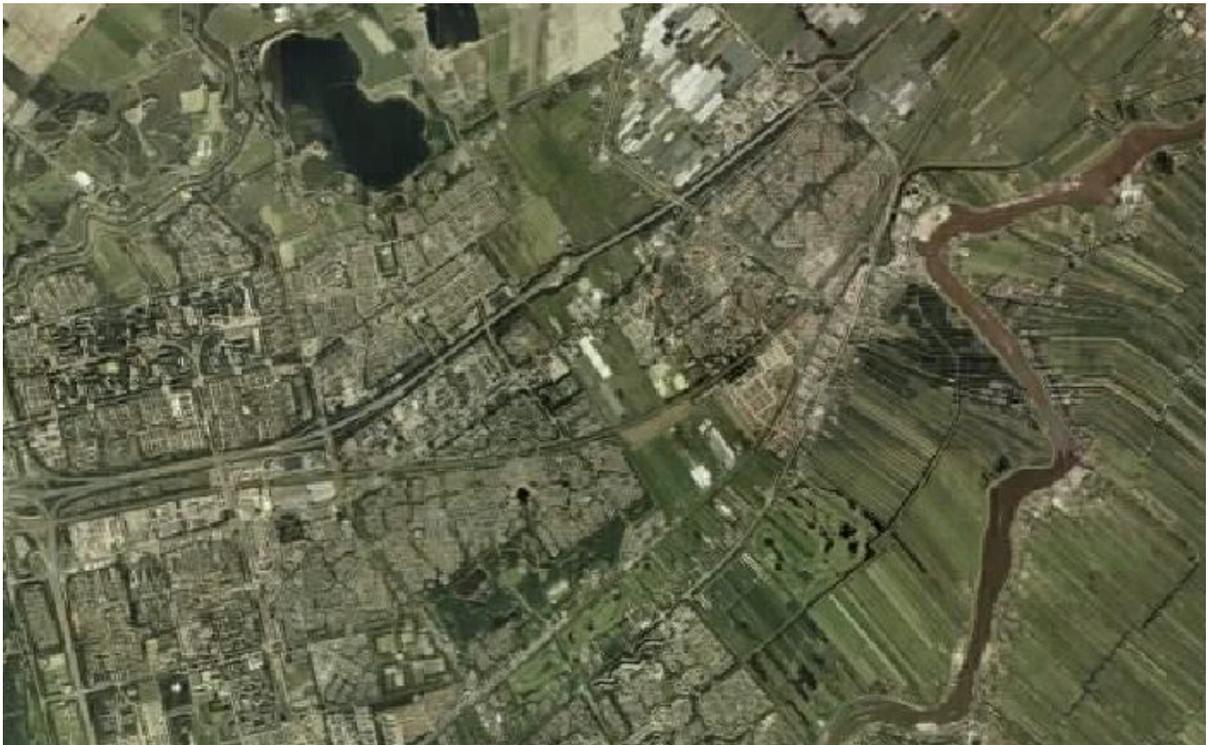


# Practical: Explorative modeling of future land use for the Randstad region of the Netherlands

*Peter H. Verburg and Jan-Peter Lesschen  
Wageningen University, the Netherlands*



*Data and CLUE software: self-extracting installation cluesmfl.exe  
Software: Arcview + Spatial Analyst Extension + SPSS*

## **Introduction: The Randstad and the Green Heart**

Randstad Holland is a name with no city. It has no official status, no mayor and no municipal council. Nor is there any organisation especially responsible for land use planning in the Randstad. No one knows exactly where the Randstad begins and ends, and there are no conclusive figures on either its surface area or population.

The term Randstad Holland was launched in the 1930s to denote a group of towns and cities located relatively close together in the west of the Netherlands. The name is thus considerably younger than the towns and cities it stands for. These are arranged roughly in the form of a horseshoe, with a diameter of some 50 to 60 kilometres. A few small towns and numerous villages are located in the open space in-between. This is largely a rural area, with space for livestock and arable farming, market gardening and leisure pursuits. And it is this that gives the Randstad a totally different character than, for example, London or Paris. The urban areas are grouped around a rural heart "the

Green Heart" of the Randstad. In the past few years, more and more studies have been using the term Randstad for more than just the ring of towns and cities in the area. Since the rural areas have increasingly been feeling the impact of urban sprawl - with more traffic on the roads, more people engaged in leisure pursuits, and encroaching suburbanisation - the trend is now to refer to the whole area - the cities and their immediate vicinity and the Green Heart - as the Randstad.

A second characteristic of the Randstad is the fact that it has more than one centre. None of the cities can claim to be the undisputed "capital". There is no hierarchy, since the four major cities - Amsterdam, Rotterdam, The Hague and Utrecht - each fulfil a separate role. Amsterdam might be the capital of the Netherlands, but it is not the seat of government. It is however the main financial and cultural centre. Rotterdam has the largest port not only in the Netherlands, but also in the world. The Hague is the seat of government, and, thanks to its central location, Utrecht is the hub of the rail system, and a conference and trade fair centre.

Most of what is said about Randstad Holland in the above text still applies. But, nearly 20 years on, the situation with regard to the Green Heart has changed somewhat. To start with, the Ministry responsible for land-use planning gave it official borders in 1990, which means that exact figures can now be provided on surface area and population of this part of the Randstad. What is more, the Green Heart Administrative Platform was set up in October 1996. Though this body does not possess local government status, it is the vehicle through which many parties can make their voices heard in both studies and consultations. And it is the future of the Green Heart that is now under discussion.

#### A. Randstad planning

Dutch planners have often been admired for developing a conurbation in the form of a number of separate cities grouped around a more or less open central area. However, the praise is not entirely deserved, since the structure of the Randstad was largely determined by geographical characteristics and historical accident. At most, the planners can be congratulated for doing their best to keep the structure more or less intact. The emergence of the Randstad in its current form proceeded roughly as follows.

1. In the Middle Ages, the central area was largely an inaccessible marshland. Settlements were built on higher ground along the dunes in the west (Haarlem, The Hague), on the sandy ground in the east (Hilversum, Amersfoort), along river banks (Utrecht and Leiden on the Old Rhine, Amsterdam on the Amstel, Rotterdam on the Rotte). Thanks to the fishing industry and shipping, the settlements on the coast and the rivers developed into centres of trade and industry. By the seventeenth century, the west of the Netherlands - Holland - had reached such pre-eminence that it also became the cultural, scientific and administrative centre.
2. The first attempts at town and country planning were made some fifty years ago. In 1958, the government published a report entitled "The West of the Country" in which, for the first time, concern was expressed at the problem of congestion in Randstad Holland. The rapid growth of the economy and rising population figures led to growing fears of unbridled urban spread in the Randstad. In 1960, the First Policy Document on Town and Country Planning in the Netherlands recommended reducing pressure on the Randstad by encouraging the development of peripheral areas in the north, south and eastern Netherlands. The Green Heart was to remain green, due to its importance as an agricultural area.
3. In the 1960s, it became evident that this policy was meeting with very limited success. The population of the Randstad continued to grow, especially in the Green Heart. Thanks to increasing car ownership, people were becoming more and more mobile and were moving in increasing numbers to towns and villages in the "green" central area. As employment was still only to be found in or near the cities, commuter traffic grew rapidly. The Second Policy

Document on Town and Country Planning in the Netherlands appeared in 1966. The aim was still to deflect the growth of the cities in the Randstad to places elsewhere, but now efforts centred on bordering regions, such as the north of North Holland, Flevoland, and the Delta area. Urban sprawl in the Green Heart was to be countered by reserving a limited number of locations for population overspill. At the basis of all these plans was the alarming prognosis that the Netherlands would have a population of 20 million by 2000. (The current prognosis is a peak of over 17 million in 2030, followed by a gradual drop).

4. In the 1970s, it became evident that these plans had not been very successful either. Local government authorities took little notice of them, while small towns and villages were still making every effort to attract both businesses and residents. As a result, the Green Heart was becoming less and less green. Migration from the cities to the smaller municipalities had taken on such proportions that the populations of the three major cities dropped by an average of more than 18% between 1970 and 1985.
5. There was clearly a need for different town and country planning policies, with central government not only setting out the main outlines, but also exerting a direct influence on planning. This was the principle underpinning the Third Policy Document on Town and Country Planning, the first instalment of which was published in 1973. It was the second instalment - the Policy Document on Urbanisation, published in 1976 - that formed the main breakthrough. Efforts to develop locations in other parts of the country were replaced by a planning policy on the Randstad, central to which was the designation of 14 towns as growth centres for new housing. With three exceptions, these growth centres were located outside the Randstad. The growth of other smaller towns and villages was to be curbed by restricting the number of new homes for which planning permission would be given. At the same time, urban regeneration projects would be launched to bring migration from the cities to a halt. In the third instalment, the Policy Document on Rural Areas, which was published in 1977, attention focused for the first time on the Green Heart's function not only as an agricultural centre, but also as a location for leisure pursuits and nature conservation.
6. In the 1980s it became evident that these policies were yielding some satisfactory results. The populations of the major cities had stabilised, and large numbers of new homes had been built in the growth centres. Without these growth centres, there would have been more urban sprawl and an even greater increase in road traffic. But the policymakers had overestimated the extent to which they could influence companies in their choice of location. Most companies were inclined to locate their operations within the Randstad - near Schiphol Airport, for instance - and this led to very dense commuter traffic to and from the growth centres. This has led to problems in and around Amsterdam in particular, where every morning and evening motorists queue to pass through the tunnels under the IJ and the North Sea Canal on their way to and from the residential areas to the north of the city. Though policy on growth centres was by no means regarded as unsuccessful, it was terminated in the course of the 1980s.

More and more credence was given to the idea that business locations and residential areas should be planned as close as possible to existing cities. This would lead to more people travelling by public transport rather than by car, and provide more opportunities to keep the Green Heart and other open areas green. These ideas formed the basis of the Fourth Policy Document on Town and Country Planning in the Netherlands, which was published in 1988, and revised in 1990 after a change of government. The VINEX, as the revised document is known, designated some twenty locations in which most of the Randstad's new housing will be built in the period up to 2005 (the VINEX locations). All of these locations are situated just outside the official boundaries of the Green Heart.

## B. The role of the Green Heart in current planning

The Green Heart was originally a somewhat vague planning concept. It was an area with no clear boundaries, in which a number of things were supposed not to happen. There were to be no new houses, greenhouses or roads. But these were exactly what were built there. The Green Heart not only became less and less green, it also shrank because certain regions were no longer regarded as being part of it. These included the Little Ring between The Hague, Delft, Rotterdam and Zoetermeer, a major part of the Haarlemmermeerpolder (where Schiphol Airport is located) and the area around the villages of Vleuten and De Meern to the west of Utrecht.

From approximately 1990, the government, working on the basis of the Fourth Policy Document (VINEX) has attempted to curb these developments. The Green Heart has been given official boundaries and it is now regarded as part of Randstad Holland. Randstad Holland thus consists of a ring of cities with a population of some 6 million, and a central area with some 670,000 inhabitants. Population density in the ring is approximately 1,680 per km<sup>2</sup>, and in the Green Heart 470 per km<sup>2</sup> (i.e. slightly higher than in the Netherlands as a whole). The largest part of the Green Heart is located in the province of South Holland, with smaller segments in North Holland and Utrecht. The Green Heart contains 43 entire municipalities and parts of 27 more.

Six Green Heart municipalities may be regarded as largely urban, i.e. Gouda, Alphen aan de Rijn (each of which have a 70,000 inhabitants), Woerden, Waddinxveen, Boskoop and De Ronde Venen (which includes Mijdrecht). The Green Heart is thus not entirely rural, nor is the ring of cities entirely urban, with more or less green belts between these cities. The population is still growing more rapidly in the Green Heart than in the Netherlands as a whole, though the difference is by no means as great as in the 1970s. A graph shows that while population growth in the Green Heart's rural municipalities has been slightly lower than the national average since the late 1980s, the populations of the urban municipalities have been growing twice as rapidly as those in the rest of the country.

Protection is no longer the sole objective of land-use planning policy for the Green Heart, which is now drawn up in close consultation with local authorities and other organisations. Apart from restrictive measures - in relation, for example, to businesses and new housing - policy largely focuses on developing the Green Heart's potential. Importantly, the area is not regarded as a single entity. Traditionally, settlements were built alongside rivers such as the Oude Rijn, the Hollandse IJssel and the Gouwe, where the land was higher and drier than the surrounding marshes. These settlements formed a basis from which marsh could be reclaimed. Nowadays, the Green Heart can be divided into a number of "segments" on the basis of these more densely populated zones. There are roughly three main areas: fens and grassland to the north, interspersed with lakes and ponds, very open fens and grassland on either side of the Lek river to the south, and a mixed area (grassland, arable farms and market gardens) to the west.

Policy needs to be adapted to these regional differences. In the residential areas, the main objective is to curb urban sprawl and the further development of infrastructure; the keynotes in the lakeland areas are nature conservation and expanding recreational facilities; in the fens and grasslands, ways need to be sought to combine livestock farming with nature and landscape conservation; and in the arable farming and market gardening areas near, for example, Zoetermeer and Aalsmeer, the emphasis is on agricultural development. New housing is not wholly excluded from the picture, but it may only be planned in specific locations within each municipality. Up to 2005, 17,000 new homes may be built in the Green Heart (as opposed to 83,000 in the VINEX locations on its periphery).

