Working Paper 10/07

LandSpaCES: A Design Module for Land Consolidation: Method and Application

Demetris Demetriou, John Stillwell and Linda See

School of Geography

University of Leeds

Leeds LST 9JT

September 2010

Contents

Abstract

Land fragmentation is a major problem in many rural areas and it is widely recognized that effective land administration and management is critical for sustainable rural development. Land consolidation programmes involving the reallocation of land between owners have been implemented in different parts of Europe. In this paper, we outline a land redistribution model called *LandSpaCES* (Land Spatial Consolidation Expert System), a module that forms the heart of an integrated planning and decision support system (IPDSS) and which utilises state-of-the-art methods and geotechnology tools to support and automate the entire process of land reallocation.

LandSpaCES is the design component of *LACONISS*, a LAnd CONsolidation Integrated Support System for planning and decision making. It is a knowledge-based system that generates alternative land reallocations under different scenarios. Two key system concepts are utilised: 'No-Inference Engine Theory (NIET)', which differentiates *LandSpaCES* from the conventional expert system (ES) development, and a parcel priority index (PPI), which constitutes the basic measure that defines the redistribution of land in terms of location. The module has been applied to a case study area in Cyprus and the results from the design module compare favourably against an independent solution derived previously by the human experts. The paper reports on how the system has been used to generate alternative redistributions of land based on different scenarios.

Keywords

Land; consolidation; design; redistribution; knowledge base; expert system; parcel; rules; decision tree; evaluation

List of Figures

List of Tables

1 INTRODUCTION

Land fragmentation refers to the situation in which landholdings consist of numerous spatially separated land parcels that may be small in size, irregular in shape, dispersed from one another and from the owner's farmstead (Van Dijk, 2003). Frequently, land fragmentation implies a defective land tenure structure which may be preventing efficient agricultural production and sustainable rural development. Fragmentation is commonplace in rural areas in the European Union (and across the world) and agricultural censuses show that it has been a problem in Cyprus in particular for several decades with average land holding size declining from 7.2 hectares in 1946 to 3.5 hectares in 2003 (Demetriou *et al.,* 2010a).

Land consolidation has been the most favoured land management approach for solving the land fragmentation problem and has been applied in many countries around the world. The European Union (EU) has provided support for consolidation schemes from the European Agricultural Fund for Rural Development (EAFRD), and the Food and Agriculture Organisation (FAO) has a long tradition of involvement in land consolidation activities (FAO, 2008). In Cyprus, land consolidation followed from a Land Consolidation Law enacted in 1969 and legislation now exists for land consolidation to be applied on a voluntary basis by agreement among the landowners. Seventy three projects have now been completed in Cyprus covering an area of 17,552 hectares or 11.22% of the total agricultural area enumerated in the 2003 agricultural census (Land Consolidation Department, 2010). In addition, 15 projects are currently running and 34 projects are under study.

The framework for carrying out the land reallocation process is defined by principles which are included in the Cyprus Land Consolidation Law. In addition, expert legal advice is provided by the Government's legal services to the Land Consolidation Department (LCD) regarding requests about specific issues and there are LCD Circulars and Advice notes. Furthermore, before a reallocation solution is approved by the Head of the LCD and then by the Land Consolidation Committee (LCC), the planner has to analyse the existing ownership data, define the minimum area and value limits of parcels owned by a landowner, and define what is meant by 'small', 'medium' or 'large' holdings (defined officially by the Head of the Department), which itself determines the number of parcels that should be allocated to a landowner. In addition, the planner will need to discuss the preferences for land parcels emerging from the reallocation system with each landowner so that he/she is willing to accept the eventual solution.

The existence of a large number of often complicated and sometimes fuzzy principles, together with many other criteria that should be taken into account, make the land reallocation problem extremely difficult for the planner and suggest the need to employ more sophisticated methods and tools in order to support the planner in the decision-making process. It is for these reasons that we are proposing a new framework for developing an integrated planning and decision support system (IPDSS) for land consolidation called *LACONISS* (LAnd CONsolidation Integrated Support System) that is briefly introduced in Section 2 and is outlined in more detail in Demetriou *et al.* (2010b).

The aim of this Working Paper, however, is to outline the design, development and evaluation of one part of the *LACONISS* framework. This component is the 'Design module' called *LandSpaCES* (Land Spatial Consolidation Expert System) which is a fundamental part of the system since it is capable of automatically generating a list of land parcels with approximate centroids (with ownership details attached) which are evaluated and then passed onto a land partitioning module which defines the precise boundaries of the parcels around the centroids that have been generated. This paper explains the concepts and methodologies underpinning the Design module and then reports on an application of the model using data from a previous land consolidation project in Cyprus. Section 3 of the paper contains an explanation of the basic steps involved in developing the knowledge-based system including system definition, knowledge acquisition, knowledge representation, knowledge base building and the definition of the inputs and outputs. Then, in Section 4, we consider classical system development issues such as the selection of the appropriate development tool and the description of the system architecture and interface for *LandSpaCES*. In addition, two basic concepts are outlined in this section. The first is 'No-Inference Engine Theory'

(NIET), an approach that differentiates *LandSpaCES* from a conventional expert system (ES); the second is the parcel priority index (PPI), which constitutes the basic measure that defines the redistribution of land in terms of location.

Following the explanation of the methods and the structure of *LandSpaCES,* the paper provides an introduction to the case study in Cyprus in Section 5 to which the system has been applied. The results produced by the design module are compared with those generated independently for the case study area by human experts in Section 6. In addition, the system was run with ten different scenarios to demonstrate how the land reallocation solutions will vary according to the values of the input parameters or 'facts' as they are called in the system. Finally, some conclusions are presented in Section 7.

2 *LACONISS*: A FRAMEWORK FOR LAND CONSOLIDATION PLANNING

The operational framework of *LACONISS* is illustrated in Figure 1. Three sub-systems are required in order to produce a land reallocation plan for a case study area that satisfies all the constraints, including the demands of the landowners and legislation. The first sub-system, *LandFragmentS* (Land Fragmentation System), represents the 'Intelligence phase' of the process and involves building an appropriate geographical information system (GIS) model for the case study area and scanning the current land tenure system to measure the extent of land fragmentation. The second sub-system, *LandSpaCES* (Land Spatial Consolidation Expert System), contains (i) a design module that integrates the GIS with an expert system (ES) and generates alternative land redistributions, and (ii) an evaluation module that uses the GIS and multi-attribute decision-making (MADM) methods to evaluate the alternative distributions. The final output of *LandSpaCES* is a map showing the centroids of parcels with their land value and ownership attributes which are then transferred to the third sub-system, *LandParcelS* (Land Parcelling System), in order to create the optimum set of boundaries for the land parcels around each of the centroids by integrating GIS with a genetic algorithm (GA) and multi-objective decision-making (MODM) methods.

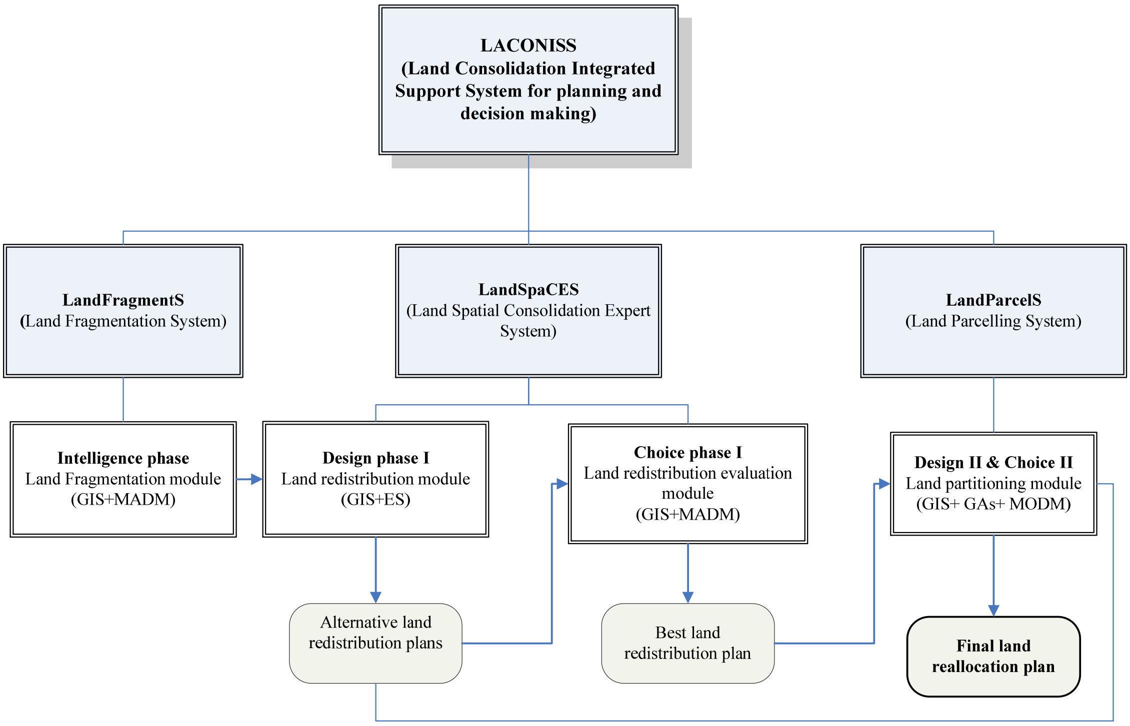


Figure 1: The operational framework of *LACONISS*

Thus, ‘Design phase I’ and ‘Choice phase I’ in Figure 1 refer to the automated generation of alternative parcel centroids (points) and their evaluation, whereas ‘Design and Choice II’ is a dual land partitioning phase that results in the final reallocation plan of land parcels (polygons). The focus of this paper is the theory and methodology underpinning the ‘Design phase 1’ module.

3 SYSTEM DESIGN

The design of a knowledge-based system involves certain tasks other than those followed in conventional software systems. This differentiation is due to the knowledge component that is embedded in the former type of system. More specifically, conventional computer programs perform tasks using conventional decision-making logic which contains little knowledge other than the basic algorithm for solving that specific problem and the necessary boundary conditions. On the other hand, knowledge-based systems collect together the small fragments of human know-how into a knowledge base, which is then used to reason through a problem, searching for a solution. This task, i.e. the design of a so-called ‘expert system’ (ES), usually consists of the following steps: system definition, knowledge acquisition, knowledge representation, knowledge-base building and the definition of the inputs and outputs. We consider each of these steps in turn in this section of the paper.

3.1 System definition

The objectives of *LandSpaCES* are to:

- automate the process of land redistribution so as to generate a complete problem solution;
- be used as a decision support tool by generating alternative land distributions;
- enhance the land redistribution process by structuring it in a systematic, standardised and transparent way using an appropriate model;
- considerably diminish the time needed by a human expert to carry out the land redistribution process; and to
- be capable of evaluating a set of alternative land redistributions using multi-attribute decision-making (MADM) methods.

The first four objectives refer to the *Design* module and the fifth to the *Evaluation* module as indicated in the *LACONISS* framework (Figure 1). In addition to achieving the objectives set out above, *LandSpaCES* may also operate as a 'what-if?' model in terms of producing alternative land redistributions using different facts (scenarios). However, LandSpaCES cannot assess the benefits in agriculture (e.g. in production, productivity, *et cetera*) as a result of the new land tenure structure achieved via land consolidation. For this purpose other models are used that may be fed by *LandSpaCES* outputs. The results can be used as inputs to the *ex-ante* evaluation of a land consolidation project based on the EU requirements and the system may also be used as a trainee tool for new and expert land consolidation technicians to understand and analyse the reasoning process underpinning land distribution.

3.2 Knowledge acquisition

Liou (1998) defines knowledge acquisition as *"the process of extracting, structuring and organizing knowledge from several knowledge sources, usually human experts, so that the problem solving expertise can be captured and transformed into computer readable form"* (p. 2-1). A variety of knowledge acquisition methods have been suggested by various authors (Breuker and Wielinga, 1983; Grover, 1983; Hart, 1986; Kidd, 1987) which tend to be a combination of approaches including: documentation (e.g. manuals, guidelines and legislation); studying past projects and their subsequent shortcomings; discussing cases with experts via personal or collaborative interviews; and observing experts applying their knowledge to current problems.

The selection of an appropriate knowledge acquisition method depends on many factors, namely the problem domain, the availability of knowledge resources and the time/cost constraints. In this research, knowledge is collected through the following resources: the first author's thirteen-year personal experience of working on land consolidation projects; informal discussions/interviews with land consolidation technicians (i.e. the prospective main users of the system) who carry out the land consolidation process over two or three decades; documentation, such as Land Consolidation Law, formal LCD guidelines and instructions, legal advice, etc., and analysis of the solution given by experts in the case study used in this research.

4.3 Knowledge representation

In order to model the land redistribution decision-making process, the problem has been split into seven sub-problems represented by decision trees to facilitate the analysis and make the decision trees effective and understandable through visual representation. One is the main decision tree illustrated in Figure 2. It represents the flow of the general decisions taken by the experts regarding land redistribution. For decisions, landowners are classified into six groups depending on the original number of parcels owned and the maximum number of parcels that may be allocated to each landowner in the new plan as shown below:

- the landowner originally has 1 parcel and he/she may receive only 1 parcel in the new plan (B1);
- the landowner originally has more than 1 parcel and he/she may receive only 1 parcel in the new plan (B2);
- the landowner originally has 2 parcels and he/she may receive up to 2 parcels in the new plan (C1);

- the landowner originally has more than 2 parcels and he/she may receive up to 2 parcels in the new plan (C2);
- the landowner originally has 3 parcels and he/she may receive up to 3 parcels in the new plan (D1); and
- the landowner originally has more than 3 parcels and he/she may receive up to 3 parcels in the new plan (D2).

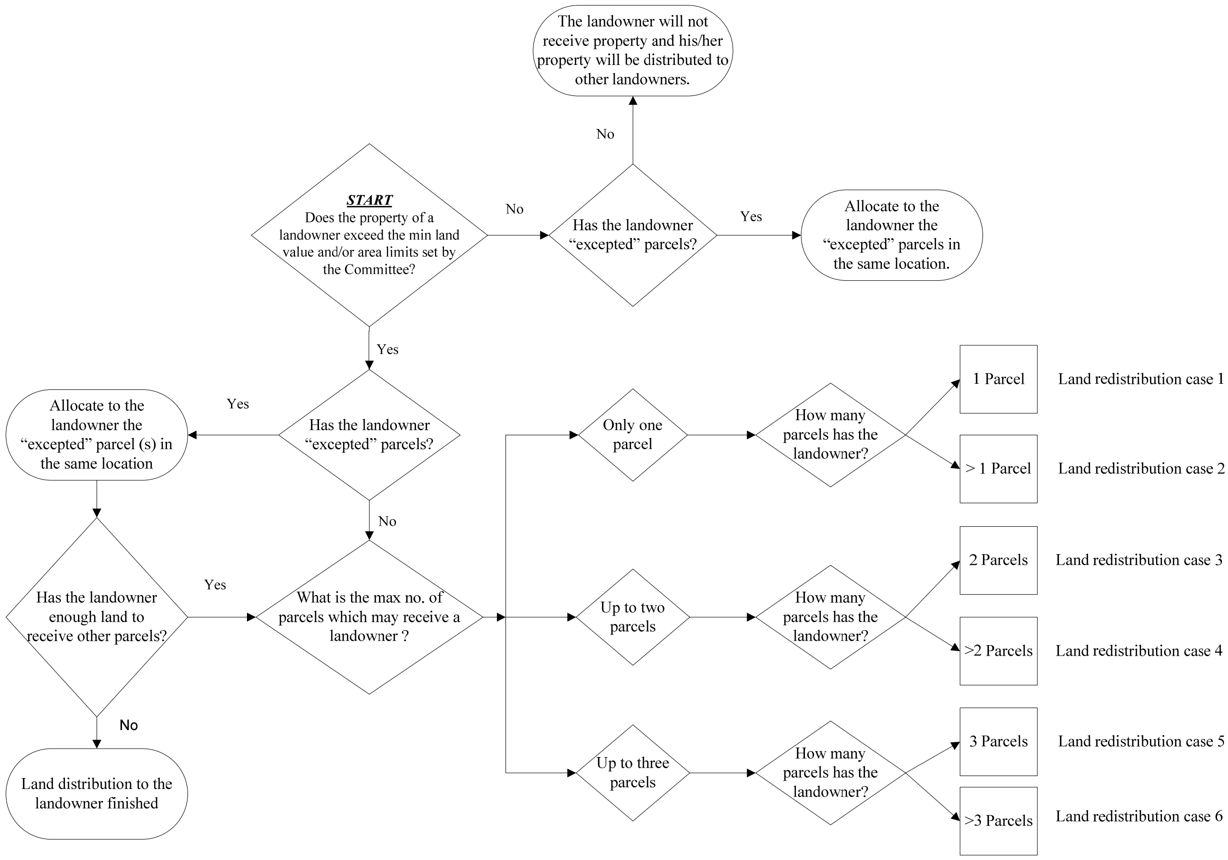


Figure 2: The main decision tree for the Design module of *LandSpaCES*

Then for each group shown above a separate decision tree has been constructed which represents a certain land redistribution case. Each tree is composed of nodes, branches and leaves. A node is represented by a diamond-shaped decision symbol containing a question which is the premise of a rule. Branches are either 'Yes' or 'No' answers to each question and leaves contain conclusions. The question-answer-conclusion chain may be split into a number of relevant sub-sets. This form of tree is called a binary tree.

The most difficult task is to formulate a process that is able to synthesise the expertise of many experts and which coincidently will be systematic, standardised and transparent. These decision trees satisfactorily represent the land redistribution decision-making process for the majority of cases and the majority of experts' knowledge. However, it is impossible to model all the land distribution conditions. Some of them may be unique for a certain project. For this reason, the expert needs to intervene accordingly to adjust the land redistribution and hence have the final say, at the planning stage, for the decisions taken.

4.4 Knowledge base building

A knowledge base contains a collection of IF-THEN rules which have been extracted from each decision tree. Decision trees and their rules follow the suggestion that the appropriate inference approach (control process) for solving the land redistribution problem is 'forward chaining' (i.e. data driven); hence rules search the solution from data and input parameters to conclusions. Rules are grouped into ten 'rule clusters', which represent the decision-making process for a certain land redistribution sub-problem. Each rule cluster is transformed to a computer language routine so as to incorporate rules and other necessary operations to solve a certain problem. Two examples of rule clusters, one for the main flowchart and one for the flowchart B2, containing rules in plain language, are presented in Tables 1 and 2 respectively.

Table 1: Rules in plain language for the main flowchart rule cluster

Rule No.	IF	THEN
1	The total area **OR** value of a landowner's property is **<** than the corresponding minimum completion limits set by the Committee **AND** the examined parcel is not "excluded"	The landowner will not receive any parcel in the new plan **AND** he will receive expiate the land value of the property **AND** the property will be available to be distributed to others
2	The total area **OR** value of a landowner's property is >= than the corresponding minimum completion limits set by the Committee **AND** the parcel examined is not "exclusive" AND the landowner has not applied to be completed	The landowner will not receive any parcel in the new plan **AND** he will receive expiate the land value of the property **AND** the property will be available to be distributed to others
3	The total area **OR** value of a landowner's property is **>=** than the corresponding minimum completion limits set by the Committee **AND** the total area is >= minimum area limit set by the Law	The landowner will receive property in the new plan
4	The total area of a landowner's property is **>=** than the corresponding minimum completion limit set by the Committee **AND** the total area is <= minimum area limit set by the Law and the landowner has applied "to be completed"	The landowner will receive property in the new plan
5	The total value of a landowner's property is **>=** than the corresponding minimum completion limit set by the Committee **AND** the total area is <= minimum area limit set by the Law and the landowner has applied "to be completed"	The landowner will receive property in the new plan
6	The total area of a landowner's property is **>=** than the corresponding minimum completion limit set by the Committee **AND** the total area is <= minimum area limit set by the Law and the landowner has not applied "to be completed"	The landowner will not receive any parcel in the new plan **AND** he will receive expiate the land value of the property **AND** the property will be available to be distributed to others
7	The total value of a landowner's property is **>=** than the corresponding minimum completion limit set by the Committee **AND** the total area is <= minimum area limit set by the Law and the landowner has applied not "to be completed"	The landowner will not receive any parcel in the new plan **AND** he will receive expiate the land value of the property **AND** the property will be available to be distributed to others
8	A landowner will not receive property in the new plan	His property will be available to be distributed to others

Table 2: Rules in plain language for the B2 flowchart rule cluster

Rule No.	IF	THEN
1	The new area will be allocated to a landowner is **<** than the minimum area limit set by the Law	The new area will be equal to the minimum area limit set by the Law
2	The area of the new parcel is <= available land in the block of that parcel	Create the new parcel
3	The number of parcels already allocated to a landowner = the maximum number of parcels may received by the certain landowner	Do not allocate him any other parcels
4	The area of the new parcel is > than the available area of the block of the parcel **AND** the Parcel Priority Index is > than the minimum of the parcels of that block	Search and allocate the examined parcel in that block and “move” the parcel(s) with less PPI in another block
5	Rule 4 can not be satisfied	Search and allocate the examined parcel in another block in which the landowner posses a parcel
6	Rule 5 can not be satisfied	Allocate the new parcel in none block to decide the user

Rule clustering is important in the design and the efficiency of the system. The entire knowledge base consists of 74 rules, of which 22 are *generic* and are included only once in the main rule cluster, and 38 are *specific* and are included in more than one rule cluster. Generic rules have a broad scope regarding decisions taken, e.g. a landowner may or may not receive property in the new plan or what is the maximum number of parcels that can be allocated to each landowner? Specific rules focus on decisions taken for each land distribution case, e.g. create a new parcel or not? Displace a new parcel in another location or not? Examples of each type of rule are as follows:

Generic rule

IF [The total area **OR** value of a landowner's property **<** the minimum completion limits set by the Committee **AND** the examined parcel is not "exempt"] **THEN** [The landowner will not receive any parcel in the new plan **AND** he/she will receive as pecuniary compensation the land value of the property **AND** the property will be available for distribution to other landowners]

Specific rule

IF [The area of the new parcel **≥** the minimum parcel area set by the Committee **AND** the area of the new parcel **<** the minimum area set by the Law] **THEN** [Set the new area equal to the minimum limit set by the Law **AND** create the new parcel]

Compared to an ES with hundreds or even thousands of rules, the knowledge base of *LandSpaCES* is small. However, most rules involve complex spatial operations and some may call procedural subroutines.

