Cranfield University

School of Applied Sciences

PhD

Academic Year 2006-2007

Marie Castellazzi

Spatio-temporal modelling of crop co-existence in European agricultural landscapes

Principal supervisor: Dr Gavin Wood
Associate supervisor: Dr Paul Burgess

August 2007

Abstract

The environmental risk of growing genetically modified (GM) crops and particularly the spreading of GM genes to related non-GM crops is currently a concern in European agriculture. Because the risks of contamination are linked to the spatial and temporal arrangements of crops within the landscape, scenarios of crop arrangement are required to investigate the risks and potential coexistence measures. However, until recently, only manual methods were available to create scenarios.

This thesis aims to provide a flexible referenced tool to create such scenarios. The model, called LandSFACTS, is a scientific research tool which allocates crops into fields, to meet user-defined crop spatio-temporal arrangements, using an empirical and statistical approach. The control of the crop arrangements is divided into two main sections: (i) the temporal arrangement of crops: encompassing crop rotations as transition matrices (specifically-developed methodology), temporal constraints (return period of crops, forbidden crop sequences), initial crops in fields regulated by temporal patterns (specifically-developed statistical analyses) and yearly crop proportions; and (ii) the spatial arrangements of crops: encompassing possible crops in fields, crop rotation in fields regulated by spatial patterns (specifically-developed statistical analyses), and spatial constraints (separation distances between crops). The limitations imposed by the model include the size of the smallest spatial and temporal unit: only one crop is allocated per field and per year. The model has been designed to be used by researchers with agronomic knowledge of the landscape. An assessment of the model did not lead to the detection of any significant flaws and therefore the model is considered valid for the stated specifications. Following this evaluation, the model is being used to fill incomplete datasets, build up and compare scenarios of crop allocations. Within the GM coexistence context, the model could provide useful support to investigate the impact of crop arrangement and potential coexistence measures on the risk of GM contamination of crops. More informed advice could therefore be provided to decision makers on the feasibility and efficiency of coexistence measures for GM cultivation.

Key words: crops in fields, crop arrangement, crop rotation, spatio-temporal modelling, landscape scale.

Acknowledgements

I would like to thank my multiple supervisors over the three years for their constant support: Prof. Joe Perry, Dr. Juliet Osborne, Dr. Kelvin Conrad at Rothamsted Research and Dr. Gavin Wood, Dr. Paul Burgess, and Prof. Joe Morris at Cranfield University.

I would particularly like to thank Prof. Joe Perry for giving me the opportunity of working on this project and for providing me with invaluable help with the design of the statistical tests, and Dr. Gavin Wood for allowing me to convert the work into a PhD.

The work undertaken in this thesis was initiated and funded by the European Union funded SIGMEA project (SIGMEA, 2005). This work would not have been achieved without the advice, data and constructive feedback from SIGMEA colleagues from INRA-Jouy-en-Josas, INRA-Grignon, INRA-Dijon, CETIOM, and SCRI.

Many thanks to Joanne Matthews, whose patience and work transformed the LandSFACTS model into proper software.

Finally, I would like to thanks my family and friends for their kind support.

Table of contents

1 Introduction 1

1.1 Background 1

1.2 Aim 4

1.3 Objectives and linked tasks 4

1.4 Deliverables 5

1.5 Thesis route map 5

2 Specifications of LandSFACTS model 9

2.1 End users 9

2.2 Model specifications 10

2.2.1 Purpose and uses 10

2.2.2 Technical aims 11

2.2.3 Modelling approach 11

2.2.4 Conclusion 12

2.3 Conclusion on specifications of LandSFACTS model 12

3 Review of the origins, metrics, and models of crop arrangements 15

3.1 Origin of crop arrangements 15

3.1.1 Environmental constraints 15

3.1.2 Agronomic constraints 15

3.1.3 Farm management constraints 17

3.1.4 Economic, policy and contracts constraints 17

3.1.5 Conclusions on spatial and temporal constraints of crops arrangements . 18

3.2 Existing spatio-temporal metrics of crop arrangements 19

3.2.1 Landscape representation 19

3.2.2 Landscape pattern metrics 20

3.2.2.1 Landscape spatial pattern metrics for categorical maps 20

3.2.2.2 Landscape spatial pattern metrics for spatial points 21

3.2.2.3 Landscape temporal pattern metrics 22

3.2.2.4 General limitations of metrics 23

3.2.3 Particularity of the agricultural landscape 23

3.2.3.1 Field unit 24

3.2.3.2 Field metrics 24

3.2.4 Conclusions on landscape pattern metrics for crop allocation model 25

3.3 Review of existing models 25

3.3.1 Main modelling approaches to crop allocation 26
3.3.2 Modelling components 27
3.3.2.1 Modelling scales 27
3.3.2.2 Model variables 28
3.3.2.3 Farmer decision-making process 29
3.3.2.4 Crop rotations 30
3.3.2.5 Mathematical approaches on processes 31
3.3.3 Conclusions on modelling crops allocations 32
3.4 Conclusion 33

4 Methodology for LandSFACTS development 35
4.1 LandSFACTS model approach: conclusions from review 35
4.1.1 Combining statistical and real variables 35
4.1.2 Crop constraints and patterns terminology 36
4.2 Datasets: from analysis to validation and examples 37
4.2.1 Burgundy dataset 39
4.2.2 Fife dataset 40
4.2.3 Beauce dataset 41
4.2.4 SmallLandSCAPE dataset 42
4.2.5 BigLandSCAPE dataset 42
4.3 Steps for LandSFACTS development 43

5 Measuring the spatio-temporal patterns of crops 45
5.1 New statistical analyses on crops' spatial and temporal patterns 45
5.1.1 Definitions 45
5.1.2 Temporal pattern of crops 46
5.1.2.1 Crop rotation phasing 46
5.1.2.2 Crop phasing 49
5.1.2.3 Crop temporal variability compared to random simulations 51
5.1.3 Spatial pattern of crops 53
5.1.3.1 Fine spatial pattern (chi-square test) 53
5.1.3.2 General spatial trend (E analysis) 57
5.1.3.3 Sensitivity of the E analysis 62
5.1.4 Conclusion on pattern analysis 65
5.2 Crop pattern analyses on landscape datasets 66
5.2.1 Temporal pattern 66
5.2.1.1 Description of crop rotations 66
5.2.1.2 Crop rotation phasing 69
5.2.1.3 Crop phasing 70

5.2.1.4 Crop temporal variability compared to random simulation 71
5.2.1.5 Conclusion on temporal pattern of crops 73
5.2.2 Spatial pattern 74
5.2.2.1 Crop spatial repartition at the study area level 74
5.2.2.2 Spatial repartition of farmers 76
5.2.2.3 Spatial repartition of crops for each farmer individually 77
5.2.2.4 Conclusion on spatial pattern of crops 77
5.2.3 Spatio-temporal pattern 78
5.2.4 Conclusion on spatial and temporal pattern analyses 80
5.3 Statistical analyses to integrate within LandSFACTS model 81
5.4 Conclusion 83

6 Mathematical representation of crop rotations 85
6.1 Mathematical classification of rotations 85
6.2 Rotations as transition matrices 86
6.3 Long-term crop proportions 88
6.4 Transition between rotations 90
6.5 Conclusion 92

7 Description of the LandSFACTS model 93
7.1 LandSFACTS model definition 93
7.1.1 Aim of the model 93
7.1.2 Distinct characteristics of the model 94
7.1.3 General modelling approach 94
7.1.4 Structure of the LandSFACTS model 95
7.2 LandSFACTS initialisation phase: inputs to CropAllocation program ... 96
7.2.1 Crops 97
7.2.2 Fields 97
7.2.3 Possible crops in fields 97
7.2.4 Crop rotations 97
7.2.5 Crop rotation for each field and the optional "RotationFields" program . 98
7.2.6 Initial crops for each field and the optional "InitialCrops" program 100
7.2.7 Crop constraints 101
7.2.7.1 Spatial constraints 101
7.2.7.2 Temporal constraints 102
7.2.7.3 Yearly crop proportion constraints 102
7.2.8 Iteration options and penalties 102
7.2.9 Simulated annealing 105
7.2.10 Simulation parameters 105

7.3 LandSFACTS simulation phase: process of CropAllocation program (allocation of crops to fields)........ 106
7.3.1 Overview of CropAllocation process 106
7.3.2 Problematic-points temporary store........ 107
7.3.3 Overcoming unauthorised crop allocation........ 108
7.4 LandSFACTS outputs and interpretations........ 109
7.4.1 Crop allocation to fields........ 109
7.4.2 Difficulty indexes of the obtained crop allocations........ 109
7.5 LandSFACTS implementation........ 110
7.5.1 Program language 110
7.5.2 Inputs and outputs format........ 110
7.5.3 Model executable and example datasets........ 110
7.5.4 LandSFACTS software........ 111
7.6 LandSFACTS current use 111
7.7 Conclusion........ 112

8 Model assessment........ 113
8.1 Model assessment in the literature........ 113
8.2 Aim of LandSFACTS assessment........ 116
8.3 Method of assessment........ 116
8.4 Assessment of conceptual model........ 117
8.4.1 Temporal and spatial units........ 117
8.4.2 Crop rotations 117
8.4.3 Control on crop spatio-temporal arrangements 118
8.4.4 Landscape as an unique scale 118
8.4.5 Conclusion........ 118
8.5 Code verification........ 119
8.6 Sensitivity analyses 120
8.6.1 Datasets for sensitivity analyses 120
8.6.2 Comparison between one or all crop alterations........ 122
8.6.3 Simulated annealing........ 122
8.6.4 Separation distances........ 128
8.7 Scenario testing for a real landscape 131
8.7.1 Setting up scenarios........ 131
8.7.1.1 Cropping system........ 132
8.7.1.2 Spatial patterns of the rotation with GM oilseed rape........ 132
8.7.1.3 Iteration parameters 135
8.7.1.4 Summary of scenarios 136

8.7.2 Results from scenarios 136

8.7.3 Conclusions 141

8.8 Recommendations on model use 142

8.9 Conclusions on model assessment 144

9 Discussions and Conclusions 147

9.1 Meeting the thesis objectives 147

9.1.1 Objective 1: origin and characteristics of spatio-temporal arrangements 147

9.1.2 Objective 2: design of the LandSFACTS model 148

9.1.3 Objective 3: assessment of the LandSFACTS model 149

9.1.4 Conclusion: from objectives to aims 150

9.2 Thesis major outputs and their advantages 150

9.2.1 Statistical analyses 150

9.2.2 Mathematical representation of crop rotations 151

9.2.3 Model on crop allocation: LandSFACTS 151

9.3 Supporting coexistence scenarios 152

9.3.1 LandSFACTS levels of use 152

9.3.2 Examples of LandSFACTS uses within coexistence context 153

9.4 Possible enhancements for coexistence scenarios 155

9.4.1 Enhancing modelling of crops arrangement 155

9.4.2 New tools specifically designed for coexistence scenarios 157

9.5 Conclusion 159

10 References: 161

Appendix A : Existing crop allocation models 171

Appendix B : Digital Appendices (CD) 174

Appendix C : Database structure 177

Appendix D : Code verification tests 178

Appendix E : Assessment of stochastic processes 179

Appendix F : Crop rotations for sensitivity analyses 183

Appendix G : Results from scenarios 185

Table of figures

Figure 1.1: Schematic diagram of the thesis "route map" 7
Figure 2.1: End-users' requirements for the LandSFACTS model 10
Figure 4.1: Examples of aggregated and homogenous spatial patterns of crops............37
Figure 4.2: Examples of aggregated and homogeneous temporal patterns of crops. 37
Figure 4.3: Location of the study sites through Europe. 38
Figure 4.4: Number, area and location of farmers' fields in Burgundy study site...........39
Figure 4.5: Landscape of the Fife dataset...40
Figure 4.6: Landscape of the Beauce dataset. 41
Figure 4.7: Landscape of the SmallLandSCAPE dataset. 42
Figure 4.8: Landscape of the BigLandSCAPE dataset...43
Figure 5.1: Crop rotation phasing: definition of chi-square test per farmer and rotation. 48
Figure 5.2: Crop phasing: definition of chi-square test per farmer and crop. 50
Figure 5.3: Crop temporal variability: definition of randomisation test per farmer and crop. 52
Figure 5.4: Crop temporal variability: two examples of randomisation points and observed value per farmer and crops, ...53
Figure 5.5: Fine spatial pattern: an example of observed values of the spatial chi-square. 54
Figure 5.6: Fine spatial pattern: definition of chi-square analysis. 55
Figure 5.7: Fine spatial pattern: examples of spatial chi-square results. 57
Figure 5.8: General spatial trend (E analysis): visual representation of the test 58
Figure 5.9: General spatial trend (E analysis): definition of the test. 59
Figure 5.10: General spatial pattern (E analysis): example of aggregation and dispersion of the groups of fields..61
Figure 5.11: General spatial pattern (E analysis): sensitivity analysis on gaps detection. 63
Figure 5.12: General spatial patterns (E analysis): sensitivity analysis on the influence of the number of groups on randomisation curves of E values 64
Figure 5.13: General spatial patterns (E analysis): Influence of the number of groups on E values. 65
Figure 5.14: Burgundy site: fine and general spatial pattern tests of the crops...............74
Figure 5.15: Burgundy site: general spatial pattern test (E analysis on farmers' fields). 76

Figure 5.16: Burgundy site: general spatio-temporal pattern (E analysis), on Farmer 7, rotation 3. 79
Figure 5.17: Burgundy site: general spatio-temporal pattern (E analysis), on farmer 9, rotation 3. 80
Figure 6.1: Mathematical classification of crop rotations. 86
Figure 6.2: Transition matrices of three types of crop rotation. 87
Figure 6.3: Calculation of long-term crop proportions from transition matrices. 89
Figure 6.4: Diagram of transitions between two individual crop rotations. 90
Figure 7.1: LandSFACTS model components and programs 96
Figure 7.2: Overview of CropAllocation program 107
Figure 8.1: Diagram of modelling steps and assessments 114
Figure 8.2: Visual test of generated random numbers within LandSFACTS model....120
Figure 8.3: Impact of simulated annealing values on the number of years with authorised crop allocations within a) Fife dataset and b) Beauce dataset ..123
Figure 8.4: Impact of simulated annealing values on overall penalties within a) Fife dataset and b) Beauce dataset. 126
Figure 8.5: Impact of simulated annealing values on the number of iteration used within a) Fife dataset and b) Beauce dataset. 127
Figure 8.6: Impact of separation distances within a) the Fife dataset and b) the Beauce dataset on the number of overall penalties. 130
Figure 8.7: Crop rotation allocation with aggregated GM rotations (Scenario A). 134
Figure 8.8: Crop rotation allocation with regularly spaced GM rotations (Scenario B). 134
Figure 8.9: Example of crop allocation for a scenario A1, (seed: 3197, year 0) 138
Figure 8.10: Example of crop allocation for a scenario B1 (seed: 27115, year 0) 139
Figure 8.11: Average penalties per fields for scenarios A1 (10 simulations) 140
Figure 8.12: Average penalties per fields for successful scenarios B1 (3 simulations). 140
Figure 9.1: Three different levels of complexity in the use of LandSFACTS model in terms of aims and outputs. 152

Table of tables

Table 3.1: Statistical methods to analyse spatial point patterns. 22
Table 5.1: Example of crops sequence in a rotation...45
Table 5.2: Example of "starting crop" numbering for a crop rotation. 46
Table 5.3: Crop rotation phasing: example data (observed state for chi-square test). 47
Table 5.4: Crop rotation phasing: two examples for the chi-square test 48
Table 5.5: Fine spatial pattern: definition of individual chi-square tests. 56
Table 5.6: General spatial patterns (E analysis): potential applications of the test. 62
Table 5.7: Summary of spatial and temporal statistical analyses 66
Table 5.8: Burgundy site: description of crop rotations. 67
Table 5.9: Burgundy site: proportion (%) of crop transitions from year n to n+1. 68
Table 5.10: Burgundy site: results of crop rotation phasing test (chi-square results) 69
Table 5.11: Burgundy site: crop phasing test (chi-square probabilities). 70
Table 5.12: Burgundy site: results from crop temporal variability test 72
Table 5.13: Burgundy site: general spatial patterns test (E analysis results on crops). ...77
Table 6.1: Transition matrix between two crop rotations 91
Table 6.2: Overall transition matrix between two crop rotations (U_{TM}) 91
Table 8.1: Summary of the Fife and Beauce datasets for the sensitivity analyses 121
Table 8.2: Comparison between the successes of crop allocation using iteration options based on the alteration of random choices (option 1 and 2.1) 122
Table 8.3: Characteristics of crop rotation allocations for scenario A and B (long-term crop proportions, levels of spatial patterns, and proportions for each rotation). 133
Table 8.4: Number of neighbouring fields with GM or non-GM rotations for scenarios A and B. 135
Table 8.5: Summary of characteristics of the four scenarios. 136
Table 8.6: The proportion of successful simulations, the number of iteration and of penalties (statistical significance shown) in obtaining crop allocation for each scenario..137

Abbreviations and symbols

Mathematical abbreviations only valid for one statistical test are not included here; their comprehensive definition is given within the definition of the test (i.e. in the same Figure).

%CV	Percentage of coefficient of variation of the proportion of fields or area of a farmer with a specific crop during as specific time period.
%R	Percentage of randomised %CV, which are higher than the observed %CV.
a-sa	Autumn set-aside
CETIOM	Centre technique interprofessionel des oleagineux metropolitains
CVp	Proportion of randomly simulated %CV values lower than the observed %CV value (values ranging from 0 to 100%). This value is used to control the temporal patterns within LandSFACTS.
E	Result value of the E analysis
E analysis	Statistical analysis of spatial patterns of crops
Ep	Percentage of randomly simulated E values lower than the observed E value. This value is used to control the spatial patterns within LandSFACTS.
E_x	Within chi-square tests: expected value
EU	European Union
GIS	Geographical Information System (spatial information)
Genesys	Gene flow model focusing on cross-pollination of rapeseed (Colbach *et al.*, 2001a; Colbach *et al.*, 2001b)
GM	Genetically Modified (crops)
ha	Hectare
H_0	Null hypothesis in a statistical test (no statistical difference between two sets of data)
H_1	Inverse of the null hypothesis in a statistical test (statistical difference between two sets of data)
ID	Unique identifier (unique integer value)
Initial crop	First crop within a crop rotation to be grown during the simulation (year 0).
INRA	Institut National de la Recherche Agronomique
km^2	Kilometres square

LandSFACTS	Landscape Scale Functional Allocation of Crops Temporally and Spatially.
le	Length of rotation
l.d.wheat	Late drilled wheat
m	General: meters
m	Within statistical tests: number of fields (same farmer and same rotation)
MAPOD	Gene flow model focusing on cross-pollination of maize (Angevin *et al.*, 2007; Angevin *et al.*, 2001)
n	Within chi-square tests: number of samples for calculating the degree of freedom of the statistical test
O_x	Within chi-square tests: expected value
OSR	Oilseed rape
p	Probability determined using a statistical test
p-sa	Permanent set-aside
R	Rye
R1 or R2	Rotations 1 or rotation 2
sa	Set-aside
sB	Spring barley
SCRI	Scottish Crop Research Institute
Seed	Within random number generation: first value used for generating pseudo-random numbers within a simulation.
SIGMEA	Sustainable Introduction of Genetically Modified crops into European Agriculture (SIGMEA, 2005)
TM	Transition matrix (in the context of crop rotation)
U_{TM}	Overall transition matrix (i.e. transition matrix between transition matrices of crop rotations)
UPS	Université de Paris Sud
W	Wheat
wW	Winter Wheat
wB	Winter barley
χ^2	Chi-square value of a chi-square test

1 Introduction

1.1 Background

Agriculture has been through multiple changes during 20th century. After the Second World War, agricultural production was accelerated by the "industrialisation" of agricultural activities in order to meet the increasing demand for food. Agriculture is now facing a new challenge, as the public and political awareness of environmental issues is increasing. Under the new European Common Agricultural Policy's Single Payment Scheme (SPS), the subventions are decoupled from food production (DEFRA, 2004b), and farmers must meet cross-compliance requirements, by keeping their land in Good Agricultural and Environmental Condition (DEFRA, 2005b). Thus, the farmer's role has shifted from being solely a food producer to include other rules such as being a steward of the environmental and the landscape.

A new environmental concern in agriculture has arisen from the possibility of growing genetically modified (GM) crops (Firbank *et al.*, 2003; Wilkinson *et al.*, 2005). The possible health consequences of introducing GM crops in the food chain are only one part of the GM issue. The environmental risks of growing GM crops in an open agricultural environment are also of great importance, particularly the risks relating to the "contamination" of non-GM crops (e.g. conventional or organic) or the spreading of undesired genes to wild relatives (herbicide resistant genes), (Timmons *et al.*, 1996). The understanding and mitigation of those risks is of prime importance for decision-making on coexistence rules for GM crops.

Genes from GM crops have two main means of dispersal: by the seeds and by pollen (Bock *et al.*, 2002). Extended seed dispersion may occur particularly during harvest activities, as a certain proportion of seeds are always lost during this activity either on the field itself or during the travel from the field to the seed store (Bock *et al.*, 2002). Then the following year, volunteers may grow up in the field, in field borders, or on road borders en route to the store. Seeds may also stay dormant for a few years before germinating. The risk of GM resurgence can therefore be present for a very long time after the last GM crop was sown (Lutman *et al.*, 2005). The volunteers may then, in turn, be a new source of GM contamination to the environment, either by seed or pollen dispersal (Squire *et al.*, 2003).

Pollen from GM crops may cross-pollinate with receptive plants, which can either be wild relatives or conventional crops of the same family (Bock *et al.*, 2002; Eastham and Sweet, 2002). GM genes may thus spread to other crops. In the example of GM Oilseed Rape (OSR) resistant to herbicide, their pollen may contaminate wild mustard flowers or another OSR field. In the first case the wild relative may acquire the pesticide resistance and thus may become a weed much harder to control; in the second case, OSR plants which are supposed to be GM-free may become "contaminated" which will reduce their market value. The issue is particularly important to organic farmers, because the purity of their product is one of their main selling points.

OSR pollen transport is mainly conducted by the wind and insects, at both short and longer distance (Bateman, 1947b; Bateman, 1947c), e.g. cross-pollination might occur at more than 5km from the source (Ramsay *et al.*, 2003). **Through both dispersal mechanisms, the landscape in which the GM is grown can determine the extent to which the genes may spread**. The term "landscape" refers to the agricultural area with its fields, cropping systems, and infrastructures. In particular, the relative spatial arrangement of the source of contamination and of the receptive area is important, because if seeds or pollen cannot reach a recipient (good soil or compatible plant) the risk of contamination is null. The pattern of seed dispersal is highly dependent on the road routes taken by the seed trailers. Pollen dispersal is much more dependent upon the landscape structure and more particularly upon the crop arrangements within the landscape, as its dispersal will be influenced by natural obstacles, such as distance between fields, or hedgerows, forest and buildings (Hunt *et al.*, 2001). The temporal arrangement of source of contamination and of receptive crops is also important, because from a contaminated seed bank, GM ferals may grow and thus become a new source of contamination several years after the GM crop was grown.

The modes of dispersion of GM genes are varied and the risks encountered are highly dependent upon the landscape and more particularly the spatial and temporal arrangements of contamination source (e.g. GM OSR) and receptive areas (e.g. conventional or organic OSR), (Bateman, 1947a; Klein *et al.*, 2006). To limit the risks of undesired GM dispersion, coexistence measures, such as separation distances between GM and non-GM crops, have been considered (DEFRA, 2006).

One method for analysing the magnitude of the environmental risk of GM genes spreading through agricultural landscapes with and without coexistence measures is to use models of gene flow at the landscape scale. One such model is Genesys, which aims to "evaluate the influence of cropping systems on transgenes escape from rapeseed crops to rapeseed volunteers in time and space" (Colbach *et al.*, 2001a; Colbach *et al.*, 2001b). For every year of simulation, the model considers the crops in every field of the landscape. Depending upon the study, the landscape spatial extent might be very large in order to encompass all risk of contaminations (i.e. several kilometres for cross-pollination). As the specific repartition of crops highly influences the spreading of GM genes, being able to run gene flow models on multiple landscapes with similar spatio-temporal characteristics of the crops would strengthen the conclusions drawn from gene flow modelling. Moreover, to test possible coexistence measures, crop arrangements meeting those conditions are required.

Currently, the creation of scenarios of crop spatio-temporal arrangement is carried out by manually altering the crops in the landscape one by one. This method is highly biased by personal decisions, time consuming and not easily reproducible. The only models simulating crop allocation have a mechanistic approach integrating large amounts of specialised information, e.g. soil nutrients, weather or farm management information. Such models have very limited usefulness for scenario building of crop arrangements, because sufficient data are rarely available.

To facilitate the generation of crop spatio-temporal arrangement scenarios for gene flow models, a research model solely aiming at allocating crops into fields to meet specified targets of crop spatial and temporal arrangements at the landscape scale, would be an asset. Such a model would simulate crop arrangements directly using an empirical and statistical approach, instead of modelling the mechanistic origin of the arrangements. Thus crop arrangement scenarios would be easily created with only minimum specialised inputs, however general knowledge of the agronomic conditions, such as crops and rotations, are still required by the user.

The subject of this thesis is precisely aimed at providing just such a model, integrating the requirements detailed above. **To summarise, this model, hereafter called LandSFACTS (Landscape Scale Functional Allocation of Crops Temporally and Spatially), will greatly facilitate and strengthen scientific investigations on the risks of GM contamination in the agricultural landscape with and without coexistence measures, by providing tailored scenarios of crops spatio-temporal arrangements. From those scientific investigations, researchers can then advise decision makers on required coexistence rules for GM crops cultivation.**

The initial motivation for the research in this thesis came from the EU-funded SIGMEA (Sustainable Introduction of Genetically Modified Crops into European Agriculture) project (SIGMEA, 2005), which is studying coexistence of GM, conventional and organic agricultural systems in European agriculture.

1.2 Aim

This research aims to support the investigation of GM crop coexistence scenarios in European arable landscapes by providing a modelling framework, the LandSFACTS model, to create and manipulate realistic scenarios of crop spatio-temporal allocations. The model uses a stochastic approach to simulate crop arrangements in fields at the landscape scale, whilst respecting empirical and statistical user-defined constraints that represent the predominant agronomic, socio-economic and political conditions.

1.3 Objectives and linked tasks

The objectives of this research project and the linked tasks are listed below.

Objective 1. Examine the origins and characteristics of spatio-temporal arrangements of crops in agricultural landscape.

Task 1.1 To review the literature on the origins and measurements of spatio-temporal arrangement of crops, in order to identify the constraints on crop arrangement and existing statistical analyses on patterns.

Task 1.2 To develop and set up relevant statistical analyses on crop patterns.

Task 1.3. To analyse the spatial and temporal pattern of crops in relevant study landscape, in order to determine parameters and statistical tests relevant to the LandSFACTS model.

Objective 2. Design the LandSFACTS model of crop arrangement with its components and processes, in order to create a flexible and generic model.

Task 2.1. To review the requirements for the LandSFACTS model and existing models in the literature.

Task 2.2. To define the system representing the LandSFACTS model, i.e. limits of the system, components and main processes involved within the system.

Task 2.3. To design and set up a flexible and generic structure for the LandSFACTS model.

Objective 3. Assess how well does the LandSFACTS model fulfil its objectives?

Task 3.1. To determine the methodology for the model assessment

Task 3.2. To assess the model approach, structure and implementations in relation to the stated specifications.

Task 3.2. To carry out a sensitivity analysis and scenario testing of the model.

1.4 Deliverables

The three deliverables listed below, will be provided by this project.

Deliverable 1: To provide constraints rules on crops arrangement and statistical analyses to characterise spatial and temporal patterns of crops from an agricultural landscape. This deliverable is presented in Chapter 4, 5, and 6. It will contribute towards fulfilling objective 1.

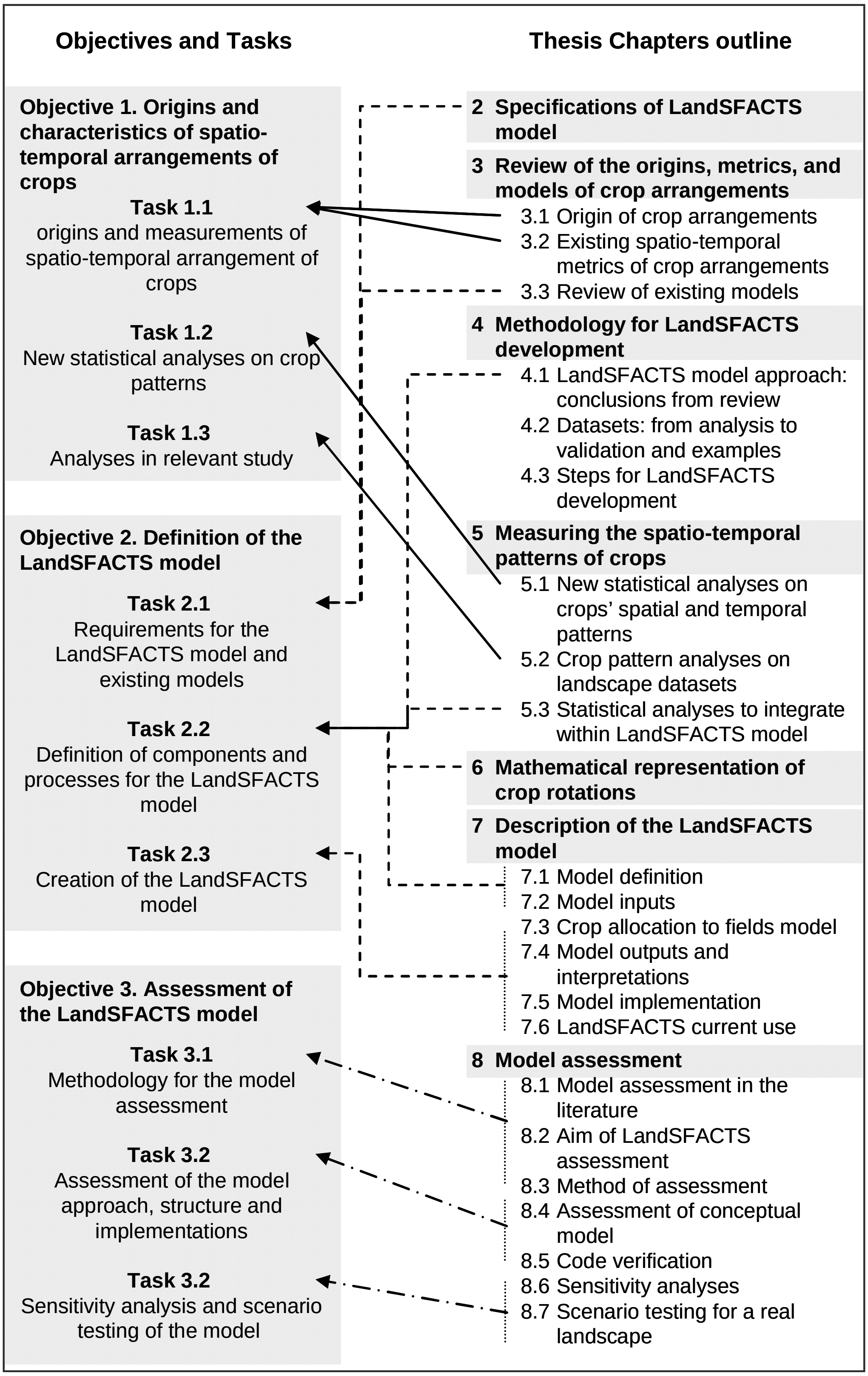
Deliverable 2: To provide the LandSFACTS model facilitating the investigation of landscape scenarios on specific spatial and temporal arrangement of crops by researchers with agronomic background. The model should integrate spatial and temporal patterns of crops, while respecting specific spatial and temporal constraints (e.g. crop rotations or isolation distances between crops). This deliverable is presented in Chapter 7. It will contribute towards fulfilling objective 2.

Deliverable 3: To provide an assessment of the LandSFACTS model against its objectives. This deliverable is presented in Chapter 8. It will contribute towards fulfilling objective 3.

1.5 Thesis route map

To reach the aim of this thesis, the three objectives laid out in Section 1.3 (Objectives and linked tasks) must be achieved. Objective 1 (examine of the origin and characteristics of crop spatio-temporal arrangements) and Objective 2 (design of the LandSFACTS model) are entwined. Objective 2 sets the framework for the project in Task 2.1 by identifying the specification of the LandSFACTS model, whereas Objective 1 provides the background to the project by identifying the origin of spatio-temporal crop arrangements from the literature (Task 1.1), by providing specifically designed tools to characterise crop arrangements were set up (Task 1.2) and by providing insights on crop arrangement in a real landscape using the new statistical tools (Task 1.3). The tools and knowledge derived from achieving Objective 1 were then incorporated within the LandSFACTS model through Task 2.2. Objective 2 was completed with the creation of the LandSFACTS model (Task 2.3). Then Objective 3, the assessment of the model, was investigated through its three tasks: (i) methodology of the model assessment (Task 3.1), (ii) assessment of the model approach structure and implementations (Task 3.2), and (iii) sensitivity analysis and scenarios testing of the model (Task 3.3).

The next chapters follow the same logical order as exemplified above. Chapter 2 defines the framework of the project by defining the end-users requirements of the LandSFACTS model (Task 2.1). Chapter 3 investigates relevant published literature on (i) the origins of crop arrangements (Task 1.1), (ii) existing statistical analyses measuring crop arrangements (Task 1.1) and finally on (iii) existing models simulating them (Task 2.1). The conclusions from the literature review (Chapter 3) inform the lay out of the methodology (Chapter 4) to be followed for the design of the LandSFACTS model (Task 2.2), and the datasets to be used for setting up and assessing the model are detailed. In Chapter 5 and 6 specifically designed tools to be integrated within the model (Task 2.2) are presented: (i) statistical analyses of crop patterns and knowledge on crop arrangement characteristics from a real landscape (Chapter 5, Task 1.2 and 1.3, and thus fulfilling Objective 1), and (ii) mathematical representation of crop rotations (Chapter 6). Then in Chapter 7, the model itself is presented with its components, simulation processes, outputs and information on their technical implementation (Tasks 2.2 and 2.3 fulfilled). This chapter concludes the objective 2. The model is then assessed to evaluate against the model specifications (Chapter 8), fulfilling objective 3. A graphical representation of the project steps and their links with meeting the objectives and tasks is shown in Figure 1.1.



<u>Figure 1.1: Schematic diagram of the thesis "route map"</u>

The line styles are different for each objective

.

2 Specifications of LandSFACTS model

To develop the required framework to support scenarios of crop arrangement, the first task of the project was to identify the requirements of the end-users within the context of GM coexistence. After identifying the end users, the specifications of the LandSFACTS model are reported in this chapter by focusing on three main points: (i) the purpose and uses of the model, (ii) the technical aims and (iii) the modelling approach. Those specifications for the LandSFACTS model are the basis of the thesis project, presented in later chapters.

2.1 End users

The targeted end-users of the LandSFACTS model are agronomic researchers working with gene flow models (e.g. Genesys for rapeseed and MAPOD for maize (Angevin *et al.*, 2007; Angevin *et al.*, 2001)), and more particularly within the SIGMEA project (SIGMEA, 2005). They are researchers at INRA-Dijon, INRA-Grignon, CETIOM, SCRI, and Bremen University. The model must specifically meet their requirement of facilitating the setting up of scenarios of coexistence measures between GM and non-GM varieties of a crop.

The main requirements for the model, described in Figure 2.1, were defined in collaboration with end-users at SIGMEA meetings and through emails from October 2004 until April 2006. Specific requirements from end-users were considered until the end of the LandSFACTS project (June 2007) in order to finely tune the model to the end-users needs. The requirements were further complemented with comments from colleagues from Cranfield University, and Rothamsted Research. Wherever possible the model was designed to be generic, to broaden its usefulness to applications other than scenarios for gene flow models.

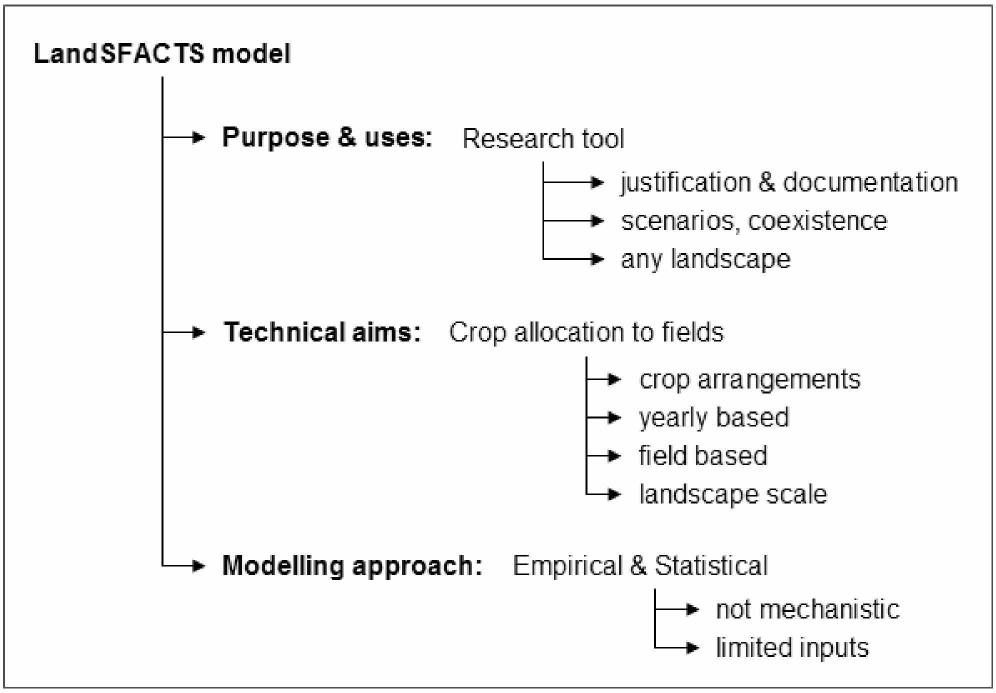


Figure 2.1: End-users' requirements for the LandSFACTS model

The end-user requirements are based on three main considerations: (i) the purpose and uses of the model, (ii) the technical aims and (iii) the modelling approach (Figure 2.1). Each of those considerations on the requirement for the model are detailed below.

2.2 Model specifications

2.2.1 Purpose and uses

The discussions with the end-users highlighted that the model must be a research tool, usable by agronomic researchers. The model could also be used by environmental consultants, although they are not targeted primary end-users. As a research tool, the model must allow the user to control the behaviour of the model, such as how the crops are allocated to fields, and also stochastic processes. Moreover, the model approach, structure and processes should be fully justified and documented.

The end-users indicated that the model needs to be able to be used for building and testing scenarios of crop allocations, and more particularly for coexistence scenarios. Hence it can be used to model the possible introduction of GM crops within agricultural landscapes and help to predict the impacts of growing GM crops with and without coexistence measures, which aim to mitigate potential risks. Therefore the model should allow the user to control (i) the crop proportions over the years (to model the introduction of a new crop or variety), (ii) crop separation distances (e.g. the distance required between GM and conventional varieties of a crop), and (iii) return period of crops (e.g. conventional variety of a crop cannot be grown the year after GM variety).

The model must be useable on any European agricultural landscape, as it is specified by the SIGMEA project (2005), which is EU funded. Therefore, the model must be able to account for the diversity in European agricultural landscapes. For example, the crops available within the model should not be intrinsic to the model. The model should be as flexible as possible, to prevent any restrictions on its future use.

2.2.2 Technical aims

The main function of the software, as specified by the end-users, was to allocate crops to fields over several years, while respecting constraints on crop arrangements. Therefore the software should be able to provide a crop allocation from one year up to 20 years. A one-year crop allocation is required for gene flow models temporally restricted to one season. For example, in Europe, maize does not survive the winter; therefore gene flows are only modelled within a year. The maximum number of years given here, 20 years, is only used as a guideline. In Europe, there is usually only one main crop grown each year, therefore the time step of the model is one year.

Crops are usually grown within fields with fixed boundaries over years. Physical boundaries, where present, can be hedgerows, barriers or roads. The field is the assumed unit of crop cultivation, as used by the gene flow model Genesys. Fields are represented by polygons with specific coordinates and the vector format is chosen as the best representation of the fields.

Lastly the end-users requested that the model must provide crop allocation at the landscape scale. The definition of the size of a landscape is not universally defined within the literature. Within the scope of analysing risks of gene flow contamination, a landscape may have from two fields up to 5,000 fields.

2.2.3 Modelling approach

The LandSFACTS model aims at providing a simple and easy way of creating scenarios of crop arrangement. The reality of farmers' decisions on crop allocation is a complex process, not completely understood or predictive, which involves environmental, agronomic and socio-economic parameters, such as nutrient flow, pest management, farm workload, and market prices. Mechanistically modelling this decision process requires a high quantity of detailed inputs, which often impedes the use of such models. Therefore the approach used in this research project was to directly simulate the crop arrangements instead of reproducing the decision-making process leading to it. To achieve this, the complex decision process leading to crop allocation is replaced by (i) stochastic decisions, (ii) empirical constraints limiting crop arrangement (e.g. return period of crops, separation distances between crops), and (iii) statistical measures of crop arrangements (e.g. general patterns). By using this approach, complex and extensive environmental and socio-economic variables and processes are replaced by a limited number of variables directly influencing the crop arrangements. By using conclusions and insights from research on farmers' decision making, the user may determine the inputs of the LandSFACTS model. The LandSFACTS model can be defined as a shortcut tool to create a unique crop allocation reproducing the conclusions on crop arrangement drawn from research on farmers' decisions.

2.2.4 Conclusion

The requirements detailed in the previous paragraphs, set up the framework in which the LandSFACTS model was developed. In summary, the end-users indicated that the model should be a research tool able to be used to build scenarios of crop arrangements at the landscape scale on any European agricultural landscape. Crops should be allocated to the fields (polygons) over the years using a one year time step, and an empirical and statistical approach should be used to directly simulate the crop arrangements and not the decision making process leading to it.

The setting up of the inputs requires from the user an extensive knowledge on the agronomic and socio-economic situation of the study site. The correct interpretation of the LandSFACTS reports requires a good knowledge of the processes occurring within the model.

2.3 Conclusion on specifications of LandSFACTS model

This chapter details the rationale and specifications for the model. They were defined in collaboration with agronomic researchers working on GM coexistence scenarios. In summary, the model, called LandSFACTS, had to be a research tool to build scenarios of crop arrangement at the landscape scale on any European agricultural landscape, within the context of GM coexistence. The model had to allocate crops to fields (polygons) over the years using a one year time step. The general modelling approach had to be empirical and statistical by directly simulating the crop arrangements, and not the decision making process leading to it. The definition of the specification of the model provided the framework, in which the LandSFACTS model had to be developed.

The next step was to analyse the system to model. Therefore, a literature review is presented in Chapter 3 on (i) the origin of the crop arrangement, and more precisely the constraints influencing it, (ii) means of statistically measuring crop patterns, and (iii) existing models on the same topic. From those analyses, the specific needs of the LandSFACTS model can be determined and the main approach of the model will be presented in Chapter 4.

3 Review of the origins, metrics, and models of crop arrangements

In the previous chapter (Chapter 2), the specifications of the LandSFACTS model compiled from end user requirements, identified the framework in which the model had to be developed. Based on those specifications, the literature was reviewed for relevant studies on (i) the origin of crop arrangements, (ii) existing metrics measuring crop arrangements, and (iii) existing models simulating crop arrangements. Each of those subjects is detailed in this chapter, and relevant conclusions are drawn for the methodology to develop the LandSFACTS model (Chapter 4).

3.1 Origin of crop arrangements

Crop arrangements in the agricultural landscape are influenced by a combination of environmental drivers, farming activities and socio-economic considerations. Those constraints on farming systems are reported in many studies (Papy *et al.*, 1988; Rellier and Marcaillou, 1990). In this chapter, four main constraints on crop cultivation will be investigated: (i) environmental, (ii) agronomic, (iii) farm management constraints, and (iv) economic, policy and contracts. Their relative effects on crop spatio-temporal arrangement will be examined.

3.1.1 Environmental constraints

Each crop and crop variety has its own range of environmental variables (Brady and Weil, 2002), in which the crop is considered to be the most profitable (best quality and highest yield). The crop cultivation within a landscape is thus influenced by its environmental constraints. The most often cited factors are the climate (including rainfall characteristics, solar radiation intensity, and temperature), the soil properties (including proportions of clay, sand, silt and organic matter) and water supply (surface or underground water). Environmental variables usually tend to vary gradually at a landscape scale, although rapid changes may be possible due to changes in topography.