In their policy document on land-use planning and the environment, the relevant provincial authorities, the Ministry of Housing, Spatial Planning and the Environment, the Ministry of Agriculture, Nature Management and Fisheries and the Ministry of Transport, Public Works and

Water Management have undertaken to enhance the "green" quality of the area. To this end, eleven projects, involving an investment of 25 billion guilders over the next 25 years will be launched in an effort to create and link nature conservation areas, develop both a waterway network for pleasure boats and a signposted cycle path network, plan the transitional areas between the cities and the Green Heart, and create new woodlands and recreational areas. Where, in the past, plans and measures tended to be introduced at random, an integrated approach will be adopted to these new projects.

### C. Further discussion

It will be obvious from the above overview that a shift has occurred from a largely defensive approach to policies in which incentives play a key role. Yet the Green Heart still gives rise to controversy. According to some critics of policy, the whole concept of the Green Heart is nothing but a myth, kept alive these forty years, which should be abandoned if Randstad Holland is to maintain its position as an international metropolis. In their view, the agriculture sector in Europe is too weak to act - on a large scale at least - as the support base of a "green revolution". Instead, nature conservation and recreation areas should be planned only in those locations which are best suited to this purpose, while the rest of the Green Heart should support the urban functions of the Randstad.

To hear the various views on the subject, the Ministry of Housing, Spatial Planning and the Environment held a round of talks with all the parties involved in late 1995. These included local authorities, water boards, representatives of the business community, and organisations involved in, for example, leisure activities and nature conservation. From these talks, it became obvious that there is still a broad support base for land-use policies that enhance the "green" quality of the Green Heart. This will therefore be the objective of land-use planning into the next century.

## **Introduction to the CLUE model**

The Conversion of Land Use and its Effects modelling framework (CLUE) (Veldkamp and Fresco, 1996; Verburg et al., 1999) was developed to simulate land use change using empirically quantified relations between land use and its driving factors in combination with dynamic modelling of competition between land use types. The model was developed for the national and continental level and applications for Central America, Ecuador, China and Java, Indonesia are available. For study areas with such a large extent the spatial resolution for analysis was coarse and, as a result, each land use is represented by assigning the relative cover of each land use type to the pixels.

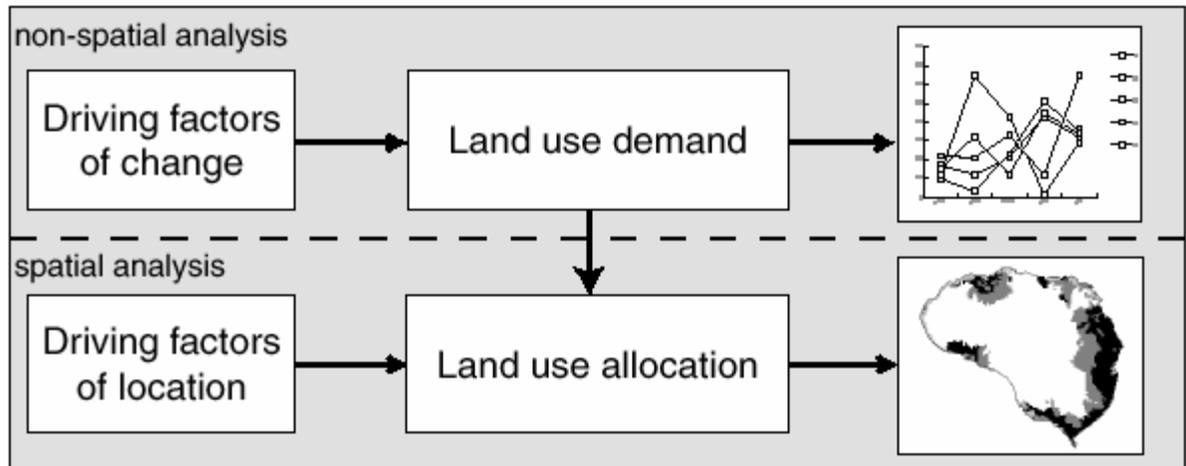
Land use data for study areas with a relatively small spatial extent is often based on land use maps or remote sensing images that denote land use types respectively by homogeneous polygons or classified pixels. This results in only one dominant land use type occupying one unit of analysis. Because of the differences in data representation and other features that are typical for regional applications, the CLUE model cannot directly be applied at the regional scale. Therefore the modelling approach has been modified and is now called CLUE-S (the Conversion of Land Use and its Effects at Small regional extent). CLUE-S is specifically developed for the spatially explicit simulation of land use change based on an empirical analysis of location suitability combined with the dynamic simulation of competition and interactions between the spatial and temporal dynamics of land use systems. More information on the development of the CLUE-S model can be found in Verburg et al. (2002) and Verburg and Veldkamp (2003).

### ***Model structure***

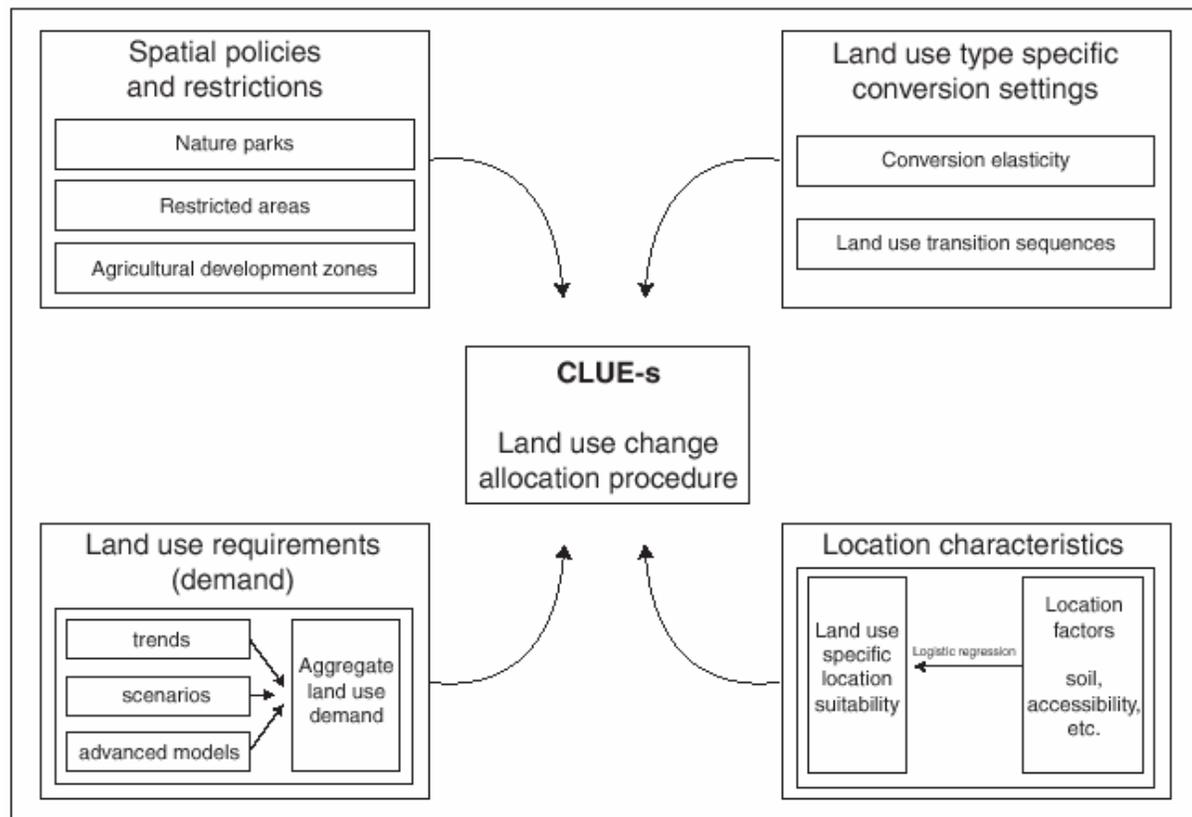
The model is sub-divided into two distinct modules, namely a non-spatial demand module and a spatially explicit allocation procedure (Figure 1). The non-spatial module calculates the area change for all land use types at the aggregate level. Within the second part of the model these demands are translated into land use changes at different locations within the study region using a raster-based system. The user-interface of the CLUE-S model only supports the spatial allocation of land use change. For the land use demand module different model specifications are possible ranging from simple trend extrapolations to complex economic models. The choice for a specific model is very much dependent on the nature of the most important land use conversions taking place within the study area and the scenarios that need to be considered. The results from the demand module need to specify, on a yearly basis, the area covered by the different land use types, which is a direct input for the allocation module.

The allocation is based upon a combination of empirical, spatial analysis and dynamic modelling. Figure 2 gives an overview of the information needed to run the CLUE-S model. This information is subdivided into four categories that together create a set of conditions and possibilities for which the model calculates the best solution in an iterative procedure. The next sections discuss each of the boxes: spatial policies and restrictions, land use type specific conversion settings, land use requirements (demand) and location characteristics.

For recent developments on the CLUE-s model and to download the latest version we refer to: <http://www.cluemodel.nl>



**Figure 1.** Overview of the modelling procedure



**Figure 2.** Overview of the information flow in the CLUE-S model

### ***Spatial policies and restrictions***

Spatial policies and land tenure can influence the pattern of land use change. Spatial policies and restrictions mostly indicate areas where land use changes are restricted through policies or tenure status. For the simulation maps that indicate the areas for which the spatial policy is implemented must be supplied. Some spatial policies restrict all land use change in a certain area, e.g., a log-ban within a forest reserve. Other land use policies restrict a set of specific land use conversions, e.g., residential construction in designated agricultural areas or permanent agriculture in the buffer zone of

a nature reserve. The conversions that are restricted by a certain spatial policy can be indicated in a land use conversion matrix: for all possible land use conversions it is indicated if the spatial policy applies.

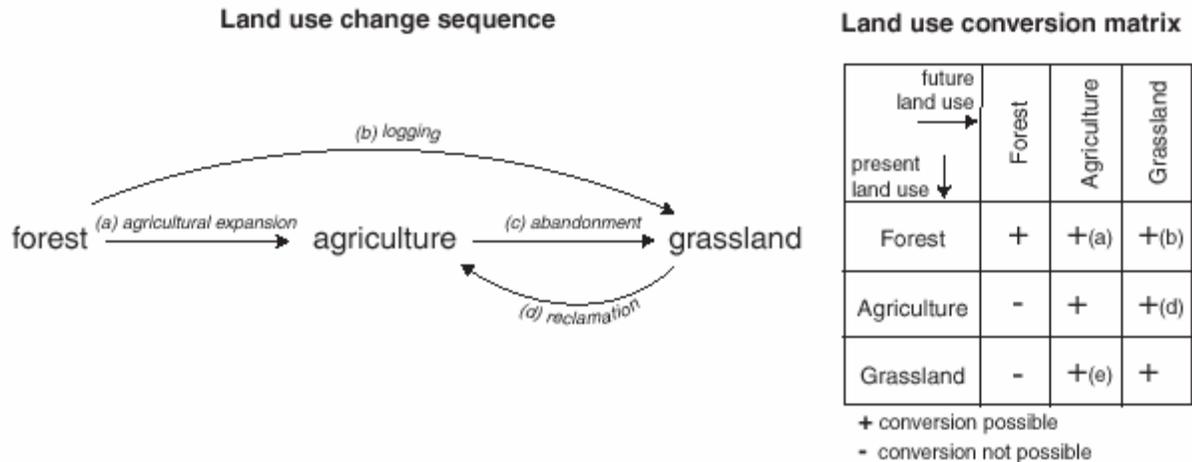
### ***Land use type specific conversion settings***

Land use type specific conversion settings determine the temporal dynamics of the simulations. Two sets of parameters are needed to characterize the individual land use types: conversion elasticities and land use transition sequences. The first parameter set, the conversion elasticities, is related to the reversibility of land use change. Land use types with high capital investment will not easily be converted in other uses as long as there is sufficient demand. Examples are residential locations but also plantations with permanent crops (e.g., fruit trees). Other land use types easily shift location when the location becomes more suitable for other land use types. Arable land often makes place for urban development while expansion of agricultural land occurs at the forest frontier. An extreme example is shifting cultivation: for this land use system the same location is mostly not used for periods exceeding two seasons as a consequence of nutrient depletion of the soil. These differences in behaviour towards conversion can be approximated by conversion costs. However, costs cannot represent all factors that influence the decisions towards conversion such as nutrient depletion, esthetical values etc. Therefore, for each land use type a value needs to be specified that represents the relative elasticity to change, ranging from 0 (easy conversion) to 1 (irreversible change). The user should decide on this factor based on expert knowledge or observed behaviour in the recent past.

