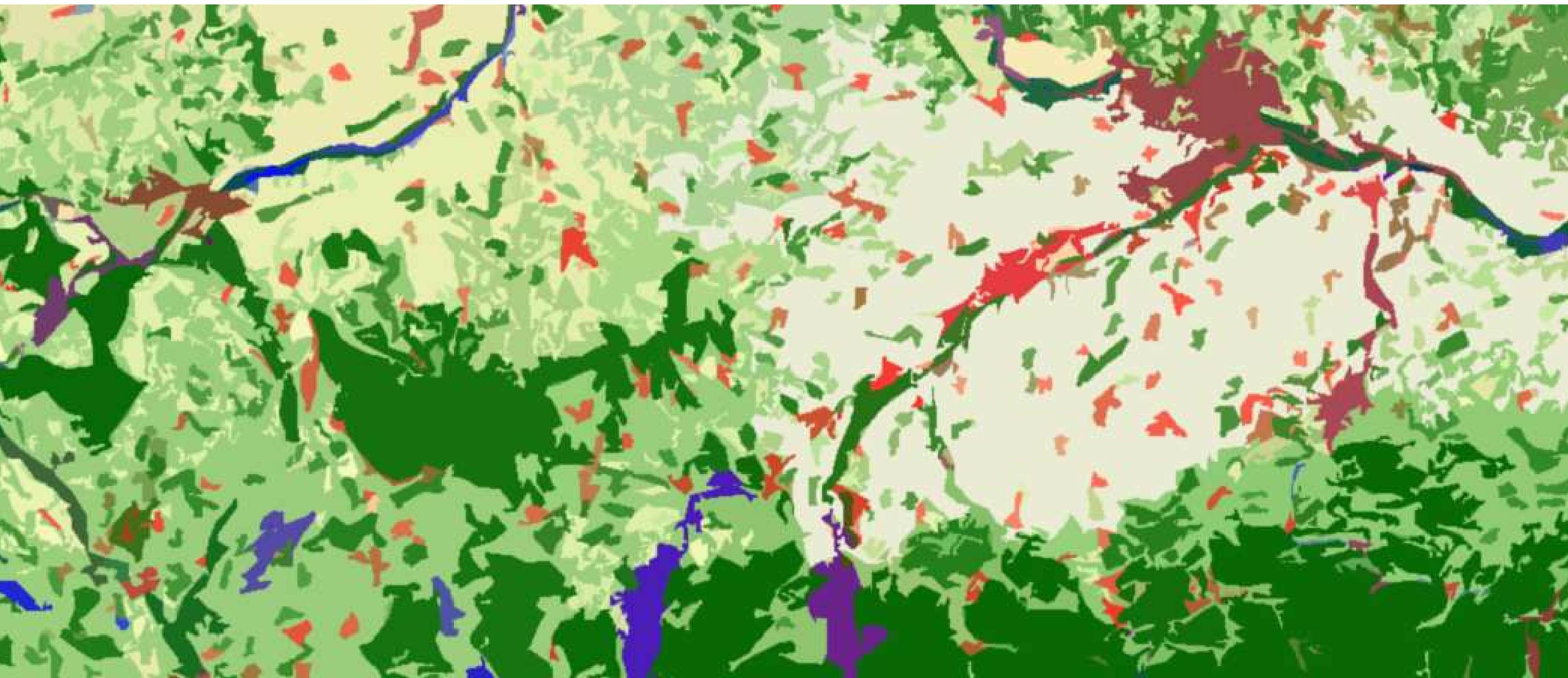
By Martijn Snelder

dr. Eric Koomen and dr. Kees Kasse

Amsterdam, August 2010

Bachelor thesis, Aarde en Economie

Vrije Universiteit



ABSTRACT

The EuClueScanner is a modelling tool under development that simulates land use changes based on scenarios of the future. The model will be available for many countries in Europe. It uses a 100 metre grid which is very detailed for a country wide land use simulation. The total land demand based on the scenarios is defined by other models. The EuClueScanner uses this externally defined demand to allocate the land use on the basis of spatially explicit explanatory factors like slope, soil water availability or the neighbourhood of urban fabric. In this research an assessment of the importance of these explanatory factors on land use in Austria is made. Multinomial regressions were performed to find the relation of these factors with the land use. The model is calibrated in multiple ways on the basis of these regressions. The model results are validated with the observed land use change between 1990 and 2000. In this period only 0.3% of the surface changed to a different land use in Austria. The model resulted to be unable to predict the land use change. Some suggestions are made to improve the model results.

TABLE OF CONTENTS

INTRODUCTION

The surface of the earth is continuously changing. Especially when men started to shape the earth to their own needs. Almost all observable differences on the land surface are the result of humans tightening or loosening their grip on that land (De Koning et al., 1998). Policymakers in Europe are well aware of the great ability of people to change the earth. Because individual interferences might not be in the benefit of others, policies are made in order to steer development in a socially beneficial direction. Meijl et al. (2006) for example describe spatial impacts of aborting agricultural subsidies and the liberalisation of the food market. Koomen et al. (2008) describe some consequences of climate change on land use. The consequences of policies and other future developments on the overall land demand per country or region are determined by macro economical or global environment models. But where will these transitions take place?

The EuClueScanner is a model under development as a tool to give an impression of the changes in land use as a result of changes in land demand. The model uses spatially explicit explanatory factors to allocate the externally defined demand for land. These factors indicate a probability of finding a certain land use type on a specific location. Factors that are used are for example slope, altitude or distance to a city. The EuClueScanner will cover most countries in Europe. Although many processes that determine land allocation are similar across countries, every country has its own characteristics. This research focuses on Austria but conclusions are useful for the model in general. The main question in this research is:

How should, in the EuClueScanner, explanatory factors be used to simulate future land use change in Austria?

With the following sub-questions:

a) Which land use changes can be observed in Austria between 1990 and 2000?
b) How can the spatial distribution of land use be explained?
c) How can the explanatory analysis best be used to simulate future land use with the EuClueScanner?

The sub-questions can also be described in three keywords (a) analysing, (b) explaining and (c) modelling. The names of paragraphs will often refer to one of these three keywords.

In order to better understand the methods that are used in this research the model is first described in the chapter *EuClueScanner*, then the used *methods* will be explained, after that the *results* will be described followed by a *discussion*, *recommendations* and the *summary and conclusions*. *Acknowledgements*, the *references* and the *annexes* can be found at the end of the paper.

THE *EUCLUESCANNER*

The EuClueScanner is raster-based model with a 100 metre grid that is able to allocate demand that is defined by other models. For a clear understanding of the methods that are used in this research some components of the EuClueScanner need to be explained. The first main component is the land use data set.

Land use data

The land use data are provided by the CORINE land cover (CLC) dataset. It contains maps showing land use in 44 different categories of 1990 and 2000 on a 100 metre grid resolution. The land use maps are defined using satellite imagery. For this research we use the JRC-9 (Joint Research Centre – 9) reclassification. In this reclassification the categories are aggregated to 9 land use classes as shown in table 1. Some original land use categories are assigned to multiple JRC-9 classes as shown by the percentages that indicate the share of the total area of the original CLC-class that is randomly reclassified into the mentioned JRC-classes. Figure 1 illustrates the effect on the JRC-9 land use map. The area in the middle is rendered to arable land and pastures in a scattered pattern. Somewhere in the middle an area is divided over arable land, pastures and semi-natural vegetation in a similar scattered pattern.

Table 1. CLC land use categories

JRC-9	JRC-9 land use	Colour	CLC-class	Name	Simulated
1	Urban fabric		1.1.1	Continuous Urban fabric	Yes
1	Urban fabric [1]		1.1.2	Discontinuous Urban fabric	Yes
2	Industry [1]		1.2.1	Industrial or commercial units	Yes
8	Infrastructure		1.2.2	Road and rail networks	No
8	Infrastructure		1.2.3	Port areas	No
5	Forest		1.2.4	Airports	No
2	Industry [1]		1.3	Mine dump and construction sites	Yes
1	Urban fabric [1]		1.4	Artificial non agricultural vegetated areas	Yes
3	Arable land		2.1.1/ 2.4.2p(50%)/ 2.4.3p(25%)	Arable land (non-irrigated)	Yes
3	Arable land		2.1.2/2.1.3	Arable land (irrigated)	Yes
3	Arable land		2.2/2.4.1/2.4.4	Permanent crops	Yes
4	Pastures		2.3/ 2.4.2p(50%)/ 2.4.3p (45%)	Pastures	Yes
5	Forrest		3.1	Forests	Yes
6	Semi-natural [2]		3.2.1/3.2.3/ 3.2.4/2.4.3p (30%)	Semi natural vegetation [2]	Yes
7	Other nature		3.2.2	Heather and moorlands	No
7	Other nature		3.3.1	Beaches, dunes and sands	No
7	Other nature		3.3.2/ 3.3.3/3.3.4	Sparsely vegetated areas	No
7	Other nature		3.3.5	Glaciers and snow	No
7	Other nature		4.1	Inland wetlands	No
7	Other nature		4.2	Coastal wetlands	No
9	Water		5.1	Inland waters	No
9	Water		5.2	Marine waters	No

[1] Note that that aggregated land use class *industry* consists of industry, commercial units, construction sites and mining dumps and that *urban fabric* can also consist of non-agricultural *vegetation.*
[2] Semi-natural vegetation includes natural grasslands, sclerophyllous vegetation, transitional woodland-crub and land principally occupied by agriculture, with significant areas of natural vegetation.

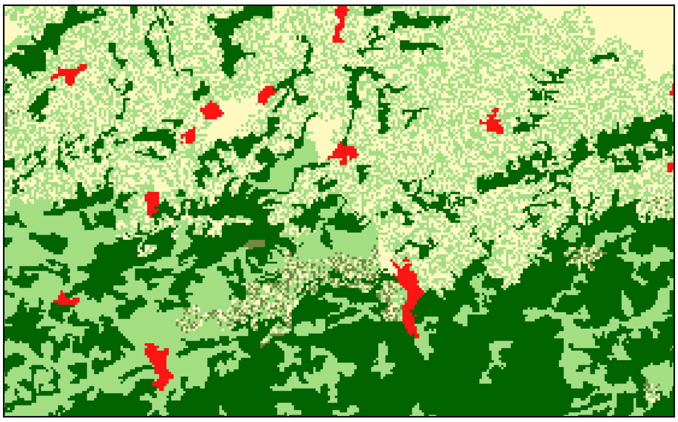


Figure 1. Illustration of the scattered result of the land use classification into multiple categories (for legend see Table 1)

Except for these scattered patterns for some land uses, only large chunks of land uses can be observed on this map. This is the result of a minimum mapping unit of 25 ha. Standalone land use units smaller than 25 are not distinguished.

The EuClueScanner does not simulate the change in all the land use classes as described in the table. These are the *static land use categories*. *Other nature* is not simulated because it is hard to model the types of nature in this category. In Austria *other nature* mainly consists of sparsely vegetated areas, glaciers and snow. Such nature types have a very location specific character. Another issue is that *other nature* is an umbrella term for all these different types of nature that behave differently. This makes it difficult to model an overall change. In addition these types of nature are expected to be relatively stable over time. For the EuClueScanner the decision has therefore been made to consider this category static. Also *infrastructure* is not modelled by the EuClueScanner mainly because of the scarce presence of large infrastructural units and the normally linear shape of infrastructure that cannot be modelled well in a raster-based program like the EuClueScanner. Finally the EuClueScanner does not model changes in water cover.

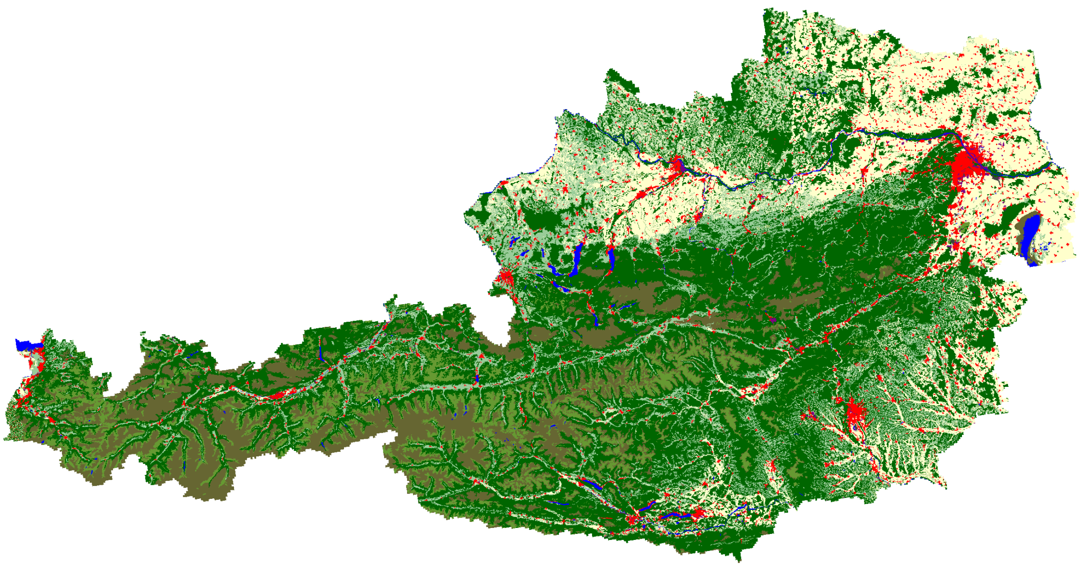


Figure 2. Land use map Austria (for legend see Table 1)

Factor data

The spatially explicit factor data are used to describe suitability for the different land use types. The factors can be divided into *neighbourhood variables* and *non-neighbourhood variables*. The values of the variables are mapped in *factor maps.* Example of two factor maps are shown in figure 3b and 3c. In this paper I might switch between the terms factors and variables because in fact the factors *are* variables.