3.1.2 Agronomic constraints

Agronomic rules aim at improving the management of crops and soil for a more efficient agriculture, i.e. higher yields at lower cost relative to time, economics, risks, and the environment. One of the oldest and most fundamental agronomic practices worldwide is crop rotation (Lawes *et al.*, 1895). It is thought to have been critical to the industrial revolution in Britain (Brunt, 1999). Crop rotation is defined as the successive growing of crops on a specific field (Wibberley, 1996). The crop sequences result from land managers decision aiming at optimising agronomic, environmental, and financial objectives while considering constraints from regulations, contracts, and risk management (Kirkegaard *et al.*, 2004; Tarim *et al.*, 2006).

There are four agronomic rules which structure crop rotations aimed at optimising the crop yields, pest and weed control, and facilitating farm management. The first rule is the return period of crops or group of crops. This rule enforces the alternation of crops in order to break the cycle of the build-up of nematodes or other soil pests (Jones and Perry, 1978). The second rule is linked to the benefits or risks of growing a specific crop immediately after another one. The benefits could arise from increased nitrogen supply, soil organic matter or water availability, improvements in soil structure, and decreased pests, diseases or weed competitions (Berzsenyi *et al.*, 2000). For example, in the UK, volunteer cereal weeds are particularly an issue if cereals are followed by autumn-sown vegetable crops (Bond *et al.*, 2006). The third rule is linked with within-year cycles, i.e. usually a crop may only be sown after the previous crop has been harvested. The sowing and harvest timing fluctuates with climatic conditions. For example, late harvesting due to low temperature or autumn rainfall can restrict autumn cultivations to such an extent that the following crops will perforce be spring- rather than autumn-sown. This constraint is exemplified by the higher prevalence of spring- over autumn-sown oilseed rape in arable rotations in Scotland in comparison to England (Champion *et al.*, 2003). The fourth rule relates to the crop proportions on a field or group of fields. Typically farmers have a limited amount of dedicated machinery and labour, therefore they often seek to spread out the work over the year, by growing a range of crop with different requirements. Growing a range of crops also spread the risk of total crop failure or dependence upon market prices and thus limit the risks of economical loss (Lockie *et al.*, 1995).

Within the UK and probably throughout the EU non-organic sectors, there is a diminution of the strict use of rotations. As margins are squeezed, market forces dominate and the trend is towards greater flexibility and less-structured rotations. No longer is most of the arable land in Britain “cultivated according to regular and well-recognised successions or “rotations” of crops” (White, 1929).

Farmers may also have to preserve distances between two specific crop cultures to avoid contamination. For example, sweet corn must be separated from grain corn by at least 300m (Sausse, 2005). Seed production requires severe distance restrictions in order to guarantee the purity of the seeds. The European Union requires a distance of 500m between oilseed rape for the seed production from hybrids and other sources of possible pollen contamination (European Union, 1966). These separation distances between crops directly influence the spatial arrangements of crops. As shown above, spatial and temporal arrangement of the crops are both taken in account for the organisation of crop allocations to fields.

3.1.3 Farm management constraints

The structure and organisation of farming systems imposes constraints on the crop arrangements. Firstly, the number, size, and shape of the fields influence the mosaic of crop arrangements. The regularity of the shape of the field is important for crop management operations (Thenail and Baudry, 2004), as machinery has a fixed width. The distance between fields and farmstead or food processors is an important factor for the accessibility of farm equipment to the fields. The type of access or road should not be underestimated, particularly if some crops require heavy and wide equipment or a harvester (Thenail and Baudry, 2004), or, in the case of the sugar beet if trucks have to collect them in the field directly. The labour and machinery resources are usually limited but can be complemented with contractor work for short periods of time, for example at harvest. Due to this limitation, farms try to spread farm workload as much as possible through the year, and fields spatially close may have the same crops in order to simplify management of those fields.

The constraints arising from the farm are generally under the control of the farmer, except for the location of food processors and factories. However farmers may be grouped into cooperatives or several farms may be managed by only one of the farmers (Orson, 2005). In those particular cases, further communal constraints may emerge. The farm management constraints further limit the spatial and temporal arrangement of crops within the landscape.

3.1.4 Economic, policy and contracts constraints

The economic constraints are independent of the farmer, they are imposed on the farming systems. The fluctuating market prices of crops influence the farmer's interest in specific crops. A crop having an increasing market price will be grown more often (temporal extension) and more widely (spatial extension). Some crop production, such as sugar beet, has to follow a quota. The variation of the level of the quota, such as is currently planned by the EU, affects widely the profitability of growing sugar beet, particularly for smaller holdings (DEFRA, 2004a). Subsidy policy also influences crop cultivations particularly with the recent shifting from a support of production of milk, meat and cereals to a reward for environmental management of their land (environmental stewardship) (DEFRA, 2005a). Contractors, such as sugar beet factories or processors using potatoes for chips are highly demanding in terms of time delivery of the products, their quantity, and their quality. Environmental legislation must also be respected such as water restrictions and the limitation of diffuse pollution into the water body (rivers or ground water) are a high priority for governments (DEFRA, 2005a). Separation distances between GM and non-GM crops in the case of coexistence, may soon be implemented by policy makers.

Policy may be set up by local government, national government or at the European level. The economic, policy and contracts constraints are subject to changes through time, which may or may not be predictable. Moreover, farmers are often bound with investments and loans for machinery or infrastructure, which may slow down adaptation to new economic, policy or contracts constraints.

3.1.5 Conclusions on spatial and temporal constraints of crops arrangements

The constraints exemplified above influence the arrangement of crops within agricultural landscapes. The temporal arrangements of crops are principally driven by agronomic constraints through crop rotation, and crop market prices. Spatial arrangements are mainly altered by environmental constraints. The spatial range of crops changes through the landscape with environmental conditions (e.g. soil characteristics or topology). Spatially close locations tend to have similar ranges of crops (spatial dependency). Landforms were noted as being an important factor on landscape patterns (Swanson *et al.*, 1988; Turner, 1990). Policy may impose specific conditions on crops' spatial patterns, particularly in the case of separation distances to avoid cross-pollination within the same crop species (e.g. in the cultivation of seeds-crops or genetically modified crop). Constraints on crops patterns are also linked to specific scales: (i) field scale for environmental and agronomic constraints; (ii) farm or group of farms scale for farm and contracts constraints; (iii) national scale for economic and policy; (iv) European scale for European policy and the Common Agricultural

Policy. Thus the constraints considered and their respective weights are dependent upon the spatial scale of the study.

The consequences of those conclusions on the design of the LandSFACTS model are detailed in Chapter 4: Methodology for LandSFACTS development, p.35.

3.2 Existing spatio-temporal metrics of crop arrangements

Many indices characterising spatial and temporal patterns within landscapes have been developed, particularly within landscape ecology (McGarigal, 2002), and specialised software facilitate their use such as Fragstats (McGarigal, 2002; McGarigal and Marks, 1995) and GRASS-r.le (Baker, 2001; Baker and Cai, 1992). Those metrics are dependent upon the representation of the landscape and the type of data analysed. In the following sections, ways of representing the landscape and metrics, and their respective relevance to the LandSFACTS model are reported.

3.2.1 Landscape representation

Depending upon the subject of study, a landscape may be thought of, and represented as, a continuous or discrete environment. A continuous representation of the landscape is commonly used to map variables without sharp boundaries, such as land covers or rainfall, whereas a discrete environment is more adequate to represent abrupt changes within the landscape, such as buildings or water courses. Both approaches have their dedicated GIS format, (i) for a continuous environment, the space is arbitrarily divided into square grid cells called raster format; (ii) for a discrete environment, specific geographical features are individually represented as points (e.g. individual trees), line (e.g. rivers), or polygons (e.g. buildings, fields). Landscape ecology research favours a continuous representation of the landscape, particularly as raster datasets are more readily available from satellite imagery (e.g. CORINE dataset).

Independently of the format of the landscape, the spatial data types may be classified into four main categories (McGarigal, 2002): spatial point, linear network, surface, categorical map.

Very often "landscape metrics" only refers to categorical map pattern (McGarigal, 2002). However, categorical maps present two difficulties; they tend to ignore variation within spatial units, and any continuous trends in the landscape (e.g. wind effect on airborne pollution) (Gustafson, 1998). As noted by Gustafson (1998) the combination of different types of data and particularly the combination of spatial points and categorical maps provide more complete information on the patterns and on the scale of patterns.

3.2.2 Landscape pattern metrics

A multitude of landscape spatial pattern metrics have been defined by a wide range of authors (Baker, 2001; Cullinan and Thomas, 1992; Fu and Chen, 2000; Gustafson, 1998; McGarigal, 2002; McGarigal and Marks, 1995; Parker and Meretsky, 2004; Remmel *et al.*, 2002). Mosaics of land use are often treated as binary data: land use class of interest and all the other ones (McGarigal, 2002). Patch metrics can either quantify the "composition" of the map with the characterisation of patch variety and abundance, or the "spatial configuration" of the patches on the map (McGarigal, 2002). The main types of metrics for categorical maps, as indicated by McGarigal (2002), are highlighted below. The main metrics available for measuring spatial point patterns are then reported. The landscape temporal pattern is then investigated and finally the limitations of the presented metrics are discussed.

3.2.2.1 Landscape spatial pattern metrics for categorical maps

The composition and abundance of landscape features or classes (e.g. land cover) can be described by composition metrics, McGarigal (2002): (i) proportional abundance of each class is a very simple but highly valuable metric; (ii) richness, measures the number of each patch type; (iii) evenness or dominance of each patch type; (iv) diversity metrics measure the richness and evenness. The Shannon's diversity index is widely used (Fu and Chen, 2000), it was developed for information theoretical measures by Shannon and Weavers (1949) and was adapted to landscape ecology by O'Neill (1988).

The next metrics are aiming at describing the spatial configuration of the landscape features (McGarigal, 2002). Some are simply descriptive of the features such as the patch size or shape, while others, such as connectivity examine the spatial relationships between elements on the landscape.

Patch size distribution and density: Most simple measurement of patch compositions.

Patch shape complexity: The most common measurements are the perimeter-to-area ratio and the fractal dimension, some less common indices exist such as patch elongation index (Fu and Chen, 2000). The complexity of the patch shape is often compared to a circle or a square, which are the simplest examples.

Core area: Interior area of a patch, which is not affected by the edges of the patch. The distance of influence of the edge is user defined and depends on patches types. It takes into account the patch size, shape and the distance of an edge.

Isolation/proximity otherwise called gaps/clustering: These metrics measure the distance between patches with similar functions. Due to the imprecision of these type of metrics, a wide range of metrics exists.

Contrast: Contrast-weighted edge density or neighbourhood contrast index, measures the sharpness between one state/or type of patch and another one.

Dispersion: The regular or irregular dispersion of patches through the landscape is measured. Common measurements are based on the nearest neighbour distances, for example their relative variability within a landscape.

Contagion and interspersion: Contagion metrics are based on landscapes in a raster format (grid of regular cells) instead of patches. The cells showing a high spatial contagion form large and aggregated distributions. On the other hand, interspersion is based on patches, and measures the intermixing of patches of different types.

Subdivision: “refers to the degree to which a patch type is broken up (e.g. subdivided) into separate patches (e.g. fragments), not the size, per se, shape, relative location, or spatial arrangement of those patches”, as they are affected by subdivision.

Connectivity: Connectivity metrics measure the degree of connectivity / continuity between patches. These measurements are particularly useful to determine “corridors” for animals and are widely studied (Baudry *et al.*, 2003).

3.2.2.2 Landscape spatial pattern metrics for spatial points

A multitude of statistical methods have been put in place to analyse spatial point patterns (Fortin *et al.*, 2002; Kabos and Csillag, 2002), the main metrics are listed in the Table 3.1.

Table 3.1: Statistical methods to analyse spatial point patterns.

(reproduced after Fortin et al (2002)).

Sampling design	Data types	
	Categorical / qualitative	Numerical / quantitative
Exhaustive census (x- y coordinates)	Nearest neighbours k-Nearest neighbours Ripley's K (uni- and bivariate) Join-count	Aggregation indices (e.g. variance / mean, etc…) *
Regular spacing	Block variance quadrat Spectral analysis Wavelet analysis Fractal dimension	Moran's I (correlation coefficient), Geary's c, Getis (global and local) Semivariance γ SADIE Mantel test (multivariate) Trend surface analysis, kriging, splines
Irregular spacing (1D and 2D)	Fractal dimension	Moran's I (correlation coefficient), Geary's c, Getis (global and local) Semivariance γ SADIE Mantel test (multivariate) Trend surface analysis, kriging, splines, voronoi polygons.

** Aggregation indices could be considered for any of the sampling designs as they do not use spatial information explicitly.*

The main difficulty is determining which metrics are the most relevant for studying a specific landscape pattern. Moreover when studying spatial data such as in a landscape, conventional statistical methods should be considered with care, particularly with regards to statistical independence and distribution of random variables (Cliff and Ord, 1981; Overmars *et al.*, 2003). To overcome the limitation of classical statistics on spatial datasets, spatial statistics such as geostatistics, which are able to take into account spatial autocorrelation, were developed. Spatial autocorrelation refers to the tendency of data to be spatially dependent on neighbouring values.

3.2.2.3 Landscape temporal pattern metrics

The measurement of temporal patterns in the agricultural landscape is almost never studied on its own. Temporal pattern is generally cited and studied with spatial pattern, in spatial and temporal pattern studies. In this case, very often landscape temporal pattern is studied by comparing landscape spatial pattern at different times (e.g. Turner (1990)). Therefore, there are no specifically designed metrics to measure landscape temporal pattern.

Recently crop successions have become more of the focus for studies since it was shown that it is important for biodiversity (Heard *et al.*, 2005). For example, flora on field margins is more strongly influenced by the cumulative effects of crop succession on neighbouring fields than by individual crops (Baudry *et al.*, 2003; Le Coeur *et al.*, 2002). However, no specific metrics are detailed in the literature. Thus for this project, metrics on temporal pattern will have to be designed to meet the needs of the study and characterise crop temporal patterns in agricultural landscapes.

3.2.2.4 General limitations of metrics

The landscape metrics are mainly affected by two factors (McGarigal, 2002): the representation of the landscape (raster or vector format) and the scale (grain -smallest unit- and extent -observed area (Dungan *et al.*, 2002)). Firstly raster format, due to the grid format, may alter greatly some of the metrics values (overestimation of perimeter because of the square cells), such as patch shape complexity, perimeter-to-area ratio or the core area metrics. Moreover the scale / grain of the raster landscape will alter the metrics further (large scale landscape will have larger, less precise square cells). Secondly, the extent of the landscape studied is also important on the measures of spatial pattern. Too small a landscape representation may miss patterns at larger scales and vice versa. Thus the scale of study must be meaningful to the phenomenon under consideration (McGarigal, 2002). Any metrics results are characterised by and dependent on the landscape format and the scale of study (extent, grain).

Furthermore, “most of the metrics are correlated among themselves” (McGarigal, 2002). This correlation is due to the limited number of primary measures of the landscape from which all other metrics are derived. In conclusion, the choice of using particular metrics should be well reviewed. The implications of each metric should be well understood to ensure that only relevant metrics are chosen.

3.2.3 Particularity of the agricultural landscape

Landscape ecology studies mainly focus on natural habitat and on the connectivity between those habitats. In an ecological context, the landscape is very often considered as a continuous environment with no clear boundaries between habitats, e.g. the boundary between a forest and a prairie is very rarely sharp. Therefore landscapes in landscape ecology are very often represented as pixels, forming a regular grid (raster format), as indicated in the previous sections.

However from the farmer's perspective, the agricultural landscape may be considered as a discrete environment (non-continuous), with fields, roads, silos, water bodies, and farm buildings (Baudry *et al.*, 2003), which are clearly spatially defined. Thus to model farmer's decision on crop allocation, the landscape should be considered as discrete, and be represented with a vector model (no pixels). Each feature of the environment has a clear boundary, and areas such as fields are represented as homogeneous polygons. As noted by Flamm and Turner (1994), this format leads to a more adequate and efficient representation of the "complexity of spatial pattern" of agricultural landscapes; this type of representation is commonly used in land use planning (Tulloch *et al.*, 2003).

In several studies, field boundary structure and composition were shown to be dependent on the type of land use of adjacent fields (Barr and Gillespie, 2000; Baudry *et al.*, 2000). Thus, the allocation of crops to fields has to be a dynamic process, with each field or landuse influencing its surroundings.

3.2.3.1 Field unit

The main feature of interest for modelling crop allocation to fields, are fields themselves. However, the definition of a "field unit" is not very clear in the literature (Goense *et al.*, 1996), and it often varies from country to country. In some regions (such as Beauce region in France), fields can be further subdivided for a short period of time (Goense *et al.*, 1996), or aggregated to form a block (islet) of fields (Thenail and Baudry, 2004). For the purpose of this study, a field is defined as an entity with non-changeable boundaries. They are often delimited by hedgerows, stone walls, rivers, roads or other barriers. It is the level at which the farmer will take decisions concerning the crops allocation and management, even if for farmers following precision agriculture the variation of yield within the fields are considered.

3.2.3.2 Field metrics

Field metrics on fields' polygons, are not often described in the literature. However, from farmers' interviews in the Mont-St-Michel Bay area, Thenail and Baudry (2004) determined the following field descriptors: (i) geometry of the field: size, shape, and compactness; (ii) spatial relationship: relative distance to farm, direct access and perimeter with woody hedgerow; (iii) physical environment: slope and hydromorphy; (iv) land tenure. Those fields' descriptors were aggregated at the islets level (groups of adjacent fields) and at the farm level. For the case study, 60% of the land use allocation could be explained by the farm descriptors (Thenail and Baudry, 2004). In the study no indices for spatial and/or temporal patterns were set up at the field or farm level, except for describing the shape of the fields which can be related to the shape of a patch.

To circumvent the issue on polygon metrics, the fields' polygons can be linked to their centroids points, and then spatial point analysis on the centroids can be carried out. However, the metrics described in Table 3.1 are not applicable, because the fields are irregularly spaced and the crops are presented as categorical information, only fractal dimension analysis is adequate. However fractal dimension metrics are not relevant to the spatial and temporal patterns of crops in agricultural landscapes. Thus new types of metrics of landscape spatial and temporal pattern of crops are required, with fields as the unit.

3.2.4 Conclusions on landscape pattern metrics for crop allocation model

The landscape pattern metrics referenced in the literature, are very largely derived from landscape ecology research. Those metrics were designed to measure spatial pattern within a continuous environment represented as a grid (raster format). Thus, they are not directly applicable to determine crops pattern, which needs to be considered at the field unit scale. Landscape temporal pattern metrics are not present in the literature. Usually temporal pattern is studied in relation to spatial pattern, and the spatial configuration of one year is compared with that of another year. In conclusion, for both spatial and temporal pattern of crops in agricultural landscape, new metrics are required to meet the specific needs of this study, which are field based metrics on categorical information (crop types).

The consequences of those conclusions on the design of the LandSFACTS model are detailed in Chapter 4: Methodology for LandSFACTS development, p.35.

3.3 Review of existing models

Models are "a simplified representation" of a complex system (Neelamkavil, 1988; Oxley *et al.*, 2004), and only need to be "good enough to accomplish the goals of the task to which it is applied" (Rykiel, 1996). Research models often aim to enhance the comprehension of the system behaviour, whereas models for policy making are rather designed to help determining the possible effects of changes in policy (Oxley *et al.*, 2004; Winder, 2003). Agricultural systems have important spatial and temporal dimensions (Kropff *et al.*, 2001), and both must be integrated to accurately model agricultural systems. Many studies focus on spatial and temporal allocation of crops; however their modelling approach is highly dependent upon the final aim of the project. In the next sections, the main modelling studies similar to LandSFACTS will be examined from the literature, along with their main modelling components: modelling scales, model variables, farmer decision-process, crop rotations, and the mathematical approaches.

3.3.1 Main modelling approaches to crop allocation

Spatial and temporal arrangements of crops in the landscape are an important parameter for environmental models at the landscape level, from studies of diffuse pollution to climate change. Therefore, the allocation of crops has been the focus of many studies and the three main approaches are (i) mechanistic models integrating farmer's decision making, (ii) statistically coherent models, and (iii) mathematical models.

Farmers' decision-making models study the mechanistic process of crop allocation. The ARABLE model (Rounsevell *et al.*, 2003a; Rounsevell *et al.*, 1998) derived from SFARMOD (Silsoe Farm Model (Audsley *et al.*, 1999)) takes a very comprehensive approach to farmer decisions, integrating driving forces such as machinery, workable hours, husbandry operations (including ploughing and baling), costs, and farmers' attitude to risk. Further studies (Joannon, 2004; Oxley *et al.*, 2002; Oxley *et al.*, 2004) are following the same lead, but with the integration of fewer driving forces. The main drawback of this approach is the quantity of data required to use the model. Moreover the model risks the integration of too many variables, thus over-complicating and over-parameterising the model. This approach is incompatible with the project requirement of developing an empirical and statistical LandSFACTS model.

Statistically coherent models present ways of manipulating governmental agronomic statistics. Some studies (Klöcking *et al.*, 2003; Mignolet *et al.*, 2004) focus on reproducing crop proportions from past agricultural statistics, by randomly allocating the crops over the landscape. Klöcking (2003) integrates (i) expert knowledge for defining crop rotations for 40 years (length of the simulation) and their spatial locations, and (ii) statistical location of each crop over the 40 years. The model coordinates the crop rotations to reach the right yearly crop proportions. Whereas Mignolet et al.'s model (2004), uses Hidden Markov Chains to determine past crop sequences integrating transition rules between crops and the statistical proportion of crops. Those models require extensive agronomic expert knowledge, and consider mainly past datasets. For the LandSFACTS model, considering (i) crop sequences as Markov chains, and (ii) target statistical proportions of crops, would be highly beneficial.

Mathematical models examine technical ways of modelling crop allocation. Detlefesen (2004) reproduces crop rotations by using network and transportation models. Klein Haneveld and Stegeman (2005) use generic multi-year linear programming models integrating the shortest forbidden crop sequences. In mathematical studies of crop allocation, the agronomic reasons and farmers' decision processes are not fully considered, and spatial distribution is overlooked. The idea of integrating unauthorised crop sequences is very valuable, and could be integrated within the LandSFACTS model.

Although existing models do not exactly correspond to the needs of the LandSFACTS model – particularly concerning the explicit integration of spatial and temporal patterns of crops on fields – very useful information may be derived from those studies, such as the use of Markov chains, statistical crop proportions and forbidden crop sequences. Further insights may be derived from their modelling components.

3.3.2 Modelling components

In this section, possible approaches to modelling components crucial for crop allocation modelling are investigated by making references to the models summarised in Appendix A. The examined modelling components are: modelling scale, model variables, farmer decision making process, crop rotations and mathematical approaches.

3.3.2.1 Modelling scales

Two types of scale influence spatial and temporal landscape modelling. At first, the scale at which the processes are modelled (basic unit) and the scale of the whole study, which in fact refers to the extent of the study.

Crop allocation may be modelled at a wide range of spatial basic units from the regional, farm, field (Dogliotti *et al.*, 2003), or to the land islet scale (as defined by the CAP regulation (Thenail and Baudry, 2004)). At smaller modelling scales, higher levels of spatial variability and spatial characteristics may be integrated with the crop allocation. For example, models running at the field scale, allow the integration of the spatial characteristics of each field (stoniness, water supply), with the farm characteristics (for example labour, crop proportion, and machinery), with regional or national characteristics. By working at a higher level, the spatial variation is decreased, for example, the study of Rounsevell (1999) considers the farm as the basic unit, with only the percentage of agricultural land use on each farm. Even less spatial variability is integrated within models working at the regional level such as the study of Mignolet (2004), which considers regionally homogeneous areas as a base unit (mean area: 425 km^2). Their crop allocation results in a relative percentage of each crop for this basic unit (regional area).

Usually for crop allocation models in European landscape, the temporal scale is annual, as only one crop is grown per year. However some studies have a higher temporal resolution, if they also model events with higher temporal variability, such as plant growth in the CropSyst model (Donatelli *et al.*, 1997; Stöckle *et al.*, 2003).

In conclusion, the basic spatial unit of the model defines the scale at which the processes are modelled and which variables are integrated, independently of the extent of the study. By working at the field level (like in the LandSFACTS model), within fields variations are not taken in account; whereas crop pattern between fields will be identifiable. The temporal basic unit for the study is a year.

3.3.2.2 Model variables

Depending on the model, and particularly on the modelling approach (e.g. mechanistic or statistical), the constraints on crops arrangements are taken into account differently. For example Audsley et al (1999) in their mechanistic “Farm scale modelling” within the IMPEL project (Rounsevell, 1999), prioritise the “soil type, climate, scale of operation, and the attitude to risk”. However a multitude of external factors may be included in the model such as (i) environmental factors: weather (rainfall, wind, solar receipts, exceptional events), nutrients inputs/outputs, management of the neighbouring areas; (ii) economic factors: subventions, level of dependence of market price, variability of market price; and (iii) available labour: working hours, number of employee, flexibility of the employee, and overall time management. A statistical model might only consider the crop proportions in the landscape as inputs.

The factors taken into account and the way they interact within the model, define precisely the scope of the model. Each model is therefore unique, and their results are highly dependent on the variables considered and on their processing.

As indicated by Kropff et al, (2001) the more complex approaches are very often the more costly and have a higher time requirement for reaching the results in comparison with simpler approaches. Simpler approaches may offer a lower level of accuracy and reliability (Kropff *et al.*, 2001), but not automatically. However using complex approaches may not always be adequate when modelling complex systems, as the multiplication of input variables and processes increases the range of errors of the outputs. The accuracy of the model outputs is independent of its complexity. The variables to take into account depend on the spatial and temporal scales of the processes modelled (Veldkamp and Lambin, 2001).

In conclusion, models variables must be adapted to the aim of the project and the modelling approach selected. Modelling crop allocation might require the integration of biophysical, land use, and socio-economic factors, which are linked to different spatial and temporal scales (Thenail and Baudry, 2004). However the accuracy or usefulness of a model is independent from its level of complexity. A very important point is that the omissions and assumptions of the model should always be clearly identified and explained (Oxley *et al.*, 2002).

3.3.2.3 Farmer decision-making process

Technical advice provided by third party agronomic experts, is very often poorly followed by farmers (Aubry *et al.*, 1998). This is not simply due to technical failing of farmers, but is mainly due to specific aims and constraints of individual farmers, such as economic and environmental constraints (Aubry *et al.*, 1998). Furthermore, a determinant factor is “risk aversion” (Audsley *et al.*, 1999), as two farmers in exactly the same conditions would manage their farm differently, this is mainly due to different approaches and attitudes to risk management and on the farmers own perception of the variability of crop yields and prices.

Often farmers must decide, organise, and execute the farm workload, consequently the decision-making process is very often implicit and internal (Wünsch, 2004). Farming systems research aims specifically at identifying and understanding the reasons of the farmers decision (Aubry *et al.*, 1998; Spedding, 1975). Integrating human drivers in models can be particularly complex (Thenail and Baudry, 2004). Many studies are still ongoing on this subject, and many models are being developed (Aubry *et al.*, 1998; Audsley *et al.*, 1999; Joannon, 2004; Oxley *et al.*, 2002).

Farmers may be modelled as “profit maximisers” (Audsley *et al.*, 1999), they also try to minimise the variability of the farm income between years. Oxley et al., (2002) implemented stochasticity into the farmer behaviour in order to model the variations between farmers due to social and cultural preferences, sensitivity to environmental conservation and their attitude towards change. Moreover they noted that a “hierarchy of nested spatial and temporal scales” affects the decision making of farmers.

Another factor to consider is that a farmer’s role is shifting from food producer to manager and safe keeper of the natural environment. This is particularly evident in the new European system of subsidies with subventions such as the “single farm payment”, which encourages farmers to enhance natural habitat around field, with field margins, hedgerows (DEFRA, 2005b).

Farmer decision-making processes are thus complex, and are integrating a wide range of constraints such as agronomic, economic, and environmental constraints, while managing risks and profits. The farmer decision-making process may be modelled as a mechanistic process or may be integrated as a stochastic variable. After Chapter 2, the modelling approach of LandSFACTS should not be mechanistic but empirical and statistical, therefore the decision-making process of farmers will be stochastically implemented.

3.3.2.4 Crop rotations

As presented in section 3.1.2: Agronomic constraints (p.15), crop rotation has a major role in the crop allocation to fields. Efficient ways of setting up new crop rotations is a constant subject of studies. Models, such as ROTOR (Bachinger and Zander, 2006), or ROTAT (Dogliotti *et al.*, 2003) are tools designed to optimise crop rotations for yield benefits. Other models, like SFARMOD (Rounsevell *et al.*, 2003b), CropSyst (Donatelli *et al.*, 1997; Stöckle *et al.*, 2003), or the study of Oxley et al (2004), aim at simulating the farmer’s decision-making process of crop allocation by integrating agronomic, environmental and farm management objectives.

Only a few studies concentrate on providing empirical or statistical tools to model crop rotations in a mathematical manner, i.e. as required for the LandSFACTS model. Klein Haneveld and Stegeman (2005) referred explicitly to some of the agronomic rules discussed above, while using a mathematical optimisation technique know as linear programming to derive crop rotations. Detlefsen (2004) presented a network model used to represent crop rotation. Their approach provides an example on how to integrate crop rotations within the LandSFACTS model.

3.3.2.5 Mathematical approaches on processes

Multiple mathematical approaches to crop allocation have been reported in the literature. The most popular and predominant approach is linear programming, which enables the optimisation of agricultural and economical parameters (Dogliotti *et al.*, 2003; Klein Haneveld and Stegeman, 2005; Rounsevell *et al.*, 2003a). As the farmer is considered as a profit optimiser, the optimum economical solution is identified, while considering the agronomic or managements constraints. Multi-agents models are also used relatively often (Le Ber *et al.*, 1998; Matthews, 2006; Parker and Meretsky, 2004). Each agent tries to meet its objectives, while respecting their constraints. The agents compete against each other. For example in the study of Le Ber et al (1998), each agent is a land cover type, and the agents compete for field allocation. Each agent aims at its target total area, while respecting its spatial constraints (e.g. soil type, slope), and respecting global constraints (e.g. percentage of each land cover type). Another common mathematical modelling technique is simulated annealing (Le Ber *et al.*, 1998). For this technique, an optimisation function is used to determine the best configuration (closest to the desired one), however, in order to avoid being blocked within local optima, sub-optimal configurations are accepted from time to time. Sub-optimal configurations, may violate local constraints (for example soil type, or slope percentage), thus the results obtained by this method have to be checked to ensure they respect important local constraints. Rules-based processes, otherwise called decision trees, may also be used (Baudry *et al.*, 2003; Oxley *et al.*, 2004). Simple rules are being followed. For example, Oxley et al (2004) represent the crop choice as a function of socio-economics, physical properties and institutional conditions. They indicated that this kind of modelling framework is better adapted to explore the possible outcomes than to predict the future. Detlefesen (2004) investigated the possibility of using a network and transportation algorithm to model crop rotations. The arcs of the network represent the decision variables and at each node the supply and demand must be satisfied. The problem is then rewritten under the form of a matrix, which then can be solved using linear programming techniques.

Le Ber (1998) compared the results from the simulation of the spatial organisation on a milk production farm, obtained by using three different modelling approaches: expert knowledge, multi-agents systems and simulated annealing. The results showed that the expert knowledge approach offered less optimisation of the results, moreover this technique was less respectful of the agronomic and farmer constraints. The results from the multi-agent systems were of higher minimal quality than the simulated annealing model, the local constraints were well respected and the results were obtained very quickly. On the other hand, simulated annealing model, was able to generate more optimal solutions but at greater expense in computer time. The results showed as well a lower variability than those generated from the multi-agent model. The simulated annealing process is conceptually much closer to the farmer thinking, moreover the connection with economical farm models would be very easily integrated.

Four main mathematical approaches were used in models close to LandSFACTS. Any of the approaches could be used for the model developed in this thesis. However, each of them would be more adapted to model some specific processes, therefore the different approaches could be concurrently used. For example decision-trees could be used to avoid some specific spatial configuration (e.g. GM oilseed rape next to conventional), and simulated annealing to optimise the spatial and temporal pattern of crops.

3.3.3 Conclusions on modelling crops allocations

The LandSFACTS model, to be designed for this thesis, must be a research tool allocating crops to fields, at the field scale, over several years. Several studies from the literature consider similar models. However each of them has a different focus, either on mechanistic approach for complex model on farmer decision-making, or a more statistical approach to produce crop allocation coherent with agronomic statistics, or they model at a different scale (e.g. farm scale, regional scale). None of them integrates the spatial and temporal patterns of crops as such.

The variables used for those models, e.g. biophysical, land use and socio-economic factors, are adapted to their aims, their modelling processes, and the spatial scale at which the processes are modelled. Their aim and modelling approach are dissimilar to the needs of the LandSFACTS model, which aims at modelling empirically and statistically crop allocations to fields. An important component of crop allocations is the decision-making done by the farmer. Some models integrate fully the decision-making, while others retain some rules and introduce some stochasticity to reproduce the individuality of farmers' behaviour, which are particularly due to diverse risk managements. Many models aim at providing support for building up crop rotations, however very few models use crop rotations in a mathematical and empirical manner. Furthermore the modelling itself may be built up around four mathematical processes as indicated by the models from the literature: linear programming, multi-agents models, simulated annealing, rules-based processes. Each of those is adapted to model specific situations, and could be used to model different parts of the crop allocation to fields model.

In summary, no existing model fulfils the LandSFACTS specifications of allocating crops to fields by directly modelling the crops spatio-temporal patterns, using an empirical and statistical approach. However several approach or tools reported in the literature are relevant to the LandSFACTS model, those are: the stochastic integration of farmers decision-making, the use of rule based constraints on crop allocation, such as forbidden crop sequences, simulated annealing techniques to increase the model efficiency. The detailed consequences of those conclusions on the design of the LandSFACTS model, are detailed in Chapter 4: Methodology for LandSFACTS development, p.35.

3.4 Conclusion

To best design the LandSFACTS model, in this chapter the literature was investigated for the origin of the crop arrangements, available statistical tests on crop arrangements, and existing models on crop allocation. The crop arrangement in agricultural landscapes results from a complex and not completely understood decision-making process of individual farmers, which integrates agronomic, environmental, economic and policy constraints. Even if mechanistic processes must not be incorporated within the model, conclusions relevant to the design of the model were identified: (i) crop rotations structure crop successions on fields; (ii) market prices of crops influences the crop choice; (iii) the spatial extent of crops can be limited by environmental conditions; (iv) spatially close fields tend to have similar ranges of crops; (v) separation distances between crops are enforced for seed production. The LandSFACTS model has to control the crop arrangements by using statistical analyses. However as none in the literature met the requirements set out for LandSFACTS (categorical information –crop types– with discrete spatial units –polygons), new statistical analyses had to be developed. The review on existing models of crop allocations confirmed that no currently available model met the LandSFACTS specifications, however the review allowed the identification of useful techniques for the LandSFACTS model, such as the use of (i) stochasticity to simulate farmer decision-making; (ii) linear programming with simulated annealing process to optimise crop allocation; (iii) rule based constraints to forbid specific configuration of crop allocation, e.g. forbidden crop sequences.

Based upon the conclusions drawn from the literature review in the current chapter, Chapter 4 lay out the methodology to develop the LandSFACTS model. New statistical measurements and mathematical representations of crop rotations are detailed in Chapter 5 and 6 respectively, before to be incorporated within the LandSFACTS model in Chapter 7.

4 Methodology for LandSFACTS development

After identifying the specifications for the LandSFACTS model (Chapter 2), the literature was reviewed (Chapter 3) to provide support for the development of the LandSFACTS model. Those reviews presented the origin, characteristics and statistical measures of the crop arrangement, and existing models close to LandSFACTS aims. From the conclusions of the reviews, this chapter presents the approach chosen for the development of the LandSFACTS model. More particularly the modelling approach and the control of spatial and temporal arrangements of crops are presented. In this chapter, datasets on crop arrangement at the landscape scale are also presented; they were used for the development of the LandSFACTS model (Chapter 5), and for its testing (Chapter 8) and later its dissemination (Appendix B).

4.1 LandSFACTS model approach: conclusions from review

As stated in Chapter 2 (Specifications of LandSFACTS model), the LandSFACTS model aims at simulating crop allocation to fields by directly modelling user-defined crop arrangements, and not the decision making process leading to it. After the literature review carried out in Chapter 3 (Review of the origins, metrics, and models of crop arrangements), no existing model meets those requirements. Therefore, the LandSFACTS model requires a new approach and structure to fulfil its objectives, which can be inspired by the conclusions from the review in Chapter 3.

4.1.1 Combining statistical and real variables

Modelling mechanistically the decision process leading to crop allocation, would result in a highly complex model and would require a huge quantity of data inputs; this approach is outside of the LandSFACTS specifications, as stated in Chapter 2. The LandSFACTS model aims at directly modelling the crop allocation, by using an empirical (based on observations from real landscapes) and statistical (quantifiable and reproducible) approach.

At first, for the statistical part, the LandSFACTS model must directly simulate the crop arrangement, using statistical tests to control the crop patterns. As no statistical analyses on crop patterns exist in the literature for categorical data (land uses) on discrete spatial units (fields as polygons), new statistical tests adapted to the data characteristics must be designed. This is investigated in Chapter 5: Measuring the spatio-temporal patterns.

However, the statistical control on crop arrangement will only provide a “loose control”, where specific rules cannot be controlled, such as on which fields crops can be grown or specific separation distances between crops. To complement the statistical measures, specific tools are needed; they are inspired by the constraints on farming systems listed in Chapter 3.1: Origin of crop arrangements and from variables of existing models listed in Chapter 3.3.2: Modelling components. Despite the non-mechanistic approach of LandSFACTS, those constraints provide rules that the model should respect. Such rules are: (i) the control of the geographical extent of the crops to reflect the conclusion drawn from the environmental constraints; (ii) the integration of crop rotations in a way that permits the consideration of fixed and flexible crop rotations, as indicated by the agronomic constraints; (iii) the addition of constraints on crop successions to complement the crop rotations, e.g. return period of crops and forbidden crop sequences; (iv) the possibility of separation distances between any specified crops (useful for GM coexistence scenarios but also for seed production for example).

This modelling approach of separating into (i) constraints and (ii) patterns the spatio-temporal crop arrangements provides a high degree of flexibility to the user to obtain desired scenarios of crops arrangements. The terminology of constraints and patterns of crops are exemplified in the next section. Statistical ways of quantifying crop patterns will be investigated in Chapter 5 (Measuring the spatio-temporal patterns). The mathematical integration of crop rotations is investigated in Chapter 6: Mathematical representation of crop rotations.

4.1.2 Crop constraints and patterns terminology

As indicated above the spatial and temporal crop arrangements are both divided into two components: (i) the constraints representing fixed rules imposed on the landscape and (ii) the patterns implementing a general trend using statistical analyses.
The exact meanings are detailed below.

spatial pattern: defined by use of statistical tests to measure the crop spatial aggregation or homogeneity (regular spatial pattern), cf. Figure 4.1.

spatial constraints: features the separation distances between crops (e.g. for seed production or possibly for coexistence system with GM crops).

temporal pattern: use of statistical tests to measure the crop temporal aggregation and homogeneity (regular temporal patterns), equivalent to the dispersion of crops through time, cf. Figure 4.2.

temporal constraints: features rules on crop successions, such as return periods and forbidden crop sequences.

A pattern is labelled random if no spatial patterns (i.e. no statistically significant aggregation or homogeneity) can be detected. The terms regular and homogeneous patterns are used interchangeably within the context of this thesis.

The spatial pattern and constraints influence the crop spatial arrangement, whereas the temporal pattern and constraints influence the crop temporal arrangement.

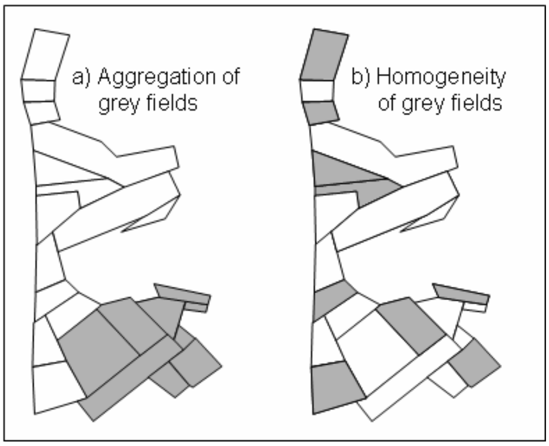


Figure 4.1: Examples of aggregated and homogenous spatial patterns of crops.

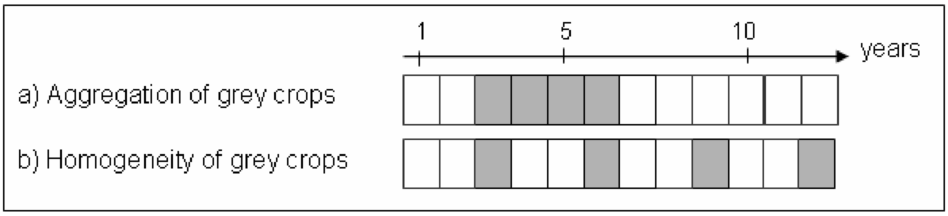


Figure 4.2: Examples of aggregated and homogeneous temporal patterns of crops.

4.2 Datasets: from analysis to validation and examples

Datasets of agricultural landscapes are required for (i) investigating existing crop arrangements (Chapter 5), (ii) devising and testing statistical analyses of crop patterns (Chapter 5), (iii) assessing the LandSFACTS software (Chapter 8), and (iv) disseminating example datasets with the LandSFACTS model (Appendix B). The dataset must be composed of a shapefile (GIS format) with the fields represented as polygons. Information on the cropping systems is required, such as crops, crop rotation, rules on crop successions, and yearly crop proportions; cropping information linked to individual fields is an advantage. Ideally the dataset should be representative of European agricultural landscapes.

Three datasets were easily available for the development of LandSFACTS: the Fife dataset (Scotland), Beauce dataset (France), and Burgundy (France), cf. Figure 4.3. Due to time constraints, those datasets were not assessed for how comprehensively they were representative of European agricultural landscapes, particularly regarding to field size, field shapes, and cropping systems. However they are sufficiently diverse to be adequate for developing the LandSFACTS model. The analyses of the spatial and temporal pattern of crops were set up and tested by using the Burgundy dataset, due to its immediate availability and its completeness of information on spatial and temporal allocation of crops (Chapter 5: Measuring the spatio-temporal patterns). The Fife and Beauce datasets were used to verify and validate the LandSFACTS model (Chapter 8: Model assessment).

Figure 4.3: Location of the study sites through Europe.

Those datasets are not in the public domain, and they are subject to confidentiality clauses. The dissemination of farmer specific information and digital data are restricted. Therefore to provide example datasets to potential LandSFACTS model users, two fictitious landscapes were created in a shapefile format: SmallLandSCAPE and BigLandSCAPE. Resemblance to any real landscape is unintentional. Both datasets are provided in a digital format in Appendix B and they were used to verify the model.
The datasets are presented below.

4.2.1 Burgundy dataset

The Burgundy study site was originally collected and used for studying oilseed rape pollen dispersal through the landscape (Colbach *et al.*, 2005), by INRA-Dijon (France) with the collaboration of Dijon-Céréales Cooperative and their member farms. This study site encompasses 72 fields surrounded by forest. The forest provides a physical barrier against pollen dispersal around the study site, thus the site can be considered as an independent unit for the evaluation of risks due to gene flow. The location of each field and their ownership is known (10 different farmers), along with the crops grown on each of them from 1994 to 1997, and the rotation that was followed in each field. The dataset is complete without any missing data, apart from the number and size of fields managed by the study site farmers outside the study area.

The most widely represented crops were oilseed rape, winter wheat, and winter barley, with 28.2, 26.6, and 23.2% respectively, of the mean crop area. The three remaining crops (spring barley, rye, and set-aside) were less-well represented on the study area comprising less than 24% together. Ten individual farmers cultivated the 72 fields comprising the study area. The number of fields per farmer ranged from one to thirteen and their total area ranges from 12.5 to 368.7 ha (Figure 4.4).

Farmer	Number of fields	Total area (ha)
0	1	12.5
1	3	59.2
2	2	18.6
3	11	157.7
4	12	208.1
5	9	210.7
6	9	210.1
7	7	181.8
8	5	49.8
9	13	368.7

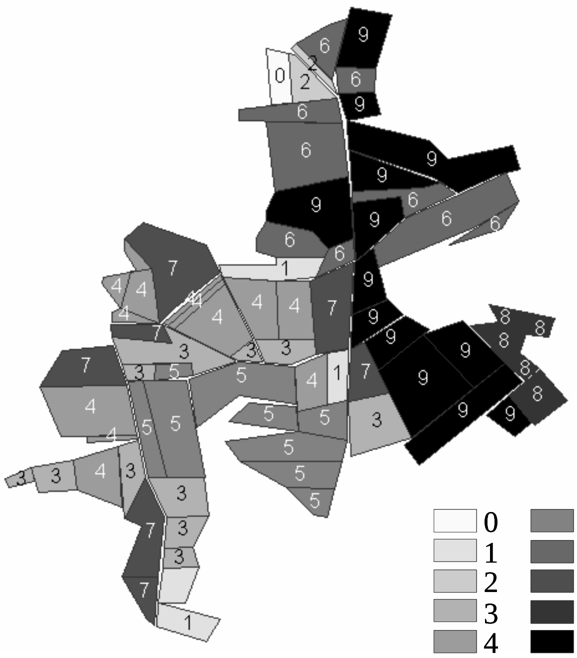


Figure 4.4: Number, area and location of farmers' fields in Burgundy study site.