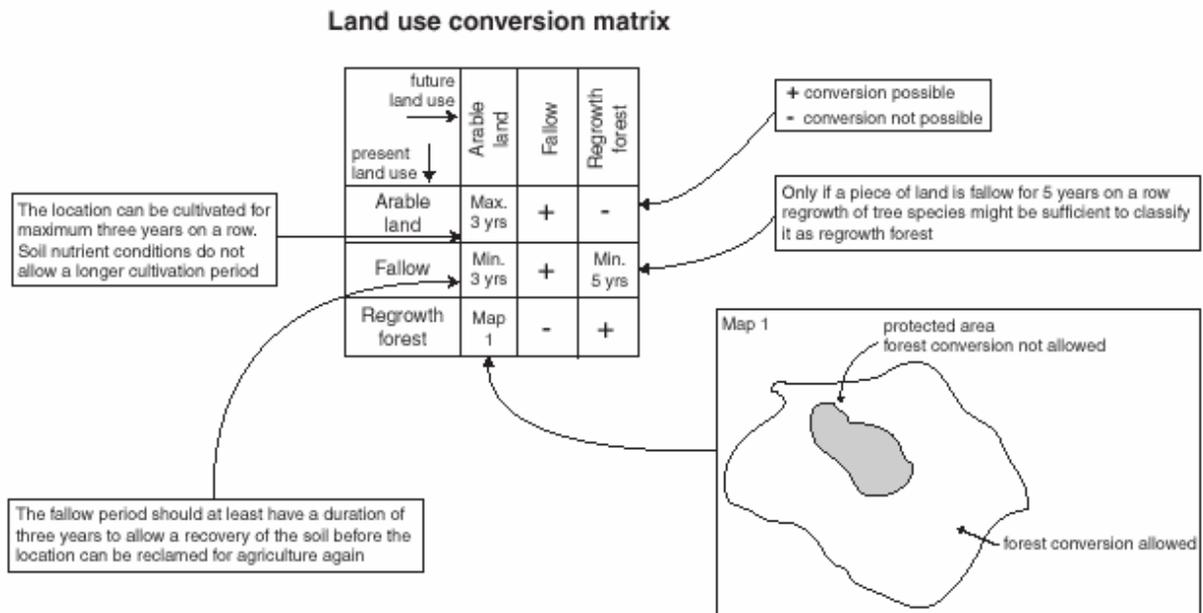
The second set of land use type characteristics that needs to be specified are the land use type specific conversion settings and their temporal characteristics. These settings are specified in a conversion matrix. This matrix defines:

- To what other land use types the present land use type can be converted or not (Figure 3).
- In which regions a specific conversion is allowed to occur and in which regions it is not allowed.
- How many years (or time steps) the land use type at a location should remain the same before it can change into another land use type. This can be relevant in case of the re-growth of forest. Open forest cannot change directly into closed forest. However, after a number of years it is possible that an undisturbed open forest will change into closed forest because of re-growth.
- The maximum number of years that a land use type can remain the same. This setting is particularly suitable for arable cropping within a shifting cultivation system. In these systems the number of years a piece of land can be used is commonly limited due to soil nutrient depletion and weed infestation.

It is important to note that only the minimum and maximum number of years before a conversion can or should happen is indicated in the conversion table. The exact number of years depends on the land use pressure and location specific conditions. The simulation of these interactions combined with the constraints set in the conversion matrix will determine the length of the period before a conversion occurs. Figure 4 provides an example of the use of a conversion matrix for a simplified situation with only three land use types.



**Figure 3.** Illustration of the translation of a hypothetical land use change sequence into a land use conversion matrix



**Figure 4.** Example of a land use conversion matrix with the different options implemented in the model

***Land use requirements (demand)***

Land use requirements (demand) are calculated at the aggregate level (the level of the case-study as a whole) as part of a specific scenario. The land use requirements constrain the simulation by defining the totally required change in land use. All changes in individual pixels should add up to these requirements. In the approach, land use requirements are calculated independently from the CLUE-S model itself. The calculation of these land use requirements is based on a range of methods, depending on the case study and the scenario. The extrapolation of trends in land use change of the recent past into the near future is a common technique to calculate land use requirements. When necessary, these trends can be corrected for changes in population growth and/or diminishing land resources. For policy analysis it is also possible to base land use requirements on advanced models of

macro-economic changes, which can serve to provide scenario conditions that relate policy targets to land use change requirements.

### ***Location characteristics***

Land use conversions are expected to take place at locations with the highest 'preference' for the specific type of land use at that moment in time. Preference represents the outcome of the interaction between the different actors and decision making processes that have resulted in a spatial land use configuration. The preference of a location is empirically estimated from a set of factors that are based on the different, disciplinary, understandings of the determinants of land use change. The preference is calculated following:

$$R_{ki} = a_k X_{1i} + b_k X_{2i} + \dots$$

where  $R$  is the preference to devote location  $i$  to land use type  $k$ ,  $X_{1,2,\dots}$  are biophysical or socio-economical characteristics of location  $i$  and  $a_k$  and  $b_k$  the relative impact of these characteristics on the preference for land use type  $k$ . The exact specification of the model should be based on a thorough review of the processes important to the spatial allocation of land use in the studied region.

A statistical model can be developed as a binomial logit model of two choices: convert location  $i$  into land use type  $k$  or not. The preference  $R_{ki}$  is assumed to be the underlying response of this choice. However, the preference  $R_{ki}$  cannot be observed or measured directly and has therefore to be calculated has a probability. The function that relates these probabilities with the biophysical and socio-economic location characteristics is defined in a logit model following:

$$\text{Log} \left( \frac{P_i}{1 - P_i} \right) = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} \dots + \beta_n X_{n,i}$$

where  $P_i$  is the probability of a grid cell for the occurrence of the considered land use type on location  $i$  and the  $X$ 's are the location factors. The coefficients ( $\beta$ ) are estimated through logistic regression using the actual land use pattern as dependent variable. This method is similar to econometric analysis of land use change, which is very common in deforestation studies. In econometric studies the assumed behaviour is profit maximization, which limits the location characteristics to (agricultural) economic factors. In the study areas is assumed that locations are devoted to the land use type with the highest 'suitability'. 'Suitability' includes the monetary profit, but can also include cultural and other factors that lead to deviations from (economic) rational behaviour in land allocation. This assumption makes it possible to include a wide variety of location characteristics or their proxies to estimate the logit function that defines the relative probabilities for the different land use types.

Most of these location characteristics relate to the location directly, such as soil characteristics and altitude. However, land management decisions for a certain location are not always based on location specific characteristics alone. Conditions at other levels, e.g., the household, community or administrative level can influence the decisions as well. These factors are represented by accessibility measures, indicating the position of the location relative to important regional facilities, such as the market and by the use of spatially lagged variables. A spatially lagged measure of the population density approximates the regionally population pressure for the location instead of only representing the population living at the location itself.

### *Allocation procedure*

When all input is provided the CLUE-S model calculates, with discrete time steps, the most likely changes in land use given the before described restrictions and suitabilities. The allocation procedure is summarized in Figure 5. The following steps are taken to allocate the changes in land use:

1. The first step includes the determination of all grid cells that are allowed to change. Grid cells that are either part of a protected area or presently under a land use type that is not allowed to change are excluded from further calculation. Also the locations where certain conversions are not allowed due to the specification of the conversion matrix are identified.
2. For each grid cell  $i$  the total probability ( $TPROP_{i,u}$ ) is calculated for each of the land use types  $u$  according to:

$$TPROP_{i,u} = P_{i,u} + ELAS_u + ITER_u$$

where  $P_{i,u}$  is the suitability of location  $i$  for land use type  $u$  (based on the logit model),  $ELAS_u$  is the conversion elasticity for land use  $u$  and  $ITER_u$  is an iteration variable that is specific to the land use type and indicative for the relative competitive strength of the land use type.  $ELAS_u$ , the land use type specific elasticity to change value, is only added if grid-cell  $i$  is already under land use type  $u$  in the year considered.

3. A preliminary allocation is made with an equal value of the iteration variable ( $ITER_u$ ) for all land use types by allocating the land use type with the highest total probability for the considered grid cell. Conversions that are not allowed according to the conversion matrix are not allocated. This allocation process will cause a certain number of grid cells to change land use.
4. The total allocated area of each land use is now compared to the land use requirements (demand). For land use types where the allocated area is smaller than the demanded area the value of the iteration variable is increased. For land use types for which too much is allocated the value is decreased. Through this procedure it is possible that the local suitability based on the location factors is overruled by the iteration variable due to the differences in regional demand. The procedure followed balances the bottom-up allocation based on location suitability and the top-down allocation based on regional demand.

Steps 2 to 4 are repeated as long as the demands are not correctly allocated. When allocation equals demand the final map is saved and the calculations can continue for the next time step. Some of the allocated changes are irreversible while others are dependent on the changes in earlier time steps. Therefore, the simulations tend to result in complex, non-linear changes in land use pattern, characteristic for complex systems.

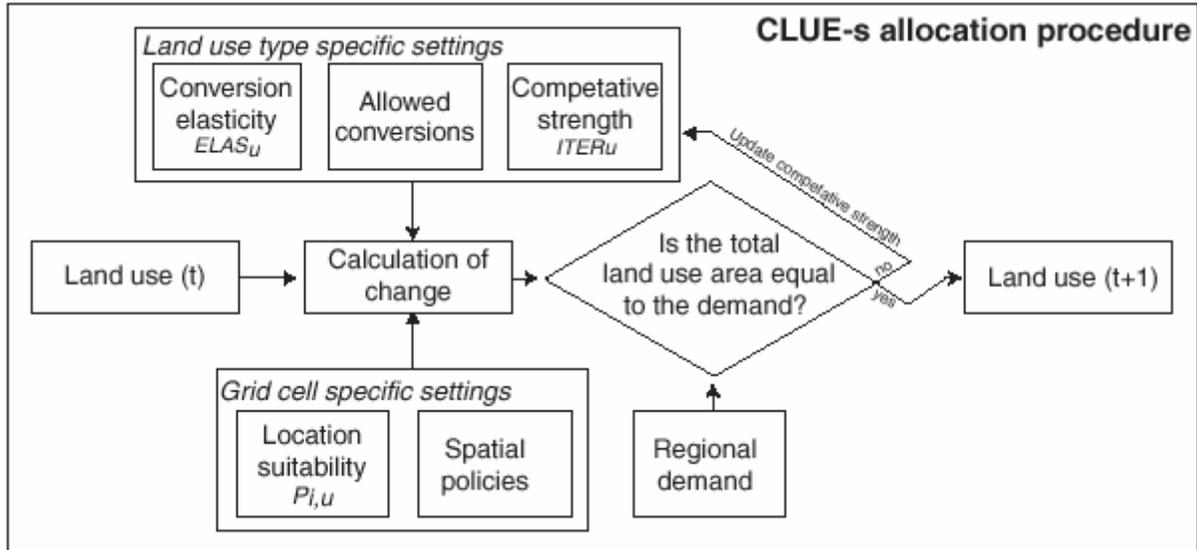


Figure 5. Flow chart of the allocation module of the CLUE-S model

### Data description of case study

In these exercises the CLUE-S model will be used for the simulation of land use change for the Randstad in The Netherlands. The selected region comprises the area between the four main cities, Amsterdam, Den Haag, Rotterdam and Utrecht (Figure 6). Nine different land uses have been distinguished in the area (Table 1). All No Data area is water, but is not taken into account. Table 2 lists the available driving factors for land use change.

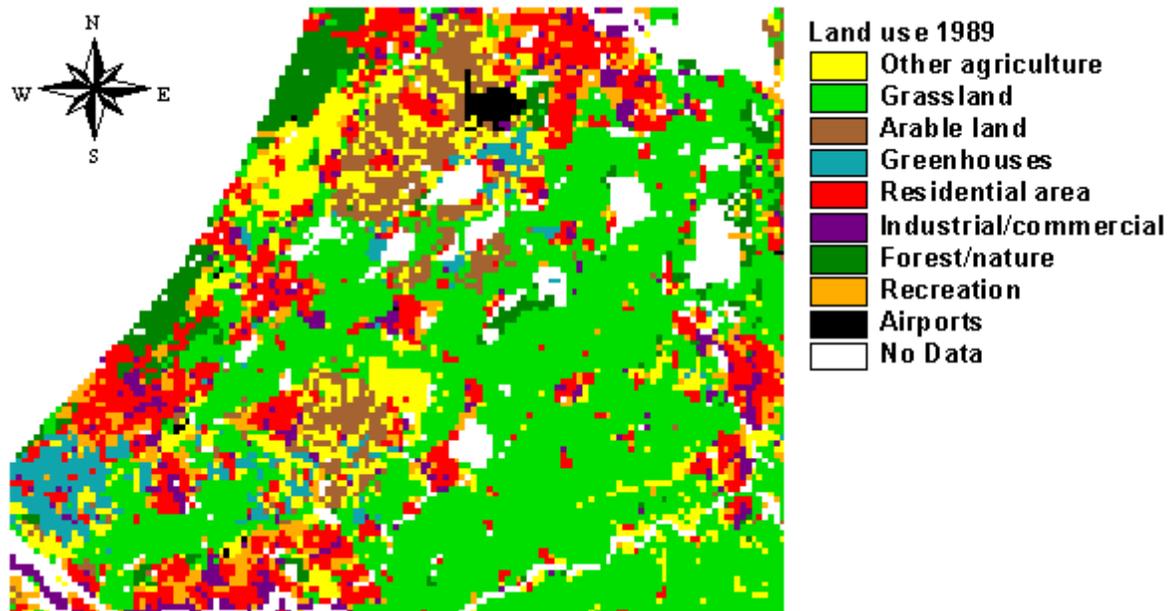


Figure 6. Initial land use map (1989)

**Table 1.** Land use types (file: cov\_all.0)

Code	Name	Description
0	Other agriculture	Agricultural land not belonging to grassland, arable land or greenhouses. This class includes horticulture, orchards, bulbs, etc.
1	Grassland	Grasslands, incl. semi-natural grasslands
2	Arable land	All arable lands
3	Greenhouses	Greenhouses
4	Residential area	Residential area and social-cultural facilities, incl. houses, roads within residential areas, schools, hospitals, churches, etc.
5	Industrial/commercial	Mining areas, industries, harbours, shopping malls, prisons, and all service industries
6	Forest/nature	Forests and natural areas, incl. peat areas, swamps, heather, etc.
7	Recreation	Parks, sport fields, kitchen garden complexes, camp sites, etc.
8	Airports	Airports

**Table 2.** Driving factors (file: sc1gr\*.fil where \* is the code of the variable)

Code	Variable	Description
0	Population density <i>Biophysical variables</i>	Population density (person/km <sup>2</sup> ) in 1989
1	Elevation	Mean elevation calculated from a high-resolution Digital Elevation Model
2	Groundwater level	<i>Groundwater dynamics characterized by depth and fluctuation through the year; higher categories are better drained soils</i>
3	Distance to water <i>Soil Characteristics</i>	Distance to the coast or main river (no network) <i>Properties of soils based on the 1:50,000 soil map</i>
4	Organic matter content	Organic matter content of the subsoil (35-125 cm)
5	Calcium content	Calcium content of the subsoil (35-125 cm)
6	Clay content <i>Socio-economic variables</i>	Clay content of the subsoil (35-125 cm)
7	Travel time to station	Travel time to nearest railway station
8	Travel time to airport	Travel time to Schiphol airport
9	Travel time to harbour	Travel time to Rotterdam harbour
10	Travel time to work	Average travel time to nearest 100,000 jobs
11	Travel time to highway	Travel time to nearest highway entrance
12	Travel time to city	Average travel time to nearest 500,000 inhabitants
13	Travel time to town	Average travel time to nearest 100,000 inhabitants
14	Travel time to recreation	Travel time to the nearest main forest/water area frequently used for recreation

A scenario for the CLUE-S model consists of a file with land requirements and a file that indicates areas where restrictions to conversion apply. Three scenarios for land use requirements (demand) are available. The first scenario file (*demand.in1*) is based on the assumption of moderate urbanisation for the period 1989 till 2008. An increase in the demand of 1.5% per year is expected for the land use types greenhouses, residential areas, industrial/commercial, recreation and airports. The demand for the land use types arable land and forest/nature will remain constant and a decrease is expected for other agriculture and grassland.