The neighbourhood variables describe the land use in the direct environment of a cell. These variables are essentially maps that are created for each of the nine land use types. They describe the frequency of occurrence of a certain land use class in the neighbourhood of the cell under consideration. Figure 3a illustrates this. The red cell has a value that is determined by the number of times that a certain land use type occurs in the surrounding rings multiplied by the weight of the rings. Notice that this weight value decreases with distance. When all the cells in the neighbourhood are of a certain land use and all the cells are multiplied with their weight factors, the factor of that land use can reach a value of 88. This method was advertised by Verburg et al. (2003)

For the statistical analysis the neighbourhood variables *urban fabric, industry, arable land, pastures*, *forests*, *semi natural vegetation* and *other nature* are used.

The *non-neighbourhood variables* that are used in the model runs that are described in this paper are *accessibility to cities* larger than 100 000 inhabitants, *accessibility to ports*, *water deficit* during the growing season, *accumulated rainfall from March to July, soil water available to plants, slope* [3], *elevation, presence of an impermeable layer, south slope* and the *Natura 2000 network.* The non-neighbourhood variables, unlike the neighbourhood variables, are very dissimilar. *Accessibility* for example can reach very large values while the binomial variables *impermeable layer* and *Natura 2000* have a value of only 0 or 1. The minimum and maximum values of the non-neighbourhood factors are listed below:

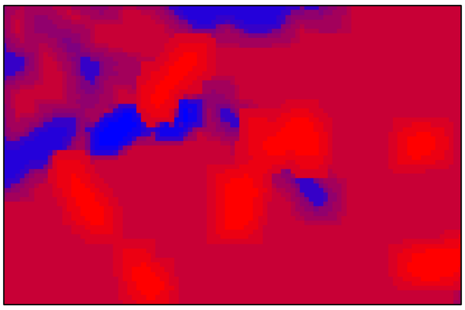
Table 2. Factor descriptives

Factor	Units	Min	Max
AccessCity	minutes	0	706
AccessPort	minutes	99	876
WaterDeficit	millimetres	-77	0
RainFall	millimetres	273	914
SoilWater	millimetres	30	220
Elevation	metres	108	3666
Slope	- [3]	1	6
ImpermeableLayer	-	0	1
Natura 2000	-	0	1
SouthSlope	-	0	1

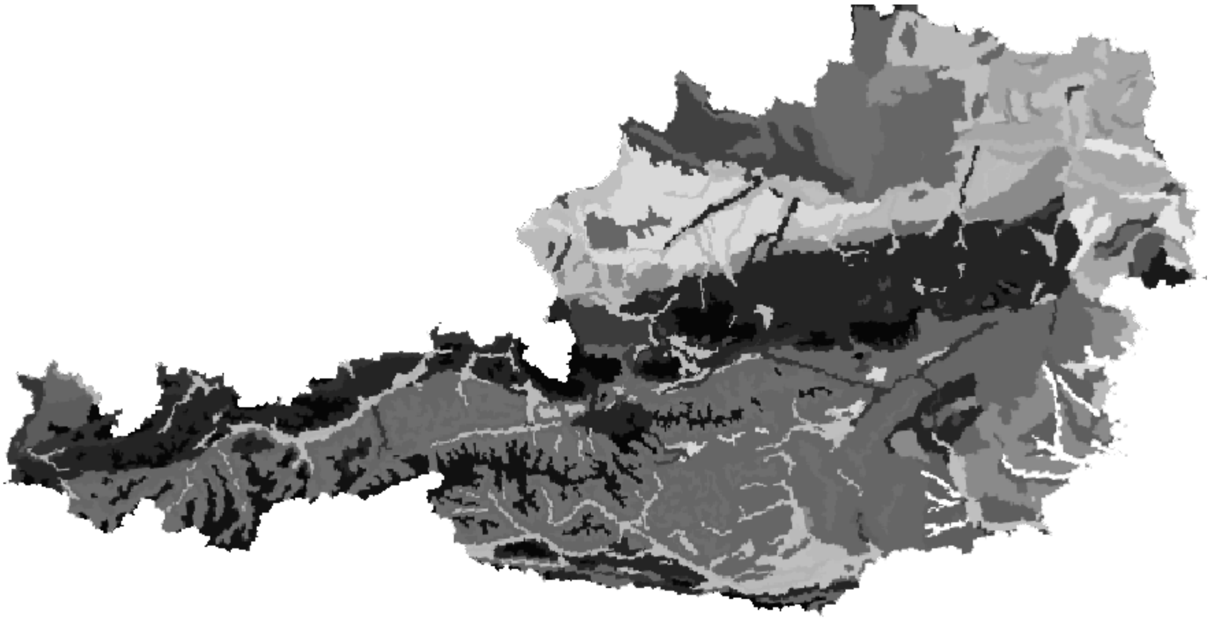
[3] Defined in 6 *slope classes* based on steepness of the slope

1	1	1	1	1	1	1
1	2	2	2	2	2	1
1	2	4	4	4	2	1
1	2	4		4	2	1
1	2	4	4	4	2	1
1	2	2	2	2	2	1
1	1	1	1	1	1	1

a)



b)



c)

Figure 3

a) An illustration of the mechanism of the *neighbourhood* factor *urban fabric*. The central cell in 3a will obtain a value that is determined by the presence of urban fabric in the surrounding cells. The values in the surrounding cells in 3a represent the weight values when urban fabric occurs in that cell.

b) This is a fragment of the neighbourhood map for urban land use. The values range from 88 (in bright red) in the centres of larger towns to 0 (in blue) for cells without any urban fabric in their neighbourhood.

c) this is an illustration of one of the *non-neighbourhood* factor maps. In this specific map, *soil water available to plants* is shown. The values range from 220 (in white) with much soil water available to plants to 30 (in black) with few water available to plants.

Demand module and land use allocation

In the previous sections the land use types and factors are introduced. The following section will explain how the EuClueScanner uses this factor data to create land use type probability maps. The land allocation of land use in the EuClueScanner depends on the probability maps created by the EuClueScanner and the overall demand that is defined externally by other models. An algorithm allocates this demand for all the land use types to the highest possible overall probability. The model allocates the land use in pure pixels. This means that every pixel which is 100x100 metres can contain only one land use and can not be a fraction of different land uses.

The probability maps are created by making a pixel per pixel calculation of the probability for all the land use types. The model uses an algorithm to make this pixel per pixel calculation. This will now be explained with three equations. The first equation describes the probability of a cell being a certain land use type as the result of a relative importance of neighbourhood factors, non-neighbourhood factors and two other model settings:

$$P_{cj} = w_j P_{cjx} + (1 - w_j) P_{cjy} + C_j - A_{jkc} \tag{1}$$

Where:

P_{cj}	is the *probability* of cell *c* being land use type *j*.
P_{cjx}	is the *probability* of cell *c* being land use type *j* using neighbourhood factors *x*.
P_{cjy}	is the *probability* of cell *c* being land use type *j* using non-neighbourhood factors *y*.
w	is the *weight of the neighbourhood variables* for land use type *j* ranging between 0 and 1.
C_j	is the *conversion (in)elasticity* of land use type *j* for cell *c* ranging between 0 and 1.
A_{jkc}	is a *value indicating if a transition is allowed* between land use type *j* and land use type *k* in cell *c* taking a value of 0 or 5.

With the *weight of the neighbourhood variables* the importance of the neighbourhood variables on the final *probability* of cell *c* being land use type *j* can be set per land use category in a text file called *neighmat.txt*. The *conversion (in)elasticity* is a value that can decrease the probability of a land use transition on the cell that is land use type *j*. This (in)elasticity can be specified per land use. The name used for this value is *conversion elasticity* but a high value actually facilitates the conservation of land use *j*. Therefore the term *conversion (in)elasticity* will be used in this paper. The conversion (in)elasticities can be set with the file *main.1*. The *values indicating if a transition is allowed* can be set with the file *allow.txt*. This value can prevent certain land use transitions. For example the transition of urban fabric to pasture is considered to be very unlikely and can thus be prevented by specifying *allow.txt*. The *value indicating if a transition is allowed* can also be used to specify local impacts of certain policies like Natura 2000. Read the EuClueScanner tutorial by Koomen et al. (2010) for a more elaborate description of the files described above.

Equation 1 describes the final probability by using probabilities based on the neighbourhood and non-neighbourhood variables. Equation 2 describes how the probability on the basis of *neighbourhood factors* is defined.

$$P_{cjx} = e^{\beta_j + \sum \beta_{xj} * X_c} / \sum e^{\beta_k + \sum \beta_{xk} * X_c} \quad (2)$$

Where:

P_{cjx} is the *probability* of cell *c* being land use type *j* using neighbourhood factors *x*.

e is the basis of a natural logarithm.

β_j is the *intercept*. It describes the relation of the explanatory factors with land use type j.

β_{xj} is *beta-coefficient.* This describes the relation of explanatory factor *x* with land use type *j*.

X_c is the *value of an explanatory factor X* for cell *c.*

β_k is an *intercept* of other land use types than *j*. It describes the relation of the explanatory factors with land use types *k.*

β_{xk} is a *beta-coefficient* of other land use types than *j.* This describes the relation of explanatory factor *x* with land use types *k*.

The *intercept* and the *beta-coefficients* will be determined on the basis of a regression analysis explained later in *methods*. The *intercepts* and *beta-coefficients* of the non-neighbourhood factors can be adjusted with the text file *alloc2.reg*. The EuClueScanner will read this external file. Some of these *alloc* files that were used to calibrate the model can be found in annex 1. The *value of an explanatory factor* for a cell is determined by the neighbourhood factor maps similar tot the one in figure 3b. Equation 2 will give a value between 0 (not probable) and 1 (probable).

Equation 3 describes the *probability* of cell *c* being land use type *j* using non-neighbourhood factors *y.* Its components are similar to equation 2:

$$P_{cjy} = e^{\beta_j + \sum \beta_{yj} * Y_c} / \sum e^{\beta_k + \sum \beta_{yk} * Y_c} \quad (3)$$

It describes the probability of cell *c* being land use type *j* using non-neighbourhood factors *y*. The *intercepts* and *beta-coefficients* of the non-neighbourhood factors are obtained from a separate statistical regression to create best results. The *intercepts* and *beta-coefficients* can be written down in the text file *alloc1.reg* which is similar to *alloc2.reg*. The *value of an explanatory factor* for a cell are determined by the non-neighbourhood factor maps like the one in figure 3c.

The outcomes of equation 2 and 3 are implemented in equation 1 to create the probability maps. The probability maps are used to allocate the total demand and this results in a EuClueScanner model output.

METHODS

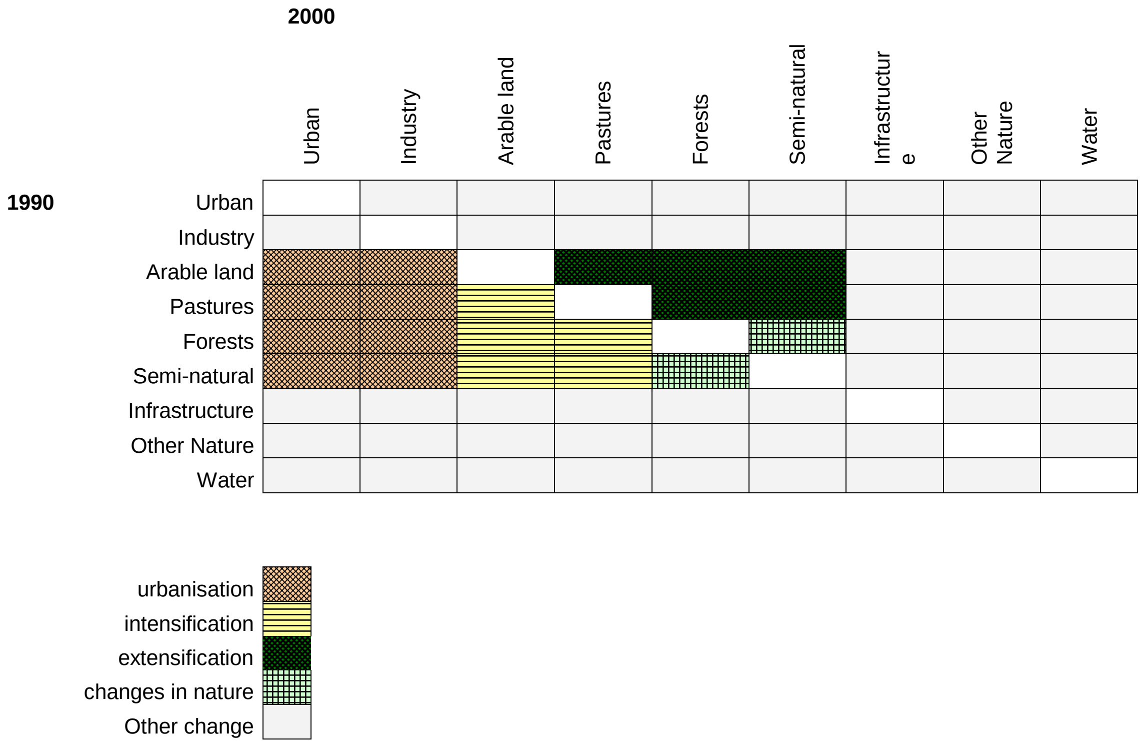
The EuClueScanner uses spatially explicit explanatory factors to create land use type probability maps. These probability maps are created on the basis of factor maps. These maps can be explained as variables that take a value dependent on its location. The EuClueScanner translates these values to the probability maps by multiplying the values with constants. Those constants are called the *beta-coefficients*. With a statistical method the *beta-coefficients* can be defined and enable the creation of the probability maps. With the probability maps a land use map can be created that is similar to the observed land use in 1990. The EuClueScanner however has some extra settings that can be used to improve the similarity of simulations with the observed land use like *conversion (in)elasticities* and *values that indicate if transitions are allowed*. The settings can be specified with the goal to find a calibration that best approaches observed land use or they can be adjusted with the intention to do other interesting findings. With these settings model results are created and they are compared with observed land use. To assess the power of the EuClueScanner in correctly simulating land use *change*, the modelled change between 1990 and 2000 is compared with the observed change. The period of 1990 to 2000 is chosen because of the available land use maps of these years.