The colouring and the numbering on the map of the study area correspond to each individual farmer listed in the table on the left side.

The main drawback of the dataset is its limited number of fields. A larger dataset with more than 1,000 fields would have been more statistically interesting. However, no datasets of this size with complete crop rotation information for each field was available at the time. Thus, this study site was selected to set up and test the methodology for analysing spatial and temporal patterns of crops.

4.2.2 Fife dataset

The Fife dataset was provided by the SCRI, after a survey carried out in 2004 (Young *et al.*, 2006) and from agricultural census (National Statistics, 2005). The CETIOM (Centre Technique Interprofessionel des Oleagineux Metropolitains) provided further information on the study area (Sausse, 2005). The dataset is composed of the shapefile of the fields (Figure 4.5), farmers and land-uses.

Figure 4.5: Landscape of the Fife dataset.

The Fife dataset has 388 fields, corresponding to an area of 24.92 km^2, managed by 5 farmers. The cropping systems are based around four constraints: temporal crop cycles, climatic conditions, and agronomic rules on crop successions, and current profitability of crops. After the survey, 114 fields out of 388 are permanent grassland. For the purpose of this project, only 10 crops are considered: wheat, winter barley, spring barley, winter oats, spring oats, winter oilseed rape, spring oilseed rape, winter GM oilseed rape, potatoes, set-aside, and other miscellaneous crops. Only a simplification of their complex cropping systems is being used in this thesis.

4.2.3 Beauce dataset

The Beauce dataset was set up for the investigation of gene flow dispersal for oilseed rape crops (Lavigne *et al.*, 2002-2006), the dataset was provided by l'Institut National de la Recherche Agronomique (INRA), l'Université Paris-Sud 11 (UPS) and Centre Technique Interprofessionnel des Oléagineux Métropolitains (CETIOM). Further information on the agronomic systems and crop rotations were provided by the CETIOM (Sausse, 2005). The shapefile with the arable fields is presented in Figure 4.6.

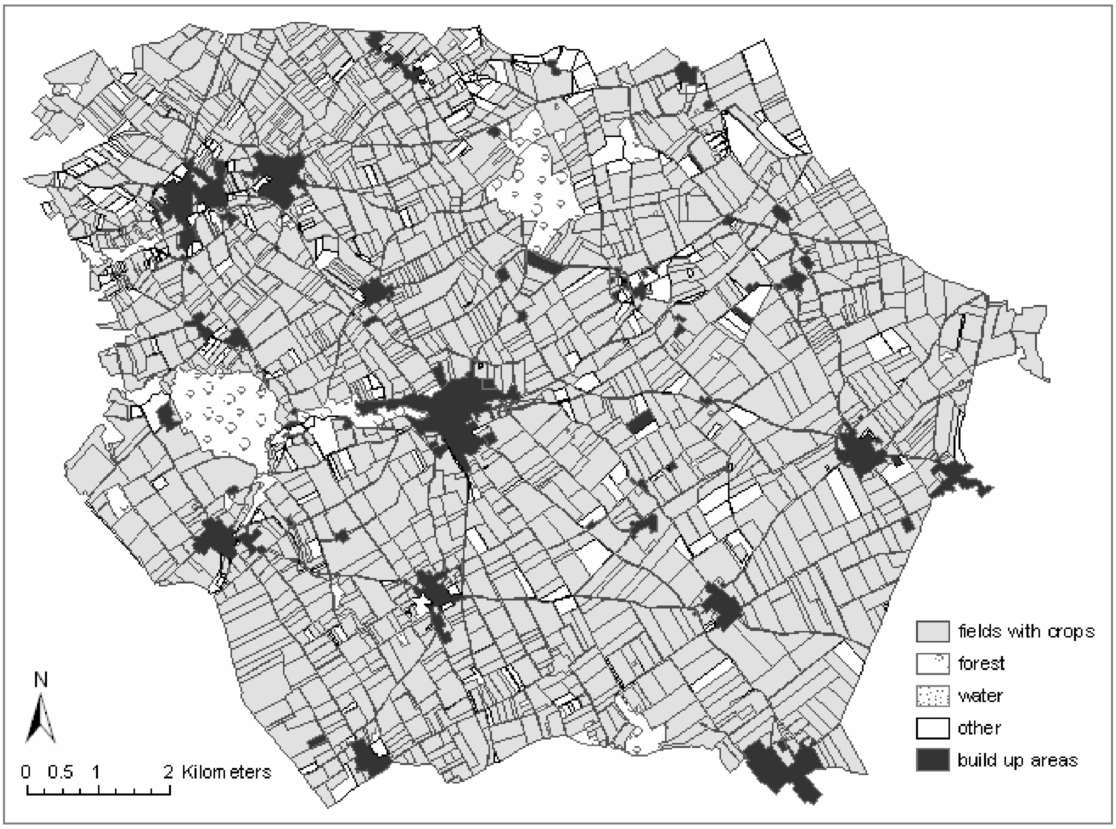


Figure 4.6: Landscape of the Beauce dataset.

The dataset is composed of 1,993 fields, over 92.23 km^2 of arable land, managed by 21 farmers. The main cultures are oilseed rape, maize, wheat, spring and winter barley, sunflower, peas and fodder. The cropping systems are varied from fixed rotations for fields with high environmental constraints (shallow soil without irrigation), to highly flexible rotations for irrigated fields or fields with seed production contracts. In order to facilitate the interpretation of the simulation results, only a simplification of their complex cropping systems is being used in this thesis.

4.2.4 SmallLandSCAPE dataset

SmallLandSCAPE is a created landscape to meet the purpose of dissemination of the LandSFACTS model. The dataset is comprised of 10 fields, Figure 4.7. Resemblance to any existing landscape is unintentional. The limited number of fields provides new LandSFACTS users with a comprehensible landscape, on which to investigate the model scope, input parameters and processes.

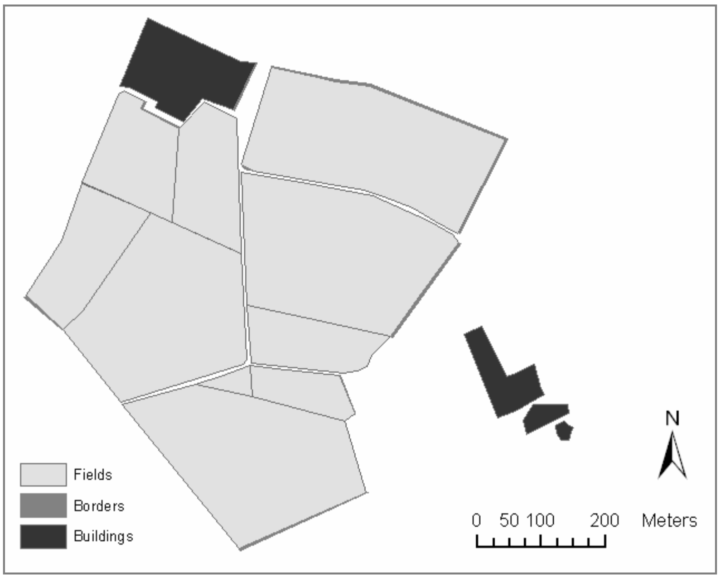


Figure 4.7: Landscape of the SmallLandSCAPE dataset.

4.2.5 BigLandSCAPE dataset

BigLandSCAPE is a created landscape to meet the purpose of dissemination of the LandSFACTS model. The dataset is comprised of 200 fields, 3 built-up area and 5 forests, Figure 4.8. Resemblance to any existing landscape is unintentional. This larger landscape provides new LandSFACTS users with the possibility to investigate further the model behaviour and usefulness, particularly regarding interactions between the fields (e.g. spatial patterns of crops).

Figure 4.8: Landscape of the BigLandSCAPE dataset.

4.3 Steps for LandSFACTS development

This chapter defines from the reviews in Chapter 3, the methodology chosen to develop the LandSFACTS model. The model should simulate stochastically the farmer decision on crop allocation (based on crop rotations) while respecting (i) spatial and temporal patterns of crops controlled by specifically designed statistical tools, and (ii) spatial and temporal constraints of the crops controlled by rules such as separation distances between crops, forbidden crop sequences, crop proportions. To create the LandSFACTS model with the structure detailed above, two sets of tools must be specifically designed: (i) new statistical analyses of crop patterns, and (ii) new mathematical representation of crop rotations.

The setting up and testing of the new statistical analyses of crop patterns required a study landscape with readily available complete data on the fields shape, owners, crops in fields over time. The Burgundy dataset met those requirements and was used for the above purpose. Two further datasets, Fife and Beauce datasets with fields shape, crops and main cropping systems were used to assess the LandSFACTS model. Then in order to freely disseminate example datasets with the model, two datasets, SmallLandSCAPE and BigLandSCAPE were created independently from any real landscape.

The next step of the project is to create new statistical measurements of crop patterns specifically designed for their integration within the LandSFACTS model (Chapter 5: Measuring the spatio-temporal patterns). Then a mathematical integration of crop rotations in a flexible and versatile format is presented (Chapter 6: Mathematical representation of crop rotations). When those two tools are set up, the LandSFACTS model can be detailed (Chapter 7: Description of the LandSFACTS model) and then assessed (Chapter 8: Model assessment).

5 Measuring the spatio-temporal patterns of crops

Statistical methods to measure spatial and temporal patterns of crops in an agricultural landscape are required in order to integrate the crop patterns within the LandSFACTS model of allocation of crops to fields. The spatial unit of the crop allocation is the field, and the crop definition is categorical (crop types). The particularities of those data types were detailed in Section 3.2.3 (Particularity of the agricultural landscape, p.23). In this chapter, statistical analyses developed for the LandSFACTS model needs are detailed, along with the conclusions drawn from their use on the Burgundy study site. The statistical analyses presented in this chapter are the subject of a published article (Castellazzi *et al.*, 2007b). Finally, after critical analysis, the most adequate statistical tests are selected for integration within the LandSFACTS model.

5.1 New statistical analyses on crops' spatial and temporal patterns

In this chapter, new methodologies to analyse the spatial and temporal pattern of crops are detailed and the Burgundy data are used as examples. The methods are detailed in literal and mathematical format; their degree of accuracy and / or precision is also reported. At first some important definitions regarding crop rotations are presented.

5.1.1 Definitions

Crop rotation: definite cyclical sequence of crops grown on a field, only one crop per year.

Starting crop: indicates the crop by which the crop rotation is starting for a specific field and a specific starting year. For each crop rotation, consecutive letters are given for each crop in the sequence (Table 5.1). A field labelled as "rotation Z, with starting crop B", indicates that the field follows rotation "Z" and the crop "B" is grown in the first year (Table 5.2).

Phasing of rotations: fields following the same rotation, having the same starting crop; through the years, the crops grown on the fields are thus temporally in phase (aggregated).

Table 5.1: Example of crops sequence in a rotation.

Crop sequences	A	B	C
Rotation Z	Wheat	Winter barley	Oilseed rape

Table 5.2: Example of “starting crop” numbering for a crop rotation.

Starting crop A	Starting crop B	Starting crop C	1994	1995	1996	1997	1998	1999	2000
X			A	B	C	A	B	C	A
			W	wB	OSR	W	wB	OSR	W
	X		B	C	A	B	C	A	B
			wB	OSR	W	wB	OSR	W	wB
		X	C	A	B	C	A	B	C
			OSR	W	wB	OSR	W	wB	OSR

W: wheat; wB: winter barley; OSR: oilseed rape

5.1.2 Temporal pattern of crops

Specific statistical analyses were set up to investigate the temporal pattern of the crops and in particular the temporal phasing of the crop rotations on different fields. The two first analyses were based on the chi-square test and the third one introduced a randomisation test.

5.1.2.1 Crop rotation phasing

The temporal phasing of crop sequences is determined by two elements: (i) the rotation and (ii) the starting crop. The crop rotation sets up the cyclical sequence of crops, whereas the starting crop defines the temporal phasing of the crop sequences. Where there is an identical rotation and an identical starting crop on several fields then there will be a high temporal phasing of the crops in the fields. In Table 5.3, the examples of a farmer labelled 3 and a farmer labelled 4 are presented. Rotation 1 of farmer 3, is followed in two of his fields, and both have the same starting crop (A); the crop rotations are thus in phase temporally. In the second example, two fields of farmer 4 are following rotation 7, with different starting crops (A and F) and therefore the crop rotations are not temporally aggregated.

Table 5.3: Crop rotation phasing: example data (observed state for chi-square test).

Farmer	Rotation	Crop sequences						Rotation length	Number of fields	Number of fields per starting crop					
		A	B	C	D	E	F			A	B	C	D	E	F
3	1	W	OSR	-	-	-	-	2	2	2					
4	7	W	W	OSR	W	wB	OSR	6	2	1					1

Burgundy study area

W: wheat; wB: winter barley; OSR: oilseed rape

To determine the crop rotation phasing, a **chi-square analysis per farmer per rotation** was performed. The analysis considered, for each individual farmer and crop rotation, the number of fields starting with the same crop (identical starting crop). The observed state, obtained from the survey (cf. Table 5.3), was tested against the expected state, which is an even distribution of the number of fields over the possible starting crops (e.g. two fields following the same six year rotation have, for each starting crop, an expected value of [2 / 6] = 0.33; cf. Table 5.4, b). The mathematical definition of the test is reported in Figure 5.1. The two examples presented in Table 5.3, are tested for the temporal pattern of the crop rotation (Table 5.4), and in these examples both have no statistically significant temporal aggregation or homogeneity at a 95% confidence interval.

The chi-square is calculated as follow:

le = length of rotation

m = number of fields (same farmer and same rotation)

i = starting crops (for a 6 years in rotation, i = A, B, C, D, E, F.)

Oi = observed value for starting crop i; number of fields starting with the same rotation year.

Ei = expected value for starting crop i; Ei = m / l

χ^2 = chi-square value

p = probability (determined using a chi-square table)

$$\chi^2 = \sum_{i=1}^{l} \frac{(Oi - Ei)^2}{Ei}$$; the probability p is then determined using a chi-square table with the chi-square value and the degrees of freedom (le – 1)

Hypothesis and probabilities:

H_0: Observed and Expected values are not significantly different

H_1: Observed and Expected values are significantly different, indicating temporal aggregation.

if $p > 0.95$ → H_0 is true (not significantly different, indicating temporal homogeneity)

if $p < 0.05$ → H_0 is false and H_1 is true (significantly different, temporal aggregation)

Figure 5.1: Crop rotation phasing: definition of chi-square test per farmer and rotation.

Table 5.4: Crop rotation phasing: two examples for the chi-square test

a) Farmer 3 – rotation 1 (Burgundy study site)

	Starting crop						Fields	Number	Degree of
	A	B	C	D	E	F	number	of years	freedom
Observed state	2.00	0.00	-	-	-	-	2.00	2	1
Expected state	1.00	1.00					Chi-square		Probability
$(Oi-Ei)^2$ / Ei	1.00	1.00					2.00		0.1573

b) Farmer 4 – rotation 7 (Burgundy study site)

	Starting crop						Fields	Number	Degree of
	A	B	C	D	E	F	number	of years	freedom
Observed state	1.00	0.00	0.00	0.00	0.00	1.00	2.00	6	5
Expected state	0.33	0.33	0.33	0.33	0.33	0.33	Chi-square		Probability
$(Oi-Ei)^2$ / Ei	1.33	0.33	0.33	0.33	0.33	1.33	4.00		0.5494

This analysis is useful to determine the temporal synchronicity between fields following the same crop rotation. However, the test requires that specific crop rotations are used in several fields. Due to the limited number of fields considered at a time (same farmer and same rotation), this analysis presents two main disadvantages: (i) low degrees of freedom (i.e. difficulty to obtain statistically significant results such as in the case of rotation 1 of farmer 3), and (ii) an incomplete study of crop temporal pattern because the synchronisation between different crop rotations is not taken into account and any rotation represented by only one field is overlooked. The low degrees of freedom may be mitigated by aggregating all the chi-square values and the degrees of freedom in the study area (as obtained above), to determine an overall probability of the existence of crop rotation phasing. To further identify temporal pattern of crops, further analysis are required.

5.1.2.2 Crop phasing

The study of the phasing of crops, regardless of their crop rotations, has the definite advantage of considering the crops of every field. To analyse the crop phasing a "**chi-square analysis per farmer and per crop**" was used. The analysis considered the proportion of fields of a farmer growing a specific crop per year (one value per farmer, per crop and per year), independent of the crop rotations. The time period used had to be as long as, or longer than, all the rotation lengths and was a multiple of all of them. The observed values, obtained from the survey, were tested against the expected values, which were an even distribution of the crops over the years. The mathematical definition of the test and example tables, are reported in Figure 5.2.

The chi-square ($\chi 2$) per farmer and per crop is calculated as follow:

le = length of rotation
fi_T = number of fields of a farmer
y = year of simulation considered; y = 1, 2, … 12.

• fi = number of fields
• c = crop considered (wheat…)
• fa = farmer considered; fa = 1, 2, … 10.

i = starting crop (for a 6 years in rotation, i = A, B, C, D, E, F).

Oc,fa,y = observed value for crop c, farmer fa and year y; proportion of a farmer' fields having a specific crop c, at a year y. Oc,fa,y = Σ [fi (fa, c, y)] (cf. Table a)

a) Observed values for farmer 1 and crop OSR (Burgundy study area)

		Simulation years (y)												%CV
fi	Farmer (fa)	1	2	3	4	5	6	7	8	9	10	11	12	
1	1	0	1	0	0	1	0	0	1	0	0	1	0	
1	1	0	1	0	0	1	0	0	1	0	0	1	0	
1	1	1	0	0	1	0	0	1	0	0	1	0	0	
fi_T =3	1	1	2	0	1	2	0	1	2	0	1	2	0	85.28

Bold: individual Oc,fa,y = proportion of fields of an farmer with crop c and at year y.
%CV: Percentage of the coefficient of variation, between years (1 value per farmer, per crop).

Ei = expected value for crop c, farmer fa; = (number of crop c within the rotation) / (rotation length) * (number of fields of farmer fa following this rotation), cf. Table b. This expected value is the same for every year due evenness of distribution through the year.

b) Expected values for farmer 1 for any one year (Burgundy study area)

Farmer (fa)	Rotation-Id	Crop sequence			Rotation length (le)	fi	Number of OSR in rotation (Nc)	A: (Nc / le) * fi	Expected value: Σ (A) for each fa
		A	B	C					
1	3	W	Wb	OSR	3	1	1	0.33	1
1	4	W	Sb	OSR	3	2	1	0.67	

p = probability (determined using a chi-square table)

$$\chi 2 = \sum_{i=1}^{l} \frac{(Oi - Ei)^2}{Ei}$$

; the probability p is then determined using a chi-square table with the chi-square value and the degree of freedom (le – 1)

Hypothesis and probabilities:

H_0: Observed and Expected values are not significantly different
H_1: Observed and Expected values are significantly different, indicating temporal aggregation.
if $p > 0.95$ → H_0 is true (not significantly different)
if $p < 0.05$ → H_0 is false and H_1 is true (significantly different, temporal aggregation)

c) Chi-square results with observed and expected values of farmer 1, OSR.

	Years												Chi-square	Probability (p)
	1	2	3	4	5	6	7	8	9	10	11	12		
Observed	1	2	0	1	2	0	1	2	0	1	2	0		
Expected	1	1	1	1	1	1	1	1	1	1	1	1	8.00	0.7133

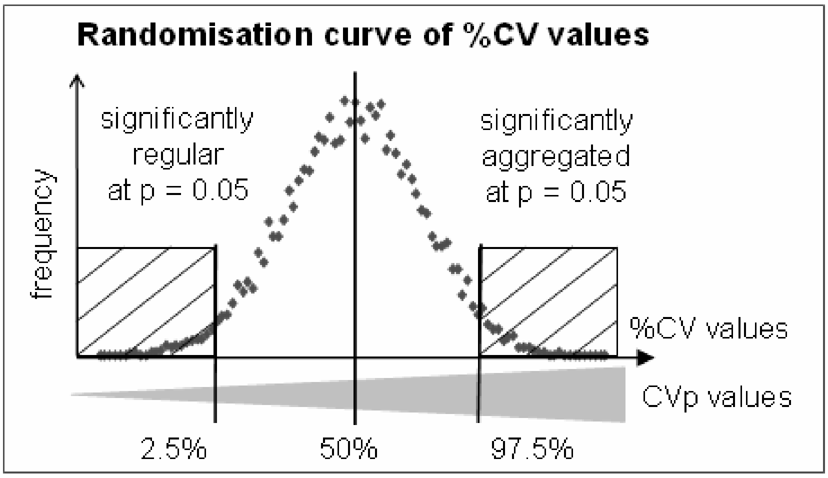
Figure 5.2: Crop phasing: definition of chi-square test per farmer and crop.

This analysis in Figure 5.2 has the main advantage of considering all crops, regardless of the crop rotations, by using a simple statistic test. However the expected values of the test, which consider an even distribution of crops through years, do not take in account the sequential structure of crop rotations, and more particularly the minimum return period of crops, e.g. sugar beet may be grown only 1 year in 3. In this analysis, the sequences of the crops are assumed to be flexible. The sequential structure of crop rotations, in the observed values, artificially increases the temporal homogeneity of crops. Hence significant homogeneity results should be considered with care and the source of the homogeneity should be investigated. However this analysis is useful for evaluating an overall temporal pattern, when the origin of the pattern does not need to be known (structures of the rotations used, choices of rotations, or farmer choice of starting crops). The next analysis integrates the structure of the crop rotation into the test, to circumvent the issue outlined above.

5.1.2.3 Crop temporal variability compared to random simulations

This analysis considers the temporal pattern of each crop of each farmer, by investigating its variability. Moreover, the sequential structure of the crop rotations is integrated into the “expected state” in order to take into account the homogeneity intrinsic to the rotation structure. The “randomisation per farmer and per crop” analysis considers the percentage of coefficient of variation (%CV) of the proportion of fields of a farmer with a specific crop during a specific time period (one value per farmer, per crop, per time period, cf. Figure 5.2 (a)). The range of expected values is obtained by calculating the %CV from 1,000 simulations of random starting crop for each crop rotation of the fields (e.g. temporal shifting of crop sequences). Those randomised values provide an estimate of the range of %CV which is physically possible with the rigidity of the crop rotations. The observed %CV, derived from the survey, may be plotted on the graph of the randomised %CV, cf. Figure 5.4. The percentage of randomised %CV, which are higher than the observed %CV (referred as %R), allows the determination of whether the observed %CV is significantly aggregated or homogeneous, cf. Figure 5.3.

Randomisation test:



Randomisation two-tailed test with the following hypothesis and probability:

Left tail: H_0: the temporal pattern of groups of fields is random.
H_1: the temporal pattern of groups of fields is homogeneous.
$p = [(100 - \%R) / 100] \times 2$
if $p < 0.05$, H_0 is rejected (significant homogeneity)

Right tail: H_0: the temporal pattern of groups of fields is random.
H_1: the temporal pattern of groups of fields is aggregated.
$p = [\ \%R / 100] \times 2$
if $p < 0.05$, H_0 is rejected (significant aggregation)

With: %R = percentage of randomly simulated %CV values higher than the observed %CV value.
CVp = proportion of randomly simulated %CV values lower than the observed

Figure 5.3: Crop temporal variability: definition of randomisation test per farmer and crop.

Two examples are shown in Figure 5.4. The oilseed rape crop of farmer 3 is not significantly different from a random temporal pattern ($p = 0.874$). However, the spring barley crop of farmer 8, presents a significant level of temporal aggregation ($p = 0.006$). Moreover the observed %CV (around 180%) corresponds to the highest level of variability obtained from 1,000 random simulations, denoting the most important level of aggregation possible with the structure of the crop rotation of farmer 8.

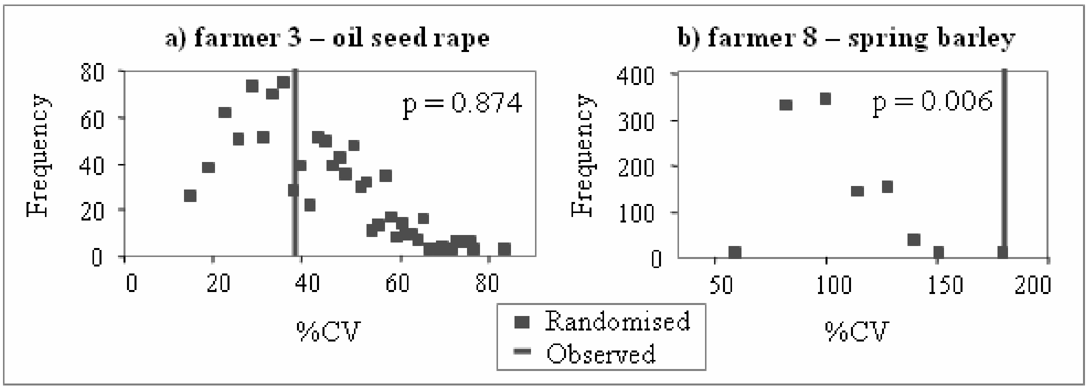


Figure 5.4: Crop temporal variability: two examples of randomisation points and observed value per farmer and crops,

This analysis takes in account the constraints of crop sequences to evaluate the significance of temporal pattern in comparison to a random pattern determined within the rotation constraints. The effect of the structure of the crop rotation itself is not studied. This analysis highlights the temporal patterns of crops, which are induced by the farmers' choices of starting crop for each crop rotation. The main disadvantage of this analysis is the use of a randomisation test, which requires more intensive and lengthy setting up and processing.

5.1.3 Spatial pattern of crops

The spatial pattern of crops is investigated by studying the spatial configuration of the crop's fields. Two approaches are presented below. The first one defines the fine spatial pattern by considering fields' neighbours, while the second one considers only general pattern by using distances separating fields.

5.1.3.1 Fine spatial pattern (chi-square test)

To determine the fine-scale spatial pattern of the crops, the neighbouring crops of each crop were considered. The definition of the neighbours is a fundamental parameter of the analysis, and should always be clearly indicated. The neighbours of a field may be either the strict neighbours: fields with a common boundary; or buffer neighbours: fields within X metres of each other. The "observed crop neighbours" is the number of times a crop in a field is neighbour of another crop in another field (a single boundary will thus be counted twice). An example is shown in Figure 5.5

a) Matrix of crops neighbours of the landscape presented in b)

		Crop in neighbouring fields				Total number of neighbours
		W	OSR	wB	sa	
Crop in fields	W	8	8	8	1	25
	OSR	8	2	3	2	15
	wB	8	3	8	4	23
	sa	1	2	4	4	11
Total number		25	15	23	11	74

b) Landscape with four crops (part of Burgundy study area)

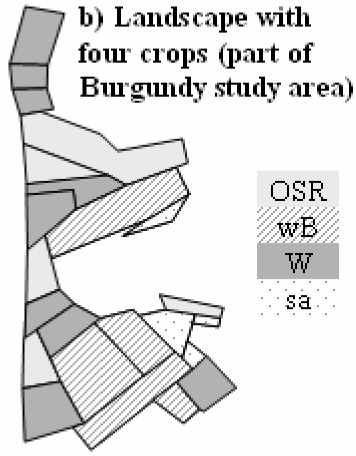


Figure 5.5: Fine spatial pattern: an example of observed values of the spatial chi-square.

In Figure 5.5, the two main constraints that exist in the table are that (i) the row and the column sum for a crop are equal, and (ii) the table is symmetrical (e.g.: 8 oilseed rape fields are neighbours to wheat fields, and 8 wheat fields are neighbours to oilseed rape fields). The observed crops neighbours are then compared with the expected number of crop neighbours, which are calculated for an even distribution, while respecting the constraints of the matrix table. The definition of the calculations of the expected values on neighbouring crops is presented in Figure 5.6.

The Spatial chi-square ($\chi 2$):

i = { 1, 2, …, n}: crop type
j = { 1, 2, …, n}: crop type neighbours
Oij = observed number of neighbours

- $N_i = \Sigma_j O_{ij}$ and $N_j = \Sigma_i O_{ij}$
- Ni = Nj when i = j
- F = Σifi ; F = total number of fields

Observed matrix of crops neighbours:

		j (crops neighbours)				
		1	2	…	n	$\Sigma_j O_{ij}$
i (crops)	1	O_{11}	O_{12}	O_{ij}	O_{in}	N_i
	2	O_{21}	O_{22}	O_{ij}	O_{in}	N_i
	…	O_{ij}	O_{ij}	O_{ij}	O_{in}	N_i
	n	O_{n1}	O_{n2}	O_{nj}	O_{nn}	N_i
	$\Sigma_i O_{ij}$	N_j	N_j	N_j	N_j	$\Sigma i N_i$

E_{ij} = expected number of crop neighbours:

if i = j: $E_{ij} = \dfrac{(f_i - 1)\times F \times (N_i)^2}{f_i \times (F-1)\times \Sigma_i N_i}$

if i ≠ j: $E_{ij} = \dfrac{G \times N_i N_j}{\Sigma_i N_i}$; with $G = (XY)^{1/2}$

$$X = \frac{[f_i \times (F-1)\times \Sigma_i N_i] - [(f_i - 1)\times F \times N_i]}{f_i \times (F-1)\times \Sigma_{K \neq i} N_K}$$

$$X = \frac{[f_j \times (F-1)\times \Sigma_j N_j] - [(f_j - 1)\times F \times N_j]}{f_j \times (F-1)\times \Sigma_{K \neq j} N_K}$$

Figure 5.6: Fine spatial pattern: definition of chi-square analysis.

The chi-square test may be carried out in five different ways, which are listed in Table 5.5. Each test is evaluating differently separate components of crop spatial pattern. Those tests are one-tailed chi-square tests, with:

H_0: Observed and Expected values are not significantly different

H_1: Observed and Expected values are significantly different, indicating non-random spatial pattern.

- if $p > 0.05$ → H_0 is true (not significantly different)
- if $p < 0.05$ → H_0 is false and H_1 is true (significantly different, temporally aggregated or homogeneous)

The differentiation of spatial aggregation from homogeneous pattern may be carried out in two ways. At first a “significantly different from random” result in the d chi-square test would indicate an aggregated spatial pattern because the neighbours of each fields, would predominantly be of alike-crops, the unlike-crops would then be highly different from random. Whereas a non-significant result would indicate homogeneity, as the number of unlike-crop neighbours would not be very different from random. The second way is to investigate the relative values of observed and expected values. Observed values higher than expected ones denote an aggregated pattern, whereas the inverse case indicates a homogeneous pattern.

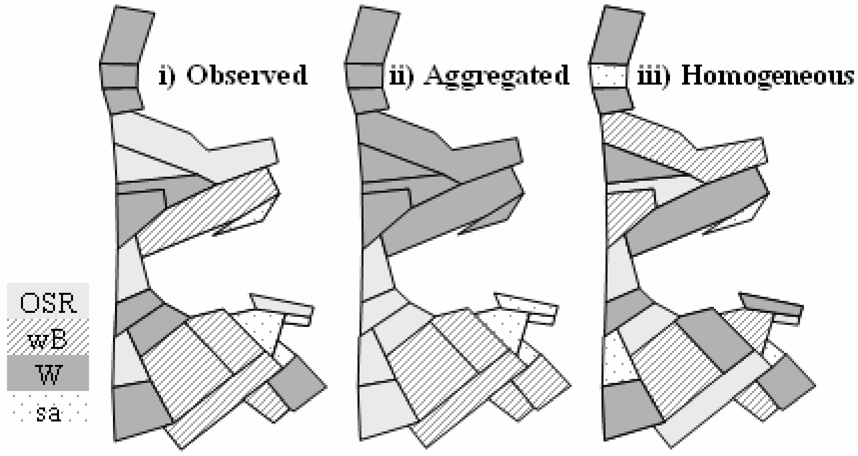
Table 5.5: Fine spatial pattern: definition of individual chi-square tests.

Tests	Process	Values considered	Output	Degree of freedom	Evaluation
a	Calculate X^2 for each Oij	On and above diagonal (j >= i)	One overall value	n(n-1)/2	Overall spatial pattern
b	Calculate X^2 for each Oij	Entire row	One value per crop	n-1	Spatial pattern per crop
c	Calculate X^2 for each Oij	Diagonal values (j = i)	One value per crop	1	Spatial pattern of alike-crop neighbours
			One overall value (sum of value of each crop)	n-1	
d	Calculate X^2 for each Oij	Above diagonal (j > i)	One overall value	(n-2) *(n-1)/2	Spatial pattern of unlike-crop neighbours
e	Calculate two X^2 for on diagonal ($O_{like} = \Sigma Oij$ with i = j and $E_{like} = \Sigma Eij$ with i = j) and above diagonal.	On (j = i) and above diagonal (j > i) separately	One overall value (sum of value of on and above diagonal)	1	Overall spatial pattern, more general than a.

Alike-crop neighbours: identical crops, which are neighbours (wheat-wheat neighbours)

Unlike-crop neighbours: different crops, which are neighbours (wheat-oilseed rap neighbours)

In Figure 5.7, results of the spatial chi-square tests are presented for the example shown in Figure 5.5 and for two examples of extreme crop patterns: aggregated and homogeneous.



	i) Observed	ii) Aggregated	iii) Homogeneous
a	non-significant	significant	significant
b	¾ N.S.	significant	¾ N.S.
c	non-significant	significant	significant
d	non-significant	significant	non-significant
e	non-significant	significant	significant

significant: significantly non random, with $p < 0.05$
¾ N.S.: three out of four crops are non-significant

Figure 5.7: Fine spatial pattern: examples of spatial chi-square results.

The definition for a,b,c,d, and e are provided in Table 5.5

The analysis shows that the landscape with the aggregated crop pattern is significantly non-random under the 5 tests. The homogeneous landscape present the same results as the aggregated landscape except for the b and d analysis, which respectively show that the spatial pattern of each crop individually is not significantly different for three crops out of four; and the spatial pattern of unlike crops is not significantly different from random. The observed example is not significantly different from random, except for one crop (set-aside) which is significantly non-random. The combination of several spatial chi-square tests described above is able to distinguish between random, aggregated and homogeneous patterns. However the chi-square test on neighbouring crops only takes into account fields which are spatially close. In the case of spatially dispersed fields, the use of the analysis described above would not be appropriate.

5.1.3.2 General spatial trend (E analysis)

To evaluate the spatial pattern of crops through the landscape, the aggregation of fields is considered. For this analysis, each field is represented by its centroid; this enables a more flexible handling of the fields' features, and does not require adjacency between fields.

A system of fields is considered (cf. Figure 5.8a) with fields grouped by one common attribute such as the same crop, the same crop rotation, or the same farmer. To be considered, a group must be composed of at least two fields, examples of groups are represented in Figure 5.8c for highly spatially aggregated group of fields and Figure 5.8d for a lower spatial aggregation.

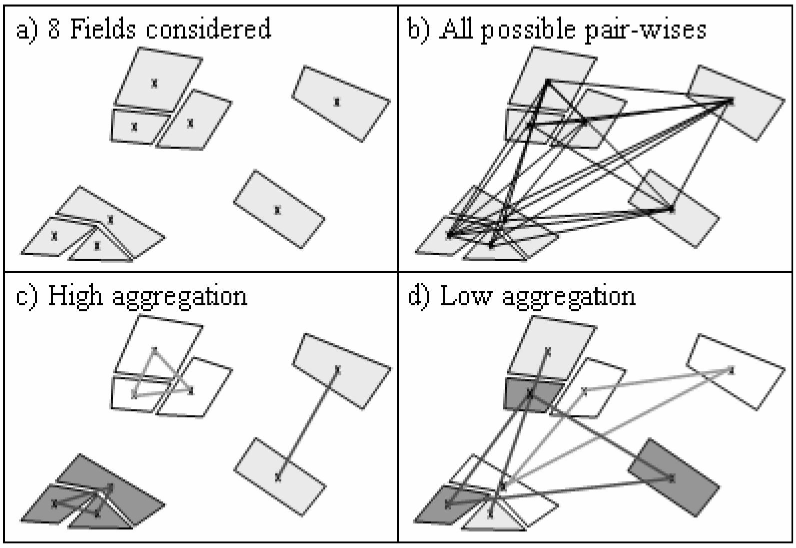


Figure 5.8: General spatial trend (E analysis): visual representation of the test

The mean distance between centroids within each group is calculated ($\frac{\sum Wi}{Ni(Ni-1)}$, refer to Figure 5.9 for a full description of the variables and calculations), and the sum of the mean distance of all the groups is determined (Σ groups = $\sum \frac{\sum Wi}{Ni(Ni-1)}$). The mean distance of the centroids of all the fields considered (Σ all fields = $\frac{\sum Wm}{m(m-1)}$) is also calculated, cf. Figure 5.8b. Then the ratio between the (Σ groups) and the (Σ all fields), tempered with the number of groups composed of more than one field, is used to determine the final values called the E analysis.

A – "E calculation":

Variables considered:

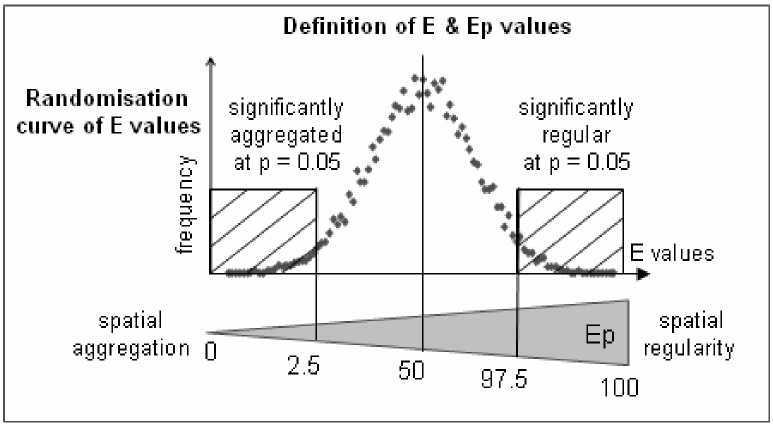
m = total number of fields considered
q = number of groups with more than one field
Ni = number of fields within a group.
W = distance between two fields' centroids.

With:

i = 1, 2, …, q
m = Σi Ni; m(m-1) / 2 = total number of pairwise
Ni (Ni-1) / 2 = total number of pairwise within a group
Σ (Wi) = sum of the distances between all centroids within a group
Σ (Wm) = sum of the distance between all the centroids of the fields considered

$$E = \frac{m(m-1)\times\sum_i\left[\frac{\sum(Wi)}{Ni(Ni-1)}\right]}{\sum(Wm)\times q}$$

B - Randomisation test:



Randomisation two-tailed test with the following hypothesis and probability.

Left tail: H_0: the spatial pattern of groups of fields is random.
H_1: the spatial pattern of groups of fields is aggregated.
p = [Ep / 100] × 2

Right tail: H_0: the spatial pattern of groups of fields is random.
H_1: the spatial pattern of groups of fields is regular (homogeneous).
p = [(100 – Ep) / 100] × 2

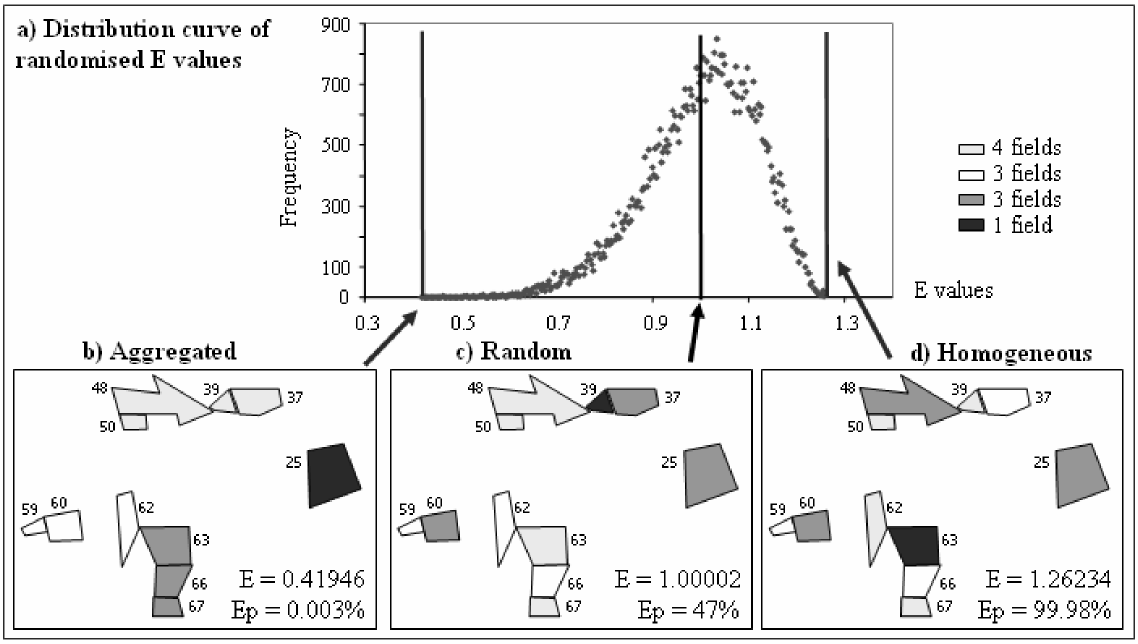
With: Ep = percentage of randomly simulated E values lower than the observed E value.

Figure 5.9: General spatial trend (E analysis): definition of the test.

A low E value indicates a high level of spatial aggregation of the fields and a high E value a low aggregation, i.e. spatial regularity. E values close to one indicate a random spatial distribution of the groups (not aggregated or homogeneous). The range of possible individual E values or numbers is influenced by the spatial disposition of the fields, the distance between the fields, and the number of fields within groups. Each set of fields has various possible E values.

To provide a “landscape independent” index, the observed E value must be compared with the range of possible E values for this specific landscape. This can be set up in two steps. At first the range and frequency of E values is obtained by randomly allocating fields to groups over 10,000 simulations, those values will be later referenced as “randomly simulated E values”. When visualised on a graph, the simulated E values show a near-normal distribution curve (cf. Figure 5.10a), and the mean of the values, corresponding to a random distribution, is tending to one. The smoothness of the curve is highly dependent on the number of fields, the number of groups of fields considered, and the spatial configuration of the fields. For example, four fields will yield only six pairwises (distance between two fields’ centroids), whereas 10 fields would yield 45 pairwises, thus increasing the range of possible values.

In the second step, the observed E value is compared to the simulated E values, to determine the likelihood of the observed E value being a random allocation of the fields to groups. Ep is the percentage of simulated E values which are lower than the observed E value. In other words, Ep indicates the probability of reaching the observed E value when grouping the fields randomly. The distribution is considered as two tailed. The left tail (lower than 1) indicates field aggregation and the right tail (higher than 1) represent field homogeneity.



<u>Figure 5.10: General spatial pattern (E analysis): example of aggregation and dispersion of the groups of fields</u>

This example is based on the groups of fields of farmer 3 on the Burgundy study area in 1995.

In the example shown in Figure 5.10, high spatial aggregation (b) is characterised by a very low E value, which occurs in only 0.0031% of the simulated E values. This observed spatial distribution is highly non-random. On the other side of the distribution curve, graph (d) represents highly homogeneous groups of fields (the member of the same group repulse each other), the E value is very high and the Ep reaches 99.98%. Nearly all the simulated E values are lower, thus this configuration is characterised as highly homogeneous. With an E value close to 1 and an Ep value close to 47%, the spatial configuration exemplified in (c) is very close to a random configuration. Those fields are not specifically grouped or homogeneous.

The E analysis can be used for answering different questions; just by changing the system considered (all the fields or the fields of a farmer) and/or by changing the way the fields are grouped (per rotation or per crop). This last cited case is a test for determining the spatio-temporal pattern, to identify if temporally grouped fields are also spatially linked. The three main cases are described on Table 5.6.

Table 5.6: General spatial patterns (E analysis): potential applications of the test.

System considered	Fields grouped by	Temporal resolution	Questions solved
All fields of 1 farmer	Same crop	1 year	Are the fields of a farmer with the same crop, spatially more correlated than a random spatial allocation of crops?
All fields of all farmers	Same crop	1 year	Are the fields with the same crop spatially more correlated than a random spatial allocation of crops?
All fields of 1 farmer	Same crop rotation, same starting crop	Rotations types	Are the fields of a farmer with the same crop rotation spatially more correlated than a random spatial allocation of crop rotations?

The E analysis is based upon a randomisation test, which conditions on the specific observed spatial distribution of fields and their shapes. Hence, the presence of particular shapes (for example long thin fields) is unlikely to interfere with the evaluation of spatial pattern. However, a formal sensitivity analysis would be required to confirm or invalidate this assumption; this analysis was not carried out due to time constraints, and is thus missing from the sensitivity analyses reported in the following section.