The second scenario file (*demand.in2*) is based on the assumption of a strong urbanisation in the period 1989 till 2008. The demand for the land use types residential area, industrial/ commercial, recreation and airports increases strongly with 3% per year. This results in a decrease of the land use

types other agriculture, grassland and arable land. The demand for the land use types nature/forest and greenhouses is expected to remain equal.

The third scenario file (*demand.in3*) will be used for validation of the model. This scenario is based on the real land use changes between 1989 and 1996. A linear increase or decrease of the demand is supposed for each land use type during this period. The time span between the two years is short for a validation, but no other comparable data set is yet available.

Three area restriction files are available. The default file is *region.fil*, which indicates all land areas for which land use change calculations should be made. The second file *region\_ehs.fil* is based on the default file but also indicates the areas that belong to the ecological main structure (EHS) of The Netherlands. Land use changes are not allowed in these areas. The third area restriction file is *region\_green.fil*, this file indicates 'The Green Heart', the open space with many grasslands in between the four major cities. No land use changes are allowed in this area, if this file is chosen.

Please *note* that not all combinations of area restriction file and demand files are possible, e.g., a strong urban expansion while protecting a large part of the landscape will not yield a solution.

## Exercise CLUE-1: Getting-to-know the user interface

This exercise makes you familiar with the user-interface of CLUE-S. The precise definition of the different parameters and input files is discussed in other exercises and in the user manual.

Download all data to directory d:/prac\_clue.

### 1.1 Start CLUE-S

CLUE-S can be started by:

Opening the directory where CLUE-S is installed with explorer and double-click *clues.exe*

The user-interface should appear on the screen (Figure c1.1).

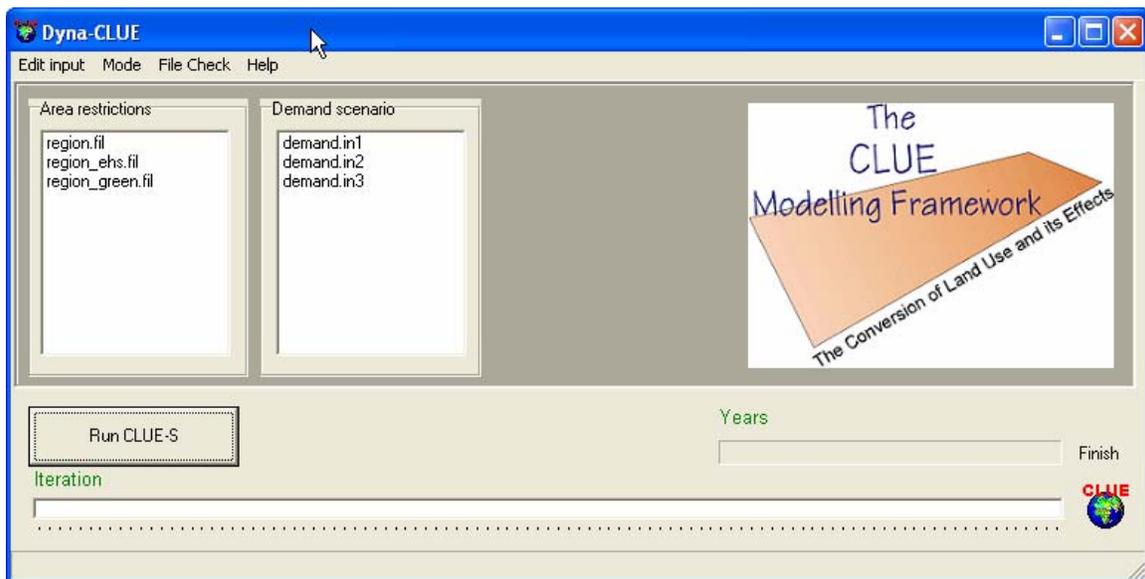


Figure c1.1. Interface of the CLUE-S model

### 1.2 Main functions

The user interface makes it possible to edit the main input files through a built-in text editor and allows the user to choose the scenario conditions. When all parameters are set the simulation can start by clicking the 'Run CLUE-S' button. Simulation results will be saved to output files that can be imported by a GIS for display and analysis.

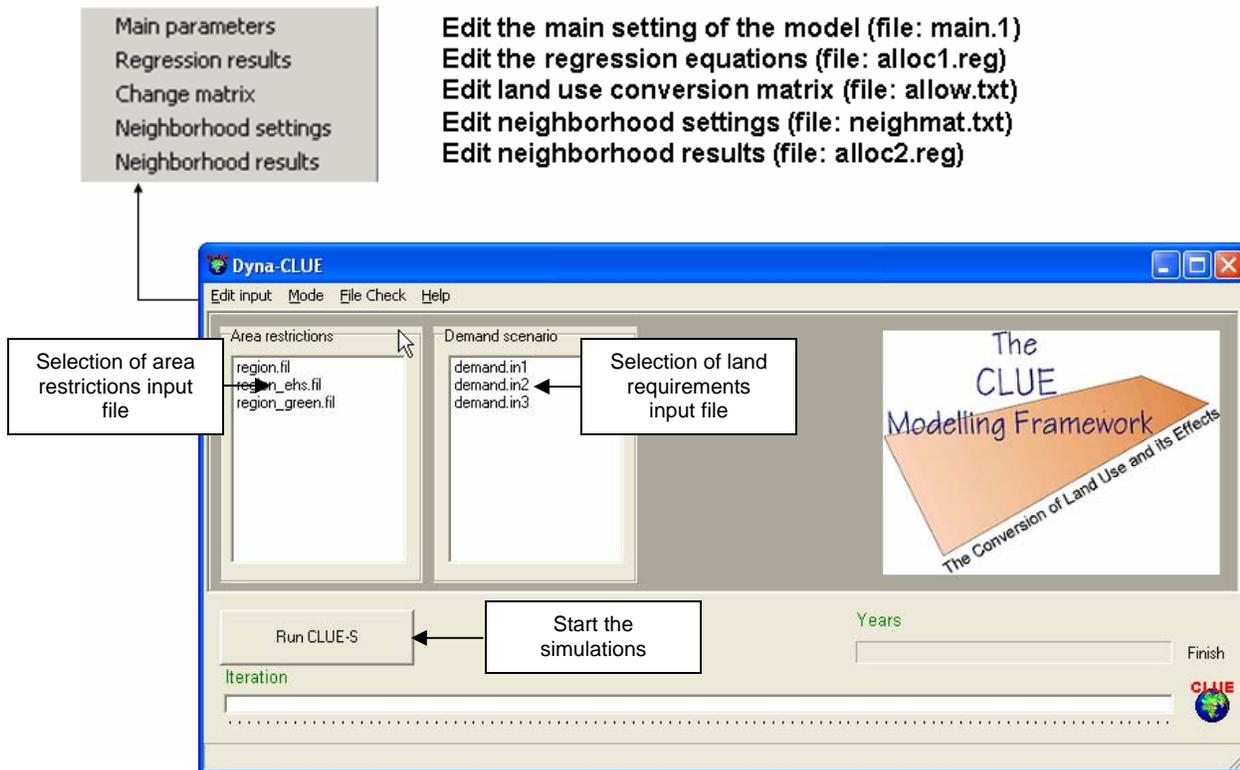


Figure c1.2. Explanation of the CLUE-S interface

### 1.3 Start the simulation

- Make sure that all input files are correctly defined
- Select *region.fil* as 'area restriction' input file

The 'area restriction' file indicates which cells of a rectangular grid are part of the case-study area and can also contain information on the locations that belong to an area with restrictions to land use conversion, e.g. a natural park.

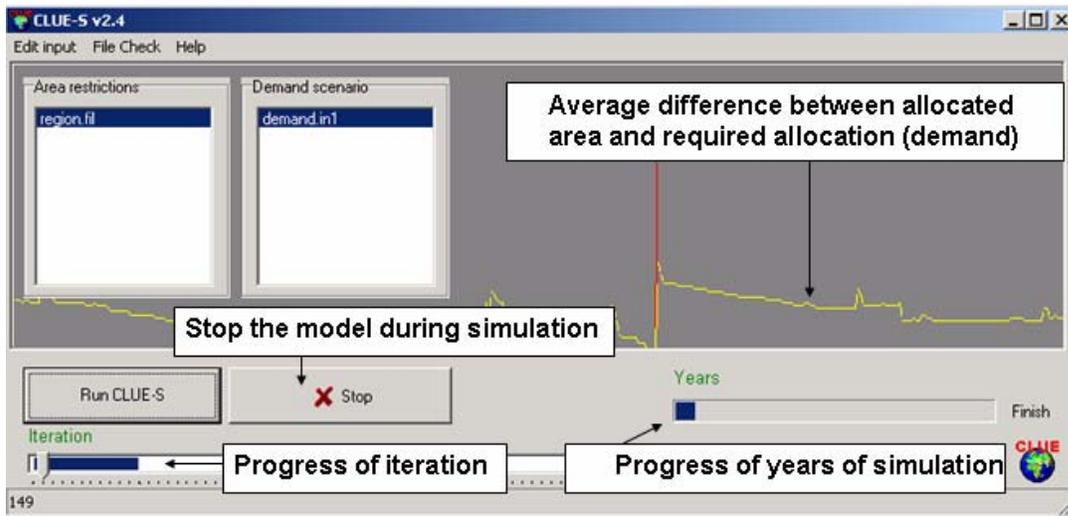
- Select *demand.in1* as 'land requirements' (demand) input file

The 'land requirements' file contains for each year that is simulated the required area of the different land use types. These claims can be calculated in other models or can be based on trend extrapolation and demographic projections. Different land requirements are possible for different scenarios.

- Click 'Run CLUE-S'.

The simulations will now start and the status bars show the progress (Figure c1.3).

**NOTE:** The status-bar for the iterative procedure shows the average difference between the allocated area of the different land use types and the required allocation of the different land use types. The simulation of one year is finished if the allocated area deviates less than the specified maximum allowed deviation. Only when one of the land use types exceeds the specified maximum deviation between allocation and requirements for one of the land use types the iterations will continue and a special indicator will appear on the screen.



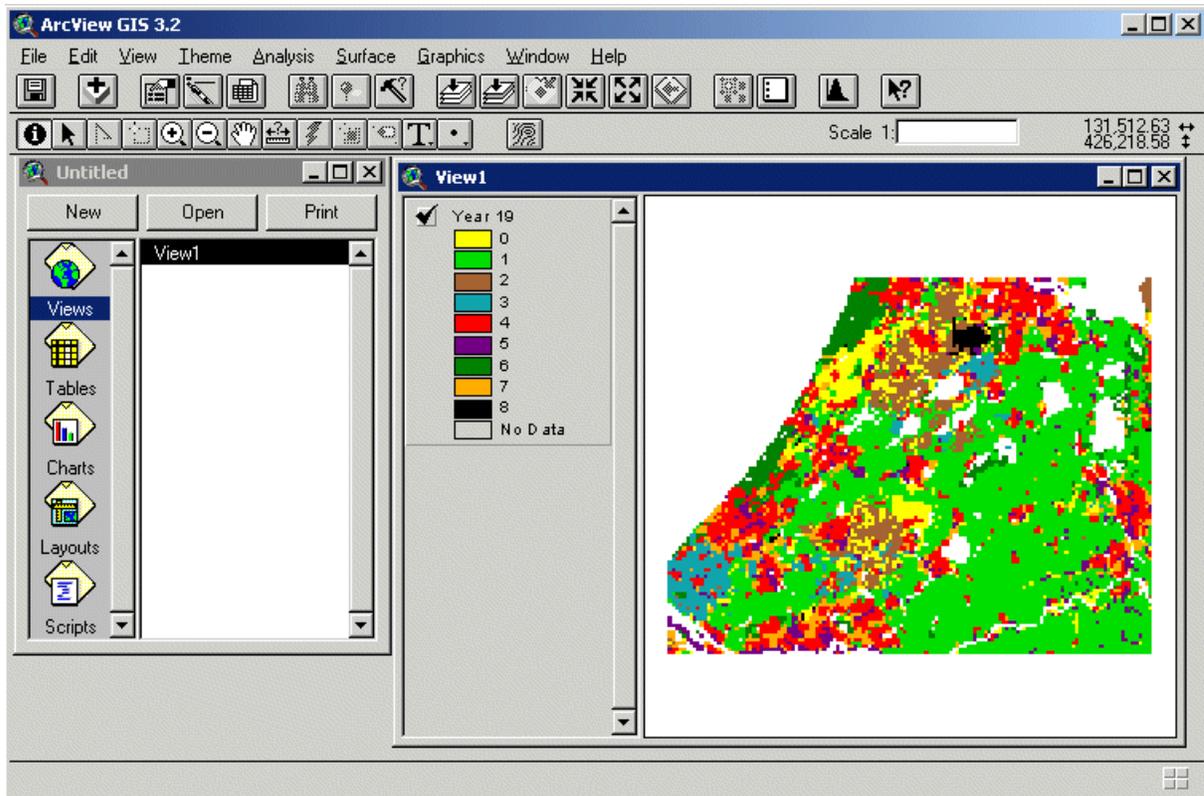
**Figure c1.3.** Explanation of the CLUE-S model run

When all simulations are made successfully the model will display the message ‘finished’ and a button that gives access to the LOG-file will appear. The log file contains information on the input files and run-time information on the iterations and may be consulted when errors occur or unexpected results are found.