3.5 Definition of the inputs and outputs

The system requires two kinds of inputs: GIS data and 'facts' which may be constraints or weights. GIS data, which are stored in the long-term memory, involve a digitized cadastral map (a feature layer) and database tables underlying the map. These data refer to parcels, landowners, ownership, etc. Facts, which are stored in the working memory, are the inputs that are used by the rules to infer new parameters or conclusions or actions. The necessary input facts for this system are the following:

- the minimum new parcel area limits according to the law for this land consolidation area;
- the minimum area and land value limits set by the Land Consolidation Committee in order to receive a landowner property in the new plan;
- the area limits for small-medium-large property sizes;
- the weights for the parcel priority index (PPI) explained later; and
- the minimum area limit to create a new parcel (for landowners that may receive more than one parcel).

These inputs are decision variables that may result in alternative land redistribution solutions.

In summary, *LandSpaCES* is a knowledge based system, driven by a set of rules, which generates outputs that include the following:

- a list of the new parcels including an ID number, area, land value, the land block in which the parcel is located and the approximate centroid location in X,Y coordinates;
- a list of the landowners of the new parcels; and
- a map showing the approximate location (represented by a centroid point) of the new parcels with the parcel ID and landowner ID.

In the next section, we consider the tools that are used to implement *LandSpaCES,* explain the system architecture and illustrate the graphical user interface.

4 SYSTEM DEVELOPMENT

4.1 Development tools

The available software for developing an ES can be divided into three main categories: expert system shells (e.g. EXSYS or XpertRule), high level or artificial intelligence (AI) programming languages (e.g. LISP, OPS5 or PROLOG) and knowledge engineering tools (e.g. CLIPS or KEE). The selection of the most appropriate and efficient tool for developing an ES is a crucial matter. Thus, a literature search was carried out for this purpose. In particular, the search focused on tools that have been used for integration with GIS and/or on well-known tools in the ES industry and academia. Based on this search, the following tools were given further consideration: Visual Rule Studio, EXSYS, XpertRule and CLIPS. A search of the Internet was carried out to obtain more information about these tools.

Visual Rule Studio is an object-oriented COM-compliant ES development environment for Visual Basic which solves the problem of software interoperability by allowing developers to package rules into component reusable objects that are used to encapsulate knowledge structure, procedures and values. Visual Rule Studio is used in the development of some GIS applications (Eldrandaly, 2003; 2006; Chau, 2003) and is easily integrated with ArcGIS via VBA and ArcObjects. This tool fulfils all the criteria to be an effective tool for this research but its promotion has been terminated.

EXSYS and XpertRule are well-known ES development tools with graphical environments. Even though they have been used in many standalone business applications, there is no evidence that they have been used for GIS applications. In addition, they are too expensive (i.e. more than £8,000) for non-funded research. On the other hand, CLIPS is a very well-known freeware ES language developed by NASA. It provides a complete and powerful environment for the construction of a rule and/or object based ES. It is maintained independently from NASA as public domain software and has been used in thousands of applications since 1985. In addition, it has been used for integration with GIS in some studies (Paliulionis, 2000; Jin *et al*., 2006). Its disadvantages include the learning time and some weaknesses in integration with GIS (Jin *et al*., 2006).

Although the great majority of ES applications have been developed using ES development tools, there are cases, either standalone or hybrid systems, which have been developed using procedural languages. Recent research carried out by Hicks (2007) discusses the 'No-Inference Engine Theory' (NIET) and argues that this method outperforms traditional ES development tools. Hicks (2007) developed a system called EZ-Xpert which can produce code for ES development based on the NIET. The code can be generated in popular procedural languages (C, Java and Visual Basic) and it can be embedded in an application. Hicks (2007) claims that his method can test over 20,000 rules per second. Other authors (Wilson, 2007; Griffin and Lewis, 1989) have developed ES using a conventional language.

Moreover, some researchers have integrated ES and GIS using conventional languages or GIS languages such as Avenue, VBA and ArcObjects (Vlado, 2002; Howard, 2003; Choi and Usery, 2004; Jin *et al.*, 2006). Such geotechnology tools have been widely used in the development of planning support systems (Geertman and Stillwell, 2000). Although conventional languages have been rarely employed for developing ES, they may be appropriate for a hybrid ES with a small knowledge base and a straightforward reasoning engine (Choi, 2002).

Thus, how do we select an appropriate ES tool? On the one hand, our conclusion from the above search is that there is a lack of specialised ES development tools capable of easily integrating ES into proprietary GIS, despite the fact that knowledge and expertise are important components of all spatial planning processes. On the other hand, undoubtedly the easiest and most efficient way to develop an ES is to use an ES shell, then use a knowledge engineering tool and then an AI language. However, these specialised ES development tools are efficient for developing standalone ES applications and not hybrid systems. The ES development tools have limitations in their flexibility and in their communication with the non-ES application, e.g. the GIS (Jin *et al*., 2006). In addition, these tools permit only loose or tight coupling integration architecture and they cannot create fully integrated systems. A solution to this problem is to use a conventional programming language that provides greater development flexibility even though it is a time-consuming task for developing an ES from scratch (Lukasheh *et al*., 2001). A comparison of the advantages and disadvantages of specialised ES development tools and conventional GIS programming tools (e.g. VBA and ArcObjects) for developing integrated GIS-ES is shown in Table 3.

Table 3: Advantages and disadvantages of specialized ES development tools and conventional GIS programming tools

Advantages	Disadvantages
ES development tools	
Ease of ES development	Time needed to learn something new
Provides all the components of a typical ES (i.e knowledge acquisition module, explanation facility, etc.)	Cost
	Programming limitations
	Impossible to have full integration with other systems
	Slower hybrid systems
	A lot of development tools have been retracted
	Some tools have limited support
Conventional GIS programming tools	
Flexibility of programming	Time consuming
Capability for fully integrated hybrid systems	Difficult to develop all the facilities of an ES (i.e explanation facility, editable knowledge base, e
System development only in one environment	
More familiar to GIS experts	
Creation of faster hybrid systems	

The above review suggests that it is better to sacrifice some of the advantages provided by ES development tools to gain the advantages offered by conventional programming languages under a common GIS development environment, that is VBA and ArcObjects in ArcGIS, to fully integrate GIS with ES techniques. However, this integration needs to be based, in terms of development, on robust theoretical foundations. This requirement is provided by the 'No-Inference Engine Theory'.

4.2 No-Inference Engine Theory

Giarratano and Riley (2005) have made some important remarks regarding the design of an ES. They emphasise that when dealing with ill-structured problems, it is possible to reach a solution using an algorithm, i.e. a step-by-step process. Such a case involves a rigid control structure, i.e. the rules are forced to execute in a certain sequence, i.e. every rule fires another rule and so on. In such a case, a major advantage of an ES is negated, i.e. dealing with unexpected inputs that do not follow a predetermined pattern and hence there is no need for an inference engine and an ES, because a procedural programming language is a more appropriate approach. Ideally, an ES is a balanced mix of strongly and weakly coupled rules, just as humans may use deductive, inductive, probabilistic and other methods to reach a conclusion.

In contrast to the above remarks, Hicks (2007) suggests The ‘No-Inference Engine Theory’ (NIET). Conventionally, the major task performed by an inference engine is conflict resolution which determines the sequence of consultation based on a predefined priority assignment in a rule, when multiple rules are activated simultaneously. However, a disadvantage of this strategy is that most of the execution time of a program is spent in performing this process. In NIET, conflict resolution is performed during the development stage, by ordering rules in a logical flow sequence. In addition, in the case of rules that present the so called ‘bartender problem’ (i.e. two or more rules have at least the first condition of their premise part in common), the conflict resolution is solved by sequencing first the most specific rules, i.e. rules with the greater number of conditions, so they can be tested first. This is a rule ordering strategy or more precisely a conflict resolution strategy which is common in the ES industry and is the default for most products. This strategy is also recommended in the manuals of some ES shell/tools such as CLIPS and M.4 (Hicks, 2007), in which the user assigns a priority value for a rule (e.g. in CLIPS this is referred to as salience).

The central concept of this approach is that it transforms the traditional declarative inference engine into a procedural solution involving a sequence of IF-THEN statements with other conventional programming operations and inputs to instantiate conditions. Inference is still performed but the components are distributed between the development and the delivery environments. In other words, in contrast to a typical ES, a rigid control structure and strong coupled rules are the characteristics of NIET-based systems in which all the inference components are performed at run-time, where a NIET system performs only inputs and logic at run-time and not conflict resolution. As noted earlier, Hicks (2007) demonstrated, using EZ-Xpert, that this simplification of the run-time task results in significant performance improvements. Actually, the concept of NIET has also been employed by other researchers (Wilson, 2007; Griffin and Lewis, 1989).

Another significant distinguishing feature of NIET is that the knowledge base and inference engine are not kept separated as in conventional ES. In NIET, there is no inference engine and the reasoning (rule base) and process (inference engine) are combined into a single unit.

Separation of the two basic components of an ES constitutes an advantage in terms of ease and efficiency in editing the knowledge base. However, this advantage is not such an important matter for *LandSpaCES* since the number of rules used (i.e. 74) is small and it is not expected that the knowledge base will be updated very frequently, given that modifications to legislation and new circulars regarding the process are rare. In addition, as noted earlier, the prominent aim of *LandSpaCES* is to demonstrate that it works satisfactorily (using ES techniques and GIS functions) in terms of performance regarding the solutions generated and not to produce an operationally perfect ES with all the facilities (e.g. editing rules, explanation facility) that are usually provided in a standalone system.

The NIET has been adopted as the development methodology of the *Design* module of *LandSpaCES* since it will be coded using a procedural language, i.e. VBA and ArcObjects. In addition to this conflict resolution strategy, the fact that rules are grouped in ten rule clusters which correspond to a part of or an entire decision tree, improves efficiency.

4.3 System architecture

The overall architecture of *LandSpaCES* is illustrated in Figure 3. Similar to other knowledge-based GIS (KBGIS), *LandSpaCES* consists of three main parts:

- a database management system (DBMS) which contains all the necessary information about a land consolidation plan. Basic database management functions can be carried out in ArcView;
- a land distribution expert system mechanism (ES) which contains the rule base (reasoning) and the process (inference engine). This is the heart of the system which makes land redistribution decision making; and
- a graphical user interface (GUI) through which the user can input data and facts, run the land redistribution process and display the system outputs.

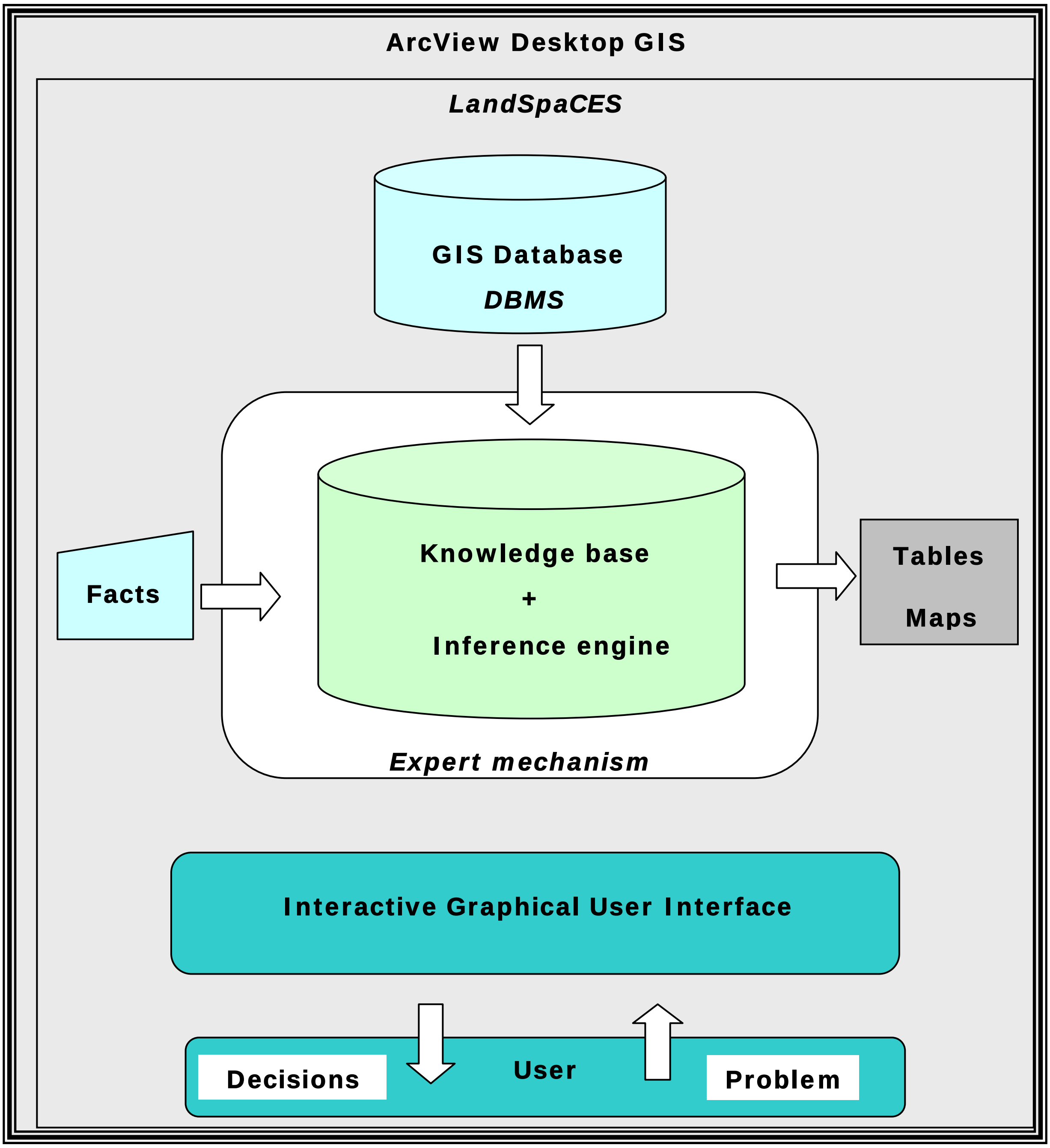


Figure 3: The overall architecture of *LandSpaCES* Design module

The methodology followed for code development was *top down design*, which is an approach successfully used by professional programmers. Top down design involves dividing a complex problem into smaller, less complex sub-problems based on the idea that the more general tasks occur near the top of the design and tasks representing their refinement occur below (Schneider, 2004). *LandSpaCES* consists of 10 Modules that contain 63 Procedures and

Functions and 10 Rule Cluster Procedures. The basic VBA and ArcObjects code for the basic flowchart and the flowchart for B2 rule clusters, which correspond to the rules shown in Tables 4 and 5 respectively, are presented in Appendix B. The hierarchy chart of *LandSpaCES* is shown in Figure 4 indicating the structure of the system and the work flow to produce a land distribution solution.

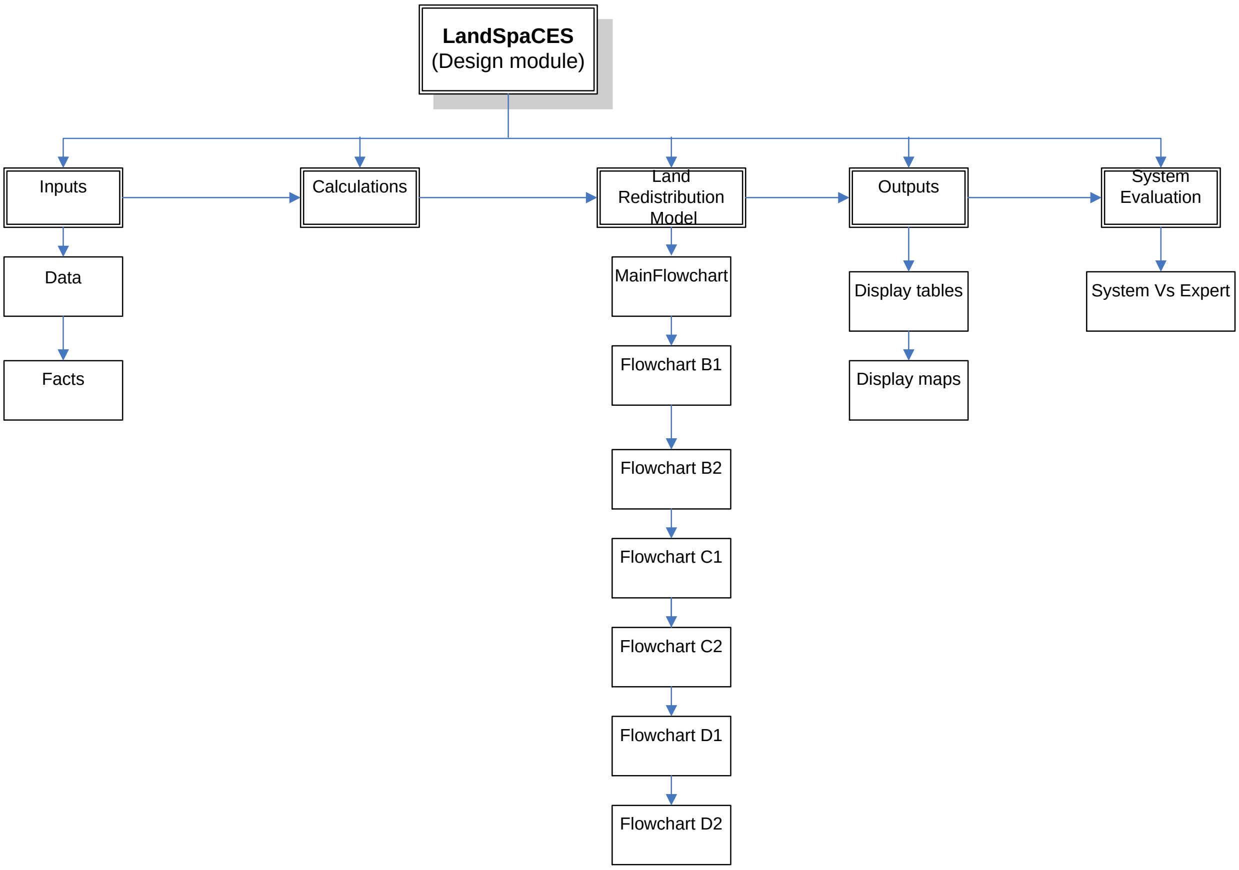


Figure 4: The hierarchy chart for the *LandSpaCES* Design module

4.4 Interface

The user interface of *LandSpaCES* is a toolbar in ArcView named 'LandSpaCES: Land Spatial Consolidation Expert System' and consists of seven icons as shown in Figure 5.

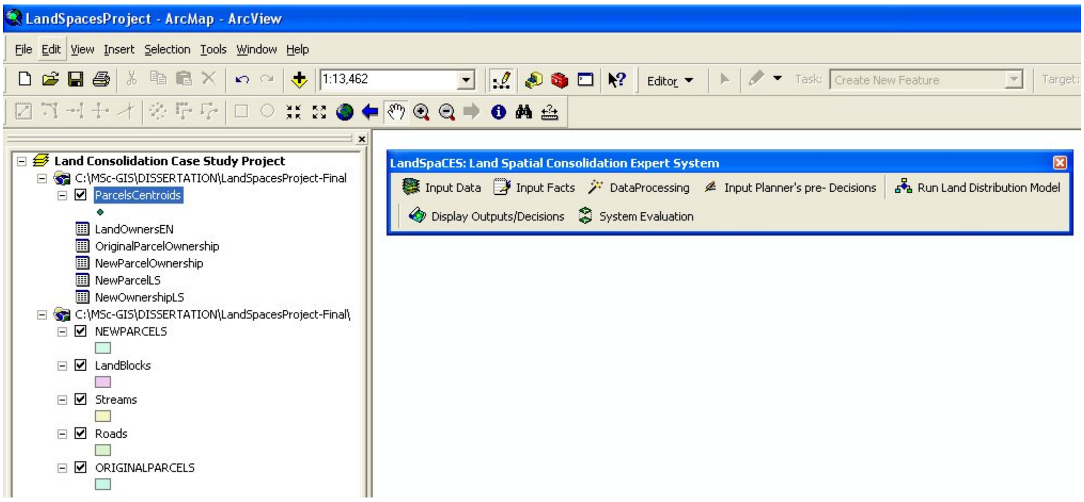


Figure 5: The *LandSpaCES* toolbar in the ArcView environment

The planner's pre-decisions, which may be input to the process, will not be developed in the context of this research. Each icon launches a separate window with a title, text boxes and command buttons. The window of each icon is shown in Figures 6 to 10. The *Input Data* window (Figure 6) is used to input the shapefiles and databases provided/created for the case study. The *Input Facts* window (Figure 7) is used to input the problem's basic facts (F1-F8). The *Run Land Redistribution Model* window (Figure 8) is used to run the land redistribution model assuming that the previous windows have been used appropriately. The planner also has to define three facts (F9-F11), namely, two weights for the PPI calculation and a minimum area limit for the creation of a new parcel for those landowners who may receive more than one parcel. The *Display Outputs/Decisions* window (Figure 9) presents the system results as two output database tables, the output map and the attribute table of the map. The *System Evaluation* window (Figure 10) executes various calculations used for system validation, i.e. the comparison of the results from the system and the human expert.