5.1.3.3 Sensitivity of the E analysis

Understanding the metric's properties and, more particularly, its sensitivity to variations in landscape's structure, is indispensable in determining the scope and the conditions of its use. Two variations are investigated: (i) the influence of an important spatial discontinuity between fields, and (ii) the influence of the number of groups. For both aims, the fields' configuration is derived from the Burgundy study area, however the fields' groups were solely designed for the purpose of the sensitivity analysis. The twelve fields considered are spatially separated into a north and a south part (cf. Figure 5.11).

For the analysis of the influence of spatial discontinuity between fields, two groups of fields were considered. 10,000 simulations were run on this specific field configuration with random allocation of the fields into groups. The random E values average at 1.0016, which is close enough to 1 to confirm the randomness of the distribution. However the randomisation curve is highly skewed with values grouped into three groups averaging at 0.35, 0.75 and 1.0. This skewness is due to the important spatial gap between the north and south part of the fields, which induces gaps in the range of possible E values. The cumulative frequency curve shows an exponential-shaped increase with decreasing spatial aggregation of fields.

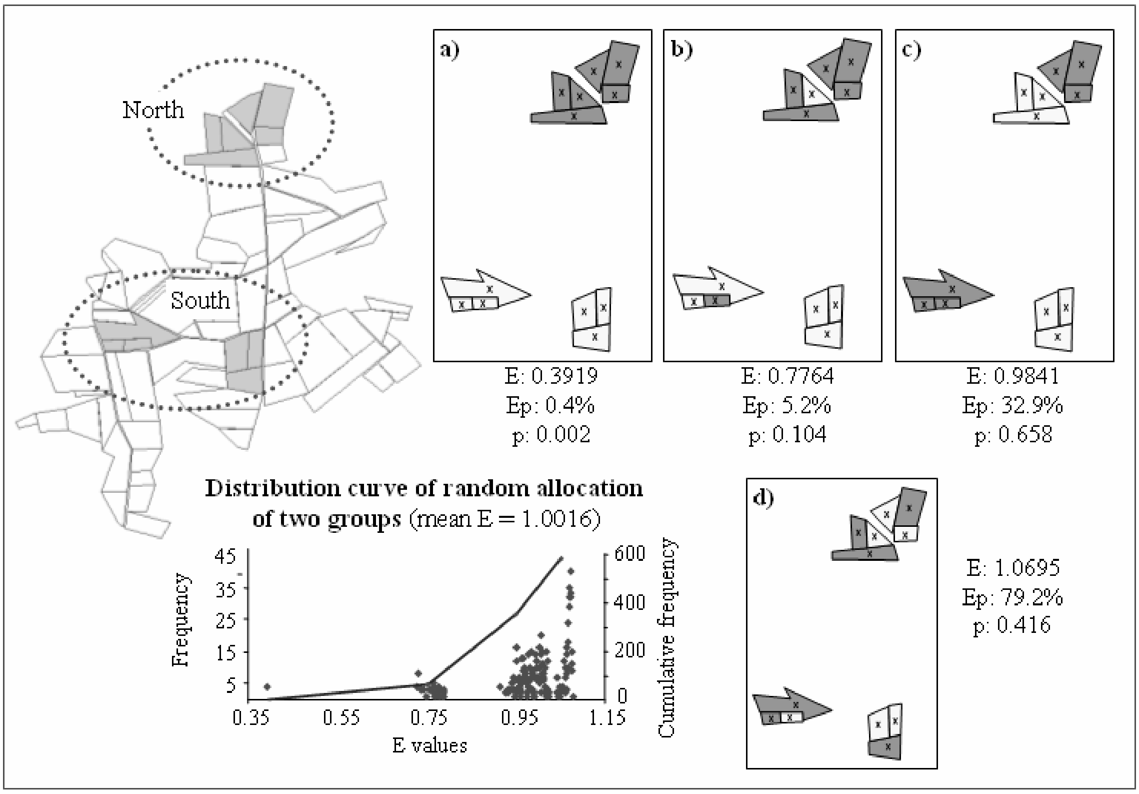


Figure 5.11: General spatial pattern (E analysis): sensitivity analysis on gaps detection.

The E analysis was calculated for specific spatial distribution of the groups of fields, represented in Figure 5.11. The most aggregated spatial configuration, with the groups of fields coinciding with the north and south delimitations, is presented in Figure 5.11 a). Only 0.4% of random E values show the same or more aggregation and the spatial aggregation is significant at the critical level of 95%. By interchanging only two fields between the two groups (cf. b), the probability of reaching this spatial configuration by chance increases and the spatial aggregation is not significant (critical level of 95%), but the value is still low (E_p = 5.2%). In the c) configuration, each group is split in two, half within the north part and half within the south part; thus each of the groups is spatially divided into two parts. The probability of this configuration occurring by chance is evaluated at 32.9%, which is relatively high. The local spatial aggregation of fields is thus not very well differentiated from the random values. However the probability of c) configuration is clearly separated from the probability evaluated for d) (79.2%), where the fields are completely mixed between groups. The E analysis is thus able to measure heterogeneity on groups which are spatially split.

The analysis on the influence of the number of groups of fields was carried out on the same spatial field configuration as indicated in Figure 5.11, however the number of groups is equal in turn to 2, 3, 4, and 6 groups. The number of fields in each group was made even between groups. The curves of randomised E values are shown in Figure 5.12. The mean of the E values is independent of the number of groups (1.00). However the increasing number of groups influences the distribution curve of E values: the range of E values is more continuous (fewer gaps) and wider. The large distance between the north and south fields, impacts on the possible E values (peaks).

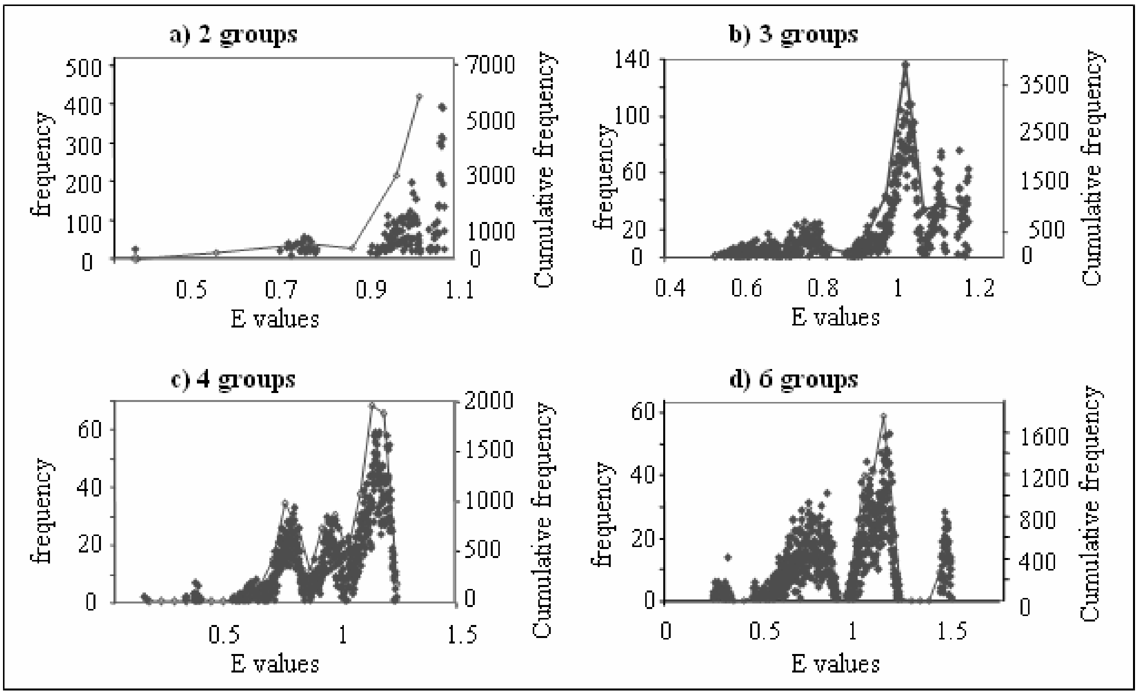


Figure 5.12: General spatial patterns (E analysis): sensitivity analysis on the influence of the number of groups on randomisation curves of E values

Both the number of individual E values and the standard deviation of E values increase logarithmically with the increasing number of groups of fields (cf. Figure 5.13). Therefore, care should then be taken when comparing E values obtained from different numbers of groups.

a) Results over 10,000 simulations	Number of groups			
	2	3	4	6
Mean E values	1.0013	1.0011	1.0003	1.001
Standard deviation of E values	0.0863	0.1286	0.1659	0.2516
Number of individual E values (+/- 0.001)	146	489	629	775

b) Relation between number of groups and number of individual values

c) Relation between number of groups and standard deviation of individual values

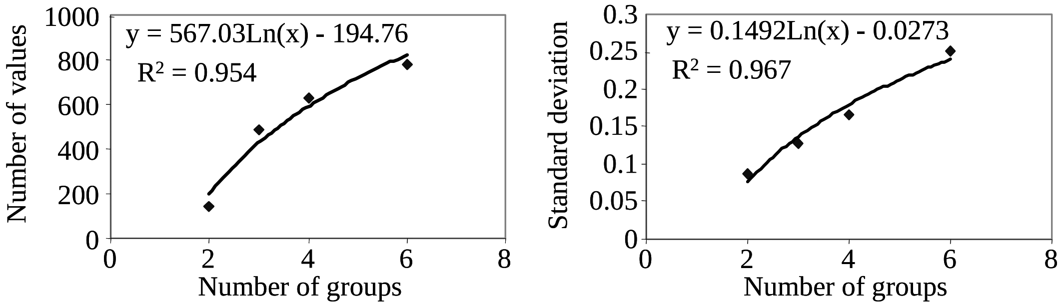


Figure 5.13: General spatial patterns (E analysis): Influence of the number of groups on E values.

In conclusion, the E analysis seems to represent well the general trend of aggregation of the groups of fields, even if groups, which are spatially split are not always very well detected.

5.1.4 Conclusion on pattern analysis

Three temporal pattern metrics, two spatial pattern metrics, and one spatio-temporal pattern metric, which is a particular case of one of the spatial pattern metrics, detailed in this chapter are briefly summarised in Table 5.7. These analyses may be used on the different study sites, when relevant. The exact application of the metrics may differ from one study to another, depending on the available datasets and on their completeness. Further analysis may also be carried out such as descriptive analyses and the analysis detailed in the literature review (cf. Section 3.2.2. Landscape pattern metrics, p.20).

Table 5.7: Summary of spatial and temporal statistical analyses

Pattern studied	Focus	Analysis
Temporal pattern	Crop rotations phasing	Chi-square test per farmer, per rotation
	Crops phasing	Chi-square per farmer, per crop
	Crops temporal variability compared to random simulations	Randomisation test of percentage of coefficient of variation of crops through time (per farmer, per crop)
Spatial pattern	Fine pattern	Chi-square on neighbouring crops
	General trends	E analysis on groups of crops
Spatio-temporal pattern	General trends	E analysis on groups of identical crop sequences

5.2 Crop pattern analyses on landscape datasets

The Burgundy study site was used to analyse the crop patterns, with the specifically designed statistical analysis described above (Section 5.1). The farmers of the site are responsible for the crop grown on their fields. Thus, each farm was considered as an independent unit for decision-making. Farmers with fewer than five fields in the study area were not included in the statistical analyses, as very low number of fields is unfavourable to statistically significant tests.

5.2.1 Temporal pattern

As indicated in Section 3.1.2 (Agronomic constraints), crop rotation is the main driver of temporal pattern. This section reports the findings on the characteristics of the crop rotation, and through three analyses, the temporal phasing of the crops in the Burgundy study area.

5.2.1.1 Description of crop rotations

In the study site, 20 unique crop rotations were followed; they are represented in Table 5.8. Crop rotations might be completely different, differ by only one crop, such as rotation 12 and 13, or the sequence of crops might be altered (cf. rotation 2 and 3). Most of the rotations (80.89%) were based on a three or six year sequences and 78% of fields were following a rotation with wheat, oilseed rape and winter/spring barley.

Table 5.8: Burgundy site: description of crop rotations.

Rotation Id	Crops Sequences A	B	C	D	E	F	Rotation length	Number of fields	Number of farmers using them
0	sa						1	7 (9.7%)	4
1	W	OSR					2	2	1
2	W	OSR	wB				3	1	1
3	W	wB	OSR				3	31 (43.1%)	8
4	W	sB	OSR				3	3 (4.2%)	2
5	W	sB	wB	OSR			4	2	2
6	W	wB	sB	sB	wB		5	1	1
7	W	W	OSR	W	wB	OSR	6	3 (4.2%)	2
8	W	wB	OSR	W	sB	OSR	6	4 (5.6%)	3
9	W	wB	OSR	W	wB	sa	6	2	1
10	W	OSR	wB	W	sa	wB	6	1	1
11	W	sB	wB	OSR	wB	OSR	6	1	1
12	W	OSR	W	OSR	W	sa	6	1	1
13	W	OSR	W	OSR	sa	sa	6	1	1
14	W	OSR	sa	sa	sa	sa	6	2	2
15	wB	OSR	sa	sa	sa	sa	6	2	2
16	sB	sa	sa	sa	sa	sa	6	1	1
17	R	OSR	R	sa	sa	sa	6	1	1
18	R	R	OSR	sa	sa	sa	6	2	1
19	W	W	sa	sa	R	OSR	6	1	1
20	W	sa	sa	sB	wB	OSR	6	3 (4.2%)	1

W = wheat; wB = winter barley; sB = spring barley; OSR = oilseed rape; R = rye; sa = set-aside

The farmers followed crop rotations which suited their individual requirements, as 62% of the crop rotations listed in Table 5.8 were used only by one farmer. However, crop rotation 3, composed of wheat/winter barley/oilseed rape, was used by 8 farmers out of 10 and was used on 43% of the fields of the study area. It was the most widely used crop rotation within the study area.

The respective proportions of each crop transition between one year (n) and the next one (n+1) are shown in Table 5.9; the transitions were weighted by the number of fields following each particular rotation.

Table 5.9: Burgundy site: proportion (%) of crop transitions from year n to n+1.

		Year n + 1						
		Wheat	Winter barley	Spring barley	OSR	Rye	sa	Total
Year n	Wheat	0.93	17.18	3.24	4.17	0	1.39	26.90
	Winter barley	1.20	0	0.28	18.75	0	0.46	20.69
	Spring barley	0	1.90	0.28	2.31	0	0.23	4.72
	OSR	23.38	0.93	0	0	0.23	1.62	26.16
	Rye	0	0	0	0.93	0.46	0.23	1.62
	sa	1.39	0.69	0.93	0	0.93	15.97	19.91
	Total	26.90	20.69	4.72	26.16	1.62	19.91	100

Percentage of the proportion of crop transitions = $\sum_i [\frac{bi \times fii}{lei}] \times \frac{100}{72}$

i: crop rotation

bi: presence of a transition from crop a to crop b in a rotation i (binary data)

lei: length of rotation i

fii: number of fields following rotation i

OSR/wheat was the most represented crop sequence on the whole study area with 23% of the fields showing this crop sequence every year (as a mean). The next most common crop sequences were winter barley/OSR, wheat/winter barley, and set-aside/set-aside. 13 crop transitions each represented less than 1% of the transitions in a year. This table also shows that some crop transitions were not used on the study area. For example, rye was never preceded by another cereal crop such as wheat, winter/spring barley. Spring barley was also never followed by wheat. However the intentionality of the unused crop sequences is unknown, it may be by chance or by design. Greater agronomic knowledge is required to be able to determine the origin and strictness of crop sequences.

The next few analyses were aimed at investigating the temporal pattern of the crops on the Burgundy study site and, more precisely, the phasing of the rotation on different fields. The methods used below were specifically designed for this study site, and are defined in Section 5.1.2. (Temporal pattern of crops, p.46).

5.2.1.2 Crop rotation phasing

The phasing of the fields' rotation was studied, by using the "chi-square analysis per farmer, per rotation" as shown in Section 5.1.2.1 (Crop rotation phasing, p.46). The results are shown in Table 5.10.

Table 5.10: Burgundy site: results of crop rotation phasing test (chi-square results)

Farmer	Rotation ID	Rotation Length	Number of fields	Number of fields per starting crop						Chi-square	Degree of freedom	Probability
				A	B	C	D	E	F			
1	4	3	2			2				4.0	2	0.1353
3	1	2	2	2						2.0	1	0.1573
	3	3	4	2	1	1				0.5	2	0.7788
4	3	3	3	2	1					2.0	2	0.3679
	7	6	2	1					1	4.0	5	0.5494
	9	6	2	2						10.0	5	0.0752
5	3	3	4	1	2	1				0.5	2	0.7788
	18	6	2				2			10.0	5	0.0752
6	3	3	3		2	1				2.0	2	0.3679
	8	6	2						2	10.0	5	0.0752
7	3	3	5	2	2	1				0.4	2	0.8187
8	20	6	3				3			15.0	5	0.0104
9	3	3	10	2	4	4				0.8	2	0.6703
			General Chi-square:							61.20	40	0.0171

Only the crops grown on the fields of farmer 8, following rotation 20, showed a significant level of temporal aggregation ($p < 0.05$), as the three fields started with the same crop out of a choice of six. Three other rotations followed by three different farmers were nearly significantly aggregated (farmer 4 and rotation 9, 5-18, 6-8) with probabilities equal to 0.0752. However most of the rotations studied did not show any significant difference from a random distribution, and no significant level of homogeneous temporal pattern ($p > 0.95$).

The overall low significance of the test might be due to the low degrees of freedom of each test, because only few fields were considered at each time. To increase the degrees of freedom and the power of the test, a general chi-square test was carried out on the dataset presented in Table 5.10. Each chi-square value and each degrees of freedom was summed, to determine an overall chi-square probability of 0.0171; indicating a significant temporal phasing of the rotations.

Consequently, the fields following the same rotations showed an overall significant temporal aggregation, even if each particular farmer-rotation combination was not significantly phased temporally. Yet, this method only considers crop rotations with more than one crop, followed by at least two fields per farmer. Thus nearly 40% of the fields were not included in this analysis. Moreover, the possibility of temporal synchronisation of independent crop rotations was not taken into account. To circumvent the restriction of considering crop rotations individually, the next analysis focuses on the crops instead of crop rotation.

5.2.1.3 Crop phasing

The analysis, “Chi-square analysis per farmer and per crop”, aims at determining if each crop type of a farmer is temporally aggregated or homogeneous. The chi-square analysis was carried out per farmer and per crop over 12 years (multiple of 3, 4 and 6 rotation lengths) independently of the crop rotations (cf. Chapter 5.1.2.2: Crop phasing, p.49), the resulting chi-square probabilities are shown in Table 5.11.

Table 5.11: Burgundy site: crop phasing test (chi-square probabilities).

Farmer	Oilseed rape	Wheat	Spring barley	Winter barley	Rye	Set-aside
3	0.9022	0.9754	-	0.7991	-	-
4	0.2843	0.1981	-	**0.0244**	-	1.0000
5	0.9174	0.9985	-	0.9985	0.1411	0.7390
6	0.5304	0.3473	0.6071	0.3882	-	0.9895
7	0.9699	1.0000	-	0.9895	-	0.9985
8	**0.0373**	0.0950	**0.0046**	**0.0046**	-	0.2330
9	0.9624	0.8228	0.2330	0.9957	-	-

“-“: less than two fields with this crop in their rotation;
bold: p < 0.05, indicating significant temporal aggregation;
shaded: p > 0.95, indicating significant temporal homogeneity.

Oilseed rape, spring and winter barley crops of farmer 8 showed significant temporal aggregation (p <0.05). This result agrees with the findings of the previous analysis (chi-square per farmer per rotation) which indicated a significant temporal aggregation of the rotation 20, composed of the same three crops listed above plus wheat and set-aside. However the other rotations followed by farmer 8, might have diluted the temporal aggregation of wheat and set-aside, even if the wheat crop was actually near significant for temporal aggregation (p = 0.095).

Winter barley of farmer 4 showed a significant temporal aggregation ($p < 0.05$). In the previous analysis, farmer 4 had fields following rotation 9, which were nearly significantly in phase ($p = 0.0752$). The significant temporal phasing of winter barley might be due to the synchronisation of other rotations with rotation 9.

In contrast to the crop rotation phasing analysis, 37% of the crop types of all the farmers showed significant temporal homogeneity ($p > 0.95$). The even repartition of the crops through the years might be due to the structure of the crop rotations themselves (minimum return period of each crop), or due to farmer choice of spreading crops over years by shifting crop rotations, in order to control the market and agronomic risks. This spreading out of the crops through years was not detected by the previous analysis. Finally 50% of the crop types of the farmers were not significantly distinct from or similar to an even temporal distribution.

In conclusion from this test, 37% of farmer's crop types were significantly spread homogeneously over the years, while only 13% showed a significant temporal heterogeneity. This analysis brings new insights on the temporal pattern of crops; however its non-consideration of the constraint of the crop sequence derived from rotation rigidity limits it. This issue is addressed by the randomisation per farmer and per crop studied in the next section.

5.2.1.4 Crop temporal variability compared to random simulation

For this test, "randomisation per farmer and crop", for each farmer, the percentage of variation over the years (%CV) of the proportion of his fields growing a crop was tested, (cf. 5.1.2.3: Crop temporal variability compared to random simulation, p.51). The observed values (%CV) were compared against 1,000 randomisations by simulating random starting crops for each rotation (%R: percentage of randomised values with higher temporal variability than the observed value), to determine if the observed value was significantly aggregated or homogeneous.

Table 5.12: Burgundy site: results from crop temporal variability test

	Oilseed rape (OSR		Wheat (wW)		Spring barley (sB)		Winter barley (wB)		Rye (R)		Set-aside (sa)	
Farmer	%CV	%R	%CV	%R	%CV	%R	%CV	%R	%CV	%R	%CV	%R
3	38.4	56.3	30.4	73.2	-	-	51.3	45.7	-	-	-	-
4	71.6	21.2	66.7	12.3	-	-	100.0	4.0	-	-	11.6	100
5	51.1	67.6	36.9	100	-	-	36.9	100	120.6	10.1	**44.8**	**2.4**
6	64.8	36.0	64.9	20.6	116.8	19.2	77.5	21.8	-	-	39.1	64.6
7	42.6	78.9	21.2	100	-	-	39.1	87.4	-	-	36.9	100
8	143.1	3.3	121.0	2.8	**180.9**	**0.3**	**180.9**	**0.3**	-	-	97.7	16.2
9	30.4	64.4	39.8	48.3	159.5	35.3	24.5	81.0	-	-	-	-

bold: significant temporal aggregation; shaded: significant temporal homogeneity.

From Table 5.12, out of 30 valid farmer-crop combinations (more than one field), five (17%) were significantly homogeneous and three (10%) were significantly aggregated at a level of confidence of 95%. Each of these categories was examined in turn.

The significantly homogeneous fields were the set-aside crop of farmer 4, the wheat and winter barley crop of farmer 5, and the wheat and set-aside crop of farmer 7 ($p < 0.05$). These results agree with the previous test. All the farmer-crop combinations with %R higher than 50% (indicating a certain degree of heterogeneity), were also identified in the "crop phasing analysis" as nearly significantly heterogeneous ($p > 0.90$) in Table 5.11.

Farmer 8 presented the highest level of temporal heterogeneity/aggregation, with spring and winter barley being significantly aggregated ($p < 0.05$, in concordance with the previous analysis). On the other hand, oilseed rape in this test was only close to significant aggregation ($p < 0.07$), instead of being significantly aggregated as in the previous test. Wheat, as in the previous test, was close to being significantly aggregated. In contrast to the previous tests, set-aside fields of farmer 5 were significantly in phase ($p < 0.05$) when compared with all possible temporal configurations (random starting crops). However, most of the farmer-crop combinations had no significant homogeneous or aggregated temporal pattern (21 combinations out of 28).

The results of this test show that even when considering the constraint of the crop sequences, five farmer-crop combinations (wheat twice, winter barley and set-aside twice, equivalent to 17% of crops) were significantly homogeneous. Some farmers were thus voluntarily synchronising the starting crop of rotations, to spread some crops over the years, whereas for three other crops (spring/winter barley and set-aside, equivalent to 10% of crops), the farmers were voluntarily aggregating them. 73% of other farmer-crop combinations did not show any significant temporal pattern.

5.2.1.5 Conclusion on temporal pattern of crops

The three temporal pattern analyses detailed above do have consistent results, even if each one of them investigated the temporal pattern of crops using a different methodology.

The crop rotation phasing analyses showed that, overall, fields following the same crop rotation were significantly temporally aggregated, as they tended to start with the same crop. Then, when each farmer's crop individually was analysed (crop phasing analyses), independently of crop rotations, more than one third of them were significantly temporally homogeneous; whereas only one eighth were significantly temporally aggregated. However as this analysis did not take in account the structure of crop rotations, the homogeneity detected might be an artefact due to return period of crops within the rotation. In the last analysis, taking into account crop rotation structure, (crop temporal variability compared to random simulation), one tenth of farmer's crops were significantly aggregated, whereas one sixth were significantly homogeneous.

Two main conclusions may be drawn on the Burgundy study site. Firstly, overall, farmers' fields following the same rotations tended to be temporally grouped. Secondly, between rotations, no clear rules of temporal pattern were detected; the degree of temporal homogeneity or aggregation was farmer and crop dependent.

Fields with the same rotation might be grouped temporally by farmers, in order to ease their management by coordination. On the other hand, fields with different crop rotations might be used to alter the temporal pattern of crops, in order to spread risks, or they might be aggregated to ease management or respond to particular market tendency.

Finally the degree of temporal phasing is thus an important component of crop patterns, and should be included in the modelling of crop pattern in the agricultural landscape. The next section investigates the existence of spatial patterns of crops on the Burgundy study area.

5.2.2 Spatial pattern

The crop spatial pattern in the study area may be studied at two levels: (i) at the study area level without taking into account the individuality of the farms, and (ii) at the farm level, within individual farms. These were both examined in turn.

5.2.2.1 Crop spatial repartition at the study area level

The spatial chi-square tests for fine pattern and the E analysis for more general pattern (cf. 5.1.3. Spatial pattern of crops, p.53) were carried out on the Burgundy study area as a whole for evaluating the yearly spatial pattern of crops from 1994 to 1997, at first using all the crops cf. Figure 5.14, and then with only the three main crops.

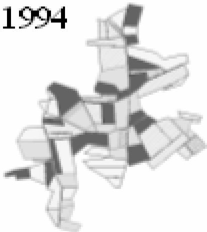
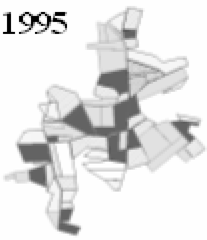
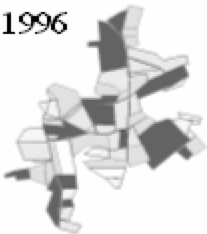
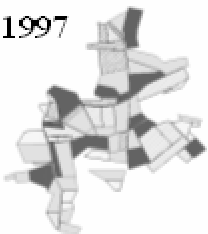
	Chi-square tests	E analysis
1994 	*a: significant* b: 1/6 significant (p-sa) c: significant *d: non-significant* e: non-significant → **Close to Homogeneity**	p = 0.9508 E = 1.0031 Ep = 47.54% **Not significant**
1995 	*a: significant* b: 2/7 significant (sB, R) c: significant *d: non-significant* e: non-significant → **Close to Homogeneity**	p = 0.2318 E = 0.91142 Ep = 11.59% **Not significant**
1996 	*a: significant* b: 2/6 significant (R, a-sa) c: significant *d: significant* e: significant → **Aggregation**	p = 0.0234 E = 0.8367 Ep = 1.17% **Significantly aggregated**
1997 	*a: non-significant* b: 1/6 significant (a-sa) c: significant *d: non-significant* *e: non-significant* → **Not significant**	p = 0.2028 E = 0.9524 Ep = 10.14% **Not significant**

Figure 5.14: Burgundy site: fine and general spatial pattern tests of the crops.

The E analysis investigated the general pattern of crops. For each year, the observed spatial configurations were within the left tail of the randomised distributions (Ep <

50%), indicating a tendency towards aggregation rather than homogeneity. The spatial patterns of crops in years 1994, 1995, and 1997 were not significantly aggregated; however in 1996 the crop pattern was significantly aggregated at the critical level of 95%.

The chi-square tests investigated the crop spatial pattern at a finer resolution. At first, the overall spatial pattern, as determined by the “a” test, showed a significant spatial non-randomness for the crops grown in 1994, 1995, and 1996. However in 1997, the crop pattern was not significantly different from a random allocation; which was reinforced by the non-significant results of test b, d and e. Thus the crop pattern in 1997 was identified as not being significantly different from random.

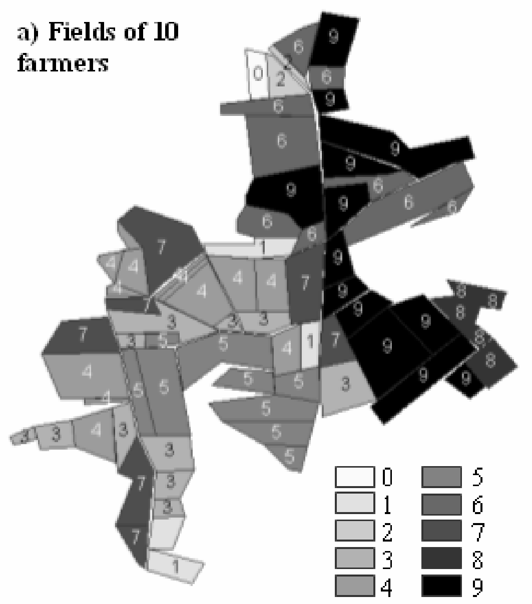
As indicated above, the general crop pattern for 1994, 1995, 1996 were identified as non-random (test a), however the non-randomness might arise from either a spatial aggregation or homogeneity. Unlike-crop neighbours (crops neighbours of a different type of crops, cf. test d) were significantly non-random in 1996, indicating an overall spatial aggregation of the crops. Unlike-crops in 1994 and 1995 were not significantly different from randomness, indicating homogeneity in the general spatial crop pattern.
The b test investigates the spatial pattern of each crop. For each year, very few crops had a significantly non-random spatial pattern: permanent set-aside in 1994; spring barley and rye in 1995; rye and autumn set-aside in 1996; and autumn set-aside in 1997.
Moreover, constant divergences, from 1994 to 1997, of the observed values from the expected values were as following: (i) some crops were more spatially aggregated than expected: wheat / oilseed rape, wheat / spring barley, and permanent set-aside / permanent set-aside; (ii) some crops were less aggregated than expected (more homogeneous): wheat / winter barley, oilseed rape / winter barley, winter barley / spring barley, spring barley / autumn set-aside. The identification of such spatial particularities of crops is indispensable for modelling realistically crop spatial pattern.

The same analysis was carried out on only the three main crops (wheat, oilseed rape and winter barley) from 1994 to 1997. The spatial pattern of the crops was not significantly different from random spatial pattern for every analysis at the fine and more general level (5 chi-square tests and E analysis).

In conclusion, when considering all the crops, no clear tendency of fine or general spatial pattern existed through the years; as three years were not significantly different from random, and one year was significantly aggregated. Moreover, the spatial patterns of the three main crops were not different from random pattern. The landscape level might not be adequate for studying the crop spatial repartition, as the spatial pattern of crops is influenced by the farmer's fields and his strategy. Studying spatial pattern at the farmer level would thus be advisable. But before that a preliminary analysis of the farms' spatial repartition is presented.

5.2.2.2 Spatial repartition of farmers

The 72 fields within the study area were part of 10 farms, which are represented in Figure 5.15, a). The pattern of the farms presents the first degree of spatial pattern in the study area.



b) Results of E analysis

Mean E	1.0013
Standard deviation E	0.0677
Observed E	0.5730
Ep (%)	< 0%
Probability (p)	0.00

c) Results of spatial chi-square tests

Overall spatial pattern (a):	significant
Spatial pattern within Farmers fields (c):	significant (except farmer 1 and 7)
Spatial pattern of different farmers fields (d):	significant
General overall spatial pattern (e):	significant

significant: significantly different from random. Farmer 0 is not considered in those tests.

Figure 5.15: Burgundy site: general spatial pattern test (E analysis on farmers' fields).

The actual spatial distribution of the fields was highly non-random as shown by the results presented in Figure 5.15 b) and c). By running 2,000 simulations of the different allocation of the fields to farmers, no configuration had an E value as low. The configuration was thus considered as highly aggregated. This example shows that the E analysis was particularly good at identifying low levels of aggregation. A very high level of aggregation, such as if all the fields of each farm were adjacent, would not be differentiated from the actual configuration. The results of the spatial chi-square tests identify that the fields overall were significantly aggregated by farmers (a, c, d, and e significant and farmer 0 was not considered). Only farmer 1 and 7 were not significantly different from random (from test c).

The spatial pattern of crops at the farm level was not conditioned by the spatial pattern of the fields of each farmer, as the location of the fields was considered as fixed. However, even if the fields of each farmer were aggregated, few fields were really adjacent. Thus the use of spatial chi-square tests per farmer is not adequate, the E analysis would then be preferred as it does not consider direct crop neighbours.

5.2.2.3 Spatial repartition of crops for each farmer individually

For investigating the spatial pattern of crops for farmers individually, only the E analysis was used, because in many cases the farmer's fields were not adjacent preventing the use of spatial chi-square tests. Most of the farmer's crops did not show any significant spatial pattern, with the exception of significant aggregation for farmer 9 in every year and farmer 4 and 7 in 1996, Table 5.13. No farmers showed a significant level of crop spatial homogeneity.

Table 5.13: Burgundy site: general spatial patterns test (E analysis results on crops).

Farmers	1994		1995		1996		1997	
	E	Ep	E	Ep	E	Ep	E	Ep
3	1.053	66.80	0.794	5.32	0.892	30.74	0.953	41.12
4	0.796	6.68	0.686	3.59	0.631	1.63	0.624	2.97
5	0.781	23.38	0.971	46.65	1.122	67.97	1.043	57.44
6	0.553	2.72	0.660	10.05	0.975	37.87	0.619	5.62
7	0.512	4.99	0.351	5.15	0.328	2.20	0.679	11.38
8	0.802	11.19	0.802	10.27	0.802	10.60	0.802	10.85
9	0.610	0.89	0.590	1.11	0.664	0.42	0.664	0.52

Shaded cells: significantly aggregated spatial pattern ($p<0.05$)

In conclusion at the farmer level, crop spatial pattern were not significantly different from random, except for one farmer (farmer 9), who consistently had an aggregated spatial repartition of crops through the years.

5.2.2.4 Conclusion on spatial pattern of crops

The analyses on the spatial pattern of crops on the Burgundy study area were carried out at two scales: the study area level and the farm level. At the study area level, when considering all the crops, no consistent spatial pattern of crops over four years was found. When considering only the three main crops, spatial patterns significantly different from random were not found. At the farmer's level, the crops were mainly not different from random, even though one farmer showed a consistent aggregated spatial pattern through the years.

In conclusion, no clear crop spatial patterns were found on this dataset, at the study area level. Moreover, the results would suggest that Burgundy farmers, with the exception of farmer 9, do not widely use the spatial blocking of fields, which is the common management of adjacent fields to ease the workload. This technique is being used more and more in England in particular (Orson, 2005).

5.2.3 Spatio-temporal pattern

To study the correlation between the spatial and the temporal patterns of crops, the spatial pattern of temporal groups of crop sequences was analysed. To study the spatial pattern, the analysis of the general trend with the E analysis was chosen instead of the fine pattern analysis with the chi-square test, as the E analysis does not require the use of neighbouring crops. The analysis would thus be more polyvalent and flexible for most landscape studied. The E analysis was carried out by following the definition set up in Section 5.1.3.2 (General spatial trend (E analysis), p.57). The fields considered follow the same crop rotation, but each group had a different starting crop (temporally in phase).

In order to study the spatio-temporal pattern of crops in a meaningful way, only farmers with many fields following the same crop rotation were investigated. In the case of the Burgundy study site, only rotation 3 of farmer 7 and 9 met those criteria. The results obtained are reported in Figure 5.16 and Figure 5.17.

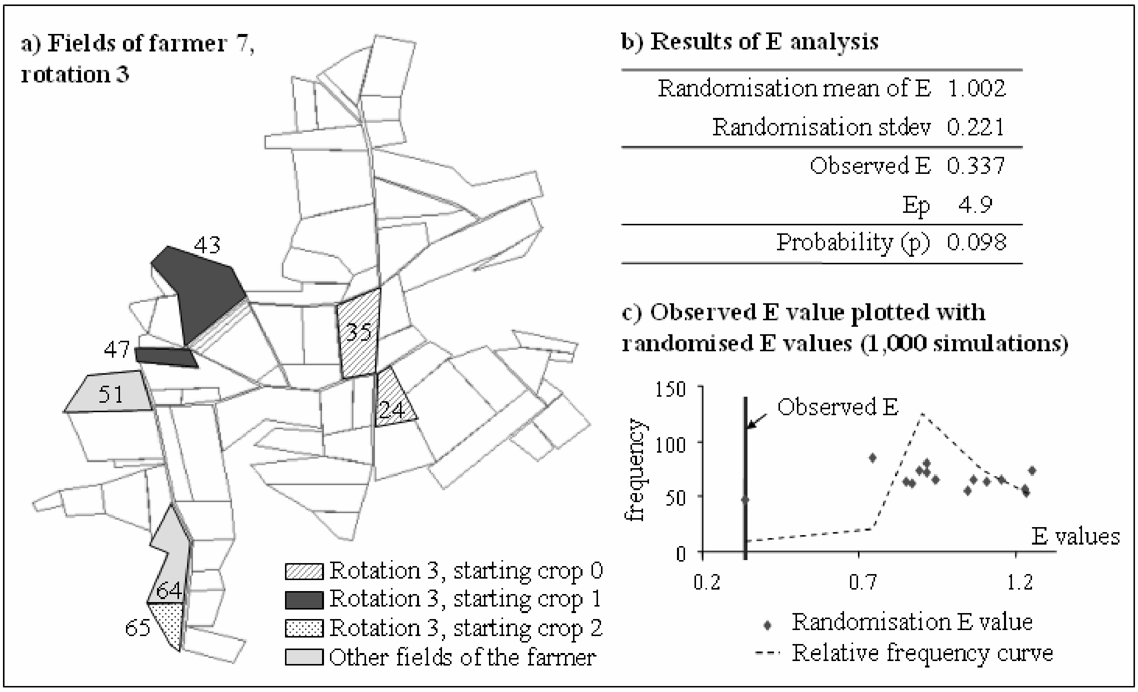


Figure 5.16: Burgundy site: general spatio-temporal pattern (E analysis), on Farmer 7, rotation 3.

Farmer 7 had five fields following crop rotation 3 (wheat / winter barley / oilseed rape), with two fields starting with wheat, two with winter barley and one with oilseed rape (cf. Figure 5.16). The groups of temporally aggregated fields showed the highest possible level of spatial aggregation, as the observed E value matched the lowest E value obtained through 1,000 random simulations. However the spatial pattern was not significantly aggregated at the critical level of 95%. This extreme configuration was occurring too often to be significantly aggregated in comparison with random allocations.

Farmer 9 had ten fields with crop rotation 3, with four starting with oilseed rape (starting crop 2), four with spring barley (1) and two with wheat (0). The observed E value corresponded to 1.3% of values obtained from 1,000 random simulation of starting crops. With a probability lower than 0.05, the crop spatial and temporal pattern of the fields of farmer 9, following rotation 3 was significantly aggregated. This specific case testifies the possibility of significant spatio-temporal aggregation of crops. This result explains the consistency through the years of the high spatial aggregation of crops of farmer 9.

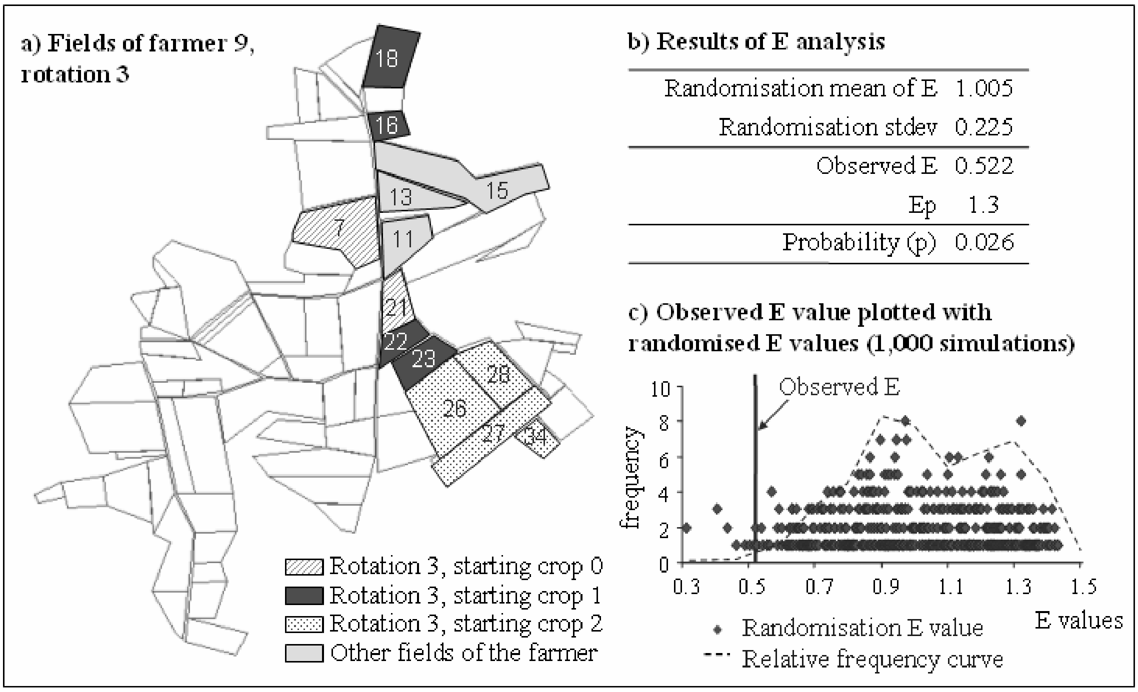


Figure 5.17: Burgundy site: general spatio-temporal pattern (E analysis), on farmer 9, rotation 3.

Thus, only farmer 9 had a significant level of spatio-temporal aggregation of crops. This farmer exemplified the possibility of positive correlation between spatial and temporal pattern of crops. However the use of this analysis is restricted to farmer’s fields, which are following the same rotation. Moreover, sufficient numbers of fields are required in order to detect results significantly different from random.

5.2.4 Conclusion on spatial and temporal pattern analyses

The analysis presented in this chapter examined the spatial and temporal pattern of crops grown on the Burgundy study site from 1994 to 1997. The analysis demonstrated the presence of significant pattern both spatially and temporally different from random, even if not always widely represented.

For the temporal pattern of crops, two main conclusions were drawn: (i) overall, farmer’s fields with the same rotation were temporally grouped, and (ii) when considering all crops, regardless of their rotations, no consistent temporal pattern was detected; they were farmer and crop dependent.

For spatial pattern of crops at the study area level, no consistent pattern through years was found when all the crops were considered. When studying only the three main crops, their spatial patterns were random. At the farm scale, the crops were mainly not different from random, except for one farmer, who showed consistently through the years a significant level of spatial aggregation. The same farmer was the only one presenting a significant spatio-temporal pattern, but thus confirming the possibility of positive correlation between spatial and temporal pattern. The analysis demonstrated a limitation of the E analyse, as it failed to spot extreme crop spatial arrangement if its random occurrence was too high. Therefore, it is recommended to use the E analyse only on high number of fields.

The spatial and temporal patterns may arise from several causes. The spatial aggregation of crops in close fields might facilitate the farmer's work (Maxime *et al.*, 1996), however this might also be due to similarity of the environmental conditions (soil types, climate, water access). Both of these causes will be reflected in the use of spatial aggregation of crops. The temporal pattern of crops arises from the crop rotations and the starting crop for each field, which are directly influenced by the farmer's needs for products and their market price, and on his risk management. These parameters are relevant to the aims of the LandSFACTS model.

5.3 Statistical analyses to integrate within LandSFACTS model

Five statistical tests were developed for measuring spatial and temporal patterns of crops. After testing on the Burgundy dataset, they were all able to identify significant patterns, except for the chi-square analysis on fine spatial pattern, which was not adapted to the dataset characteristics. Only the most adapted and versatile tests should be integrated within the LandSFACTS model to provide control to the user on the spatial and temporal patterns of crops. A comparison between the tests is detailed below.