This method will now be explained in more detail. The first step in this method is to analyse the land use changes between 1990 and 2000. Similar analyses with the EuClueScanner are performed by Diogo and Koomen (2010) and Pegels (2010).

Analysis of land use change

A transition matrix was created to get a quantitative overview of the land use and the changes in land use in Austria between 1990 and 2000. These changes are generalized into four distinctive *land use change processes*. The aggregation of the transitions into four transition processes is illustrated with the following table:

Table 3. Categorisation of land use change processes



The transformation of arable land, pastures, forests and semi-natural vegetation into urban fabric and industry is described as *urbanisation*. The transformation of pastures into arable land and the transformation of forest and semi-natural vegetation into pastures and arable land are described as agricultural *intensification*. The transformation of arable land into pastures and the transformation of arable land and pastures into forest and semi-natural vegetation are described as agricultural *extensification*. Transitions of forest to semi-natural vegetation and back are labelled as *changes in nature*.

The process *changes in nature* does not cover all the changes in nature because *other nature* is not modelled by the EuClueScanner. The land use change of all the static land use classes are classified as *other change*. The changes of urban fabric and industry to other land uses are modelled by the EuClueScanner but the process is not analysed as a separate land use change process because the apparent land use change is for a large part related to different categorising in 1990 and 2000. The difference in categorising was observed by validating the changes with areal photos. First a *transition map* has been created in order to get a spatial overview of the changes. To check the validity or the reason of land use transitions they are compared with areal images via Google Earth. Besides these land use transitions some of the locations of the land use change processes are observed on areal photos to get an idea of the nature of a transition.

Explaining land use

The calibration of the EuClueScanner is done on the basis of the land use in 1990. The land use of 1990 therefore serves as a proxy for suitability of land. It is for example no coincidence that urban fabric can be found in accessible valleys, that arable land is found on large fertile plains and forests cover the steep hill flanks.

To determine the influence of the available factor data on land use of 1990 many multinomial regression analyses were performed. First a selection of the available data is made which excludes insignificant or poor variables and choices are made between correlating variables. Correlating variables need to be left out in order to avoid incorrect assessments. Variables that had a Pearson correlation coefficient higher than 0.60 were considered correlated with an exception of *elevation* that has a correlation of 0.72 with *accessibility to cities* and a correlation of 0.68 with *slope*. An exception was made for elevation because of its strong explanatory value. These correlation coefficients are still acceptable.

With the multinomial regression an attempt is made to assess the influence of the available factors on the land use. Land use change however is the result of many different forces driving the decisions to change the land and an exact prediction is hard to establish. The *discussion* will be based on research of real driving forces of spatial allocation of land use change.

Modelling land use

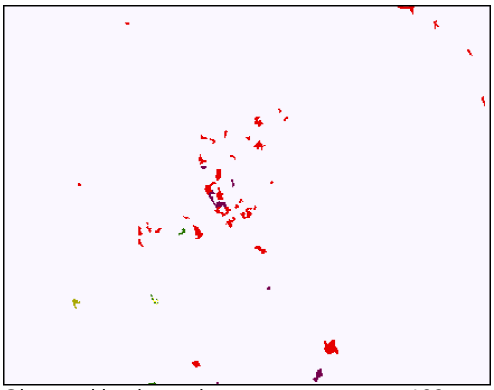
The beta-coefficients obtained with the regression analysis are used to specify model settings. To assess the spatial impact of the explanatory factors various model runs were performed. The starting state of the model runs is 1990 so for this year the land use for all the model runs is the same. Then the model will be run from 1990 to 2000. The demand for land uses in the model runs for 2000 is set equal to the actual demand of 2000. The *total* allocated land should therefore be almost the same as that of 2000. The only difference in the total allocated land is the change of the *static land use categories* between 1990 and 2000 that are not modelled by the EuClueScanner but in reality have changed. For example new infrastructure between 1990 and 2000 cannot be modelled by the EuClueScanner because this land use category is considered static. This leads to a small overestimation of total land use in the non-static land use categories, urban fabric, industry, arable land, pastures, forest and semi-natural vegetation.

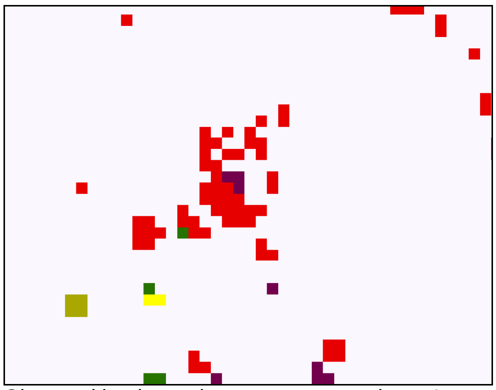
The modelled land use and the land use change were compared with the observed land use and land use change. A good way to observe the land use allocation purely on the basis of the explanatory factors is to change the *conversion (in)elasticity* and the *values indicating if a transition is allowed*. Model runs with high conversion elasticity and ‘allowing’ all transitions will be called *elastic*. To assess the predictive value of neighbourhood and non-neighbourhood factors separately the model had been run with different *weights of the neighbourhood variables*. In total, 26 successful model runs had been made on the basis of 9 multinomial regressions tuning these settings.

A numerical pixel per pixel comparison is performed between the model runs and the observed land use. Because the model is developed to estimate the spatial allocation of *future* land use changes the land use *conversions* are compared with the observed land use conversions too. To be able to do this transition maps had to be created. A pure pixel match of exactly the same land use conversions between 1990 and 2000, for example from forest to pasture, is very hard to establish on the 100 metre grid. To determine ill-predicted land use changes that were spatially close to the observed land use changes these simulated transitions were aggregated to a kilometre grid like illustrated in figure 4. First the conversions are generalised to the 5 *land use change processes*. If multiple transition processes took place on the same kilometre grid cell only one of those processes can be shown on the map so a choice had to be made between the transitions. The following order of importance was defined reflecting the anticipated impact on the landscape:

urbanisation > *intensification* > *extensification* > *changes in nature* > *other changes*

This means that if *urbanisation* takes place on a kilometre grid cell this always will be shown. *Other change* will only be shown if there are no other land use change processes in that kilometre grid cell. The aggregation processes is illustrated in the following figure:


Observed land use change processes at a 100 metre grid


Observed land use changes aggregated to a 1 kilometre grid


Simulated land use change processes by neighbourhood factors (on the same location)

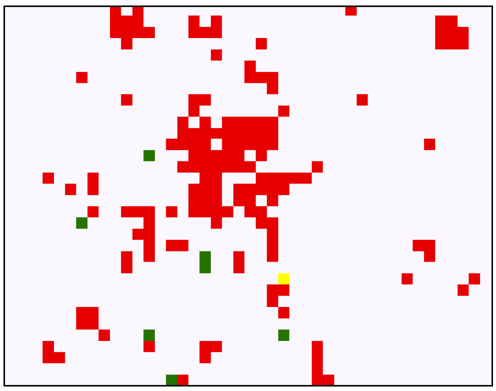

Simulated land use changes aggregated to a 1 kilometre grid

Figure 4. Illustration of the aggregation of land use change process

In addition to the *numerical* pixel-per-pixel comparison modelled land use maps and transition maps are also visually compared with observed land use and land use transitions.

RESULTS

In this chapter the land use changes between 1990 and 2000 will first be shown in a transition matrix. To obtain a clearer overview the transitions are aggregated to land use change processes and visualised in a transition map. Some locations of land use change are examined more closely in order to better understand the processes. Spatial patterns will be related to explanatory factors. These regression results are used to specify the model settings. The maps resulting from the model runs will be thoroughly examined and compared. The numerical pixel-per-pixel comparisons are assembled and some results are shown. In addition to the numerical results the modelled land use maps are assembled in annex *2*. In order to observe the country wide pattern of land use changes the transition maps of two model runs are shown together with a map showing the observed transitions. Finally the performance of two model runs in simulating urbanisation is illustrated with maps showing the observed and modelled urbanisation in the region of Vienna between 1990 and 2000.

Analysis of land use change

Almost 27000 ha, which is 0.32% of the surface of Austria, changed to a different land use class between 1990 and 2000. The land use transitions are shown in table 4. Urbanisation is the most important process driving land use change in Austria. Between 1990 and 2000 urban fabric increased with 7781 ha and industry with 2156, while arable land lost 8062 ha and pastures 3284 ha. 44% of the total amount of transitions can be considered urbanisation. 11% is related to agricultural extensification while only 3% of this loss on agricultural land is compensated elsewhere. 30% of the total change had been transformations from forest to semi-natural vegetation and back. Transitions from semi-natural vegetation to forest are often related to the regrowth of forests probably after been cut down for timber. A map showing the land changes is shown in figure 5. The size of the locations of change are exaggerated to a 1 kilometre resolution to give a clearer overview.

Table 4. Land use transitions in Austria between 1990 and 2000

1990 \ 2000	Urban	Industry	Arable land	Pastures	Forests	Semi-natural	Infrastructure	Other Nature	Water	Area	%
Urban	317455	78	2	8	6	0	32	0	0	317581	3.78%
Industry	234	15239	184	345	590	67	156	0	322	17137	0.20%
Arable land	4148	2635	1512278	842	522	163	312	0	74	1520974	18.12%
Pastures	2881	301	250	1191703	1311	143	12	0	43	1196644	14.26%
Forests	440	878	178	431	3751930	4322	11	117	55	3758362	44.77%
Semi-natural	185	162	0	15	3669	582875	82	6	0	586994	6.99%
Infrastructure	0	0	0	0	0	0	5567	0	0	5567	0.07%
Other Nature	19	0	11	0	105	129	0	920907	82	921253	10.97%
Water	0	0	9	16		6	0	0	69713	69744	0.83%
Area	325362	19293	1512912	1193360	3758133	587705	6172	921030	70289	8394256	
%	3.88%	0.23%	18.02%	14.22%	44.77%	7.00%	0.07%	10.97%	0.84%	100.00%	

urbanisation	11630	44%
intensification	874	3%
extensification	2981	11%
changes in nature	7991	30%
Other change	3113	12%
Total change	26589	

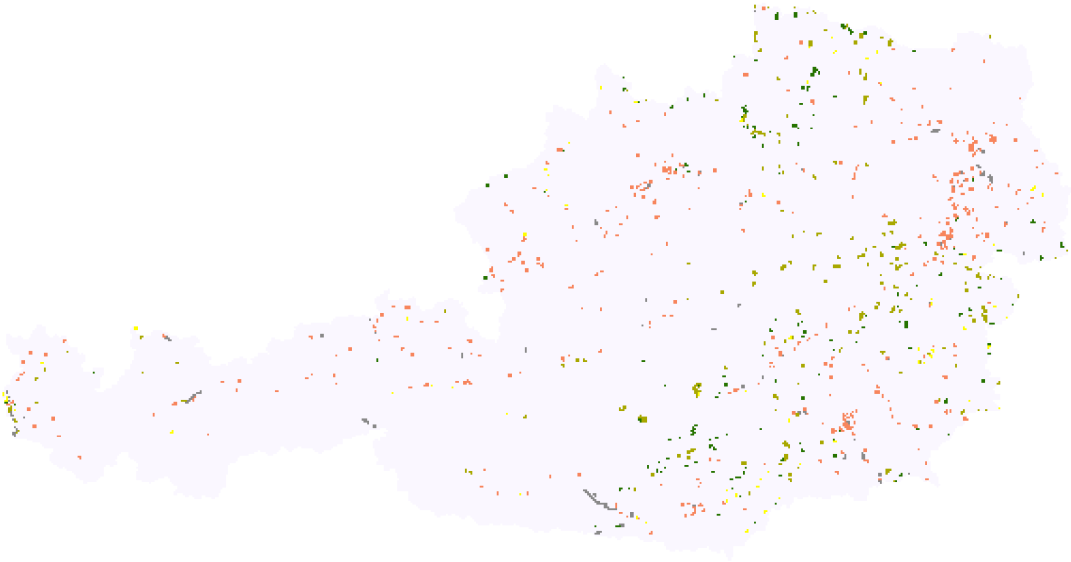


Figure 5. Transition map showing land use change processes in Austria between 1990 and 2000. (for legend see Table 3 or 4)

Most *urbanising* on the land use maps happens at the borders of current urban land use and close to some of the bigger cities. A large amount of the urbanisation can be observed around Vienna, Graz and Linz. The urbanisation around Salzburg is of a whole different character. There the most important urbanisation process was the opening of multiple golf courts which are also classified as urban fabric. Most urbanisation between the land use map of 1990 and 2000 happens in a blocky pattern at the edge of a city and normally not further than about 1 or 2 kilometres of former urban land. Urbanisation often takes place exactly in agricultural parcels as shown in the following figure:

Figure 6. Illustration of an actual urbanisation pattern northeast of Vienna (Screenshot of Google Earth)

Agricultural *intensification* only happens on a very small scale in Austria. Of the 0.3% land use change in Austria only 3% is categorised as agricultural intensification. The observed intensification consists for the largest part of very local cuttings of forest like the transition matrix (table 4) shows. Some agricultural *extensification* has taken place in Austria and mainly consists of the conversion of pasture and arable land into forest. The *changes in nature* are probably the result of forest cuttings for timber and the regrowth of trees. Such land use transitions occur in the central highlands of Austria. The high Alps in the west are apparently not used for timber.

Explaining land use

A lot of regressions were performed in an attempt to explain land use as good as possible with the available data. These regressions are used to explain the land use of 1990. The results will later be used to calibrate the model.

