EU-ClueScanner tutorial;
using the 1km application for DG Environment

Commissioned by:
European Commission DG Environment
Project reference: ENV.G.1/SER/2008/0050

Authors   Eric Koomen, Maarten Hilferink, Martin van der Beek, Marta Perez Soba, Peter Verburg
Date   March 25, 2010
Version   1.0 (final)
Status   External document
Reference   gnp08019

Table of contents

1 Introduction

The EU-ClueScanner model is part of a multi-model framework that aims to produce policy-relevant information related to land-use change. It can support the creation of ex-ante assessments of policies that impact spatial developments. The framework is thus relevant to many different DGs of the Commission to help answer questions such as: what spatial developments are likely to occur in the coming decades? what possible impacts would a foreseen policy have on land-use patterns in Europe and how will these affect issues such as biodiversity for instance? how will specific environmental protection policies and/or climate change adaptation measures influence spatial patterns? The final report documenting the development of the Eu-ClueScanner model discusses the way the model was applied for DG Environment to answer this type of questions (Perez-Soba et al., 2010).

The purpose of this tutorial is to help those who will operate the EU-ClueScanner get started with the model. In addition it will explain the basic functionalities of the model and specify how specific components of the model can be adjusted to define new policy alternatives. The tutorial also points at relevant documents for background information and more advanced manipulation of the model. The text only refers to the EU-ClueScanner application developed for DG Environment. This application is created at a 1 kilometre resolution and thus referred to as *EuClueScanner1km* within the model. The model also contains reference to a more detailed 100 metre version of the model (*EuClueScanner100m*) that is currently being developed for the Joint Research Centre. Both model versions are similar in applied methodology and base data, but differ in resolution, calibration and included simulation runs. The 1km version includes a set of policy related scenarios of the future, whereas the 100m version only contains a reference and validation run.

This first chapter provides an initial introduction to the model in general and the application for DG Environment in particular. It helps users getting started with the model and describes the most important model components. Chapter 2 describes how existing policy alternatives can be viewed and explains the main elements of the graphical user interface (GUI). It discusses how various types of data can be viewed, how the calculation process can be traced and how land-use results and indicator values can be retrieved. Chapter 3 helps the user to define new policy alternatives. First through describing relative simple operations, such as the replacing of spatial data sets and the editing of the current demand specification, and then by highlighting more complex modelling tasks such as the manipulation of basic simulation settings. The fourth and final chapter discusses briefly how possible improvements can be made to the EU-ClueScanner model. This involves issues such as revise the calibration or linking the model to other models. The appendices at the end of this tutorial contain in-depth information on various model elements and include amongst others, an overview of the basic land-use classes and included spatial data sets.

1.1 The EU-ClueScanner in a multi-model framework

The EU-ClueScanner is a land allocation model positioned at the heart of a multi-scale, multi-model, framework. It bridges sector models and indicator models and connects Global and European scale analysis to the local level of environmental impacts, see Figure 1.

External, global models account for interactions between Europe and other world regions. The dynamics of the global economy are typically described by GTAP or derived LEITAP model, whereas global climate-land use interactions are incorporated in an integrated assessment model (IMAGE). This configuration was used in the EURURALIS project (Eickhout et al., 2007; Van Meijl et al., 2006) and the applications created for the DG Environment.

Results from the global level models relate to the demand for goods and commodities and are, in Europe, delivered at Member State level. The output of the global-level models is translated into a land demand in km^2 for the specific land-use types distinguished in the EU-ClueScanner model. This translation is performed in a newly developed demand module that is implemented in the GeoDMS model script and thus part of the EU-ClueScanner model.

Land allocation in the EU-ClueScanner model is based on the approach taken in the well-known Dyna-CLUE model (Verburg et al., 2006; Verburg and Overmars, 2009). Especially for DG Environment this approach has now been programmed in the GeoDMS environment using the numerical algorithm of the Land Use Scanner model to optimize its performance for use on desktop computers. The Land Use Scanner is another well-established land-use model with many applications within Europe (Dekkers and Koomen, 2007; Koomen et al., 2008). Combining the strengths of both models ensures a consistent, state-of-the-art and flexible modelling core.

In addition, a series of indicator models corresponding to the demands of the policy alternatives are implemented. Indicator models use information derived from economic models (e.g. LEITAP) and the land allocation models to arrive at a balanced set of indicators focussing on the land-use and environmental domains. To assess specific issues (such as, for example, flood risk) additional detailed spatial data sets are used in the implemented indicators. Another interesting option for many ex-ante assessments is the possibility to link the pan-European analysis at 1 km^2 resolution to more detailed models for specific case studies that either focus on specific regions or on specific policy domains.
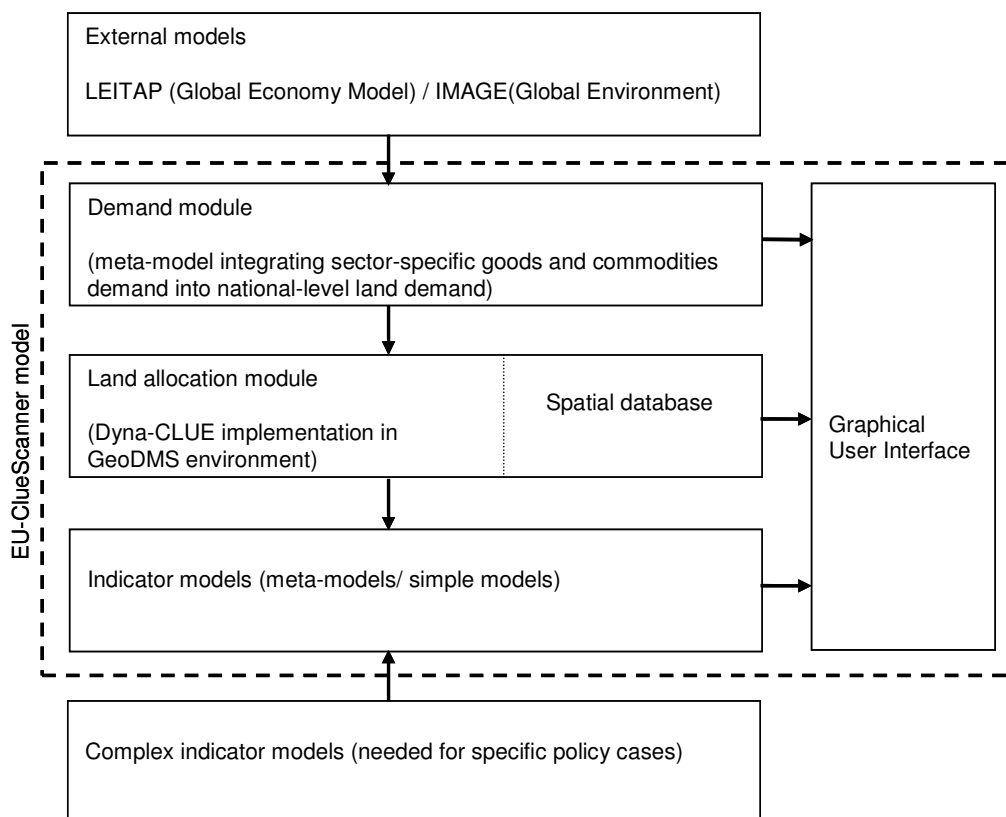

Figure 1 The EU-ClueScanner as the the core of a multi model framework.

1.2   DG environment application

The EU-ClueScanner model is a very flexible environment that allows the construction of many different applications. These applications can differ on the initial land-use that is used as starting point for simulation, the land-use types that are simulated, the demand that is specified for each type of land use, the location characteristics that are included and other model-specific setting relating to, for example, the possible conversions of specific land-use types. This section lists the most important characteristics of the EU-ClueScanner application developed for DG Environment. It specifies the included land-use typology and included policy alternatives.

Land-use typology
The Corine Land Cover database for the year 2000 (CLC2000) was selected as base data for the DG Environment application as it is the most recent, consistently developed and available pan-European source of land-use related information. The CLC2000 data was aggregated to a 1km^2 spatial resolution representing 18 main land-use types for inclusion in the land-use model. Table 1 describes the thematic integration of the initial land-cover classes in CLC2000 to this set of 18 main classes. The table, furthermore, lists for which land-use types future patterns are simulated. For the remaining land-use types the location is kept stable.

Although the CLC data technically refer to land cover we use the more common term 'land use' in this report as it better corresponds to phenomena such as abandoned farmland that are simulated by the model. The latter type of land use, for example, allows for the simulation of forest regeneration and is not based on the CLC data. It arises from simulation when agricultural demand is insufficient to maintain agricultural use. Based on the conversion settings specific location may develop into semi-natural vegetation (3) or forest (10).

Policy alternatives
European policies can be relevant for land-use change in two ways. First, there is a group of policies that influences the *demand for land*, e.g. stimulation of different types of agriculture through the Common Agricultural Policy. A second group of policies influence *land-use configurations*, for example by excluding or favouring some regions for a specific type of land use. For the development of these policy alternatives one future scenario is selected as reference: the Global Co-operation (B1) scenario. This well-known scenario originating from the IPCC-SRES framework sketches the climatic and socio-economic boundary conditions within which the policy alternatives are developed.

Within the DG Environment application seven policy alternatives are included next to a baseline (see the summary in Table 2). The first set of policy alternatives deals with different implementation options of the proposed Renewable Energy Directive (Directive 2009/28/EC) and considers potential changes in the demand of land (through biofuel production) that can be associated with this policy. In addition, several spatial policy alternatives are defined that influence land-use configuration, each focusing on a separate important policy theme relevant for the environment:
-    Biodiversity alternative: strengthening the green environment (i.e. nature and landscape);
-    Soil and climate change alternative: protecting soil and adapting to climate change.

The chosen policy packages fit within the proposed modelling framework and are able to illustrate key policy issues and trade-offs for the EU. The policy options selected are not necessarily likely to be implemented in reality. Their inclusion in the land-use simulations merely aims to show the potential of the modelling framework to assess the impact of such explicit policies. Thus answering *what-if?* type of questions. An extensive description of the reference scenario and policy alternatives can be found in the final report of this project (Perez-Soba et al., 2010).

Table 1 Thematic resolution of the model and their relation with the Corine Land Cover types.

Nr.[1]	CLC (main) class[2]	Description	Simulated[3]
0	1.1/1.2/1.3/1.4	Built-up area	Yes (0)
1	2.1.1/2.4.2p(50%)/ 2.4.3p(25%)	Arable land (non-irrigated)	Yes (1)
2	2.3/ 2.4.2p(50%)/ 2.4.3p (45%)	Pasture	Yes (2)
3	3.2.1/3.2.3/ 3.2.4/2.4.3p (30%)	(semi-) Natural vegetation (including natural grasslands, scrublands, regenerating forest below 2 m, and small forest patches within agricultural landscapes)	Yes (3)
4	4.1	Inland wetlands	No (9)
5	3.3.5	Glaciers and snow	No (9)
6	2.1.2/2.1.3	Irrigated arable land	No (4)
7	New	Recently abandoned arable land (i.e. "long fallow"; includes very extensive farmland not reported in agricultural statistics, herbaceous vegetation, grasses and shrubs below 30 cm)	Yes (5)
8	2.2/2.4.1/2.4.4	Permanent crops	Yes (6)
10	3.1	Forest	Yes (7)
11	3.3.2/ 3.3.3/3.3.4	Sparsely vegetated areas	No (9)
12	3.3.1	Beaches, dunes and sands	No (9)
13	4.2.2	Salines	No (9)
14	4.1/4.2.1/4.2.3/ 5.1/5.2	Water and coastal flats	No (9)
15	3.2.2	Heather and moorlands	No (9)
16	New	Recently abandoned pasture land (includes very extensive pasture land not reported in agricultural statistics, grasses and shrubs below 30cm)	Yes (8)

Notes:

[1]This column shows the individual land-use types distinguished in the DG Environment application. Please note that the model can contain up to 18 classes depending on its application, but in this case 16 land-use types are distinguished (the numbers 9 and 17 are not present in this model version). This set of land-use categories is referred to as the LU18 classification in the model.

[2]For reference purposes the corresponding (main) land-cover classes from the CLC2000 dataset are shown. See Appendix 1 for a complete overview of these land cover classes.

The 'p' after certain CLC-classes denotes that only part of it is assigned to the DG Environment class. This partial assignment has been done randomly according to the mentioned probabilities. So each cell with CLC type 2.4.2 (representing complex cultivation patterns) has a probability of 50% to be classified as arable land and a probability of 50% to be classified as pasture. A random selection was chosen as that corresponds well with actual land-use patterns that can be observed from aerial photographs.

[3]The numbers between brackets refer to the coding generally used within the model (referred to as Clue10).

Table 2 Overview of the included land-use simulation runs following the reference scenario and supplemented policy alternatives and their short names used in the model.

Nr.	Characteristic	Short name
1	Reference scenario: Global Co-operation (B1)	ReferenceB1
2	Biofuel mandate in OECD countries without EU (unrestricted land conversion of forests into agricultural land)	BFP5nonEU
3	same as 2) with EU	BFP5nonEUandEU
4	same as 3) with full protection of all existing forests	BFP5nonEUandEU_noF
5	Biodiversity alternative	BiodiversityB1
6	Biodiversity alternative with alternative 3 as reference	BiodiversityB1_high
7	Soil and climate change alternative	SoilCCB1
8	Soil and climate change alternative with alternative 3 as reference	SoilCCB1_high

1.3 Getting started

Before installing first please check the system requirements. The following hard- and software components are needed to successfully install and run the model. You need to have local administrator rights to be able to do the installation yourself. If you do not have these rights you should ask your system administrators to assist with the installation procedure.

Hardware requirements
1. Processor: Intel Pentium or Pentium compatible;
2. Internal RAM: 1Gb (4 GB recommended);
3. Internal hard disk with at least 20 GB free for the GeoDMS Program Files (ca. 15 MB) and data files;
4. Screen: High resolution with supporting video card (dual-screen recommended).