For measuring the temporal patterns of crops, three tests were designed, the two first one based on a chi-square test and the last one on a randomisation test. The "crop rotation phasing" test considers the starting crop of different fields with the same rotation. This test would not be useable across crop rotations, and is not versatile enough to be integrated within the LandSFACTS model. The "crop phasing" test considers the phasing of the crop regardless of the crop rotations, by considering the proportions of fields with each crop for every year. However, this analysis does not take into account the constraints induced by the crop rotations and the return period of the crops. The "crop temporal variability compared to random simulations" test considers the temporal patterns of all the crops, regardless of the rotations. The temporal pattern is measured as the percentage of variability of the number of fields over the years. The observed crop allocation is compared with random initial crops for all the fields, therefore the constraints of the rotations are taken in account. This later test is more versatile than the two previous tests on temporal patterns. It would be even more useful if instead of recording the number of fields with each crop, the area of each crop was taken in account.

For measuring the spatial patterns of crops, two tests were designed, one based on a chi-square test and the second on a randomisation test. The chi-square test measuring fine spatial pattern considers the neighbouring crops of each crop, which are compared to the expected number of crop neighbours from an even distribution. The randomisation test on more general patterns of crops, the E analysis, considers the distances between the centroids of fields with the same crops. The observed crop allocation is compared with a random allocation of the crops to the fields. The randomisation process can take into account restrictions of the spatial extent of the crops. The E analysis has the main advantage of being useable on fields, which are not spatially continuous. The E analysis is also relevant for measuring spatio-temporal patterns. Therefore, the E analysis has a greater versatility than the chi-square test on fine patterns.

Moreover, in general, randomisation tests provide several advantages over chi-square tests. First, the chi-square tests aim solely at determining if an hypothesis is significantly true or not, i.e. whether the observed values are significantly different from the expected values. Randomisation tests, however, tend to be more: (i) versatile as they compare values measuring the degree of pattern, and (ii) adapted to inflexible constraints. The randomisation curve takes in account the constraints of the landscape, for example in the temporal test, the structure of crop rotations is respected. Moreover, the randomisation curves provide a continuous scale from both extreme patterns, i.e. from aggregated to regular patterns, against which an observed pattern may be compared. Therefore, both randomisation tests, i.e. "E analysis" for spatial pattern and "Randomisation test of percentage of coefficient of variation of crops through time" for temporal pattern, were more adapted, and were thus chosen for integration within the LandSFACTS model (Chapter 7).

5.4 Conclusion

In this chapter, five specifically designed statistical analyses were presented to measure the temporal and spatial patterns of crops for integration within the LandSFACTS model. The temporal pattern tests are (i) the crop rotation phasing test, based upon a chi-square test per farmer, per rotation, (ii) the crop phasing test, based upon a chi-square test per farmer per crop, and (iii) the randomisation test of percentage of coefficient of variation of crops through time. The third test is particularly adapted for integration within the LandSFACTS model, as it can be used across all the farmers, rotations, and on any landscape due to the portability of randomisation tests. The spatial pattern tests are (i) the fine pattern test, based upon a chi-square test on neighbouring crops, and (ii) the general trend test, also called the E analysis based on groups of identical crop sequences. The E analysis is particularly adapted for integration within the LandSFACTS model, as it provides a general overview of the trend of crop spatial patterns, and is usable on any landscape due to the portability of randomisation tests. The E analysis is also useful to test spatio-temporal patterns of the crops, i.e. coordination between temporal and spatial aggregations.

The statistical tests were carried out on the Burgundy study site. The following conclusions were drawn: (i) farmers tend to grow the same crops in fields with identical rotations (temporally grouped); (ii) the temporal patterns of the crops tended to be farmer and crop dependent; (iii) spatial patterns of the crops were not consistent between years; (iv) one occurrence of strong spatio-temporal patterns of crops was detected, indicating that a farmer was growing identical crops every year for his/her spatially close fields. The scale dependency of the pattern was noted. In conclusion, the new statistical tests were successful at characterising the crop patterns of a real landscape, and the “E analysis” for spatial patterns and the “Randomisation test of percentage of coefficient of variation of crops through time” for temporal patterns are particularly useful to characterise and simulate crops patterns within LandSFACTS model. Their main advantage, in comparison with the other tests, is their use of the randomisation test, which provides a reference (randomisation curve) to evaluate specific crop patterns. The integration of these statistical tests within the LandSFACTS model is further investigated in Chapter 7.1.3: General modelling approach, p.94.

In the next chapter, a new mathematical representation of crop rotations to integrate within the LandSFACTS model is presented.

6 Mathematical representation of crop rotations

Crop rotation is defined as the successive growing of crops on a field (Wibberley, 1996), and rules underlying them are complex, as reported in Section 3.1.2: Agronomic constraints (p.15). To integrate crop rotations into the LandSFACTS model, the decision process leading to crop sequences should not be explicitly modelled, (cf. Chapter 2: Specifications of LandSFACTS model). The complexity of crop rotations needs to be represented in a simple and systematic mathematical structure. Only few studies considered crop rotations in a mathematical manner (cf. Section 3.3.2.4: Crop rotations p.30). In order to achieve the mathematical integration of crop rotation, a systematic classification of the rotation is presented, before proposing a mathematical and statistical structure for representing crop rotations. This chapter is the subject of a publication currently submitted (Castellazzi *et al.*, 2008) .

6.1 Mathematical classification of rotations

The proposed classification of crop rotations is exemplified by a typical arable five-year rotation (Figure 6.1) for medium to heavy soils in the East Anglian region in eastern England (Clarke *et al.*, 2000), described by Jim Orson (Orson, 2005). The classification into four categories is based upon variability in the pathways (flexibility), and the length of the rotation.

The first type of rotation is the “fixed rotation” (Figure 6.1a). Each crop follows a pre-defined order with no possibility of deviation (see, for example, (Colbach *et al.*, 2005). The rotation can also be defined as “cyclical, and its rotation length is fixed, i.e. four years in this example. The second rotation (Figure 6.1b) is a “flexible” and may be represented as a multi-pathway network. For at least one crop within the rotation, the farmer makes a choice between several crops. As with the previous category, this type of rotation is “cyclical” with a fixed rotation length. The third category of rotations (Figure 6.1c) is again flexible, multi-pathway, and cyclical. However the rotation length is variable, as for example, the return period of the wheat 2 can be either four or five years. The fourth category (Figure 6.1d) encompasses less structured rotations, with great flexibility, cyclical structure with a highly variable rotation length. The pathways increase exponentially with years.

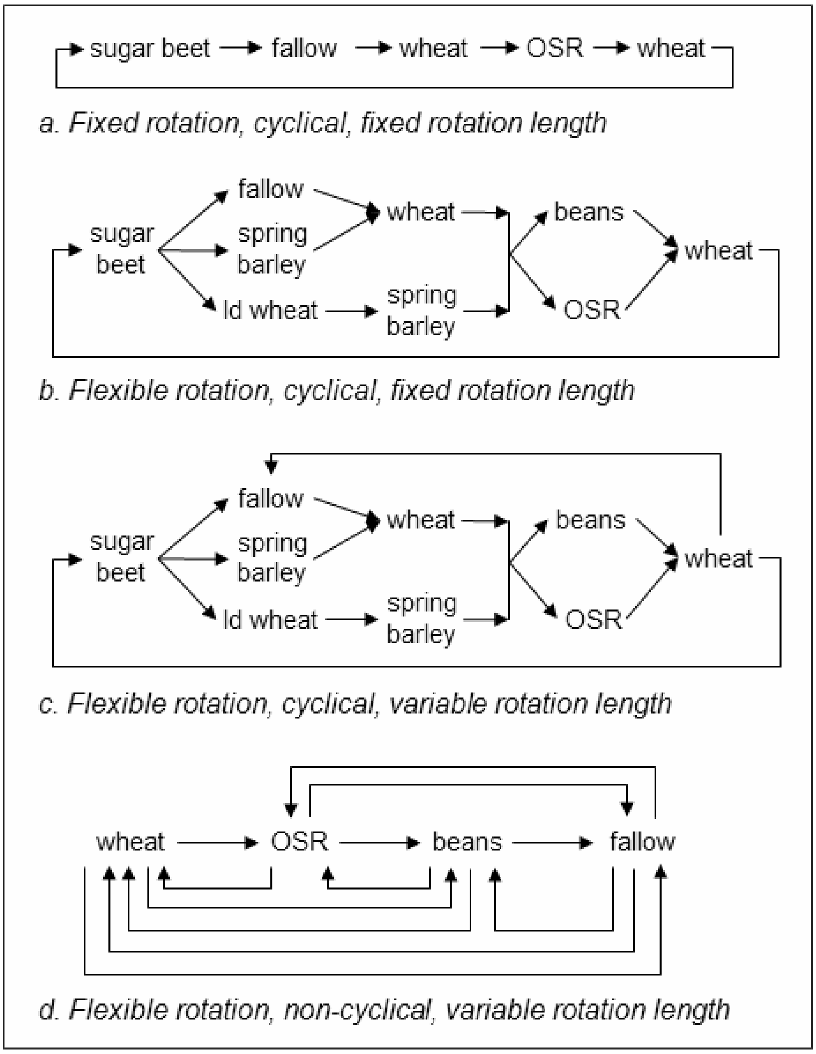


Figure 6.1: Mathematical classification of crop rotations.

In some regions rotations are highly fixed (like in Burgundy, whereas in areas more susceptible to market variation, only the main rotation principles are followed (like in eastern England, or the Fife agricultural area in Scotland). It is important that each type of rotations must be usable in the LandSFACTS model; thus, each type of crop rotation must be represented in the same format, despite their differences.

6.2 Rotations as transition matrices

Crop rotations are mainly sets of rules dictating crop sequences, where the primary driving rule is the influence of the crop in the current year on the crop choice for the next year. Therefore, it is possible to represent crop rotations as Markov chains, also called a stochastic matrix or transition matrix (Cox and Miller, 1965). A crop transition matrix T is a square matrix with as many rows (i) and columns (j) as distinct crops. A crop may be considered as two distinct crops, if there is a need to separate their occurrence within the rotation, e.g. first and second wheat. The element in row i and column j , called Tij, represents the probability under the rotation that, given the current crop i in a field, crop j has the probability Tij to be grown the next year in the field. The sum of the elements of each row must be equal to 1.

As an example, transition matrices of the rotations presented in Figure 6.1, are reported in Figure 6.2.

a) Transition matrix of fixed rotation, cyclical, fixed rotation length

		Current year				
		Sugar beet	Fallow	Wheat 1	OSR	Wheat 2
Previous year	Sugar beet	0	1	0	0	0
	Fallow	0	0	1	0	0
	Wheat 1	0	0	0	1	0
	Oilseed rape (OSR)	0	0	0	0	1
	Wheat 2	1	0	0	0	0

b) Transition matrix of flexible rotation, cyclical, fixed rotation length

		Current year								
		Sugar beet	Fallow	Spring barley 1	l.d. wheat	Wheat 1	Spring barley 2	Beans	OSR	Wheat 2
Previous year	Sugar beet	0	0.3	0.35	0.35	0	0	0	0	0
	Fallow	0	0	0	0	1	0	0	0	0
	Spring barley 1	0	0	0	0	1	0	0	0	0
	l.d. wheat	0	0	0	0	0	1	0	0	0
	Wheat 1	0	0	0	0	0	0	0.5	0.5	0
	Spring barley 2	0	0	0	0	0	0	0.5	0.5	0
	Beans	0	0	0	0	0	0	0	0	1
	OSR	0	0	0	0	0	0	0	0	1
	wheat 2	1	0	0	0	0	0	0	0	0

c) Transition matrix of flexible rotation, cyclical, variable rotation length

		Current year			
		Wheat	Fallow	OSR	Beans
Previous year	Wheat	0	0.3	0.2	0.5
	Fallow	0.2	0	0.3	0.5
	Oilseed rape (OSR)	0.5	0.25	0	0.25
	Beans	0.25	0.25	0.5	0

Figure 6.2: Transition matrices of three types of crop rotation.

The fixed rotation (a) has only binary probabilities, as only one crop is possible after another one. The distinction between the two wheat crops imposes the return period of the crop, and conditions the rotation length. The other two transition matrices (b and c) represent more flexible crop rotations, i.e. several crops possible for a given year. The probability of choosing a crop between several, is directed by the values in the matrix (value lower than 1).

In a crop rotation, the choice of a new crop does not always depend upon the previously grown crop, for example potatoes may only be grown every eighth year. Fixed crop rotations can incorporate those restrictions, however flexible rotations cannot. Therefore, for constraints over several years, the flexible transition matrices should be complemented with further temporal constraints, such as return period of crop, or maximum repetition of a crop on a field or by forbidden specific crop sequences.

By representing crop rotations as transition matrices, the complex decision making of integrating environmental variables (rainfall, temperature…), farm management, and market prices within a mechanistic model is replaced by a simple empirical approach based on statistical probabilities. This simplification provides a simple basis for modelling crop rotations stochastically, without requiring a multitude of parameters. However, even if only probability values are required within the transition matrices, those values need to be chosen carefully. To represent crop rotations realistically, the probability values should be derived from results of interdisciplinary research in agronomy, farm management, environment, socio-economics or agronomic statistics. In future developments of the transition matrices, this approach could be elaborated by integrating specific variables within the matrices to influence crop choices (e.g. climatic data, relative crops market prices).

The transition matrices can be used to calculate the long term proportions of crops from each crop rotation. This tool is very useful to control simulated crop proportions of a grower or over the whole landscape in a model.

6.3 Long-term crop proportions

For a fixed rotation, the crop proportion over the long-term is equal to one over the number of crops, e.g. for the rotation a in Figure 6.1, the long-term proportion of every crop is 1/5 = 0.20. However, as wheat occurs twice, its long-term proportion equals 0.4 (2 * 1/5). For flexible crop rotations, the calculation is more complex and uses the properties of transition matrices to estimate it, cf. Figure 6.3. By multiplying the matrix by itself many times (which can be called the "burn-in period", default value: 200 times), the resulting matrices converge towards a steady state. A steady state may only be reached if the probabilities in the transition matrix are between 0 and 1 exclusively. In the case of a transition matrix containing 0 and 1, such as in fixed rotations, the steady state is evaluated by carrying out further multiplications (called a "saving period", default value: 100 times) and averaging the resulting matrices. After each multiplication, a rounding check is carried out, to force the rows of the matrix to sum to 1 exactly.

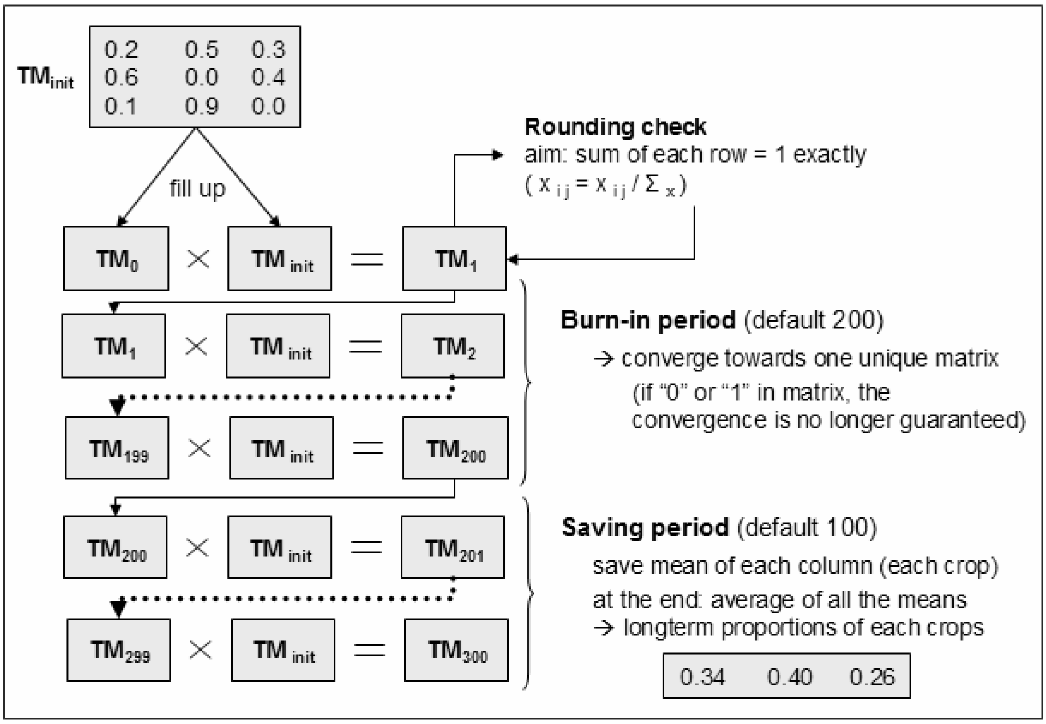


Figure 6.3: Calculation of long-term crop proportions from transition matrices.

As an example, consider the crop rotation b in Figure 6.1, represented by the transition matrix b in Figure 6.2. The matrix multiplication converged to the following long-term crop proportions: $p_{\text{sugar beet}} = 0.2$, $p_{\text{fallow}} = 0.06$, $p_{\text{spring barley 1}} = 0.07$, $p_{\text{late drilled wheat}} = 0.07$, $p_{\text{wheat 1}} = 0.13$, $p_{\text{spring barley 2}} = 0.07$, $p_{\text{beans}} = 0.1$, $p_{\text{oilseed rape}} = 0.1$, $p_{\text{wheat 2}} = 0.2$. The result for sugar beet can be simply exemplified. As sugar beet is always the crop grown in the first year of the rotation, it occurs every five year, e.g. a long-term crop proportion of 0.2.

Knowledge of the long-term crop proportions of rotation is useful when aiming at controlling the crop proportions over several fields or even at the landscape level. For example to achieve 30% of wheat over the whole landscape over 10 years, the proportion of wheat over all the rotations in fields must reach this value. Therefore by varying the crop proportions of each rotation, their proportions over the whole landscape can be controlled and modelled.

6.4 Transition between rotations

The use of transition matrices can be further extended to represent the transitions between the crop rotations themselves. The crop rotation on a field may change over time due to fluctuations in market prices of crops or in environmental conditions such as climate change. For example a farmer might want to alternate between a three years fixed crop rotation, R1 (wheat, oilseed rape, beans), and a two years fixed crop rotation, R2, (wheat, oilseed rape), Figure 6.4.

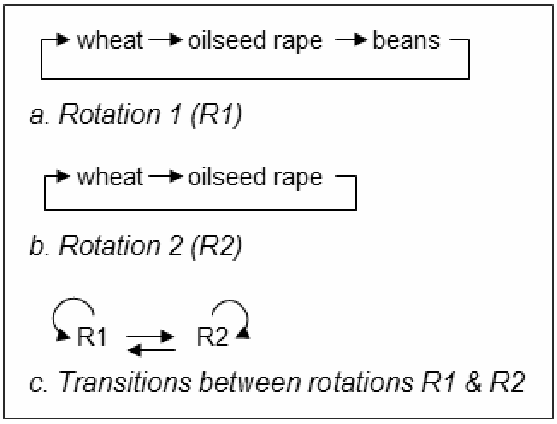


Figure 6.4: Diagram of transitions between two individual crop rotations.

The choice between the two rotations might be driven by the relative market prices of the three crops, the incidence of pests or disease, or even climatic conditions. The transitions between the two rotations can be represented as a stochastic process, and thus by transition matrices. The square matrix holds the probabilities of transitions between one rotation to itself, or to the other rotation (Table 6.1). The probability of rotation 1 after itself is noted as the probability r, and rotation 2 after rotation, 1 - r, as the sum of each row must equal 1.

Table 6.1: Transition matrix between two crop rotations

Previous year		Current year	
		Rotation R1	Rotation R2
	Rotation R1	r	1 - r
	Rotation R2	1 - s	s

The system now consists of three transition matrices, i.e. rotation 1, rotation 2, transitions between rotation 1 and 2, which is mathematically cumbersome. This system can be simplified by combining the three transition matrices into one, denoted U_{TM} (Table 6.2).

Table 6.2: Overall transition matrix between two crop rotations (U_{TM})

			Current year				
			(R1)			(R2)	
			Wheat	OSR	Beans	Wheat	OSR
Previous Year	(R1)	Wheat	0	r	0	(1 - r) / 2	(1 - r) / 2
		OSR	0	0	r	(1 - r) / 2	(1 - r) / 2
		Beans	r	0	0	(1 - r) / 2	(1 - r) / 2
	(R2)	Wheat	(1 - s) / 3	(1 - s) / 3	(1 - s) / 3	0	s
		OSR	(1 - s) / 3	(1 - s) / 3	(1 - s) / 3	s	0

The overall transition matrix U_{TM} is 5 × 5, composed of four blocks that represent the transitions between individual crops of rotation R1 (3 × 3 top left block), R2 (2 × 2 bottom right block), R1 to R2 (3 × 2 top right block) and R2 to R1 (2 × 3 bottom left block). The entries of the transition from R1 to R1 and R2 and R2, are a copy of the individual transition matrices R1 and R2, multiplied by the probability of remaining within their respective rotations. The entries of the transition from R1 to R2 and R2 to R1 represent the probability of a change from a crop in one rotation to a crop in a different rotation. For simplicity it is assumed that when such an event occurs the crop in the new rotation is chosen at random, although this is not a strictly necessary condition. In this case of random choice, each of those entries is constructed by dividing the probability of changing from one rotation to the other by the number of crops in the new rotation. This process may easily be generalised to transitions between more than two rotations; the single overall matrix U_{TM} that results will always be square with the number of rows and columns equal to the sum of the number of crops over all the individual rotations.

The transition between rotations could be linked to some external driving trend, such as climatic change, for example rotation R2 may become more likely than rotation R1 over the next 50 years. In those conditions, fields with rotation R1 would switch to R2 over time. Here, for simplicity, we disallow a reversion to rotation R1 once a change has been made from R1 to R2, so $s = 1$. But now, r is a function of time. For example, consider the situation where a transition from R1 to R2 would have been unthinkable at the beginning of the present century, but the probability of which increases steadily year by year until, by 2050, it becomes inevitable. This may be modelled by the equation: $r = (2050 - Y)/50$, where Y represents the current year. It is easy to substitute this variable value for r into software that implements an algorithm to model such change.

6.5 Conclusion

In this chapter, the crop rotations, as presented in Section 3.1.2 Agronomic constraints (p.15), were classified into four types ranging from strict rotations (i.e. fixed sequence of crops) to flexible rotations (i.e. non-cyclical and variable rotation length). To assure the usefulness of the LandSFACTS model, all the above types of crop rotation must be equally mathematically handled in the model. Instead of modelling the decision-making process of farmers, the choice of crops to grow was stochastically modelled. By assuming that choosing a crop only depends upon the previously grown crop, Markov chains (transition matrices) can be used. The transition matrices define the probability of growing a crop after any other crop within the rotation. For any restrictions on crop successions spanning over more than two years, the transition matrices should be complemented by constraints, e.g. return period of crop. From transition matrices, long-term crop proportions can be calculated, thus providing a tool to control them within the LandSFACTS model. The use of transition matrices can be further extended to model transition between crop rotations.

In the next chapter (Chapter 7), the integration of the rotations within the LandSFACTS model is detailed, alongside with the description of whole model.

7 Description of the LandSFACTS model

The specifications for the LandSFACTS model were detailed in Chapter 2. In brief, the LandSFACTS model is a research tool to facilitate the setting up of scenarios of crop spatio-temporal arrangement at the landscape scale, within a GM-conventional coexistence context. After analysing published literature relevant to the aims of the LandSFACTS model (Chapter 3), two lacking topics were defined (Chapter 4) and developed: (i) statistical metrics to measure crops spatio-temporal patterns (Chapter 5), and (ii) mathematical integration of crop rotations (Chapter 6).

In this chapter, by using conclusions drawn from all the previous chapters, the model is assembled. At first, the model will be defined with its distinct characteristics and the general modelling approach (Section 7.1). Then the model inputs, i.e. agronomic inputs and model parameters, are detailed (Section 7.2), followed by the description of the main process of the allocation of crops to fields over the years (Section 7.3), i.e. the "CropAllocation" program. Afterwards the model outputs, i.e. crop allocation and difficulty indexes of finding authorised allocation, are detailed and their interpretations explained (Section 7.4). The chapter is concluded with details on the implementation of the model, including the program language, data format, model executable and the availability of the LandSFACTS software (Section 7.5). This chapter has been partly published in conference proceedings (Castellazzi *et al.*, 2007a), and will be the subject of a peer-reviewed article.

7.1 LandSFACTS model definition

7.1.1 Aim of the model

The specifications of the LandSFACTS model have been determined from multiple discussions with end-users in Chapter 2. The aim was for the LandSFACTS model to be a scientific research tool, which allocates crops into fields, to meet user-specified crop spatio-temporal arrangements, using an empirical and statistical approach. The model must meet the needs for creating GM coexistence scenarios, such as spatial and temporal separation distances between crops, with the aim of being used by researchers with agronomic knowledge of the landscape studied.

7.1.2 Distinct characteristics of the model

The LandSFACTS model has three distinct characteristics in comparison with published models; they are reported below.

- The model's crop allocation to fields is aimed at reaching a user-specified spatio-temporal arrangement of crops, using empirical and statistical approach. Therefore, a substantial part of the decision-making process leading to the crop arrangements is not taken in account.
- The model must be useable on any European landscape, thus no agronomic information is intrinsic. The model should only provide the structure to input site-specific agronomic rules.
- The model aims at allocating a crop to every field for every year of simulation. Thus the spatial and temporal unit of the model is the field and a year respectively. The fields are represented as polygons with boundaries unchangeable through time, field merging or divisions are not considered.

7.1.3 General modelling approach

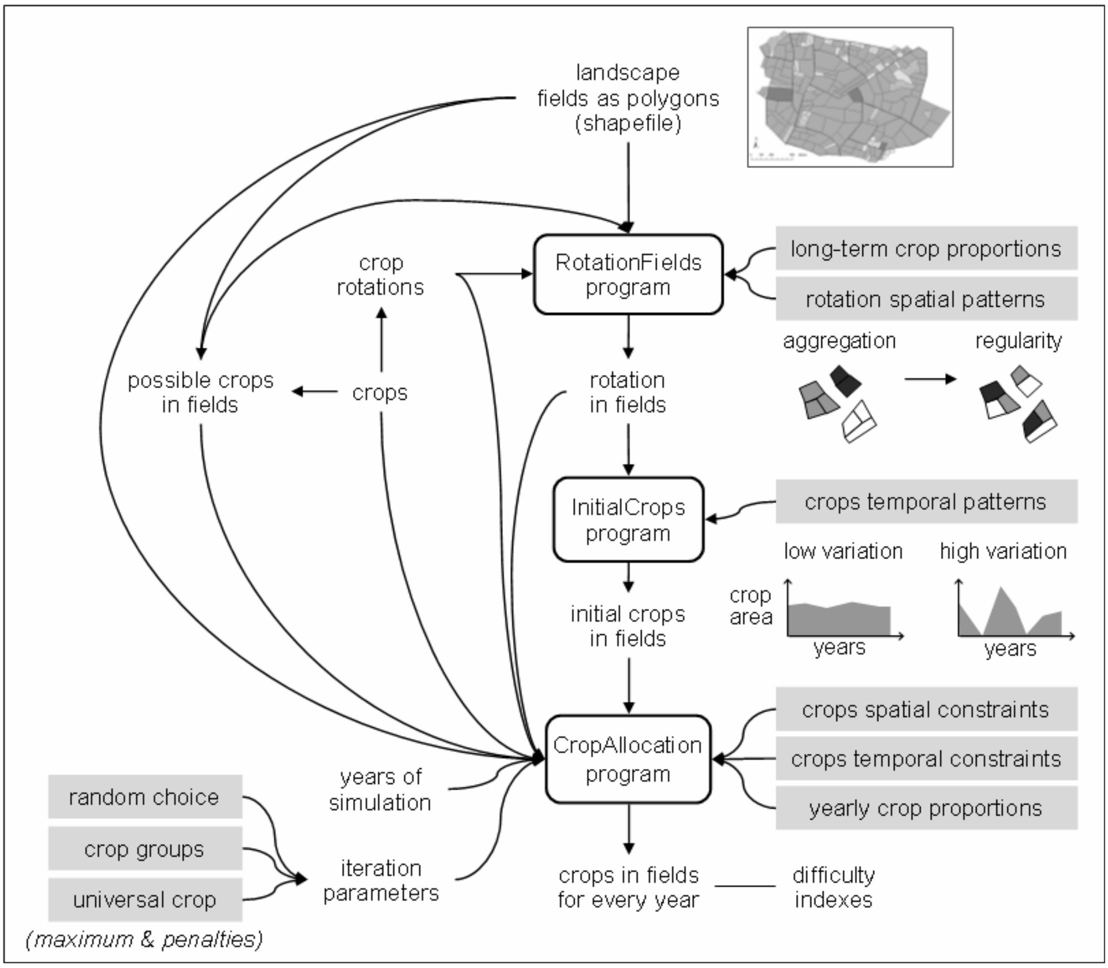
As reported in Chapter 2, the model will simulate directly the crop arrangements by using an empirical and statistical approach. Therefore the core modelling variables are kept to the strict minimum of the crops, the fields, and the crop rotations. Further variables, aimed at controlling the crop arrangements, are the three types of constraints: (i) spatial constraints, imposing separation distances between crops, (ii) temporal constraints, imposing return period and maximum repetition of crops on fields and forbidden crop sequences, and (iii) yearly crop proportion constraints limiting the area proportions over the whole landscape. These variables were derived from the review on the constraints influencing the crop arrangements (cf. Section 3.1: Origin of crop arrangements, p.15).

In addition to the strict constraints, general trends in spatio-temporal patterns of the crops are controlled by using the statistical analyses developed in Section 5.1 (New statistical analyses on crops' spatial and temporal patterns, p.45). The patterns result from the coordination between crop rotations of fields. The spatial patterns, e.g. dispersion of wheat over the whole landscape or its higher concentration to specific areas, are mainly directed by the spatial repartition of the rotations in fields. The temporal pattern mainly results from the coordination of the initial crops in the fields, e.g. if all the rotations in fields start with the same crop or if they are shifted in time.

The constraints and patterns of crops do not have to be modelled at the same time. The pattern influences the general trend on the whole landscape and does concern all the crops in all the fields at the same time. On the other hand, the constraints have a more localised influence, for example if two fields, one with GM and one with non-GM crop are too close, or, if the current crop sequence on a specific field is forbidden. The pattern should thus be imposed on the landscape only once to influence the rotation spatial repartition and the initial crops, whereas the constraints could be checked for every year of crop allocation. The crop rotations also have a major influence on the crop temporal arrangement, as it initially dictates the crop successions.

7.1.4 Structure of the LandSFACTS model

To use the LandSFACTS model two steps are required: (i) an initialisation phase - the preparation of the input data for crop allocation to fields (i.e. simulation phase); and (ii) a simulation phase – crops are allocated to fields through years, while respecting the user-defined spatial and temporal arrangement of the crops (crop constraints, iteration parameters, etc.). To support the initialisation phase particularly in case of missing data or new scenario testing, two programs are available to help with the rotation allocation to fields (RotationFields) and with the initial crops in each field (InitialCrops). The simulation phase (i.e. crop allocation to fields) is solely comprised of the “CropAllocation program”. The individual inputs, outputs and links of the three programs are presented in Figure 7.1**Error! Reference source not found.**.



<u>Figure 7.1: LandSFACTS model components and programs</u>

The elements with grey background are optional components. All the components and processes presented in the diagram are part of the "initialisation phase", except for the CropAllocation program which represents the "simulation phase", and the LandSFACTS outputs (crops in fields for every year and difficulty indexes).

In the following section, the initialisation phase with the inputs to the "CropAllocation program" is described (Section 7.2), then the process controlling the crop allocation to fields is detailed (Section 7.3). The outputs of the simulation (Section 7.4) and then details on the technical implementation of the model (Section 7.5) are reported.

7.2 LandSFACTS initialisation phase: inputs to CropAllocation program

The LandSFACTS model integrates multiple elements to provide a crop allocation meeting all the user requirements. As shown in **Error! Reference source not found.**, the model inputs are interrelated. The definition of the crops and the list of fields are the two core inputs, on which other inputs are built. The spatial extent of crops is controlled with the list of "possible crops in fields". The crop rotations have the crops as their main components. Then a crop rotation must be allocated for each field, along with the crop to grow in the first year. Specific constraints on crops' temporal and spatial arrangement and yearly crop proportions are available. Further inputs control the behaviour of the model during the iteration process. In this chapter, the inputs of the simulation (crop allocation) are detailed successively. They are all compulsory for the model, except if specifically stated otherwise.

7.2.1 Crops

Crops are the smallest unit, which are yearly allocated to every field. All the crops to be allocated within the simulation must be set up at the start.

7.2.2 Fields

Fields are the spatial unit on which the crops are being allocated, and they must have fixed boundaries through all the years. They must also be simple polygons (defined as a closed line with no line crossing), with known centroid coordinates if the spatial patterns are to be controlled, and with known area if the crop proportions are to be controlled (long-term proportions or yearly crop proportions).

7.2.3 Possible crops in fields

The spatial extent of crops might need to be limited to specific fields. For example, maize is preferably grown on low slope levels to limit soil erosion and close to water sources to allow the irrigation of the crop. For each field, the available crops must be specified, and only those crops will subsequently be allowed to be grown in those fields. Therefore only crop rotations which have all their crops authorised on the field can be chosen.

7.2.4 Crop rotations

The crop rotations are integrated within the model as transition matrices (cf. Chapter 6: Mathematical representation of crop rotation p.85). The transition matrices regulate the probability of growing a crop depending only upon the previous crop in the field. The long-term crop proportions must be calculated for each transition matrix.

7.2.5 Crop rotation for each field and the optional "RotationFields" program

The model requires one crop rotation per field. The allocation of a crop rotation per field, limits the possible spatial repartition of the crops over the landscape, particularly if the available crop rotations incorporate different crops. The crop rotations for each field might be provided by a survey carried out on the studied landscape. However, if the exact location of the rotation in the landscape is unknown or if new allocation of the rotations in the landscape must be tested, the user may use the "RotationFields" program.

The "RotationFields" program allocates the rotations to the fields. The user may specify either or both of the following parameters: (i) desired long-term proportion of any or all crops by area and the standard deviation permitted from the target, and (ii) the desired spatial patterns of the crop rotations. The spatial pattern is controlled by the statistical analysis based on the "E analysis" defined in Section 5.1.3.2 (General spatial trend (E analysis), p.57). Two other parameters are indispensable: the maximum number of iterations to obtain the allocation and the choice of using or not using weighted rotations to optimise the rotation allocation. The weighted rotation option is a preliminary step within the program, which alters the probability that a field is allocated any particular rotation, without considering the area of the field. (Areas are always considered within subsequent steps of the program). If the weighted rotation option is not used then each field has an equal probability of being allocated any of the rotations. If the weighted rotation option is selected, then each rotation is given a random weight, which affects accordingly its probability for being chosen for any field. The weighted rotations option is only useful if fields have relatively similar areas. It should also be noted that the use of the weighted rotation option may sometimes provide extreme allocations (e.g. if rotation x is given a weight of 95%, this rotation will be over-represented within the whole landscape).

Program inputs

- Possible crops in fields
- Rotation definitions as transition matrices
- Target long-term proportions of any crops + standard deviation permitted – not compulsory
- Target interval for the spatial pattern of crop rotations (Ep values, cf. Section 5.1.3.2: General spatial trend (E analysis), p.57) – not compulsory
- Number of randomizations for creating the randomization curve for the statistical test

- Maximum number of iterations
- The use, or not, of “weighted rotation” to optimize the allocation process

Program approach

The program goes through the following steps in order:

1. Crop rotations with a forbidden crop (i.e. long-crop proportion target equal to 0) are not considered within the program.
2. Each rotation is given an equal probability to be chosen when allocating a rotation to every field.
3. If “weighted rotations” was chosen, the above equal probabilities will be altered into uneven probabilities of choosing the rotations.
   - Each rotation is given a random weight (the weight of all rotations adds up to 1).
   - Calculate the long-term proportions of each crop with the current random weighing
   - Check how many crop proportion targets are met.
   - If more targets are met than the current best --> this current random weighting replaces the current best weighing.
   - If less targets are met than the current best --> this current random weighing is deleted, and a new random weighting is created. This loop keeps on iterating until the maximum number of iterations is reached.
   - If all targets of long-term crop proportions are met or if the maximum number of iteration is reached, the program proceeds to step 4.
4. For each field, the possible rotations are determined (using possible crops in fields)
5. For each field, a rotation is randomly allocated by using the rotation weights.
   - If all targets of crop long-term proportions and spatial patterns are met --> the rotation allocation is accepted
   - if not, the program goes to step 5 until the maximum number of iterations is reached.

Program outputs

- A rotation for each field
- Crop long-term proportions over the landscape of the rotation allocation
- Spatial pattern value (E and Ep) of the rotation allocation
- Report on the iteration process

7.2.6 Initial crops for each field and the optional "InitialCrops" program

The initial crops determine the crops, from the field crop rotation, to be grown in the first year of simulation. The coordination of the initial crops between fields influences the temporal patterns of the crops. An initial crop must be specified for every field. If they are not, then a random allocation option is available. If the random allocation option is activated, the "CropAllocation program" will randomly choose an initial crop when starting.

If the user wants to coordinate the initial crops between fields towards a specific crop temporal pattern, the InitialCrops program is available. The program will randomly allocate an initial crop to each field, and check if the current crops temporal patterns meet the requirements. The statistical analysis of the crops temporal pattern is based on the "Randomisation test of percentage of coefficient of variation of crops through time" defined in Chapter 5.1.2.3 (Crop temporal variability compared to random simulations, p.51).

Program inputs

- Number of randomisations for creating the randomization curve for the statistical test (default: 1,000 randomisations)
- Number of years on which the coefficient of variation is calculated (default: 100 years)
- Two choices of randomisation processes: after a failed allocation of initial crops to fields, the failed allocation has one initial crop altered (improve_ iteration) or all are re-randomised (random_iteration)
- Maximum number of iterations until allocation is accepted.
- Crop rotations as transition matrices
- Long-term crop proportions for each rotation
- A crop rotation per field
- Field areas
- Target interval for the temporal pattern of initial crops

Program approach

The program goes through the following steps chronologically:

1. The randomization curve for the temporal pattern analysis is created by using randomly allocated initial crops to each field.
2. The temporal pattern of the current initial crops is calculated

3. If the current CVp value is within the targets → the current initial crop is accepted and the program stops.
4. If not, the iteration process starts:
5. If improve_ iteration was selected → a randomly chosen field has a new random initial crop.
6. If random_ iteration was selected → all the fields have new random initial crops.
7. The temporal pattern of the current initial crops is calculated
8. If the current CVp value is within the targets → the current initial crop is accepted and the program stops.
9. If not: the program goes to step 5.

Program outputs

- An initial crop for each field
- Temporal pattern value (%CV and CVp) of the rotation allocation
- Report on the iteration process

7.2.7 Crop constraints

Three broad types of constraints can control further the crop arrangements: spatial constraints, temporal constraints and the yearly crop proportion constraints. Each type of constraint is checked with the proposed crop allocation. Crop allocations can only be accepted if they meet all of the constraints. The setting up and use of crop constraints are not compulsory.

7.2.7.1 Spatial constraints

The spatial constraints aim at enforcing separation distances between two individual types of crops grown in fields. Two fields are considered as neighbours, if the shortest distance between their outside boundaries is within the specified distance. The main GIS function used is the “positive buffer” function, e.g. fields boundaries expanded outwards by the separation distance. If two crops with spatial constraints are too close to each other, one of them will have to be changed. The program allows setting priorities to crops alteration. For example for a specific coexistence scenario, the presence of conventional crops might prevent GM crops being grown in the neighbourhood; to integrate this constraint within LandSFACTS, the GM crop is given the highest priority in being altered in case of conflict with conventional crops. The number of individual spatial constraints is not limited. This constraint type is checked for every single yearly crop allocation, over the whole landscape.

7.2.7.2 Temporal constraints

The temporal constraints aim at enforcing rules on the crop succession. Three types of temporal constraints are available: (i) return period of crops, or group of crops, on a field, i.e. temporal separation between crops; (ii) maximum successive growing of a crop, or group of crops on a field; (iii) forbidden crop sequence. Those constraints are inspired by classical agronomic and rotational recommendations. The temporal constraints are linked to individual fields in order to reflect the pluralism of individual farmers' decisions. The yearly crop allocations are checked for their agreement with the temporal constraints in relation to the precedent crop allocations.

7.2.7.3 Yearly crop proportion constraints

A target area proportion can be set up for every crop and year, with an authorised standard deviation. Specifying a target for all crops and years is not compulsory. Every yearly crop allocation must meet the specified targets for the year and the crops.

7.2.8 Iteration options and penalties

The iteration options control the behaviour of the model, when the program attempts to overcome an unauthorised crop allocation, by changing some of the crop allocations.
Four iteration options are available:

- Option 1. In this option, all fields, whether problematic or not have their crop randomly altered. If this option is used during the first year of simulation, a new random initial crop is chosen for every field. For any other year of simulation, a new choice of crop is made within the transition matrices for every field. This option is not an optimisation process.
- Option 2.1. In this option, one randomly chosen problematic field has its crop randomly altered. If this option is used during the first year of simulation, the initial crop of one problematic field is randomly chosen. For any other year, a new choice of crop is made within the transition matrix of a problematic field. The choice of the problematic field to alter is detailed in Section 7.3.2 (Problematic-points temporary store, p.107).
- Option 2.2. In this option, the one randomly chosen problematic field (see Option 2.1) has its crop exchanged with a crop from the same crop group set up by the user (e.g. crops with the same function within a rotation).

- Option 2.3. In this option, the one randomly chosen problematic field (see Option 2.1) has its crop exchanged with the universal crop. The universal crop cannot be linked with any temporal or spatial constraints.

For each iteration option, the user specifies the maximum number of iterations allowed. If all options are enabled, they will be carried out successively until the crop allocation for the current year is accepted.

The order of the options is intentional, as each option provides a more specific crop alteration than the previous options, which increases the probability of finding an authorised crop allocation. Option 1 is the only option without optimisation, i.e. a complete new crop allocation for every field is generated every time; whereas the other three options are optimising, as they improve on a “current” crop allocation by altering only the crop of one problematic field. Option 1 is useful to provide a completely new random allocation without any optimisation. Although very often option 1 will not find an authorised crop allocation, however it has a specific use. For example, allowing 10 iterations with option 1 as a precursor to any optimisation iteration options (2.1, 2.2, and 2.3), means that the “best” crop allocation out of 10 random ones (from Option 1) would be used for optimisation. Therefore the optimisation process has more chance to be started from a “normal” crop allocation instead of an extremely bad one. This technique would also decrease output variability between simulations with the same inputs (e.g. standard deviation of overall penalties or number of iterations used).

Both Option 1 and Option 2.1 provide new random decision(s) within the crop rotations. As they fully respect the crop rotations, both options can be set up to high number of maximum iteration without altering the quality of the crop allocation. Option 1 and Option 2.1 differentiate by their process. Option 2.1 is based upon an optimisation process, as it tries to improve upon a current crop allocation by altering only one field with an unauthorised crop, whereas Option 1 is non specific, as it alters all the crops, regardless of their current agreement to constraints. Due to its optimisation technique, Option 2.1 is more efficient and reliable than Option 1 to find an authorised crop allocation.

Option 2.2 uses the optimisation technique by altering the crop of only one problematic field, as Option 2.1. However the crop rotations are not respected, as the problematic crop is exchanged for a crop of the same crop group set up by the user. For example a group could be crops with the same function within a rotation, or a-like crops.

Option 2.3 also uses the optimisation technique, by replacing a problematic crop by a "universal crop" chosen by the user. The universal crop option is the last chance to find an authorised crop allocation. A universal crop may be fallow for example, i.e. if no crops can be grown on a specific field, it is left as fallow for a year. In order to be authorised on any field, the model obliges the user at choosing only a crop, which is not linked to any spatial or temporal constraints. Thus the universal crop will always agree with all the constraints. There is one exception, due to the possibility of imposing a yearly crop proportion for the universal crop. Therefore using the universal crop might not improve a crop allocation if any of the crops does not meet its targeted yearly crop proportion. For example, if GM oilseed rape should be at least in 5% of the arable area, and a GM oilseed rape field is spatially too close to conventional one (i.e. spatial constraint), by exchanging the GM crop for the universal crop (e.g. fallow), the area of GM oilseed rape may fall below 5% and thus not be authorised. For some scenario, it might be useful to use as the universal crop an "unreal crop", e.g. "flag crop", in order to keep track of which fields couldn't comply with the crop allocation constraints.

Depending upon the aims of the simulation and the constraints imposed on the crop allocation, the maximum iteration options must be adjusted. The impact of iteration options on the authorised crop allocation and on the difficulty indices of the simulation, are further explored in Section 8.6 (Sensitivity analyses, p.120). Section 8.8 (Recommendations on model use, p.142) presents recommendations for their setting up.

A penalty value may also be set up for each iteration option. The penalty value will be applied to every field on which an iteration option was successfully used (improved crop allocation). Thus, the simulation keeps a track of how often the crops in specific fields are changed to reach the desired crop patterns. The penalties to the fields will allow comparison between simulations, for example, to evaluate the difficulty of obtaining an authorised crop allocation if the mandatory separation distance is increased between GM oilseed rape and conventional varieties.