#### 1.4 Display of the simulation results

All results of the simulation are saved in the installation directory. To display the simulation results it is needed to use a GIS package. For these exercises we will use the ArcView package with the Spatial Analyst Extension. Follow the steps below to display a land use map generated by the CLUE-S model:

- Start ArcView and make sure that the Spatial Analyst extension is installed and activated: File | Extensions | Check ‘Spatial Analyst’ | OK
- Open a new View by selecting ‘Views’ from the project window and click the ‘New’ button.
- Import the file with simulation results that you want to display: File | Import Data Source. Select import file type: ASCII Raster | OK.
- Select the file with simulation results. Go to the installation directory of CLUE-S. **Note:** the installation directory should not contain spaces or special characters, since this might disturb the import of ASCII files in ArcView. Set ‘List Files of Type:’ at ‘All Files (\*.\*)’. The simulation results are stored in files called: *cov\_all.\** where \* indicates the year after the start of the simulation. Select the file you want, e.g. *cov\_all.19* and click ‘OK’.
- Specify a name and directory where you want to store the resulting grid, e.g., year19.
- The program will prompt: ‘Cell values as integers?’. Click: ‘Yes’.
- The program will prompt: ‘Add grid as theme to the View?’. Click: ‘Yes’.
- The result of the simulation can now be seen on the screen and analysed using ArcView (Figure c1.4).



**Figure c1.4.** Simulation result displayed in ArcView

It is now possible to change the graphical presentation by changing the colours of the map into colour that are easily associated with the different land use type. For the simulation of The Randstad the suggested colours of Table 1.1 can be used.

**Table 1.1.** Land use types with suggested colours

Land use code	Land use type	Suggested colour
0	Other agriculture	Yellow
1	Grassland	Light green
2	Arable land	Brown
3	Greenhouses	Blue-green
4	Residential area	Red
5	Industrial/commercial	Purple
6	Forest/nature	Dark green
7	Recreation	Orange
8	Airport	Black

→ Prepare and compare maps of the results of the simulation for 1989, 2000 and 2008.

## Exercise CLUE-2: Testing model options

### 2.1 Scenario conditions

The CLUE-S model has a number of parameters that need to be specified before a simulation can be made. The setting of these parameters depends on the assumptions made for a particular scenario. In the next exercises we will explore four different scenario conditions:

- 1) Land requirements
- 2) Spatial policies (area restrictions)
- 3) Conversion elasticity
- 4) Land use conversion sequences

Different scenarios allow the comparison of different possible developments and give insight in the functioning of the model. Such analysis is most easy by visual comparison or through the calculation of the differences between the two scenarios in a GIS.

➔ Run the model with the baseline scenario: use the original settings of the ‘main parameters’, ‘regression results’ and ‘conversion matrix’, select *region.fil* and *demand.in1*. Import the results (e.g. for the start and end of the simulation, year 0 and year 19) and compare them in ArcView. ArcView maps can be exported to other graphical formats by: File | Export. Select the Windows Bitmap (\*.bmp) format for the best results.

### 2.2 Land requirements

The land requirements are input to the model. For each year of the simulation these requirements determine the total area of each land use type that needs to be allocated by the model. The iterative procedure will ensure that the difference between allocated land cover and the land requirements is minimized. Land requirements are calculated independently from the CLUE-S model itself, which calculates the spatial allocation of land use change only. The calculation of the land use requirements can be based on a range of methods, depending on the case study and the scenario. The extrapolation of trends of land use change of the recent past into the near future is a common technique to calculate the land use requirements. When necessary, these trends can be corrected for changes in population growth and/or diminishing land resources. For policy analysis it is also possible to base the land use requirements on advanced models of macro-economic changes, which can serve to provide scenario conditions that relate policy targets to land use change requirements.

Two different files with land requirements are provided with the model for the period from 1989 to 2008. The first scenario file (*demand.in1*) is based on the assumption of moderate urbanisation for the period 1989 till 2008. An increase in the demand of 1.5% per year is expected for the land use types greenhouses, residential areas, industrial/commercial, recreation and airports. The demand for the land use types arable land and forest/nature will remain constant and a decrease is expected for other agriculture and grassland.

The second scenario file (*demand.in2*) is based on the assumption of a very strong urbanisation in the period 1989 till 2008. The demand for the land use types residential area, industrial/ commercial, recreation and airports increases strongly with 3% per year. This results in a decrease of the land use types other agriculture, grassland and arable land. The demand for the land use types nature/forest and greenhouses is expected to remain equal.

➔ Select the land requirement scenario *demand.in2* and run the model keeping all other settings equal to the first run of the model. Analyse the results with ArcView through displaying the land use pattern at the start of the simulation and at the end of the simulation.

**NOTE:** Each simulation, the model will overwrite the results of a previous simulation. If you want to save the results, rename the output files or move the output files to another directory.

### 2.3 Define a demand scenario

The two previous demand scenarios are not very sophisticated, because they are just linear trends based on an urbanisation rate. In the introduction text of this chapter and in the summary of the fifth national policy document on spatial planning of the Dutch government (Ministry of VROM, 2002) policies for spatial planning are mentioned. Use this information and other information from websites (for suggestions see the reference list), to compose a more realistic and sophisticated land requirement scenario.

➔ Define your own scenario for the period 1989-2008 by generating a new land requirements input file for CLUE-S. You should be able to give reasons for your choices for each land use type. Follow the steps below to create your scenario:

- Open Microsoft Excel to facilitate the calculations. You can start to make your changes based on one of the pre-defined demand files, e.g. demand.in1.
- Specify for each year (1989-2008) the land requirements of the different land use types in a table following the specifications below:
  - Each row indicates a year; each column a land use type following the order of the land use coding.
  - Make sure to include also the land requirements for 1989 (year 0). These should be similar to the land use map of 1989. If you base your calculations on the number of pixels in the map you should multiply these values by 25, because the map has a resolution of 500 meter, and the demand is given in hectares.
  - The total land area required should equal the area in the *region.fil* file, i.e. 279,625 ha. The sum of the values on each row should equal this amount for each year.

	A	B	C	D	E	F	G	H	I	J
1	0	1	2	3	4	5	6	7	8	
2	1989	32425	132775	16325	10200	40875	12900	15350	16900	1875
3	1990	30889	132696	16404	10489	41382	13250	15396	17218	1900
4	1991	29354	132618	16482	10779	41889	13600	15443	17536	1925
5	1992	27818	132539	16561	11068	42396	13950	15489	17854	1950
6	1993	26282	132461	16639	11357	42904	14300	15536	18171	1975
7	1994	24746	132382	16718	11646	43411	14650	15582	18489	2000
8	1995	23211	132304	16796	11936	43918	15000	15629	18807	2025
9	1996	21675	132225	16875	12225	44425	15350	15675	19125	2050
10	1997	21458	131472	16791	12286	44958	15580	15675	19355	2050
11	1998	21244	130704	16707	12348	45498	15814	15675	19587	2050
12	1999	21031	129920	16623	12409	46044	16051	15675	19822	2050
13	2000	20821	129120	16540	12471	46596	16292	15675	20060	2050
14	2001	20613	128304	16457	12534	47155	16536	15675	20300	2050
15	2002	20407	127473	16375	12596	47721	16784	15675	20544	2050
16	2003	20203	126625	16293	12659	48294	17036	15675	20791	2050
17	2004	20000	125760	16212	12723	48873	17292	15675	21040	2050
18	2005	19800	124879	16131	12786	49460	17551	15675	21292	2050
19	2006	19602	123982	16050	12850	50053	17814	15675	21548	2050
20	2007	19406	123067	15970	12914	50654	18082	15675	21807	2050
21	2008	19212	122136	15890	12979	51262	18353	15675	22068	2050
22										

- When all values have been defined, select the values (without land use type names and year numbers) and paste the contents into a text editor (e.g. Notepad). Insert a line at the top of the file with the number of lines (years) for which the land requirements are specified (20 in our example).

```
20
32425 132775 16325 10200 40875 12900 15350 16900 1875
31452 132135 16403 10490 41382 13250 15396 17218 1900
30509 131434 16482 10788 41895 13609 15442 17541 1925
29593 130674 16561 11094 42414 13977 15489 17871 1951
28706 129854 16641 11409 42940 14356 15535 18207 1977
27844 128975 16721 11733 43473 14745 15582 18549 2003
27009 128036 16801 12066 44012 15145 15628 18898 2030
26199 127037 16881 12409 44558 15555 15675 19253 2057
25413 125979 16963 12761 45110 15977 15722 19615 2084
24650 124862 17044 13124 45670 16410 15769 19984 2112
23911 123685 17126 13496 46236 16855 15817 20360 2140
23194 122448 17208 13880 46809 17311 15864 20743 2168
22498 121151 17291 14274 47390 17780 15912 21133 2197
21823 119794 17374 14679 47977 18262 15960 21530 2226
21168 118376 17457 15096 48572 18757 16007 21935 2256
20533 116898 17541 15525 49174 19266 16055 22347 2286
19917 115358 17625 15966 49784 19788 16104 22767 2316
19320 113757 17710 16419 50402 20324 16152 23195 2347
18740 112094 17795 16886 51027 20875 16200 23631 2378
18178 110368 17880 17365 51659 21440 16249 24076 2410
```

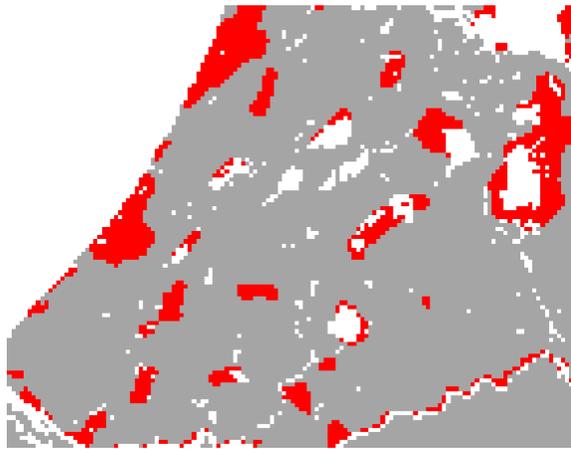
- Save this file in the installation directory as *demand.in\** where \* can be defined by the users, e.g. *demand.in4*.
- Restart the CLUE-S model; it is now possible to select the new land requirement file and simulate the land use changes.
- Analyse the results.

### Exercise CLUE-3: Spatial policies (area restrictions)

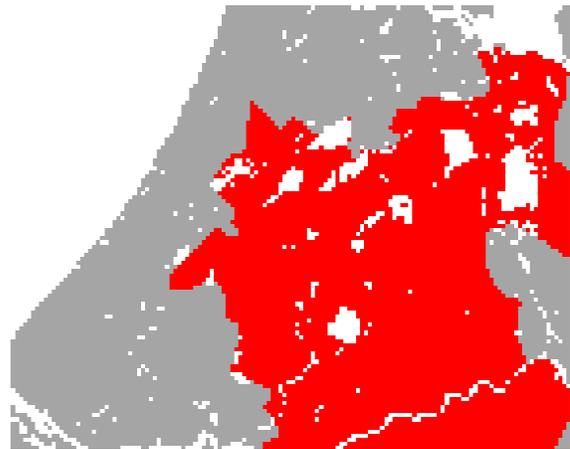
This option indicates areas where land use changes are restricted through spatial (land use) policies or tenure status. Maps that indicate the areas for which the spatial policy is implemented must be supplied. Some spatial policies restrict all land use change in a certain area, e.g., when in a forest reserve all logging is banned. Other land use policies restrict a set of specific land use conversions, e.g., residential construction in designated agricultural areas. In this exercise we will only address policies that restrict all land use changes in designated areas.

With this version of the model for the Randstad three area restriction files are supplied that can be selected through the user-interface. Each file contains a map designating the areas where land use change is restricted. The maps are shown in the figures below, but can also be imported in ArcView as ASCII Raster file similar to the procedure used to import the results of the simulations. The files are located in the installation directory.

<i>Region.fil</i>	Default file that indicates all land areas for which land use change calculations should be made.
<i>Region_ehs.fil</i>	Area restriction file based on the default file but also indicates the areas that belong to the ecological main structure (EHS) of The Netherlands. Land use changes are not allowed in these areas.
<i>Region_green.fil</i>	Area restriction file that indicates 'The Green Heart', the open space with many grasslands in between the four major cities. No land use changes are allowed in this area.



Ecological main structure



Green heart

➔ Run the CLUE-S model with the different area restriction files keeping all other settings equal to the first run of the model. Compare the results with the initial situation (1989, year 0) and compare the impact of the different area restrictions.