Software requirements
1. Win32 Operating System: XP or Vista;
2. To edit meta info and claim data: GUI for working with Microsoft Access database files (*.mdb, version 2003 or later);
3. To edit configuration (*.dms) files: an ASCII text editor (the Crimson Editor 3.70 is recommended).

Downloading and installing the four components
To install the software and project data, you need to download and install 4 folders:
1. **GeoDms573** (by default C:\Program Files\ObjectVision\GeoDms573) contains the compiled GeoDMS software. The software can be installed by running the file: http://www.objectvision.nl/OutGoing/EuClueScanner/GeoDms573-SetupW32.exe This can be done directly from this web location, or after downloading the file to a temporary location on your PC. Info on this installation can be found at http://www.objectvision.nl/geodms/ under the menu item Software -> Installation Instructions.
2. **SourceData** (by default assumed to be located in C:\SourceData\EuClueScanner) contains external data that is read-only in the context of the Eu-ClueScanner project. Download http://www.objectvision.nl/OutGoing/EuClueScanner/EUCS-SD-RC.rar and extract its contents in C:\SourceData or any other read-only location. Enable file compression to save disk space before extracting. See the Getting started section below for more information on extracting files and enabling file compression. Without file compression, the SourceData takes 10 GB. File compression reduces this size to 1.4 GB on disk.
3. **ProjectData** (choose any location, for example C:\Users\XXX\projects\EuClueScanner, further referred at by '%projdir%') contains data and model scripts that are created and changed within the context of the project. Download http://www.objectvision.nl/OutGoing/EuClueScanner/EUCS-prj-RC.rar and extract its contents in your project base folder (in this example: in C:\Users\XXX\projects).
4. **LocalData** (by default assumed to be located in C:\LocalData\EuClueScanner) contains intermediate and final results of the calculations, including the 'CalcCache' storage location that the GeoDMS creates. Pre-calculated results for all configured scenarios are provided here to enable you to request Indicator values without having to (re)calculate a scenario. Also the pre-calculated results of the demand model runs are provided. Download http://www.objectvision.nl/OutGoing/EuClueScanner/EUCS-LD-RC.rar and extract its contents in C:\LocalData or any other read-only location.

Folder sizes and their advised management

The system components have been classified into these four groups on the basis of advised management policies in order to minimize the tasks of version control, exchange and synchronization.

Table 3 Main folder characteristics

Folder	Approximate size on disc (compressed)	Required rights	Advised Management
ProgramFiles	14 MB	Read/Execute	Secure installation procedure
SourceData	1.4 GB	ReadOnly	Secure installation procedure. Compression on.
ProjectData	5.5 MB	ReadWriteCreate	Incremental Backup / version control
LocalData including LocalData\CalcCache	350 MB and CalcCache size (several GB)	ReadWriteCreate	Compression on. Remove old files during disk cleanup; manage like a Downloads Cache folder

Additional tips and tricks

1. In order to save disk space it is **strongly recommended** to enable file compression in the LocalData and SourceData folders before extracting the .rar files. This can be done in the Windows Explorer by right-clicking on the created (and still empty) folder, selecting 'Properties', clicking the 'Advanced' options on the 'General' tab and then ticking the 'Compress contents to save disk space' box. For additional instructions on how to enable FileCompression, see 'Keep files Compressed' at http://www.objectvision.nl/geodms/docs/CalcCache.htm#h6

2. You can extract .rar files with WinRar (WinRar 3.80 was used for compressing the data).

3. You can start the GeoDMS software from the Start button or execute %ProgramFiles%/ObjectVision/GeoDms573/GeoDmsGui.exe. The first time you start this, the program doesn't know which configuration to load and will present a File Open Dialog (also accessible from the MainMenu->File->Open Configuration File). You'll have to select %projdir%/cfg/default.dms to open the default model configuration. When you close a modelling session you normally should **not save** any changes into the current default.dms unless you are absolutely sure you have added valuable new elements to the model. Otherwise you create barely changed copies of the original model and loose disk space.

4. If you choose to extract the SourceData or the LocalData in a different folder than C:\SourceData resp. C:\LocalData, you have to tell the GeoDMS where to look in the general settings. These and other more complex settings are managed from the 'Tools' option in the main EU-ClueScanner menu by selecting 'Options' and then the 'General Settings' tab (see below). Here you can provide the chosen location in the Paths section. Close and restart the GeoDmsGui.exe to make the changed values effective.
These settings are stored in a config ini file, which is part of the project configuration.

5. General information on how to use the GeoDMS user interface can be found at http://www.objectvision.nl/geodms/ under the menu item User Guide -> GeoDMS GUI. For details on the Declarative Model Script (the .dms files) look at the menu item Modelling.
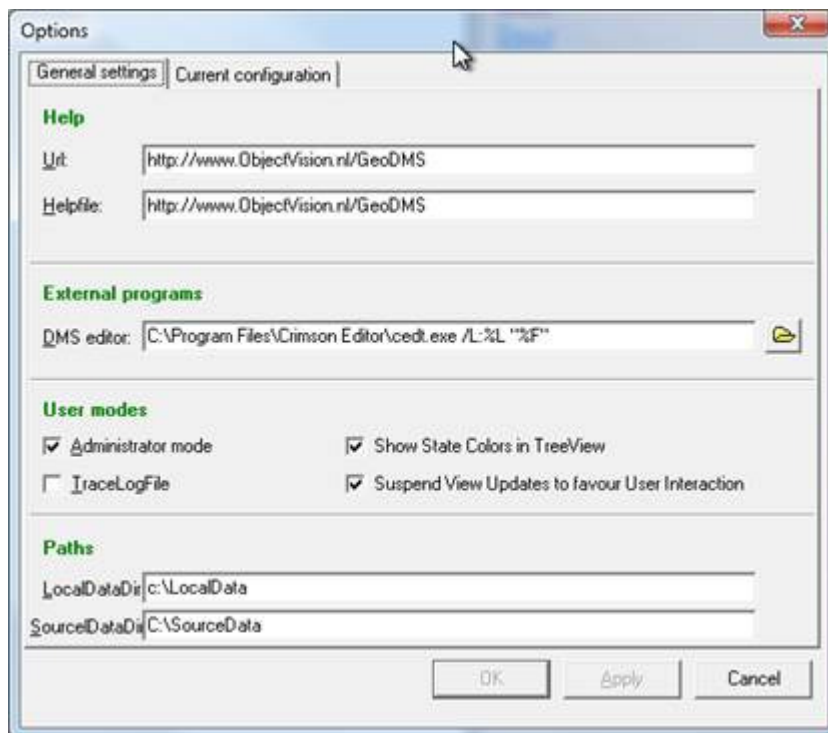
Figure 2 General settings tab in the Main menu.

Advanced options
The GeoDms executes a configurable external editor when you double-click on messages that refer to model scripts or press Ctrl-E. By default, the Crimson Editor is configured. The appropriate line number and file name are provided to the editor by means of the parameters added to the reference to the editor in general setting tab in the Main menu (/L:%L "%F"). The Crimson Editor is downloadable from http://www.crimsoneditor.com/; please check whether the pathname included in the 'General Settings' tab corresponds to the location where the editor is installed). You can download and install GeoDMS syntax highlight files for this editor from:
https://geodms.svn.sourceforge.net/svnroot/geodms/dev/res/Crimson Editor with subversion, or from http://geodms.svn.sourceforge.net/viewvc/geodms/dev/res/Crimson%20Editor/ You can also configure your own favourite editor as an external program under the 'General Settings' tab.

1.4 Main model components

The EU-ClueScanner model has the following main components:
- a *demand module* that specifies the amount of land that has to be allocated per land-use type, normally following scenario conditions;
- *location specific characteristics* (locspecs) that influence the suitability of a location for a land-use type, typically reflecting scenario-specific restrictions or stimulating policies that are added to the statistically derived suitability definition;
- *conversion settings* that specify which land-use conversions are allowed and, optionally, where these conversions may take place;
- weights for *suitability factors* (SF) and *Neighbourhood enrichment* (NE) that together comprise the statistically derived suitability map for each land-use type;
- a *dynamic transition potential* that specifies the conversion elasticity (ease of conversion) for each land-use type; and

- an *output generator* specifying how the above-mentioned components are combined into a simulation run.

The relation between these model components is depicted graphically in Figure 3, whereas Appendix 4 describes their characteristics in more detail. The main components of the model are also present in the Dyna-CLUE model and have been introduced in previous publications (e.g. Verburg et al., 2006; Verburg and Overmars, 2009). Based on the settings specified for these components and the current land use, the model simulates land-use changes for each year in the simulation period (in our case 30 years). The scenario-specific settings for these components are documented in the appendices included with the final report of the project that commissioned the implementation of this land-use model (Perez-Soba et al., 2010). Based on the projected land-use patterns a number of indicators is calculated related to land use, specific themes (e.g. erosion, biodiversity) and socioeconomic conditions. The next chapters will describe how these model components can be viewed and edited to inspect existing scenarios and create new ones.
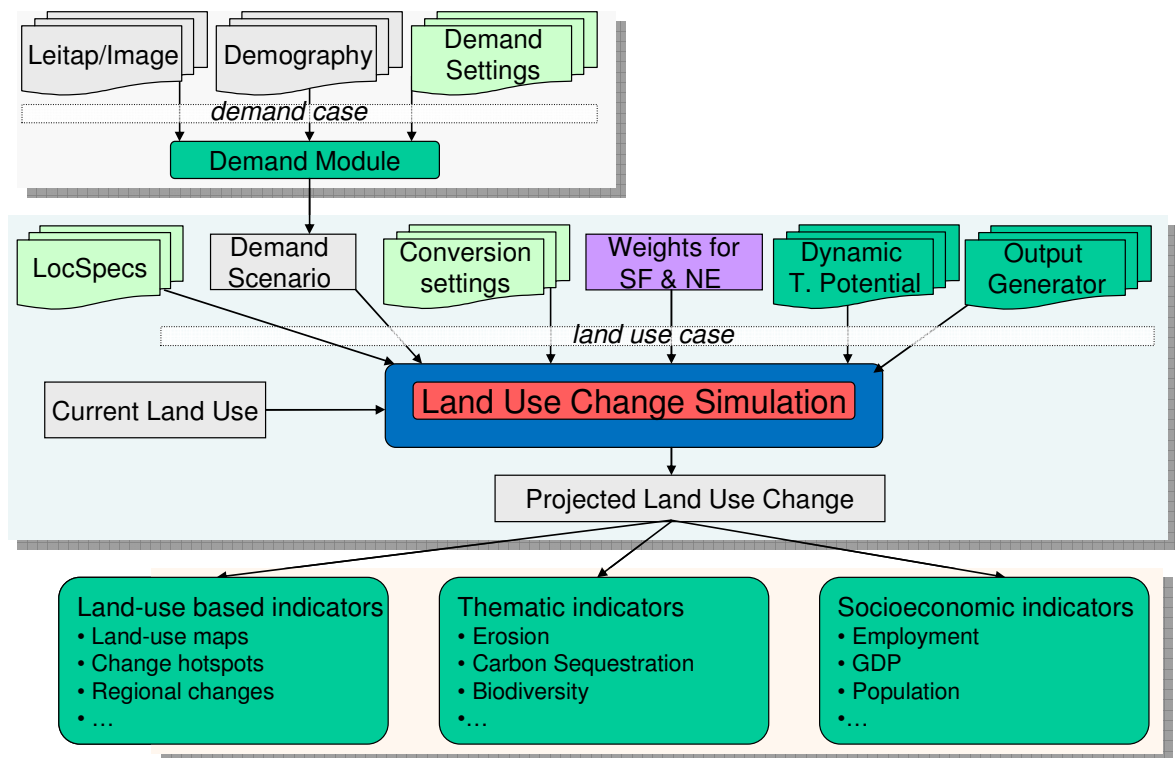

Figure 3 Main model components.

The main model components are an integral part of the EU-ClueScanner model. Their policy alternative specific settings can be edited in various in Microsoft Access databases or small parameter files as is described in Chapter 3. For convenience sake the location of these editable settings files is provided in the table below.

Table 4 Location of main model components

Component	Editable settings files and their storage location
Demand module	demand.mdb in projects\EuClueScanner\data\demand specifies how the demand should be calculated per land per year; the results of these calculations are stored in demand.in1 files in the \LocalData\EuClueScanner\Demand*PolicyAlternative-Name* (e.g. ReferenceB1)*CountryName* folder
Locspecs	locspecX.fil (when present) in \SourceData\EuClueScanner*PolicyAlternative-Name\CountryName*
Conversion settings	allow.txt in \SourceData\EuClueScanner*PolicyAlternativeName* *CountryName*, this file may refer to spatially explicit allow driver maps (e.g. X1.tif) stored in \SourceData\EuClueScanner\Drivers
Weights for SF and NE	alloc1.reg and alloc2.reg in \SourceData\EuClueScanner*PolicyAlternativeName*\ *CountryName*
Dynamic transition potential	main.1 in \SourceData\EuClueScanner\ *PolicyAlternativeName* *CountryName*, this file may refer to spatially explicit dynamic driver maps (e.g. sc1gr33.1) stored in \SourceData\EuClueScanner\DynamicDrivers
Output generator	part of a crucial land-use allocation GeoDMS script (\Projects\EuClueScanner\cfg\ Default\ClassicClueWrap.dms) that determines which and how intermediate results of the simulation should be processed.

2 Viewing existing policy alternatives

This chapter helps the modeller to get familiar with the main components of the EU-ClueScanner model. First the main elements of the Graphical User Interface (GUI) are explained. Subsequently the main components of the model are described. These relate to the available spatial data (land-use and factor data) and policy alternatives (claim data, runs, results, tracing calculation process and indicators). Note that this chapter is partially based on an adaptation of the first two chapters of the GeoDMS GUI user guide (van der Beek, 2009), which can be found at http://www.objectvision.nl/GeoDMS/ under User Guide -> GeoDMS GUI. More in-depth information on the concepts and the syntax of the GeoDMS configuration language is available at the same website under Modelling- > Modeller's guide.