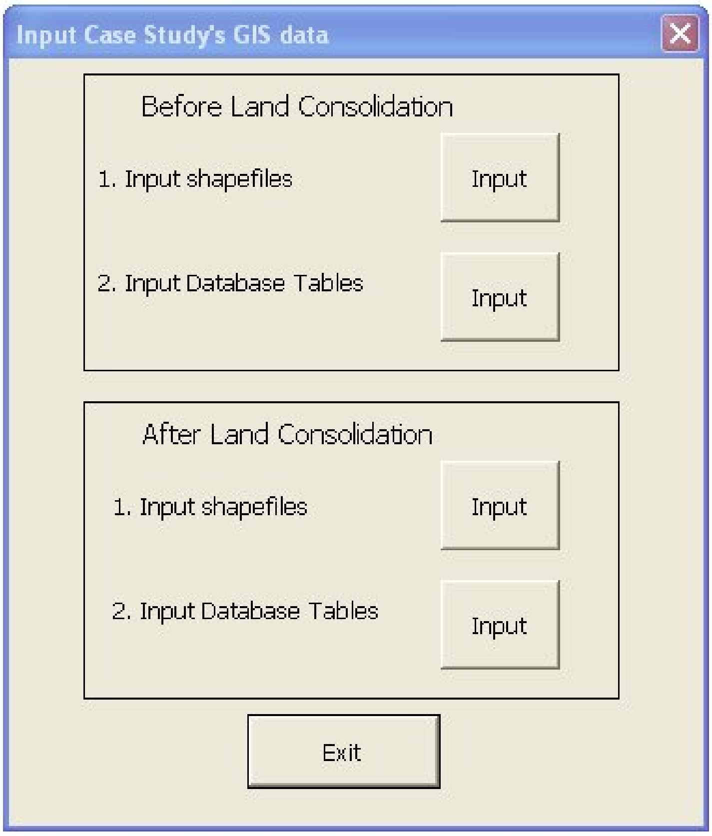


Figure 6: The Input Case Study's GIS data window

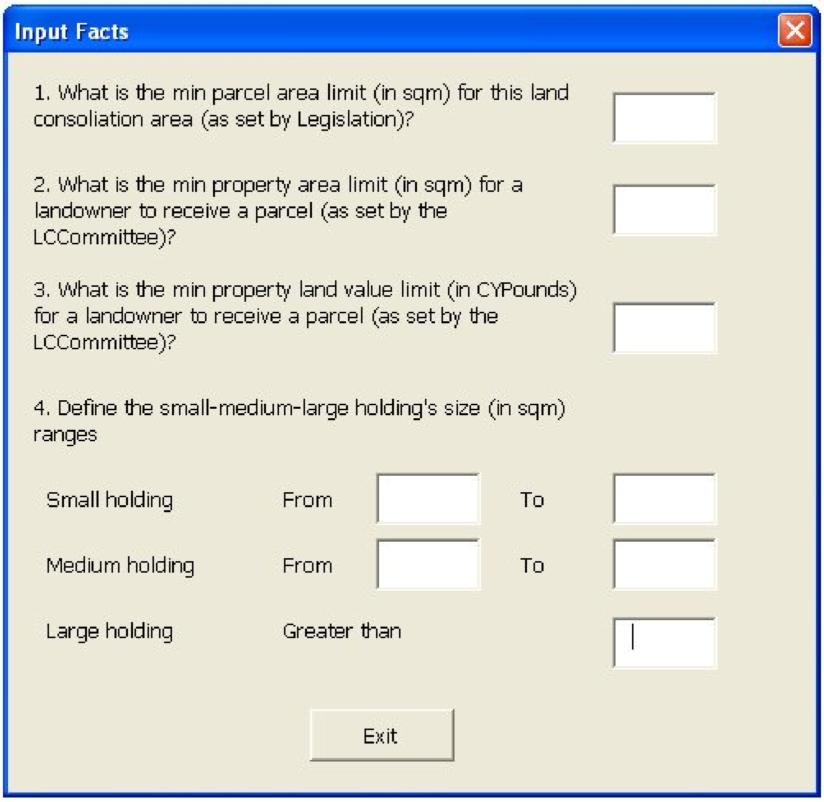


Figure 7: The Input Facts window

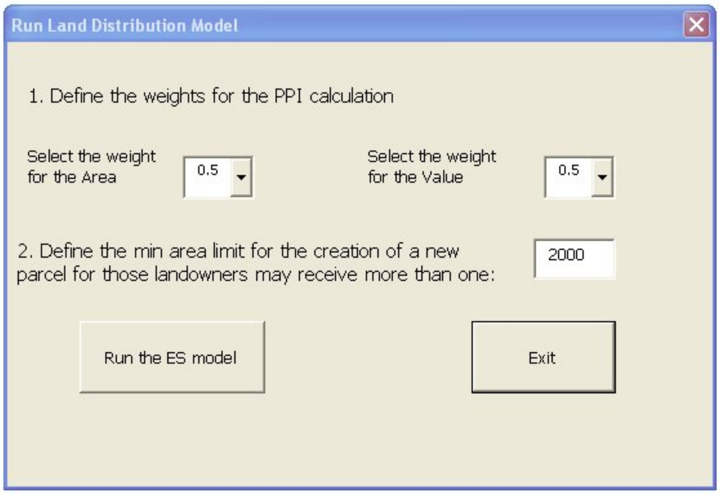


Figure 8: The Run Land Distribution Model window

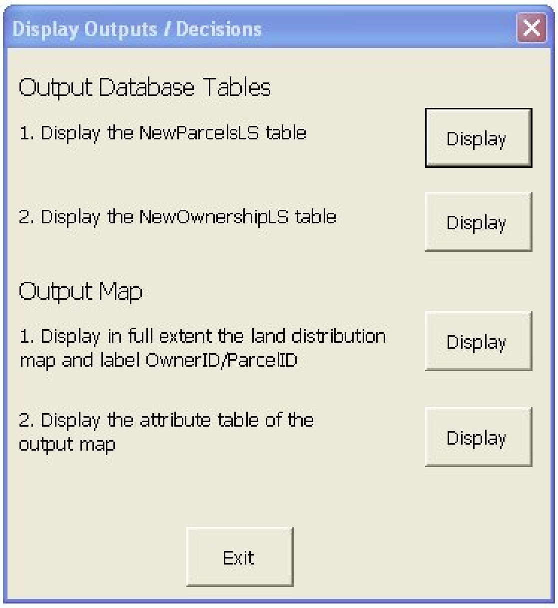


Figure 9: The Display Outputs/Decisions window

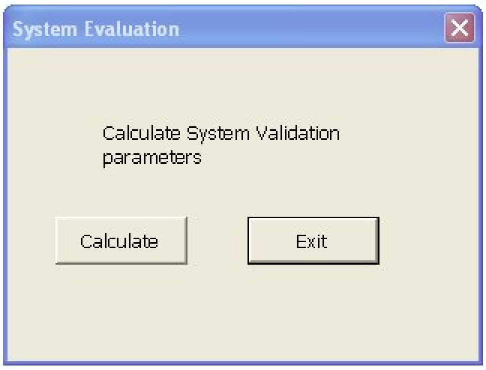


Figure 10: The System Evaluation window

4.5 The parcel priority index (PPI)

One of the crucial matters to be considered when building the land redistribution model is the way in which the preferences of the landowners are incorporated. It is accepted that the most important concern for landowners in a land consolidation project is the location of the new parcels which they will receive. It is also well known by land consolidation planners that each landowner wishes to receive his/her property in the location of his/her 'best parcel', then to the next 'best parcel' and so on (Sonnenberg, 1998; Cay *et al*., 2006; Ayranci, 2007). Thus, the question focuses on what parcel is considered to be the 'best' for a landowner. Practice has shown that the 'best parcel' is perceived as that with the largest area and/or the highest land value per hectare or a combination of these two factors. A parcel with these characteristics may contain the farm buildings as well. Area and land value (either market price or agronomic value) are the two fundamental factors commonly used as the basic land reallocation criteria. Consequently, a measure that takes into account these two factors may satisfactorily represent the preference of a landowner in terms of the location he/she wishes to receive his/her new parcels. For these reasons, a measure called the parcel priority index (PPI) has been introduced in this research and will operate as a factor for:

- representing the preference of landowners by setting in rank order all the parcels of a project and then separately the parcels of each landowner defining preferences regarding the location they wish to receive the new parcels;
- representing the priority of a landowner-parcel pair in the land distribution process in terms of allocating a parcel in a certain location or not. In other words, the higher the PPI the higher the priority, hence the higher possibilities for a landowner to receive his property in the desired location(s). Thus, PPI is a crucial factor that determines if a parcel (and hence its landowner) will be 'displaced', during land distribution, to another location so as to satisfy the first, second etc. preference of a landowner; and
- ensuring equity, transparency and standardisation of the process since the redistribution of parcels in a new location is based exclusively on this factor.

The power of the PPI is that it treats the two basic entities of reallocation, i.e. landowners and parcels, as one common entity, the parcel-landowner. In particular, the PPI combines the characteristics of the parcel-landowner entity and defines the entity's priority for reallocation into a single number between 0 and 1. Currently, it is calculated based on two factors: the area and the land value of the parcel. Initially, the PPI is calculated separately for each parcel based on a linear scale transformation that takes into account the minimum, maximum and mean values of the relevant dataset. The transformation is based on the assumption that the minimum, maximum and mean values of a dataset correspond to scores of 0, 1 and 0.5, respectively. The mean is involved in the transformation in order to avoid great variations between the PPI values when extreme values are present. Thereafter, the overall PPI is calculated based on the relevant weight assigned by the user for each factor. Weights represent the importance of each factor in the land distribution process and influence the location of the new parcels. The relevant formulas are given below:

PPI for land area

$$PPIA_i = \frac{(A_i - MinA) * 0.5}{MeanA - MinA} \quad \text{(if } A_i \leq MeanA\text{)} \tag{1}$$

$$PPIA_i = \frac{(A_i - MeanA) * 0.5}{MaxA - MeanA} + 0.5 \quad \text{(if } A_i > MeanA\text{)} \tag{2}$$

where $PPIA_i$ is the PPI based on the area of parcel i, A_i is the area of parcel i, and $MinA, MaxA, MeanA$ are the corresponding area values for all the parcels in the dataset.

PPI for land value

$$PPIV_i = \frac{(V_i - MinV) * 0.5}{MeanV - MinV} \quad \text{(if } V_i \leq MeanV\text{)} \tag{3}$$

$$PPIV_i = \frac{(V_i - MeanV) * 0.5}{MaxV - MeanV} + 0.5 \quad \text{(if } V_i > MeanV\text{)} \tag{4}$$

where $PPIV_i$ is the PPI based on the land value of parcel i, V_i is the land value of parcel i, and $MinV, MaxV, MeanV$ are the corresponding land values for all the parcels in the dataset.

Overall PPI

$$PPI_i = WA * PPIA_i + WV * PPIV_i \tag{5}$$

where PPI_i is the overall PPI for parcel i, and WA, WV are the weights for area and land value respectively that should sum to 1.

Ideally, the PPI may take into account more reallocation criteria (than parcel area and land value) regarding a parcel and a landowner. For example, other criteria that may be used for parcels are land use and morphology. Similar to parcels, the characteristics of a landowner might be profession (farmer or not farmer, full-time or part-time farmer etc.), age, residence (in the village or not), number of children and so on. Then, the overall PPI may be calculated in a similar manner.

Based on the above, the PPI is the crucial factor for the land redistribution process since it determines the priority of a parcel-landowner entity in the reallocation. Thus, a parcel-landowner cannot ensure a location in the new plan (and they may be 'displaced' at any time during the process) until the land redistribution process has terminated. In the next section, *LandSpaCES* is applied to the case study area and the results are evaluated.

5 CASE STUDY IN CYPRUS

5.1 Selection of an appropriate project

A number of criteria were used to assist the selection of an appropriate case study. First, it was necessary to identify a land consolidation project where there was enough data available to cover the range of different land reallocation cases but not so much data as to prevent reasonably quick processing and testing of the prototype system during its development and evaluation. Second, the land consolidation project should be a typical project in terms of applying legislation and practices followed by experts. In other words, the case study should ideally reflect the problem domain that the system will solve. A good way of meeting this criterion is through comparison of the situation before and after implementation of the land consolidation plan. Better results from the process should indicate better application of the legislation and planning practice (Table 4). This criterion is important for the evaluation of the expert system. Third, the available data should be in computerised form as far as possible so as to avoid lengthy data encoding. It is noted that the LCD still does not use a complete GIS; it uses an old database system in combination with CAD and surveying software systems. Finally, the project should be carried out by a team of land consolidation experts in which the researchers had no involvement. Thus, the relevant land reallocation solution used for comparison with the model results will be completely independent. This criterion is also important for the development of the expert system. A search on the completed land consolidation projects showed that the *'Chlorakas'* project fulfils the above requirements; hence it was selected as the case study.

Table 4: Comparative statistics before and after land consolidation (LCD, 1993)

Item	Before	After	(%) increase or decrease
Aggregate number of owners	278	204	- 26,6
Local residents	198	159	- 19,7
Residents of neighbouring villages and of Pafos town	60	34	- 43,3
Residents of other towns	10	5	- 50,0
Residents of distant villages	5	3	- 40,0
Overseas residents	5	3	- 40,0
Unknown residence	0	0	--
Total area (ђa)	195,0	187,8	- 3,7
Area held in whole ownership (ha)	161,0	187,5	+ 16,4
% of area held in whole ownership	82,6	99,8	+ 20,8
Area held in undivided shares	34,0	0,3	- 99,1
% of area held in undivided shares	17,4	0,16	- 99,1
Average size of ownership (ha)	0,70	0,92	+ 31,4
Number of plots or shares	436	27,0	- 93,8
Number of plots	347	268	- 22,8
No. of plots held in whole ownership	302	266	- 11,9
Number of plots held in shares	45	2	- 95,5
Number of shares	134	4	- 97,0
% of plots held in shares	13,0	1,5	- 88,5
Average no. of plots or shares per owner	1,6	1,3	- 18,8
Average size of plot / share (ha)	0,4	0,7	+ 75,0
Length of roads (km)	4,8	12,7	+ 164,6
Area served by roads (ha)	110,0	187,5	+ 70,4
% of area served by roads	56,4	99,8	+ 77,0
Plots served by roads	132	268	+ 103,0
% of plots served by roads	38,0	100,0	+ 163,2

5.2 Study area

The land consolidation area is a part of the broader region of Chlorakas village in the District of Pafos which is located at an altitude of 70 m above mean sea level and at a distance of 3km north of the town of Pafos (Figure 11). The village administrative boundaries cover a total area of 492 hectares of lowland while the extent of the consolidated area is 195 hectares.

Figure 11. Location of the case study in the district of Pafos

The main crops cultivated in the area are citrus, vines, vegetables and bananas. The project began in March 1971 and was among the first applied in Cyprus. It was completed in June 1974 and a cadastral map showing the parcel layout, roads etc. before and after land consolidation, is illustrated in Figure 12. Comparative statistics before and after land consolidation implementation in Table 4, show a considerable reduction in land fragmentation. In particular, the average size of ownership increased by 31.4%, the number of plots fell by 22.8%, the number of shares declined by 97.0%, the average number of parcels or shares per landowner reduced by 18.8%, etc. Legislation was strictly applied with a limited number of justified exceptions from the rules. These results illustrate that the project was successful and are therefore appropriate for comparison with the results of *LACONISS*.

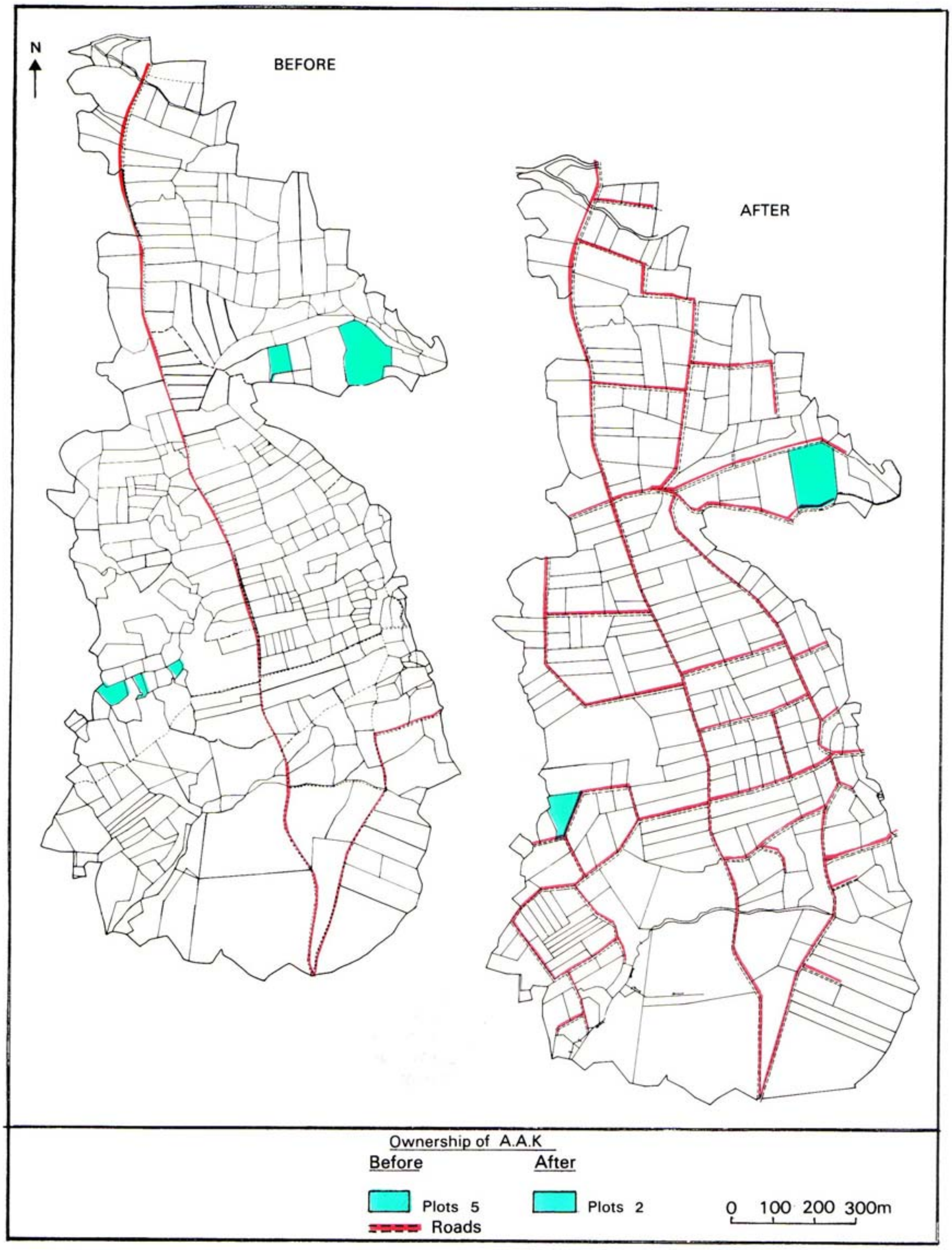


Figure 12. Study area before and after land consolidation

5.3 Data collection

The following data were provided by the LCD for the study area: databases; cadastral maps; and documents. Furthermore, many discussions were undertaken with some members of the team that carried out the Chlorakas project to clarify aspects related to the data and to get other useful information not included in the data.

Five database files were provided that were initially cleaned and renamed for the purpose of this research. The final structure of the collected databases is shown in Table 5.

Table 5: The original databases

	Database Fileneme		Database Fields	Field Type	Field Description
1	LandOwnersEN	1	**Owner_ID**	Number	Primary Key; it is a unique number for each landowner
		2	Owner_Code	Text	It is a code for each landowner used by the Department
		3	Owner-Name	Text	Name and Surname of an owner
		4	Total_Old_Area_Owned	Number	The original total area (in sqm) of the property owned by an owner in the study area
		5	Total_Old_Value_Owned	Number	The original total value (in Cyprus pounds) of the property owned by an owner in the study area
		6	Total_New_Area_Owned	Number	The total area (in sqm) of the property received by an owner in the new plan
		7	Total_New_Value_Owned	Number	The total value (in Cyprus pounds) of the property received by an owner in the new plan
2	OriginalParcels	1	**Parcel_ID**	Text	Primary Key;it is a unique number for each original parcel
		2	Parcel_Sheet_Number	Text	The cadastral sheet number in which a parcel belongs
		3	Parcel_Area	Number	The official registered area of a parcel
		4	Parcel_Value	Number	The value (CyP) of a parcel as it has been defined by the Valuation Committee
3	OriginalParcelsOwnership	1	Owner-ID	Number	it is a unique number for each landowner
		2	Owner_Code	Text	It is a code for each landowner used by the Department
		3	Parcel_ID	Text	it is a unique number for each original parcel
		4	Share_Numerator	Number	The numerator of the fraction of a share for a parcel
		5	Share_Denominator	Number	The denominator of the fraction of a share for a parcel
		6	Only_Trees	Boolean	The property includes only trees
4	NewParcels	1	**Parcel_ID**	Text	Primary Key;it is a unique number for each new parcel
		2	Parcel_Area	Number	The area of a new parcel
		3	Parcel_Value	Number	The value of a new parcel
5	NewParcelOwnership	1	Owner-ID	Number	it is a unique number for each landowner
		2	Owner_Code	Text	It is a code for each landowner used by the Department
		3	Parcel_ID	Text	it is a unique number for each new parcel
		4	Share_Numerator	Number	The numerator of the fraction of a share for a parcel
		5	Share_Denominator	Number	The denominator of the fraction of a share for a parcel

The following three cadastral maps were provided:

- a 1:5,000 cadastral map showing the original cadastral situation before land consolidation (in reduced size in Figure 13);
- a land consolidation plan at the same scale drawn by hand which shows the cadastral situation after the implementation of land consolidation (in reduced size in Figure 14); and
- a 1:2,500 cadastral map after the registration of the new cadastral status by the Land Surveys Department. A sheet of this map is shown in Figure 15 (in reduced size).