- Is strict protection of The Green Heart needed to keep an open green area? **NOTE:** The assumption that no changes will occur in the EHS is unrealistic because these areas are designated for nature development and change of agricultural into nature is expected. The model has also the possibility to restrict only certain conversions such as residential construction. This will be explained in more detail in Exercise 6.
- Which land use type is most affected by the protection of the ecological main areas? (**\*EX10**)

**NOTE:** Each simulation, the model will overwrite the results of a previous simulation. If you want to save the results, rename the output files or move the output files to another directory.

#### Exercise CLUE-4: Conversion elasticity

The conversion elasticity is one of the land use type specific settings that determine the temporal dynamics of the simulation. The conversion elasticity is related to the reversibility of land use changes. Land use types with high capital investment or irreversible impact on the environment will not easily be converted in other uses as long as there are land requirements for those land use types. Such land use types are therefore more ‘static’ than other land use types. Examples of relatively static land use types are residential areas, but also plantations with permanent crops (e.g., fruit trees). Other land use types are more easily converted when the location becomes more suitable for other land use types. Arable land often makes place for urban development while expansion of agricultural land can occur at the same time at the forest frontier.

These differences in behaviour towards conversion of the different land use types can be approximated by the conversion costs. However, costs cannot represent all factors that influence the decisions towards conversion such as nutrient depletion, esthetical value etc. Therefore, in the model we have assigned each land use type a dimensionless factor that represents the relative elasticity to conversion, ranging from 0 (easy conversion) to 1 (irreversible change). The user should specify this factor based on expert knowledge or observed behaviour in the recent past. An extended explanation of the possible values of the conversion elasticity and how behaviour changes when the land requirements increase or decrease in time is given below.

- 0: Means that all changes for that land use type are allowed, independent from the current land use of a location. This means that a certain land use type can be removed at one place and allocated at another place at the same time, e.g. shifting cultivation.
- >0...<1: Means that changes are allowed, however, the higher the value, the higher the preference that will be given to locations that are already under this land use type. This setting is relevant for land use types with high conversion costs.
- 1: Means that grid cells with one land use type can never be added and removed at the same time. This is relevant for land use types that are difficult to convert, e.g., urban settlements and primary forests. A value of one stabilizes the system and prevents that in case of deforestation other areas are reforested at the same time.

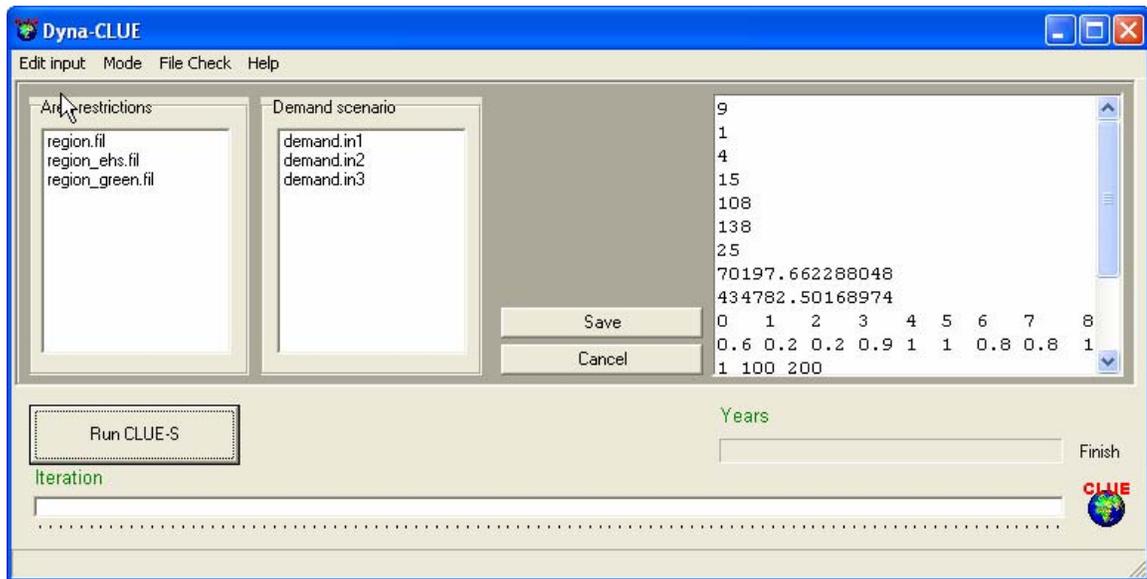
The conversion elasticity is specified in the ‘Main Parameters’ input file that can be edited through the user interface (click ‘Main Parameters’ in the ‘Edit Input’ menu) or with a text editor in the file ‘main.1’ in the installation directory. The conversion elasticity of each land use type is specified on line 11 of this file (an explanation of all other parameters in this file can be found in the user manual). The first conversion elasticity corresponds with land use type 0, the second with land use type 1, etc.

**Table 4.1.** Current settings of the conversion elasticities

Code	Land use type	Conversion elasticity
0	Other agriculture	0.6
1	Grassland	0.2
2	Arable land	0.2
3	Greenhouses	0.9
4	Residential area	1
5	Industrial/commercial	1
6	Forest/nature	0.8
7	Recreation	0.8
8	Airports	1

→ Run the CLUE-S model with alternative settings for the conversion elasticity. Change the conversion elasticity by:

- Click on 'Main Parameters' under 'Edit Input'. The main parameters can now be edited.
- Line 11 contains the conversion elasticity settings of the different land use types in the same order as the land use type coding. Change these values to new values.
- 



- Click on 'Save'.
- Run the model after selecting the 'Area restrictions file' and the 'Land requirements' file for the base scenario.
- Display the results with ArcView.
- Compare the differences in spatial pattern of land use change as result of the changes in conversion elasticity. Try to explain why your changes resulted in a different land use pattern.

**NOTE:** Each simulation, the model will overwrite the results of a previous simulation. If you want to save the results, rename the output files or move the output files to another directory.

### Exercise CLUE-5: Land use conversion sequences

Not all land use changes are possible and some land use changes are very unlikely (e.g., arable land cannot be converted into primary rain forest). Many land use conversions follow a certain sequence or cycle, e.g. fallow land and forest regrowth often follow shifting cultivation. The conversions that are possible and impossible are specified in a land use conversion matrix. For each land use type it is indicated in what other land use types it can be converted during the next time step. Depending on the definition of this conversion matrix and the time-steps chosen, complex land use sequences are possible.

The land use conversion matrix can be edited by clicking ‘Change matrix’ under ‘Edit Input’. It is also possible to use a text editor (e.g. Notepad) to edit the file *allow.txt* in the installation directory. The rows of this matrix indicate the land use types during time step  $t$  and the columns indicate the land use types in time step  $t+1$ . If the value of a cell is 1 the conversion is allowed while a 0 indicates that the conversion is not possible. The rows and columns follow the number code of the land use types.

Example: in the matrix below all conversion are possible except the conversion from residential area, industrial/commercial and airports into other agriculture, grassland, arable land or forest/nature.

	Other agriculture	Grassland	Arable land	Greenhouses	Residential area	Industrial/commercial	Forest/nature	Recreation	Airports
Other agriculture	1	1	1	1	1	1	1	1	1
Grassland	1	1	1	1	1	1	1	1	1
Arable land	1	1	1	1	1	1	1	1	1
Greenhouses	1	1	1	1	1	1	1	1	1
Residential area	0	0	0	1	1	1	0	1	1
Industrial/commercial	0	0	0	1	1	1	0	1	1
Forest/nature	1	1	1	1	1	1	1	1	1
Recreation	1	1	1	1	1	1	1	1	1
Airports	0	0	0	1	1	1	0	1	1

Code	Land use type
0	Other agriculture
1	Grassland
2	Arable land
3	Greenhouses
4	Residential area
5	Industrial/commercial
6	Forest/nature
7	Recreation
8	Airports

➔ Run the baseline scenario for the Randstad with a different setting of the conversion matrix (keeping all other settings equal) and analyse the differences in outcome with ArcView. Compare the results.

**Note:** Some land use conversion settings will have no effect because they are overruled by the conversion elasticity and land requirement settings.

Each simulation, the model will overwrite the results of a previous simulation. If you want to save the results, rename the output files or move the output files to another directory.

## Exercise CLUE-6: Area restricted land use conversions

Certain land use conversions are only allowed in specific areas due to policy constraints: e.g. the conversion of agricultural land into greenhouses is only allowed in The Westland. The CLUE-S model has an option in the conversion matrix to include these area specific restrictions. A map indicates where the changes are allowed. The map should be named *sc1gr\*.fil* with \* denoting a unique number, which follow the highest of the driving factor files. An example:

	LU0	LU1	LU2
LU0	1	1	20
LU1	1	1	1
LU2	1	1	1

In this example the conversion of land use type 0 into land use type 2 is only allowed in areas indicated in file *sc1gr20.fil*. This file has a similar format as the area restriction file with a value 1 in areas where this conversion is allowed and a value 0 for areas where this change is not allowed. **Note:** the number of driving factor files (line 4) in the main parameter file (*main.1*) has to be changed according to this setting.

For this exercise you will simulate a policy scenario in which nature conservation and development is strongly stimulated by the government. In this scenario it is assumed that the government wants to increase the nature area in the Randstad by 15,000 ha for the period 1989-2008. However, this growth should take places in the areas designated for the ecological main structure (EHS).

➔ Run the simulation for the scenario with strong nature development in the EHS areas. Follow the steps below:

- Adapt the land use requirement file *demand.in1* as such that a gradual increase with 15,000 ha of forest/nature at the expense of grassland and arable land is simulated for the period 1989-2008. Make sure that the total demand for all land use types remains equal.
- Import the file with the EHS areas (*region\_ehs.fil*) in ArcView (File | Import Data Source | ASCII Raster).
- Reclassify the -9998 values into 1 (Analysis | Reclassify | Unique) to create a file that indicates areas where the specific land use conversions are allowed.
- Save the theme by: Theme | Save data set and give the data set a name
- Export this data set (map) to an ASCII file (File | Export Data Source | ASCII Raster) and name it *sc1gr15.fil*.
- Edit the conversion matrix as such that the conversion to forest/nature is only allowed in areas of the EHS.
- Adapt the main parameter file (*main.1*) by changing the number on line 4, which indicates the number of driving factor files.
- Run the simulation with *region.fil* and the adapted demand file.

➔ Run the simulation again with the same demand scenario, but now without the area restrictions for the conversion to forest/nature (edit the conversion matrix).

➔ Analyse the results and compare the maps. Is the designation of EHS areas necessary for concentrated nature development?

## Exercise CLUE-7: Creating a new application

In the previous exercises you explored the CLUE-S model and learned how to adapt the different input files and scenarios. The results of the simulation were not very realistic, because the model has not been validated. The conversion matrix, the conversion elasticities and the regression results are all rough estimations. In this exercise you will create yourself a new application for the Randstad for simulation of the period 1989 till 1996. A second land use map for 1996 is available. This gives the opportunity to validate the model that you have created, which will be done in Exercise CLUE-8.

### 7.1 Selection of location factors

One of the most important input files is the regression result. This file contains the parameters of the logistic regression for each land use type. The CLUE-S model uses a statistical analysis to define the suitability of locations for different land use types. The suitability of a location is a function of a number of case-study specific location factors, such as soil quality, accessibility, socio economic conditions, etc. For this exercise you will have to make a statistical analysis to relate these location factors to the suitability of the different land use types by a logit model. The driving factors can be positive related with the land use type (e.g. population density with residential areas) or negative (e.g. population density with forest/nature). Table 7.1 presents the available ArcView ASCII grid files of the land use types and location factors. These files can be imported in ArcView through File | Import Data Source | ASCII Raster.

**Table 7.1.** Available land use type and driving factor files

Code	Filename	Description
0	Cov0.0	Other agriculture
1	Cov1.0	Grassland
2	Cov2.0	Arable land
3	Cov3.0	Greenhouses
4	Cov4.0	Residential areas
5	Coc5.0	Industrial/commercial
6	Cov6.0	Forest/nature
7	Cov7.0	Recreation
8	Cov8.0	Airports
0	Sc1gr0.fil	Population density
1	Sc1gr1.fil	Elevation
2	Sc1gr2.fil	Groundwater level
3	Sc1gr3.fil	Distance to water
4	Sc1gr4.fil	Organic matter content
5	Sc1gr5.fil	Calcium content
6	Sc1gr6.fil	Clay content
7	Sc1gr7.fil	Travel time to station
8	Sc1gr8.fil	Travel time to airport
9	Sc1gr9.fil	Travel time to harbour
10	Sc1gr10.fil	Travel time to work
11	Sc1gr11.fil	Travel time to highway
12	Sc1gr12.fil	Travel time to city (500,000 inhabitants)
13	Sc1gr13.fil	Travel time to town (100,000 inhabitants)
14	Sc1gr14.fil	Travel time to recreation

➔ Prepare for each land use type a hypothesis of the potential location factors (from Table 7.1) that might influence the suitability of a location for that land use type. Select for each land use type 3 to 5

driving factors. As help you can import the current land use map (*cov\_all.0*) and the driving factors in ArcView to see their spatial distribution and to analyse the relations between the driving factors and land use types.