2.1 Tree view and main GUI elements

The graphical user interface (GUI) provides the modeller with a range of windows to view data layers, look-up background information, inspect simulation results and follow the simulation process. Figure 4 presents an overview of the many different windows in a typical application. Which windows are shown in a session depends on the options that are selected (ticked) in the main menu under View.
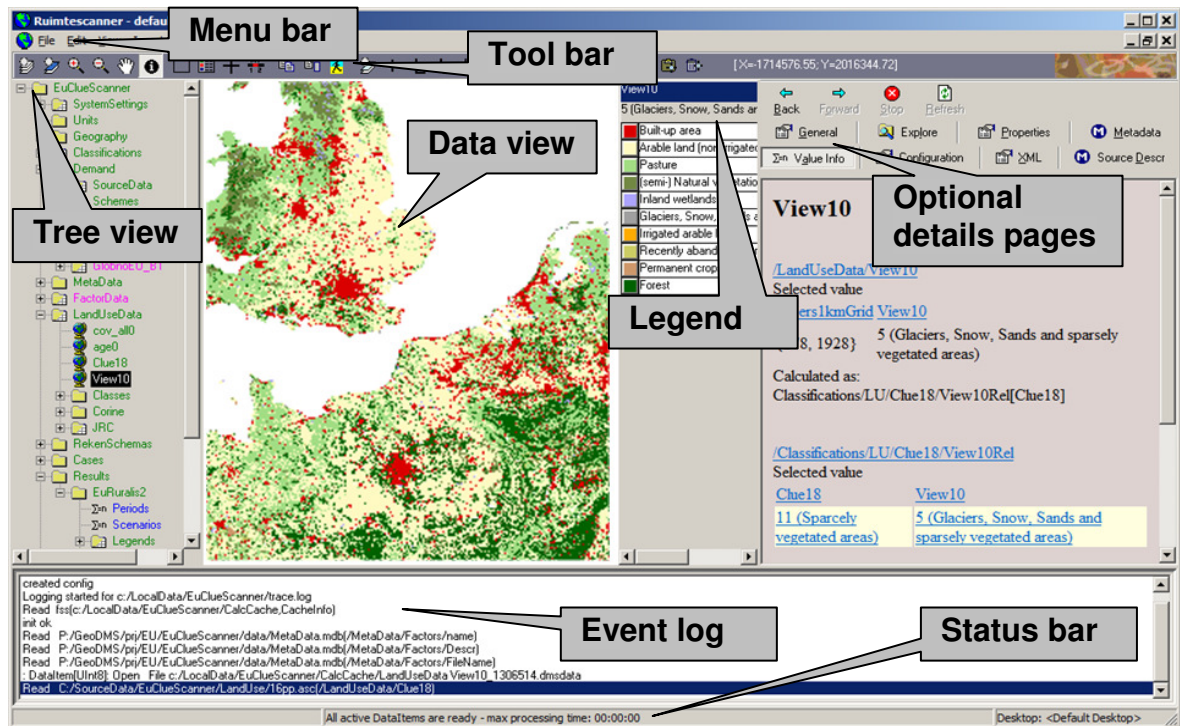

Figure 4 Tree view and main components of the GeoDMS interface.

The GUI contains the following elements:
- the Menu bar with several pull down menus;
- an initially empty dark-grey Tool bar that contains window-specific tools;
- on the left hand side a Tree view that allows navigation through the available data sets and model result;

- a Data view that displays data as tables or maps;
- a map Legend that appears with map in the Data views;
- various Details pages containing technical and background information that appear or disappear by simultaneously clicking the ALT and 1 keys
- an Event log that presents information on the processed steps. This event log appears/disappears by simultaneously clicking the ALT and 2 keys
- a Status bar that presents hints and status information about ongoing processes.

The Tree view is the main navigation component within the application. It is comparable to the Windows Explorer in Windows and allows easy access to the huge collection of data and model result items that are available in the model. The Tree view presents the hierarchical structure of the EU-ClueScanner configuration. By default the root and the first level items are shown (the root item is expanded). Each item in the tree (called a tree item) is presented with a name and an icon. The icons are related to the default viewer for the corresponding tree item. In section 2.2.1 these icons are explained.

The selected item in the Tree view is the active item in the application. This is an important concept, as many functions of the application work on the active tree item. By clicking the right mouse button a pop-up menu can be activated, with a set of menu options that work on this active tree item. This also applies to most main menu options and to the Detail pages. The pop-up menu options, most main menu options and the details pages are described in Chapter 4 and 5 of the GeoDMS user guide.

Icons used in the Tree view
The main purpose of the icons is to show which default viewer is used for an item. Double clicking or pressing the enter key on a selected tree item activates this viewer. The following icons are used to indicate this default viewer:

 Data item that can be viewed in a map. This implies the domain of the data item has a geographic relation (see GeoDMS modeller's guide for how to configure a geographic domain unit). Dependent on the geographic domain, the data is visualized in a grid, point, arc or polygon layer.

 Data item that cannot be visualized in map (has no geographic relation) but can be viewed in a Datagrid.

 Data item that contains a classification (a set of classes). A classification can be made or edited with the classification and palette composer, see GeoDMS user guide Chapter 10

Not all tree items are data items. For the non-data items (these items cannot be viewed with a primary data viewer), the following icons are in use:

 A tree item with no data item and subitems that also have no data items. The icon is used for containers grouping other containers or units.

 A tree item with no data item, but with data items as subitems. All subitems of this item with the same domain unit can be viewed in a Datagrid.

 A tree item with no data item and no subitems, e.g. a unit

The next sections explore the five main components of the tree view: land-use data, other spatial (factor) data sets, the demand specification, models runs and indicators.

2.2 Land-use data

The EU-ClueScanner contains two representations of the initial year (2000) land use. The land-use data for the initial year are based on a spatial and thematic aggregation of the Corine Land Cover 2000 dataset (see Section 0). The land-use data can be viewed by browsing through the Tree view. Start by clicking on the + in front of the container (directory) called *LandUseData*. This container includes data-layers that describe the land use in the year 2000. The small globes indicate that the data item can be visualised in a map view.

Maps can be drawn in the data view area in three different ways:
1. by double-clicking on any of the data layer names. This option allows multiple data layers to be drawn in the same map view window.
2. by activating (clicking on) the data layer and then simultaneously pressing the Ctrl-M key combination. This option will open a new map view window for every data layer.
3. by giving a right mouse-click on the data layer and selecting the Map View or the Default option from the appearing menu. The other options allow showing the data as a table (which does not make sense here) or histogram. The latter option shows the frequency distribution of the different land-use types.

The 2000 land use is shown in two different thematic aggregations: one using 18 classes and one using 10 classes to provide an easier to interpret map. The included classes are described in the legend that appears on the right hand side of the map. The last column in this legend contains the count for each class. As the application has a 1 km^2 resolution the count values can be read as surface areas in km^2. More background information on the data is available in a meta-data sheet that can be retrieved as one of the Details pages.

Right mouse-clicking on the legend area allows the user to select, amongst others, the statistics function. This option provides some basic statistics on the selected dataset, e.g. minimum and maximum value per grid cell and the total area (sum) covered by this land-use type. The edit palette pop-up menu option allows you to change the classification and colour as is explained in more detail in the GeoDMS user guide.

The tool-bar now contains several standard GIS-functions such as:

zoom in, zoom out, pan, get info, zoom to full extent,

copy visible area to clipboard, toggle scale bar.

The objective of these and other functions is indicated with a mouse-over text. For a full description of all tools, including options to export specific views and datasets as bitmaps or other data formats, see the GeoDMS user guide.

Please note that this basic data set is stored as a single ascii grid in the LandUse folder in the SourceData directory. This data set can be copied from here to be used in, for example, a GIS environment. Alternatively all data items that can be viewed in a map, such as the land-use data sets, can be exported by right-clicking on the attribute in the tree view and selecting the option export primary data. The user can then choose to export data in various formats, such as ascii grid or bitmap (bmp) with world file (reference to a coordinate system). It is necessary to first draw the map in the map window before this option is selected. These export options will create a copy of the whole data set, using its full extent. When the user only wants to export the part of the map that is visible in the map window, the tool bar option 'copy visible area to clipboard' should be selected.

2.3   *Factor data*

The Eu-ClueScanner also contains a wide range of spatial data sets that describe specific themes such as accessibility, geomorphology, climate and land use in neighbouring cells. These datasets were collected for the EURURALIS project (Eickhout et al., 2007; Van Meijl et al., 2006) and are used as independent factors in the statistical calibration of the model.

The data can be requested in the EU-ClueScanner from the FactorData container and can be shown as indicated above. They are shown with a corresponding legend file and meta data (a standardised description) in the details pages. Appendix 2 provides an overview of the included datasets. The factor data are stored as ascii grids in the 'Drivers' folder in the SourceData directory.

2.4   *Demand*

The container *Demand* specifies the total amount of land (in km^2) that has to be allocated per year, per land-use type, per region. These land-use claims or demands are calculated in the demand module based on the information from external sources such as the LEITAP model and can be viewed in the Demand container. Demands are calculated per demand scenario (in our case the reference scenario and three biofuel policy alternatives) and specified for each of the 24 claim regions. These regions normally consist of individual countries, only the Baltic countries and Belgium and Luxemburg are aggregated. Please note that for each region the exact demand needs to be specified per land-use type per year as the model is not able to allocate land when more (or less) land is claimed than is available in a region.

The regional demand can be viewed in the tree view by selecting Demand -> runs -> *ScenarioName* (e.g. ReferenceB1) -> per_country -> *CountryName* (e.g. Austria) -> Output. Doubleclicking (or Ctrl+D) provides a tabular overview of the land-use claims. Please note that the StableAreas class in this table refers to the total amount of land covered by the land-use types that are allocated by the model (see Table 1).

The demand data are calculated by the demand module and stored as text files (with a .in1 extension) in the Demand folder in the LocalData directory. This folder contains a subfolder per demand scenario (e.g. ReferenceB1) that in turn contains subfolders per demand region. Appendix 4 explains the layout of these files. The functioning of the demand module is explained in Section 3.2.

2.5   *Runs*

The term *runs* refers to individual simulation runs. These represent a unique combination of claims (demand) and local suitability settings and in our case correspond to the eight policy alternatives that are included in the model. The definition of the runs can be inspected in runs -> EuClueScanner1km -> runs -> *PolicyAlternativeName* (e.g. ReferenceB1) -> *CountryName* (e.g. Austria). This is a reflection of all possible model settings and may be a bit confusing at first. The most important components of the case definition are described in Chapter 3. It is clear from these folders that the model is specified at the level of individual countries. In fact, the main parameters of a simulation run (case) are stored as main.1 text files in country-specific folders in policy-alternative folders (e.g. ReferenceB1) folder in the SourceData directory. Appendix 4 describes the layout of these files.

To actually produce and view intermediate and final simulation results several options exist in the

country-specific policy alternative folder:

-    DynaClue -> TimeSteps -> *Year* (e.g. P2010) -> ResultingState -> LandUse starts the
     simulation of land use in a specific year. Obviously this simulation process takes longer
     when a longer period is selected. This TimeSteps folder contains also many other
     important dynamic model settings per individual simulation year.
-    simulation_results -> LandUse shows land use in the final year in various representations
     that are explained in the legend file and meta data in the details pages.

The Event log shows the progress of calculations during simulation. It lists, amongst others, the
amount of data items that still have to be calculated and thus gives an indication of the additional
time needed to complete a calculation.

The Runs -> EuClueScanner1km -> ResultsGen -> *ScenarioName* (e.g. ReferenceB1) folders
allow the generation of complete sets of results for all countries in Europe for each year in the
simulation period. Activating the Generate data item starts this calculation process. As the
generation of results for the whole of Europe takes several hours, this option should be used with
care. The results are stored as separate .tiff files for each simulation year in the folder:
\LocalData\EuClueScanner\Results*ScenarioName*. It is also possible to generate results for a
specific year, by selecting the appropriate year in the LandUsePerYear folder. Please note that the
administrator mode should be switched on to activate the latter option.

Pre-processed results from model runs that have already been stored can be viewed in the
indicators folder in the tree view, as is discussed in the following section. Results of pre-calculated
simulation runs are read from the \LocalData\EuClueScanner\EuClueScanner1km\Results.use
directory in subfolders with the name of the policy alternative. These results consist of the basic
(Clue10) simulated land use for each individual year and are stored as TIFF files.

Scenario results that have been generated in \LocalData\EuClueScanner\Results*ScenarioName*
can be moved to \LocalData\EuClueScanner\EuClueScanner1km\Results.use\<ScenarioName> to
become available for subsequent inspection or the calculation of indicator values. This can be done
using, for example, the Windows Explorer. The new results will then be shown in the indicators
folder in the tree view when the 'IsCalculated' box is for this policy alternative is ticked in the
metadata.mdb in the \projects\EUClueScanner\data folder (see Section 3.3 for more details on
this). This separation between initially generated results and final results used for indicator
calculation is done to prevent an accidental partial overwrite of pre-generated data that would
corrupt already approved results.

Temporary results that are produced by updating data items during a modelling session are stored
in the LocalData\EuClueScanner\CalcCache directory. This is the storage facility for all
(intermediate) results that are produced during simulation. The results stored here will
automatically be retrieved when an attribute is called upon from the tree view. Please note that this
directory can become extremely voluminous, so it is advisable to delete files not recently used in
the CalcCache when the model becomes too slow or after the model has crashed.

2.6   Indicators

The *indicators* folder contains, for each pre-calculated policy alternative, the land-use results in
various representations, derived thematic indicator maps (related to specific policy themes) and a
set of socio-economic indicators that originate from the LEITAP and demographic models that feed
into the EU-ClueScanner. The latter type of indicators is derived from the LEITAP model and only
available at the National level. This extensive set of indicators helps the user interpret simulation
results. A full overview of the available indicators is presented in the final report of this project

(Perez-Soba et al., 2010) and can be obtained by opening the indicators subfolders for a policy alternative.