Table 5 shows the regression results of neighbourhood factors. The coefficients show the positive or negative relation with the neighbourhood factors compared to the reference category forest. Intuitively all the land uses are expected to be found in the proximity of the same land use and therefore show a positive beta-coefficient with the neighbourhood factors of their own land use. The positive values of land use classes with neighbourhood factors of other land use classes however show that certain land uses are also expected to be close to other land use types.

Table 5. Beta-coefficients of neighbourhood variables per land use category

Factor	Land use class				
β	Urban	Industry	Arable land	Pastures	Semi-natural
Intercept	-2.350	-3.432	-2.597	-2.235	0.161
β					
urban	0.136	0.018	0.041	0.035	-0.029
forest	-0.071	-0.082	-0.045	-0.039	-0.082
industry	0.028	0.204	0.039	0.018	-0.041
pasture	0.011	-0.026	0.045	0.094	-0.013
arable	0.021	-0.004	0.103	0.044	-0.002
$semi-natural$	-0.072	-0.049	-0.008	0.006	0.091

The reference category is Forest. These regression was performed by Vasco Diogo (Significance and pseudo R-squares not known)

The regression results of the non-neighbourhood factors are shown in table 6. As shown in equation 3, the influence on probability of land use *j* by factor *y* in a cell is dependent on the *beta-coefficient* of factor *y* for this land use multiplied with the *value* of factor *y* in a cell. The possible values of the factors *y* differ greatly as shown earlier in table 2 and these factor descriptives are also added to table 3. In combination with the beta-coefficients they give a better indication of the importance of the relation of a factor with a land use type.

Accessibility to cities, *elevation* and the *slope* prove to be very important in determining the spatial allocation of land use change in Austria. The strong relation between land use change and the explanatory factors *slope* and *accessibility to cities* is illustrated in figure 7. Besides the relation with land use *change* the correlation of these factors with the current land use is also very evident. Compare figure 2 and figure 7 to observe the relation between land use and slopes. Clearly the cities lay deep in the valleys. Arable land can mainly been found in the flat lands of Upper Austria, Lower Austria and

Table 6. Beta-coefficients of non-neighbourhood variables per land use category

Factor	Factor descriptives		Land use class				
β			Urban	Industry	Pastures	Arable land	Semi-natural
Intercept			5.257	4.289	3.463	2.326	-8.320
β	Min	Max					
AccessCity	0	706	-0.019	-0.036	0.003	-0.002	-0.005
AccessPort	99	876	-0.004	-0.005	0.002	-0.002	0.006
SoilWater	30	220	-0.002	-0.007	0.007	0.001	0.002
Elevation	108	3666	-0.003	-0.004	-0.005	0.000	0.005
Slope	1	6	-1.284	-1.105	-1.238	-0.727	-0.676
ImpermeableLayer	0	1	-0.558	-0.921	-0.442	-0.066	0.228
Natura2000	0	1	-1.559	-1.452	-1.131	-0.805	0.552
SouthSlope	0	1	0.504	-0.222	0.401	0.615	0.418

The reference category is Forest. All variables are significant.
Pseudo R-squares: Cox and Snell; 0.613, Nagelkerke; 0.660, McFadden; 0.361

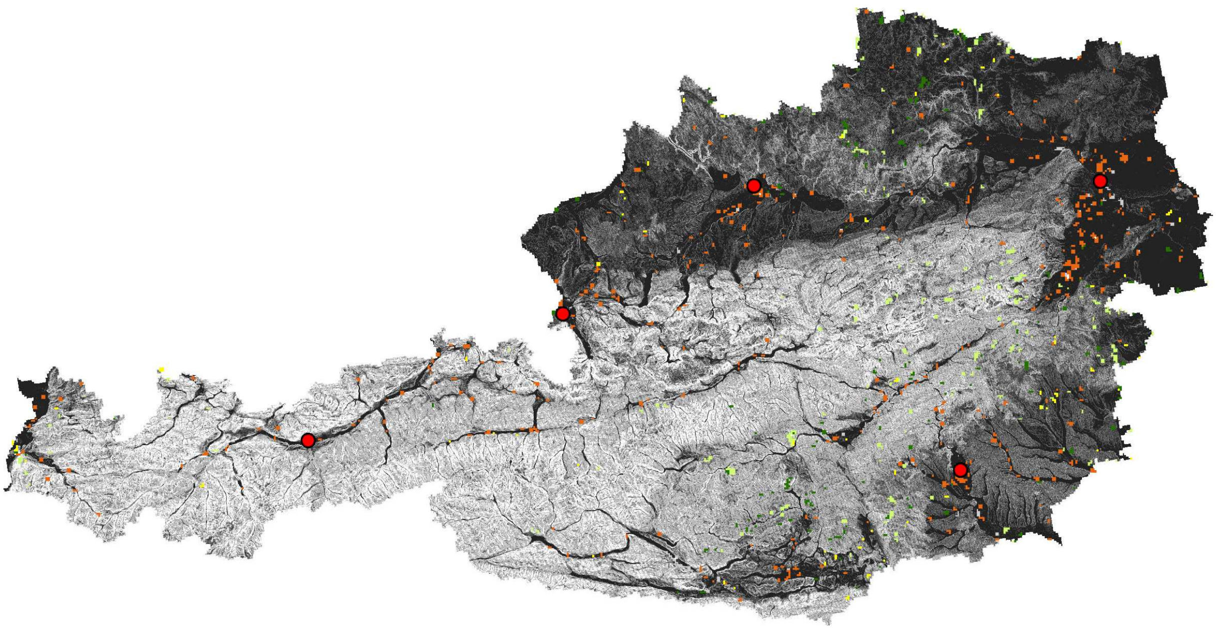


Figure 7. Overlay of the slope factor map, the location of the 5 cities with more than 100 000 inhabitants and the transition map. The values in the slope factor map ranges from class 0 in black (flat), to slope class 6 in white (steep). The 5 cities are indicated with the red dots. (For the legend of the land use change processes see table 3 and 4)

Burgerland. Pastures lay on the gentle slopes, forests on the steeper slopes and semi-natural vegetation higher on the higher lands. The tops of the highest mountains are covered with glaciers and snow and are categorised as *other nature*.

In total 9 different regression results were used for the model runs of which 5 are described in this paper. An overview the *alloc* files that were used is shown in annex 1. Neighbourhood and non-neighbourhood factors were used as dependent variables in separate and combined regressions. The combined regression resulted in the regression which will be called *combined.* Such regressions however deliver unrealistic results. Because the fine-grained neighbourhood factors adequately capture the clustered occurrence of land-use patterns they are very strong explanatory variables and they tend to overshadow the often much coarser non-neighbourhood variables leading to unreliable beta estimates. This strong spatial autocorrelation is a known issue in land-use change analysis (Overmars et al., 2003). For example *accessibility to cities* takes a positive value in the combined regression, (while it has a negative value in regression without neighbourhood variables) meaning that urban fabric is likely to be located where the distance from the cities increases. Apparently the distance to cities serves as a kind of balancing factor for the neighbourhood factor *urban fabric* that allocates urban fabric in the neighbourhood of the cities. The regression called *Vasco* is the original regression result that is used to calibrate the EuClueScanner. In the regression *elevation* the factor elevation is added to improve the allocation of semi-natural vegetation in the model run. Although *elevation* gives good model results, the factor *accumulated rainfall from March to July* correlated with *slope* and *elevation*. To remove this correlation the regression *final* was performed. The factor elevation in *final* however still shows some degree of correlation with slope and *accessibility to cities* as explained in *methods.* In Austria natural factors like *slope* and *elevation* play an important role in the spatial allocation of land use. To assess the explanatory power of natural factors the regression *natural* was also performed.

Modelling land use

The effect of explanatory factors can be observed with the land use maps created by the EuClueScanner. To asses the explanatory power of the factors alone, multiple *elastic* model runs were performed. Seven of these elastic model runs are listed in table 7. In addition eight *inelastic* model runs are listed that give model results that are much closer to the observed land use of 2000. These 15 model runs in total are based on just 5 different regression results but the model results change because of adaptations in the *neighbourhood-non-neighbourhood variable* ratio (*w),* the *conversion (in)elasticities* (*C)* and the *values indicating if a transition is allowed* (*A*).

One of the model results is shown in figure 8. In this example the model settings were *elastic*. As a result of these setting a lot of land has been reallocated which is in fact unrealistic but gives some clear insights in the performance of explanatory factors. The results in table 7 show that the *percentage correctly modelled* is much lower for the elastic model runs than for the inelastic model runs. The land use maps resulting from the elastic model runs can be found in annex *2*. Compare them with close attention to observe the influence of the factors on the spatial allocation of land use. Because these *elastic* model runs reallocate much land, the chance that they correctly model land use *change* is much larger. This can be observed in the column *percentage of change correctly modelled* in table 7.

The difference of the correctly modelled land between the elastic model runs ranges between 63% for two model runs and 77% for a model run on the basis of neighbourhood variables. Regression *combined* shows that combining neighbourhood and non-neighbourhood variables in one multinomial regression does not deliver good results. Three elastic model runs were performed with the regression results of *Vasco*. Leaving out the non-neighbourhood factors (so setting w = 1) gives the best results and correctly models 77%. Attempts to improve the explanatory value of the non-neighbourhood factors resulted in the regressions *elevation, natural* and *final.* They indeed increased the percentage

Table 7. Model results

	Regression	nr	w	C	A	% correctly modelled	% of change correctly modelled
Elastic model runs	*combined*	1	diff	+	-	63%	24%
	Vasco	2	0	+	+	65%	20%
		3	1	+	+	77%	14%
		4	0.7	+	+	69%	20%
	elevation	5	0	+	+	70%	22%
	natural	6	0	+	+	69%	19%
	final	7	0	+	+	69%	21%
Inelastic model runs	*Vasco*	8	0.3	-	-	99.25%	0.5%
	elevation	9	0.3	-	--	99.33%	0.7%
		10	0.7	-	--	99.35%	1.0%
		11	1	-	--	99.35%	1.8%
	natural	12	0	-	-	99.30%	0.2%
		13	0.8	-	-	99.35%	0.9%
	final	14	0	-	-	99.18%	0.5%
		15	0	--	-	99.34%	0.3%

nr: model run number.
w: weight of neighbourhood variables. 0: only non-neighbourhood variables. 1: only neighbourhood variables. 0.3, 0.7 and 0.8: partly neighbourhood and partly non-neighbourhood variables. diff: a mix of different weights per land use type.
C: conversion (in)elasticity. +: elastic (0). -: inelastic (0.9). --:very inelastic (1.5).
A: allow conversions. +: allow all. -: don't allow from urban. --: don't allow from urban and industry.
% correctly modelled based on a pixel per pixel comparison with observed land use in 2000.
% of change correctly modelled based on a pixel per pixel comparison of land conversions from 1990 to 2000.

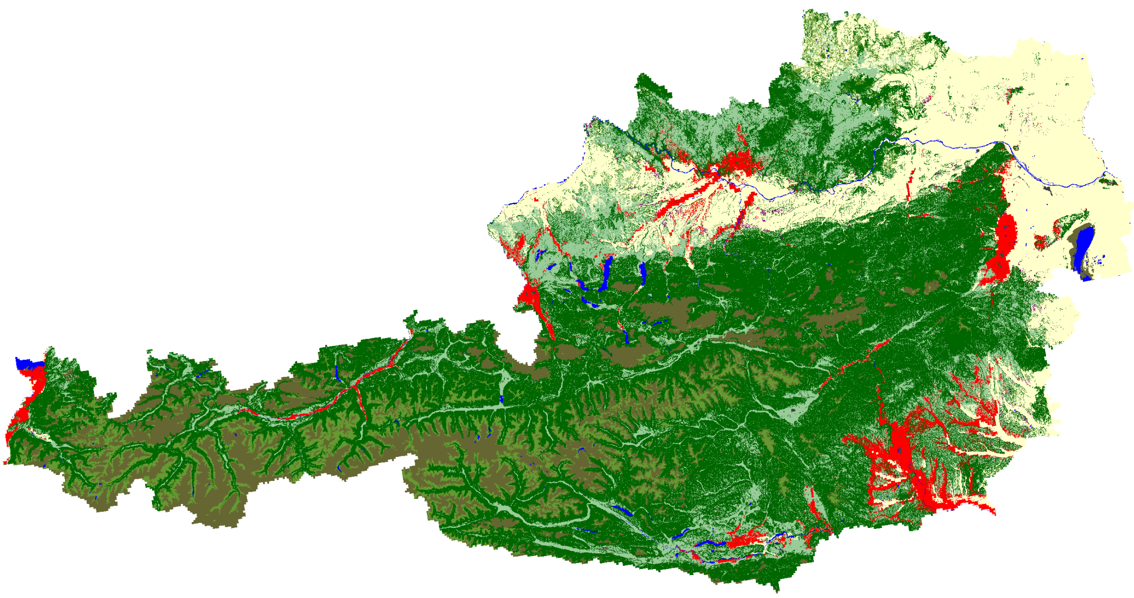


Figure 8. Model run *nr 7* (for legend see table 1)

correctly modelled land use from 65% (*nr* 2) to 69% and 70% (*nr* 5, 6 and 7). Even with less explanatory factors, simply because *elevation* and *accessibility to ports* were included in these regressions. This however was not enough to be able to improve the model results that are based on solely neighbourhood factors.

Let us look at the elastic results in more detail in table 8. Here the performance per land use category of two elastic model runs is shown. On the left is the model run on the basis of neighbourhood factors (*nr* 3) and on the right the model run with non-neighbourhood factors that performed best in land use allocation (*nr* 5). Although model run *nr* 5 performs well in modelling semi-natural vegetation forest and arable land, the factors are not precise enough to predict relatively scarce land uses like urban fabric and industry.

Table 8. Correctly simulated land use change per land use category in two model runs.