Furthermore, the model records the number of times each field has an unauthorised crop allocation, and which constraint it failed. The most problematic fields and constraints may therefore be spotted, providing a tool to alter the scenarios either to facilitate the allocation or to increase the difficulty.

7.2.9 Simulated annealing

The simulated annealing process is a generic algorithm for optimisation, which aims at increasing the probability of reaching a desired target, by preventing the program to be blocked at a local minimum. For example, the logic behind this technique can be compared to the situation of walking in a labyrinth, i.e. when blocked at a dead-end (local minimum), it is necessary to walk back to a previous intersection and take a new pathway. In the case of the LandSFACTS model, the program may be blocked at an unauthorised crop allocation, which cannot be improved further by altering only one crop without going back to a “worse allocation” in order to find a new pathway towards the desired allocation. The exact process is explained below.

The program tries to overcome an unauthorised crop allocation by altering the “currently best” crop allocation using the iteration options. When only one crop is altered at a time (any option except 1), the program may not be able to improve the “currently best” crop allocation by altering only a crop. A worse crop allocation would need to be accepted as the “currently best” (step backwards) to unblock the program and thus increase the chance of finding an authorised crop allocation.

The value of simulated annealing influences the speed and chances of reaching an acceptable crop allocation. A low value, such as 1 or 2, would accept “worse” situations very often, and the optimisation process would be slow (many iterations required) or even nonexistent. A high value, such as 1,000, would provide a way out of local minima only after having checked 1,000 crop alterations. Such a high value requires an even higher maximum number of iterations to be allowed. Finding the right balance for the simulated annealing value is important, to avoid local minima, while not slowing down the iteration process.

The use of the simulated annealing process complies with the requirement that an authorised crop allocation (i.e. successful allocation) must meet all the constraints specified by the user.

7.2.10 Simulation parameters

The simulation parameters are comprised of the number of year of crop allocations required, and the constraints that need to be checked for this simulation (they can be disabled, if necessary).

7.3 LandSFACTS simulation phase: process of CropAllocation program (allocation of crops to fields)

The simulation phase is comprised of only one program called “CropAllocation”. The program allocates the crops to the fields, by using the crop rotations of the fields and their initial crops, and by respecting temporal and spatial constraints, and the yearly crop proportion.

7.3.1 Overview of CropAllocation process

The first step of the program is to check the coherence of the inputs. For example the rotations assigned to the fields (section 7.2.5) should all be defined as rotations (7.2.4) and with calculated long-term crop proportions.

The model starts by assigning the initial crops to the fields, Figure 7.2. If random initial crop was specified, it is carried out. This proposed crop allocation is checked for its agreement with the spatial constraints (separation distances between crops), and the yearly crop proportion for the initial year (year 0). If the proposed crop allocation respects them, it is authorised and saved. The program now consider the next year. For each field, a new crop is randomly chosen using the transition matrix of their assigned rotation and considering the crop allocated for the previous year. The proposed crop allocation is checked for its agreement with the temporal constraints, yearly crop proportions, and spatial constraints. If the proposed crop allocation respects them all, it is authorised, saved, and the program moves to the next year. The program will keep on allocating the crops to the fields, for all the required years.

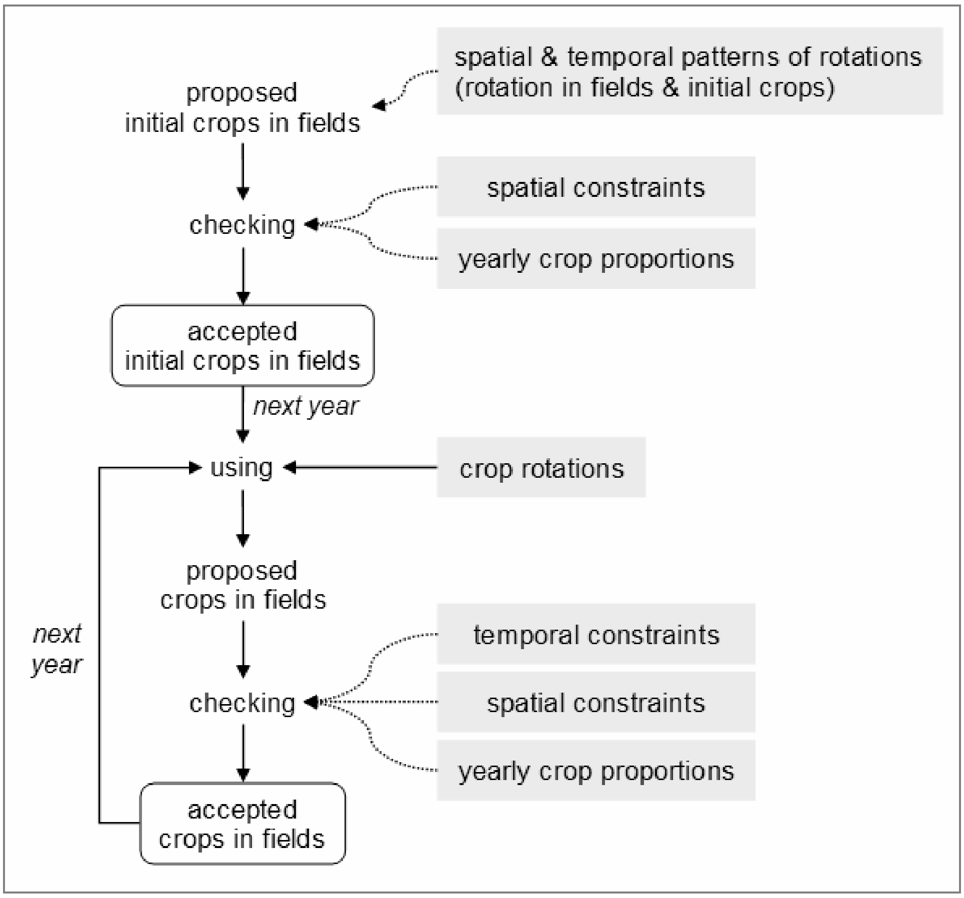


Figure 7.2: Overview of CropAllocation program

7.3.2 Problematic-points temporary store

For every iteration, the crop allocations are checked for their agreement with the constraints, the problematic-points temporary store records the fields with unauthorised crop allocations, i.e. not complying with constraints. The store is reset before each new iteration. The problematic-points store aims (i) at assessing the number of failed constraints of the current crop allocation over the whole landscape (sum of the points of all the fields); (ii) at identifying the fields that should have their crop changed.

The problematic-points are calculated over all the constraints. For the temporal constraints, if a crop cannot be authorised on a field due to previous crops, a point is added to the field problematic-point temporary store. For the yearly crop proportions, if any crop proportion over the whole landscape is not respected, all the fields have one point added to their problematic-point temporary store. In addition, the difference between the current crop proportion and the targets are recorded. For the spatial constraints, if two crops are closer to each other than authorised, both fields have one point added to their problematic-point temporary store.

The sum of the problematic-points for all the fields provides an overall estimation of the closeness to the desired allocation. The closer the sum is to null, the "better" is the allocation. This sum is indispensable to compare two crop allocations and deciding which one is closer to the desired allocation.

The problematic-points are also useful to determine the fields, which should have their crop altered. The fields with the highest values in their problematic-point temporary store are given a higher probability of having their crop altered. For this purpose, some fields have their problematic-points altered to prioritise the resolution of temporal problems over spatial problem. The points of a field are set to null in the following circumstances: if the field does not meet a spatial constraint and (i) the field's crop does not have the highest level of priority of being altered (cf. chapter 7.2.7.1: Spatial constraints, p.101); (ii) the field's crop has the highest level of priority of being altered but this field meets all the non-spatial constraints when other fields do not.

7.3.3 Overcoming unauthorised crop allocation

During the simulation, if a crop allocation does not meet all the constraints, the allocation is "unauthorised". The program then must alter this crop allocation, until it is "authorised". It is the iteration process. The current crop allocation to be improved is labelled as "currently best".

A new crop allocation is proposed by altering the "currently best" one. The alteration is done by using the iteration options controlling the behaviour of the model (7.2.8: Iteration options, p.102). If only one crop is to be altered at a time (any iteration option except 1), the program chooses one field randomly accordingly to the problematic-points (7.3.2).

This new crop allocation is checked for its compliance with the constraints. If all the constraints are met, the crop allocation is authorised, and thus this iteration process is stopped. If all the constraints are not met, the program must decide if the new allocation is closer to the desired landscape than the "currently best" one. If the sum of its problematic-points temporary store is lower or equal to the one of the "currently best", the new allocation overwrites the "currently best". If not, it is deleted. Then the program goes to the next iteration.

Every time, the alteration of a crop in a field improves the crop allocation, a penalty is applied to the field. The penalties to fields are independent from the problematic-point temporary store. The penalties to fields provide an evaluation of the difficulty of obtaining an authorised crop allocation (i.e. how many fields had to have their crops altered, and how many times).

If no crop allocation is accepted after the maximum number of iterations is used, the simulation is prematurely stopped. The model reports to the user, the number of years of successful allocation and indicates which year failed (no crop allocation are given for the failed year), and which constraints caused the failure.

7.4 LandSFACTS outputs and interpretations

The model provides two main outputs: the crop allocation to fields and the difficulty indexes. Both are detailed below.

7.4.1 Crop allocation to fields

The major output of the LandSFACTS simulation is a crop for every field and every successful year. Only crop allocations agreeing with all the specified constraints, are considered as successful and thus reported to the user. A log file is also provided to document the inputs, iteration processes and outputs of the simulation.

7.4.2 Difficulty indexes of the obtained crop allocations

The difficulty of obtaining a crop allocation is evaluated by using three main indices: (i) overall penalties to fields, (ii) number of iterations used, (iii) number of conditions which had to be overcome during the iteration process. The calculation of the indices is explained in section 7.2.8 (Iteration options, p.102). The overall penalties to fields index aim at providing an overall evaluation of the difficulty of obtaining the crop allocations. The number of iterations used gives an indication on how difficult it was to find a correct crop allocation. The last index provides an insight on how many constraints had to be overcome during the iteration process.

The difficulty indexes provide an evaluation of the difficulty that the software has to generate a crop allocation meeting all the user-specified constraints. These indexes are particularly useful to compare scenarios, for example to determine if an increased separation distance between two crops affects the difficulty of finding a valid crop allocation. It should be noted that the value for each difficulty index is specific to one particular simulation (i.e. they depend on random choices made within the model). Therefore to estimate the difficulty of a specific scenario, LandSFACTS should be run many times (e.g. a strict minimum of 10 times for statistically significant estimation, 100 times would be more relevant but is not always feasible due to lengthy run times) with the same inputs, in order to provide a range of difficulty indices that could be analysed subsequently.

7.5 LandSFACTS implementation

7.5.1 Program language

The LandSFACTS model is available as three independent C++ programs, written using Bloodshed Dev-C++, version 4.9.9.1. a Bloodshed Software (open source software, available on http://sourceforge.net), under a Microsoft Windows environment. The programs were built in a modular format designed to facilitate the implementation of further developments. The programs are comprehensively commented upon to facilitate program debugging and further developments.

7.5.2 Inputs and outputs format

The input tables and log files with the program results are text files delimited with tabulations, ".txt" extension. The data are within a database structure, cf. Appendix C. The database structure has the advantage of being a concise and effective way of storing data, as it prevents redundancy and incoherence in the dataset. User-friendly software facilitating the data input for the LandSFACTS C++ programs is presented in Section 7.5.4.

7.5.3 Model executable and example datasets

The model executable and the source code are available in Appendix B in CD format. An example dataset is also provided, with the data in the required text file format.

7.5.4 LandSFACTS software

To facilitate the use of the LandSFACTS C++ programs, they are embedded within the LandSFACTS software, developed by Joanne Matthews as part of the SIGMEA project (2005). The software provides graphical interfaces with wizards to facilitate the data entry by the user, and manages the coherence and dependencies of the data, by using a database. The software provides many extra tools, to only cite a few: determination of the areas, centroids, and neighbours of the fields from the landscape shapefile, display of crops and rotation names instead of ID, automatic calculation of the long-term crop proportions of each rotation and at the landscape level, checking of coherence between inputs.

The LandSFACTS software is written in Python version 2.4.3, and uses SQLite version 3.3.6 for database. The user interfaces are build up using PyQt version 4.1.1 based on Qt version 4.1.2. The help file is created using HelpMaker, version 7.3.

The LandSFACTS software version 1.6 was released on the 8 June 2007, as open source software under the GNU Public Licence. The software is complemented with a (i) help file detailing users inputs, data interpretation, project examples and technical information, (ii) example datasets and projects as explained in the help file, they are based upon the SmallLandscape and BigLandscape shapefiles, and (iii) a tutorial in Microsoft PowerPoint format detailing the building up of a new project. They are currently available within the Rothamsted Research website: http://www.rothamsted.ac.uk/pie/LandSFACTS/. They are included in Appendix B in digital format.

7.6 LandSFACTS current use

The LandSFACTS software and thus model is currently being used by researchers to investigates scenarios of the introduction of GM crops within European landscapes, within the SIGMEA project (2005). These current users are part of the following research organisations: INRA (Institut National de la Recherche Agronomique, France), CETIOM (Centre technique interprofessionel des oleagineux metropolitains, France), UPS (Universite Paris Sud, France), CSL (Central Science Laboratory, UK).

7.7 Conclusion

The LandSFACTS model as presented in this chapter, aims at allocating crops to fields to meet user-defined crop arrangements. The modelling approach and processes were kept as simple as possible, in order to allow future users of the software to be fully aware of the processes behind the crop allocations. The control of the crop arrangements is divided into two main sections, inputs controlling (i) temporal arrangement of crops: crop rotations, temporal constraints, initial crops of fields regulated by temporal patterns ("InitialCrops" program) and yearly crop proportions; and (ii) spatial arrangements of crops: possible crops in fields, crop rotation in fields regulated by spatial patterns ("RotationFields" program), and spatial constraints. The above inputs are based upon the definition of the crops and fields. Further inputs are required to control the behaviour of the model in the search for the desired crop allocation (iteration parameters and simulated annealing parameter), and to record the difficulty in obtaining it (penalties to fields). The "CropAllocation program" is based upon a linear programming technique, complemented with a controlled simulated annealing process. For the first year, the program allocates the crops to the fields using the initial crops in fields, this proposed crop allocation is authorised if it agrees with all the spatial, temporal constraints and yearly crop proportions. In this case, the program uses the crop rotation to determine the next year's proposed crop allocation. If a crop allocation is not authorised, it is improved by following the iteration parameters set up by the user, until either the proposed allocation is authorised or the program runs out of iterations allowed. In this latter case, the program is stopped. The model outputs the authorised crop allocation to fields and a report on the difficulty of obtaining it.

The three programs composing the LandSFACTS model are available as stand-alone C++ programs, with inputs and outputs as text files. However to facilitate their use, they are embedded within the LandSFACTS software, developed by Joanne Matthews as part of the SIGMEA project (2005). The software provides user-friendly interfaces with wizards guiding the user's inputs, and facilitating data management. The software is currently being used by researchers investigating the introduction of GM crops within European landscapes, within the SIGMEA project (2005).

In the next chapter, the model is assessed to determine if it meet the requirements set up by the users as specified in Chapter 2.

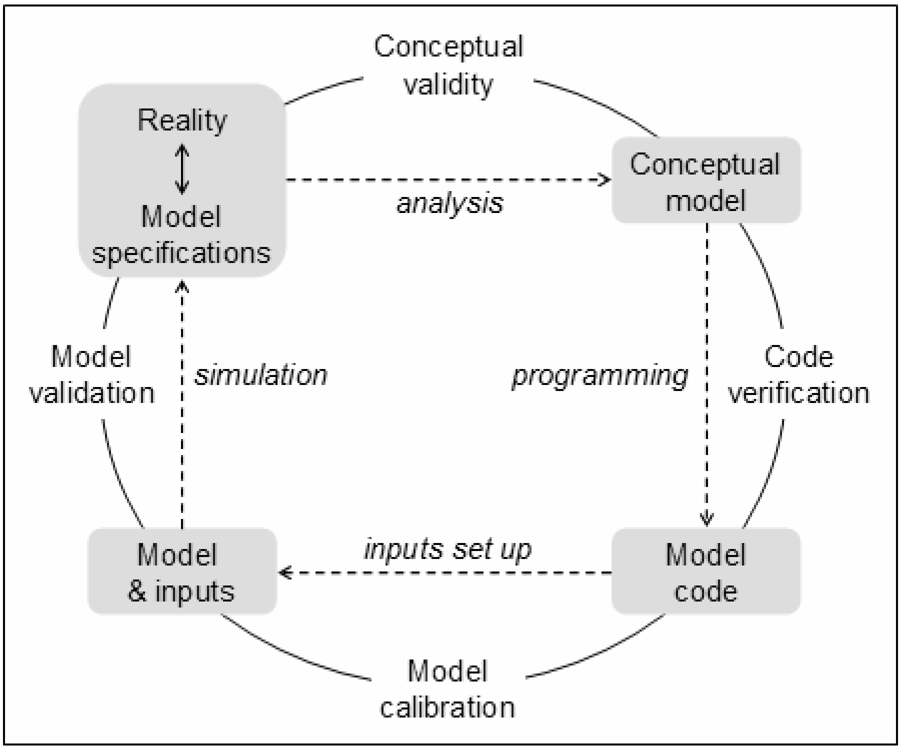
8 Model assessment

The LandSFACTS model as presented in Chapter 7 is the result of the work carried out in this thesis. The first step of the project was the identification of the model specifications detailed with the end users (Chapter 2). The knowledge derived from reviewing published research work (Chapter 3) allowed to determine a methodology for the development of the model (Chapter 4). Specifically designed tool were required: (i) new statistical analyses for spatial and temporal patterns of crops (Chapter 5), and (ii) a new mathematical representation of crop rotation (Chapter 6). Both tools were integrated within the LandSFACTS model, as reported in Chapter 7. The model and its software are available in Appendix B. After designing and developing the model, a further step is required: the model must be assessed for its fitness to purpose.

Models become more credible, and thus more likely to be used, after their fitness for purpose has been assessed and is clearly documented to the users of the model (Rykiel, 1996). The form of the assessment is dependent upon the aims of the model, and the approach chosen. The chapter describes a review of model assessment from the literature, after which an aim is identified for the LandSFACTS assessment. Then the main steps of the assessment were carried out and reported in sections on the assessment of the conceptual model, the code verification, the sensitivity analyses, and finally a case study. The model assessment is concluded by general recommendations on the model.

8.1 Model assessment in the literature

The model assessment must follow the steps of development of the model (Refsgaard and Henriksen, 2004; Rykiel, 1996; Sargent, 2003; Schlesinger *et al.*, 1979), cf. Figure 8.1.



<u>Figure 8.1: Diagram of modelling steps and assessments</u>

Adapted from Refsgaard and Henriksen (2004), Rykiel (1996), and Sargent (2003)

By representing the system reality within the scope of the model specifications, a conceptual model is built up. This step can be assessed for its conceptual validity. Then the conceptual model is translated into model code using programming. The code needs to be verified. The user may then input parameters within the model, also called calibration. And finally the model can be simulated using those inputs in order to obtain the model results. The assessment of this final step is the model validation.

Model assessment and particularly the definitions of the terms: validation, verification, and calibration are the focus of much on-going debate (Oreskes *et al.*, 1994; Pontius *et al.*, 2004; Refsgaard and Henriksen, 2004; Rykiel, 1996). The controversy arises from both semantic and conceptual philosophy.

Oreskes et al (1994) argue that verification is only possible on closed systems. As most earth science models are unable to encompass a whole system, they cannot be considered as closed, thus their verification is impossible. Within this context, verification can only be tested for the correct implementation of the conceptual model or algorithms to model codes (Hoover and Perry, 1989). In order to prevent unanticipated circumstances, Rykiel (1996) identifies two types of code verification errors: (i) mechanical, e.g. programming errors, usually solved by program debugging and (ii) logical, e.g. the logic employed within the programs. Code verification is very difficult for large models, therefore generally only the common circumstances of use of the model are being verified (Rykiel, 1996).

As stated by Oreskes et al (1994), validation does not aim at establishing if the model accords with reality. A model can be considered valid, if it doesn't have any known or detectable flaws (Oreskes *et al.,* 1994), and if it fulfils its specific purpose (Sargent, 2003). Sargent (2003) noted that the cost and time required for model validation over the complete scope of the model is often too expensive. Therefore, models are often only tested on a limited range of intended applications, and are thus only partially validated, or validated until proof of the contrary.

In order to improve communication concerning the validation processes, Rykiel (1996) advocates that three main elements should always be cited when reporting any validation processes: (i) the model's purpose, (ii) the criteria that must be met for validating the model, (iii) the context in which the model is valid. If those elements are not specified, the usefulness of the validation process is null, as it is unknown for what the model has been validated.

Many different techniques of validation are reported within the literature (Rykiel, 1996; Sargent, 2003), including qualitative and quantitative measures. The validation techniques are very often specifically designed for each model, the main techniques are: comparison to other models, degeneration tests, extreme condition tests, face validity, historical data validation, internal validity, parameter variability or sensitivity analysis, traces, predictive validation. Depending on the type of model, the available data, and time and cost constraints, one or several of the above validation techniques are used. For example, Pontius et al (2004) advocated four steps for the validation of land-use models: (i) budgets of the source of error; (ii) to compare the model to a Null model (no changes between the initial landscape and the predicted one), (iii) to compare the model to a Random model (random changes); (iv) to perform the analysis at multiple scales. The error budget is mainly composed of two elements: the errors of location and the errors of quantity (Pontius *et al.*, 2004). The determination of the respective part in the error budget of the location and quantity error is essential for identifying how to improve the model further. Another example is presented by Baudry (2003), who used random allocation of the crops into fields, in order to compare with the impact of using the "agronomic rules" on the crop allocation. Joannon's (2004) model validation aimed at determining if their crop allocation model could help explain the choices of farmers, and had no predictive purpose. The validation had the two following steps: at first only the rules of crop succession were investigated, then the second step considered the crop allocation simulated against the real crop allocation.

Model assessment is still the subject of much debate and discussion. However for determining the scope and application of models, it is invaluable. Detailing the aim of the assessment process and the techniques that are used for a particular model assessment, as advised by Rykiel (1996) and Pontius et al (2004), is indispensable for a good communication and understanding within the scientific community. Ideally model assessment should be carried out by outsiders of the model development, particularly to assure the independence of the assessment (Sargent, 2003). However, the assessment of the model in this thesis should provide useful information on the scope of the model, on known restrictions and highlight possible enhancements of the model. In the next section, the precise aim of the LandSFACTS assessment is presented.

8.2 Aim of LandSFACTS assessment

The full assessment of the LandSFACTS model is outside of the scope of this thesis, due to time constraint. Therefore the aim of this assessment was the evaluation of the credibility of the model within its normal scope of use. The normal use of the model is defined as using the LandSFACTS model for obtaining crop allocation meeting all the crop arrangements specified by the user. The LandSFACTS model is intended to be used (i) to create scenarios of crop arrangements, (ii) to fill up incomplete datasets, and (iii) to investigate the impacts of constraints on crop arrangements. The model should not be used to forecast or predict crop allocation, it is only a scenario building tool, respecting the users' specifications of spatio-temporal crop arrangements. Considering the circumstances detailed above, the LandSFACTS model is valid if (i) the obtained crop allocations meet all the constraints specified by the user; and if (ii) the iteration processes follow the parameters set up by the user.

8.3 Method of assessment

The model assessment was carried out in three steps: (i) evaluation of the conceptual model, (ii) code verification, and (iii) sensitivity analyses on iteration parameters and separation distances. The assessment was also complemented with a case study. The crop allocation obtained from the model could not be compared with historical data, as the model is not a forecasting tool. For the same reason, the correct calibration of the model, i.e. choice of the inputs, is the responsibility of the user.

The LandSFACTS model is composed of three programs: RotationFields, InitialCrops and CropAllocation. The first two programs provide an alternative way for the user to set up specific inputs for the crop allocation. Neither interferes with the processes within the CropAllocation program, which allocates crops to fields. Therefore their assessment was limited to code verification. The CropAllocation program, which simulates the crop allocation to fields, is a complex program involving stochastic decision making and constraint checking. Therefore, it was the main focus of the assessment.

8.4 Assessment of conceptual model

After Sargent (2003), a conceptual model is valid if (i) the assumptions and theories behind the model are correct and if (ii) the structure, logic, mathematical relations of the model are "reasonable". In relation to those points, the conceptual model was examined through the following topics: (i) temporal and spatial unit of crop allocation, (ii) crop rotations, (iii) control on crop spatio-temporal arrangements, (iv) landscape as a unique scale.

8.4.1 Temporal and spatial units

The main assumption of the model is its restriction to allocating only one crop to one field for every year. Two issues are linked to this restriction: the fixed boundaries of the fields over the years, and the limitation to one crop per year. The assumption of the model is the fixed boundaries of the fields over the years. In agricultural landscape, field boundaries are redefined over the years by merging with other fields, subdivision into smaller fields, or both at the same time. The LandSFACTS software does not incorporate this degree of complexity, and it is recognised as being an issue for crop allocation in some landscapes. The temporal restriction to one crop per year, is valid for European agriculture, where harvesting is usually annual.

8.4.2 Crop rotations

The crop rotations represented as transition matrices and complemented with the temporal constraints, direct the crop allocation to fields over the years. This method allows the modelling of fixed and flexible crop rotations. However, the model only integrates constant probabilities of transitions: they cannot evolve through time. This possibility of evolution would be important for modelling evolving landscapes, such as for climate change scenarios. A mitigation measure can be currently used by using master transition matrices, which would regulate the transition from one crop rotation to another crop rotation (cf. Chapter 6.4: Transition between rotations, p.90).

8.4.3 Control on crop spatio-temporal arrangements

The LandSFACTS model integrates various ways of directly influencing the spatial and temporal arrangement of crops within the landscape. The model controls the crop arrangement with the crop rotations to fields, spatio-temporal constraints of crops, using empirical tests; and spatio-temporal patterns of crops, using statistical measures. The limits of the crop patterns statistics were reported in Chapter 5.1 (New statistical analyses on crops' spatial and temporal patterns, p.45).

Moreover the control of the spatial and temporal patterns is currently outside of the main simulation (RotationFields and InitialCrops program). Therefore the patterns are fixed over the years, which can be considered as a draw back in evolving landscapes, where crop rotations are highly variable (not fixed).

8.4.4 Landscape as an unique scale

The analyses of spatio-temporal patterns on the Burgundy dataset cf. 5.2 (Crop pattern analyses on landscape datasets, p.66), showed a scale dependency in the crop patterns. Therefore the model should have provided the possibility of controlling crop patterns at different scales, such as farmers (group of fields) and groups of farmers (e.g. cooperatives). Due to time constraints, the control on the crop patterns is only available at the landscape scale, instead of multiple scales. The current conceptual model is not wrong, but could benefit by integrating different scales of interactions.

8.4.5 Conclusion

The conceptual model of LandSFACTS has some important restrictions, such as the limitation to allocating only one crop per year, fixed field boundaries, constant crop rotations, and a unique scale for crop pattern control. The model would greatly benefit from overcoming them. However, the model does not aim at forecasting real crop allocation; it is only a tool to create scenarios of crop arrangement. Therefore, as long as the restrictions of the model are clearly identified and communicated to the user, the conceptual model can be considered valid within its restrictions.

8.5 Code verification

Code or program verification aims at ensuring that the implementation of the conceptual model into computer programs is correct (Sargent, 2003). No computer program of consequent length can be fully verified, however, they should be tested as fully as possible. The LandSFACTS programs were verified at several stages. At first, during the program development, the process of each line of code was checked. Then general tests were carried out to verify the agreement between expected and obtained results from the program. The main technique used was the “degenerative test”. For this test, inputs were carefully chosen to test how the model behaved in specific circumstances. For example, to test if the model forbade correctly the return period of crops (e.g. a minimum of one year gap may be required between wheat crops), a crop rotation with this forbidden sequence (e.g. rotation with continuous wheat) was allocated to all the fields. The crop allocation obtained should not have two wheat crops consecutively grown on any field. All the constraints imposed on the crop allocation have been checked one by one during their development. Further tests were carried out by combining different constraints, and checking their simultaneous integration within the software. A list of some of the general tests carried out is reported in Appendix D.

A source of possible errors is the stochastic processes occurring within the model. The stochastic decisions are based on a pseudo random generator from the standard C++ library GCC (Gnu Compiler Collection) version 3.3.1. A pseudorandom number generator is an algorithm generating a sequence of numbers, which approximate random number properties. With the same “seed”, i.e. number to initiate the generator, the sequence of random numbers is identical. This particularity allows exact replication of simulations, which is particularly useful for debugging a program or to investigate the influence of variables. Depending on the user choice, the generator is started either with a seed based on computer time or on a specific seed provided by the user. The first random number generated is never used, to avoid biased results (time seed of consecutive simulation could be very close or even identical). The random numbers generated were visually tested for randomness, cf. Figure 8.2. Over 1,000 sequentially generated random numbers, they visually seemed to be spread out from 0 to 1 included. No specific patterns in the number generation were recognised; therefore, this random generation was acceptable for the purpose of LandSFACTS model. However, for future improvement, a more robust random generator might be required.

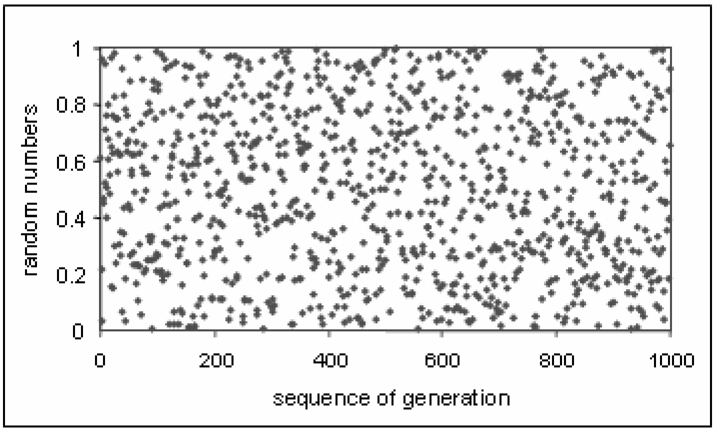


Figure 8.2: Visual test of generated random numbers within LandSFACTS model

Stochastic decisions are occurring at three instances within the program: when (i) choosing randomly an initial crop; (ii) choosing randomly a crop based on the previous crop using the transition matrix; (iii) choosing a field from all the problematic fields, to alter its crop. Each of those options was tested as indicated in Appendix E. The randomness of the stochastic process within the model appeared satisfactory for the desired level of the model requirements.

8.6 Sensitivity analyses

The sensitivity analyses were used to identify the impacts of the model's parameters upon the difficulty of obtaining a crop allocation. Ideally, all the parameters of the model should have been analysed, however due to time constraints, only the three following were chosen: (i) comparison between one or all crop alterations (i.e. iteration option 1 or 2.1), (ii) simulated annealing value, and (iii) the distances for the spatial constraints. They were chosen, as their impact on the difficulty of obtaining a landscape was not easily predictable (i.e. straightforward). The investigation on the simulated annealing also provided an insight on how to set up this variable. And the analyses on a wide range of distances for spatial constraints was used to help understand their impact on the difficulty of obtaining a crop allocation.

8.6.1 Datasets for sensitivity analyses

For the sensitivity analyses, two datasets were used: the Fife and the Beauce study area. For the statistical significance of the sensitivity analysis, datasets with a high number of fields were indispensable. Thus, the Fife study area with 388 fields and the Beauce study area with 1,993 fields were advantageous. The sensitivity analysis did not aim at replicating an existing landscape, but at analysing how the model behaved. Therefore, the agronomic information was adapted to the needs of the analysis, and did not reflect reality within the study areas, Table 8.1. Both datasets have many differences, e.g. number of fields, fields shapes, rotations, separation distances, therefore their results should not, and cannot be directly compared. They provide two independent platforms, on which to test the sensitivity of the model.

Table 8.1: Summary of the Fife and Beauce datasets for the sensitivity analyses

	Fife	Beauce
Number of fields	388	1,993
Crops	13 crops, including GM and conventional oilseed rape	11 crops, including GM and conventional oilseed rape
Crop rotation(s)	2 rotations (cf. Appendix F): - permanent grassland - all crops (11), flexible rotation, probabilities adapted to yearly crop proportions	1 rotation (cf. Appendix F): “all crops” rotation, flexible rotation for 10 crops, probabilities adapted to yearly crop proportions
Rotations in fields	114 fields are permanent grassland (29% of fields), the other fields have the all crops rotation	All fields with “all crops” rotation
Initial crops in fields	random	random
Spatial constraints	100m between GM and conventional oilseed rape. If too close, the GM crop must be altered first.	200m between GM and conventional oilseed rape. If too close, the GM crop must be altered first.
Temporal constraints	- After GM oilseed rape, no conventional oilseed rape - cereals: up to three years in a row - winter crops: up to three years in a row	- wheat: two years in a row maximum - after GM oilseed rape, no conventional oilseed rape the next year or the year after
Yearly crop proportions	- winter conventional oilseed rape: 0.18 constant over years - wheat: 0.22 constant over years - set-aside: 0.08 constant over years - spring conventional oilseed rape: 0.04; 0.04; 0.03; 0.02; 0.01 - GM oilseed rape: 0; 0; 0.01; 0.02; 0.03	For every year: - conventional oilseed rape: 0.15 - GM oilseed rape: 0.05 - wheat: 0.3
Years of simulations	5	5

8.6.2 Comparison between one or all crop alterations

The Fife dataset was used to compare the probability of finding authorised crop allocations with two different sets of iteration options: only 10,000 maximum iterations for option 1 (all fields have their crop randomly altered) or only 10,000 maximum iterations for option 2.1 (one problematic crop has its crop randomly altered). Each set was run for 50 simulations of 5 years. Option 2.1 was run firstly without any simulated annealing (a value of -1), and secondly with simulated annealing using a value of 50. The results are presented in Table 8.2.

Table 8.2: Comparison between the successes of crop allocation using iteration options based on the alteration of random choices (option 1 and 2.1)

Iteration option	1 (all fields)	2.1 (one field)	2.1 (one field)
Simulated annealing	not applicable	not used	50
Percentage of successful simulations	0	72%	100%
Average of successful years	0	3.6	5

When all the fields had their crops altered at each iteration (option 1), the program was unable to find any crop allocation meeting all the constraints. However if only one crop was altered at a time (option 2.1), with no simulated annealing values set, 36 out of 50 simulations were completely successful (authorised crop allocation for the 5 years). On average, only 3.6 years of successful crop allocations were found out of 5. If the simulated annealing value was set to 50 to avoid local minima, the program found authorised crop allocation for all five years of all the 50 simulations (5 successful years out of 5).

In conclusion, the use of the optimisation algorithm (option 2) increased the chances of finding an authorised crop allocation over a non-optimisation algorithm (option 1). The advantage of using a simulated annealing value was also highlighted; the impact of this variable was further analysed within the next section.

8.6.3 Simulated annealing

The simulated annealing option aimed at preventing the program getting blocked in a local minimum before finding a crop allocation for all the required years. It was most useful with iteration option 2.1 labelled as “one randomly chosen problematic field had its crop randomly altered”, (7.2.8: Iteration options, p.102). Both datasets were tested with 10,000 maximum iterations per year for option 2.1 (penalty value = 1), over 5 years and with various simulated annealing values. For every simulated annealing value, the Fife dataset was run 50 times, and the Beauce dataset 5 times. The limited number of replicates was due to time constraints, e.g. the Beauce dataset, with its 1,993 fields required at least 5 hours per simulation.

Simulated annealing values and successful simulations

For both datasets, when not using the simulated annealing option, not all required years had an authorised crop allocation, i.e. on average, only three years out of five were authorised for the Fife dataset, and no years for the Beauce dataset, cf. Figure 8.3.

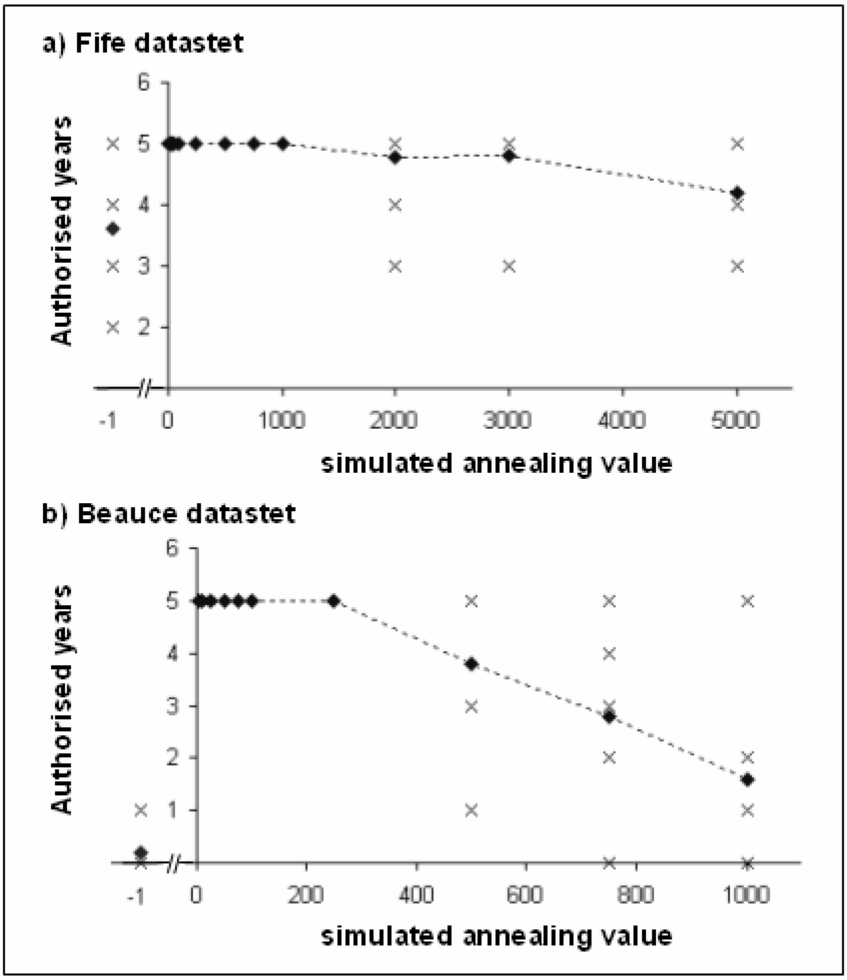


Figure 8.3: Impact of simulated annealing values on the number of years with authorised crop allocations within a) Fife dataset and b) Beauce dataset

The dotted line represents the average value of authorised years over 50 simulations each. The cross points are individual values for the 50 simulations. They show the values obtained for each specific simulated annealing value (up to 6 individual values:

from no year authorised up to the 5 years authorised). The points on the left side of the vertical axis represent simulation without the simulated annealing option.

With simulated annealing values ranging from 3 to 1,000 on the Fife dataset, crop allocations were found for each year. With higher simulated annealing values, simulations were less successful at finding authorised crop allocation for the 5 years. In average, for a simulated annealing of 5,000, the simulations failed finding an authorised allocation for the last year (authorised year = 4). The results from the Beauce datasets followed the same pattern of 100% of authorised crop allocation for 5 years for simulated values ranging from 3 to 250. With higher simulated values (500 to 1,000), the proportions of authorised years found decreased sharply to reach only one year in five with authorised crop allocation for 1,000 simulated annealing values.

The differences in results between the two datasets, i.e. the simulated annealing values and the sharpness of the decrease, might be due to several combined factors. The constraints on the Fife dataset are less restricting than on the Beauce dataset, particularly in regards to (i) the number of fields where GM can be grown (Fife: 388-114 fields; Beauce: 1,993 fields), and (ii) separation distances (Fife: 100m; Beauce: 200m). Moreover, the maximum number of iterations was fixed at 10,000 (one field has its crop changed at each iteration) for both datasets, regardless of their characteristics. For example, for a simulated annealing value of 1,000, every time the simulated annealing option is used (a worse crop allocation temporarily accepted to unblock the program), it is 1,000 iterations over 10,000, which have been used, i.e. 1/10th of the possible iterations. Therefore, every time the simulated annealing option is used, less iterations are available to reach an authorised allocation afterwards. This is more critical for the Beauce dataset (for which on average only 1.5 years produced valid, authorised data) with its higher number of fields with potential problems (cf. above) than for the Fife dataset (for which all 5 years produced authorised data). This exemplifies that the ratio of the simulated annealing parameter to the maximum number of iteration is a determining factor that should be adapted to the potential number of fields with problematic crop allocations. With a higher number of maximum iterations for both dataset, it would be expected that the decrease in successful simulation would occur at higher simulated annealing values. Further simulations and case studies would be required to capture the relationships between them, and set up rules for choosing appropriate values of the simulated annealing parameters.

In conclusion, simulated annealing values were indispensable for the optimisation processes aiming at authorised crop allocations. However, after a specific simulated annealing value threshold, the simulated annealing process did not facilitate the search for authorised crop allocations. The threshold seemed to depend upon the number of potential problematic fields (dependent upon the difficulty of the constraints) and, the maximum number of iterations available.

Simulated annealing values and overall penalties

The analysis of the relationship between the overall penalty and the simulated annealing parameter value was important to determine if or how the simulated annealing parameter value influenced the overall penalties of the crop allocations. Only the simulations successful for all the years were considered for this analysis, i.e. simulated annealing values between 3 to 1,000 for the Fife dataset, and between 3 to 250 for the Beauce dataset (cf. previous section: Simulated annealing values and successful simulations). Every time a crop was successfully changed on a field (improving the previous allocation by using or by not using the simulated annealing option), a penalty of 1 is added to the field. The overall penalties are the sum of all the penalties of all the fields for all the simulated years.

For both datasets, the overall penalties decreased exponentially with increasing simulated annealing values, before they stabilised at around 290 penalties for the Fife dataset and 640 for the Beauce dataset, cf. Figure 8.4. The standard deviation decreased dramatically with increased simulated annealing values. The shape of the curves can be explained by the fact that with small simulated annealing values, altered crop allocation are accepted very often (i.e. many worse allocation accepted as better). Therefore, more penalties are applied to fields. With higher simulated annealing values, the iteration process has more iterations to find an altered crop allocation which improves crop allocation. Moreover, the proportion of the number of iterations accepted using the simulated annealing option over the overall number iteration needed (overall penalties) is more stable with higher simulated annealing values, thus decreasing the variability in overall penalties between simulation runs.

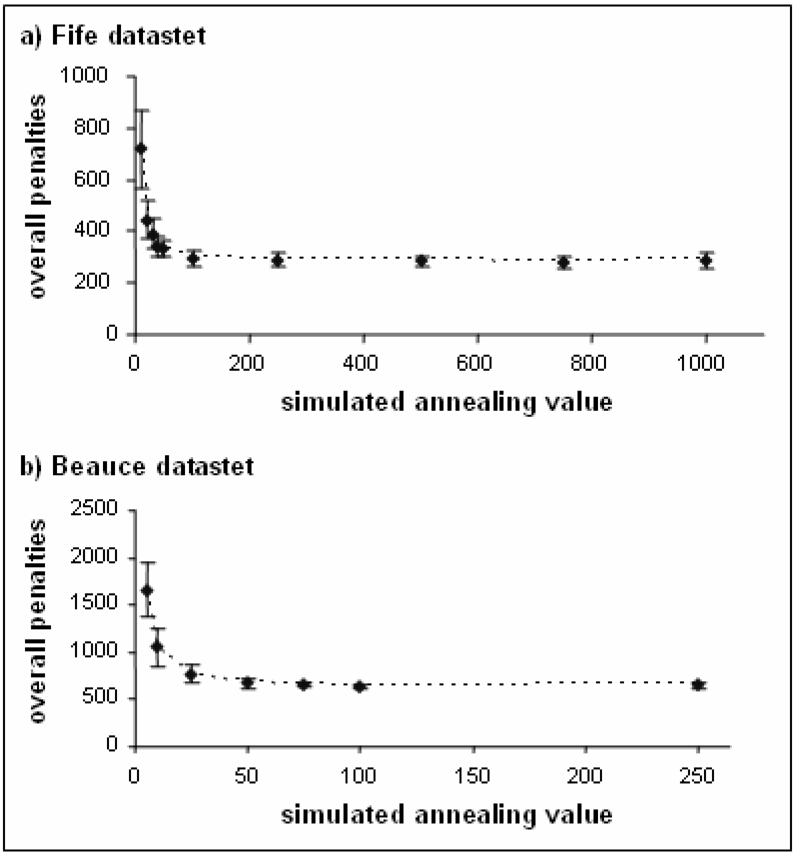


Figure 8.4: Impact of simulated annealing values on overall penalties within a) Fife dataset and b) Beauce dataset.

The error bars represent the standard deviation from 50 simulations. All of the points represented had the 5 years of authorised crop allocations.

Therefore after the analyses, to find the lowest and the most stable overall penalties between many runs of the same simulation, a high simulated annealing value is recommended. However, as demonstrated earlier a high simulated annealing value might decrease the probability of finding authorised crop allocations.