Indicators can either be viewed as maps and tables. Regional aggregations usually exist at the national level and lower Nuts2 and Nuts3 statistical regions. The indicators can be viewed by selecting them in the tree view. Background information on the origin, input data and applied methodology can be found in the associated meta-data sheets that can be viewed by clicking Alt+1 and selecting the MetaData tab in the Details pages. Figure 5 shows an example of an indicator map retrieved within the user interface.

The simulated land-use patterns for the different policy alternatives are essential input for the calculation of the various indicators. This input is read from the pre-calculated results folder (Results.use) in the the LocalData\EuClueScanner\ EuClueScanner1km directory. It is thus a prerequisite for indicator calculation to first perform the land-use simulation for any new policy alternatives and copy the results to the Results.use folder.

A comparison of land-use maps can be created with the ExploreTools/CompareLanduse template. It allows the user to compare two individual land-use maps that can be selected from pull-down lists. The ExploreTools can be activated by clicking Ctrl+A when the administrator mode is switched on. A so-called case generator window appears that first asks the user to specify a name for this map comparison and then (after clicking next) to select two land-use maps from two pull-down menus that will be compared on a pixel-by-pixel basis. The resulting maps are available in a new folder in the tree view under runs -> CompareLandUse that contains four maps:
- map1 is the land-use map that was selected first
- map2 is the land-use map that was selected subsequently
- diff1 is a map showing the initial (map1) land use for those locations where land use changed between the two years.
- diff2 is a map showing the final (map2) land use for those locations where land use changed between the two years.

Please note that the created difference maps are created temporarily within a modelling session. They are not stored when the model is closed. Should you wish to store specific maps you created you can do so by exporting them through the various options in the Tool bar. If you have specific map comparison you want to evaluate in each session it may be wise to define that comparison option as an indicator in the appropriate GeoDMS script file (see Section 4.3).
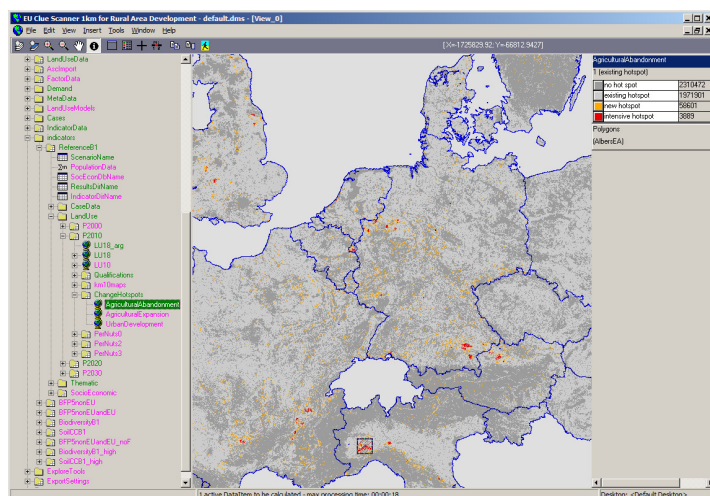

Figure 5 Example of a land-use based indicator (hot-spot of agricultural abandonment).

3   Defining new policy alternatives

Creating a new model application to simulate land use according to, for example, a new reference scenario or policy alternative normally requires the following steps:
1. specification of the application in terms of land-use related themes, thus answering the question how the proposed policy alternative may influence land-use patterns;
2. specification of the application in terms of model inputs, thus define which model components should be used to simulate the anticipated spatial processes;
3. implementation of the new settings and data in the model and simulation;
4. visualisation and interpretation of the model results.

The development of a new policy alternative is an iterative process that requires intensive cooperation between policy makers and scientists with a good overview of data and model potentials (i.e. land-use modelling experts). This is especially important in the specification of the application (step 1 and 2), but also for the evaluation of initial simulation outcomes (step 4); policy makers should evaluate whether or not the results are in line with their expectations. They may request alternative model specifications when the simulated land-use patterns conflict their anticipations or other policy objectives.

The implementation of a proposed new policy alternative should start with obtaining a clear understanding of its land-use related impacts. Initial discussions between policy makers and modellers/scientists should thus focus on clarifying the intentions and likely consequences of proposed policy alternatives in terms of land-use change. When these cause-effect relations are clear, specification in terms of model inputs can be proposed by the modeller. A basic question in this respect is: will the policy influence the demand for land (e.g. through intervention in the macro-economic processes as is the case with a reform of the Common Agricultural Policy) or will it mainly affect the locations of future developments. The former case will normally require the application of a regional economic model such as LEITAP, whereas the latter types of spatial policy can be represented in spatial data sets (e.g. maps of areas where specific land-use restrictions will apply). The policy alternatives that were developed for DG Environment also resulted from several discussion sessions that focussed on the objectives of specific policies and the potential of the land-use modelling framework to accommodate their likely consequences. The actual implementation of the final set of policy alternatives has been extensively documented in an appendix to the final report (Perez-Soba et al., 2010). This description is essential reading for those who consider developing a new policy alternative.

New applications request a modification of the modelling set-up in some form or other. Roughly ranging from simple to complex, these may include:
- adding new thematic data (e.g. policy maps, revised accessibility);
- adding new land-use datasets, e.g. CLC 2006 when it becomes available;
- developing and adding new indicators;
- extending the study area with, for example, Switzerland, Balkan or Turkey;
- link the land-use model with other models (transport, hydrology, economics);
- changing the resolution (e.g. 100m grid);
- revise the calibration based on the new land-use data or other additional data sets;

The table below organises the main tasks related to the implementation of new model applications under two main headings: 1) defining and running new policy alternatives; and 2) more extensive improvements of the modelling framework required for specific new applications. It indicates the complexity of these tasks and lists the model components that should be edited and the profile of the user that is likely to be involved. Most of the basic tasks needed to modify or implement a new

scenario are feasible by an experienced user within reasonable time (depending on the type of scenario between 1 hour and 10 days). It is important to note that new and more complex applications may involve additional relevant partners (e.g. to run hydrological or global agro-economic models).

Please note that the complexity of the tasks only refers to the work involved in adjusting the EU-ClueScanner model. **In many cases the most substantial part of the work will have to be performed outside the actual model.** Collection and preparation of pan-European datasets is, for example, an extremely time consuming issue. Even the collection of base data for a single new member state and harmonising it with the currently available datasets can easily take several months for a GIS-expert.

Table 5 Indication of complexity (*=simple, ** = advanced, *** = complex) of various tasks related to the application of the EU-ClueScanner framework.

Tasks	Complexity	Model component	User profile
1. Defining and running new policy alternatives			
Replacing spatial data sets (adding pre-processed data relating to e.g. policy, accessibility)	*	file manager	programmer
Editing current demand specification	*	Dbase	programmer
Combining existing scenario components (demands, transition potential definitions, etc.) into new policy alternative	*	Dbase/ DMS-script	policy maker
Adding new demand specification (e.g. adding new external LEITAP or demographic projections)	*	Dbase/ DMS-script	programmer
Adding new spatial data sets (adding pre-processed data and referring to it in policy alternatives)	*	DMS-script/ DMS-script	programmer
Edit basic simulation settings (e.g. Conversion Settings Matrices)	*	DMS-script/ parameter files	programmer/ scientist
2. Improving the EU-ClueScanner model			
Replacing land-use base data (e.g. inserting CLC2006) and adjusting demand settings	*	DMS-script	scientist
Revise calibration, renew specification of weights of suitability factors and neighbourhood enrichment, based on pre-processed data for one country	**	e.g. SPSS/ DMS-script	scientist (with expertise in statistics)
Add/edit indicator definitions including graphic representation (strongly dependent on the type of indicator)	**	DMS-script	programmer
Link model to other models e.g. to assess specific impacts (strongly dependent on type of linking)	**		programmer
Edit allocation process: - change Transition Potential definitions - change rules of Neighbourhood enrichment - adjust post-processing rules (succession) - edit Output Generation Definition	***	DMS-script	programmer/ scientist
Change set of land use classes, including revision of demand and other simulation settings	***	DMS-script	programmer/ scientist
Extend study area (e.g. include Switzerland or Balkan), including adding of demand and other model settings	***	DMS-script	programmer/ scientist
Change resolution including respecification of model settings	***	DMS-script	programmer/ scientist

This chapter aims to introduce the relatively simple tasks related to the defining and running of new policy alternatives, whereas the subsequent chapter discusses some of the advanced tasks involved in improving the modelling framework. The complex tasks listed in the table are beyond the scope of this tutorial as they require a thorough knowledge of the modelling environment. The same goes for adaptations to the modelling shell (GeoDMS) such as: adding new spatial analytical

or other functionality, supporting new file formats or updating the model to a new windows version. Such changes will normally have to be programmed in the GeoDMS-source code (in C++) and are thus the domain of software programmers.

To define new policy alternatives a relatively small set of files can be manipulated relating to:
- the spatial data sets describing policy maps or driving forces;
- the regional land demand associated with the policy alternatives; and
- model settings (including the statistically calibrated suitability factors, transition potential and location specific characteristics).

This chapter first explains how existing spatial data sets can be replaced. Subsequently the editing and adding of new elements (demand specifications, data sets) is explained. Finally more complex issues such as the changing of the basic model settings and the creation of completely new policy options are discussed.

3.1   Replacing spatial data sets

The simplest way to include new information in the existing policy alternatives is through replacing existing spatial data sets. This may be necessary to include revised policy maps or updated accessibility maps. To replace an existing dataset, a number of actions have to be taken as is described below.

First the to-be replaced data set needs to be identified and located. Appendix 2 provides an overview of the included factor data sets. These are stored as TIFF files in the SourceData\EuClueScanner\DRIVERS folder on your computer. The exact location of the Sourcedata folder can be found by in the 'General setting' tab (see Section **Error! Reference source not found.**). Alternatively the location (Source description) of a specific file can be retrieved by opening the details pages (by selecting this option from the 'View' menu or by pressing the 'alt' and the '1' buttons), while viewing the dataset. In the viewer on the right you can read the source of the dataset.

It is necessary that the new data set has the same grid specifications (extent, location, and cell size) and is created in the same 1972-WGS Albers projection (see Appendix 3) as the old data set it is replacing. It should also have the same values units (e.g. kilometres or per cent) as the old file to make sure that it has an equivalent impact. When the units are, for example, accidentally specified as meters instead of kilometres the data set will have 1000 times stronger impact. When the replacing data set has exactly the same name as the original file it will be directly used in calculation. The adding of new data sets is discussed in Section 3.5.

3.2   Editing current demand specification

The demand for land is calculated in the model with the demand module. This module translates base data related to demographic development and the demand for agricultural commodities into a demand for land in km^2 per land-use type. The base data comes from external sources and is stored in a Microsoft Access database (demand.mdb) located in projects\EuClueScanner\data\demand. This database consists of collection of different tables related to *demographic data*, a*gricultural commodities*, and additional *demand settings*.

The *demographic data* is derived from scenario-based projections of the PHOENIX model (Hilderink, 2004) and stored in four scenario-specific tables in the database (demography_A1, demography_A2, demography_B1, demography_B2). The demand for *agricultural commodities* is taken from the LEITAP/IMAGE combination (Van Meijl et al., 2006) and stored in several scenario-specific tables (using the original names of the files coming from the LEITAP modellers, e.g.

GTAP_CLUE_090615_DGEnv_Ref). The demand database also contains scenario-specific *demand settings* (e.g. the amount of land in m^2 per inhabitant) that are used to translate the base data into land demand. These settings may differ for different scenarios and in our case include 7 different variants that are included as separate tables in (with the prefix demand_settings) in the demand.mdb. Table 6 briefly explains the demand setting parameters. The demographic and agricultural demand base data and demand settings are stored at the national level.

The current demand settings can be edited fairly easily in Microsoft Access to, for example, test the impact of changes in agricultural policy. This should, however, be done with great caution as the current settings have been made by a wide range of experts in the EURURALIS project. Note also that changes in Microsoft Access cannot be undone. It is thus advisable to save a copy of the original database. Ideally, revised demand settings should be stored in a new table with a meaningful new name within the demand.mdb. A new policy alternative can then be created that makes use of this new demand specification (see Section 3.4).

Table 6 Demand setting parameters

Demand setting	Explanation
isEU15	Parameter indicating whether or not the other demand settings parameters apply to EU15 countries. The demand settings for the new accession states are specified in the second row of the column and may differ on all of the following demand settings.
RatioPermanentRotationalCrop	Ratio of permanent crops to arable land, used to calculate the demand for permanent crops
CeasingSubsSetAside	Starting year of decrease in set aside subsidies (or 0 when decrease is absent)
PeriodSubsSetAsideDiminishing	Period in years over which set-aside subsidies are diminishing
IncreasingSubsSetAside	Starting year of increase (or 0 when decrease is absent)
PeriodIncreaseSpread	Period in years over which increase is spread
MaxPercentSetAside	Maximum percentage of land set aside
PercentIncrease	Percentage of the increase
MaxFallowPastures	Percentage of maximum fallow for pastures
MinFallowPastures	Percentage of minimum fallow for pastures
BiofuelSetAside_2000_2010	Percentage of biofuel allowed on set-aside land in 2000-2010
BiofuelSetAside_2010_2020	Percentage of biofuel allowed on set-aside land in 2010-2020
BiofuelSetAside_2020_2030	Percentage of biofuel allowed on set-aside land in 2020-2030
TrendDensity	Yearly increase in the amount of urban land in m^2 used per inhabitant
corrPasture	Minimum percentage used to define a lower limit to the area including fallow (areaInclFallow)

3.3   Combining existing scenario components

Individual simulation runs consist of a socio-economic scenario and superimposed policy alternatives. These are defined in the table Scenarios that is stored as part of a Microsoft Access database (metadata.mdb) located in \projects\EUClueScanner\data. This table can be viewed and edited with Microsoft Access (see the example in Figure 8). You can add runs (new simulation alternatives) by inserting new rows into the Scenarios table. Changes to the contents of this table become effective after reloading the model configuration from the GeoDmsGUI (File->Open). A case definition consists of several components that are discussed below.