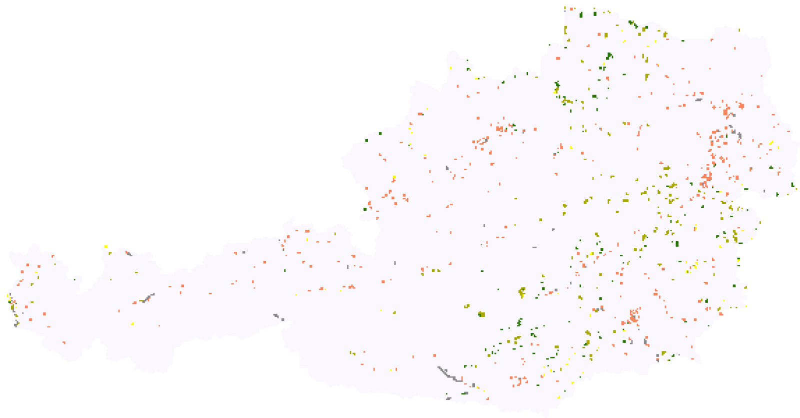
Neighbourhood factors (model run *nr* 3)		Non-neighbourhood factors (model run *nr* 5)	
Land use	Correct	Land use	Correct
Urban fabric	65%	Urban fabric	23%
Industry	41%	Industry	0%
Arable land	68%	Arable land	66%
Pastures	54%	Pastures	38%
Forest	83%	Forest	78%
Semi-natural	78%	Semi-natural	72%

Take a look at table 7 again. In the *inelastic* model run comparison, the original calibration with the regression results of *Vasco* and a default *neighbourhood factor weight* of 0.3, was compared with other model specifications. Some settings gave better results than *Vasco* but none of them were outstandingly good. The part of correctly simulated land in this model runs is very high because most of the land uses are retained since the starting state of the model run (1990). Therefore the most important conclusions should be drawn from the percentage of change that is correctly modelled as shown in the right column. Although some attempts are made to include non-neighbourhood factors in the model runs, the best model results are made with solely neighbourhood factors. The neighbourhood factors correctly model 1.8% of the observed change between 1990 and 2000 in model run *nr* 11. Although this is the best result, 1.8% still seems a poor prediction. The model runs based on the regression *elevation* perform best in land use and land use change simulation based on non-neighbourhood factors, but because of the large correlation in the factors of *elevation, final* will be considered as the best model run based on non-neighbourhood factors.

So far for the exact match. How does the inelastic model run based on neighbourhood factors (*nr* 11) and the best inelastic model run based on non-neighbourhood factors (*nr* 14) perform more generally? When analysing the processes aggregated to a 1 kilometre grid, the model run based on neighbourhood variables is able to approximate 15 percent of the land use change. This is based on a pixel per pixel comparison of the transition maps shown in figure 9. The model run based on non-neighbourhood variables is able to approximate 5 percent of the land use change. An attempt to further generalise the land use simulation to a 5 kilometre grid did not produce good results. Next to the maps in figure 9, the modelled and observed change per ha (on a 100 metre grid) are shown. The total change simulated by the models (figure 9b and 9c) seems low compared to the large number of land use change process locations on the map when comparing it with the observed land use change process map (figure 9a). The *modelled* changes however are more dispersed than the *observed* land use change processes. This is clearly because the minimum mapping unit of 25 ha on the land use maps of 1990 and 2000. Only large chunks of change or changes in these chunks will be observed as change on these maps. The EuClueScanner however does *model* smaller changes in land use.

a) Observed land use change processes in ha

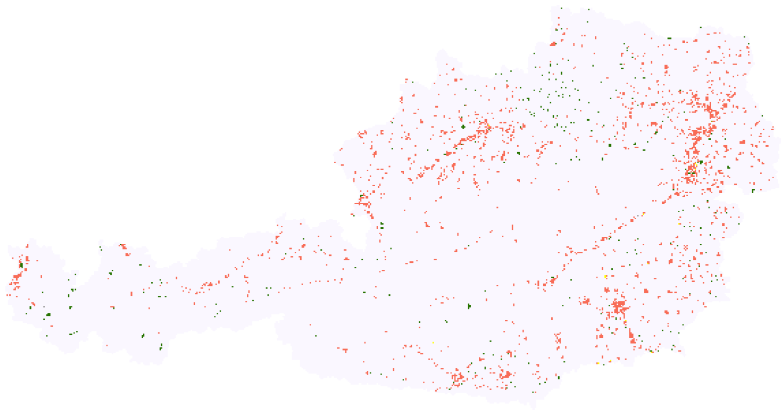
urbanisation	11630
intensification	874
extensification	2981
changes in nature	7991
Other change	3113
Total change	26589



---

b) Simulated with model run nr 11 in ha

urbanisation	9560
intensification	120
extensification	1058
changes in nature	2
Other change	0
Total change	10743



---

c) Simulated with model run nr 14 in ha

urbanisation	8033
intensification	1606
extensification	6683
changes in nature	1357
Other change	0
Total change	17679

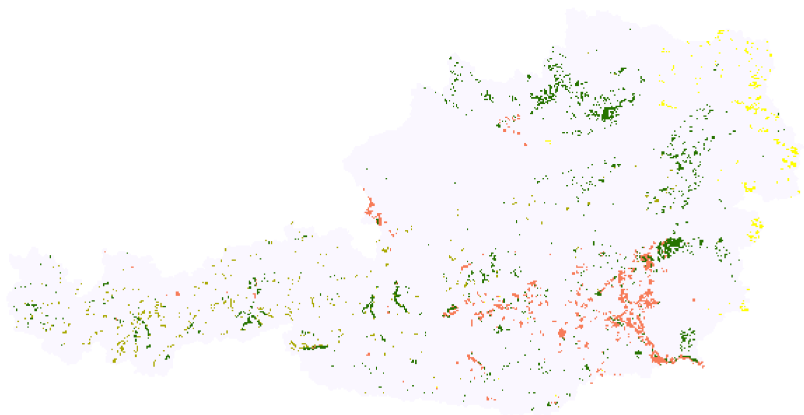
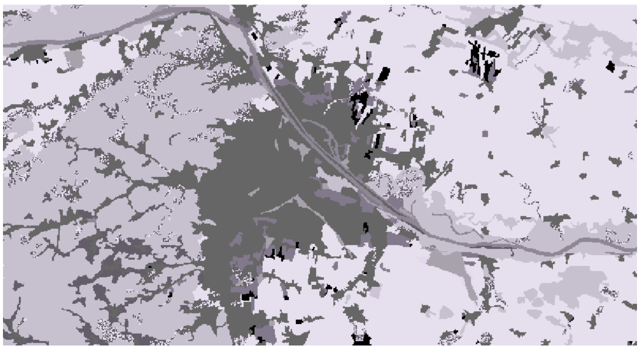


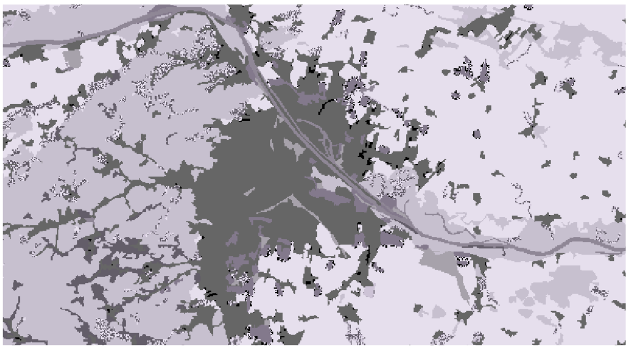
Figure 9. Transition maps showing observed and simulated land use change processes with a 1 km resolution in.

Urbanisation is the most important land use change process and some urbanisation simulations will now be compared to observed processes in further detail. On the basis of the neighbourhood factors (*nr* 11) the EuClueScanner was able to predict 5% of urbanising pixel-per-pixel. Aggregating the processes to a 1 kilometre grid enables the neighbourhood factor based model run to predict 33%. Most observed urbanising on the land use maps happens at the borders of current urban land use and close to some of the bigger cities. Especially Vienna was growing rapidly between 1990 and 2000.

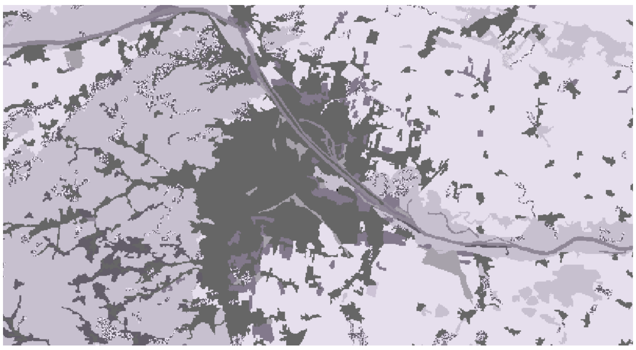
Figure 10a shows urbanisation between 1990 and 2000 surrounding Vienna. Figure 10b and Figure 10c show simulations of land use changes. The urbanisation happens in a blocky pattern following the agricultural parcels. The neighbourhood factors are able to approximate the changes but the *non*-neighbourhood factors do not predict any urbanisation in Vienna.



a) Observed urbanisation



b) Simulated urbanisation with neighbourhood factors (model run *nr* 11)



c) Simulated urbanisation with non-neighbourhood factors (model run *nr* 14)

Figure 10. Urbanisation between 1990 and 2000 in the region of Vienna. Urbanisation is shown in black.

DISCUSSION

The EuClueScanner is under development as a tool to enable policy makers to observe expected changes in land use as the result of pursued policy. How should the EuClueScanner be improved in order to be able to fulfil this task? In the following section the performance of the EuClueScanner will be analysed and discussed. In this paper the independent variables which are used to predict land use are constantly called the explanatory factors. But are these really the factors explaining land use? In the following part the real factors driving land use allocation will be analysed per land use class and a comparison will be made with the current explanatory factors.

Urban fabric

Houses together with public urban areas form the land use class urban fabric. The *elastic* model results (see figure 8) based on non-neighbourhood factors show that the EuClueScanner has great difficulties modelling urban land use. Especially the non-neighbourhood factors dramatically failed to properly allocate urban fabric. The factor *accessibility to cities* only defines the location of 5 cities in Austria. Another issue is that statistically the probability of finding arable land on the lower grounds is larger than the probability of finding urban fabric. The reason might be found in the fact that urban fabric is also found in the higher valleys in the Alps while arable land can largely be found outside the alpine region. Because of this the cities are replaced to higher grounds which is of course not realistic.

In reality the location choice of housing is determined by factors like access to job centres environmental amenities, clean air, scenic views and preserved natural habitat and recreation opportunities including access to parks and open space and the presence of nearby retail and service facilities (Kim et al., 2005). According to Brian Arthur (1988) growth of cities is largely determined by the presence of industry.

Is one of the factors described above available in the Corine land use data? The only factor that is available at the moment is the location of chunks of industry. At the moment these data can only be used as a neighbourhood factor. According to Kanatschnig and Weber (1998) the average travel distance to work in Austria is between 20 and 30 kilometres so the explanatory value of the neighbourhood factor industry on the existence of urban fabric is very limited. These neighbourhood factors only reach 3 cells or 300 metres. A similar land use change model, The Land Use Scanner, already uses factors that describe the number of jobs in the proximity of a certain cell. The Land Use Scanner is described by Hilferink and Rietveld (1998). The factor would greatly improve the ability of the EuClueScanner to simulate the existence of urban areas. Urban fabric should be proportionately distributed with the number of jobs in a reasonable travel distance of the workplaces.

There is already a model under development that tries to model location decisions of housing and firms based on the factors influencing location choice. This model does not only predict land use change itself but it predicts land use change as a result of decisions made by individual agents. These agents have predefined preferences and make decisions by weighing the housing location suitability factors. This decision making process is very detailed and at the moment it will be impossible to model it in this way on the scale of a country like Austria. The area of study of this model, however, happens to be a small part in the Rhine valley of Austria. It is described by Loibl et al. (2007).

Industry

The EuClueScanner was also not able to accurately allocate firms using the available non-neighbourhood factors. So what drives the location decisions of firms in reality? Production facilities are influenced by the distance to the market what could result in spatial clustering on certain locations. Other types of firms cluster in order to obtain agglomeration economies. Finally the government uses zonal regulations to prevent extensive urban sprawl (McCann, 2001).

Firms however also diverge over the country. For example mining and quarrying is highly dependent on the physical environment. In oligopolistic environments firms will move away from each other. In other markets the location of the company is not important. Finally there are in contrast to agglomeration economies, agglomeration diseconomies like traffic and high land prices on desirable locations. These diseconomies have a centrifugal effect (McCann, 2001).

Although firms scatter over the country the EuClueScanner has to rely on the clustering processes that are observed in location choice behaviour of firms because other locations are dependent on some many factors that cannot be taken into account by the model. Some suggestions for improvement of the current explanatory factors will be made in the *recommendations.*

Arable land and pastures

The EuClueScanner was able to model the rough contours of arable land quiet well with approximately 65% correctly simulated for the *elastic* model runs. *Slope* and *soil water available to plants* happened to be important explanatory factors. In addition of slope and soil water available to plants the *type and deepness of soil* might be other important factors describing the agricultural value of land.

Only 36% of pastures was correctly simulated with the non-neighbourhood factors. This result is not very impressive but the overall distribution of pastures is well simulated as could be seen when comparing model run *nr 7* and real land use in annex *2*.

A technical difficulty in simulation land use is the scattered pattern of pastures due to the categorising of land use types in the CLC-data as desribed in the chapter *EuClueScanner* and illustrated in figure 1. The difficulty to predict pastures and arable land with the used factors might also have a non-technical explanation. There is a large degree of government intervention in Austria which declines the importance of geographical factors on land use. Although an important part of Austria is still used for agricultural purposes, the extreme topography and hard subsoil makes the largest part of Austria relatively unsuitable for agriculture uses. Technological improvements increased productivity and focused productivity on more fertile and accessible land after the Second World War (MacDonald et al. 2000). However, since the seventies of the twentieth century the farmers receive direct payments from the government to be able to maintain themselves. Conservation of cultural landscapes is the main argument for these subsidies (Tasser 2005). The subsidies are part of the Common Agricultural Policy. This classifies rural areas on the basis of certain factors like altitude and slope. According to this classification 70% of the agricultural land nowadays is ‘disadvantaged’. 21% of arable land and 85% of pastures is located in mountainous areas. The grants are inversely proportional to the suitability of the agricultural land. This means that payments are higher if the land is less suitable (Schneeberger, 2003).