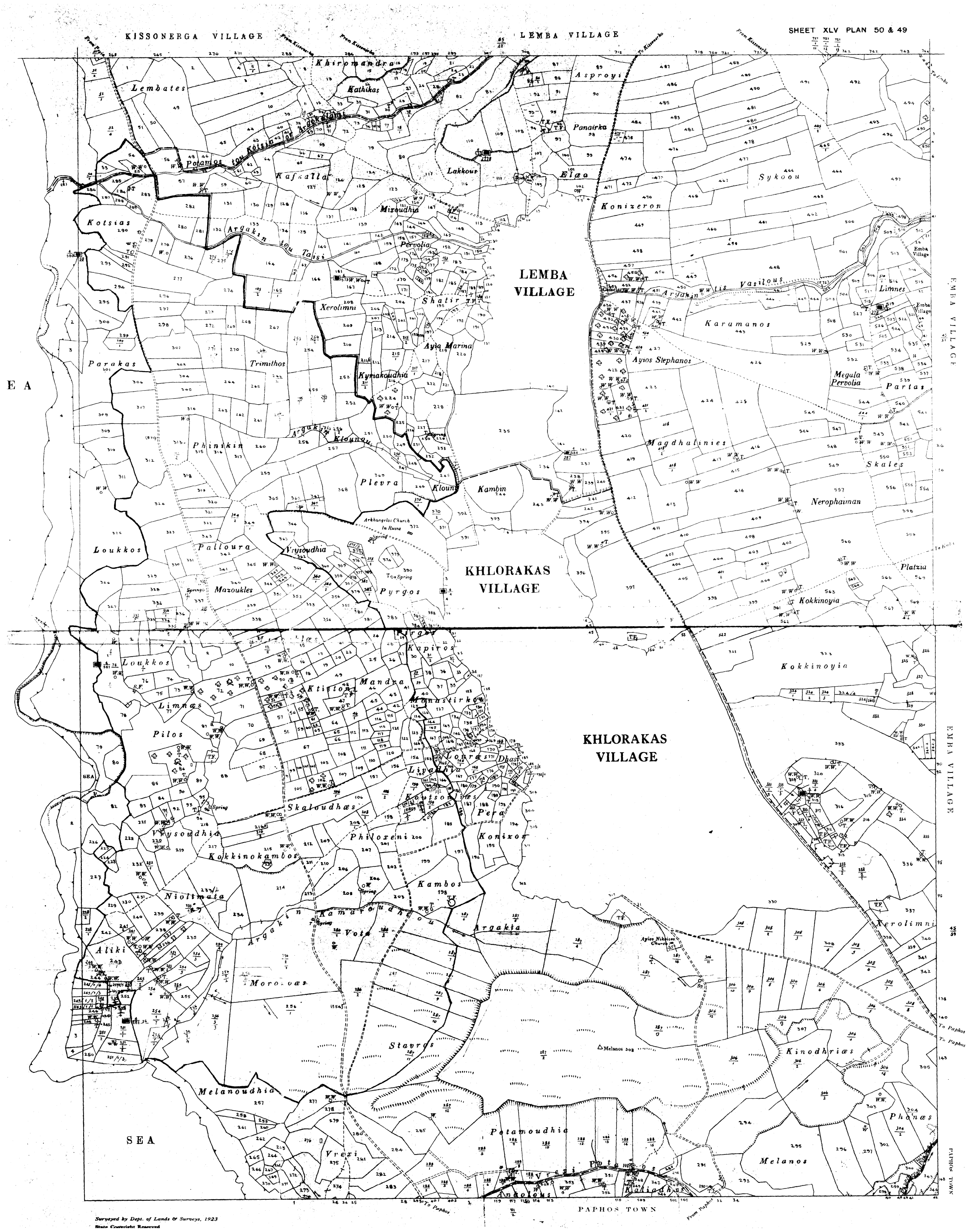


Figure 13: The original cadastral map

Figure 14: The original land consolidation plan

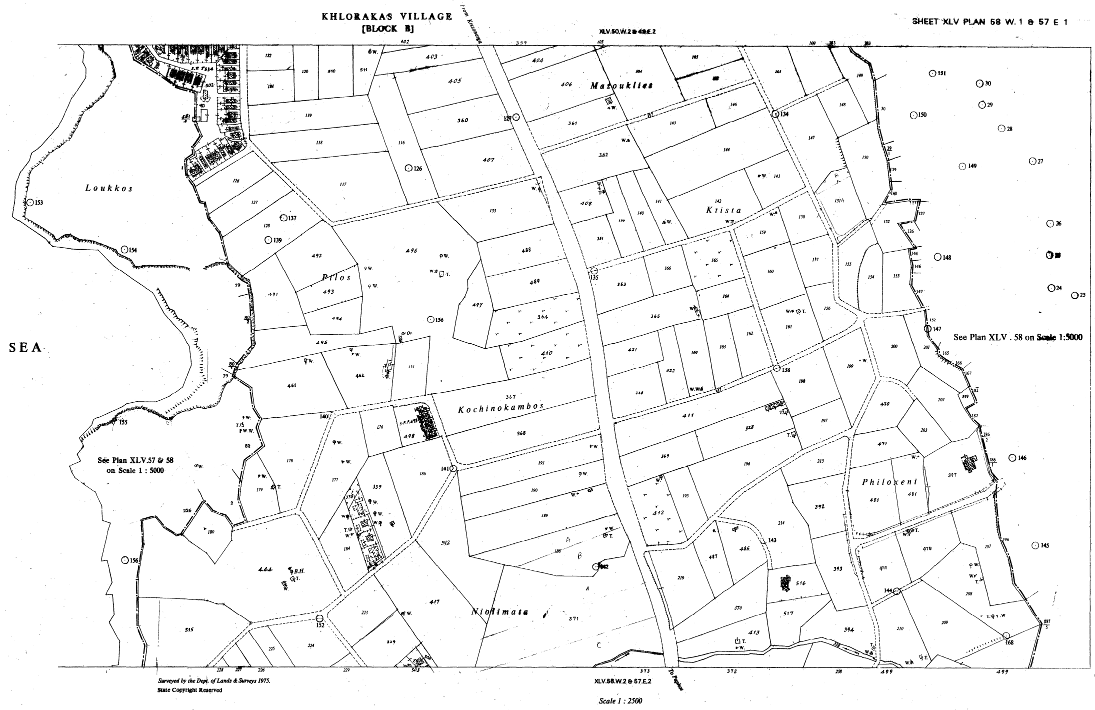


Figure 15: The official cadastral map after land consolidation

The following two kinds of documents were collected:

- catalogues in which the cadastral situation is recorded before and after land consolidation indexed by the name of the landowner(s) and by the Owner_ID of the parcel. Part of each catalogue is shown in Tables 6 and 7 respectively; and
- photocopies of some documents (e.g. the proceedings of the meetings of the Land Consolidation Committee) which contain useful information about the decisions of the Committee regarding the land consolidation project. Some Committee decisions are used as 'facts' in *LandSpaCES*, e.g. the minimum parcel area and the value limit of each parcel that a landowner should have so as to receive a property in the new plan.

5.4 Building the geodatabase

The geodatabase consists of two datasets that contain the information before (original data) and after (the human expert's solution) land consolidation, respectively. Both datasets are composed of layers and database tables as shown in Table 8. Dataset 1 consists of three layers and two database tables. It is used as input to the system to create its outputs, i.e.

output layers and database tables. Dataset 2 consists of one layer and two databases tables. It is also used as an input to the system for system evaluation.

Table 6: The landowner’s catalogue before land consolidation (in Greek)

ΚΑΤΑΛΟΓΟΣ ΙΔΙΟΚΤΗΤΩΝ ΠΕΡΙΟΧΗΣ ΑΝΑΔΑΣΜΟΥ T3CHLORAKA
ΜΕ ΠΑΛΑΙΟ+ΕΚΤΑΣΗ+ΑΞΙΑ+ΜΕΡΙΔΙΟ

Σελίδα: 1

Α/Α	ΚΩΔΙΚΑΣ	ΟΝΟΜΑΤΕΠΩΝΥΜΟ /ΔΙΕΥΘΥΝΣΗ	Φ./Σ.	ΤΕΜΑΧΙΟ	ΕΚΤΑΣΗ	ΑΞΙΑ ΕΥΡΩ σ.	ΜΕΡΙΔΙΟ	ΕΚΤ.ΜΕΡ.	ΑΞΙΑ ΜΕΡΙΔ. ΕΥΡΩ σ.
1	1	ΑΓΓΕΛΙΚΗ ΒΡΥΩΝΗ ΧΡΙΣΤΟΔΟΥΛΟΥ	45/50	319	3.344	941,44	ΟΛΟ	3.344	941,44
								3.344	941,44
2	2	ΑΓΓΕΛΙΚΗ ΝΙΚΟΛΑ ΧΡΙΣΤΟΔΟΥΛΟΥ	45/58	92	2.006	538,21	ΟΛΟ	2.006	538,21
			45/58	94	74	8,54	2/5	30	3,42
								2.036	541,63
3	3	ΑΓΓΕΛΙΚΗ ΚΛΕΑΝΘΟΥΣ ΚΩΝΣΤΑΝΤΙΝΟΥ	45/58	204	6.689	1.623,17	ΟΛΟ	6.689	1.623,17
								6.689	1.623,17
4	4	ΑΘΗΝΑ ΝΙΚΟΛΑ Χ"ΓΙΑΝΝΗ	45/50	316	9.030	2.421,94	ΟΛΟ	9.030	2.421,94
								9.030	2.421,94
5	5	ΑΛΕΞΑΝΔΡΟΣ ΠΑΠΑΚΩΣΤΑ	45/50	370/1	3.010	730,43	ΟΛΟ	3.010	730,43
								3.010	730,43
6	6	ΑΛΕΞΗΣ Χ"ΕΥΣΤΑΘΙΟΥ ΚΑΤΣΑΡΙΔΗΣ	45/50	283	8.361	2.228,87	ΟΛΟ	8.361	2.228,87
			45/50	284					
			45/50	286	1.226	313,53	ΟΛΟ	1.226	313,53
			45/50	287	3.586	915,81	ΟΛΟ	3.586	915,81
			45/50	288					
			45/50	289					
								13.173	3.458,21
7	7	ΑΛΕΞΑΝΔΡΑ ΝΙΚΟΛΑ ΑΛΕΞΑΝΔΡΟΥ	45/50	342	9.030	2.542,40	ΟΛΟ	9.030	2.542,40
								9.030	2.542,40
8	8	ΑΝΑΣΤΑΣΙΑ ΧΡΙΣΤΟΔΟΥΛΟΥ ΓΙΑΓΚΟΥ	45/58	164	1.244	238,35	ΟΛΟ	1.244	238,35
								1.244	238,35
9	9	ΑΝΑΣΤΑΣΙΑ ΓΕΩΡΓΙΟΥ ΑΡΓΥΡΟΥ	45/50	335	1.672	405,79	ΟΛΟ	1.672	405,79
			45/50	335ΕΝΤΟΣ			31/56		
			45/58	4	5.016	1.302,81	ΟΛΟ	5.016	1.302,81
			45/58	5					
								6.688	1.708,60
10	10	ΑΝΑΣΤΑΣΙΑ ΑΓΗΣΙΛΑΟΥ Χ"ΙΩΑΝΝΟΥ	45/50	355	2.675	649,27	ΟΛΟ	2.675	649,27
			45/58	10	6.689	1.255,82	ΟΛΟ	6.689	1.255,82
								9.364	1.905,09
11	11	ΑΝΑΣΤΑΣΙΑ ΓΕΩΡΓΙΟΥ ΓΙΑΝΝΙΚΚΟΥ	45/50	258	9.030	2.652,60	3/6	4.515	1.326,30
								4.515	1.326,30
12	12	ΑΝΑΣΤΑΣΙΑ ΧΑΡΙΛΑΟΥ ΝΕΟΦΥΤΟΥ	45/50	305	6.020	1.534,32	1/8	753	191,79
			45/50	314	8.026	2.045,20	1/2	4.013	1.022,60
			45/58	230	5.351	1.025,16	1/2	2.675	512,58
			45/58	240					
			45/58	236	2.676	977,32	1/2	1.338	488,66
			45/58	236ΕΝΤΟΣ			1/6		
								8.779	2.215,63

Table 7: The parcels catalogue after land consolidation (in Greek)

ΚΑΤΑΛΟΓΟΣ ΝΕΩΝ ΤΕΜΑΧΙΩΝ ΤΗΣ ΠΕΡΙΟΧΗΣ ΑΝΑΔΑΣΜΟΥ T3CHLORAKA
ΕΥΡΕΤΗΡΙΟ ΚΑΤΑ ΑΥΞΟΝΤΑ ΑΡΙΘΜΟ ΤΕΜΑΧΙΟΥ(ΜΕΤΑ ΤΗ ΚΤΗΣΗ ΚΑΤΟΧΗΣ)

Σελίδα: 1

Α/Α	ΑΡΙΘΜΟΣ ΝΕΟΥ ΤΕΜΑΧΙΟΥ	ΦΥΛΛΟ / ΣΧΕΔΙΟ	Ο	ΕΚΤΑΣΗ ΤΕΜΑΧΙΟΥ (τ.μ.)	ΑξΙΑ ΤΕΜΑΧΙΟΥ	ΜΕΡΙΔΙΟ	ΣΕΛΙΔΑ ΣΤΟΝ ΤΥΠΟ 3	ΟΝΟΜΑΤΕΠΩΝΥΜΟ ΙΔΙΟΚΤΗΤΗ ΔΙΕΥΘΥΝΣΗ
1	1		00	2.676	594,59	ΟΛΟ	59	ΕΛΕΝΗ ΣΑΒΒΑ Χ"ΕΥΣΤΑΘΙΟΥ
2	2		00	7.358	1.990,52	ΟΛΟ	6	ΑΛΕΞΗΣ Χ"ΕΥΣΤΑΘΙΟΥ ΚΑΤΣΑΡΙΔΗΣ
3	3		00	3.010	686,00	ΟΛΟ	143	ΜΕΛΑΝΗ ΓΕΝΕΘΛΙΟΥ ΚΩΝΣΤΑΝΤΗ
4	4		00	3.010	686,00	ΟΛΟ	85	ΘΕΚΛΑ ΓΕΝΕΘΛΙΟΥ ΚΩΝΣΤΑΝΤΗ
5	5		00	3.010	686,00	ΟΛΟ	177	ΣΟΦΙΑ ΓΕΝΕΘΛΙΟΥ ΚΩΝΣΤΑΝΤΗ
6	6		00	8.695	1.946,95	ΟΛΟ	177	ΣΟΦΙΑ ΓΕΝΕΘΛΙΟΥ ΚΩΝΣΤΑΝΤΗ
7	7		00	5.518	1.424,55	ΟΛΟ	6	ΑΛΕΞΗΣ Χ"ΕΥΣΤΑΘΙΟΥ ΚΑΤΣΑΡΙΔΗΣ
8	8		00	18.562	5.460,69	ΟΛΟ	50	ΔΕΣΠΟΙΝΑ ΛΟΥΚΑ ΧΡΙΣΤΟΥ
9	9		00	3.679	838,92	ΟΛΟ	50	ΔΕΣΠΟΙΝΑ ΛΟΥΚΑ ΧΡΙΣΤΟΥ
10	10		00	3.678	798,77	ΟΛΟ	102	ΚΑΤΕΡΙΝΑ ΑΓΑΘΟΚΛΗ ΙΩΑΝΝΟΥ
11	11		00	6.187	1.343,39	ΟΛΟ	89	ΙΛΙΑΔΑ ΧΑΡΑΛΑΜΠΟΥΣ Ν. ΚΥΠΡΙΑΝΟΥ
12	12		00	2.675	580,92	ΟΛΟ	195	ΧΡΙΣΤΟΣ ΛΟΥΚΑ ΧΡΙΣΤΟΥ
13	13		00	2.675	580,92	ΟΛΟ	114	ΚΥΡΙΑΚΟΣ ΛΟΥΚΑ ΧΡΙΣΤΟΥ
14	14		00	9.030	1.889,71	ΟΛΟ	30	ΒΡΥΩΝΑ ΑΓΑΘΟΚΛΗ ΠΑΝΑΡΕΤΟΥ
15	15		00	16.221	3.647,86	ΟΛΟ	74	ΕΛΕΝΗ ΧΡΥΣΟΣΤΟΜΟΥ ΣΠΥΡΟΥ
16	16		00	9.365	2.637,23	ΟΛΟ	45	ΓΕΩΡΓΙΟΣ ΛΕΩΝΙΔΑ ΧΡΙΣΤΟΔΟΥΛΟΥ
17	17		00	6.689	1.551,84	ΟΛΟ	172	ΡΕΒΒΕΚΑ ΣΤΥΛΙΑΝΟΥ ΧΡΙΣΤΟΔΟΥΛΟΥ
18	18		00	10.367	2.110,12	ΟΛΟ	119	ΚΥΡΙΑΚΟΥ Χ"ΝΙΚΟΛΑ ΑΧΙΛΛΕΑ ΠΑΦΟΣ
19	19		00	7.190	1.069,58	ΟΛΟ	159	ΞΑΝΘΙΠΠΗ ΙΩΑΝΝΟΥ ΑΝΤΩΝΙΟΥ
20	20		00	5.351	1.363,46	ΟΛΟ	128	ΜΑΡΙΑ ΚΛΕΑΝΘΟΥΣ ΚΩΝΣΤΑΝΤΙΝΟΥ ΠΑΦΟΣ
21	21		00	5.016	1.278,46	ΟΛΟ	197	ΧΡΙΣΤΟΔΟΥΛΟΣ ΧΑΡΑΛΑΜΠΟΥΣ ΠΑΠΑΝΙΚΟΛΑ Η' ΠΑΡΠΑΣ
22	22		00	4.347	1.108,03	ΟΛΟ	156	ΝΙΚΟΛΑΣ ΧΑΡΙΛΑΟΥ ΝΕΟΦΥΤΟΥ
23	23		00	10.367	2.771,78	ΟΛΟ	166	ΠΑΝΑΓΙΩΤΑ ΕΥΣΤΑΘΙΟΥ Χ"ΦΙΛΙΠΠΟΥ
24	24		00	4.849	1.300,67	ΟΛΟ	152	ΝΙΚΟΛΑΣ ΣΠΥΡΟΥ ΕΛΛΗΝΑΣ
25	25		00	9.197	2.466,79	ΟΛΟ	131	ΜΑΡΙΑ ΧΑΡΑΛΑΜΠΟΥΣ ΠΑΠΑΝΙΚΟΛΑ ΠΑΦΟΣ
26	26		00	4.013	1.076,42	ΟΛΟ	86	ΘΕΚΛΑ ΚΥΡΙΑΚΟΥ ΚΟΜΜΑΤΟΥ
27	27		00	10.367	3.177,14	ΟΛΟ	175	ΣΑΒΒΑΣ ΙΩΑΝΝΟΥ ΑΓΑΘΟΚΛΗ ΠΑΝΑΡΕΤΟΥ
28	28		00	6.689	1.729,10	ΟΛΟ	28	ΒΙΚΤΩΡΙΑ ΣΟΦΟΚΛΕΟΥΣ ΚΑΡΑΓΙΑΝΝΗ
29	29		00	10.034	3.203,63	ΟΛΟ	54	ΔΗΜΗΤΡΗΣ ΧΑΡΑΛΑΜΠΟΥΣ ΠΑΠΑΝΙΚΟΛΑ
30	30		00	8.695	2.776,48	ΟΛΟ	55	ΔΗΜΗΤΡΑ ΧΡΙΣΤΟΦΟΡΟΥ ΔΗΜΗΤΡΙΟΥ

Table 8: The input data files

Type of data	**Data before land consolidati (Dataset 1)**	**Data after land consolidatio (Dataset 2)**
Layers	OriginalParcels.shp	NewParcels.shp
	Roads.shp	
	Streams.shp	
Database tables	OriginalParcelOwnership.dbf	NewParcelOwnership.dbf
	LandOwners.dbf	

Dataset 1: Data before land consolidation

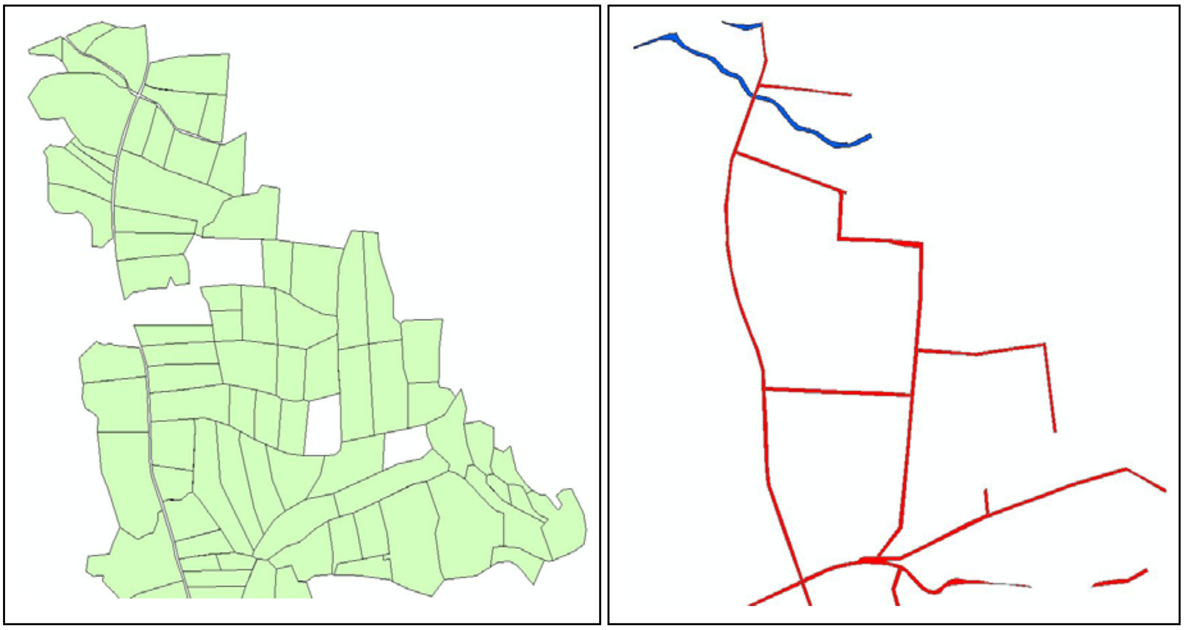
The first feature class named *"OriginalParcels"* has polygon geometry created to represent the original parcels. The shapes of the original parcels were drawn through screen digitizing of the scanned original cadastral map. A part of this layer is shown in Figure 16a. This process was carried out using the functions of the Editor Toolbar of ArcMap. Once the original parcels were created, their attributes, which were contained in the OriginalParcels database table, were copied into the feature class. A feature class in ArcMap is basically a kind of a database table (i.e. with columns as fields and rows as records) and a feature is a type of row. In other words, feature classes are composed of features and tables are composed of rows. A feature, in a feature class, may represent a parcel or a road, and a row in a database table represents attribute information which is related to that feature.