The column *ItemName* contains a name given by the user to characterise the new alternative. This is the name that will appear in the tree view of the model.

The *IsActive* and *IsActive100m* columns denote whether the new case is part of the DG

Environment application (with a 1km grid) or the JRC application (with a 100m resolution) and should be ticked accordingly.

The *IsCalculated* and *IsCalculated100m* columns indicate whether pre-calculated results already exist. These should be stored in the LocalData\EuClueScanner\ EuClueScanner1km\Results.use folder when available. Please note that this folder name should correspond exactly to the ItemName in the Scenarios table to be found. Section 2.5 describes how results for a new case can be obtained and stored.

The column *MainName* determines from which folder the main.1 files are read that are stored in the SourceData directory. Appendix 4 provides a short description of the format of these main simulation settings files (called Main.1)

The column *ParamName* determines from which folder the other scenario specific parameter files are read. These files include:
- alloc1.reg, which relates Suitability Factors to land-use types;
- alloc2.reg, which relates Neighbourhood Enrichment factors to land-use types;
- neighmat.txt, which defines the size and shape of the neighbourhoods and the weight assigned to the neighbourhood function;
- allow.txt, which specifies succession and/or refers to succession maps;
- locspecX.fil, which refer to location-specific characteristics (e.g. spatial policies) per land-use type when applicable (only in specific policy alternatives).

See Appendix 4 for descriptions of these file formats. Please note that these model specification files in the current EU-ClueScanner configuration are unique per country and thus stored in separate folders per country. The files also differ for policy alternatives that aim to influence the spatial configuration of land-use types (in our case the Biodiversity and Soil and Climate alternatives).

The Column *DemandName* in the Scenario table identifies for each case which demand set to use. The DemandName is read from the separate DemandSets table in the metadata.mdb. This link has been defined with the relationships option in Microsoft Access.

Finally, the *LandUseTypes* and *PeriodSet* items in the table specify which set of land-use types and period of simulation should be specified. For the DG Environment 1km application the Clue10 typology (see Table 1) should be selected and the period of simulation is normally 2000 to 2030.

New runs can be created by combining the above components. It is, for example, possible to combine the spatial policy alternatives (Biodiversity and Climate Change) with a different demand scenario than is currently implemented (e.g. by selecting the BFP5nonEU demand scenario).

Id	ItemName	IsActive	IsActive100m	IsCalculated	IsCalculated100	MainName	ParamName	DemandName	LandUseTypes	PeriodSet
2	ReferenceB1	☑	☑	☑	☐	B1	B1	%LocalDataProjDir%/Demand/ReferenceB1	Clue10	P2000_2030
3	BFP5nonEU	☑	☐	☑	◼	B1	B1	%LocalDataProjDir%/Demand/BFP5nonEU	Clue10	P2000_2030
4	BFP5nonEu_SK	☐	☐	☐	☐	B1_G1C2E2_L2	B1	%LocalDataProjDir%/Demand/BFP5nonEU	Clue10	P2000_2030
5	BFP5nonEUandEU	☑	☐	☑	☐	B1	B1	%LocalDataProjDir%/Demand/BFP5nonEUandEU	Clue10	P2000_2030
6	BFP5nonEUandEU_SK	☐	☐	☐	☐	B1_G1C2E2_L2	B1	%LocalDataProjDir%/Demand/BFP5nonEUandEU	Clue10	P2000_2030
11	BiodiversityB1	☑	☐	☑	☐	B1_biod	B1_biod	%LocalDataProjDir%/Demand/BiodiversityB1	Clue10	P2000_2030
12	SoilCCB1	☑	☐	☑	☐	B1_SCC	B1_SCC	%LocalDataProjDir%/Demand/SoilCCB1	Clue10	P2000_2030
13	BFP5nonEUandEU_noF	☑	☐	☑	☐	B1	B1	%LocalDataProjDir%/Demand/BFP5nonEUandEU_no	Clue10	P2000_2030
14	BiodiversityB1_high	☑	☐	☑	☐	B1_biod	B1_biod	%LocalDataProjDir%/Demand/BiodiversityB1_high	Clue10	P2000_2030
15	SoilCCB1_high	☑	☐	☑	☐	B1_SCC	B1_SCC	%LocalDataProjDir%/Demand/SoilCCB1_high	Clue10	P2000_2030
16	Validation2000	☐	☑	☐	☐	Validation2000	Validation2000	%LocalDataProjDir%/Demand/Validation2000	Jrc14	P1990_2000
17	ReferenceJRC	☐	☑	☐	☐	ReferenceJRC	ReferenceJRC	%LocalDataProjDir%/Demand/ReferenceB1	Jrc14	P2000_2030
(ber)		☐	☐	☐	☐			%LocalDataProjDir%/Demand/Validation2000		

Figure 6 Screenshot of the Scenarios table in the metadata.mdb.

3.4   Adding a new demand specification

To be able to understand how a new demand specification can be added it is important to understand the way the relations between the different tables in the metadata.mdb are defined. Figure 7 provides an overview of these relations that can be shown and edited in Microsoft Access with the option tools -> relationships. The figure makes clear that policy alternatives (called Scenarios in the metadata.mdb) rely on demand sets that combine information from both the GTapSets and DemandSettingSets tables. This section describes different ways to make new demand specifications: add new demand sets, add new demand settings sets or add new demand base data (such as the LEITAP demand sets stored in the GTapSets table).
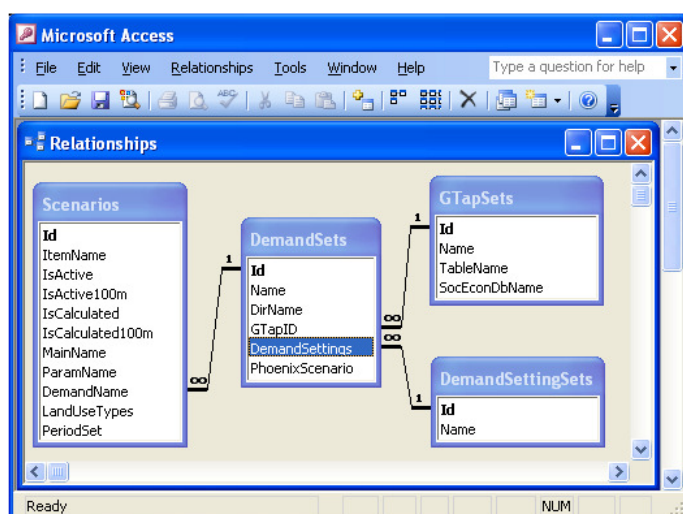
Figure 7 Relationships in the metadata.mdb

Add new demand sets
It is possible to add a new demand sets to the model by adding an extra record (with a new unique ID) to the DemandSets table in the metadata.mdb (see Figure 8). In this record the following components should be defined:
- Name, a short name for the new demand set that will show up in the tree view;
- DirName, the name of the directory where the demand data will be stored when the demand module is run;
- GTapID, the name of the LEITAP model run that contains the base data for the demand for agricultural commodities. This name is read from the separate GTapSets table in the metadata.mdb through a link that has been defined with the relationships option in Microsoft Access. The actual LEITAP results are stored as separate tables in the demand.mdb (see Section 3.2).
- DemandSettings, the name of the appropriate DemandSettingsSet. This name is read from the separate DemandSettingsSet table through a link has been defined with the relationships option in Microsoft Access. The actual demand settings are stored as separate tables in the demand.mdb (see Section 3.2).
- PhoenixScenario, referring to the name of the required demographic scenario (stored as separate table in demand.mdb with the prefix demography).

The addition of a new demand set is fairly easy when existing components are used as is shown in the figure below for a new demand set called *test*. When new demand settings or demand base data related to agriculture or demography are to be included, more actions are required as is described in the next subsection.

Figure 8 Screenshot of the DemandSets table in the metadata.mdb.

To apply the new demand sets in land-use simulation, the following steps are necessary:
1. A (new) policy alternative (model run) should be defined in the Scenarios table that refers to the new demand set (see Section 3.3);
2. The new or revised policy alternative should be applied in the demand module in the EU-ClueScanner to calculate the demand per land-use type, per year for each country following the changed settings and existing demand base data. This calculation starts after activating the Result data item in the Tree view under Demand -> runs -> *PolicyAlternativeName* -> FileGenerator. A demand file for an individual country can be created by activating the data item demandFile in Demand -> runs -> *PolicyAlternativeName* -> per_country -> *CountryName* -> Results.

The demand specification resulting from running the demand module is stored as a Demand.in1 text file in the folder \LocalData\EuClueScanner\Demand*PolicyAlternativeName\CountryName*. Please note that existing demand specification files are not directly replaced when the demand for an already existing policy alternative is calculated anew. In this case a new Demand.in1.tmp file is created, that will only replace the initial file when the EU-ClueScanner model is closed. Upon closing the model, the initial files are renamed to Demand.in1.old to allow the user to recover the initial demand settings when necessary. Appendix 4 provides a short description of the format of the demand files.

When the demand files have been generated for a new policy alternative, the simulation of future land use can be started as was described in Section 2.5

Adding new demand settings
New demand settings can be added by creating a new table in the demand.mdb. To make sure that the right structure is used, it is best to copy an existing demand settings table and paste it (with structure and data) as new table within the demand.mdb. The copied table should be given a sensible short name with the prefix demand_settings. Please note that certain characters (e.g. spaces, slashes are not allowed as tree item names in the GeoDMS and should thus be avoided. The demand settings of the new table can then be adjusted to, for example, change the amount of urban land per inhabitant (see also Section 3.2).

To be able to use the new demand settings created in the demand.mdb in the specification of a new demand set the following steps are necessary:
1. First it is necessary to create a reference to the new demand settings in the metadata.mdb. This should be done by adding a new record to the DemandSettingsSet table in the metadata.mdb with the exact name of the new demand settings table excluding its demand_settings prefix. When these names do not match the demand module in the

EU_ClueScanner will not be able to combine data from the two databases and produce an error.

2. The new settings can then be combined with the other demand base data sets (e.g. a demographic scenario from Phoenix model) in the DemandSets table to create a (new) demand set. Please note that the DemandSets table has to be reopened in order to read the newly defined DemandSettingsSet. The new or revised demand sets can then be used to define a new policy alternative as was explained above.

Adding new demand base data

New base data referring to demographic developments or the demand for agricultural commodities can be added to the demand.mdb. Such new demographic or agricultural projections can be created with the PHOENIX or LEITAP/IMAGE runs or be inferred from other sources. To be able to directly use the new demand input it is necessary that exactly the same table format is used as in the current demand base data. This means that, for example, the same regional division (mostly countries) and units (inhabitants) are used.

For the LEITAP/IMAGE model a structured file format is agreed upon. The results of a LEITAP/IMAGE run are delivered as .csv files (comma separated value files commonly used for data export that in this case use a semicolon as separator) with the following attributes:
- Country;
- Year;
- PermanentPasture;
- AreaArableLand;
- AreaArableBiofuels;
- AreaPerennialBiofuels.

New LEITAP model run results should be stored as .csv files in the \projects\EUClueScanner\data\ Demand folder. They can be imported in the demand.mdb (through the options external data -> textfile) as one table for each .csv file. The .csv filename should correspond with the name of the table. In the GtapSets table in the metadata.mdb a record is then added for each imported csv table in the demand.mdb. In this GtapSets table, each Gtap set has an ID (autonumber), a name, a tablename (which needs to match with the name of the table in the demand.mdb) and a SocEconDbName. This last attribute refers to additional databases for these LEITAP results with economic indicators, stored in the \projects\EUClueScanner\data\indicator_econ folder. In the DemandSets table the name of the GtapSet is selected to specify which LEITAP model run results are used.

After the .csv files are imported in the existing structure in the demand.mdb database they can be used to calculate new demand specification files when they are referred to in a new demand set specification (i.e. a new record in the DemandSets table as was described above). In this version of the EU-ClueScanner the demands are calculated at the national level. The module could also function at sub- or supra-national level. This implies the data used in the module is available or can be (des)aggregated to the requested level. Technically it is possible to calculate the demands at other regional levels (as long as the number of regions does not become too large). But some adaptations to the basic model configuration are needed for which some technical support may be required for non-expert users.

3.5   Adding new spatial data sets

New spatial data sets (with a new name) referring to, for example, accessibility or spatial policy can be added to the model. Typically, these will be created in a GIS environment. They should have the same grid specifications (extent, location, and cell size) and be created in the same 1972-WGS

Albers projection (see Appendix 3) as the other data sets available in the model. They can be stored together with all other datasets in SourceData\EuClueScanner\Drivers. After placing data in this location, references have to be made to it in the EU-ClueScanner. In the Microsoft Access-database 'Metadata.mdb', which can be found in the Projects\EuClueScanner\data directory, a row can be added to the FactorData table specifying the correct FileName and other characteristics. The data set can now be included in the definition of simulation runs (e.g. as factor in the calibration of the model, see Section 4.2, or as a new allow driver map, see Appendix 4) or be used in indicator calculations. Note that spatial datasets that are used to increase or decrease the local suitability values for a land-use type in the context of a policy alternative are applied in simulation as part of a single locspec file (see Appendix 4).

3.6   Editing basic simulation settings

The basic simulation settings are stored in the main.1 and other parameter files. These text files can be edited with any text editor. Appendix 4 describes the content of these files and, where appropriate, specifies alternative settings. As consequences of editing these parameter files can be substantial a thorough understanding of their functioning is mandatory. This can be obtained by reading the additional material referred to in Appendix 4. Especially the available tutorials for the classic CLUE model are extremely helpful in getting to know the basics of this model.