For the model this means that correctly simulating land use changes in arable land and pastures is difficult. Luckily the policy also preserves the land as it is, so not much land use transitions should be expected as long as the policies are preserved. New political parties could however change such policies. In that case a land use allocation that is more proportionate to their suitability as described by the non-neighbourhood factors can be expected.

Forest and semi-natural vegetation

Without human interference most land will change into forest. In Austria the rougher landscapes are covered with forests. When climate gets hostile enough forests are replaced by smaller vegetation classified as semi-natural vegetation. When the climate gets even more hostile for plants the alpine peaks, bare rocks and snow cover will be exposed. Altitude resulted to be an important explanatory factor for semi-natural vegetation. Including this factor in the regression improved the prediction of semi-natural vegetation from 44% to 72%. Semi-natural vegetation will often be found on higher grounds than forest. More surprising is that *forest* is expected to be on the steeper slopes according to the regression results. This is could be explained by the fact that forests occupy the mountain flanks while semi-natural areas can be found high on the mountains. Meanwhile, the lower flatter grounds are covered with arable land, pastures and urban fabric.

RECOMMENDATIONS

An important goal of this research is to provide some recommendations on the basis of the observed land use model results in Austria. The recommendations are based on the *results* and the *discussion*.

Only use neighbourhood factors or improve data

When there are no plans to improve the non-neighbourhood variables it is best to leave them out as explanatory factors and only use neighbourhood factors to simulate future land use change. When the quality of the non-neighbourhood variables *do* increase the EuClueScanner should be recalibrated again to find out the new explanatory value of these factors.

Simulate on the basis of land use change processes

Although the current non-neighbourhood factors of land use change are able to simulate land use quiet well, the performance on land use *change* in Austria is not yet promising. The model was not able to make good simulations on both a pixel to pixel comparison and the country-wide pattern (see figure 9c). One reason is the relatively short period of 10 years. The match with the observed change might increase with the years. However, a very important other reason seems to be that the land use changes between 1990 and 2000 are largely independent of probabilities based on the distribution of land uses in 1990. A calibration on the basis of *change processes* might give better results.

From pure pixel determination to proportions

The EuClueScanner allocates pure pixels. The result is that only a very small part of the land use change will exactly be predicted. It might be better if the EuClueScanner is able to assess a chance of land use change per pixel. A consideration needs to be made about the technical possibilities. One possibility might be to allow a cell to contain more than one type of land use. The cell can than be divided in percentages of the land uses categories. The cells will describe the relative proportion of the land use types present in the cell. This method is also used in the *Land Use Scanner*. The model is described by Hilferink and Rietveld (1999).

Improving the simulation of urbanisation

When factors *are* improved this creates some opportunities for improvement of land use change simulation. Most improvement can be made in the urbanisation process. It is the most frequently occurring land use change process and it mainly occurs on the edges of present urban fabric making it relatively easy to predict.

Although the neighbourhood factors explain land use change best as shown by the comparison with the observed change between 1990 and 2000, they do not account for important geographical factors influencing the suitability of a certain location for land use change. Including these geographical factors however proved not to improve the simulation results. I imagined some ways to include these factors but finally I would not suggest adding them to the model. The steep slopes are mainly covered with forest. The valleys are normally covered with urban fabric, pastures and arable land. The chance of finding urban land or industry in the neighbourhood of arable land or pastures is much more probable than finding urban fabric or industry in the neighbourhood of forest or semi-natural vegetation like shown in the beta-coefficients. Therefore urbanisation simulated on the basis of neighbourhood factors will be more likely in the valleys than in on the steeper slopes.

The current neighbourhood factors have a drawback though. They only predict land use change in a distance of 300 metres of a city because this is the range of a neighbourhood factor. The neighbourhood variables will predict urbanisation in locations surrounded by urban fabric. The observed urbanisation however does not necessarily takes place on these locations. Much urbanisation occurs between the edges of the city to approximately 1 to 2 kilometres away of the edges. This distance however reaches too far for the neighbourhood factors. An adjustment of the neighbourhood factors could solve this distance problem. A flat neighbourhood factor with all rings having the same weight might already do the job. Neighbourhood factors reaching for example 1.5 km however become huge. It might slow down the calculation process. Therefore I would recommend further research to consider the best technical solution. In addition of this neighbourhood factor two extra variables might predict the polarised urbanisation on a country-wide scale:

- a factor that describes the number of jobs on an average travel distance of around 20 to 30 km
- a factor that describes the number of other households of around 20 to 30 km

This average travel distance is based on the findings of Kanatschnig and Weber as described in the *discussion*.

Most industry is located next to current industrial zones or they are developed on completely new zones on relatively unpredictable locations. Urbanisation in general normally takes place in a blocky pattern clearly related to agricultural parcels. The parcel polygons are probably hard to include in the raster-based EuClueScanner but a prediction of urbanisation with land use change probabilities equal per parcel or per owner might give realistic and useful results. This might however be beyond the scope of the EuClueScanner.

The previous suggestions and points of attention lead to the following conclusions.

- The transitions to urban fabric could, with all improvements, be based on the neighbourhood of *urban fabric*, the number of jobs in an average travel distance, the number of households in an average travel distance and the chance of urbanisation could be evenly distributed over a parcel or land use owner.
- The transition to industry and could than mainly be based on the neighbourhood of *industry*, the number of jobs in an average travel distance, the number of households in an average travel distance and the chance of urbanisation could be evenly distributed over a parcel or land use owner.

SUMMARY AND CONCLUSIONS

In this research the EuClueScanner was calibrated and the modelled results were validated with the observed land use in 1990 and 2000. Only a small part of the Austrian land (0.3%) changed to another land use type. The demand change in this period is used as input for the EuClueScanner and it performs a simulation on the basis of neighbourhood and non-neighbourhood explanatory factors. Although both types of factors are able to explain and simulate the land use, the neighbourhood factors perform best in land use *change* simulations in Austria with 1.8% of the land use changes correctly simulated based on a pixel per pixel comparison on a 100 metre grid. Aggregating these changes to land use change processes on a 1 km grid increased the match to 15% for a model run based on neighbourhood variables. The addition of non-neighbourhood factors to the model always had a negative impact on the model results.

An important reason for the mismatch of non-neighbourhood factors on land use change might be the fact that future land use changes are largely independent on the current land use allocation. The non-neighbourhood factors are calibrated to simulate the land use of 1990. The current land use however does not explain current *processes* of land use changes. Some land use change processes are very hard to predict in Austria. The current policy for example aims to preserve the agricultural land uses. The result is that economic forces do not drive agricultural land use changes. The few changes that do occur are very randomly distributed over the country. Also the exact locations of forest clearings and regrowth are hard to simulate, but a global simulation should be possible. The impact of this back and forth transitions are however much smaller than the almost irreversible land use change process urbanisation.

Most improvement for land use change simulation can be made in the urbanisation process. With 44% on the total change is urbanisation the most important land use change process in Austria. Urbanisation mainly occurs on the edges of present urban fabric making it relatively easy to predict. The extension patterns of urban fabric and industry are often related to the shape of agrarian parcels. The urbanisation process is not evenly spread over the country but occurs mainly in already urban regions. Adding and adapting some explanatory factors might improve the model results.

When there are no plans to add or adapt these factors it is suggested to leave out the non-neighbourhood variables as explanatory factors. To improve the usefulness of the simulation it should also be considered to change from pure pixel determination to land use proportions per cell representing the probabilities of finding a certain land use.

ACKNOWLEDGEMENTS

To be able to conduct the research *Eric Koomen* explained me the context in which this research takes place. Because the there were many technical difficulties I am very grateful for his accessibility and quick mail replies. Eric Koomen also thoroughly read my report and gave some very constructive comments to improve the readability of the paper. *Vasco Diogo* has conducted similar studies and could help me out with very specific questions. I also frequently used the results of a multinomial regression that he had performed because I was not able to reproduce it. *Maarten Hilferink* found a way to export maps with the unfinished EuClueScanner which enabled me to continue the research.

REFERENCES

Arthur, W.B. (1988) Urban systems and historical path-dependence. Stanford university, Stanford

Koning, G.H.J. de, A. Veldkamp, P.H. Verburg, K. Kok and A.R. Bergsma (1998) CLUE: A tool for spatially explicit and scale sensitive exploration of land use changes. In: Stoorvogel, JJ., Bouma, J., Bowen, W.T. (eds.), Information technology as a tool to assess land use options in space and time. Proceedings of an international workshop, Lima, Peru. 1997. Quantitative Approaches in Systems Analysis no 16.

Diogo, V., Koomen, E. (2010) Explaining land-use changes in Portugal 1990-2000. In: Painho, M., Santos, M.Y. and Pundt, H. (eds.) Geospatial Thinking; Proceedings of AGILE 2010, Guimarães, Portugal.

Hilferink, M., Rietveld, P., (1998) Land Use Scanner: An integrated GIS based model for long term projections of land use in urban and rural areas. Journal of Geographical Systems 1: 155-177

Kanatschnig, D., Weber, G. (1998) Nachhaltige Raumentwicklung in Österreich. On behalf of the Bundesministeriums für Umwelt, Jugend und Familie, Österreichischen Institut für Nachhaltige Entwicklung, Vienna.

Kim, T., Horner, M.W., Marans, R.W. (2006) Life Cycle and Environmental Factors in Selecting Residential and Job Locations. Housing studies 20: 457-473

Koomen, E., Diogo, V., Hilferink, M., Beek, M. (2010) EU-ClueScanner100m; model description and validation results. Commissioned by: European Commission DG Environment. Project reference: ENV.G.1/SER/2008/0050. Vrije Universiteit / Object Vision, Amsterdam.

Koomen, E., Loonen, W., Hilferink, M (2008) Climate-Change Adaptations in Land-Use Planning; A Scenario-Based Approach. Springer, Dordrecht.

Loibl, W., Tötzer, T., Köstl, M., Steinnocher, K. (2007) Simulation of polycentric urban growth dynamics through agents. In: Koomen, E. and Stillwell, J. (eds.). Modelling land-use change: progress and application. GeoJournal Library, Vol. 90, Springer, Dordrecht, pp. 219-235

Loonen, W., Koomen, E. (2009) Calibration and validation of the Land Use Scanner allocation algorithms. Netherlands Environmental Assessment Agency (PBL), Bilthoven, pp. 23

MacDonald, D., Crabtree, J.R., Wiesinger, G., Dax, T., Stamou, N., Fleury, P., Gutierrez Lazpita, J., Gibon, A. (2000) Agricultural abandonment in mountain areas of Europe: Environmental consequences and policy response. Journal of Environmental Management 59: 47-69

McCann, P. (2001) Urban Regional Economics. University Press, Oxford

Meijl, H. van, Rheenen, T. van, Eickhout, B. (2006) The impact of different policy environments on agricultural land use in Europe. Agriculture, Ecosystems & Environment 114: 21-38

Overmars, K.P., de Koning, G.H.J. and Veldkamp, A. (2003) Spatial autocorrelation in multiscale land use models. Ecological Modelling 164: 257-270.

Pegels, M. (2010) Een decennium aan veranderingen in het Verenigd Koninkrijk. Onderzoek naar de ruimtegebruiksveranderingen in het Verenigd Koninkrijk tussen 1990 en 2000. Bachelor thesis. Vrije Universiteit, Amsterdam

Schneeberger, W. (2003) Instrumente und Maßnahmen zur Erhaltung des
Berggebiets in Österreich. Institut für Agrarökonomik der Universität für Bodenkultur, Vienna

Tasser, E. (2005) Kein allheilmittel; almersließungen auf dem prüfstand. EURAC, Bolzano, Italy

Verburg, P.H., Nijs, T.C.M. de, Eck, J.R. van, Visser, H., Jong, K. de (2003) A method to analyse neighbourhood characteristics of land use patterns. Computers, Environment and Urban Systems 28: 667-690

ANNEXES

Two Annexes are added to the paper. Annex 1 contains an explanation and the actual *alloc* files that were used to specify the EuClueScanner model settings. Annex 2 shows the inelastic model results that were listed in table7.

Annex 1: Alloc files

Table 12 and 13 show the *alloc1.reg* and *alloc2.reg* files that were used for the model runs. The alloc files are simple notepad files and are based on a multinomial regression. Table 9, 10 and 11 contain respectively the land use IDs, non-neighbourhood factor IDs and neighbourhood factor IDs. The following part describes the *alloc1.reg* file per line with between brackets the variable signs used in equation 3. The *alloc2.reg* file has a similar format.

1: The value in first line represents land use *ID*. (j)

2: The value on the second line is the *intercept* for this land use class. (β_y)

3: The value on the third line represents the *amount* of variables used for this land use class.

4: Then the *beta-coefficients* of the variables are shown with their *factor IDs*. (β_{yj})

n: Under these coefficients a blank line precedes the next land use class. Etc.

The Pseudo R-squares give an indication of how much of the land use can be explained with these variables.