Simulated annealing values and simulation time

An important consideration, when running simulations is the time required to obtain an authorised crop allocation. The time required is directly linked with the number of iterations that have to be performed before finding an authorised crop allocation, for a given set of conditions to respect. Therefore, to optimise the time cost, the limitation of the number of iterations required is an important consideration.

In Figure 8.5, the numbers of iterations used to find authorised crop allocation are shown in relation to their simulated annealing values. For both datasets, the number of iterations decreased with increasing simulated annealing value down to an optimised value, 50 for the Fife dataset and 10 for the Beauce dataset, before it steadily increased. With small simulated annealing values, crop allocations regardless of their number of unauthorised crops, are regularly accepted during the iterations, thus preventing any optimisation process. Whereas with large simulated annealing values, "worse crop allocations" are not very often accepted, thus the program has a large number of iterations available to find a better allocation from the current crop allocation. In the latter situation, the process is more optimised but may require more iterations, i.e. time. To optimise the time required for simulations the simulated annealing values should be chosen to correspond to the dip in number of iterations.

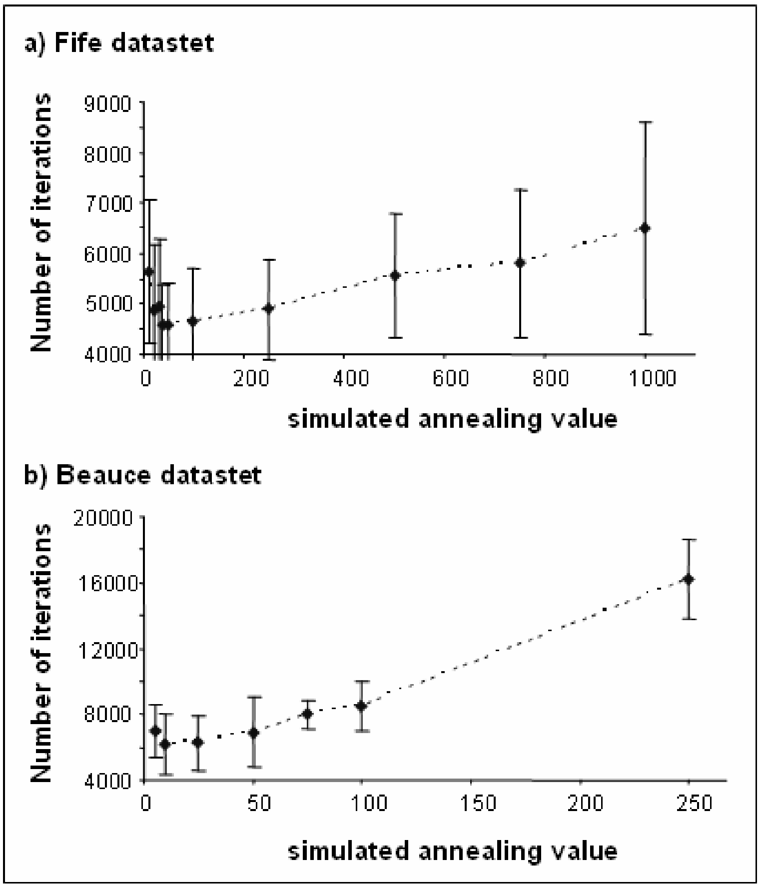


Figure 8.5: Impact of simulated annealing values on the number of iteration used within a) Fife dataset and b) Beauce dataset.

In conclusion, the simulated annealing values for a given maximum number of iterations influenced (i) the probabilities of finding authorised crop allocations, (ii) the overall penalty incurred to find authorised crop allocation, (iii) the variation in overall penalties between similar runs, and (iv) the time required to find authorised crop allocations. Therefore, the simulated annealing values should be chosen very carefully. Moreover as the optimum value was dependent upon the dataset (e.g. landscape, crops, and the constraints on the crop allocation), ideally, the above analyses should always be carried out to determine the optimum simulated annealing value, before producing any results from the LandSFACTS model. As a very rough guideline, the maximum number of iterations should be, at the very least, double the expected number of fields with problematic crop allocation, or the number of fields simulated, whichever is the greater. Furthermore, the simulated annealing parameter value should be set initially to 0.01 times the maximum number of iterations. This guideline could be used to run the first few simulations, then, depending on the results (no authorised crop allocation found, or successful crop allocation but too time consuming), both values should be refined further, using the conclusion drawn from the above sensitivity analysis. For example, in the case of unsuccessful simulations, the first step is to increase the maximum number of iterations. More extensive sensitivity analyses would be required to provide more authoritative guidelines on the setting up of the simulated annealing parameter value.

8.6.4 Separation distances

The impact of separation distances on the crop allocation of a landscape should be dependent upon: (i) the proportions of the targeted crops in the landscape, (ii) the spatial patterns of the targeted crops in the landscape, (iii) the distance to be respected, and upon (iv) the mosaic of the fields, e.g. size of fields, field shapes, and adjacency of fields. The impact of increased separation distances on the difficulty of finding authorised crop allocation was studied on two datasets.

Both datasets were set up to run for one year with fixed proportions of the crops being separated (Beauce: 15% conventional oilseed rape, 5% GM oilseed rape; Fife: 3% conventional and 3% GM), and for a wide range of separation distances (Beauce: from 0 to 300m with 6 values, Fife: from 100 to 2,000m with 8 values). For both datasets, the maximum number of iterations (option 2.1) was fixed at 10,000 per year and the simulated annealing value at 50. Fifty replicates were run for each scenario, except for the Beauce dataset where 40 replicates were run for the 100 and 200m scenario, and 10 replicates for the 300m scenario (restrictions due to higher run time). Crop allocations meeting the constraints were found for each scenario (Figure 8.6), except for the Beauce dataset with the 300m separation distance, for which 50% of the simulations were unsuccessful at finding authorised crop allocation. An authorised crop allocation could have been found with higher maximum iterations; however, this was not tested to respect the consistency between the scenarios and because of the time limitations for the analyses.

For both datasets, with increasing separation distances, the overall penalties for finding authorised crop allocations increased (Figure 8.6). The form of this increase was different for the two datasets. The Beauce dataset showed a linear increase in overall penalties with increasing separation distances (r^2 = 0.94), whereas the Fife dataset demonstrated an exponential increase (r^2 = 0.99). When only the lowest separation distances of the Fife datasets are plotted, the relationship could be explained with a linear line (r^2 = 0.99). Therefore, it is proposed that if the Beauce dataset was simulated with higher separation distances and with higher iteration maximum, the overall penalty might increase exponentially with increasing separation distances, as is shown by the Fife dataset. However, this proposal would have to be tested.

In the case of the Fife dataset, the standard deviation of the overall penalties increased with increasing separation distances. The large variation in iteration numbers is probably due to the initial random choice of the crop allocation; if the program randomly chose a crop allocation with few unauthorised crop allocations, less iteration were required to obtain the desired authorised crop allocation, than if the initial choice had many unauthorised crop allocations. The impact of the initial crop allocation would be higher with harsher constraints. For identical separation distances, for example 100m, the average overall penalty for the Beauce dataset was 130 (standard deviation: 22), and 3.5 (standard deviation: 3.48) for the Fife dataset. This important difference could be caused by several contributory factors. Firstly, the difference between the datasets results arises from the number of fields with potential unauthorised crops. The finding of authorised GM oilseed rape allocations with a given separation distance from conventional oilseed rape would be more difficult in Beauce than in Fife, for the following main reason. Beauce has more fields that can grow GM than Fife (overall number of fields, and number of fields with oilseed rape in their rotations), thus more oilseed rape fields are within the separation distances, i.e. more fields might need to have their crop changed (for which more iterations would be required). Despite this difference in the difficulty of finding a crop allocation, both datasets were run with the same number of maximum iterations: 10,000. This number of iterations is enough for the Fife dataset, but not for the Beauce dataset, as is shown by the unsuccessful simulation for separation distances higher than 200m. A further factor would impact on the differences between the datasets: the shape and mosaics of the crops, which might influence the number of fields within the separation distances. Unfortunately, this factor was not quantified within the assessment.

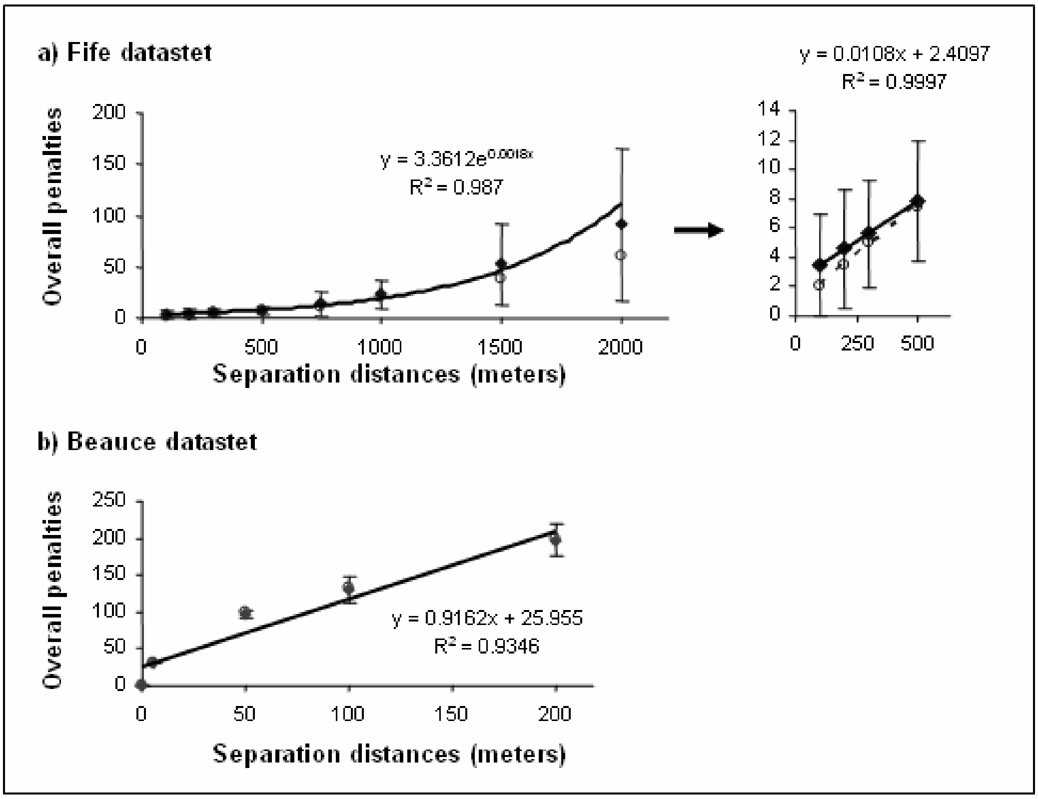


Figure 8.6: Impact of separation distances within a) the Fife dataset and b) the Beauce dataset on the number of overall penalties.

To thoroughly understand the impact of separation distances on the difficulty of finding authorised crop allocation, further analyses would be required. A possible study would be to alter the crop proportions over the landscape to quantify its influence on finding crop allocation with separation distances. Another study concentrating on the impact of field shape and mosaic upon crop allocation would also be useful; this would require the qualification and quantification of the fields in the landscapes. These studies would provide an insight into the feasibility of separation distances for landscapes with different proportions of the crops and different types of field shape.

8.7 Scenario testing for a real landscape

The scenarios exemplified in this section only demonstrate the use of the LandSFACTS model. This study did not aim to provide any complete evaluation of the proposed scenarios, and did not aim to replicate real landscape situations.

The scenarios aimed to investigate the impact of spatial patterns of GM oilseed rape and of specific separation distances between GM and conventional oilseed rape, on the difficulty of obtaining authorised crop allocations. To reach this aim, four scenarios were developed: one for each extreme of spatial pattern of oilseed rape varieties (A: aggregated and B: regular), to be tested with (A1 and B1) and without (A0 and B0) the separation distance. The spatial pattern of oilseed rape varieties (as described in the following section) was controlled by imposing the pattern on the allocation of the crop rotations to fields. Scenarios without the separation distances were useful to construct a baseline, against which to assess the scenarios with the separation distance. In summary, the four scenarios were:

- A0: spatial aggregation of GM oilseed rape rotation, no spatial constraints,
- A1: spatial aggregation of GM oilseed rape rotation, 50m spatial constraints,
- B0: spatial regularity of GM oilseed rape rotation, no spatial constraints,
- B1: spatial regularity of GM oilseed rape rotation, 50m spatial constraints.

The setting up of the scenarios is presented below, followed by the analyses of the results.

8.7.1 Setting up scenarios

The scenarios were based upon the Beauce dataset as the high number of fields facilitates statistical analyses and increases the credibility of the results. However, using the complete cropping system of the Beauce dataset would have unnecessarily increased the difficulty of both setting up and interpreting the results. Therefore only a very simplified version of its cropping system was used, i.e. its nine main crops (oilseed rape, wheat, spring and winter barley, peas, maize, sunflower, set-aside and other cereals), flexible crop rotations complemented with temporal constraints, and no spatial restrictions on the extent of crops (all crops available to all fields). The separation distance between GM and conventional oilseed rape was set to 50m, this small distance was chosen to optimise the simulation time and increase the chance of finding authorised crop allocations. If a GM and conventional oilseed rape were too close to each other, the GM crop had to be changed. Because this set up could result in no GM oilseed rape being allocated to fields, it was necessary to control the yearly proportion of GM oilseed rape. Within the LandSFACTS model, the spatial pattern of crops is controlled by imposing spatial pattern on the crop rotations, i.e. when rotations are allocated to fields. This required that a GM oilseed rape crop must be present in at least one crop rotation, but also that at least one crop rotation must not contain this crop. The cropping system and setting up of spatial patterns are detailed below.

8.7.1.1 Cropping system

The cropping system of the dataset was reduced to three crop rotations. The first one was permanent set-aside. The other two had flexible crop proportions: one with and one without GM oilseed rape. For both rotations, the probability of transition from any crop to oilseed rape (GM or conventional depending upon the rotation) equalled 0.2, to wheat 0.3, to sunflowers, maize, peas or spring barley 0.05, to set-aside, winter barley or other cereals 0.1. The flexible crop rotations were complemented with the following temporal constraints: (i) wheat could only be grown two years in a row, (ii) oilseed rape had a return period of 3 years maximum (i.e. at least one year gap between two oilseed rape crops), and (iii) conventional oilseed rape could not be grown if GM oilseed rape was in the field two years ago. Due to the high flexibility of the crop rotations, the proportion of the main crops was controlled by using the yearly crop proportion constraint: 15% of conventional oilseed rape, 5% of GM oilseed rape and 30% of wheat. The allocation of a rotation to each field is presented in the next section, as it imposed the spatial pattern of the crops. Because the rotations were very flexible, the choice of the initial crop did not impose a temporal pattern on the crops. Therefore, the initial crops were set as randomly chosen from the crops of the rotation of each field at the start of the simulation; this initial allocation could then have been altered during the iteration process in order to meet all the specified constraints, e.g. crop proportions or temporal constraints.

8.7.1.2 Spatial patterns of the rotation with GM oilseed rape

The allocation of the rotations to the fields was done by using the RotationFields program. The constraints were 15% (0.2 authorised deviation for all proportions, i.e. for 15%, acceptable values were between 12% and 18%) of the area with conventional oilseed rape, 5% with GM oilseed rape, 30% with wheat, 10% with set-aside, and the fields with the GM rotation should be spatially aggregated for scenario A (Ep lower than 0.05, with 1,000 randomisation points) and spatially regular for scenario B (Ep higher than 99.95, with 1,000 randomisation points). The crop proportions within the rotations and for the whole landscape were chosen carefully, in order to be mathematically compatible. In the above circumstances, any rotation allocation, which had the rotation with GM allocated for around a quarter of the fields, and the rotation with conventional oilseed rape for three quarters of the fields, were respecting the constraints on crop proportions.

The resulting crop rotation allocations are presented in Figure 8.7 and Figure 8.8, and they met the requirements set up for crop proportions and spatial patterns of the rotations, cf. Table 8.3. The proportions of GM and conventional oilseed rape rotations were slightly different between the two scenarios (more fields with the GM rotation, i.e. 1.4% of the landscape area, for scenario A than for scenario B), but they both met their targets. Closer proportions of the crops between the two scenarios would have been better to isolate the influence of spatial patterns of the rotations from crop proportions. The difference in number of fields between scenarios was due to the variable field areas.

Table 8.3: Characteristics of crop rotation allocations for scenario A and B (long-term crop proportions, levels of spatial patterns, and proportions for each rotation).

Crops	Scenario A (aggregation)	Scenario B (regularity)	Targets
Conventional oilseed rape	0.145	0.155	0.15 (0.2)
GM oilseed rape	0.055	0.041	0.05 (0.2)
Wheat	0.3	0.294	0.3 (0.2)
Sunflowers	0.05	0.049	-
Maize	0.05	0.049	-
Set-aside	0.1	0.119	0.1 (0.2)
Peas	0.05	0.049	-
Winter barley	0.1	0.098	-
Other cereals	0.1	0.098	-
Spring barley	0.05	0.049	-
E values	0.987222	1.01899	-
Ep values	0	100	Scenario A: < 0.05 Scenario B: > 99.95
Number of fields with GM oilseed rape rotation (in brackets the area in km^2)	578 (25.4)	391 (19.0)	-
Number of fields with conventional oilseed rape rotation (in brackets the area in km^2)	1415 (66.8)	1558 (71.3)	-

Figure 8.7: Crop rotation allocation with aggregated GM rotations (Scenario A).

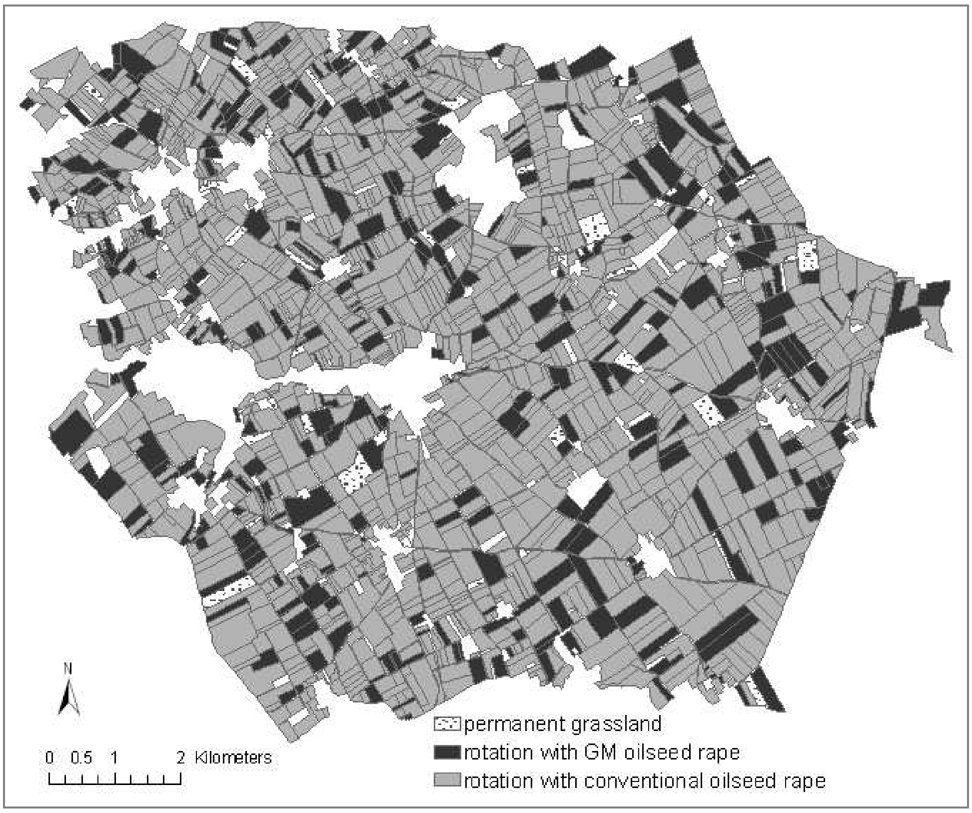


Figure 8.8: Crop rotation allocation with regularly spaced GM rotations (Scenario B).

The difference in spatial patterns of the GM oilseed rape was visually identifiable. The spatially aggregated GM oilseed rape fields (scenario A, Figure 8.7) were concentrated in the middle of the study area, while on the "regular landscape" (scenario B, Figure 8.8) they were spread out through the landscape. Moreover, the number of GM fields being neighbours of other GM fields was visually much higher in the aggregated landscape than in the regular one, this was corroborated by results in Table 8.4. In scenario A, fields with GM oilseed rape had twice the probability of being within 50m of another field with GM oilseed rape than in scenario B, but they also had a higher probability of conflicting with conventional oilseed rape rotations (GM rotation within 50m distance from conventional rotation). It should be noted, that for scenario A due to the higher number of fields with the GM rotation, more fields were available for growing GM oilseed rape. Therefore, finding authorised crop allocation for scenario A could be artificially facilitated.

Table 8.4: Number of neighbouring fields with GM or non-GM rotations for scenarios A and B.

Rotations within 50m distance		Scenario A	Scenario B	A – B
GM rotation	GM rotation	1074	524	550
GM rotation	Conventional rotation	5234	3929	1305
Conventional rotation	Conventional rotation	6377	7703	-1326
Any rotation	Permanent grassland	0	529	-529

More extreme spatial patterns could be found, by using a randomisation curve with more than the current number of points (1,000). A higher number of randomisation values would increase the chance of finding extreme patterns, i.e. increasing the tail of the distribution. However, finding a desired rotation allocation with a large number of randomisation points would require more computer time. With the definition of the cropping systems, the allocation of the rotations to fields, and the choice of the initial crops, the agronomic parameters of the scenarios were set up. The next section presents parameters controlling the behaviour of the LandSFACTS model.

8.7.1.3 Iteration parameters

The iteration parameters control how the model alters the proposed crop allocations, when it does not meet the constraints specified by the user, in order to find authorised crop allocations (meeting all the constraints). The iteration parameters for the scenarios were set as 100 maximum iterations with all fields randomly altered (option 1, no penalty), 10,000 maximum iterations with random alteration of a problematic field (option 2.1, penalty equals 1), and 500 maximum iterations with the universal crop (option 2.3, penalty equals 100) labelled "flag crop" with a simulated annealing value of 50. The iteration parameters were used to determine the best allocation out of 100 full random allocations, before optimising it by using a new random choice within the transition matrix of the rotations. As a last resort, the universal crop could be allocated. The parameters were identified from the results of the sensitivity analysis.

8.7.1.4 Summary of scenarios

The four scenarios described in Table 8.5, using the inputs parameters already detailed in this section were simulated over three years. Each scenario was run 10 times with different random decisions (i.e. random numbers generated using time based seeds). The results are presented in the following section.

Table 8.5: Summary of characteristics of the four scenarios.

		Spatial patterns of GM oilseed rape rotation	
		aggregation	regularity
Spatial constraints	none	A0	B0
	50m	A1	B1

8.7.2 Results from scenarios

The statistical summary of the difficulty of obtaining authorised crop allocation for each scenario is presented in Table 8.6 (refer to Appendix G for detailed data). The difficulty of obtaining crop allocations with the scenarios without separation distances (scenario A0 and B0) was not significantly different from each other ($p = 0.01$, cf. T-test results). These two scenarios always provided the desired crop allocation with around 354 (+-33) as overall penalties and around 4,000 iterations. It can therefore be concluded, that the spatial patterns of the rotation with GM oilseed rape did not interfere with the difficulty of finding an authorised crop allocation when there were no spatial constraints (only temporal constraints and yearly crop proportions).

Table 8.6: The proportion of successful simulations, the number of iteration and of penalties (statistical significance shown) in obtaining crop allocation for each scenario.

Scenarios ID	Proportion of successful simulations		Number of iterations used		Overall penalties	
	Average	Stdev	Average	Stdev	Average	Stdev
A0	100	0	3973	939	355	34
B0	100	0	4139	742	353	32
A1	100	0	9903	1438	671	58
B1	30	48.3	18090	2866	1488	536
B1s	100	0	15588	3589	713	64
Student T-test, with p = 0.01. and 18 degree of freedom (except for B1s: 11 degrees)						
A0 vs B0	-		Not significantly different		Not significantly different	
A1 vs B1	-		Significantly different		Significantly different	
A0 vs A1	-		Significantly different		Significantly different	
B0 vs B1	-		Significantly different		Significantly different	
B0 vs B1s	-		Significantly different		Significantly different	
A1 vs B1s	-		Significantly different		Not significantly different	

B1s: only successful simulations (3 replicates out of 10)

Stdev: standard deviation

The scenarios with separation distances of 50m between GM and conventional oilseed rape were significantly different ($p < 0.01$ for their overall penalties) from their respective scenario without separation distances. With a separation distance of 50m, scenario A1 (spatially aggregated GM oilseed rape fields) had authorised crop allocation for all the simulations, whereas in the case of scenario B1 (spatially regular GM oilseed rape fields) 70% of the simulations (7 out of 10), failed to find an authorised crop allocation for the third year. The comparison of the number of iterations and overall penalties of scenario A1 and B1, showed that they are statistically different ($p < 0.01$). However as scenario B1 had 7 replicates without complete crop allocations (only two years out of three), the three successful simulations were grouped in a sub-sample called B1s. A significantly higher number of iterations was required to find authorised crop allocation when the GM crops were spatially regular ($p < 0.01$). However, the number of fields which had to have their crops altered (quantified with the overall penalties) was not significantly different between A1 and B1s. The results from B1s were only based upon three replicates, which had successful crop allocations over ten replicates, i.e. they were the three most efficient run out of ten. Therefore, they do not fully represent scenario B1. To increase the number of successful simulations of scenario B1, higher maximum iterations would have to be needed, for example, the maximum number of iterations of option 2.1 (one new random crop) could have been increased from 10,000 to 20,000. However considering that for those scenarios, one iteration required around 1.5 seconds of computer time, a simulation could last up to 25 hours (20,000 iterations * 3 years). The higher number of simulations required to find authorised crop allocations for scenario B1 in comparison to scenario A1, also demonstrated that spatially aggregating crops, which had required separation distances, did tend to lead to successful coexistence. A higher number of replicates and iteration maximum would have provided a more complete evaluation. As noted in Section 8.7.1.2 (Spatial patterns of the rotation with GM oilseed rape, p.132), the results could be biased by the higher number of fields available to grow GM oilseed rape in scenario A in comparison with scenario B. Therefore, more locations of GM oilseed rape could be tested by the model during the iteration processes in scenario A than B. The evaluation of the impact of this difference in available fields would require further investigation.

An example of crop allocations for each scenario is presented in Figure 8.9 and Figure 8.10. The presented crop allocations were chosen as they had overall penalties close to the median of their scenario groups.

Figure 8.9: Example of crop allocation for a scenario A1, (seed: 3197, year 0)

Figure 8.10: Example of crop allocation for a scenario B1 (seed: 27115, year 0)

To identify the fields, which had their crops successfully altered to meet the constraints, the overall penalties to fields table (OvFP table) was available in the output log file of the LandSFACTS program. The average penalties to fields for all successful simulations of the scenarios (average of 10 simulations for scenario A1 and of 3 simulations for scenario B1) are presented in Figure 8.11 and Figure 8.12. For both scenarios, nearly all fields had to have their crop altered at least once to meet the temporal constraints, yearly crop proportions or separation distance. The fields with the highest penalties on average (the darkest colour) were all with the GM rotation. This was due to the crop priority set up, i.e. if a GM and conventional crops were too close to each other, the GM crop was altered. Therefore, fields with GM crops had more chance of having their crop altered.

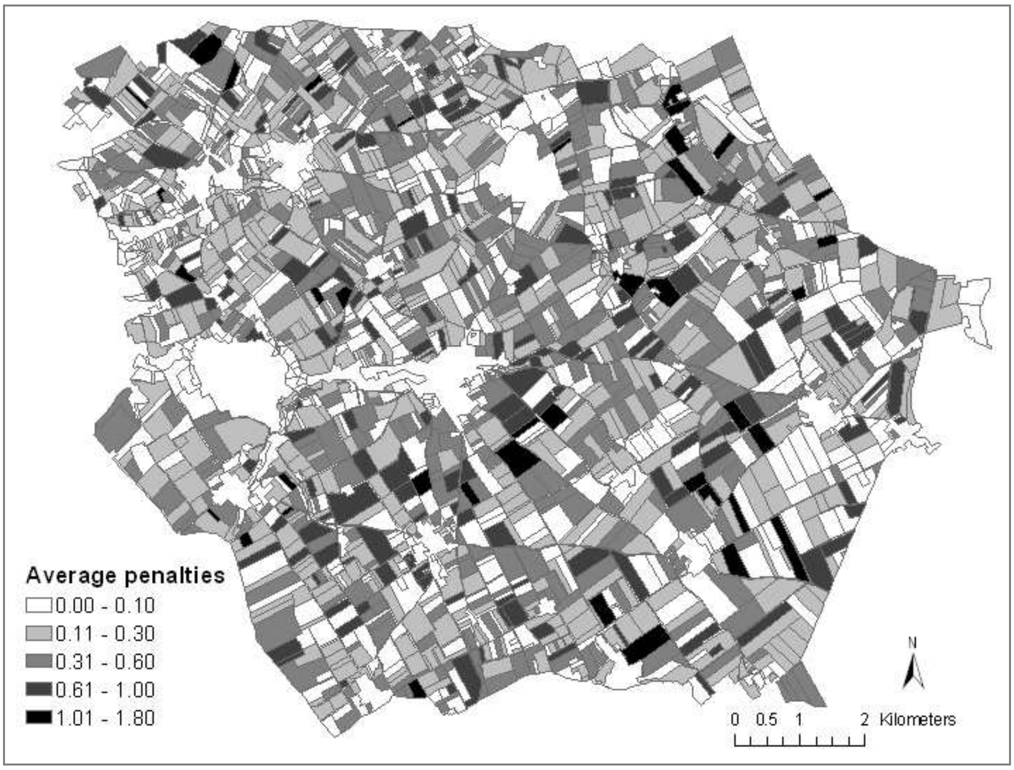


Figure 8.11: Average penalties per fields for scenarios A1 (10 simulations)

Figure 8.12: Average penalties per fields for successful scenarios B1 (3 simulations)

Due to the low number of successful replicates for those scenarios (10 for scenario A, and only 3 for scenario B), only tentative conclusions were deduced from the average penalties to fields. Higher number of simulations would provide a better identification of the most problematic fields (more statistically significant), and therefore would allow the investigation of the characteristics of those fields. For example, a specific crop rotation, field's shape, or number of neighbours could be common denominators for most problematic fields. Identifying the origin or main causes that increased the difficulty of finding a successful crop allocation, would provide an insight into how coexistence measures would be most beneficial, and identifying possible pitfalls (e.g. fields with a specific shape could increase the difficulty of finding a crop allocation meeting specific constraints).

8.7.3 Conclusions

The case study presented in this Section 8.7 did not aim at providing a complete evaluation of the proposed scenarios, but to provide a full example of the use of the LandSFACTS model. By studying the above scenarios, it could be concluded that the spatial patterns of crops and the constraint of separation distances did influence the difficulty of obtaining crop allocation. Spatially aggregated crops, in comparison with spatially regular crops, facilitated the search of authorised crop allocations for the studied scenarios. However to validate those conclusions, further studies would be required to investigate the correlation between the flexibility in crop allocation (e.g. more fields available to grow GM oilseed rape) and the difficulty of finding an authorised crop allocation. Further scenario testing would also be needed to evaluate the exact interactions between separation distances, spatial patterns and yearly crop proportions, on the difficulty of finding crop allocation. The identification of why some configurations of parameters hindered the ease of finding crop allocations, would provide some insight into how to best set up coexistence measures adapted to specific landscapes. The scenarios used here only considered relatively small separation distances (50m) for testing purposes. Analyses with larger distance, i.e. hundreds of meters and even several kilometres, would also be required for real coexistence scenarios.

The testing of the above scenarios also demonstrated the use of the LandSFACTS model. More particularly the usefulness of controlling spatial patterns of the crop rotations was exemplified.

8.8 Recommendations on model use

Through the conceptual model assessment, sensitivity analysis and the scenarios, several recommendations for the use of the model were identified. The recommendations are ordered in the following sections: (i) general recommendations, (ii) recommendations on designing scenarios (landscape, constraints, and multiple runs), and (iii) recommendations on setting up model parameters.

At first as a general recommendation, the model should only be used within its stated purpose as in Section 7.1.1: Aim of the model (p.93). More particularly the model only provides crop allocations that meet the user specifications of crop spatio-temporal arrangements. Therefore, the user is responsible for the inputs provided to the software, in terms of their agronomic and socio-economic adequacy and their relevance to the scenarios studied. In particular, the extent of the simulated landscape should be adequate for the overall study purposes. For example, modelling gene flow of crops might require a landscape as small as two fields to study small-scale flows, or up to more than 2,000 fields (c. 10km^2) for a larger-scale flow study. In addition, using the model requires a thorough understanding of the constraints affecting the studied landscape, and expert knowledge on the study landscape should be sought for realistic scenarios. To optimise further the use of the model, a good understanding of the model structure is recommended. For example, as the model only considers one crop per field per year, agricultural systems with complex crop successions within individual years cannot be easily simulated. Two solutions could circumvent this issue, (i) every intra-annual crop succession could be considered as one crop, or (ii) the model time step (a year) could be reallocated to be a smaller time step (e.g. a month or 10 days). Again it should be noted that the model does not attempt to forecast or predict real crop allocation and therefore it must not be used for this purpose.

Secondly, to design the scenario the three following points should be considered: (i) the landscape itself, (ii) the choice of the constraints, (iii) the relevance of multiple runs. The spatial delimitation of the simulated landscape might influence the crop allocation of the fields close to the boundary, particularly since spatial constraints between crops (separation distances) must be respected. To limit this “edge effect”, the simulated landscape should be spatially extended to include the surrounding fields by at least twice the largest separation distance. If such information is not available, the outer fields of the landscape (twice the separation distance), should be considered as having potentially flawed crop allocation. The model and in particular the spatial statistical test

(Section 5.1.3.2: General spatial trend (E analysis), p.57), were designed and tested on a limited set of landscapes. Therefore, the use of the model on widely different mosaics of fields (with, for example, a wide range of fields areas) or fields shapes (with, for example a large proportion of long thin fields) should be done cautiously, and such results should be carefully checked.

The constraints to be imposed on the landscape should be chosen very carefully. Whereas a large number of constraints and highly restrictive constraints might be useful to reproduce the complexity of an existing landscape, they increase the difficulty of finding an authorised crop allocation. Moreover, such levels of complexity are not necessary or relevant for all scenarios testing. When setting up the constraints, the number of fields within the landscape should be considered, as small number of fields will limit the potential location(s) of crops and thus their potential areas (e.g. reaching exactly 20% of wheat in a landscape with 10 fields might only be possible with specific areas for each field). Furthermore, the coherence between all the constraints should be checked to prevent the case that no authorised crop allocation exists (which must lead to unsuccessful simulations). Incoherence between crop rotations, temporal constraints, and yearly crop proportions should be investigated with particular care.

Depending upon the aim of the scenario, the focus may often be on obtaining one or more authorised crop allocations or on comparing scenarios (e.g. on the impact of different separation distances on the possible crop proportions). In the first case, running the simulation only once for every scenario might be sufficient. However, for the latter case, the obtained crop allocation might be less important than the difficulty of obtaining it (i.e. penalties). As noted in the sensitivity analyses (Section 8.6, p.120), running the model with identical inputs will provide for each run a unique crop allocation and index value of the difficulty of obtaining it. Therefore to obtain an accurate estimation of the difficulty of obtaining crop allocations for a specific scenario, it is recommended to base conclusions on as many runs of the model as possible. A minimum of 10 runs is suggested as a rule of thumb to meet minimal statistical requirements, and this minimum number should be drastically increased for landscapes with larger numbers of fields. This analysis on multiple runs of the model, rather than single runs, should provide a better overview of the difficulty of the scenario for crop allocation.

Thirdly, the choice of the simulation parameters (iteration options) in relation to the aims of the scenario, and the expected difficulty of the scenario is crucial. As defined in

Section 7.2.8: Iteration options (p.102), the four iteration options to find authorised crop allocation have different processes to improve crop allocation, i.e. no optimisation using new random choices (Option 1), optimisation using new random choice (Option 2.1), crop group (Option 2.2) or universal crop (Option 2.3). The iteration options must be chosen carefully by the user, because they will influence the crop allocation obtained, particularly for option 2.2 and 2.3 as they do not follow the rules set up in the crop rotations. If simulation time is not an issue, high maximum iterations for option 1 and 2.1 provide the advantage of respecting the crop rotations. However option 2.2 and 2.3 provide the possibility of improving the crop allocation more quickly, but crops outside of the dedicated crop rotation of the fields might be used. As a general guideline, the maximum number of iterations (all optimisation options together) should, at the very least, double the expected number of fields with problematic crop allocation, or equal the number of fields simulated, whichever is the greater. Precise recommendations are not possible, as they depend upon unforeseen interactions between the landscape and the constraints simulated. It is recommended to follow those guidelines for the first run of the simulation and then adapt the maximum number of iterations depending upon the difficulty of reaching an authorised allocation (if no valid allocation is found, the maximum should be increased). If optimisation options are used (option 2.1, 2.2, 2.3), simulated annealing values should be used to avoid the optimisation process being blocked at a local minimum as reported in Section 8.6.3 (Simulated annealing, p.122). The value should be set initially to 0.01 times the maximum number of iterations (Section 8.6.3: Simulated annealing, p.122). When comparing scenarios of crops allocation, the same iteration parameters should be used, except if the scenarios investigate the impact of iteration parameters themselves (e.g. when investigating the influence of changing problematic crop by a crop with the same function in a rotation, using crop groups).

8.9 Conclusions on model assessment

In this chapter, the LandSFACTS model was assessed to provide an evaluation of the adequacy of the model to meet its stated purposes. The review of the conceptual model highlighted several shortcomings, which could be the focus of future improvements. These included the limitation of one crop per field, fixed boundaries of fields over time, constant crop rotations and crop patterns only imposed at the landscape scale. If those restrictions are clearly communicated to potential users, the conceptual model can be considered as adapted to its purposes.

The code of the model (three programs) was verified, and is deemed reliable for its normal conditions of use. However, it should be noted that code verification can never be exhaustive thus hidden errors might still be present.

The sensitivity analyses investigated the impact of model parameters on the difficulty of obtaining authorised crop allocation. The studied model parameters were the iterations options including simulated annealing values and the impact of separation distances. The analyses on iteration options demonstrated the efficiency of the optimisation algorithm (iteration option 2.1) over the non-optimisation one (iteration option 1). The simulated annealing values used for the optimisation algorithm were reported to increase the probability of finding an authorised crop allocation and, when found, the difficulty and variation in overall penalties and time required. Therefore the choice of simulated annealing value is important for efficiently improving crop allocations.

The sensitivity analysis on separation distances showed an increased difficulty of obtaining authorised crop allocation with increasing separation distances. The relationship was either linear or exponential depending upon the datasets. It is proposed that these differences were due to differences in the number of fields, crop proportions, and fields sizes and shape. Further analyses are required to determine the impact of their interactions. A complete sensitivity analysis was not carried out, due to time constraints.

To complement the assessment of the model, complex scenarios on the impact of spatial patterns of crops were tested. The study reported that if crops constrained by separation distances were spatially aggregated by using their crop rotations, the difficulty of finding authorised crop allocation was significantly lower than in the case of a regular pattern. The scenarios also exemplified that the crop allocations obtained met the specified conditions, and that the reports on the difficulty of obtaining the allocations were helpful in differentiating between coexistence scenarios. Further scenario testing would be useful to investigate the interactions between spatial patterns, spatial constraints with larger separation distances, proportions of the rotations, and yearly crop proportions.

Through the assessment of the LandSFACTS model, three main recommendations were highlighted. Firstly, the model should only be used within its stated purpose, i.e. to allocate crops to fields that meet user specified crop spatio-temporal arrangement. Moreover, to use successfully the model, the user should have a thorough understanding of the cropping system within its study area, the internal structure of the model, and in particular the iteration parameters. Secondly, multiple runs of the model are required to obtain a better accuracy on the difficulty of obtaining authorised crop allocation. And thirdly, the iteration options should be carefully chosen and be adapted to the aim of the studies.

As no flaws were detected during the assessment of the model, it can be said that the LandSFACTS model appears valid for the stated specifications set out by the potential users (Chapter 2). However only a limited set of assessments were carried out and further investigation would be required to fully validate the model, such as a full code verification and a sensitivity analysis on all the variables of the model.

9 Discussions and Conclusions

The previous chapters presented the work carried out to meet the aim of the thesis, i.e. providing a tool to support scenario building of crop arrangement within the context of GM coexistence. In this final chapter, the following points are discussed: (i) an overview on how the thesis aim was reached by fulfilling the objectives, (ii) the major thesis outputs and their advantages, (iii) discussions on the levels of use of the LandSFACTS model and how the model supports coexistence scenarios research, and (iv) examples on how best to enhance this support.

9.1 Meeting the thesis objectives

This thesis aims to provide a research framework for building up scenarios of crop arrangement within the context of GM coexistence, through the design of the LandSFACTS model. To reach this aim, the thesis was centred around three objectives as presented in Chapter 1:

(i) Objective 1, the examination of the origin and characteristics of spatio-temporal arrangement of crops;

(ii) Objective 2, the design of the LandSFACTS model; and

(iii) Objective 3, the assessment of the LandSFACTS model for its stated purpose.

The success in meeting these three objectives within this thesis is presented in the following section.

9.1.1 Objective 1: origin and characteristics of spatio-temporal arrangements

The first objective aimed to determine the origin and characteristics of spatio-temporal arrangements of crops in the literature and in real landscapes. This objective was met through the work presented in Chapter 3 and 5. The origins and existing measurements of crop arrangements as described in the published literature are summarised in Chapter 3. The study highlighted the complexity of the farmers decision process leading to crop allocation and notably the high number of constraints considered, e.g. environmental, agronomic, economic, and policy constraints. The review showed that no statistical analyses were currently available to characterise crop patterns, and therefore specifically designed statistical analyses would need to be integrated within the LandSFACTS model. In Chapter 5, five new statistical analyses of crop spatial and temporal patterns were developed and used on a real landscape to determine characteristics of crop patterns. In general, these statistical analyses successfully quantified the crop patterns in the real landscape. Two pattern tests, one spatial ("E analysis") and one temporal ("Crop temporal variability compared to random simulations") were particularly relevant for integration within the LandSFACTS model, due to their ability to quantify patterns, their broad applicability and ease of use. Overall the patterns detected indicated a strong scale (farmer and landscape) and crop dependency, and a significant spatio-temporal aggregation of crops (i.e. spatially close fields with similar crops). This work satisfactorily achieved the first objective. Having developed and evaluated a range of tools for describing crop arrangements, these tools were then used to design the LandSFACTS model as required by Objective 2.

9.1.2 Objective 2: design of the LandSFACTS model

Objective 2 aimed to design the LandSFACTS model of crop arrangement with its components and processes. The first step was the definition of the model specification. This was carried out in collaboration with end-users, agronomic researchers working with gene flow models on GM coexistence scenarios. As reported in Chapter 2, the model had to be a research tool, which allocated crops to fields to meet user-defined crop arrangements, by using an empirical and statistical modelling approach.

The modelling approach chosen for LandSFACTS model was based upon: (i) the knowledge on the origin and characteristics of crop arrangement (Chapter 3, Objective 1), and (ii) the review of existing models relevant to LandSFACTS aim (Chapter 3, Objective 2). As presented in Chapter 4, the approach was centred around (i) the stochastic modelling of crop rotations, (ii) the spatial and temporal constraints, which rule crop arrangement (Objective 1), and (iii) the spatial and temporal patterns of crops, which are statistical analyses controlling the general trend of crop patterns (tests derived from Chapter 5, Objective 1).

The use of matrices was the chosen method for representing the rotations as probabilities of transitions from one crop to another one. The stochastic modelling of crop rotations as transition matrices is reported in Chapter 6.

The full description of the LandSFACTS model is presented in Chapter 7. The model is divided into two steps: (i) the setting up of the inputs, and (ii) the allocation of the crops to fields ("CropAllocation" program). In the first step, the following are defined: crops, fields boundaries, spatial extent of crops, a rotation for each field (possibility of imposing spatial patterns using "RotationFields" program), the initial crops in fields (possibility of imposing temporal patterns using "InitialCrops" program), spatial constraints (separation distances between crops), temporal constraints (return period of crops, forbidden crop sequences), yearly crop proportions, and model and iteration parameters (to control the behaviour of the model to find authorised crop allocation).

The crop allocation to fields is carried out year by year, by using a linear programming methodology based on optimisation and simulated annealing processes. The model generates, as outputs, the authorised crop allocations and indices representing on the difficulty of obtaining a valid allocation. The three console programs of the LandSFACTS model were embedded into the LandSFACTS software to facilitate the use of the model (SIGMEA, 2005). By completing the LandSFACTS model, Objective 2 is achieved.