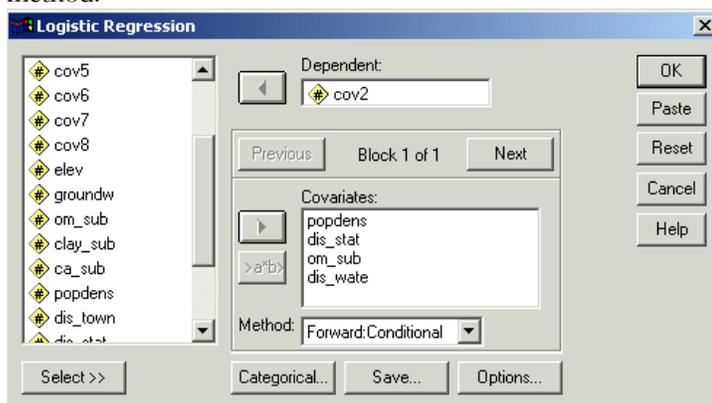
## 7.2 Performing the statistical analysis

In this exercise you will import the land use maps and the driving factors into SPSS for statistical analysis. A file called *stat.txt* contains all data for land use types and driving factors in tabular format for each grid cell. Normally, you can create the file yourself through the file conversion tool. Each column contains the values of a file (first column: first file according to Table 7.1, etc.); each row contains the values for a single grid cell for all files.

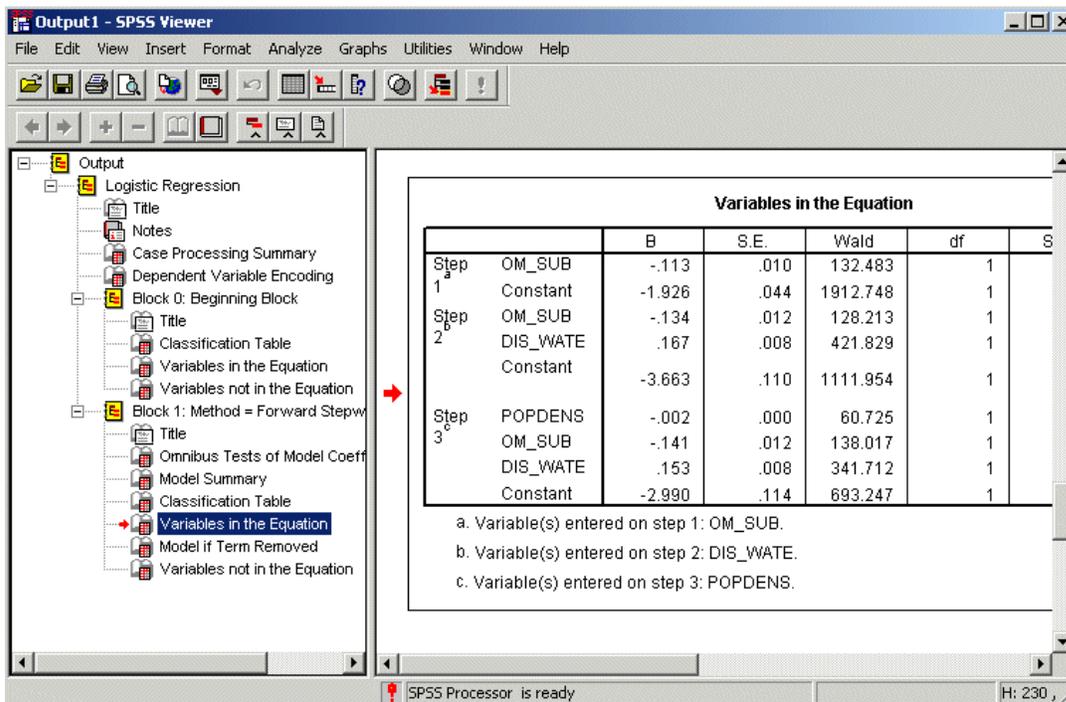
➔ Import the *stat.txt* data file into SPSS by starting the SPSS software and selecting: 'File | Read Text Data' and select the just produced file *stat.txt* and go through the import wizard. This will import all data. The names of the columns are not imported and should be specified by the user. The order of the columns equals the order of the files specified in the Table 7.1. In SPSS it is possible to edit the names in the 'Variable View' tab. This will facilitate the analysis.

➔ Perform a logistic regression analysis to test the hypotheses you have specified for the different land use types.

- Click: Analyze | Regression | Binary Logistic.
- Select a land use type as dependent variable and the location factors you have selected as independent variable.
- Select a method for regression: 'Enter' if you want to include all selected variables or one of the other methods for stepwise regression. Select for this exercise the 'Forward: Conditional' method.



- Click the 'Save' button and check the probabilities option to save the predicted probabilities.
- Click the 'Options' button to set the conditions for stepwise regression. In the 'probability for stepwise' section you can indicate the significance values for Entry and Removal during the stepwise procedure. For large data sets these values can be set at respectively 0.01 and 0.02.
- Run the regression by clicking 'OK'.
- Output will be generated in a special window. Use the help function if you have problems interpreting the output.



**Figure 7.1.** Output section of the logistic regression with SPSS

The ‘Variable in the equation’ table in the ‘Block 1’ output section contains the resulting logit model (Figure 7.1). The  $\beta$  values for the driving factors are the values in the column B for the last step, in the example of Figure 7.1 the  $\beta$  value for population density (POPDENS) is -0.002. By default only 3 decimals are shown, this number can be increased in the Cell properties window. Select the specific cells and go to ‘Format | Cell Properties’.

### 7.3 Evaluating the goodness of fit with the ROC method

The ROC characteristic is a measure for the goodness of fit of a logistic regression model similar to the  $R^2$  statistic in Ordinary Least Square regression (Pontius and Schneider, 2001). A completely random model gives a ROC value of 0.5; a perfect fit results in a ROC value of 1.0.

➔ Calculate the ROC value for your regression model.

- Click: Graphs | ROC Curve. Select the observed land use as ‘State variable’ with value 1 for occurrence of this land use type and the ‘Predicted probability’ as ‘Test variable’. Click ‘OK’ (Figure c7.2).
- The results will show a ROC curve and the ‘area under the curve’ that is the test result (Figure c7.3).

When the ROC value is very low (e.g. below 0.700), you should reconsider your driving factors and do the logistic regression again with other driving factors. When you are satisfied with the ROC value, write down the resulting  $\beta$  values for the driving factors and the constant ( $\beta_0$ ).

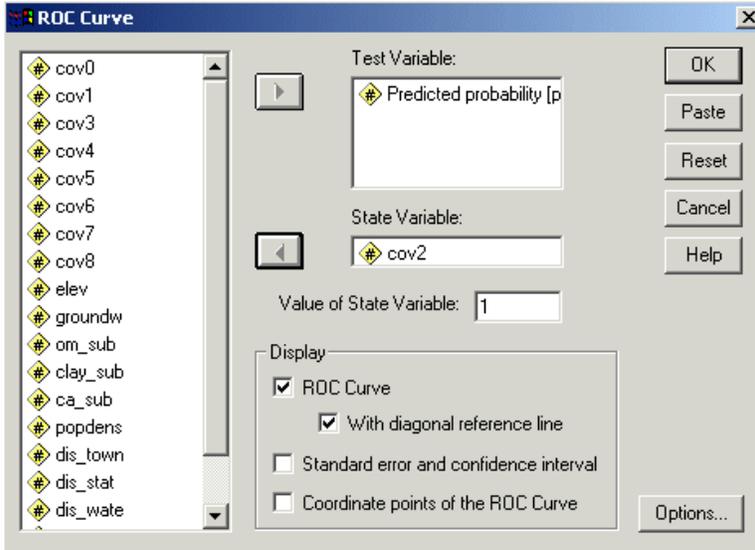


Figure c7.2. Window for ROC curve configuration

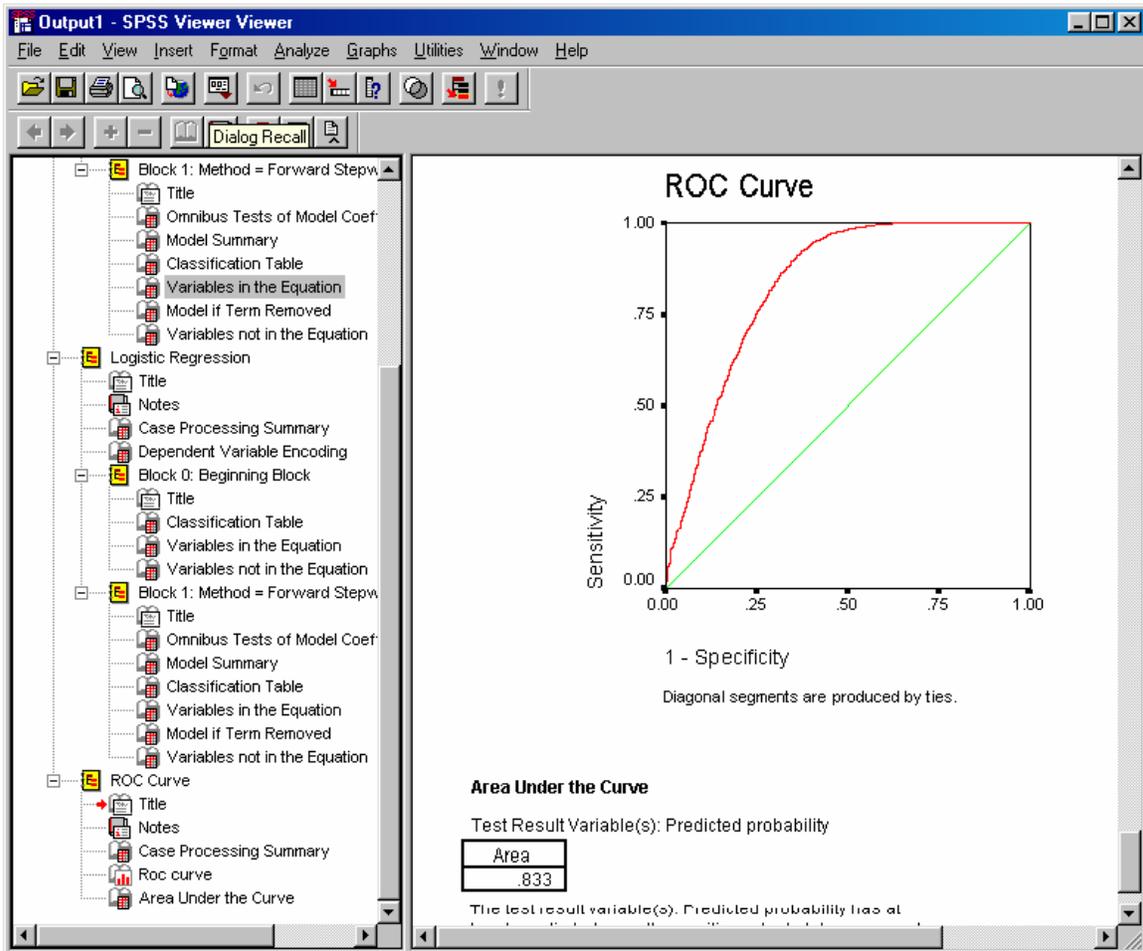
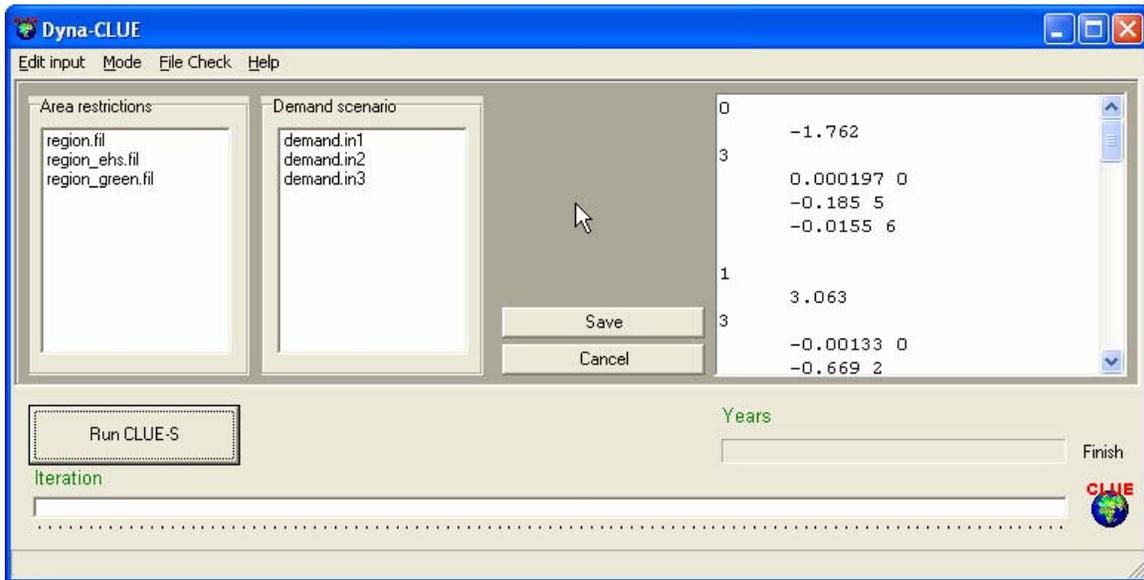


Figure c7.3. ROC curve of a logistic regression with SPSS

## 7.4 Using the regression results as input to the CLUE-S model

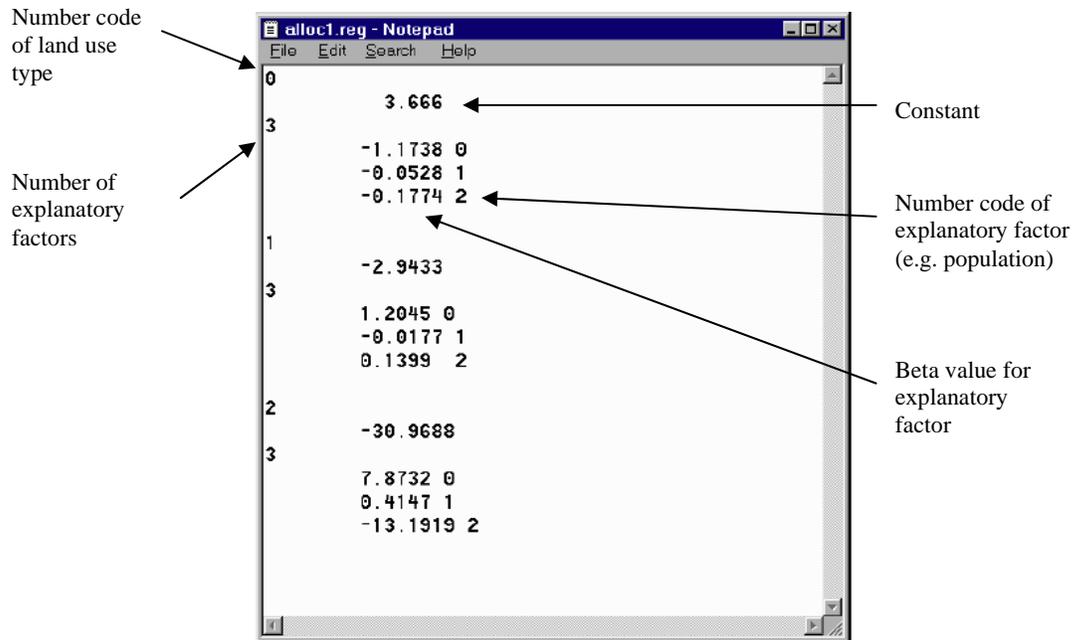
The regression results are used by the CLUE-S model to determine the suitability of the locations for the different land use types. For each land use type considered a separate regression model is used. The result of the analysis in SPSS should therefore be translated to input for the CLUE-S model. The input file for the regression equations is called *alloc1.reg* and is located in the installation directory and can be edited through the user interface of CLUE-S (Figure c7.4).