Table 9. *Land use IDs*

j	Land use
0	Urban fabric
1	Industry
2	Arable land
3	Pastures
4	Forest
5	Semi-natural vegetation

Table 10. Non-neighbourhood *factor IDs*

y	Factor
1	accessibility to cities
3	accessibility to ports
9	Water deficit during the growing season
32	presence of an impermeable layer
39	accumulated rainfall from march to july
45	soil water available to plants
65	slope
66	south slope
67	Natura 2000 network
68	elevation

Table 11. Neighbourhood *factor IDs*

x	Factor
0	Urban fabric
1	Industry
2	Arable land
3	Pastures
4	Forest
5	Semi-natural vegetation
7	Other nature

Table 12. Alloc1.reg (*y*)

	Regression				
	Combined	*Vasco*	*Elevation*	*Natural*	*Final*
	0	0	0	0	0
	9.661	5.964	5.167	3.346	5.257
	7	8	8	6	8
	0.0005 1	-0.00043 1	-0.000333 1	-0.002 45	-0.000316667 1
	0.006 9	-0.00007 3	-0.000067 3	-0.004 68	-0.0000666667 3
	-0.003 45	-0.004 39	-0.001 45	-1.287 65	-0.002 45
	-0.282 65	0.001 45	-1.306 65	-0.425 32	-0.003 68
	0.041 32	-1.482 65	-0.002 68	-1.832 67	-1.284 65
	-0.228 67	-0.333 32	-0.535 32	0.521 66	-0.558 32
	0.532 66	-1.599 67	0.492 66		-1.559 67
		0.490 66	-1.540 67	1	0.504 66
	1			1.794	
	7.817	1	1	5	1
	6	6.303	5.270	-0.007 45	4.289
	-0.000133 1	8	9	-0.005 68	8
	-0.011 9	-0.00073 1	-0.000633 1	-1.138 65	-0.0006 1
	-0.008 45	-0.00010 3	-0.0001 3	-0.712 32	-0.0000833333 3
	-0.291 65	-0.007 39	-0.003 39	-1.957 67	-0.007 45
	-0.131 32	-0.004 45	-0.007 45		-0.004 68
	-0.328 67	-1.376 65	-1.157 65	2	-1.105 65
		-0.657 32	-0.003 68	4.155	-0.921 32
	2	-1.681 67	-0.840 32	6	-1.452 67
	1.818	-0.133 66	-0.236 66	0.007 45	-0.222 66
	6		-1.513 67	-0.005 68	
	0.008 9	2		-1.242 65	2
	-0.003 45	4.779	2	-0.498 32	3.463
	-0.399 65	8	3.482	-1.101 67	8
	-0.181 32	-0.00015 1	9	0.397 66	0.00005 1
	0.057 67	0.00002 3	0.000017 1		0.0000333333 3
	0.515 66	-0.008 39	0.00003 3	3	0.007 45
		0.012 45	0.000 39	1.526	-0.005 68
	3	-1.505 65	0.007 45	6	-1.238 65
	1.438	-0.139 32	-1.253 65	0.002 45	-0.442 32
	6	-1.124 67	-0.005 68	0.000 68	-1.131 67
	0.000167 1	0.380 66	-0.414 32	-0.711 65	0.401 66
	0.008 9		0.386 66	-0.042 32	
	-0.003 45	3	-1.095 67	-0.903 67	3
	-0.249 65	0.707		0.629 66	2.326
	-0.108 32	7	3		8
	0.487 66	-0.00012 1	0.144	5	-0.0000333333 1
		0.00003 3	9	-6.631	-0.0000333333 3
	4	0.003 45	-0.000033 1	5	0.001 45
	0	-0.863 65	0.00000 3	0.005 68	0.000 68
	1	-0.015 32	0.005 39	-0.675 65	-0.727 65
	0 1	-0.626 67	0.001 45	0.275 32	-0.066 32
		0.610 66	-0.727 65	0.622 67	-0.805 67
	5		-0.002 68	0.381 66	0.615 66
	-1.844	4	-0.196 32		
	3	0	0.598 66		4
	-0.263 65	1	-0.560 67		0
	-0.085 32	0 1			1
	0.355 66		4		0 1
		5	0		
		-7.423	1		5
		8	1 0		-8.320
		0.00012 1			8
		0.00003 3	5		-0.0000833333 1
		0.008 39	-9.206		0.0001 3
		-0.005 45	9		0.002 45
		-0.004 65	-0.00010 1		0.005 68
		-0.594 32	0.000117 3		-0.676 65
		0.581 67	0.002 39		0.228 32
		0.182 66	0.002 45		0.552 67
			-0.669 65		0.418 66
			0.005 68		
			0.212 32		
			0.399 66		
			0.528 67		
Pseudo R-Squares					
Cox and Snell	0.864[4]	Not self performed	,618	,600	,613
Nagelkerke	0.931		,666	,646	,660
McFadden	0.758		,366	,348	,361

[4] These pseudo R-squares are calculated for a combined regression with neighbourhood variables.

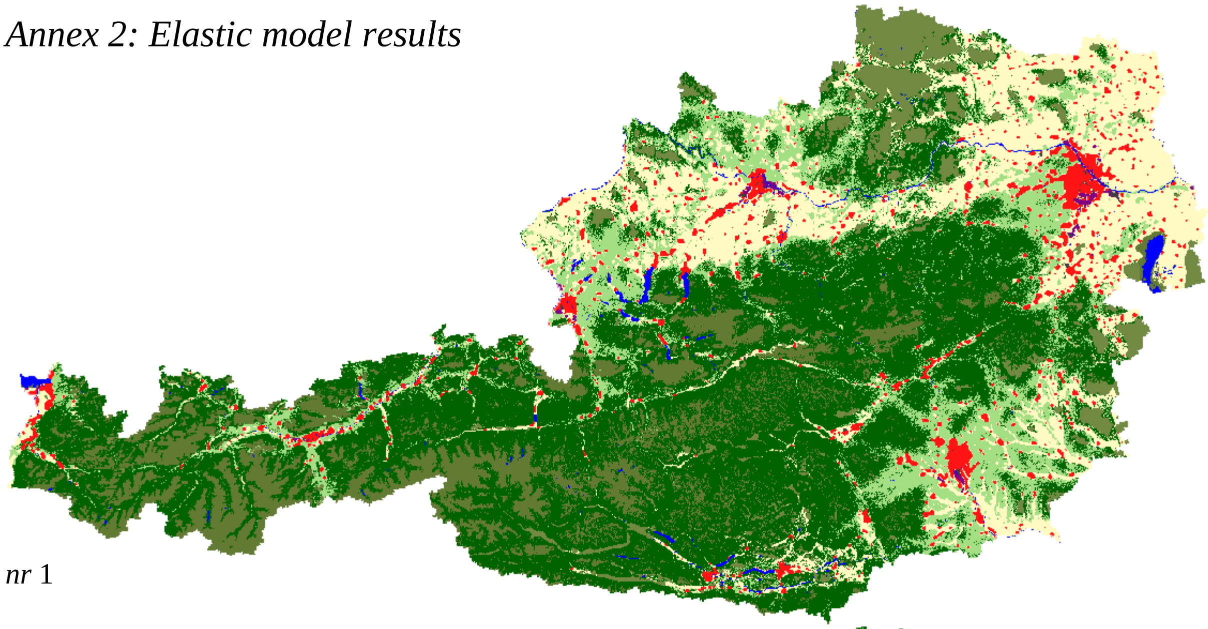
Table 13. Alloc2.reg (x)

Regression				
Combined	*Vasco*	*Elevation (Vasco)*	*Natural (Vasco)*	*Final (Vasco)*
0	0	0	0	0
9.661	-2.350	-2.350	-2.350	-2.350
5	6	6	6	6
-0.197 4	0.136 0	0.136 0	0.136 0	0.136 0
-0.154 7	-0.071 4	-0.071 4	-0.071 4	-0.071 4
-0.115 3	0.028 1	0.028 1	0.028 1	0.028 1
-0.109 2	0.011 3	0.011 3	0.011 3	0.011 3
-0.200 5	0.021 2	0.021 2	0.021 2	0.021 2
	-0.072 5	-0.072 5	-0.072 5	-0.072 5
1				
7.817	1	1	1	1
5	-3.432	-3.432	-3.432	-3.432
-0.175 4	6	6	6	6
-0.124 7	0.018 0	0.018 0	0.018 0	0.018 0
-0.144 3	-0.082 4	-0.082 4	-0.082 4	-0.082 4
-0.112 2	0.204 1	0.204 1	0.204 1	0.204 1
-0.129 5	-0.026 3	-0.026 3	-0.026 3	-0.026 3
	-0.004 2	-0.004 2	-0.004 2	-0.004 2
2	-0.049 5	-0.049 5	-0.049 5	-0.049 5
1.818				
5	2	2	2	2
-0.077 4	-2.597	-2.597	-2.597	-2.597
-0.054 7	6	6	6	6
0.010 3	0.041 0	0.041 0	0.041 0	0.041 0
0.063 2	-0.045 4	-0.045 4	-0.045 4	-0.045 4
-0.035 5	0.039 1	0.039 1	0.039 1	0.039 1
	0.045 3	0.045 3	0.045 3	0.045 3
3	0.103 2	0.103 2	0.103 2	0.103 2
1.438	-0.008 5	-0.008 5	-0.008 5	-0.008 5
5				
-0.068 4	3	3	3	3
-0.042 7	-2.235	-2.235	-2.235	-2.235
0.063 3	6	6	6	6
0.010 2	0.035 0	0.035 0	0.035 0	0.035 0
-0.021 5	-0.039 4	-0.039 4	-0.039 4	-0.039 4
	0.018 1	0.018 1	0.018 1	0.018 1
5	0.094 3	0.094 3	0.094 3	0.094 3
-1.844	0.044 2	0.044 2	0.044 2	0.044 2
5	0.006 5	0.006 5	0.006 5	0.006 5
-0.049 4				
0.037 7	4	4	4	4
0.017 3	0	0	0	0
0.023 2	1	1	1	1
0.127 5	0 1	0 1	0 1	0 1
	5	5	5	5
	0.161	0.161	0.161	0.161
	6	6	6	6
	-0.029 0	-0.029 0	-0.029 0	-0.029 0
	-0.082 4	-0.082 4	-0.082 4	-0.082 4
	-0.041 1	-0.041 1	-0.041 1	-0.041 1
	-0.013 3	-0.013 3	-0.013 3	-0.013 3
	-0.002 2	-0.002 2	-0.002 2	-0.002 2
	0.091 5	0.091 5	0.091 5	0.091 5

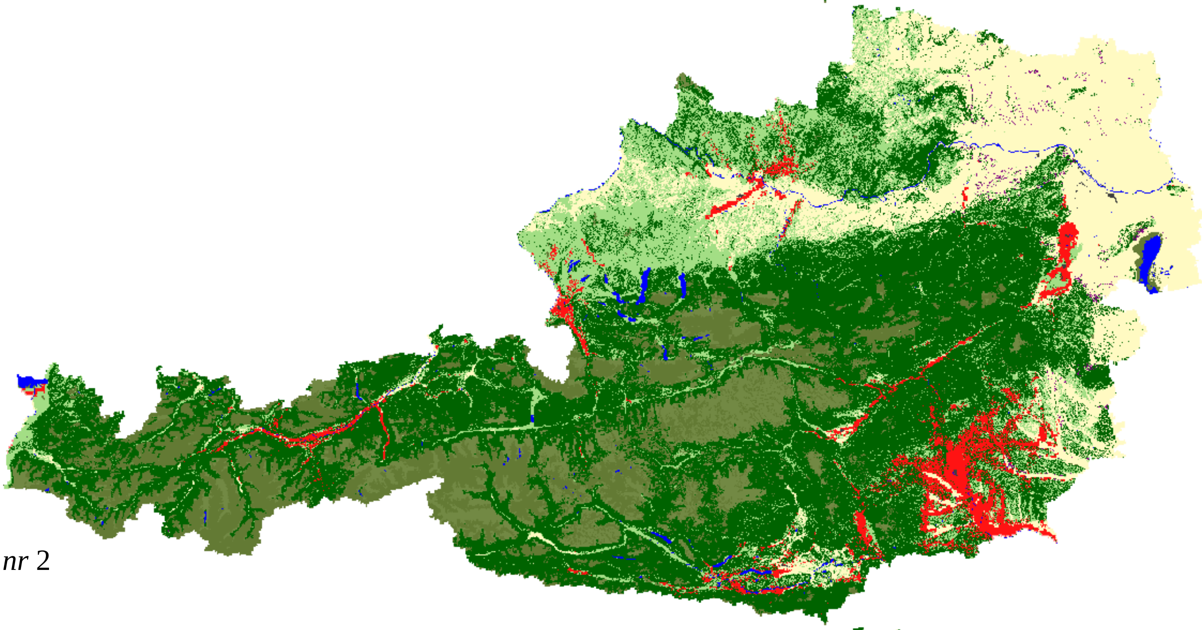
Pseudo R-Squares					
Cox and Snell	,864[5]	Not self performed	“	“	“
Nagelkerke	,931				
McFadden	,758				

[5] These pseudo R-squares are calculated for a combined regression with non-neighbourhood variables.