The second and third feature classes named *"Roads"* and *"Streams"* were created in the same way using the same scanned map. They both have polygon geometry to represent the surface of the road network proposed by the expert (some roads already existed) and the streams respectively. A part of both layers (roads illustrated in red and streams in blue colour) are shown in Figure 16b to illustrate this.

These three feature classes with the "Landowners" and "OriginalParcelOwnership" database tables constitute the data input to the system. The system creates new tables and feature classes, some of which are its final outputs.

Dataset 2: Data after land consolidation

The *"NewParcels"* feature class, with polygon geometry, represents the new parcels allocated by the expert; that is, the new land consolidation plan. It was created in the same way using a scanned image of the actual land consolidation plan and by copying the attributes of the NewParcels table to it. This feature class and its related database tables "NewOwnership" and "LandOwners" were used as inputs to evaluate the system outputs.



(a) Parcels (b) Roads and streams

Figure 16: The original parcels and roads/streams layers (part of the study area)

5.5 Establishing relationships

A *relationship* was established among layers and database tables for each dataset. Both relationships are shown graphically in Figure 17.

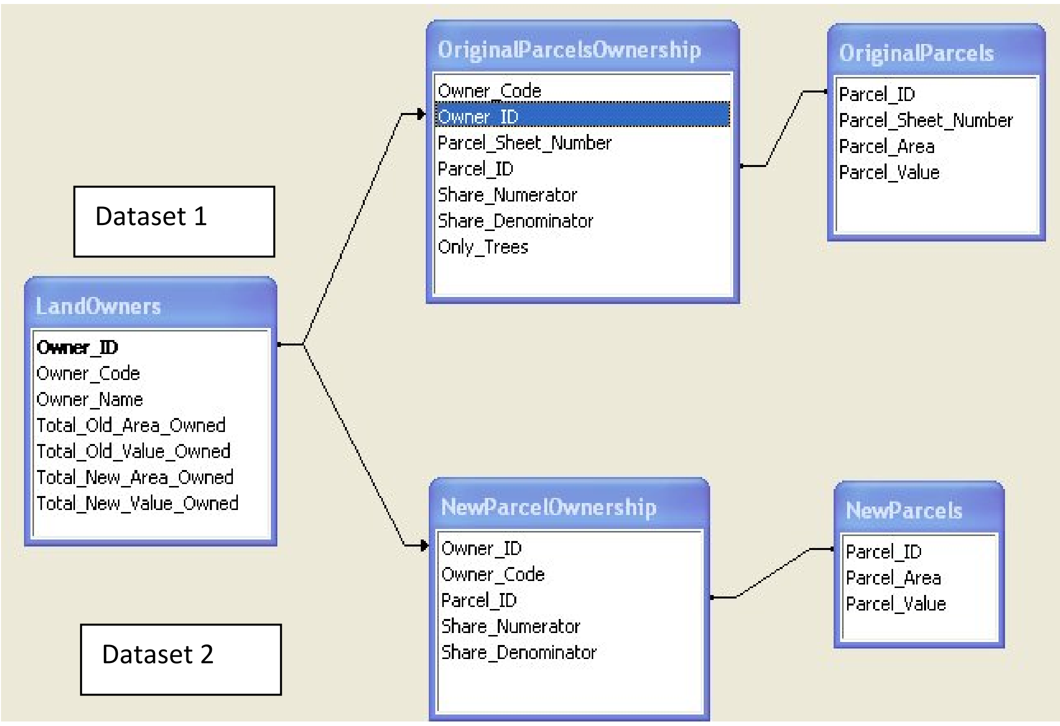


Figure 17: The relationships for each dataset

A more detailed representation of the established relationships for Dataset 1 is shown below:

$$\text{LandOwners} \xrightarrow[\text{(One-to-many relationship)}]{\text{(Key field: Owner\_ID)}} \text{OriginalParcelOwnership} \xrightarrow[\text{(Many-to-many relationship)}]{\text{(Key field: Parcel\_ID)}} \text{OriginalParcels}$$

As shown in Figure 17, the OriginalParcelOwnership table is a junction table between LandOwners and the OriginalParcels tables. A junction table is a table that contains common fields from two or more tables. It is on the many side of a one-to-many relationship with each of the other tables. As an example, a landowner with a unique Owner_ID may have many ownerships, i.e. parcels, shares in parcels or trees in a parcel. The attributes of these parcels are contained in the OriginalParcels layer. When a parcel is shared, i.e. it belongs to more than one landowner or when trees standing on a parcel may be owned by another landowner, this type of ownership is called “dual ownership”. In a similar way, the following relationships can be established for Dataset 2:

$$\text{LandOwners} \xrightarrow[\text{(One-to-many relationship)}]{\text{(Key field: Owner\_ID)}} \text{NewParcelOwnership} \xrightarrow[\text{(Many-to-many relationship)}]{\text{(Key field: Parcel\_ID)}} \text{NewParcels}$$

5.6 Data quality

Data quality is used to give an indication of how good the data are in terms of their overall suitability for a specific purpose or how free they are from errors and other problems (Heywood *et al*., 2002). The two datasets mentioned above that have been used to build and evaluate the system, need to be complete**,** compatible**,** consistent and applicable for the task that will be performed. Also, other parameters of data quality such as accuracy, precision, bias and resolution should be taken into account. Definitions for these quality parameters are found in Heywood *et al*. (2002) or Bernhardsen (2002) and data quality issues are discussed by Howard (1998).

Once we are aware about the possible errors in the data, we have to think about the possible sources of these errors. In this case study, both attribute and spatial errors are possible. Attribute errors may be inherited in the database files and may be introduced during data processing. Spatial errors may be introduced during the conversion of maps to digital format via digitizing.

The data quality parameters noted above provides a useful checklist of quality indicators. It is possible to construct a data lineage, i.e. a record of data history that presents essential information about the development of data from their source to their present format. Instead of creating a complete lineage for this project, the main data quality parameters are briefly discussed.

Completeness refers to a lack of errors of omission in a database. Even though some data were lacking such as the landowners' preferences and personal information such as profession, age etc., the data were adequate to evaluate (verify and validate) the land distribution model and the overall system. The data cover the whole study area and the data model contains all the necessary cadastral information.

Compatibility refers to the capability of different datasets or different layers and databases to be combined to produce outputs. The external boundary of the land consolidation area was slightly modified in the NewParcels layer to be exactly compatible with the corresponding boundary of the OriginalParcels layer.

Consistency refers to the absence of apparent contradictions in a database. Some spatial and attribute inconsistencies were detected in the data. In particular, some polygons (i.e. parcels) existed on the map but not in the attribute table and *vice versa*. Also, landowners who did not receive property in the new plan were not included in the databases.

Applicability is used to describe the appropriateness or suitability of data for a set of commands, operations or analyses. The geodatabase created for this particular problem and the execution of special operations such as overlays, intersections, etc. for further analysis is definitely applicable since the cadastral data model has been used for many years.

Accuracy (which is the inverse of error), concerns spatial and attribute data. The accuracy of spatial data (which is similar to resolution) is limited approximately to 2-3 metres since the scale of the hard copy maps is 1:5,000. This fact affects the accuracy of the coordinates of the original parcels and consequently their locations, shape and area. Also, the calculated parcel area differs from the registered parcel area (as it is recorded on the titles). However, both accuracy errors do not constitute a problem for the purpose of this research. The accuracy of the attribute data is assured since they have officially been provided by the Land Surveys Department (the Government's cadastral and mapping agency).

Bias in a GIS dataset is the systematic variation of data from reality. Bias is a possible error (e.g. in digitizing) for this project although it is not expected to affect the outputs. Finally, precision is the recorded level of detail of a dataset. The precision of the recorded parcels' coordinates does not affect the precision of outputs.

This section has presented the case study characteristics and datasets that have been used (a) in the development of *LACONISS* and (b) as a benchmark for evaluation. The task of building and checking the geodatabase was laborious and took a relatively long period of time because the GIS model was built almost from scratch. The quality of the GIS model is adequate so as to ensure the reliability of system development and the evaluation that follows.

6 SYSTEM EVALUATION

In general, system quality issues are hierarchically ranked as follows: evaluation, assessment, credibility, validation and verification (O'Keefe *et al*., 1987). Evaluation reflects the benefits in terms of value for money to the users, sponsors and the organization more generally. Assessment is the set of issues that consider the 'fit' between the system and the user independently and the quality of decisions made. Credibility is defined as the extent to which a system is believable or the level of confidence in the system results (Balci, 1987). Validation involves the correctness of the software in respect to the user needs and requirements, i.e. building the right system. Thus, validation is more concerned with the quality of the decisions made by the system. On the other hand, verification is defined by Adrion *et al*. (1982) as the consistency, completeness and correctness of the software. In other words, verification means testing that the system has been built correctly in terms of eliminating errors and making sure that it corresponds to the predefined specifications.

In practice, software developers rarely deal with more than verification and validation, the so called 'V&V' (O'Keefe *et al*., 1987). Verification and validation are part of an iterative process in system development. Verification is generally carried out during the development of a system component and during the composition of various system components while validation can be performed for each system component and detailed efforts can be carried out for the complete system. Taking into account that the system under development is a prototype, it is considered adequate to test the system quality based on two these aspects: verification and validation.

6.1 Verification

Normally, in the case of an ES, verification is focused on the knowledge base and not on the inference engine since it is usually provided (and hence verified) by an ES shell. However, in this research, the knowledge base and inference engine are not separated so they are both verified. Usually verification involves the following checks: consistency, completeness, correctness and redundancy. Consistency generally means using consistent variable names across all of the rules. Completeness refers to problems that are related to the structure of a rule such as unreferenced and illegal attributes and conclusions or unreachable premises and dead-end conclusions. Correctness focuses on the violation of the structure of the rule-base and in particular having conflicting rules i.e. two or more rules have the same IF attributes, but come to contradictory conclusions; subsumed rules, i.e. two or more rules have the same conclusions but one contains additional constraints on the situations in which it will succeed (Nguyen *et al*., 1987); and circular rules, i.e. a chain of reasoning begins with some condition and then returns to that same condition. Redundancy occurs when two or more rules succeed in the same situation and have the same conclusions (Nguyen *et al*., 1987), or the reasoning chain contains a redundant rule.

The fact that *LandSpaCES* was developed using conventional programming tools and not an ES shell considerably alters the anomalies that can occur with rules and introduces other anomalies that need to be considered. Since *LandSpaCES* consists of Procedures, Functions and Rule Cluster Procedures, the verification process is separated into two parts: one for the Procedures and Functions which carry out customised and necessary tasks such as calculations, creation of tables, fields, etc. and one for the Rule Cluster Procedures which comprise the systems' expert mechanism, i.e. the rules and the inference engine. The verification process followed is shown graphically in Figure 18.

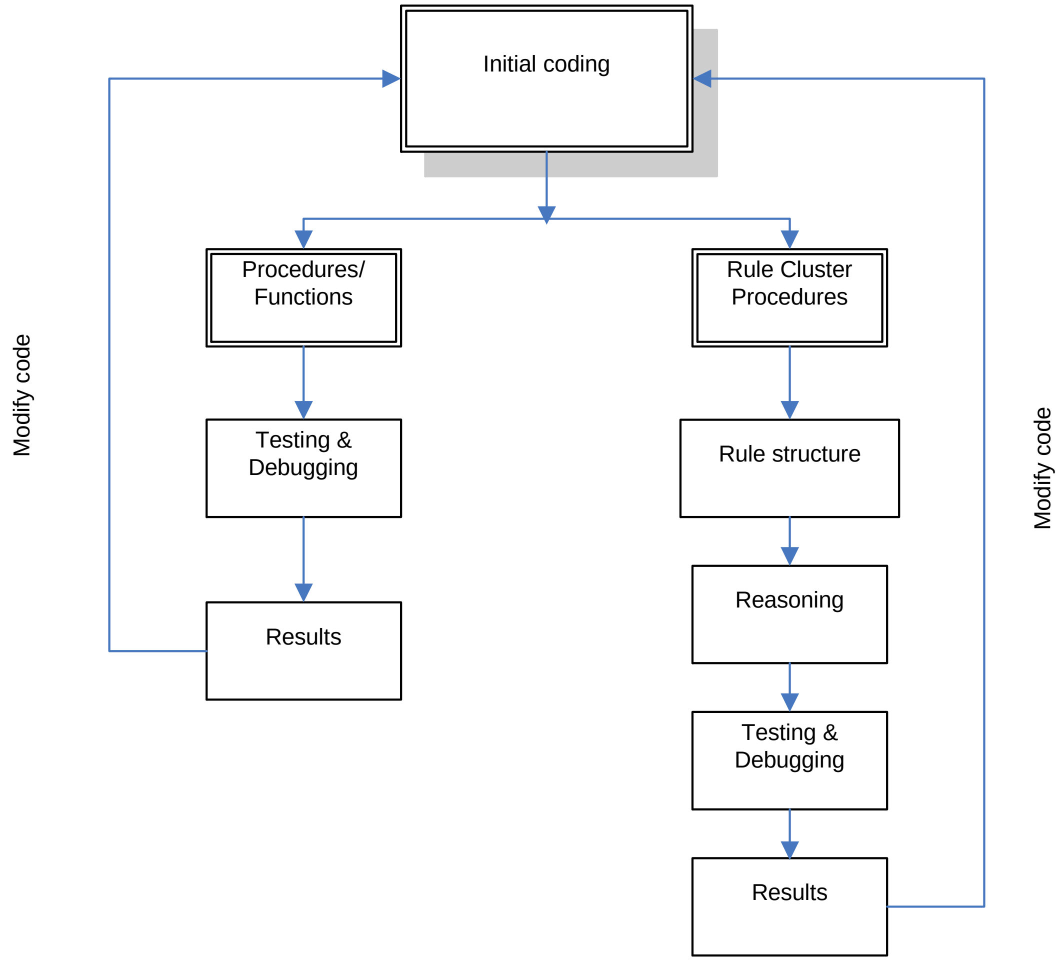


Figure 18: The verification process for *LandSpaCES*

Verification of Procedures and Functions

Verification of Procedures and Functions is an iterative process which is initially carried out individually for each Procedure-Function to verify that it works properly and that it produces the correct results. Afterwards, Procedures and Functions which are called by other Procedures are also verified when the calling Procedure is verified. As shown in Figure 18, after the initial coding, the next step is testing and debugging. Testing is the process of finding errors in a program and debugging is the process of correcting errors that are found. The last step is to investigate the results. If the results are not correct, then a detailed review and inspection of the code is necessary to detect and correct the errors by modifying the code. This is an iterative process until the code runs properly and produces the correct results.

Verification of Rule Cluster Procedures

As shown in Figure 18, verification of Rule Cluster Procedures involves two more components than the process followed for the Procedures-Functions: rule verification structure and line reasoning (inference engine) verification. As noted earlier, rule verification structure consists of four aspects: consistency, completeness, correctness and redundancy. The structure of rules in a Rule Cluster Procedure has to be consistent with the rules written in plain language which in turn have to be consistent with the decision trees. Thus, a cross-verification is carried out between them to assure the four aspects of rule structure verification are satisfied. Because the total number of rules is relatively small (i.e. 74 rules) and because development is via conventional programming tools rather than an ES shell (which may provide rule checking facilities), all checks were carried out by careful inspection, review and comparison of the system results with the known or expected results.

Verification of reasoning was carried out by careful cross-reviews and inspection of the order of the rules in the code (conflict resolution strategy) and the order of the rules in the decision trees. Afterwards, the exploration of the system results compared with human expert results revealed a possible line of reasoning problems. Common errors detected by the process of Rule Cluster verifications were:

- wrong numbering of the new parcels;
- parcel allocation to the wrong landowner;
- allocation of more parcels than the maximum permissible to a land owner;
- landowners with no allocated parcel while they should have been allocated one;
- generation of an inappropriate number of parcels for a rule cluster; and
- allocation carried out in an unexpected manner.

All the above errors were corrected. Verification ensures that the code works and produces correct outputs. An example of *LandSpaCES* output follows.

Output tables

There are two output tables: *NewParcelLS.dbf* which includes the attributes of the newly created parcels and *NewOwnershipLS.dbf* which includes the attributes of the owners. A part of each of these tables is shown in Figures 19 and 20, respectively. The two tables can be joined with a one-to-one relationship using the Parcel_ID as the key field.

Attributes of NewParcelLS

OID	Parcel_ID	Area	Value	X_Centroid	Y_Centroid	Block	PPI
0	1	3267.23	535.49	144928.16	352915.18	8	0.324893
1	2	6535.44	923.26	145424.05	351810.65	19	0.512819
2	3	8822.7	1377.6	144995.97	353021.67	8	0.546201
3	4	2940.9	415.47	145537.59	352893.32	7	0.270247
4	5	8822.7	1446.12	145009.81	352743.21	12	0.550143
5	6	4411.35	754.4	145201.49	353051.63	7	0.55375
6	7	7841.75	1107.91	144798.5	352601.23	10	0.527578
7	8	5554.49	1171.08	145235.75	351924.33	17	0.523973
8	9	3267.23	510.22	145453.67	351990.21	19	0.316531
9	10	3021.99	809.07	145237.53	352875.09	7	0.390352
10	11	9476.34	1129.78	144998.63	353476.69	4	0.53401
11	12	3920.87	102.04	145613.08	352933.42	7	0.324038
12	13	5228.16	816.36	145343.09	351967.56	17	0.491692
13	14	8822.7	1471.87	145540.78	351941.7	18	0.551625
14	15	5881.8	481.07	144866.67	351908.01	25	0.531603
15	16	9476.34	1479.65	145291.3	353023.41	7	0.554141

Record: 1 Show: All Selected Records (0 out of 259

Figure 19: A part of the *NewParcelLS.dbf* table

Attributes of NewOwnershipLS

OID	Parcel_ID	Owner_ID	Share_N	Share_D
0	1	1	1	1
1	2	3	1	1
2	3	4	1	1
3	4	5	1	1
4	5	7	1	1
5	6	12	1	1
6	7	21	1	1
7	8	27	1	1
8	9	29	1	1
9	10	31	1	1
10	11	36	1	1
11	12	38	1	1
12	13	40	1	1
13	14	41	1	1
14	15	42	1	1
15	16	45	1	1

Record: 1 Show: All Selected

Figure 20: Part of the *NewOwnershipLS.dbf* table

Output maps

A part of the output map (ParcelsCentroids layer) is illustrated in Figure 21. Each point represents the centroid (i.e. the approximate location) of each new parcel. The numerator and the denominator of the fraction shown above each centroid represent the Owner_ID and Parcel_ID for a new parcel, respectively. The map does not distinguish the points by land owner at this stage; nor does it display the PPI, the value or the area of the parcels. The case study area is divided into blocks which are defined as areas surrounded by roads or project boundaries and the Block_IDs are the red underlined numbers In Figure 21. A part of the attribute table of the output map is shown in Figure 22.

Order of rule cluster procedures

An interesting kind of verification was the running of the module using a different order of the rule cluster procedures. In fact, this involves executing the redistribution process using a different ordering of landowners/parcels because each rule cluster procedure represents a different group of landowners/parcels, that is, a different land redistribution case (Cases 1-6). This kind of verification aims to investigate the consistency of the results with those obtained by the ‘normal’ sequence of land redistribution cases as defined by the main decision tree. Thus, we run the program for another two cases: firstly, by inverting the ranking of the rule cluster procedures and secondly by randomly defining the sequence of rule cluster procedures, compared to the ‘normal’ sequence. As a result, we may compare three sets of outputs in terms of sequence of rule clusters, called ‘normal’, ‘inverse’ and ‘random’ cases. Table 9 shows the outputs of the three cases for some basic validation criteria used later in system validation.

Table 9: *LandSpaCES* verification in terms of the order of rule clusters

Validation criterion	**‘Normal’**	**‘Inverse’**	**‘Random’**
Number of landowners who received property	210	210	210
Number of landowners who received a ‘completed’ parcel	31	31	31
Total number of new parcels created	268	269	271
Number of landowners who received a different number of new parcels	---	1	3
Number and percentage of new parcels received by the same landowners in the same blocks and in the same location (centroid)	---	206 (76.86%)	228 (85.07%)

The main findings from this comparison are as follow. There is absolute consistency between the three cases for the first two criteria i.e. the number of landowners who received property and ‘completed parcels’. This is expected since these kinds of decisions are extracted from generic rules (which use the same facts), before the actual beginning of the land redistribution process. On the other hand, the number of the new parcels created and the number of parcels allocated to each landowner differ slightly from those of the ‘normal’ case because they are extracted from specific rules which are based on facts that may change during the evolution of the solution (i.e. the running of the program). In other words, the current redistribution decisions at a given time, as the program evolves, may be influenced by the order of the rule clusters, which in turn affects the decisions. For example, the available land in a block varies from time to time during the evolution of the solution. So, the program may allocate an extra parcel (which of course cannot exceed the maximum number of parcels that may be allocated to a landowner) to a landowner compared to the 'normal' case where the certain solution could not allow this. Similarly, the allocation of parcels in terms of location (block and centroid) are strongly consistent, i.e. 76.86% and 85.07% for the 'inverse' and 'random' case respectively, compared to the 'normal' case. The differences of 23.14% and 14.93% respectively do not constitute an inconsistency of the program because, in these cases, the order of rule clusters normally influences the decision regarding location due to the dynamic character of the land redistribution process.