4 Adjusting the EU-ClueScanner model

Before discussing possible adjustments to the existing EU-ClueScanner model configuration, it is good to highlight two specific model characteristics: the *administrator mode* and the GeoDMS *script files* that comprise the model.

For more advanced users a specific *administrator mode* exists that allows the inspection of basic model settings (e.g. units, standard regional divisions, classification schemes for the visualisation of data), many in-between steps in simulation, auxiliary data files and templates used for the creation of indicator values etcetera. This mode is active when the administrator mode tick box is clicked in the General Settings tab in the Options window that appears by clicking Ctrl+Alt+O (see Figure 2). By now you may be expert enough to use this mode.

Most of EU-ClueScanner components are programmed in the GeoDMS *script files* that can be edited with any text editor. The user interface only offers limited functionality to edit these files, as previous experience has learnt that direct editing of the underlying script files offers more flexibility, a more compact and robust programming environment and a better overview of the context of these files and their relation with other model components. The GeoDMS files define the actual modelling application and offer an open and flexible environment to manipulate its many components. The definition of indicator calculations, the inclusion of new spatial datasets and model run characteristics are also defined in the GeoDMS files. For a complete description of the naming conventions, available operators and functions, potential configuration errors and many other aspects of GeoDMS scripting, see the modellers guide at http://www.objectvision.nl/GeoDMS/ under modelling.

4.1 Replacing land-use base data

The current EU-ClueScanner model uses the CLC2000 data to represent current land use. In the near future a revised version (CLC2006) is expected to be released, that would allow the incorporation of a more recent starting year for simulation. To replace the current land use data set a few steps are necessary.

First a reclassified version of the new land-use dataset must be created that uses the same land-use classification and has exactly extent and projection as the current land-use base data set. Table 1 specifies how the current, extensive set of CLC types is reclassified to the LU18 classification used in the model. It is important that this reclassification scheme is followed closely to prevent the introduction of inconsistencies with previous simulation results. The reclassified data should then be stored as an ascii grid (called 16pp.asc), closely following the projection and data file format specifications listed in Appendix 3. This new data set can then replace the current version stored in \SourceData\EuClueScanner\LandUse.

A second, more complex step is adjusting the demand settings (see also Section 3.2). This calls for a thorough analysis of the current specification and a revision of those settings that are not appropriate anymore in the demand_settings tables in the demand.mdb. This may, for example, relate to the year set aside subsidies cease. Exactly which settings should be changed and how this should be done, depends on the year the new base data refer to and the scenario and policy alternative under consideration.

Then, the start and end year of simulation are specified in the main.1 file for each country and policy alternative in the in \SourceData\EuClueScanner folder. This setting dates back to the initial

version of the DynaClue model and is not used any longer, so there is no need to adjust it. In the EuClueScanner model the simulatioin period is now specified in the PeriodSet column in the Scenarios table of the metadata.mdb. A new period may have to be referred to with a new name in this column. When this is done, the characteristics of this new period have to be specified in the default.dms scriptfile in the folder \Projects\EuClueScanner\cfg\default. This new specification can be based on already existing examples. In fact, copying the lines of script relating to an existing period and editing these may the most efficient strategy.

Finally, the spatial data sets related to the age of land use and dynamic simulation should be considered carefully. The age of land use is stored in the age0.asc file in the folder \SourceData\EuClueScanner\LandUse and may have to be updated. Ideally this calls for a detailed transition analysis comparing land use in the initial base year (2000) and the new base year. Based on this, locations where land use remained the same will increase in age (with the difference in years between the two observation moments), whereas changed locations will receive a new age of 0 years. Also the dynamic driver files (stored in \SourceData\EuClueScanner\DynamicDrivers) contain explicit mentioning of specific years and may have to be adjusted. Again this depends on the new period of simulation and the characteristics of the policy alternative under consideration.

When new land-use base data become available, the user can also to consider updating the calibration. However, as the driving forces that influence land-use patterns are considered to be relatively stable over time this is not essential.

4.2   Revise calibration

Calibration is the process of tuning a model in such a way that sensible results can be obtained. This is normally done through a statistical (logistic regression) analysis that describes the relation between a number of spatially explicit explanatory variables (e.g. accessibility, soil type and climatic variables) and an observed (to be explained) land-use pattern (Pontius Jr. et al., 2008). These explanatory variables are called factor data in the model.

In case of the Eu-ClueScanner a separate calibration was performed for each individual country, to reflect the fact that driving forces behind land-use change may differ between countries in specification and explanatory power. This means that different factors are used to explain land-use patterns per country; in general about 10 to 20 different factors are used for a country from the total list of over 50 factors described in Appendix 2. In addition, explanatory variables are used that describe land use in the direct environment (or neighbourhood) of a grid cell.

Binomial logistic regression was applied to estimate the probability of occurrence of each simulated land-use type individually for each country, based on a set of explanatory variables representing different driving forces. An advantage of binomial logistic regression is that different explanatory variables can be used to explain the occurrence of different land-use types, something that is more difficult to realise in multinomial logistic regression that can also be applied to calibrate land-use models (Loonen and Koomen, 2009). In addition, binomial regression also allows the user to revise the existing calibration of a single land-use type or add reference to a new land-use type, where multinomial regression would call for a complete new calibration effort.

The Eu-ClueScanner application for DG Environment simulates the location of seven types of land use (see Table 1), most of which are steered by a statistical calibration of underlying factor data. Only recently abandoned arable land and pastures are governed fully by transition rules that are applied in subsequent modelling step and are thus not relying on the statistical calibration. The current calibration used Corine Land Cover data for 2000 and thus describes the importance of various driving forces in explaining the land-use patterns of that year. We assume, however, that

the observed relations also hold true for the simulation of future land use. It should, furthermore, be noted that the importance of the factor data was estimated separately from the importance of the neighbourhood-related variables. The results of the statistical analysis are used to describe the importance of different driving forces and are stored in files called alloc1.reg (for factor data) and alloc2.reg (for neighbourhood-related data) in the Sourcedata\EuClueScanner\ ScenarioName directory. See Appendix 4 for a more extensive description of their layout and content. The overall weight of the neighbourhood data in relation to the factor data is specified for each land-use type individually in the neighbourhood settings file (neighmat.txt). All these model specification files will normally be the same for different scenarios, but they are stored in scenario specific folders to allow for the inclusion of different calibration values per scenario.

A revision of the calibration becomes necessary when new land-use or factor data are added to the model, or when a land-use type is introduced. In this case a new statistical analysis should be performed that uses an observed land-use (e.g. of the new base year) as to be explained variable and a set of explanatory variables. Standard statistical software such as SPSS can be used to do this type of analysis. To keep calculation times within reasonable bounds it is advised to limit the number of observations, depending on the calculation power and available internal memory in the computer. The binomial logistic regression specification used to calibrate the current model version used a balanced sample (with approximately equal amounts of 0 and 1 values for the dependent variable) of about 10,000 observations per land-use type. The sample should preferably be structured with a minimum distance between the observations to limit spatial autocorrelation. In addition it is suggested to limit the number of to be explained variable values (land-use types) to about 9 and to make sure that all land-use types have sufficient observations (at least 2% of the total amount of observations) to prevent the creation of statistically insignificant results. A thorough statistical background is required to perform such analyses and assess the value of their outcomes.

4.3   Edit or add indicator definitions

The indicators are defined in a set of GeoDMS script files that are stored in the folder \projects\EuClueScanner\cfg\Default. Initial reference to the indicators is made in the IndicatorData.dms file. This file also defines some of the units and legends that are used in their calculation and representation. Their actual calculations are defined in the indicator-specific GeoDMS files in the IndicatorData subfolder. Additional spatial data sets that are needed for the calculation of the indicator values are stored in the \SourceData\EuClueScanner\Indicators folder. These data sets are independent of the policy alternatives and are thus stored in this general folder that also contains the factsheets describing, for example, the origin and methodology of the indicators.

The definition of the indicators can be edited by manipulating the corresponding GeoDMS files. Obviously this requires a thorough understanding of the current indicator definitions and the basics of the GeoDMS scripting language. Straightforward changes can, however, already be made by relatively novel users of the GeoDMS language.

New indicators can be added by defining them in separate GeoDMS files and referring to them in the IndicatorData.dms file through an include statement (e.g. #include <NewIndicator.dms>, see similar references in IndicatorData.dms). The GeoDMS language offers a large set of operators that allow the user to define new quantitative evaluation measures. The operators include more complex spatial analysis options generally found in more advanced GIS-packages that relate to, for example, combining, comparing and overlaying individual data layers, defining contiguous areas with the same raster values, calculating distances or network analysis. ,A full list of these operators is available in the GeoDMS Modeller's guide that can be found under modelling on: http://www.objectvision.nl/GeoDMS/. Examples of other indicators developed within the same

modelling environment are described elsewhere (Bubeck and Koomen, 2008).

4.4   Link EU-ClueScanner to other models

The EU-ClueScanner model can be linked with other spatial models related to either derive more specific input for simulation (e.g. changed regional agricultural demand from CAPRI, or revised accessibility maps from TRANSTOOLS) or provide additional impact assessments (e.g. through coupling with a hydrologic model). These model couplings require a clear vision on the anticipated level of integration and a careful consideration of the thematic, temporal and spatial resolution of the involved models. A model coupling can be a straightforward exchange of output and input data when the considered land-use types (thematic resolution), time period (temporal resolution) and regional divisions or grid cell size (spatial resolution) are aligned. Substantial efforts are, however, required when the models need to be adjusted. This would, for example, be the case when additional agricultural land-use types have to be inserted in the land-use model to allow the input from an agro-economic model such as CAPRI.

Although coupling of the modelling framework to alternative detailed indicator models is possible it may not always be recommended. Many indicator models are based on detailed understanding of processes at the micro-level (e.g. causing greenhouse gas emissions) and are therefore subject to scaling errors when applied to a 1 km spatial resolution. It is thus important to choose indicator models that are suited for and sensitive to the information provided by the EU-ClueScanner framework at the spatial and temporal scale of analysis. Also a good fit with the thematic content of the different land-use classes is necessary. For example, due to the limited differentiation currently possible in land-use intensity, no specific indicator on the agricultural biodiversity is included in the modelling system given the dependence of this indicator on detailed changes in land-use intensity.

Appendix 1 Overview of original CLC2000 classes

Corine land cover classes

1. Artificial surfaces
1.1 Urban fabric
- 1.1.1. Continuous urban fabric
- 1.1.2. Discontinuous urban fabric

1.2 Industrial, commercial and transport units
- 1.2.1. Industrial or commercial units
- 1.2.2. Road and rail networks and associated land
- 1.2.3. Port areas
- 1.2.4. Airports

1.3 Mine, dump and construction sites
- 1.3.1. Mineral extraction sites
- 1.3.2. Dump sites
- 1.3.3. Construction sites

1.4 Artificial, non-agricultural vegetated areas
- 1.4.1. Green urban areas
- 1.4.2. Sport and leisure facilities

2. Agricultural areas
2.1 Arable land
- 2.1.1. Non-irrigated arable land
- 2.1.2. Permanently irrigated land
- 2.1.3. Rice fields

2.2 Permanent crops
- 2.2.1. Vineyards
- 2.2.2. Fruit trees and berry plantations
- 2.2.3. Olive groves

2.3 Pastures
- 2.3.1. Pastures

2.4 Heterogeneous agricultural areas
- 2.4.1. Annual crops associated with permanent crops
- 2.4.2. Complex cultivation patterns
- 2.4.3. Land principally occupied by agriculture
- 2.4.4. Agro-forestry areas

3. Forest and seminatural areas
3.1 Forests
- 3.1.1. Broad-leaved forest
- 3.1.2. Coniferous forest
- 3.1.3. Mixed forest

3.2 Shrub and/or herbaceous vegetation associations
- 3.2.1. Natural grassland
- 3.2.2. Moors and heathland
- 3.2.3. Sclerophyllous vegetation
- 3.2.4. Transitional woodland shrub

3.3 Open spaces with little or no vegetation
- 3.3.1. Beaches, dunes, and sand plains
- 3.3.2. Bare rock
- 3.3.3. Sparsely vegetated areas
- 3.3.4. Burnt areas
- 3.3.5. Glaciers and perpetual snow

4. Wetlands
4.1 Inland wetlands
- 4.1.1. Inland marshes
- 4.1.2. Peat bogs

4.2 Coastal wetlands
- 4.2.1. Salt marshes
- 4.2.2. Salines
- 4.2.3. Intertidal flats

5. Water bodies
5.1 Inland waters
- 5.1.1. Water courses
- 5.1.2. Water bodies

5.2 Marine waters
- 5.2.1. Coastal lagoons
- 5.2.2. Estuaries
- 5.2.3. Sea and ocean

Appendix 2 Overview of spatial data sets

This appendix provides an overview of two types of spatial data sets:
- the 52 factor data sets that are used as independent factors in the statistical calibration of the model. These are called upon in the alloc1.reg and alloc2.reg files. The datasets are made available from the EURURALIS 2.0 project and are described in more detail in the meta data sheets (adapted from: Overmars et al., 2006) that are included as part of the details pages in the model.
- the 13 allow driver maps (X1-X13) that specify the spatially explicit settings for the conversion matrix. These maps are called upon in the allow.txt files. The meta data sheets that are included as part of the details pages in the model describe these maps in more detail.