**Figure c7.4.** Editor for the regression results

This file is structured as follows (Figure c7.5):

- Line 1: Number code for land use type (e.g. grassland).
- Line 2: Constant of regression equation for land use type ( $\beta_0$ ).
- Line 3: Number of explanatory factors (sc1gr#.fil files) in the regression equation for that land use type.
- Line 4 and further: On each line the beta coefficients ( $\beta_1$ ,  $\beta_2$ , etc.) for the explanatory factor and the number code of the explanatory factor.



**Figure c7.5.** Structure of the regression results file (*alloc1.reg*)

➔ Use the output of your regression analysis to change the input file for CLUE-S. Use the coding system of the land use types and location factors, similar to the coding used in Table 7.1 and the filenames *sc1gr\*.fil* where \* is the code for the driving factor. Save your changes.

To visualize the result of the logistic regression you can create probability maps based on the regression equations. Use the ‘Calculate probability maps’ button in CLUE-S under ‘Mode’, and click ‘Run CLUE-S’. **Note:** you have to select an area restriction and demand scenario file, although these do not influence the probability map. The probability map for each land use type is stored in *prob1\_\*.1* where \* indicates the land use code. These maps can be imported as floating point grids in ArcView ‘File | Import Data Source | ASCII Raster’. **Note:** do not import as integers.

## 7.5 Change other input data and settings

The parameters for the regression results have been imported, but other parameters still have to be adapted. The conversion matrix, the conversion elasticities and the main parameters will be adapted according to your own settings. The current conversion matrix is rather simple and some unlikely land use changes are still possible.

1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
0	0	0	1	1	1	0	1	1
0	0	0	1	1	1	0	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
0	0	0	1	1	1	0	1	1

Code	Land use type
0	Other agriculture
1	Grassland
2	Arable land
3	Greenhouses
4	Residential area
5	Industrial/commercial
6	Forest/nature
7	Recreation
8	Airports

➔ Edit the conversion matrix as such that it will be more realistic for the simulation of land use change between 1989 and 1996.

The next parameters to change are the conversion elasticities (Table 6.2). The CLUE-S model is quite sensitive for these conversion elasticities. The conversion elasticity is specified in the 'Main Parameters' input file that can be edited through the user interface. The conversion elasticity of each land use type is specified on line 11 of this file. The first conversion elasticity corresponds with land use type 0, the second with land use type 1, etc.

**Table 7.2.** Current settings of the conversion elasticities

Code	Land use type	Conversion elasticity
0	Other agriculture	0.6
1	Grassland	0.2
2	Arable land	0.2
3	Greenhouses	0.9
4	Residential area	1
5	Industrial/commercial	1
6	Forest/nature	0.8
7	Recreation	0.8
8	Airports	1

➔ Edit the conversion elasticities for the simulation of land use change between 1989 and 1996.

As last you have to change some settings in the main parameter file, besides the conversion elasticities. Each line of the main parameter file describes one or more parameters of the model (Table 7.3). Due to the changes in the regression equations, the maximum number of independent variables in a regression equation (line 3) maybe has to be adapted. Also the end year of the simulation (line 13) should be adapted. The other parameters can remain the same. Only you should change the number of driving factors (line 4), when you add an extra driving factor.

**Table 7.3.** Main parameters (see manual for further explanation)

Line	Description	Format
1	Number of land use types	Integer
2	Number of regions	Integer
3	Max. number of independent variables in a regression equation	Integer
4	Total number of driving factors	Integer
5	Number of rows	Integer
6	Number of columns	Integer
7	Cell area	Float
8	xll coordinate	Float
9	yll coordinate	Float
10	Number coding of the land use types	Integers
11	Codes for conversion elasticities	Float
12	Iteration variables	Float
13	Start and end year of simulation	Integers
14	Number and coding of explanatory factors that change every year	Integers
15	Output file choice	1, 0, -2 or 2
16	Region specific regression choice	0, 1 or 2
17	Initialization of land use history	0, 1 or 2

➔ Run the CLUE-S model with the new settings for the period 1989-1996. Use the *region.fil* and *demand.in3* for the simulation. Import the result (*cov\_all.7*) to ArcView and compare it with the actual land use of 1996 (*cov1996.asc*). When you are not satisfied with the result you can change one or more of the parameters and do the simulation again. You might even create and add a new driving factor.

### Exercise CLUE-8: Validation

In exercise CLUE-7 you compared visually the result of the simulation with the real land use pattern in 1996. This is a rather subjective manner of validation. No generally agreed upon method for the quantitative evaluation of the goodness of fit for simulation models has evolved yet. However, in this exercise you will use a more objective manner according to a multiple resolution procedure for the model goodness of fit. This method has been developed by Costanza (1989). It is based on the measuring of similarity of the patterns, and the idea that measurement at one resolution is not sufficient to describe complex patterns. The method yields indices that summarize the way the fit changes as the resolution of measurement changes. An expanding 'window' is used to gradually degrade the resolution of the comparison.

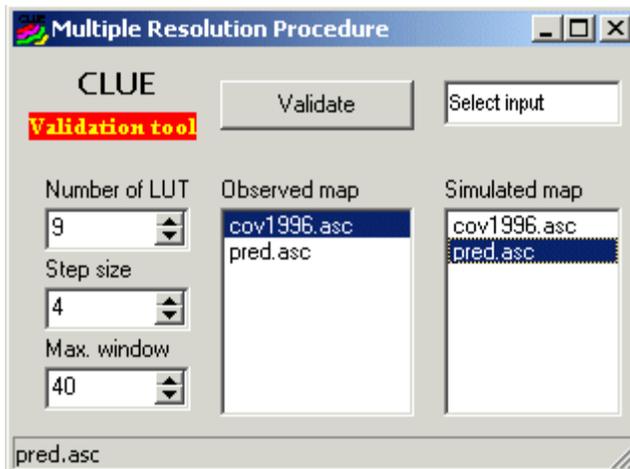
With this method the near misses, besides the direct hits, receive some weight and tell whether the pattern was being relatively well matched. The fit for each sampling window is estimated as 1 minus the proportion of cells that would have to be changed to make the sampling windows each have the same number of cells in each category, regardless of their spatial arrangement. A formula was developed that retains more of the information about the relative proportions of categories within the sample window in estimating fit while decreasing the resolution.

$$F_w = \frac{\sum_{s=1}^{t_w} \left[ 1 - \frac{\sum_{i=1}^p |a_{1i} - a_{2i}|}{2w^2} \right]_s}{t_w}$$

$F_w$  is the fit for sampling window size  $w$ ,  $w$  the dimension of one side of the (square) sampling window,  $a_{ki}$  the number of cells of category  $i$  in scene  $k$  in the sampling window,  $p$  the number of different categories (i.e., land use types) in the sampling windows,  $s$  the sampling window of dimension  $w$  by  $w$  which slides through the scene one cell at a time and  $t_w$  the total number of sampling windows in the scene for window size  $w$ .

A small program has been developed to calculate these fits for different window sizes. This program (*validation.exe*) can be found in the installation directory of the CLUE-S model. First the observed map and the simulated map, which should be *.asc* files, have to be selected. The cell size and extent of these files should be exactly the same. Then the number of land use types, the step size of the expanding window and the maximum window size have to be indicated (Figure c8.1). After running the procedure an output file *output.txt* is created, it contains two columns with the window size and with the calculated fit.

➔ Run the validation program for your simulation result of exercise 7. First rename the file *cov\_all.7* to *pred.asc*. Open *validation.exe* and select observed map (*cov1996.asc*) and simulated map (*pred.asc*). Set the number of land use types at 9, step size at 4 and maximum window size at 40. Then click 'Validate' to run the procedure. Copy the results of *output.txt* to Excel and make a graph of the results.



**Figure c8.1.** Validation program with multiple resolution procedure

To determine an overall degree of fit between two maps the information in the plot of window sizes versus fit must be summarized. A weighted average of the fits at different window sizes is a possible way of summarizing the overall fit that allows more weight to be given to smaller window sizes while not totally ignoring the large window sizes. For this purpose the following formula can be used:

$$F_t = \frac{\sum_{w=1}^n F_w e^{-k(w-1)}}{\sum_{w=1}^n e^{-k(w-1)}}$$

$F_t$  is a weighted average of the fits over all window sizes,  $F_w$  the fit for sampling windows of linear dimension  $w$ ,  $k$  a constant, and  $w$  the dimension of one side of the (square) sampling window. The value of  $k$  determines how much weight is to be given to small versus large sampling windows. A default value for  $k = 0.1$ .

➔ Calculate the weighted average of the fits ( $F_t$ ) for your results based on the formula above.

The calculated fit of goodness of the model is only a relative value that can be used to compare the results of different simulations. However, the model performance should be compared because a large pattern is always simulated correctly. It should also be compared with a certain default, especially for simulations where not many cells are changing. In that case a high value for  $F_t$  can be obtained, without giving information about how good the model is. Therefore you have to compare the calculated fit with the fit of a certain 'null model', for example the initial conditions. This indicates if the model has any additional value.

➔ Run the validation procedure but now for the initial land use map (*cov\_all.0*). Copy the results to Excel and plot the data together with calculated fit for the predicted land use map. Calculate also the  $F_t$  value for the initial land use map. What is your conclusion; does your model have any additional value?

➔ Compare the validation results with other students. Discuss the outcomes and try to improve the model by adapting the regression results (driving factors), conversion matrix and/or conversion elasticities. Then run the CLUE-S model again and calculate the new fit.

### Exercise CLUE-9: Create your own scenario

Write a storyline of a scenario for the development of the Randstad region for the period 1996 (most recent data) to 2016. The storyline should include the background of the expected developments in the region (e.g., large urbanization is expected because Netherlands has become the Capital of the United States of Europe). The storyline should contain aspects that describe the quantity of change and regulations and policies that determine the location and nature of change.

Translate your storyline to model conditions including:

- the demand for the different land use types
- restricted areas
- elasticities to change (e.g., policies that restrict the re-location of certain land use types)
- conversion settings (including spatial policies aimed at certain conversions).

→ Use the regression equations your validated model (exercise CLUE-8) to run the model for 20 years.

→ You must reconfigure your model to start the simulation in 1996:

1. delete file: cov\_all.0; rename file cov1996.asc into cov\_all.0
2. reconfigure the demand file in such a way that it contains 21 lines (one for each year of the simulation, the first line containing the areas of the different land use types in 1996 (calculate these based on the areas in the map cov1996.asc using ArcView taking note of the cell area of 25 hectares.
3. Change the years of simulation in the main parameter file to 1996 2016.

→ Run your scenario and analyze the outcome.

→ Prepare 1 map, table, diagram or graph based on your results that shows clearly the most important information relevant to policy makers that are not interested in land use modeling itself. Prepare a short powerpoint presentation to present your results.

### References of CLUE exercises

- Ministry of VROM, 2002. Fifth national policy document on spatial planning of the Dutch government.
- Pontius, R.G. and Schneider, L.C. 2001. Land use change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, 85: 239-248.
- Veldkamp, A. and Fresco, L.O. 1996. CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. *Ecological modelling*, 91: 231-248.
- Verburg, P.H., Veldkamp, A., de Koning, G.H.J., Kok, K. and Bouma, J. 1999. A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. *Ecological modelling*, 116: 45-61.
- Verburg, P.H., Soepboer, W., Veldkamp, A., Limpiada, R. and Espaldon, V. 2002. Modeling the spatial dynamics of regional land use: the CLUE-S model. *Environmental Management*, 30(3): 391-405.
- Verburg, P.H. and Veldkamp, A. 2004. Projecting land use transitions at forest fringes in the Philippines at two spatial scales. *Landscape Ecology* 19(1): 77-98 .
- Costanza R., 1989. Model goodness of fit: a multiple resolution procedure. *Ecological Modelling* 47: 199-215.