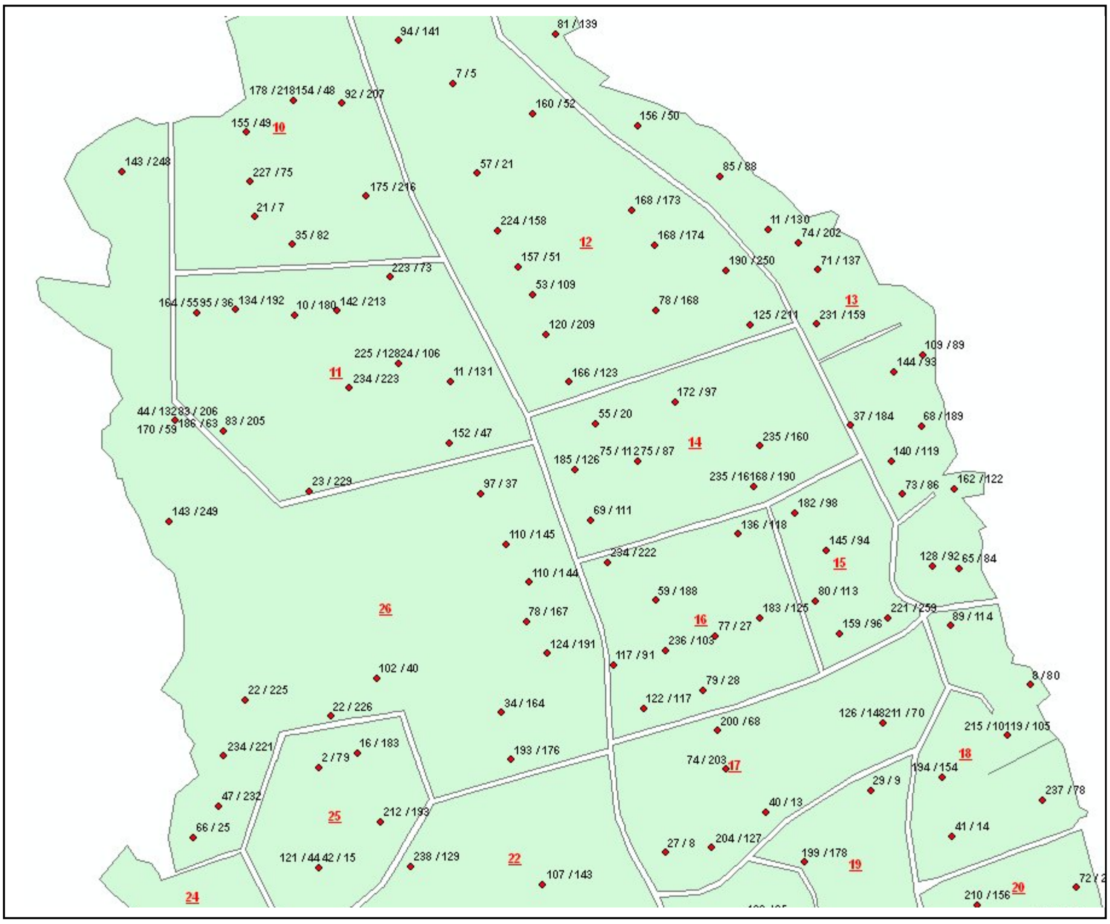
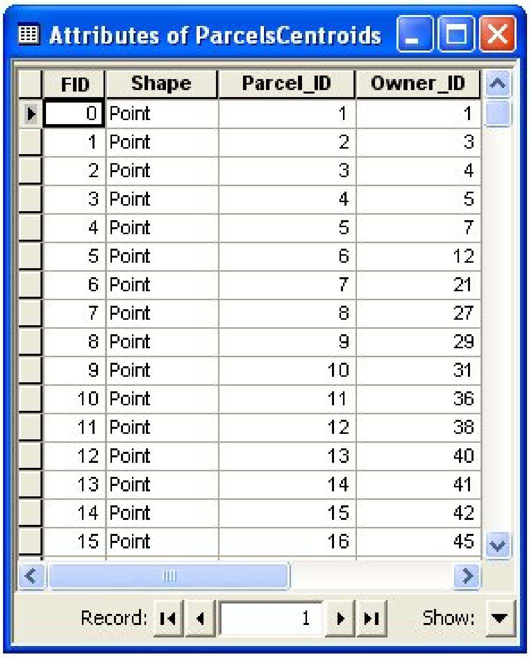


Figure 21: Part of the output map of the case study area showing the new parcel locations as points

The above findings lead to the conclusion that the different order of rule clusters may slightly change the number of parcels allocated to a landowner and may also influence the location of the new parcels. However, both cases are not considered as inconsistencies in the module – they actually constitute alternative land redistribution solutions.

At this stage, the solutions generated input a list of owners and characteristics from a database table in the same order. It is hypothesized that changing this order will change the solution but only to a small degree. This will be investigated further in future.


Attributes of ParcelsCentroids

FID	Shape	Parcel_ID	Owner_ID
0	Point	1	1
1	Point	2	3
2	Point	3	4
3	Point	4	5
4	Point	5	7
5	Point	6	12
6	Point	7	21
7	Point	8	27
8	Point	9	29
9	Point	10	31
10	Point	11	36
11	Point	12	38
12	Point	13	40
13	Point	14	41
14	Point	15	42
15	Point	16	45

Record: 1 Show:

Figure 22: Part of the attribute table relating to the output map

6.2 Validation

Validation is inherently more complex than verification; hence it should be properly structured at the outset of the process. O'Keefe *et al*. (1987) propose a framework consisting of several approaches to structure validation (e.g. establishing criteria, criterion versus construct validity, maintaining objectivity and reliability) and methods (components validation, test case, Turing tests, simulation, control groups, sensitivity analysis, *et cetera*) for carrying out system validation. Similarly, Sojda (2006) reviews a number of validation methods such as Gold Standard, real time-historical datasets, panel of experts, sensitivity analysis and component validation. Also, among various system validation methods, a survey of ES developers (O'Leary, 1991) showed that the 'test' case is the dominant method for the systematic validation of ES. Based on this literature, the 'establishing criteria' approach and the ‘test case’ method have been selected as the most appropriate to structure and carry out the validation, respectively.

Establishing validation criteria

The simplest approach to establish the criteria for validating a system is to define the output level of expertise that the system should perform, remembering that there is no ‘optimum solution’ in this evaluation context. A system, for example, may perform at the level of an expert, better than an expert or at the level of a good trainee. As laid out in the system definition, the objective of *LandSpaCES* is to perform as near as possible to a human expert and in some aspects to perform better than an expert. A common method used to determine the level of expertise of a system is to get the system to ‘sit an exam’, i.e. test it with a case study (test case method) and measure its success in solving the specific problem. A minimum acceptable range can be defined by the sponsors or the users. Many developed ESs have explicit levels of performance as an evaluation criterion. Typically, a number of case studies are presented to the system, and the number of ‘correct’ answers, compared to those of an expert, are tallied. The system is then determined to be, for example, ‘90% correct’ or ‘95% perfect’. In this research, an acceptable performance level must be defined, since this is one of the major research questions, i.e. if the decisions taken by the system can be reasonably close to those of the human expert.

Given the issues discussed above, the following nine performance criteria were defined for *LandSpaCES*:

- number of landowners who received property;
- number of common landowners who received property;
- number of landowners who received a ‘completed’ parcel; a completed parcel is a new parcel which is the unique that is allocated to a landowner when the total size of holding is less than the minimum size provided by legislation for the new parcels in a certain land consolidation area. For example, a landowner may have a holding with a total area of 2000m^2 while the minimum size for the new parcels according legislation should be 2 donums i.e. 2676m^2. Thus, the landowner needs to pay for the extra land of 676 m^2 in order to receive a completed parcel.

- number of common landowners who received a ‘completed’ parcel;
- total number of new parcels created;
- number of new parcels created per owners’ group;
- number of new parcels received by each landowner;
- number of new parcels received by each landowner in common blocks; and
- number of new parcels received by each landowner in a common location.

The above validation criteria cover the most important decisions made by the expert regarding the new land distribution plan. Thus, they can really be used to evaluate the overall system performance when compared to the human expert solution. Results which involve calculations and not decisions are not included in the validation criteria (e.g. area and land value of new parcels).

Validation using a case study

O’Keefe *et al*. (1987) note four guidelines to follow for selecting a case study. First, the problems to be encountered by the system should be reflected in the cases by specifying a prescribed input domain. Second, a sufficient number of test cases covering a variety of input data are required to elicit the range of parameters necessary to test the system and to be able to establish some statistical measures of significance. Third, the nature of the problem investigated by the system should help establish the characteristics of the cases. *LandSpaCES* is validated using a real land consolidation case study by comparing its results/decisions with those taken by the human expert.

The level of system performance

The level of expertise of the system is measured for each validation criterion by comparing the decisions made by the system and the human expert. Table 10 shows the system performance for each validation criterion (CR1-CR9) and Table 11 shows the system performance for each land distribution group.

Table 10: *LandSpaCES* performance based on nine validation criteria

Validation criterion	***LandSpaCES***	**Human expert**	**System performance %**
Number of landowners received propert (CR1)	210	204	98.04
Number of the common landowners received property (CR2)	Agreement in 204 out of 204 landowners		100.00
Number of landowners received a 'completed' parcel (CR3)	31	24	70.83
Number of common landowners receive 'completed' parcel (CR4)	Agreement in 24 out of 24 landowners		100.00
Total number of new parcels created (CI	268	267	99.63
Number of the new parcels created per owners' group (CR6)	See details in Table 5 below		69.23-100
Number of the new parcels received by landowner (CR7)	Agreement in 219 out of 253 landowners		86.56
Number of new parcels received by eacl landowner in common blocks (CR8)	Agreement in 210 out of 267 new parcels		78.65
Number of new parcels received by eacl land owner in a common location (CR9)	Agreement in 167 out of 267 new parcels (See Figure 17)		62.55

Table 11: Number of new parcels created per land distribution group

Land distribution group	LandSpaCES	Human expert	Difference	System performance %
Completed parcel	31	24	+7	70.83
B1	80	79	+1	98.73
B2	26	26	0	100.00
C1	32	32	0	100.00
C1b*	18	16	+2	87.50
C2	18	26	-8	69.23
D1	26	24	+2	91.67
D2	37	38	-1	97.37
Total	268	265+2** =267		

* Group C1 has been programmatically split into two parts for efficiency reasons.
**The human expert allocated two exempted parcels from rules that are not classified in the land distribution groups.

The results of the system performance are very encouraging. The system reproduces the human expert decisions from 62.55 to 100% for the nine validation criteria. In particular, the system allocated property to 210 landowners whereas the human expert allocated 204 (CR1). The difference of six more landowners that have been allocated property by the system is in accordance with the difference of seven more 'completed parcels' created by the system (CR3). This is due to the fact that these landowners presumably had not applied (as regulations require) to receive a 'completed parcel'; that is, they did not wish to be allocated a property in the new plan and they received some compensation. However, the system was not provided with this information since it was unknown and a relevant procedure to take this into account will be developed as a future system enhancement. Despite these slight differences the system and the human expert fully agreed with 'completed parcel' in 204 (CR2) and 24 (CR4) common landowners, respectively. CR5 shows that the system and the expert created almost exactly the same number of new parcels, i.e. 268 and 267 respectively, although there was a notable variation in the number of parcels created for each land distribution group as shown in Table 11.

System performance for each land distribution group (CR6) ranges from 69.23 to 100%. The difference of seven or more ‘completed parcels’ created by the system has already been explained. The difference of eight or more parcels allocated by the human expert compared to the system for group C2 is due to the fact that some landowners have been allocated more parcels than the maximum number provided by the principle of ‘small-medium-large’ holdings size, since the Head of the Department may accept some exceptions from this rule in some justified cases. Again, the system had not been fed with this information and hence strictly applied the rules. The differences for the other land distribution groups are small (i.e. ranging from 0 to 2 parcels) for the same reason. Also, two parcels allocated by the expert cannot be classified in any distribution group because they have been allocated as exceptions to the rules.

CR7 shows that for 219 out of 253 (i.e. 86.56%) cases, the system and the expert agreed to allocate the same number of parcels, i.e. from 0 to 3 in the new plan. CR8 and CR9 refer to the location of the new parcels. In particular, 210 out of 267 parcels (i.e. 78.65%) allocated by the system for each landowner in common blocks agreed with those allocated by the expert. In addition, 167 out of 267 parcels (i.e. 62.55%) have been allocated by the system in exactly the same location (as shown in Figure 16) as that allocated by the expert. The same location (identified as a red dot) means that the centroid of a new parcel falls within the boundaries of the same parcel allocated by the expert. A blue dot is where a parcel centroid produced by *LandSpaCES* is not the one identified by the human expert. The boundaries of the land parcels defined by the human expert have also been added in Figure 23 to indicate the location of each parcel centroid in relation with the boundaries of the corresponding parcel which has been designed by experts in the case study. Each of these parcel centroids will be the starting point for the generation of parcels boundaries with *LandParcelS*.

Although the results showed that system performance against the human expert is good, further improvements could be made by adding some more rules to the knowledge base. Also, testing with more case studies may extract more robust conclusions. However, it should be noted that human experts’ decisions cannot be considered as optimal although this is the standard way of evaluating ES since the aim of such a system is to emulate the human reasoning process utilised for solving a narrow problem domain. So, actually

LandSpaCES has transformed the land redistribution, from a semi-structured, with sometimes subjective decisions taken by experts and the very slow semi-computerised current manner of the process to a systematic, transparent and effective process. Thus, it could be said that the model solution may be better since it overcomes the drawbacks of the current semi-manual method followed.

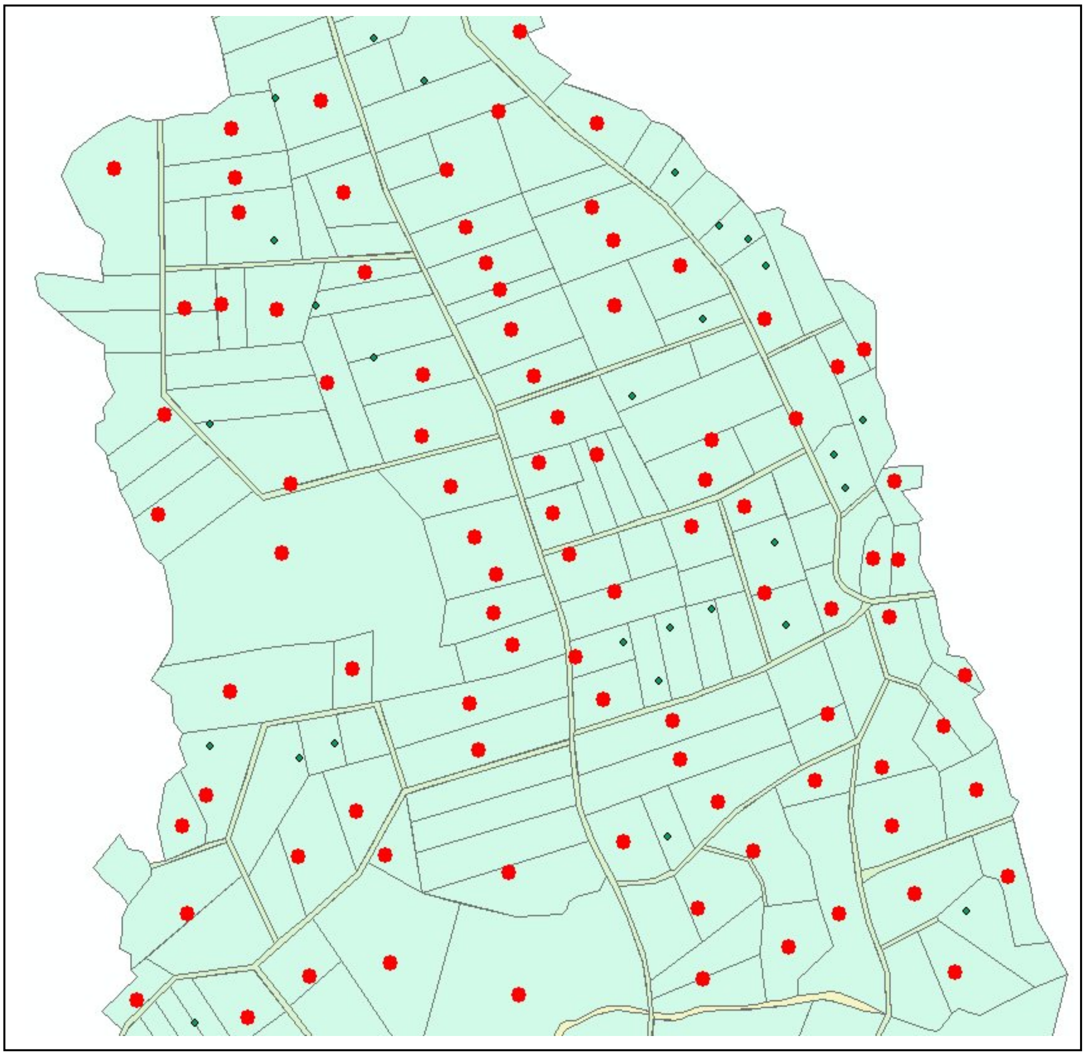


Figure 23: Agreement between those parcel centroids generated by the system and those defined by the human expert (red dots)

Performance in terms of time

One area where the system definitely outperforms the human expert is the amount of time taken to complete the process. A small survey carried out based on ten expert land consol*i*dation technicians showed that an individual expert needs about 30 working days to solve this particular case study land distribution problem. *LandSpaCES* solves the same problem in 6 minutes which is an impressive reduction in time for this task. System superiority over the human expert would be even greater if other data were taken into account, e.g. the landowners' preferences, land use, the landowners' personal data (residences, age, occupation), which were currently lacking from the system database. The system is not able yet to directly model certain kinds of information that the experts had been provided with, namely: the landowners' preferences and demands via official applications; the pre-decisions of the planning team and the Head of the Department to violate (based on justification) a relevant legislation provision; land use; or landowners personal status i.e. full-time or part-time, residence *et cetera*. Thus, despite this type of information not being available to LandSpaCES, the system performed very well.

6.3 Generation of alternative solutions

The primary aim of the Design module of *LandSpaCES* is to be able to generate alternative land distributions by changing the input facts. To test the sensitivity and the reliability of results, the system was run for ten different sets of facts (scenarios) generating ten alternative land redistributions. Then, the interaction between the facts and the results was tested in terms of logic and the variation of changes. The input 'facts', which involve eleven different land distribution variables and the results for three main land distribution criteria, are shown in Table 12. It should be noted that the values of facts need to be feasible with respect to a particular project; otherwise the results will also be infeasible and unrealistic.

A general picture derived from Table 12 is that changing the facts can generate quite different solutions. A1's facts are those used by the expert for the case study used earlier to test the system. By changing F2 and F3 slightly in A2 compared to A1, the three basic outputs changed, that is, they were reduced because the higher F2 and F3 are, the less parcels that are created, the less landowners that receive property and the less landowners that are 'completed'. This result is what was expected and it is also confirmed in A3. By only changing F9 and F10, i.e. the PPI weights in A4 compared to A1, no change in results occurred (except in C1 with one less landowner receiving property) because the PPI influences the location of the new parcels. This result is also expected. A5 involves modifying facts (compared to A1) F4 to F8 representing the 'small-medium-large' holdings rule, that in turn, defines the maximum number of parcels that may be allocated to each landowner. Thus, this change causes an increase (14 parcels) in the number of parcels created (output C1). Outputs C2 and C3 remain stable since they are not affected by facts F4 to F8 but are only affected by fact F1.

Table 12: Facts and outputs for ten alternative land distributions

Alternative No	Facts											Outputs/Criteria		
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	C1	C2	C3
A1	2676 (2)	1000	100	2676 (2)	5351 (4)	5352 (4)	10702 (8)	10703 (8)	0.5	0.5	2000	268	210	31
A2	2676(2)	1500	150	2676 (2)	5351 (4)	5352 (4)	10702 (8)	10703 (8)	0.5	0.5	2000	264	206	27
A3	2676 (2)	2000	200	2676 (2)	5351 (4)	5352 (4)	10702 (8)	10703 (8)	0.5	0.5	2000	259	201	22
A4	2676 (2)	1000	100	2676 (2)	5351 (4)	5352 (4)	10702 (8)	10703 (8)	0.3	0.7	2000	267	210	31
A5	2676 (2)	1000	100	1338 (1)	4014 (3)	4015 (3)	6690 (5)	6691 (5)	0.5	0.5	2000	282	210	31
A6	1338 (1)	750	75	1338 (1)	4014 (3)	4015 (3)	6690 (5)	6691 (5)	0.5	0.5	1000	302	210	9
A7	1338 (1)	500	50	1338 (1)	4014 (3)	4015 (3)	6690 (5)	6691 (5)	0.5	0.5	1000	305	213	12
A8	1338 (1)	500	50	1338 (1)	2676 (2)	2677 (2)	5352 (4)	5353 (4)	0.5	0.5	1000	314	213	12
A9	2676 (2)	1000	100	2676 (2)	5351 (4)	5352 (4)	10702 (8)	10703 (8)	0.7	0.3	2000	268	210	31
A10	6690 (5)	4000	350	6690 (5)	9366 (7)	9367 (7)	14718 (11)	14719 (11)	0.5	0.5	4500	204	182	89

Facts

F1	The minimum parcel area limit (in m^2) for this land consolidation area as set by legislation
F2	The minimum holding's size limit (in m^2) for a landowner to receive a parcel in the new plan as set by the Committee
F3	The minimum holding's land value limit (in CyP) for a landowner to receive a parcel in the new plan as set by the Committee
F4	The lower limit (in m2) of a "small" holding size
F5	The upper limit (in m2) of a "small" holding size
F6	The lower limit (in m2) of a "medium" holding size
F7	The upper limit (in m2) of a "medium" holding size
F8	The lower limit (in m2) of a "large" holding size
F9	The weight for parcel area for the calculation of the PPI (Parcel Priority Index)
F10	The weight for parcel land value for the calculation of the PPI (Parcel Priority Index)
F11	The minimum residual area limit (in m^2) for the creation of a new parcel for those landowners may receive more than one
	Note: the number in brackets represents the area in donums (1 donum=1338 m^2)

Outputs/Criteria

C1	Total number of new parcels created
C2	Number of landowners received property in the new plan
C3	Number of landowners "completed"

Note: the number in brackets represent the are in donums (1 donum=1337.78m^2)

A6 involves a tremendous modification to the project variables by changing all the facts except F9 and F10. F1 decreased to half the value (i.e. 1 donum) compared to A1 which is by itself a big change. The other facts (F2 to F8 and F11) change appropriately taking into account the new value of F1. Results showed a significant increase (34 parcels) in the number of new parcels created (C1) and a remarkable decrease (22 parcels) in the number of landowners who received a ‘completed’ parcel (C3). This result is also feasible because fact F1 is crucial for outputs C1 and C3. C2 remains stable because facts F2 and F3 are almost the same in the alternative concerned and A1.