ID	Name	Description	Coverage
Factor data set used in calibration of the model			
1	ACCESS1_06M	Timecost to cities > 100.000 (seconds)	missing Turkey (Cyprus and Malta are reasonably ok)
2	ACCESS2_06M	Timecost to cities > 500.000 (seconds)	missing Turkey and Cyprus and Malta
3	ACCESS3_06M	Timecost to ports > 15.000 kTon/year (seconds)	missing Turkey and Cyprus and Malta
4	ACCESS4_06M	Timecost to cities > 650.000 (seconds)	missing Turkey and Cyprus and Malta
5	ACCESS5_06M	Euclidean distance to nearest road of level 0 or 1 (metres)	missing Turkey and Cyprus and Malta
6	ACCESS6_06M	Timecost to major airports (seconds)	missing Turkey and Cyprus and Malta
7	ACCESS7_06M	Timecost to airports & ports (seconds)	missing Turkey and Cyprus and Malta
8	claycont_06pc	Soil clay content (%)	missing Iceland, Turkey, Cyprus and Malta
9	ddw_shortage	Water deficit growing season (mm)	ALL
10	dem_final	Elevation based on SRTM3 NASA-data (m)	ALL
11	envmap01	ALN: Alpine north (0/1)	missing Iceland, Turkey, Cyprus
12	envmap02	BOR: Boreal (0/1)	missing Iceland, Turkey, Cyprus
13	envmap03	NEM: Nemoral (0/1)	missing Iceland, Turkey, Cyprus
14	envmap04	ATN: Atlantic north (0/1)	missing Iceland, Turkey, Cyprus
15	envmap05	ALS: Alpine south (0/1)	missing Iceland, Turkey, Cyprus
16	envmap06	CON: Continental (0/1)	missing Iceland, Turkey, Cyprus
17	envmap07	ATC: Atlantic central (0/1)	missing Iceland, Turkey, Cyprus
18	envmap08	PAN: Pannonian (0/1)	missing Iceland, Turkey, Cyprus
19	envmap09	LUS: Lusitanian (0/1)	missing Iceland, Turkey, Cyprus
20	envmap10	ANO: Anotolian (0/1)	missing Iceland, Turkey, Cyprus
21	envmap11	MDM: Mediterranean mountains (0/1)	missing Iceland, Turkey, Cyprus
22	envmap12	MDN: Mediterranean north (0/1)	missing Iceland, Turkey, Cyprus
23	envmap13	MDS: Mediterranean south (0/1)	missing Iceland, Turkey, Cyprus
24	EUAC120_2006	# of people that reach a location from their home within 120 minutes (persons)	EU15
25	EUAC30_2006	# of people that reach a location from their home within 30 minutes (persons)	EU15
26	EUAC60_2006	# of people that reach a location from their home within 60 minutes (persons)	EU15
27	Geomorf01	Average height difference of 0-20 m: flat (0/1)	missing Turkey
28	Geomorf02	Average height difference of 20-80 m: rolling (0/1)	missing Turkey
29	Geomorf03	Average height difference of 80-200 m: hilly	missing Turkey

ID	Name	Description	Coverage
		(0/1)	
30	Geomorf04	Average height difference of 200-400 m: mountainous (0/1)	missing Turkey
31	Geomorf05	Average height difference of > 400 m: very mountainous (0/1)	missing Turkey
32	IL_2006	Presence of an impermeable layer within the soil profile (0/1)	missing Iceland, Turkey, Cyprus, former Yugoslavia
33	landsc_06	ORN- LandScan population, maximum set to 4000 (persons)	ALL
34	mean_temp_06	Mean yearly temperature	ALL
35	Peat_06	Presence of peat in European soil database of JRC (0/1)	missing Iceland, Turkey, Cyprus
36	poppot_1mi06	Population potential with 12.5 km inflection point, max set to 1 million (persons)	missing Iceland, Turkey
37	poppot_log06	Log (population potential with 12.5 km inflection point)	missing Iceland, Turkey
38	poppot_sum06	Gaussian population potential with 12.5 km inflection point (persons)	missing Iceland, Turkey
39	rain_wc_5m	Accumulated rainfall March-July (mm)	ALL
40	rain_wc_yr	Accumulated rainfall per year (mm)	ALL
41	Salinity	Presence of saline soils according to European Soil Database of JRC (0/1)	missing Iceland, Turkey, Cyprus and Malta
42	slope_final	Slope based on NASA's SRTM3 elevation data (degrees)	ALL
43	soildepth_06	Soil depth based on European soil database of JRC (cm)	missing Iceland, Turkey, Cyprus, Malta, Switzerland, Yugoslavia
44	stoniness100	Stoniness based on European soil database of JRC (0/1)	missing Iceland, Turkey, Cyprus and Malta
45	Swap	Soil water available to plants based on European soil database of JRC (mm)	EU25+2 MINUS nor/swe/fin and Cyprus and Malta
46	sz_landsc_rur	ORN- LandScan population, maximum set to 100 (persons)	EU25+2 MINUS nor/swe/fin and Cyprus and Malta
47	t_min0_1000	Count of months a year with average temperature < 0 degrees C (months)	missing Iceland, Turkey, Cyprus
48	t_plus15_1000	Count of months a year with average temperature > 15 degrees C (months)	missing Iceland, Turkey, Cyprus
49	wr_06	Soil with water restriction (too much water) based on JRC European soil database (0/1)	missing Iceland, Turkey, Cyprus AND and Malta
50	ac_cst_c_m06	Distance to coast for Malta and Cyprus (m)	only Cyprus and Malta
51	ac_pnts_c_m06	Distance to city/airport/port for Malta an Cyprus (m)	only Cyprus and Malta
Allow driver maps used in specification of land-use succession			
52	X1	Natura2000 (0 or outside = 1; transition will only be allowed outside protected areas)	EU27 for other countries incomplete/ unreliable depending on base data
53	X2	Erosion sensitive areas (0/1)	idem
54	X3	Natura2000 and erosion sensitive areas (0/1)	idem
55	X4	Natura2000, erosion sensitive areas and peat (0/1)	idem
56	X5	Erosion sensitive and peat areas (0/1)	idem
57	X6	Natura2000 and peat areas (0/1)	idem
58	X7	Natura2000 and river flood prone areas (0/1)	idem

ID	Name	Description	Coverage
59	X8	Succession abandoned arable to semi-natural (succession code denoting that this transition is enforced after a specified number of years; coded as 1000+number of years).	idem
60	X9	Succession abandoned pasture to semi-natural (succession code)	idem
61	X10	Succession semi-natural to forest (succession code)	idem
62	X11	Succession abandoned arable to semi-natural for Biodiversity scenario (succession code)	idem
63	X12	Succession abandoned pasture to semi-natural for Biodiversity scenario (succession code)	idem
64	X13	Succession semi-natural to forest for Biodiversity scenario (succession code)	idem

Appendix 3 Spatial data set format details

The EU-ClueScanner uses data of different resolutions (1km and 100m), different projections (Albers and Lambert) and different data file formats (ascii grid and tiff). This appendix specifies these aspects of the data. The EU-ClueScanner 1km (the version for DG Environment) uses data related to a 1km grid defined in the Albers projection. The EU-ClueScanner 100m also contains data related to a 100m grid defined in the currently more common Lambert projection. The specifications of the latter data are also provided to help with potential data exchanges between the different model versions. The EU-ClueScanner 100m contains a lookup matrix that relates the cells of the 100m grid on Lambert coordinates to the cells of the 1km grid on Albers coordinates. This matrix is used to reproject 1km factor data for use in the model application at a 100m grid.

EU-ClueScanner 1km

Projection
The initial data sets in the 1km version of EU-ClueScanner use the WGS-1972 Albers projection with the following specifications:

PROJCS["WGS_1972_Albers",
GEOGCS["GCS_WGS_1972",
DATUM["D_WGS_1972",
SPHEROID["WGS_1972",6378135.0,298.26]],
PRIMEM["Greenwich",0.0],
UNIT["Degree",0.0174532925199433]],
PROJECTION["Albers"],
PARAMETER["False_Easting",0.0],
PARAMETER["False_Northing",0.0],
PARAMETER["Central_Meridian",22.65],
PARAMETER["Standard_Parallel_1",32.5],
PARAMETER["Standard_Parallel_2",54.5],
PARAMETER["Latitude_Of_Origin",51.4],
UNIT["Meter",1.0]]

Ascii format specification
Most spatial data sets in the model are stored as TIFF, but also grids in ascii file format can be read. This file format is supported by, for example, ArcGIS. These files have a header that describes the main characteristics. The standard values for the EU-ClueScanner application for DG Environment are:

ncols	5856
nrows	4092
xllcorner	-3,932,373.5
yllcorner	-1,869,965.125
cellsize	1000
NODATA_value	-9999

In which:
ncols represents the number of columns;
nrows represents the numbers of rows;
xllcorner represents the x-lower left coordinate;
yllcorner represents the y-lower left coordinate;
cellsize represents the cellsize in meters; and

NODATA_value represents the value that the model interprets as a cell with no value.

EU-ClueScanner 100m

Projection
In the 100m model, all datasets use the ETRS 1989 Lambert projection with the following specifications:

PROJCS["ETRS_1989_LAEA_L52_M10",
GEOGCS["GCS_ETRS_1989",
DATUM["D_ETRS_1989",
SPHEROID["GRS_1980",6378137.0,298.257222101]],
PRIMEM["Greenwich",0.0],
UNIT["Degree",0.0174532925199433]],
PROJECTION["Lambert_Azimuthal_Equal_Area"],
PARAMETER["False_Easting",4321000.0],
PARAMETER["False_Northing",3210000.0],
PARAMETER["Central_Meridian",10.0],
PARAMETER["Latitude_Of_Origin",52.0],
UNIT["Meter",1.0]]

Data format specifications
Due to size constraints, the 100m model only supports TIFF files stored with smallest possible bit size (i.e. 1 bit for binary data, 8 bits for other types) and LZW compressed, with the following specifications:

      Cell size: 100 m
      Size: 60,000 columns * 45,300 rows
      Pixel Size = (100;-100)
      Corner Coordinates:
      Upper Left (1,500,000; 5,430,000)
      Lower Left (1,500,000; 900,000)
      Upper Right (7,500,000; 5,430,000)
      Lower Right (7,500,000; 900,000)

It is possible to store grid data as TIFF with ArcGIS software package, using ArcToolbox[TM] *Data Management Tools -> Raster -> Raster Dataset -> Copy Raster.* The file extension (.tif) must be specified while naming the file. The projection and format specifications referred above can be defined using the *Environments* options, namely *General Settings (Output coordinate system* and *Extent*) and *Raster Analysis Settings* (*Cell size*). The most straightforward way to define the projection is using the file ETRS_1989_LAEA_L52_M10.prj, which can be found e.g. in the folder SourceData/JRC/landuse. After this, the created TIFF file has to be exported using File -> Export map, in order to define the bit size and compression mode. To export from ArcGIS as TIFF in a compressed (LZW) format may turn out to be difficult. We solved this by specifying LZW as the default option in the ArcGIS preference definition file for TIFF files (C:\Program Files\Common Files\ESRI\Raster\defaults\tiffs.pdf). Notwithstanding its name this is a plain text file that can be edited in any text editor.

When importing the dataset to the model, be sure to include a copy of the projection file in the folder.

Appendix 4 Model specification files

This appendix explains the contents of the most important model specification files. Most of these files are stored in country-specific folders in the policy-alternative folders (e.g. ReferenceB1) in the SourceData directory. Only the demand data that are calculated by the demand module are stored in the Demand folder of the LocalData directory.

The model specification files of the EU-ClueScanner are based on the well-known CLUE model. For people that want to get familiar with the basic elements (including the specification files) of this land-use model, reference is made to the tutorial available for the CLUE model. This can be downloaded from the website www.cluemodel.nl under the download caption. The original Clue help file is included in the Projects\EuClueScanner\doc folder. This extensive set of hyperlinked HTML pages is a valuable source of background information on the basics of this model.

main.1
The main.1 file contains all important parameters that determine the configuration of the simulation. All main.1 files contain the following parameters:

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	X-lower left coordinate	Float
9	Y-lower left 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
18	Neighbourhood calculation choice	0, 1 or 2
19	Location-specific preference addition	Integers

A more extensive description of these parameters is provided below:
1. The number of land-use types that are distinguished in simulation. A maximum of 12 different land-use types can be identified.
2. The number of regions, default is 1 and the maximum number of regions is 3. Normally a study area consists of only one region, but for study areas that are divided in parts with a very different behaviour of land-use types, more regions can be used. Please note that in the EU-ClueScanner separate models are specified for individual countries that are not subdivided in smaller regions.
3. The maximum number of independent variables in a regression equation. This is the number of driving factors of the regression equation with most variables.
4. The total number of driving factors that are addressed, this means the number of driving

factor files per year.

5. Number of rows of the input grids.
6. Number of columns of the input grids.
7. The cell area of the grid cells, this should be in the same units as in the demand file. In our case this is 1 km^2.
8. X-coordinate of the lower left corner, this is also indicated in the header of the ArcView and ArcGIS ASCII files.
9. Y-coordinate of the lower left corner, this is also indicated in the header of the ArcView and ArcGIS ASCII files.
10. The internal number coding of the land-use types used in the simulation process. This should start with 0 for the first land-use type (= Clue10-type 0 introduced in Table 1), 1 for the second land-use type, etc. This line of script is also used to state which land-use types are simulated together (i.e. have the same demand for land). This is the case with the recently abandoned arable land (Clue10-type 5), forest (Clue10-type 7) and recently abandoned pasture land (Clue10-type 8) in the example below that are all treated as natural vegetation (Clue10-type 3).
11. The codes for the allowed changes and behaviour of the land-use types (conversion elasticities) have to be indicated for each land-use type, following the order of line 10. This must be a value between 0 and 1.
    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.
12. Iteration variables: three numbers should be specified:
    Iteration mode: with the options: 0, which means that convergence criteria are expressed as a percentage of the demand; or 1, meaning that these criteria are expressed as absolute values (units of demand)
    First convergence criterium: average deviation between demanded changes and actually allocated changes (default for %: 0.35; otherwise minimal 1 grid cell divided by the number of land-use types).
    Second convergence criterium: maximum deviation between demanded changes and actually allocated changes (default for %: 3; otherwise should minimally be the percentage of 1 cell change of the land-use type with the smallest demand, e.g. for a land-use type with a demand of 200 ha and a cell area of 4 ha, the value should be minimally 4/(200/100) = 2).
13. Start year and end year of the simulation.
14. Number of dynamic driving factors, e.g. population density. This number should be followed by the coding of these explanatory factors, e.g. 2 9 11, which means that driving factors 9 and 11 are dynamic and have input files for each year. These file codes correspond to the numbers listed in the FactorData table in the metadata.mdb. Appendix 2 also lists these numbers. When no dynamic drivers are used a '0' should be paced on this line.
15. Choice for the type of output file:
    1. ArcView-ArcGIS headers will be printed in output files;
    0. No headers in output files (suited for e.g., Idrisi).
    -2. No headers in output files, suppress information in log-file on iteration procedure.
    2. ArcView-ArcGIS headers will be printed in output files, suppress information in log-file on

iteration procedure.