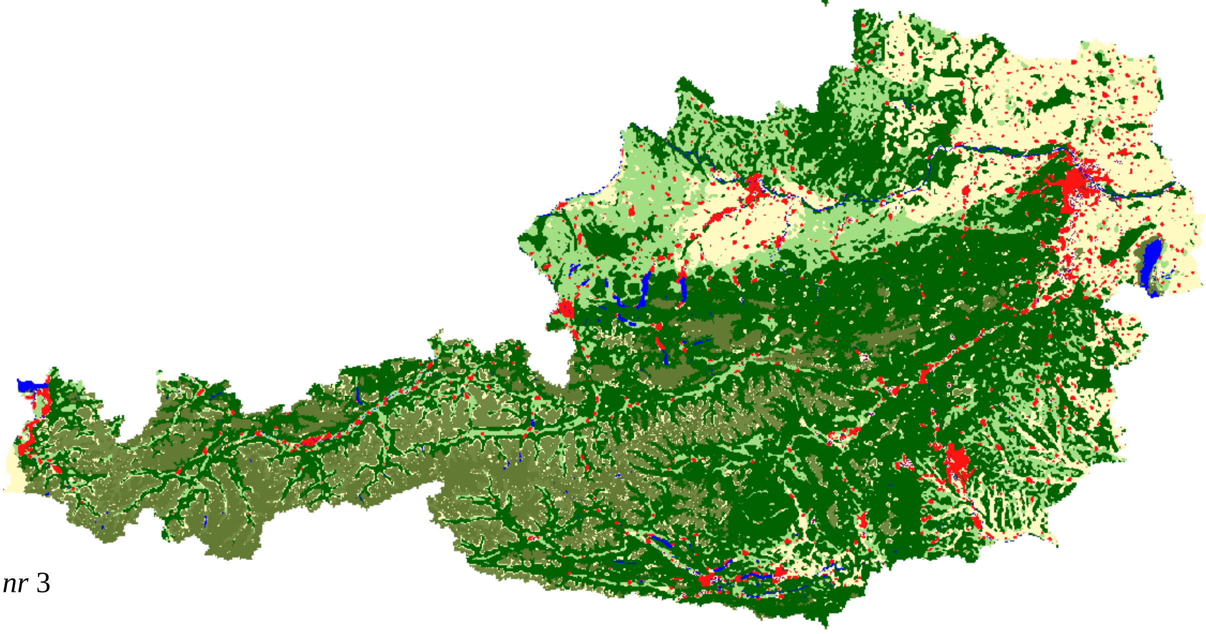
Annex 2: Elastic model results



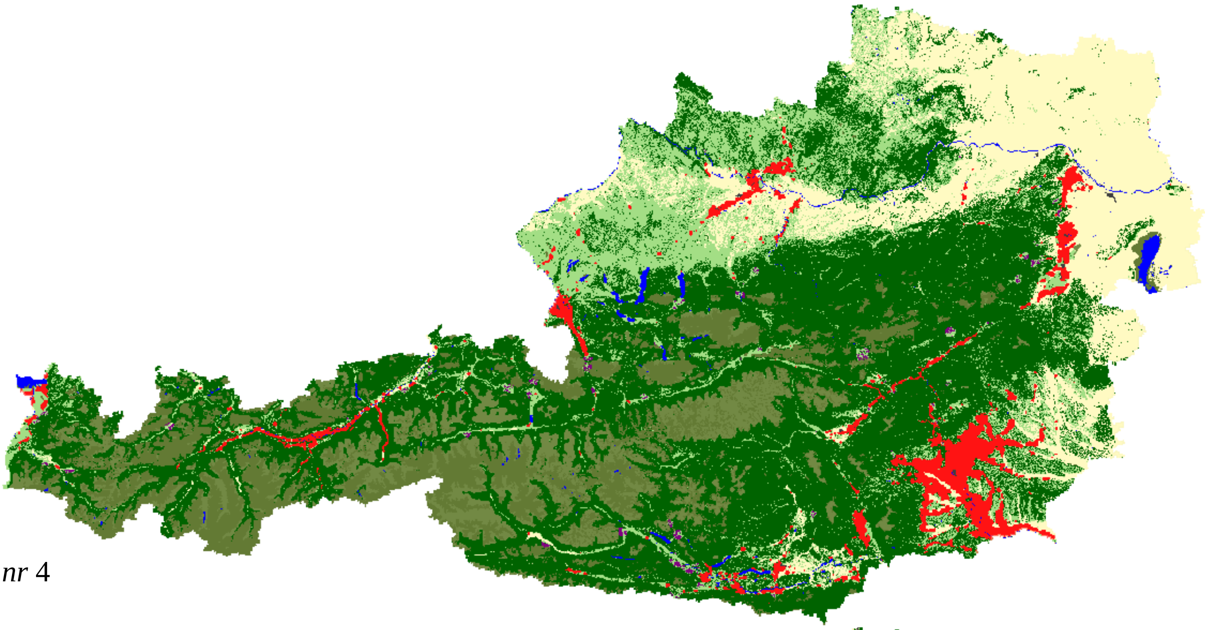
nr 1



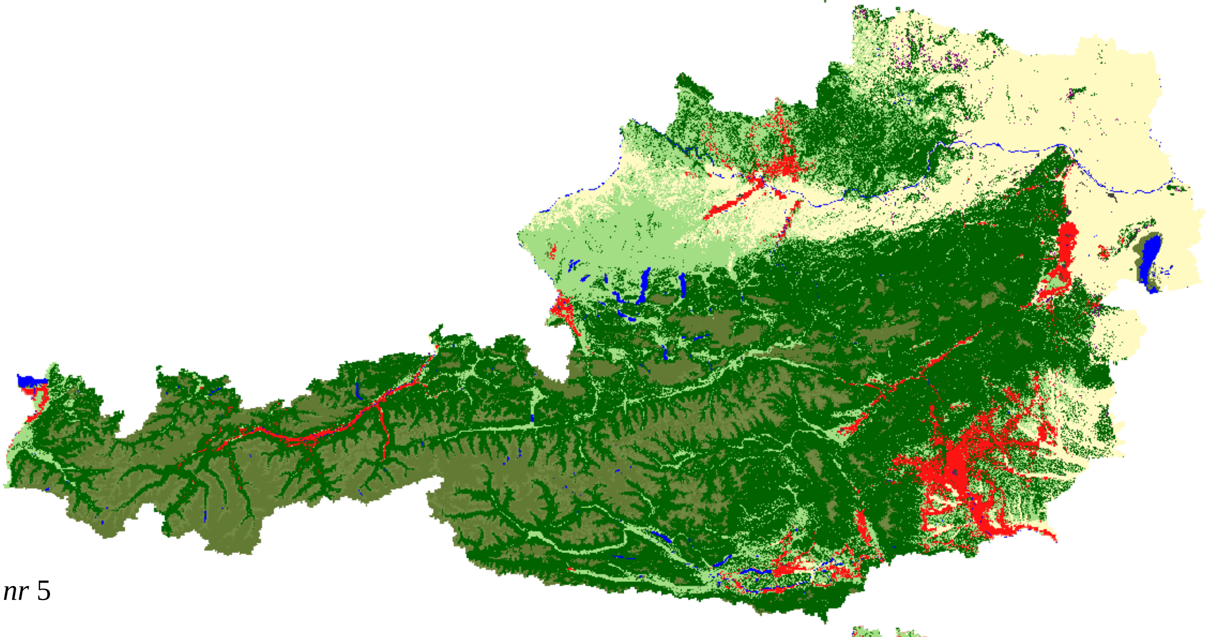
nr 2



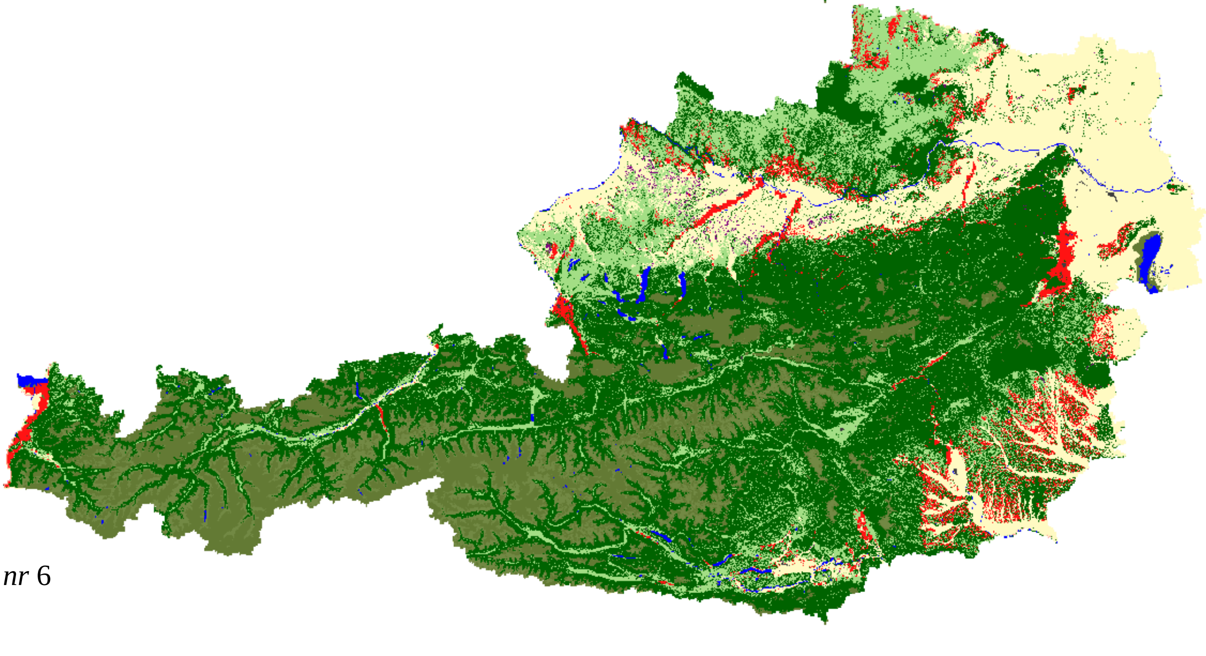
nr 3



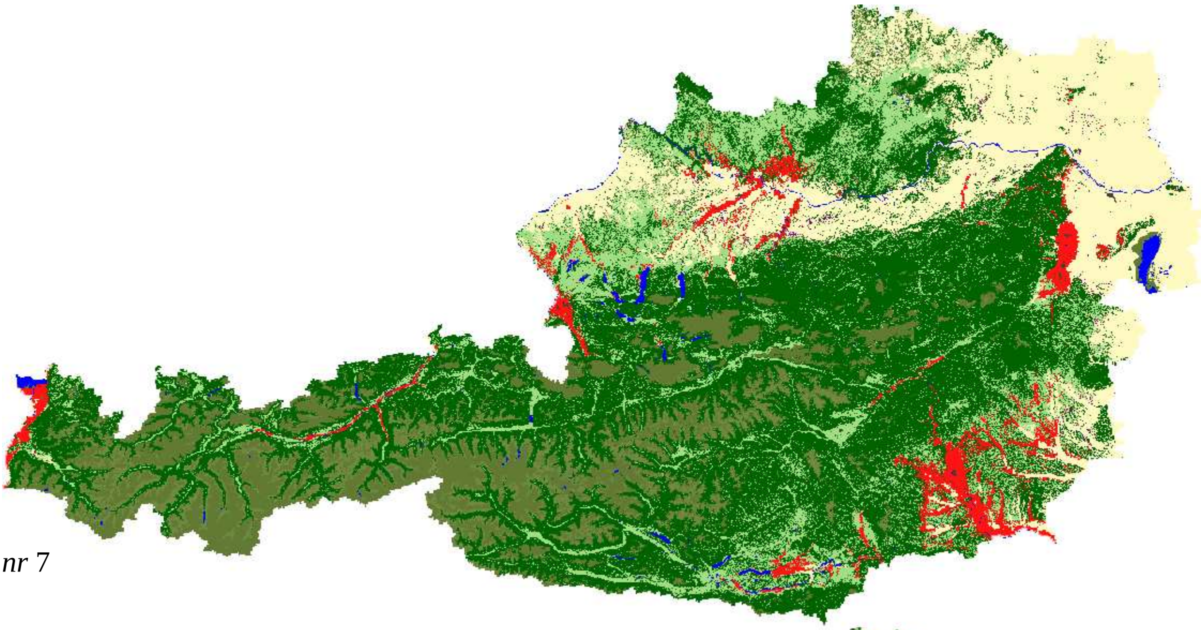
nr 4



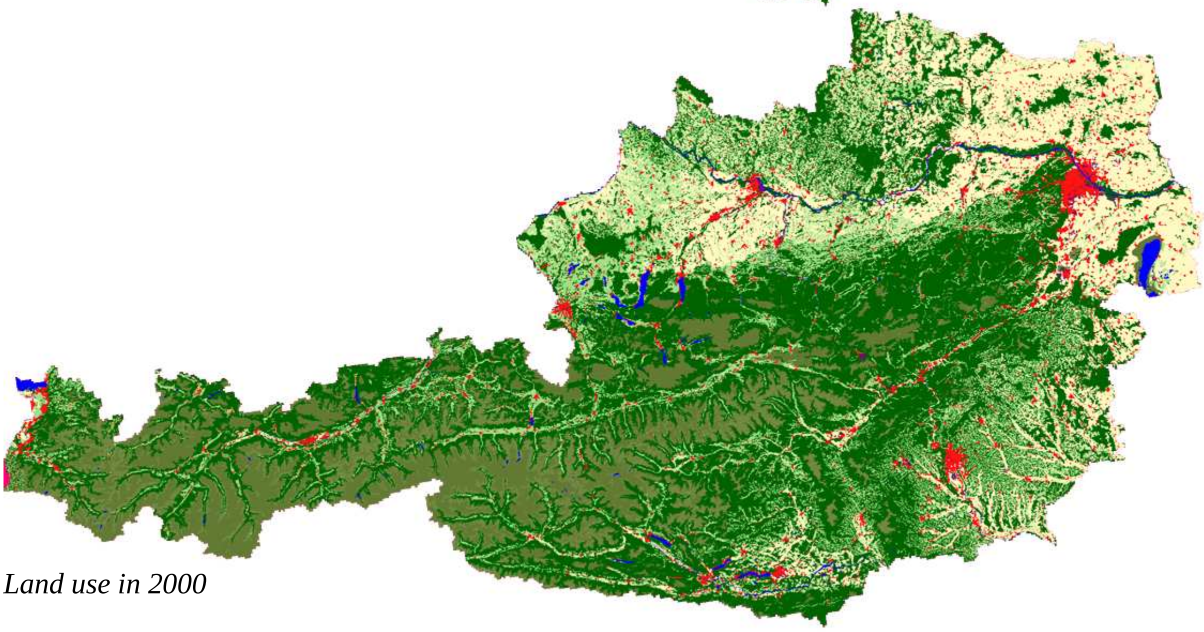
nr 5



nr 6



nr 7



Land use in 2000