In addition, by slightly changing F2 and F3 (A7) compared to the previous solution (i.e. A6), all outputs were also slightly changed (increased by 3, C1, C2 and C3) compared with the previous results. This indicates that F2 and F3 are also crucial in the land distribution process because they may change the basic outputs. Changing F4 to F8 and keeping the same facts as in A7, only C1 is changed as happened between A5 and A1. A9 involves inversely changing facts F9 and F10 compared to A4. No change occurred since as mentioned earlier, PPI affects only the location of the new parcels which is not represented in certain outputs.

Finally, A10 involves great changes since it considers that the area under consolidation is non-irrigated (i.e. the minimum acceptable F1 is 5 donums), whereas in all the previous solutions based on the fact that the case study refers to an irrigated land consolidation area (i.e. the F1 can be at minimum either 1 or 2 donums). Thus, by dramatically changing all the facts except for F9 and F10, results in a completely different picture since all outputs are significantly changed: C1 (decreased by 64), C2 (decreased by 28) and C3 (increased by 58) compared with the base alternative i.e. A1.

The above results indicate that the system is robust in generating various alternative land distributions by using different sets of facts. These solutions can then be passed to the Evaluation module of *LandSpaCES* for assessment using MADM methods and more than the three criteria used in this section.

7 CONCLUSIONS

One of the distinguishing features of this research is that land redistribution is treated as a semi-structured decision-making problem and not an optimization problem as considered by most existing studies. In this context, our paper has demonstrated that the integration of GIS with ES methods can efficiently solve the problem of land redistribution by producing results for a real case study that are very close to the decisions arrived at ‘manually’ by human experts. In contrast to studies that have treated the problem as an optimization process, the results are realistic and applicable.

The integration of GIS with ES is not an easy task, since there is a lack of specialised external or embedded tools in proprietary GIS for this purpose. Therefore, the selection of appropriate tools was crucial. The decision to employ a conventional programming platform provided by VBA and ArcObjects, rather than an ES development tool, leads to two potential criticisms, i.e. the obvious disadvantages and the lack of theoretical foundations for building such a system. However, the disadvantages of not using an existing ES development tool are overcome by the great advantages offered by VBA and ArcObjects to develop a fully integrated system. The second issue is addressed via the use of the 'No-Inference Engine Theory' which contains a theoretical foundation for developing an ES using a conventional language.

In addition, other significant system development issues concern the land redistribution itself, which is based on a series of rules extracted from decision trees that are based on principles defined by legislation in force and other official documents. However, two important matters arose: the representation of landowners' preferences and the subjectivity of the decisions made by experts regarding the location of the new parcels allocated to a landowner. To handle these two issues, the innovative concept of a parcel priority index (PPI) was introduced, which is a powerful measure representing: the preferences of landowners regarding the location they wish to receive their new parcels; the priority of a pair of landowner-parcels in the land distribution process in terms of allocating a parcel to a certain location or not; and ensuring equity, transparency and standardisation of the process. The power of PPI is that it treats the two basic components of reallocation, i.e. landowners and parcels, as one common parcel-landowner entity, and that it may take into account many reallocation criteria regarding parcels and landowners.

Whilst the above considerations refer to innovations of this research regarding system development, the actual quality of the system can be revealed through evaluation. Evaluation proved that *LandSpaCES* is a robust and reliable system that has fulfilled its aims. In particular, it is capable of automatically generating a complete land redistribution solution and to be used as a decision support tool by generating a set of alternative land distributions for various sets of facts. Furthermore, the time performance is impressive showing that it has tremendously diminished the time needed by a human expert to carry out the land redistribution process. *LandSpaCES* transforms the land redistribution process, which is semi-structured, complex and time consuming, into an efficient, systematic, standardised and transparent methodology.

Despite the encouraging performance of *LandSpaCES*, there are two limitations: the knowledge base and the ES facilities. As was expected, it is very difficult to model all of the land redistribution reasoning of a human expert. Investigation into the differences between the system and human expert results showed that some more rules could be added into the model to improve its performance. However, some of these rules involve the combination of complex operations that would require further programming and extra time, which is not possible in the context of this research. Another system limitation is the lack of some facilities offered in a typical ES. In particular, since the knowledge base is not separated by the inference engine, it cannot be edited (e.g. new rules cannot be added easily or existing rules cannot be edited) by a user. Instead, programming skills are necessary to carry out this task. Also, the system does not offer an explanation facility which is a very important part of a decision making system in order to explain the decisions in a step by step manner. Both limitations can be overcome by further research and work.

Although LandSpaCES has some limitations that could be tackled by adding more rules to the knowledge base or incorporating additional data into the model, it is a valuable contribution to solving the land redistribution process in terms of automation, effectiveness, equity and transparency, which has potential applicability to many other countries once the relevant country-specific knowledge base is developed.

References

Adrion, W., Branstad, M., Cherniavsky, J., 1982. Validation, verification and testing of software. *ACM Computing Surveys* 14:2, 159-192.

Ayranci, Y., 2007. Re-allocation aspects in land consolidation: A new model and its application. Asian Network for Scientific Information. *Journal of Agronomy*, 6(2), 270-277.

Balci, O., 1987. Credibility assessment. In: Balci, O., (ed.), *Proceedings of the 1987 Eastern Simulation Conference*. The Society for Computer Simulation, La Jolla, CA.

Bernhardsen, T., 2002. *GIS: An introduction*, 3rd ed. Wiley, New York.

Breuker, A., Wielinga, J., 1983. *Analysis Techniques for Knowledge Based Systems*. *Part 2: Methods for Knowledge Acquisition, Report 1.2*, ESPRIT Project12, University of Amsterdam.

Cay, T., Ayten, T., Iscan, F., 2006. An investigation of reallocation model based on interview in land consolidation. *Shaping the change*. XXIII FIG Congress, Munich, Germany, October 8-13.

Chau, K., 2003. Manipulation of numerical coastal flow and water quality models. *Environmental modelling and software*, 18, 99-108.

Choi, J., 2002. A rule-based expert system using an interactive questions-and-answer sequence. Accessed on 15th September 2007 at URL: http://www.cobblestoneconcepts.com/ucgis2summer2002/choi/choi.htm

Choi, J., Usery, E., 2004. System integration of GIS and a rule-based expert system for urban mapping. *Photogrammetric Engineering and Remote Sensing*, 70, 2, 217-224.

Demetriou, D., Stillwell, J., See, L., 2010a. Land consolidation in Cyprus: Why is an integrated planning and decision support system required? Submitted to *Land Use Policy.*

Demetriou, D., Stillwell, J., See, L., 2010b. A framework for developing an Integrated Planning and Decision Support System for Land Consolidation. Submitted to *Environment and Planning B.*

Eldrandaly, K., 2003. A COM-based Spatial Decision Support System for Industrial site selection. *Journal of Geographic Information and Decision Analysis*, 7(2), 72-92.

Eldrandaly, K., 2006. A COM-based expert system for selecting the suitable map projection in ArcGIS. *Expert Systems with Applications*, 31, 94-100.

FAO, 2008. *Opportunities to Mainstream Land Consolidation in Rural Development Programmes of the European Union*. FAO-Land Tenure Policy Series, Rome.

Geertman S., Stillwell J., 2000. Geoinformation, geotechnology and geoplanning in the 1990s. *Working Paper 00/01*. School of Geography, University of Leeds, Leeds.

Giarrantano J., Riley G., 2005. Expert Systems: Principles and programming. *Course Technology*. Canada, 1-188, 353-380

Griffin, N., Lewis, F., 1989. *A rule-based inference engine which is optimal and VLSI implementable*. Tools for Artificial Intelligence. Architectures, Languages and Algorithms. IEEE International Workshop, 23-25 Oct, pp 246 - 251

Grover, D., 1983. A pragmatic knowledge acquisition methodology, *International Journal on Computing and Artificial Intelligence*, 1, 436-438.

Hart, A., 1986. *Knowledge Acquisition for Expert Systems*. Kogan Page, London.

Heywood, I., Cornelius, S., Carver, S., (2002) *An Introduction to Geographical Information Systems*. 2nd ed. Pearson Education Limited, Harlow.

Hicks, R., 2007. The no inference engine theory-Performing conflict resolution during development. *Decision Support Systems,* 43(2), 435-444.

Howard, V., 1998. Unit100: Data Quality measurement and assessment. NCGIA. Accessed on 1 April 2010 at

URL: http://www.ncgia.ucsb.edu/giscc/units/u100/u100_f.html

Howard, H., 2003. Development of an Expert System for Cartographic Design Education. *PhD Thesis,* Graduate Scholl, University of Kansas.

Jin, Z., Sieker, F., Bandermann, S., Sieker, H., 2006. Development of a GIS-based Expert System for on-site storm water management. *Water Practice & Technology,* 1, 1.

Kidd, L., 1987. *Knowledge Acquisition for Expert Systems: A Practical Handbook*. Plenum Press, New York, London.

Land Consolidation Department, 1993. *Land Consolidation in Cyprus*, Ministry of Agriculture and Natural Resources of Cyprus Nicosia.

Land Consolidation Department, 2010. *Annual Report (2009) for Land Consolidation*. Nicosia, Cyprus.

Liou, Y., 1998. Expert System Technology: knowledge acquisition. In: Liebowitz, J., (eds.) 1998*. The Handbook of Applied Expert Systems*. CRC Press, MA, USA, 2-1 - 2-11

Lukasheh, A., Droste, R., Warith, M., 2001. Review of ES, GIS, DSS and their applications in landfill design management. Waste Management and Research, 19, 177-185.

Nguyen, T., Perkins, W., Laffery, T., Pecora, D., 1987. Knowledge base verification. *AI Magazine,* 8(2), 65-79.

O' Keefe, M., Balci, O., Smith, E., 1987. Validating expert system performance. *IEEE Expert*, 2 (4), 81-89.

O' Leary, D., 1991. Design, development and validation of expert systems: A survey of developers. In: *Verification, Validation and Testing of Expert Systems*. John Wiley, New York.

Paliulionis, V., 2000. Architecture of an Intelligent GIS. *Informatica*, 11(3), 269-280.

Schneider, D., 2004. *An Introduction to Programming Using Visual Basic 6.0*. Fourth Edition. Pearson Education Inc, New Jersey, 880pp.

Sojda, R., 2006. Empirical evaluation of decision support systems. *Environmental Modelling & Software,* 20(2), 269-277.

Sonnenberg, J., (1998). New method for the design of the reallocation plan in land consolidation projects. FIG conference. Accessed on 1st September 2007 at URL: http://www.sli.unimelb.edu.au/fig7/Brighton98/Comm7Papers/TS30-Rosman.html

Van Dijk, T., 2003. Dealing with Central European Land Fragmentation. Eburon, Delft.

Vlado, V., 2002. Knowledge based GIS for site suitability assessment. Accessed on 15th September 2007 at

URL: http://gis.esri.com/library/userconf/proc02/pap1185/p1185.htm

Wilson, L., 2007. Rule based programming. In: C. Dr. Dobb's Portal: The world of software development. CMP Technology. Accessed on 1st September 2007 at URL: http://www.ddj.com/199702471

Appendix A: Decision trees

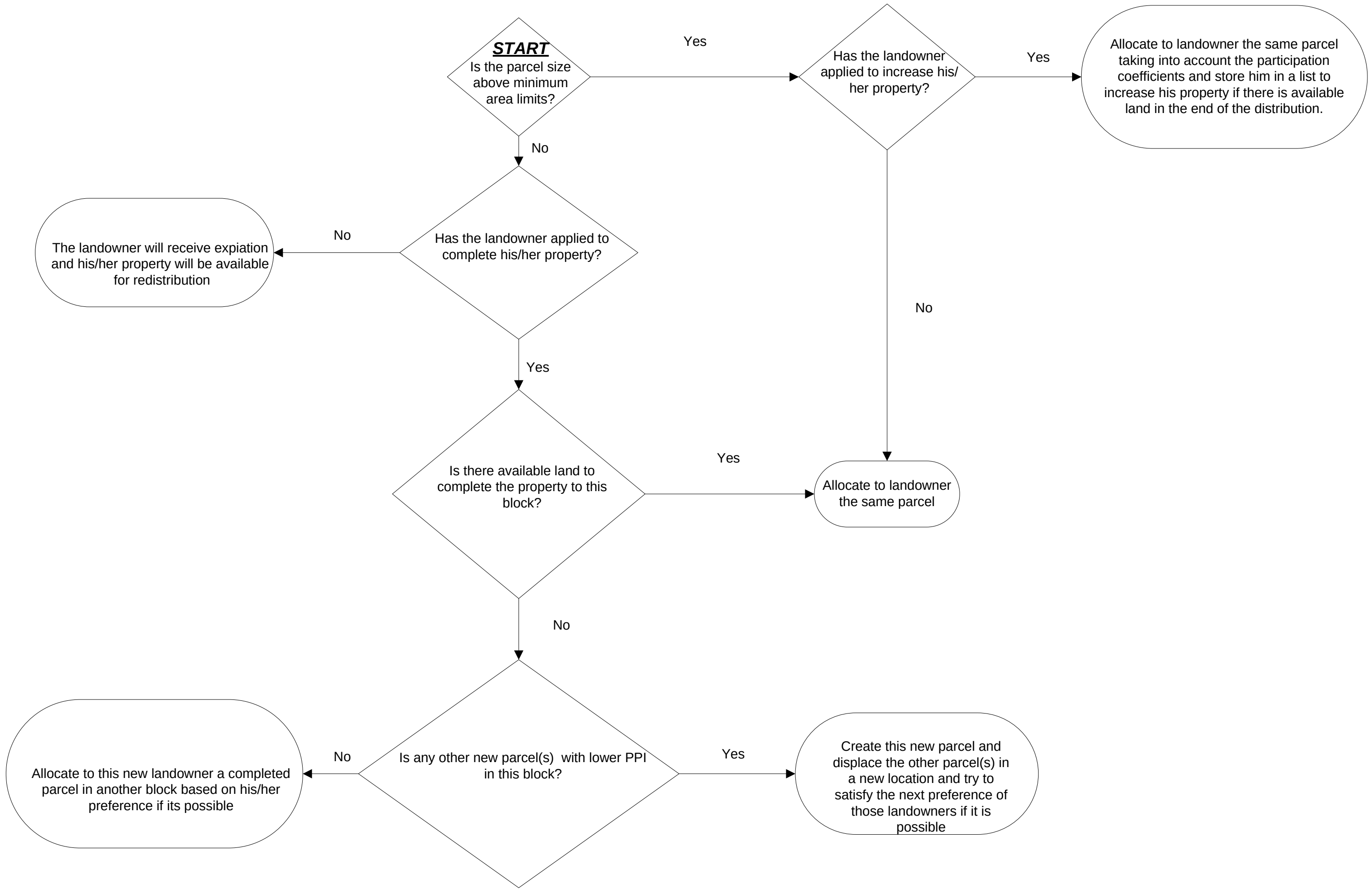


Figure A.1: Flowchart B1: Landowner has only one parcel and may take only one

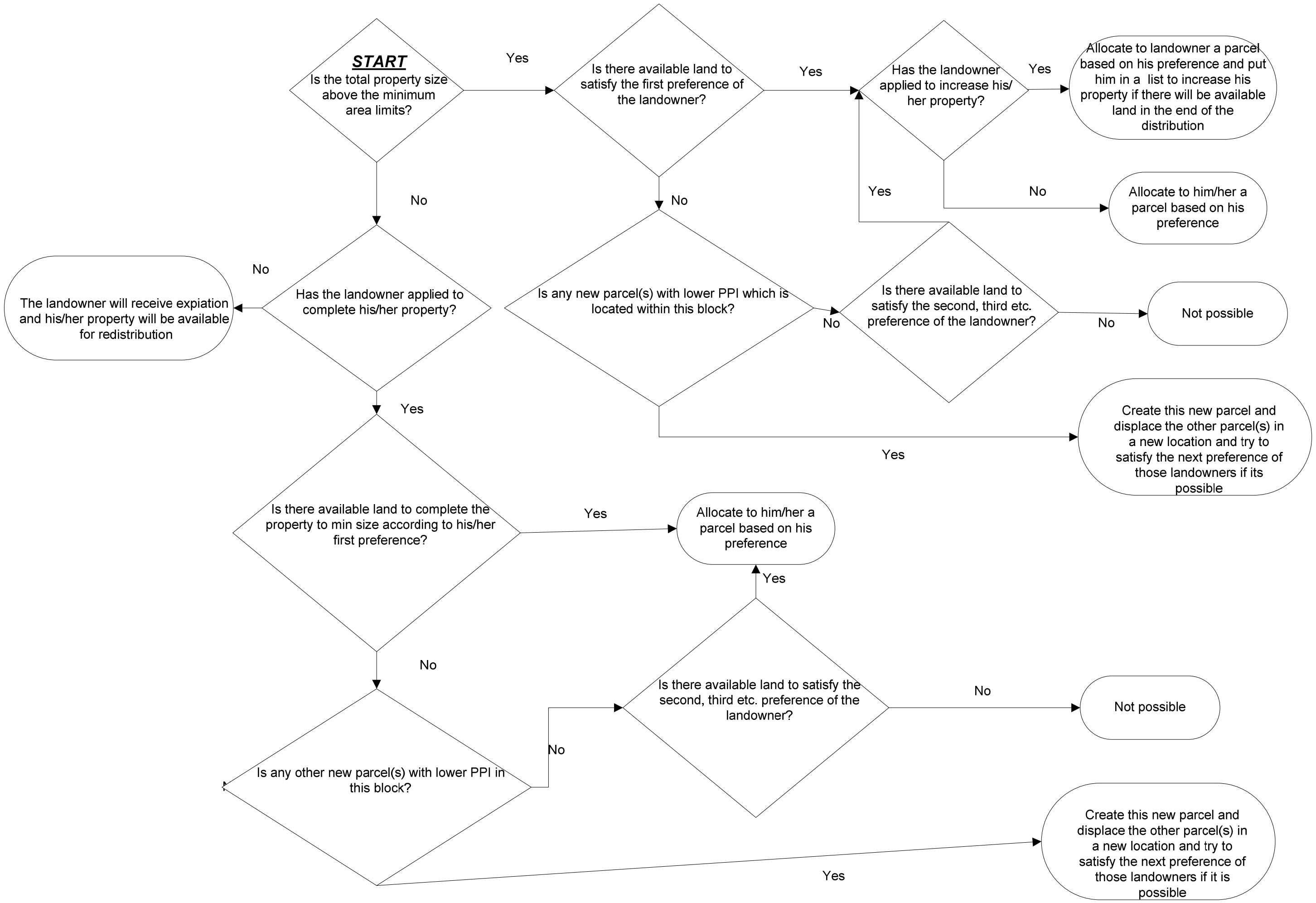


Figure A.2: Flowchart B2: Landowner has more than one parcel and may take only one