Note: If the ArcView-ArcGIS file type is chosen, all input grid files should contain an ArcView-ArcGIS header.

16.   Choice for a region specific regression:

0. No different regressions for different regions

1. Different regressions for different regions; demands for each region specifically defined

2. Different regressions for different regions; demand for all regions aggregated.

The default is value 0.

17.   When temporal dynamics are taken into account it is necessary to know the land-use history. Often this is not known and a random number has to be assigned. One can choose between three options:

0. The initial land-use history will be read from file age.0, which should contain a grid with, for every pixel, the number of years that the pixel is used for the current land-use type.

1. A random number will be assigned to all pixels to represent the number of years that the current land use is already found at that location according to the standard seed for the random number generator. With this option two runs will result in the same random number, which is useful for comparison.

2. A random number will be assigned to all pixels to represent the number of years that the current land use is already found at that location with a different random number generator. Two consecutive runs will start with different land-use histories.

For option 1 or 2 an additional number should be added that indicates the maximum number of years that can be generated by the randomizer.

18.   Choice for using the neighbourhood function:

0. Neighbourhood function is not used.

1. Neighbourhood function is used in simulation.

2. Only the influences are calculated, the influence files are saved directly, no simulation.

19.   Variables for location specific preference addition. The first number is a switch: activate the function (1) or not (0). If the switch is set to 1, it should be followed by, for each land-use type, the fraction of the preference addition that is added to the probability as calculated in the regression.

Below is an example of a main.1 file used in the EU-ClueScanner model.

```
line 1     10
line 2     1
line 3     18
line 4     64
line 5     294
line 6     577
line 7     1
line 8     -979373.5
line 9     -537965.125
line 10    0 1 2 3 4 3 5 3 3 6
line 11    1 0.5 0.3 0.8 1 0.3 0.9 0.8 0.3 1
line 12    1 21 84
line 13    2000 2030
line 14    5 33 34 39 40 46
line 15    2
line 16    0
line 17    0
line 18    1
line 19    1 0 0.2 0.2 0 0 0 0.2 0 0 0
```

demand.in1

The required amount of land (demand) per year, per land-use type, per region is stored in demand.in1 text files. These files contain for each year that needs to be simulated the demand in km^2 (see example below). The first line specifies the number of years for which demand is included in this file. In our case this is 31 (the initial year and 30 subsequent simulation years). After that, every line contains the demands for the respective land-use types for one year. The order of the land-use types should be the same as in the main parameter file (main.1) and the second line of the file contains the demands for year 0. Note: the total demand for all land-use types together should be equal to the total land area from the region file and should remain constant for each year.

Below is an example of the first 7 lines of a demand.in1 file for one demand region (normally country). The demand for land-use type 0, for example, increases, every year with about 5 km. The example also makes clear that several land-use types have the same demand. These are types denoted with a '3' (natural vegetation) in line 10 of the main.1 file. It, furthermore, shows that the demand for the last land-use type (Clue10-type 9, or all non-simulated land uses) remains stable

```
31
3525 14472 11922 43538 0 43538 699 43538 43538 9928
3530.18 14342.1 12000.3 43600.1 0 43600.1 683.21 43600.1 43600.1 9928
3535.36 14212.3 12088.9 43652.1 0 43652.1 667.42 43652.1 43652.1 9928
3540.54 14082.4 12186.3 43695.2 0 43695.2 651.63 43695.2 43695.2 9928
3545.72 13952.5 12291.3 43730.7 0 43730.7 635.84 43730.7 43730.7 9928
3550.9 13822.7 12402.9 43759.5 0 43759.5 620.05 43759.5 43759.5 9928
etc.
```

Remember, these files are calculated within the EU-ClueScanner by the demand module. The files are not intended to be edited by the user. Instead the input parameters of the Microsoft Access demand database (demand.mdb) located in projects\EuClueScanner\data\demand should be edited.

alloc1.reg

The alloc1.reg file contains the part of logistic regression results that relate to the importance of the suitability factors in explaining observed land-use patterns. This file has the following format (see also the example below):

Line 1: Number code for land-use type (here 0 = urban land use).
Line 2: Constant of regression equation for this land-use type.
Line 3: Number of explanatory factors in the regression equation for that land-use type.
Line 4 and further: On each line the beta coefficients (-0.00022 etc.) for the explanatory factor and the number code of the explanatory factor. The number codes of these factors (*sc1gr#.fil* files) can be found in the FactorData table in the metadata.mdb. Appendix 2 also lists these numbers. In the example below, line 4 refers to ACCESS1_06M: the timecost to cities > 100.000 (seconds).

```
0
  2.27795883488721
5
  -0.000220538796588075 1
  -6.21309139227147E-5 3
  -0.000111193665261406 7
  0.00529359739246736 33
  -0.0815179240484676 42

1
  0.375937349034798
5
  -0.00407050607853913 39
  -0.26956963232153 42
  0.0151641381333413 45
  0.00614536533423984 46
  0.319236906526724 48
```

alloc2.reg

The alloc2.reg file contains the part of logistic regression results that relate to the importance of the neighbourhood enrichment factors in explaining observed land-use patterns. This file has the same format as alloc1.reg:

Line 1: Number code for land-use type.
Line 2: Constant of neighbourhood regression equation for land-use type (ß0).
Line 3: Number of explanatory factors (land-use types) in the regression equation for that land-use type. For example the presence of urban area and arable land in the neighbourhood are explanatory factors for the land-use type urban area, so the number of explanatory factors is 2.
Line 4 and further: On each line the beta coefficients (ß1, ß2, etc.) for the explanatory factors and the number code of the explanatory factor. In using the neighbourhood function, the explanatory factors are the land-use types and therefore the number codes are the number codes for the land-use types.

Land-use conversions can be explained partly by the occurrence of land uses in the neighbourhood. For example, new urban area is more likely to develop at the fringe of existing urban area than elsewhere. Especially in the context of urban growth, it is useful to include neighbourhood interactions as a driving factor (Verburg et al., 2004). To characterise the neighbourhood of a location the enrichment factor is often used. This measure is defined by the occurrence of a land-use type in the neighbourhood of a location relative to the occurrence of this land-use type in the study area. The weight that is assigned to the neighbourhood characteristics is assigned in the neighbourhood settings file (neighmat.txt).

neighmat.txt

In this file the size and shape of the neighbourhoods and the weight assigned to the neighbourhood function is indicated for each land-use type individually. The file has the following format:

Line 1 Weight (0-1) assigned to the neighbourhood function for each land-use type.
Line 2 and further. For each land-use type:
- The radius of the neighbourhood (1 = 1 grid cell on each side of the central cell).
- Followed by the shape of the neighbourhood (with weights (0-1) for individual cells).

Below is an example of the first part of such a file for the EU-ClueScanner model. It shows that for the first land-use type only the first ring of cells surrounding the central cell is selected as neighbourhood. The second reference to this first land-use type indicates that each cell receives an equal weight.

```
0.3 0 0 0.2 0 0 0 0.2 0 0

1
1 1 1
1 0 1
1 1 1

1
1 1 1
1 0 1
1 1 1
```

allow.txt

The allow.txt file specifies how succession of land-use types takes place. It may refer to succession maps that specify locations where certain successions may take place. It is a Y x Y matrix where Y equals the number of land-use types. So when 10 land-use types are simulated by the model it will be a 10x10 matrix. Rows denote the present land-use type and columns the potential future land-use type. A value of 1 indicates that the conversion is allowed, a value of 0 indicates that the conversion is not allowed. Other codes refer to maps that specify whether succession is allowed for individual (1x1 km) locations.

The example below shows a typical allow.txt file. For clarity's sake the related Clue10 land-use type codes (0-9) are included next to the file. The value 52 in row 1 on line 2 in this example denotes that the conversion of the second land-use type (arable land) into the first (urban) is conditional on allow driver map 52. This map represents the Natura2000 areas with a value of 0, indicating that urbanisation is not allowed here (see the FactorData table in the metadata.mdb or Appendix 2 for a complete list of all allow driver maps). The remaining areas contain a value of 1 indicating that conversion is allowed. The conversion settings for the different policy alternatives in the DG-Environment application of the EU-ClueScanner are provided in a separate annex to the final report.

to: from land use	0	1	2	3	4	5	6	7	8	9
0	1	0	0	0	0	0	0	0	0	0
1	52	1	1	0	0	1	52	0	0	0
2	52	61	1	0	0	0	52	0	1	0
3	0	61	52	1	0	0	52	60	0	0
4	0	0	0	0	1	0	0	0	0	0
5	52	61	52	58	0	1	52	0	0	0
6	52	61	52	0	0	1	1	0	0	0
7	0	61	52	0	0	0	52	1	0	0
8	52	61	52	59	0	0	52	0	1	0
9	0	0	0	0	0	0	0	0	0	1

locspecX.fil

The implementation of policy themes is, to some extent, done by location specific modification of the suitability of the land for a specific land-use type. The suitabilities reflect an index of the potential land rent that can be attained at a specific location for a specific land-use type. Scenario settings (subsidies and taxes) influence these suitabilities. Financial support may, for example, be considered for maintaining agriculture in mountainous areas. This leads to a higher probability that agricultural land use is conserved in these areas than could be expected from the statistical regression between land use and (biophysical and socio-economic) driving factors that apply for the study area as a whole. Such modifications are reflected in the location-specific addition factors (locspec). These location-specific addition factors for different policies are combined in one map for each land-use type. See the appendix to the final report for an overview of the way various policies are combined in the current application.

When this option is used, for each land-use type a file named locspec#.fil (# indicates the land-use type number) should be supplied in the country-specific folders of the relevant policy-alternative folders (e.g. ReferenceB1) in the SourceData directory. These additions are stored in maps that provide a value in each grid cell that should preferably lie between 0-1 in order to fit the range of the regression results. Note that the switch on line 19 in the main.1 file should be set to 1 and weight factors should be specified. The function is not active when the switch in main.1 has the value 0.

References

Bubeck, P. and Koomen, E. (2008) The use of quantitative evaluation measures in land-use change projections; An inventory of indicators available in the *Land Use Scanner*. Spinlab Research Memorandum SL-07. Vrije Universiteit Amsterdam/ SPINlab, Amsterdam.

Dekkers, J.E.C. and Koomen, E. (2007) Land-use simulation for water management: application of the Land Use Scanner model in two large-scale scenario-studies. Chapter 20. In: Koomen, E., Stillwell, J., Bakema, A. and Scholten, H.J. (eds.), *Modelling land-use change; progress and applications*. Springer, Dordrecht, pp. 355-373.

Eickhout, B., Van Meijl, H., Tabeau, A. and Van Rheenen, T. (2007) Economic and ecological consequences of four European land use scenarios. *Land Use Policy* 24: 562-575.

Hilderink, H.B.M. (2004) Populations and Scenarios: Worlds to Win? RIVM report 550012001/2004. RIVM

Koomen, E., Loonen, W. and Hilferink, M. (2008) Climate-change adaptations in land-use planning; a scenario-based approach. In: Bernard, L., Friis-Christensen, A. and Pundt, H. (eds.), *The European Information Society; Taking Geoinformation Science One Step Further.* Springer, Berlin, pp. 261-282.

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

Overmars, K., Verburg, P.H., Bakker, M., Staritsky, I., Hellman, F. and Schulp, N. (2006) EURURALIS 2.0; Technical documentation CLUE-s. Wageningen University research, Wageningen.

Perez-Soba, M., Verburg, P.H., Koomen, E., Hilferink, M., Benito, P., Lesschen, J.P., Banse, M., Woltjer, G., Eickhout, B., Prins, A.-G. and Staritsky, I. (2010) Land use modelling - implementation; Preserving and enhancing the environmental benefits of "land-use services". Final report to the European Commission, DG Environment. Alterra Wageningen UR/ Geodan Next/ Object Vision/ BIOS/ LEI and PBL, Wageningen.

Pontius Jr., R.G., Boersma, W., Castella, J.-C., Clarke, K., De Nijs, T., Dietzel, C., Duan, Z., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C.D., McConnell, W., Pijanowski, B.C., Pithadia, S., Sood, A.M., Sweeney, S., Trung, T.N., Veldkamp, T.A. and Verburg, P.H. (2008) Comparing the input, output, and validation maps for several models of land change. *Annals of Regional Science* 42(1): 11-37.

van der Beek, M. (2009) User guide Ruimtescanner; version 5.64. Object Vision bv, Amsterdam.

Van Meijl, H., Van Rheenen, T., Tabeau, A. and Eickhout, B. (2006) The impact of different policy environments on agricultural land use in Europe. *Agriculture, Ecosystems & Environment* 114(1): 21-38.

Verburg, P.H., De Nijs, T.C.M., Ritsema van Eck, J., H.Visser and de Jong, K. (2004) A method to analyse neighbourhood characteristics of land use patterns. *Computers, Environment and Urban Systems* 28(6).

Verburg, P.H. and Overmars, K. (2009) Combining top-down and bottom-up dynamics in land use modeling: exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. *Landscape ecology* 24: 1167-1181.

Verburg, P.H., Rounsevell, M.D.A. and Veldkamp, A. (2006) Scenario-based studies of future land use in Europe. *Agriculture, Ecosystems & Environment* 114(1): 1-6.