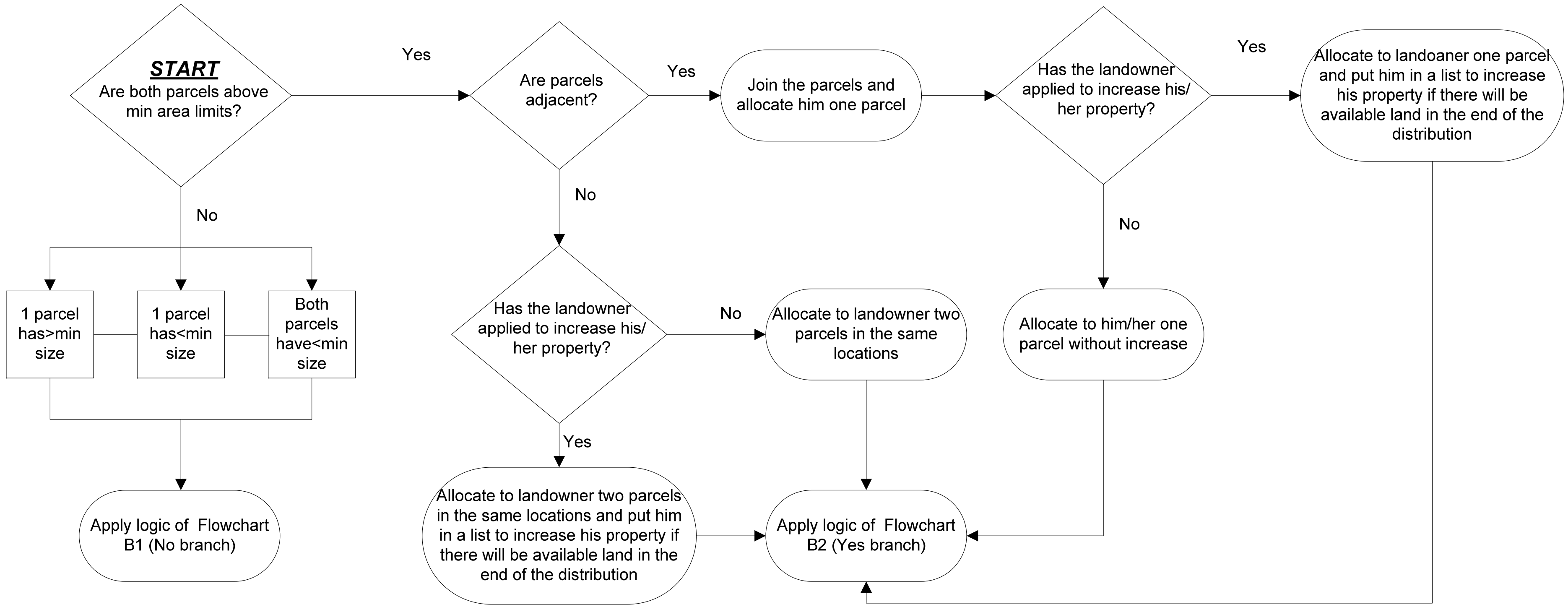


Figure A.3: Flowchart C1: Landowner has two parcels and may take up to two

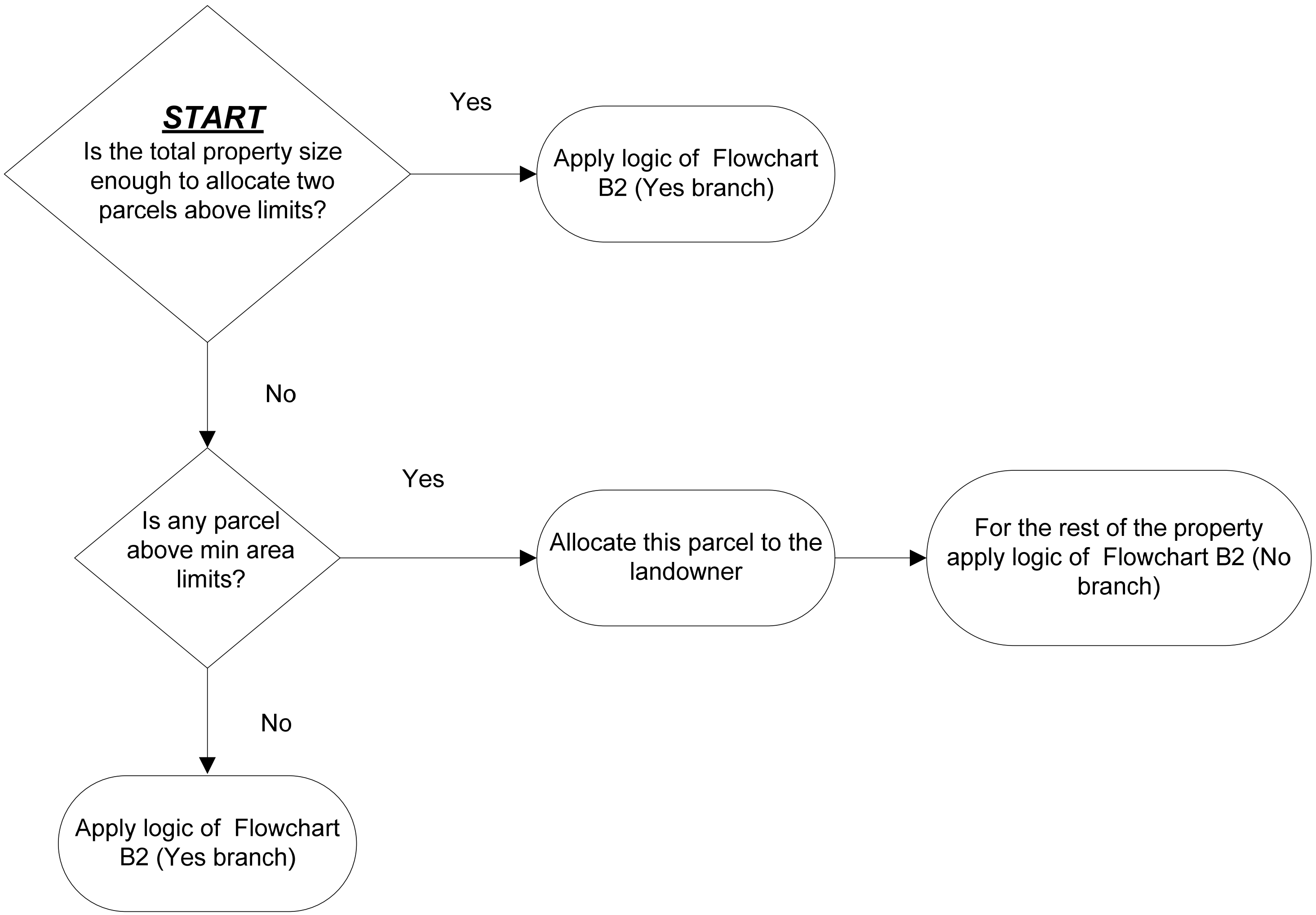


Figure A.4: Flowchart C2: Landowner has more than two parcels and may take up to two

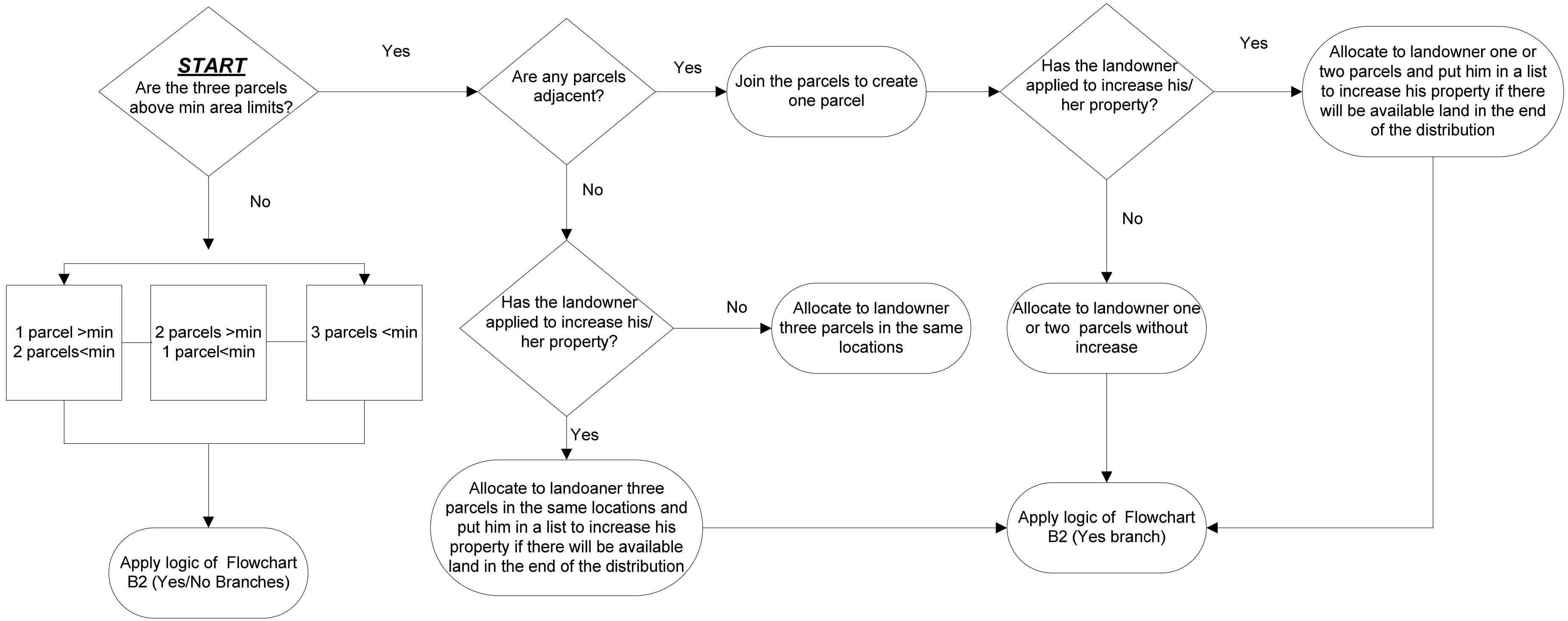


Figure A.5: Flowchart D1: Landowner has three parcels and may take up to three

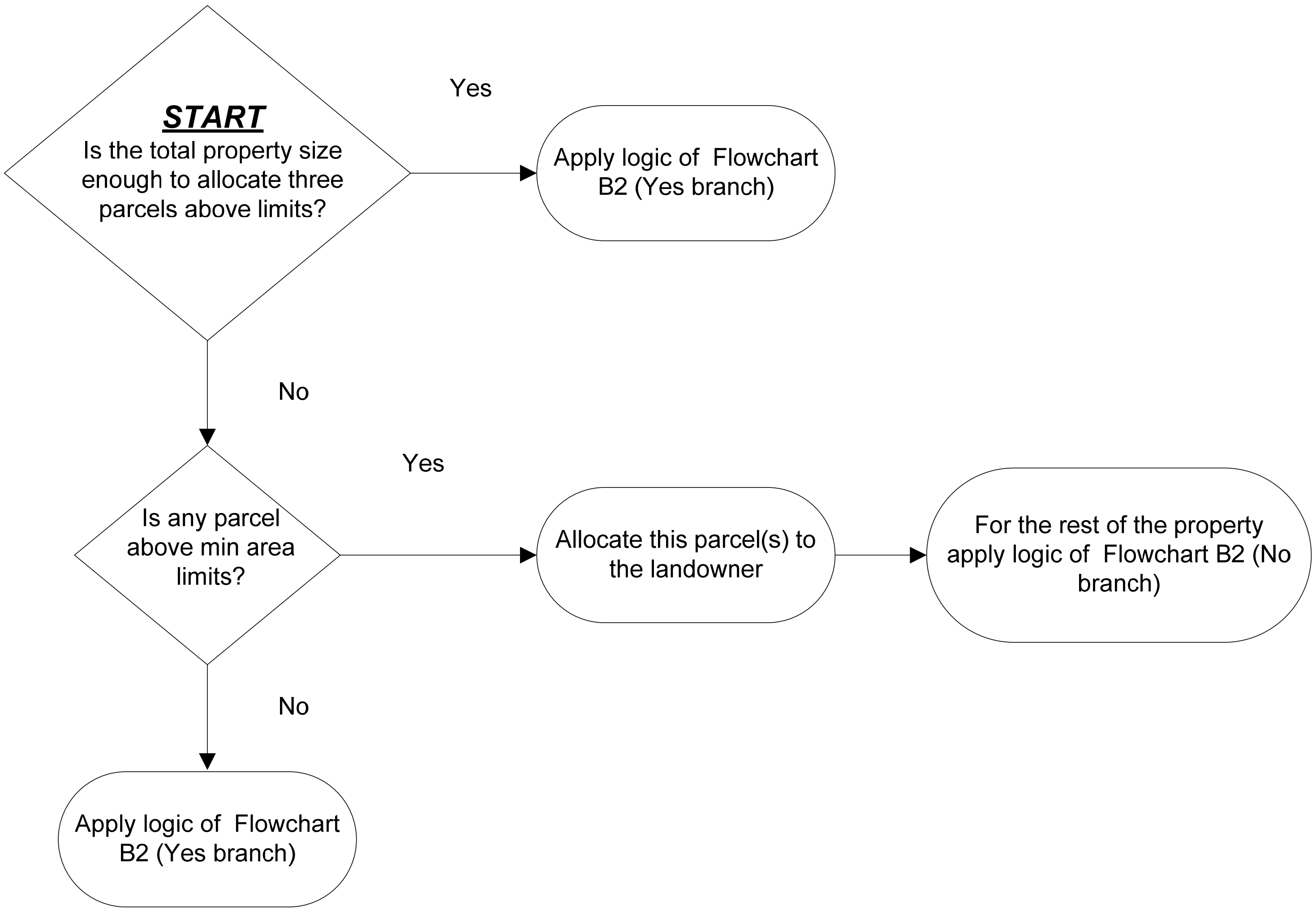


Figure A.6: Flowchart D2: Landowner has more than three parcels and may take up to three

Appendix B: Basic code two rule cluster procedures

B1: Basic code for the main flowchart rule cluster procedure

```
'Start the loop through Owners-Ownership-OriginalParcel tables
Call CalculatePPI

Set pOwnersRow = pOwnersCursor.NextRow
Do Until pOwnersRow Is Nothing
'Get relation1
            Set pDataSetRel = pRelClass1
            Set pRelSet1 = pRelClass1.GetObjectsRelatedToObject(pOwnersRow)
            pRelSet1.Reset

Set pOwnershipRow = pRelSet1.Next
Do Until pOwnershipRow Is Nothing
 'Get relation2
            Set pDataSetRel2 = pRelClass2
            Set pRelSet2 = pRelClass2.GetObjectsRelatedToObjectSet(pRelSet1)
            pRelSet2.Reset

Set pFeature = pRelSet2.Next
Do Until pFeature Is Nothing
Rule 1
   If _
    pOwnersRow.value(intPosTotalAreaField) < CDbl(InputFacts.txtCommMinArea) _
    And pOwnersRow.value(intPosTotalValueField) < CDbl(InputFacts.txtCommMinValue) _
    And pFeature.value(intPosException) = "NO" Then
    pOwnersRow.value(intPosReceiveField) = "NO"
    pOwnersRow.Store
   End If
'Rule 2
   If _
    pOwnersRow.value(intPosCompletionField) = "NO" _
    And pFeature.value(intPosException) = "NO" Then
    pOwnersRow.value(intPosReceiveField) = "NO"
    pOwnersRow.Store
   End If
'Rule 3
   If _
    pOwnersRow.value(intPosTotalAreaField) >= CDbl(InputFacts.txtCommMinArea) _
    Or pOwnersRow.value(intPosTotalValueField) >= CDbl(InputFacts.txtCommMinValue) _
    And pOwnersRow.value(intPosTotalAreaField) >= CDbl(InputFacts.txtLawMinArea) Then
    pOwnersRow.value(intPosReceiveField) = "YES"
    pOwnersRow.Store
   End If
'Rule 4
   If _
    pOwnersRow.value(intPosCompletionField) = "YES" _
    And pOwnersRow.value(intPosTotalAreaField) >= CDbl(InputFacts.txtCommMinArea) _
    And pOwnersRow.value(intPosTotalAreaField) <= CDbl(InputFacts.txtLawMinArea) Then
    pOwnersRow.value(intPosReceiveField) = "YES"
    pOwnersRow.Store
   End If
'Rule 5
   If _
    pOwnersRow.value(intPosCompletionField) = "YES" _
    And pOwnersRow.value(intPosTotalValueField) >= CDbl(InputFacts.txtCommMinValue) _
    And pOwnersRow.value(intPosTotalAreaField) <= CDbl(InputFacts.txtLawMinArea) Then
    pOwnersRow.value(intPosReceiveField) = "YES"
    pOwnersRow.Store
```

```
   End If
'Rule 6
   If _
     pOwnersRow.value(intPosCompletionField) = "NO" _
     And pOwnersRow.value(intPosTotalAreaField) >= CDbl(InputFacts.txtCommMinArea) _
     And pOwnersRow.value(intPosTotalAreaField) <= CDbl(InputFacts.txtLawMinArea) Then
     pOwnersRow.value(intPosReceiveField) = "NO"
     pOwnersRow.Store
   End If
'Rule 7
   If _
     pOwnersRow.value(intPosCompletionField) = "NO" _
     And pOwnersRow.value(intPosTotalValueField) >= CDbl(InputFacts.txtCommMinValue) _
     And pOwnersRow.value(intPosTotalAreaField) <= CDbl(InputFacts.txtLawMinArea) Then
     pOwnersRow.value(intPosReceiveField) = "NO"
     pOwnersRow.Store
   End If
'Rule 8
   If _
     pOwnersRow.value(intPosReceiveField) = "NO" Then
     pFeature.value(intPosEliminate) = "YES"
      pFeature.Store
   End If

Set pFeature = pRelSet2.Next
Loop

Set pOwnershipRow = pRelSet1.Next
 Loop

Set pOwnersRow = pOwnersCursor.NextRow
Loop
```

B2: Basic code for the B2 flowchart rule cluster procedure

```
'Start the loop through Owners-Ownership-OriginalParcel tables

Set pOwnersRow = pOwnersCursor.NextRow
Do Until pOwnersRow Is Nothing
'Get relation1

'Update PPI values for the current land owner
Dim OwnerID As Double
OwnerID = pOwnersRow.value(intPosOwnersField)
Call UpdatePPIs(OwnerID)

Set pDataSetRel = pRelClass1
Set pRelSet1 = pRelClass1.GetObjectsRelatedToObject(pOwnersRow)
pRelSet1.Reset

Set pOwnershipRow = pRelSet1.Next
Do Until pOwnershipRow Is Nothing

'Get relation2
 Set pDataSetRel2 = pRelClass2
 Set pRelSet2 = pRelClass2.GetObjectsRelatedToObjectSet(pRelSet1)

' Assumes pRelSet2 is an ISet
Dim pArray As IArray
Dim pFeature As IFeature
```

```
Dim pFeature2 As IFeature
Dim lFldIdx As Long
Dim bInserted As Boolean
Dim f As Long

' The field no to sort the features
lFldIdx = intPosPPI

' Setup a new array to hold the sorted features
Set pArray = New esriSystem.Array

' Loop through the features in the Set and use an Insert Sort
' to add them to the Array
pRelSet2.Reset

Set pFeature = pRelSet2.Next
While Not pFeature Is Nothing

  bInserted = False

For f = 0 To pArray.count - 1
            Set pFeature2 = pArray.Element(f)
             If pFeature2.value(lFldIdx) < pFeature.value(lFldIdx) Then
                        pArray.Insert f, pFeature
                        bInserted = True
                        Exit For
            End If
Next f

If Not bInserted Then pArray.Add pFeature
  Set pFeature = pRelSet2.Next
Wend

' Now we can loop through the array and the features will be sorted in descending
' order of the specified field no. lFldIdx

For f = 0 To pArray.count - 1
  Set pFeature = pArray.Element(f)

'Apply Flowchart B2

  If _
     pOwnersRow.value(intPosNumParcelsField) > 1 _
     And pOwnersRow.value(intPosMaxNumParField) = 1 _
     And pOwnersRow.value(intPosReceiveField) = "YES" Then
     'Get the OwnerID
     Dim o As Double
     o = pOwnersRow.value(intPosOwnersField)

     Dim ParcelID As String
     ParcelID = pFeature.value(intPosParcelID)

     Dim dblShareFactor As Double
     dblShareFactor = GetShareFactor(o, ParcelID)

     'Get the the OriginalArea owned(A) and calculate the new area (NewA)
     Dim a As Double
     Dim NewA As Double
     a = pOwnersRow.value(intPosTotalAreaField)
     NewA = Round(Calculations.CalculateNewOwnerArea(a), 2)
  ' RULE 1
     If NewA < InputFacts.txtLawMinArea Then
```

```
    NewA = InputFacts.txtLawMinArea

    'Get the the OriginalValue owned(V) and calculate the new area (NewV)
    Dim V As Double
    Dim NewV As Double
    V = pOwnersRow.value(intPosTotalValueField)
    NewV = Round(Calculations.CalculateNewOwnerValue(V), 2)

    'Get the max num of parcels may received by a land owner
    intMaxNumPar = pOwnersRow.value(intPosMaxNumParField)

'Get the Centroid of that Parcel
'Get the LabelPoint(centroid) of each New Parcel

'Call GetCenX and GetCenY functions to get the coordinates of the centroids
Dim CoorX As Double
Dim CoorY As Double
CoorX = CreateParcelsCentroids.GetCenX(pFeature)
CoorY = CreateParcelsCentroids.GetCenY(pFeature)

'Find the BlockID in whcih that Parcel-Centroid belongs to
'Call Function GetBlockIDofParcels (pParcelFeature)
Dim BlockID As Integer
BlockID = CreateParcelsCentroids.GetBlockIDofParcels(pFeature)

'Sum the Area of all NEWPARCELS belongs to that Block to estimate the current occupied area
'Call function SumBlockArea(BlockID)
Dim CurrentBlockArea As Double
CurrentBlockArea = SumBlockArea(BlockID)

'Call Function TotalBlockArea to get the total area of a block
Dim AvailableBlockArea As Double
AvailableBlockArea = GetBlockArea(BlockID) - CurrentBlockArea

'Get the PPI of the parcel
Dim PPI As Double
PPI = pFeature.value(intPosPPI)

Dim ParcelArea As Double
ParcelArea = pFeature.value(intPosArea) * dblShareFactor

Dim ParcelValue As Double
ParcelValue = pFeature.value(intPosValue) * dblShareFactor

'Check if there is available area to locate the parcels
Dim NoParcels As Integer
NoParcels = 0


 'RULE 2
 If NewA <= AvailableBlockArea Then

 'Check if the number of parcels received by an owner reach the maximum accepatble
 'and stop the process

   'Create the new records in the NEWPARCELSLS table
   Call AddRecNewParcelLS(NewA, NewV, CoorX, CoorY, BlockID, PPI)

   'Create the new records in the NEWPARCELSLS table
   Call AddRecNewOwnershipLS(o)

   'Call the CreateLabelPoints procedure to create the centroids on the map
```

```
   Call CreateParcelsCentroids.CreateLabelPoints(pFeature, o)

     NoParcels = NoParcels + 1
'RULE 3
     If NoParcels = intMaxNumPar Then

     Exit For 'i.e. Exit Loop of parcels

     End If
 'RULE 4
 ElseIf NewA > AvailableBlockArea Then
 'i.e. if there is not available area in that Block
 'Get the minPPI value for the certain block

     Dim minPPI As Double
     minPPI = GetMinBlockPPI(BlockID)

   If PPI > minPPI Then
   Dim FlagP As Boolean
   FlagP = True
   Call Reallocation(BlockID, PPI, NewA, NewV, CoorX, CoorY, o, pFeature, NoParcels, FlagP)
  'RULE 5
   If FlagP = False Then
      'Check if there is available land in the next parcels'BlockID
       Call GetNextBlockID(o, BlockID, NewA)
       'Store the new parcel in the NewParcelLS and NewOwnershipLS tables
  'RULE 6
       Call AddRecNewParcelLS(NewA, NewV, 0, 0, BlockID, 0)
       'Note:NEWPARCELS Centroid X,Y are 0 because the new parcel is not located in an exact
       'location in the block. Also, PPI is 0 since this parcel-owner Randomly located to this block
       Call AddRecNewOwnershipLS(o) 'Get a BlockID which has availabe land for the NewParcel
   End If

     If NoParcels = intMaxNumPar Then
       Exit For 'i.e. Exit Loop of parcels
     End If

      End If
   End If
 End If
End If

NextParcel:
Next f

 ' Clean up the array
 pArray.RemoveAll
 Set pArray = Nothing

Set pOwnershipRow = pRelSet1.Next
Loop

Set pOwnersRow = pOwnersCursor.NextRow
Loop
```