9.1.3 Objective 3: assessment of the LandSFACTS model

Objective 3 aimed to provide an assessment of the LandSFACTS model, and is presented in Chapter 8. The first step was to define the aim of the assessment and its coverage. The full assessment of the LandSFACTS model (e.g. sensitivity analyses for all the model variables, complete code verification) was outside of the scope of the thesis due to time constraints. Thus the assessment concentrated on evaluating the adequacy of the model for its stated purposes, by (i) reviewing the conceptual model, (ii) verifying code, (iii) analysing the sensitivity to iteration parameters and separation distances, and (iv) using case study scenarios.

The investigation of the conceptual model highlighted some important restrictions of the model. This included the consideration of only one crop per field, an inability to alter the field boundaries and the crop rotations over the years, and ability to only control the crop patterns at the landscape scale. Overcoming those limitations would greatly benefit the model, however if they are clearly identified and communicated to potential users, the conceptual model can be considered as valid within those restrictions.

The code verification did not identify any invalidating flaws; however, as with the verification of any other program, hidden errors might still be present. The sensitivity analyses provided further insight on how to set up the iteration parameters, and in particular, the influence of the simulated annealing values for obtaining a higher probability of finding authorised crop allocations, while limiting the difficulty of obtaining them and the time required. Increasing separation distances between crops was found to increase the difficulty of obtaining authorised crop allocation. The exact type of correlation (linear or exponential) seemed dependent upon the crop proportions and the fields' sizes and shapes.

The study of scenarios of the impact of crop spatial patterns on the difficulty of finding valid crop allocations exemplified the relevance and usefulness of the model. The analyses carried out were limited to relatively small separation distances (50m) for testing purposes, and analyses with larger distance would also be required for real coexistence scenarios. After the assessment carried out on the LandSFACTS model, it seems to fulfil the specifications set out by the potential users (Chapter 2). However only a limited set of assessments were carried out, and further investigation would be required to fully validate the model. The third and last objective of this thesis was thus completed, by providing an assessment of the LandSFACTS model.

9.1.4 Conclusion: from objectives to aims

The three objectives of the thesis were completed, i.e. the LandSFACTS model (Objective 2) was built by incorporating spatio-temporal arrangements as highlighted in Objective 1, and the model was assessed as fit to meet its purpose (Objective 3). Therefore the aim of the thesis to provide a research framework for building up scenarios of crop arrangement within the context of GM coexistence, through the design of the LandSFACTS model can be considered as achieved.

9.2 Thesis major outputs and their advantages

In the following sections, ways of using the LandSFACTS model to support scenarios testing within the GM coexistence context are discussed. By accomplishing the thesis objectives, three major and stand-alone outputs were achieved.

9.2.1 Statistical analyses

The first output is the statistical analyses developed in this thesis (Castellazzi *et al.*, 2007b). They have their own added value independently from the LandSFACTS model. The statistical tools allow crop patterns in landscapes to be characterised statistically; no similar tools are available in the published literature. Statistical characterisation of crop patterns could be useful within the context of studies focusing on the impact of cropping systems, such as bird ecology studies. The spatial test on general trend of crop patterns

(E analysis) could even be used on any categorical data linked to discrete features (e.g. polygons or points).

9.2.2 Mathematical representation of crop rotations

The second output is the mathematical representation of crop rotations as transition matrices (Castellazzi *et al.*, 2008). Transition matrices provide a unique way of representing crop rotations as a stochastic process in a flexible mathematical format. Fixed and flexible crop rotations can then be equally handled within mathematical models. Long-term proportions of the crops can also be calculated by using the properties of transition matrices. A further advantage of this method is the simplicity of the approach, which is usually easily understood by users without requiring wide mathematical background. The transition matrix representing crop rotations could be useful for any models requiring the mathematical integration of crop rotations.

9.2.3 Model on crop allocation: LandSFACTS

The third output is the LandSFACTS model itself (Castellazzi *et al.*, 2007a). The model provides crop allocations, meeting user-defined specifications of crop arrangements, along with an evaluation of the difficulty of obtaining the allocation. The modelling approach is mainly statistical (in the control of crop rotation and crop patterns) and empirical (constraints), by modelling directly the crop arrangements instead of the mechanistic process leading to it (i.e. farmer decision making). The advantage of the model is its limited number of inputs, which are all easily available. Multiple tools are available to control the crop arrangements and using them all is not compulsory. The user can thus choose and control the constraints on crop arrangements to meet their needs. Another advantage of the LandSFACTS model is its structured and referenced process to allocate crops to fields. Each step of the process can be traced, recorded, and most importantly can be justified. The use of the model, instead of a manual process for allocating crops to fields, provides the user with the possibility of obtaining many crop allocations with the crop arrangement (i.e. running the model many times with the same inputs), and also with the possibility of comparing scenarios with slightly different crop arrangements (i.e. running the model by altering one input at a time). The model processes were deliberately kept simple, in order to provide a tool whose processes were easily and quickly understood by potential users. To facilitate the use of the model, the LandSFACTS software provides user-friendly interfaces and help files detailing the model inputs, processes and outputs interpretations.

9.3 Supporting coexistence scenarios

The aim of this thesis was to provide a framework to support the creation of scenarios of crop arrangement within the context of coexistence of GM crops with conventional and organic related crops. In this section, the levels of use of the models are described, before examples relevant to coexistence scenarios are provided.

9.3.1 LandSFACTS levels of use

The LandSFACTS model can be used at three different levels of complexity depending upon the aim of the study (Figure 9.1). At the simplest level, the model is useful to fill up an incomplete dataset to provide a single crop allocation. For example, if only the main crop rotations are known on a given landscape, without any information of their spatial location among fields, the LandSFACTS model can provide a crop for each field for each desired year. At the second level, the model can be used to build up scenarios of crop allocations where, for the same scenario, many unique crop allocations are provided. Multiple allocations with the same characteristics of crop arrangement, as provided by LandSFACTS, are indispensable to differentiate between the impact of crop arrangement (i.e. general patterns, coexistence measures, or any constraints) and specific crop allocation (i.e. location of each crop). At the highest level of complexity, the LandSFACTS model is useful to compare scenarios of crop allocations, by using the difficulty indexes of obtaining a crop allocation, i.e. penalties to fields.

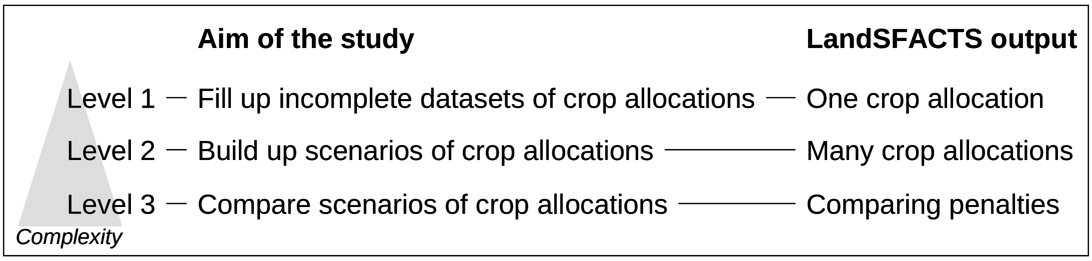


Figure 9.1: Three different levels of complexity in the use of LandSFACTS model in terms of aims and outputs.

The multiple tools of the LandSFACTS model to control crop arrangements, i.e. crop rotations, spatial patterns of rotations, temporal patterns of crops, spatial and temporal constraints and yearly crop proportions, are not all compulsory, i.e. depending upon the aim of the project and available information on the study area, different combinations of inputs can be used. The model can provide a crop to every field, with the minimum input of a rotation for all the fields.

The choice of the inputs must be adapted to the aim of the study. For example testing the interactions between field shapes and separation distances between crops would not require the same type of inputs as testing the feasibility of growing a GM crop in a specific real landscape with specific coexistence rules. The cropping system of the first scenario would probably be extremely simplified, whereas for the second scenario it would need to be as close as possible to the agronomic characteristics of the landscape studied.

The three levels of use of the LandSFACTS model and the flexibility in the corresponding input choices (in the context of GM coexistence) are discussed below.

9.3.2 Examples of LandSFACTS uses within coexistence context

With its three levels of use, the LandSFACTS model on its own, or linked with gene flow models, can provide useful support to evaluate coexistence scenarios of GM crop and related non GM crops. Potential studies are the investigation of (i) the relationship between agricultural landscape characteristics and risks of contamination, (ii) the impact of coexistence measures on the risk of contamination, (iii) the "physical feasibility" of coexistence measures, and (iv) the economic cost of coexistence scenarios. Each of them is further exemplified below.

(i) Risks of contamination from GM crops to non GM related crops are dependent upon the arrangement of the crops in the landscape (Bateman, 1947a). Therefore understanding the relationship between crop arrangement and risks, would provide the knowledge to evaluate the risks for a given agricultural landscape, without having to carry out a full study using detailed surveys and gene flow models. To understand this relationship, gene flow models would have to be run with multiple crop arrangements controlled through the LandSFACTS model. Examples of crop arrangements to test are different crop rotations, crops spatial patterns, or crop temporal patterns.

(ii) The identification of the most appropriate coexistence measures to control the risks of contamination from GM crops to non GM related crops is fundamental to avoid uncontrolled contamination. For example, appropriate separation distances between crops are still being investigated (Sanvido *et al.*, 2007). In order to evaluate the cumulative risks at the landscape scale, it is necessary to consider coherent crop arrangements over several years. The LandSFACTS model can generate temporally coherent crop arrangements respecting coexistence measures, which can then be tested with gene flow models. Using this technique, determinant thresholds in the coexistence measures could be uncovered, e.g. optimal separation distances between crops. Moreover, as the LandSFACTS model can provide many crop allocations, based on the same crop arrangement rules, the effect of coexistence measures (i.e. potential policy) can be separated from the impact of a specific configuration of the crop. Conclusions drawn from such studies would strengthen the adequacy and credibility of coexistence measures.

(iii) Coexistence measures aim to limit the risk of contamination from GM crops to non-GM related crops under a specific threshold. However the coexistence measures may or may not be "physically feasible" on any given landscape, due to specific field configurations, cropping systems, or crop proportions (refer to Chapter 8), (Perry, 2002). For example, if an agricultural landscape currently has 30% of its area with conventional oilseed rape, would any farmer be able to grow GM oilseed rape if a minimum of 500m is required between GM and conventional crops? The "physical feasibility" can be investigated with the LandSFACTS software, by imposing coexistence measures on the current cropping system of a given landscape, and identify whether an arrangement is indeed possible – including a measure of the difficulty in finding a fit (using the penalties). Only if coexistence measures on a given landscape are "physically feasible" would it be necessary to evaluate the economic and farm management feasibility in the given landscape. The LandSFACTS model would provide a first screening of the feasibility of coexistence measures in a given landscape.

(iv) The economic cost of coexistence measures is an important factor affecting the farmer's decision on growing GM crops (Bock *et al.*, 2002; Messean *et al.*, 2006). For example the cost can be linked to (i) involuntarily contaminated non-GM crops which have to be sold as GM, or (ii) decreased area for GM cultivation if an internal edge with a conventional variety of the crop has to be grown within each GM field. The LandSFACTS model provides support to examine both cases. In the first case, the number of non-GM crops within a specific distance of any GM fields could be identified. In the second case, the edge area of the GM fields, which have to be cultivated with conventional crop, could be estimated. An evaluation of the economical cost in a given landscape would then be facilitated.

As presented in the above sections, the LandSFACTS model could provide useful support for studies investigating GM coexistence measures. The conclusion drawn by researchers from gene flow models would be strengthened. Thus, more informed advice could then be provided to decision makers on the feasibility and efficiency of coexistence measures for GM cultivation.

9.4 Possible enhancements for coexistence scenarios

The LandSFACTS model could be enhanced to provide further support for scenarios within the GM coexistence context, in the two following ways: (i) by enhancing the modelling of crop arrangement, and (ii) by providing new tools specifically designed to answer coexistence scenario needs.

9.4.1 Enhancing modelling of crops arrangement

The main role of LandSFACTS in supporting the testing of coexistence models is its ability to model crop arrangement. Therefore by enhancing LandSFACTS modelling of cropping systems, the usefulness of the model will be increased and widened. Four main enhancements of crop arrangement modelling are: (i) the control of crop patterns over a range of scales, (ii) the annual control of the spatial pattern of crops, (iii) the evolution of crop rotation over time, and (iv) decision on crop allocation based upon the field's status. They are presented below.

(i) Crop patterns are different depending upon the scale of study, as reported in Chapter 5. For example, farmers may want to group their crop in time and space, in order to facilitate more efficient farm management (e.g. less traffic and less regulation); as a result the area of crops grown every year may fluctuate at the farm scale. However at the landscape scale (many farms), the proportion of the main crops may be relatively constant over the years (low temporal variation). Currently, the LandSFACTS model provides support for modelling crop patterns at only one scale, i.e. the landscape scale if the whole landscape is modelled, or the farm scale if only the fields of a farm are modelled. The LandSFACTS model would be enhanced by allowing the user to control crop patterns over at least two scales simultaneously and independently, i.e. at both the farm and landscape scale; furthermore the control over a third level, e.g. group of farmer (cooperative) scale, would be advantageous as cooperatives might influence the crop managements (particularly relevant in French agricultural landscapes).

(ii) Crops may be spatially aggregated (aggregated spatial pattern) because of environmental characteristics (close fields have the same characteristics), or because of a farmer's decision to group the management of fields that are close spatially. Currently, within the LandSFACTS model, the spatial pattern of crops is being controlled by limiting the available crops for each field, or by requiring a specific spatial pattern on the crop rotations, or both. Those controls are imposed only once at the start of the simulation. Moreover, the latter control is not efficient, if the crop to be controlled appears in all the rotations. This control also becomes inefficient in the case of very flexible crop rotations; to address this problem in the LandSFACTS model, the spatial pattern of crops would need to be controlled directly for every year of crop allocation. This could be done by checking the spatial pattern of crops for every year of simulation.

(iii) Crop rotations can evolve over time to adapt to new environmental conditions, such as climate change, or new market conditions. Rotations might either be adapted by altering a few crops, or they might be replaced by a totally different rotation adapted to the new circumstances. Currently, the LandSFACTS model only considers one unchangeable crop rotation for each field. Several crop rotations can be considered within the model, if they are presented as one main rotation controlling the probabilities of switching between sub-rotations (refer to Chapter 6). It would be useful to provide the user with support in creating this ‘main rotation’ from the individual sub-rotations (normal conventional rotations). Furthermore, by linking the probabilities of transition from crop to crop, to external variables, i.e. temperatures or rainfall levels, variations within crop rotations could be controlled. LandSFACTS would then be able to simulate scenarios of crop allocation in evolving landscapes, i.e. landscapes responding to external factors such as climate change or introduction of new crops.

(iv) The farmer’s choice of growing a crop in a field is dependent upon the status of the field, e.g. nutrient availability, organic matter content, water balance, pests invasions, and weeds growth. The status of fields are different for each field and each year. Within the LandSFACTS model the decisions on crop succession are solely directed by crop rotations, which might be altered to meet constraints on crop arrangement. The probabilities of transition from one crop to another could be regulated by “fields status” variables, which would be updated yearly for each individual field, depending at least upon the crops that are grown. This enhancement would provide more responsive landscapes to agronomic or environmental variables at the field level. However, this type of enhancement would complicate the model processes, and more importantly would introduce mechanistic processes within the model, and thereby contradict the main modelling approach of LandSFACTS, which is statistical and empirical. Therefore, any mechanistic addition to LandSFACTS model should be integrated as an “extension” outside of the core of the model, in order to avoid the confusion between the different modelling approaches.

The enhancement of the modelling of cropping system within LandSFACTS would provide more adapted support to scenarios of GM coexistence. Other specific tools are presented in the next section.

9.4.2 New tools specifically designed for coexistence scenarios

For the specific purpose of coexistence scenarios, the LandSFACTS model provides control over the separation distances between crops, crop temporal successions, and yearly crop proportions. However, there are further tools, which could be created to provide even more support to coexistence scenarios. These include (i) the ability to control field shapes, (ii) the handling of field margins, (iii) discard areas in fields (buffers), and (iv) the control of farms spatial distribution and the integration of silos for mixing grains. They are presented below.

(i) Field shape, size and orientation change from one region to another. However, in each case, they usually define homogeneous areas where crops may be grown. Field boundaries are mostly constant over the years, although they can be merged or subdivided when farm lands are reorganised. The sizes, shapes and number of neighbouring fields are important factors for the risks of GM contamination (Bateman, 1947a; Klein *et al.*, 2006) and the feasibility of coexistence rules (Damgaard and Kjellsson, 2005). Currently the LandSFACTS model only considers fixed field boundaries over the years for a given landscape. Being able to alter field boundaries would increase the versatility of the model. More importantly being able to control the field shape, size, number of neighbours would provide a powerful tool to investigate the interactions between the field characteristics and the risks of GM contamination. To achieve this, the model would need to create new landscapes with user defined characteristics. This research area is currently being investigated by the team at the MIA unit (Département de Mathématiques et Informatique Appliquées) in INRA-Jouy-en-Josas (Adamczyk *et al.*, 2006). The integration within the LandSFACTS model of their newly developed statistical analyses characterising field shapes and models of field mosaics creation, would greatly enhance the potential uses of the model for coexistence scenario investigation.

(ii) Margins around fields or along roads are fertile areas of land on which feral plants (GM offspring from crops grown in neighbouring fields in the previous years) may grow. This is particularly the case with oilseed rape, whose ferals can regularly be spotted on roads margins (Charters *et al.*, 1999; Garnier *et al.*, 2006). The management of the feral population is crucial in limiting the risks of GM genes spreading through the landscape (Cresswell and Osborne, 2004; Ellstrand, 2003; Stewart *et al.*, 2003). Currently the LandSFACTS model does not consider margins at all. By providing the creation and handling of margins and their management, the model would provide support to scenario testing of the impact of adding, removing margins, or altering their management in any given landscape.

(iii) Discarding crops on the edges of GM fields is being considered as an additional coexistence measure (Damgaard and Kjellsson, 2005). For example, a buffer of 5m of conventional oilseed rape may need to be grown around GM oilseed rape, and would be downgraded as GM contaminated oilseed rape. This coexistence measure would reduce the risk of pollen contamination to neighbouring oilseed rape fields. Currently the LandSFACTS model does not consider such measures. The model could be enhanced by allowing the user to set up a “discard buffers” for specific crops. The addition of such a tool within the LandSFACTS model would increase its support to coexistence scenarios by controlling another type of coexistence measure.

(iv) A further coexistence measure to keep the GM contamination below a specific threshold, is to dilute potentially contaminated grains with non contaminated grains (Ceddia *et al.*, 2007), i.e. by physically mixing grains within silos. This dilution could occur at the scale of “groups of farms”, i.e. cooperatives. The grouping of farmers into cooperatives would impact upon the potential contamination at the silo level. The ability to alter the spatial distribution of groups of farmers in the LandSFACTS model, would provide support in investigating the impact of the spatial distribution of cooperatives on the potential contamination at the silo level.

The enhancements reported within the above sections would increase the support provided by the LandSFACTS model to scenarios of GM coexistence, by enhancing the modelling of crop arrangement (e.g. crop patterns over multiple scales, evolving cropping systems over time), and by providing specifically designed tools to address coexistence issues (e.g. altering field boundaries, margins and discard areas, and farm’s spatial grouping). The LandSFACTS model was implemented to allow such enhancements; therefore, their integration within the model should not be a problem. However, by increasing the complexity of the model, the model might become more difficult to use, and could confuse potential users. Such enhancements should thus be clearly identifiable within the model and should be developed as optional tools within the LandSFACTS model.

9.5 Conclusion

This thesis meets the aim of providing a framework for simulating crop arrangements for GM coexistence scenarios. The three objectives, i.e. the origin and characteristics of spatio-temporal arrangement, the design and assessment of the LandSFACTS model have been met and the main output, the LandSFACTS model, has been deemed valid for its stated purpose. To design the model, two further tools were created. These were statistical analyses on crops spatio-temporal patterns and mathematical representations of crop rotations. This chapter has also provided examples of the support provided by the LandSFACTS model to coexistence scenarios, including support for the investigation of the impact of crop arrangement on risks of GM contamination, of potential coexistence measures on the level of risks, of the physical feasibility of coexistence measures, and of the economic cost of coexistence measures. This chapter provided suggestions for possible improvements on the control of crop arrangements (e.g. enhanced control of crop patterns and rotations) and the integration of further tools adapted for coexistence scenarios needs (e.g. control of fields boundaries, fields margins, discard areas, spatial distribution of farmers and silos).

The LandSFACTS model has been designed to be used within the context of coexistence of GM crops in European agricultural landscapes. However, the LandSFACTS model is not restricted to this context. The model could be used on any agricultural landscape worldwide as long as only one crop is grown per field every year. The model could be useful to any model requiring land-uses as an input, in order to control and statistically characterise their crop arrangement inputs. For example, there are examples where models of soil erosion, organic farming, plant disease, or animal-plant interactions could benefit from the investigation of their sensitivity to crop arrangements. The model is also potentially useful to investigate the introduction of crops other than GM. A recent new environmental concern is linked to the use of bio-energy crops, such as willow, within classical agronomic landscapes. The impact of such alteration of land uses (i.e. from arable land to “forest”) on the wildlife is currently under focus (RELU-Biomass, 2007). The LandSFACTS model would be useful to investigate potential scenarios of bio-energy crop adoption.

10 References:

Adamczyk, K., Kiêu, K., Monod, H. and Stoica, R. 2006. Tessellation model for agricultural landscape. 6th French-Danisch Workshop on Spatial Statistics and Image Analysis in Biology. 23-25 mai 2006, Skagen.

Angevin, F., Klein, E., Choimet, C., Gauffreteau, A., Lavigne, C., Messean, A. and Meynard, J.M. 2007. Modelling the impacts of cropping systems and climate on maize cross pollination in agricultural landscapes: The MAPOD model. European Journal of Agronomy, Accepted.

Angevin, F., Klein, E., Choimet, C., Meynard, J.M., de Rouw, A. and Sohbi, Y. 2001. Modélisation des effets des systèmes de culture et du climat sur les pollinisations croisées chez le maïs : le modèle MAPOD. Rapport du groupe 3 du programme de recherche "Pertinence économique et faisabilité d'une filière 'sans utilisation d'OGM' ". INRA - FNSEA, 56 pages.

Aubry, C., Papy, F. and A., C. 1998. Modelling Decision-Making Processes for Annual Crop Management. Agricultural Systems, 56 (1), pp.45-65.

Audsley, E., De Baets, A. and Rosato, P. 1999. Farm scale modelling. In: M. D. A. Rounsevell, ed. Spatial modelling of the response and adaptation of soils and land use systems to climate change - an integrated model to predict European land use (IMPEL). pp.40-64.

Bachinger, J. and Zander, P. 2006. ROTOR, a tool for generating and evaluating crop rotations for organic farming systems. European Journal of Agronomy, 26 (2), pp.130-143.

Baker, W.L. 2001. The r.le Programs: A set of GRASS programs for the quantitative analysis of landscape structure. [Internet], Available from: <http://grass.itc.it/gdp/landscape/r_le_manual5.pdf> [Accessed 21/05/2006].

Baker, W.L. and Cai, Y.M. 1992. The R Le-Programs for Multiscale Analysis of Landscape Structure Using the Grass Geographical Information-System. Landscape Ecology, 7 (4), pp.291-302.

Barr, C.J. and Gillespie, M.K. 2000. Estimating hedgerow length and pattern characteristics in Great Britain using Countryside Survey data. Journal of Environmental Management, 60 (1), pp.23-32.

Bateman, A.J. 1947a. Contamination in Seed Crops .3. Relation with Isolation Distance. Heredity, 1 (3), pp.303-336.

Bateman, A.J. 1947b. Contamination of Seed Crops .1. Insect Pollination. Journal of Genetics, 48 (2), pp.257-275.

Bateman, A.J. 1947c. Contamination of Seed Crops .2. Wind Pollination. Heredity, 1 (2), pp.235-246.

Baudry, J., Bunce, R.G.H. and Burel, F. 2000. Hedgerows: An international perspective on their origin, function and management. Journal of Environmental Management, 60 (1), pp.7-22.

Baudry, J., Burel, F., Aviron, S., Martin, M., Ouin, A., Pain, G. and Thenail, C. 2003. Temporal variability of connectivity in agricultural landscapes: do farming activities help? Landscape Ecology, 18 (3), pp.303-314.

Berzsenyi, Z., Gyorffy, B. and Lap, D. 2000. Effect of crop rotation and fertilisation on maize and wheat yields and yield stability in a long-term experiment. European Journal of Agronomy, 13 (2-3), pp.225-244.

Bock, A.-K., Lheureux, K., Libeau-Dulos, M., Nilsagård, H. and Rodriguez-Cerezo, E. 2002. Scenarios for co-existence of genetically modified, conventional and organic crops in European agriculture, Report to European Commission, 133 pages.

Bond, W., Davies, G. and Turner, R. 2006. The biology and non-chemical control of Volunteer Cereals. [Internet], Henry Doubleday Research Association, Ryton, Available from: <http://www.gardenorganic.org.uk/organicweeds/downloads/volunteer%20cereals.pdf> [Accessed 21/05/07].

Brady, N. and Weil, R. 2002. The Nature and Properties of Soils. Upper Saddle River, New Jersey, 960 pages.

Brunt, L. 1999. An arbitrage model of crop rotation in 18th Century England. In: L. Brunt, J. Humphries, A. Offer and D. Stead, eds. Discussion papers in Economic and Social History. University of Oxford, p.25.

Castellazzi, M., Wood, G.A., Burgess, P.J., Morris, J., Conrad, K.F. and Perry, J.N. 2008. A systematic representation of crop rotations. Agricultural Systems, 97, pp.26-33.

Castellazzi, M.S., Matthews, J., Wood, G.A., Burgess, P.J., Conrad, K.F. and Perry, J.N. 2007a. LandSFACTS: Software for Spatio-temporal Allocation of Crops to Fields. Proceedings of 5th Annual Conference of the European Federation of IT in Agriculture, Glasgow, UK.

Castellazzi, M.S., Perry, J.N., Colbach, N., Monod, H., Adamczyk, K., Viaud, V. and Conrad, K.F. 2007b. New measures and tests of temporal and spatial pattern of crops in agricultural landscapes. Agriculture Ecosystems & Environment, 118 (1-4), pp.339-349.

Ceddia, M.G., Bartlett, M. and Perrings, C. 2007. Landscape gene flow, coexistence and threshold effect: The case of genetically modified herbicide tolerant oilseed rape (Brassica napus). Ecological Modelling, 205 (1-2), pp.169-180.

Champion, G.T., May, M.J., Bennett, S., Brooks, D.R., Clark, S.J., Danieis, R.E., Firbank, L.G., Haughton, A.J., Hawes, C., Heard, M.S., Perry, J.N., Randle, Z., Rossall, M.J., Rothery, P., Skellern, M.P., Scott, R.J., Squire, G.R. and Thomas, M.R. 2003. Crop management and agronomic context of the Farm Scale Evaluations of genetically modified herbicide-tolerant crops. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 358 (1439), pp.1801-1818.

Charters, Y.M., Robertson, A. and Squire, G.R. 1999. Investigation of feral oilseed rape populations. Report to DEFRA. No 12, 93 pages.

Clarke, J., Moss, S. and Orson, J. 2000. The future for grass weed management in the UK. Pesticide Outlook, April 2000, pp.59-63.

Cliff, A.D. and Ord, J.K. 1981. Spatial processes: Models and Applications. London, Pion, 266 pages.

Colbach, N., Clermont-Dauphin, C. and Meynard, J.M. 2001a. GENESYS: a model of the influence of cropping system on gene escape from herbicide tolerant rapeseed crops to rape volunteers - I. Temporal evolution of a population of rapeseed volunteers in a field. Agriculture Ecosystems & Environment, 83 (3), pp.235-253.

Colbach, N., Clermont-Dauphin, C. and Meynard, J.M. 2001b. GENESYS: a model of the influence of cropping system on gene escape from herbicide tolerant rapeseed crops to rape volunteers - II. Genetic exchanges among volunteer and cropped populations in a small region. Agriculture Ecosystems & Environment, 83 (3), pp.255-270.

Colbach, N., Fargue, A., Sausse, C. and Angevin, F. 2005. Evaluation and use of a spatio-temporal model of cropping system effects on gene escape from transgenic oilseed rape varieties - Example of the GENESYS model applied to three co-existing herbicide tolerance transgenes. European Journal of Agronomy, 22 (4), pp.417-440.

Cox, D.R. and Miller, H.D. 1965. The Theory of Stochastic Processes. London, UK, Chapman & Hall, 398 pages.

Cresswell, J.E. and Osborne, J.L. 2004. The effect of patch size and separation on bumblebee foraging in oilseed rape: implications for gene flow. Journal of Applied Ecology, 41 (3), pp.539-546.

Cullinan, V.I. and Thomas, J.M. 1992. A Comparison of Quantitative Methods for Examining Landscape Pattern and Scale. Landscape Ecology, 7 (3), pp.211-227.

Damgaard, C. and Kjellsson, G. 2005. Gene flow of oilseed rape (Brassica napus) according to isolation distance and buffer zone. Agriculture, Ecosystems & Environment, 108 (4), pp.291-301.

DEFRA 2004a. Economic, social and environmental implications of EU sugar regime reform. Rural Business Unit of University of Cambridge, and the Royal agricultural college, 108 pages.

DEFRA 2004b. Single Payment Scheme: Information for farmers and growers in England. [Internet], Available from:

<http://www.defra.gov.uk/farm/capreform/pubs/update050208.htm> [Accessed 21/05/07].

DEFRA 2005a. Environmental Stewardship. [Internet], Available from: <http://www.defra.gov.uk/erdp/schemes/es/default.htm> [Accessed 07/09/2005].

DEFRA 2005b. Single Payment Scheme: Cross Compliance, Handbook for England. [Internet], Available from: <http://www.defra.gov.uk/farm/capreform/pubs/ccompl.htm> [Accessed 21/05/07].

DEFRA 2006. Consultation on proposals for managing the coexistence of GM, conventional and organic crops. [Internet], Available from: <http://www.defra.gov.uk/corporate/consult/food-farming/gmnongm-coexist/consultdoc.pdf> [Accessed 21/05/07].

Detlefesen, N. 2004. Crop rotation modelling. Proceedings of the EWDA-04 European workshop for decision problems in agriculture and natural resources, Silsoe Research Institute, England, p.5-14.

Dogliotti, S., Rossing, W.A.H. and van Ittersum, M.K. 2003. ROTAT, a tool for systematically generating crop rotations. European Journal of Agronomy, 19 (2), pp.239-250.

Donatelli, M., Stockle, C.O., Ceotto, E. and Rinaldi, M. 1997. Evaluation of CropSyst for cropping systems at two locations of northern and southern Italy. European Journal of Agronomy, 6, pp.35-45.

Dungan, J.L., Perry, J.N., Dale, M.R.T., Legendre, P., Citron-Pousty, S., Fortin, M.J., Jakomulska, A., Miriti, M. and Rosenberg, M.S. 2002. A balanced view of scale in spatial statistical analysis. Ecography, 25 (5), pp.626-640.

Eastham, K. and Sweet, J. 2002. Genetically modified organisms (GMOs): The significance of gene flow through pollen transfer. Environment issue report, No 28, Report to European Environment Agency, 75 pages.

Ellstrand, N.C. 2003. Current knowledge of gene flow in plants: implications for transgene flow. Philosophical Transactions of the Royal Society B: Biological Sciences, 358 (1434), pp.1163-1170.

European Union 1966. Council Directive 66/402/EEC of 14 June 1966 on the marketing of cereal seed.

Firbank, L.G., Heard, M.S., Woiwod, I.P., Hawes, C., Haughton, A.J., Champion, G.T., Scott, R.J., Hill, M.O., Dewar, A.M., Squire, G.R., May, M.J., Brooks, D.R., Bohan, D.A., Daniels, R.E., Osborne, J.L., Roy, D.B., Black, H.I.J., Rothery, P. and Perry, J.N. 2003. An introduction to the Farm-Scale Evaluations of genetically modified herbicide-tolerant crops. Journal of Applied Ecology, 40 (1), pp.2-16.

Flamm, R.O. and Turner, M.G. 1994. Alternative Model Formulations for a Stochastic Simulation of Landscape Change. Landscape Ecology, 9 (1), pp.37-46.

Fortin, M.-J., Dale, M.R.T. and ver Hoef, J. 2002. Spatial analysis in ecology. in Encyclopedia of Environmetrics. A. El-Shaarawi and W. W. Piegrorsch, eds. Chichester, John Wiley & Sons, Ltd. 4, pp.2051-2058.

Fu, B. and Chen, L. 2000. Agricultural landscape spatial pattern analysis in the semi-arid hill area of the Loess Plateau, China. Journal of Arid Environments, 44, pp.291-303.

Garnier, A., Deville, A. and Lecomte, J. 2006. Stochastic modelling of feral plant populations with seed immigration and road verge management. Ecological Modelling, 197 (3-4), pp.373-382.

Goense, D., Hofstee, J.W. and vanBergeijk, J. 1996. An information model to describe systems for spatially variable field operations. Computers and Electronics in Agriculture, 14 (2-3), pp.197-214.

Gustafson, E.J. 1998. Quantifying Landscape Spatial Pattern: What is the State of the Art? Ecosystems (1), pp.143-156.

Heard, M.S., Rothery, P., Perry, J.N. and Firbank, L.G. 2005. Predicting longer-term changes in weed populations under GMHT crop management. Weed Research, 45 (5), pp.331-338.

Hoover, S.V. and Perry, R.F. 1989. Simulation. Reading, Addison-Wesley, 696 pages.

Hunt, J.C.R., Higson, H.L., Walklate, P.J. and Sweet, J.B. 2001. Modelling the dispersion and cross-fertilisation of pollen from GM crops, Report to DEFRA, 39 pages.

Joannon, A. 2004. Coordination spatial des systemes de culture pour la maitrise de processus ecologic. Ph.D Thesis, INRA Paris-Grignon. 393 pages.

Jones, F.G.W. and Perry, J.N. 1978. Modeling Populations of Cyst-Nematodes (Nematoda Heteroderidae). Journal of Applied Ecology, 15 (2), pp.349-371.

Kabos, S. and Csillag, F. 2002. The analysis of spatial association on a regular lattice by join-count statistics without the assumption of first-order homogeneity. Computers & Geosciences, 28, pp.901-910.

Kirkegaard, J., Christen, O., Krupinsky, J. and Layzell, D. 2004. Break crop benefits in temperate wheat production. Proceeding for the 4th International Crop Science Congress, Queensland, Australia.

Klein, E.K., Lavigne, C., Picault, H., Michel, R. and Gouyon, P.H. 2006. Pollen dispersal of oilseed rape: estimation of the dispersal function and effects of field dimension. Journal of Applied Ecology. 43, 141-151 pages.

Klein Haneveld, W.K. and Stegeman, A.W. 2005. Crop succession requirements in agricultural production planning. European Journal of Operational Research, 166, pp.406-429.

Klöcking, B., Strobl, B., Knoblauch, S., Maier, U., Pfutzner, B. and Gericke, A. 2003. Development and allocation of land-use scenarios in agriculture for hydrological impact studies. Physics and Chemistry of the Earth, Parts A/B/C, 28 (33-36), pp.1311-1321.

Kropff, M.J., Bouma, J. and Jones, J.W. 2001. Systems approaches for the design of sustainable agro-ecosystems. Agricultural Systems, 70 (2-3), pp.369-393.

Lavigne, C., Angevin, F., Adamczyk, K., Benoit, M., Colbach, N., Klein, E., Le Bail, M., Le Ber, F., Mari, J.-F., Messean, A., Meynard, J.-M., Mignolet, C., Monod, H. and Sausse, C. 2002-2006. Modélisation de la dispersion de transgènes à l'échelle de paysages agricoles. Rapport final du projet soutenu dans le cadre de l'ACI « Impact des OGM dans les agro-écosystèmes ».

Lawes, Sir Bennet, J., Gilbert and Sir Henry, J. 1895. The Rothamsted Experiments. Edinburgh, Blackwood.

Le Ber, F., Dury, A., Chevrier, V. and Benoit, M. 1998. Simuler l'organisation spatiale d'un territoire agricole: differentes approches. In Actes du Colloque SMAGET, Clermont-Ferrand, CEMAGREF, p.249-259.

Le Coeur, D., Baudry, J., Burel, F. and Thenail, C. 2002. Why and how we should study field boundary biodiversity in an agrarian landscape context. Agriculture Ecosystems & Environment, 89 (1-2), pp.23-40.

Lockie, S., Mead, A., Vanclay, F. and Butler, B. 1995. Factors Encouraging the Adoption of More Sustainable Crop Rotations in South-East Australia - Profit, Sustainability, Risk and Stability. Journal of Sustainable Agriculture, 6 (1), pp.61-79.

Lutman, P.J.W., Berry, K., Payne, R.W., Simpson, E., Sweet, J.B., Champion, G.T., May, M.J., Wightman, P., Walker, K. and Lainsbury, M. 2005. Persistence of seeds from crops of conventional and herbicide tolerant oilseed rape (Brassica napus). Proceedings of the Royal Society B-Biological Sciences, 272 (1575), pp.1909-1915.

Matthews, R. 2006. The People and Landscape Model (PALM): Towards full integration of human decision-making and biophysical simulation models. Ecological Modelling, 194 (4), pp.329-343.

Maxime, F., Nicoletti, J., Leroy, P. and Papy, F. 1996. Donner de la souplesse au choix d'assolement par des rotations-cadres. Colloque Aide à la décision et choix de stratégies dans les entreprises agricoles, Laon, Inra et Conseil régional de Picardie, p.85-99.

McGarigal, K. 2002. Landscape pattern metrics. in Encyclopedia of Environmentrics. A. El-Shaarawi and W. W. Piegrorsch, eds. Sussex, England, John Wiley and Sons. 2, pp.1135-1142.

McGarigal, K. and Marks, B.J. 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. USDA For. Serv. Gen. Tech. Rep. PNW-351.

Messean, A., Angevin, F., Gómez-Barbero, M., Menrad, K. and Rodríguez-Cerezo, E. 2006. New case studies on the coexistence of GM and non-GM crops in European agriculture. Report to European Commission, EUR 22102 EN, 112 pages.

Mignolet, C., Schott, C. and Benoit, M. 2004. Spatial dynamics of agricultural practices on a basin territory: a retrospective study to implement models simulating nitrate flow. The case of the Sein basin. Agronomie, 24, pp.219-236.

National Statistics 2005. Scottish agricultural census summary sheets by geographic area: June 2004. [Internet], National Statistics publication, Available from: <http://www.statistics.gov.uk/> [Accessed 21/05/07].

Neelamkavil, R. 1988. Computer Simulation and Modelling. Chichester, John Wiley and Sons, 307 pages.

O'Neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., DeAngelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christenson, S.W., Dale, V.H. and Graham, R.L. 1988. Indices of landscape pattern. Landscape Ecology, 1 (3), pp.153-162.

Oreskes, N., Shraderfrechette, K. and Belitz, K. 1994. Verification, Validation, and Confirmation of Numerical-Models in the Earth-Sciences. Science, 263 (5147), pp.641-646.

Orson, J. 2005. Discussion on crop rotation in South-England. J. N. Perry and M. S. Castellazzi. personnal communication.

Overmars, K.P., de Koning, G.H.J. and Veldkamp, A. 2003. Spatial autocorrelation in multi-scale land use models. Ecological Modelling, 164, pp.257-270.

Oxley, T., Jeffrey, P. and Lemon, M. 2002. Policy Relevant Modelling: Relationships Between Land Use, and Farmer Decision Processes. Integrated Assessment, 3 (1), pp.30-49.

Oxley, T., McIntosh, B.S., Winder, N., Mulligan, M. and Engelen, G. 2004. Integrated modelling and decision-support tools: a Mediterranean example. Environmental Modelling and Software, 19, pp.999-1010.

Papy, F., Attonaty, J.M., Laporte, C. and Soler, L.G. 1988. Work Organization Simulation as a Basis for Farm-Management Advice (Equipment and Manpower, Levels against Climatic Variability). Agricultural Systems, 27 (4), pp.295-314.

Parker, D.C. and Meretsky, V. 2004. Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. Agriculture Ecosystems and Environment, 101, pp.233-250.

Perry, J.N. 2002. Sensitive dependencies and separation distances for genetically modified herbicide-tolerant crops. Proceedings of the Royal Society of London Series B-Biological Sciences, 269 (1496), pp.1173-1176.

Pontius, J.R.G., Huffaker, D. and Denman, K. 2004. Useful techniques of validation for spatially explicit land-change models. Ecological Modelling, 179 (4), pp.445-461.

Ramsay, G., Thompson, C.E. and Squire, G.R. 2003. Quantifying landscape-scale geneflow in oilseed rape. RG0216, Report to DEFRA, 48 pages.

Refsgaard, J.C. and Henriksen, H.J. 2004. Modelling guidelines - terminology and guiding principles. Advances in Water Resources, 27 (1), pp.71-82.

Rellier, J.P. and Marcaillou, J.C. 1990. Modeles de raisonnement en conduite de culture et consequences pour les systemes d'aide a la decision. Agronomie, 10 (6), pp.487-498.

RELU-Biomass 2007. Social, economic and environmental implications of increasing rural land use under energy crops. [Internet], Available from: <http://www.relu-biomass.org.uk/index.php> [Accessed 20/08/2007].

Remmel, T.K., Csillag, F., Mitchell, S.W. and Boots, B. 2002. Empirical distributions of landscape pattern indices as functions of classified image composition and spatial structure. Symposium on Geospatial Theory, Processing and Applications, Ottawa.

Rounsevell, M.D.A. 1999. Spatial modelling of the response and adaptation of soils and land use systems to climate change - an integrated model to predict European land use (IMPEL), Report to Commission of the European Communities, 241 pages.

Rounsevell, M.D.A., Annetts, J.E., Audsley, E., Mayr, T. and Reginster, I. 2003a. Modelling the spatial distribution of agricultural land use at the regional scale. Agriculture, Ecosystems and Environment, 95, pp.465-479.

Rounsevell, M.D.A., Annetts, J.E., Audsley, E., Mayr, T. and Reginster, I. 2003b. Modelling the spatial distribution of agricultural land use at the regional scale. Agriculture Ecosystems & Environment, 95 (2-3), pp.465-479.

Rounsevell, M.D.A., Evans, S.P., Mayr, T.R. and Audsley, E. 1998. Integrating Biophysical and Socio-economic Models for Land Use Studies. Proceedings of the ITC - ISSS Conference on Geo-information for Sustainable Land Management, Enschede, p.368.

Rykiel, E.J. 1996. Testing ecological models: The meaning of validation. Ecological Modelling, 90 (3), pp.229-244.

Sanvido, O., Widmer, F., Winzeler, M., Streit, B., Szerencsits, E. and Bigler, F. 2007. Definition and feasibility of isolation distances for transgenic maize cultivation. Transgenic Research, in press.

Sargent, R.G. 2003. Verification and validation of simulation models. Proceedings of the 36th conference on Winter simulation, Washington D.C., p.37-48.

Sausse, C. 2005. Discussions on crop rotation in study area within SIGMEA project. J. N. Perry and M. S. Castellazzi. personnal communication.

Schlesinger, S., Crosbie, R.E., Gagne, R.E., Innis, G.S., Lalwani, C.S., Loch, J., Sylvester, R.J., Wright, R.D., Kheir, N. and Bartos, D. 1979. Terminology for Model Credibility. Simulation, 32 (3), pp.103-104.

Shannon, C.E. and Weaver, W. 1949. The mathematical theory of communication. Urbana, IL, University of Illinois Press., 55 pages.

SIGMEA 2005. Sustainable Introduction of Genetically Modified Crops into European Agriculture. [Internet], EU-funded, Contract no.: 502981, Available from: <http://sigmea.dyndns.org/> [Accessed 01/08/2005].

Spedding, C.R.W. 1975. The study of agricultural systems. in Study of Agricultural Systems. G. E. Dalton, eds. London, Applied Science Publishers, pp.3-19.

Squire, G.R., Begg, G.S. and Askew, M. 2003. The potential for oilseed rape feral (volunteer) weeds to cause impurities in later oilseed rape crops, Report to DEFRA, RG0114, 27 pages.

Stewart, C.N., Halfhill, M.D. and Warwick, S.I. 2003. Transgene introgression from genetically modified crops to their wild relatives. Nat Rev Genet, 4 (10), pp.806-817.

Stöckle, C.O., Donatelli, M. and Nelson, R. 2003. CropSyst, a cropping systems simulation model. European Journal of Agronomy, 18, pp.289-307.

Swanson, F.J., Kratz, T.K., Caine, N. and Woodmansee, R.G. 1988. Landform Effects on Ecosystem Patterns and Processes. Bioscience, 38 (2), pp.92-98.

Tarim, S.A., Manandhar, S. and Walsh, T. 2006. Stochastic constraint programming: A scenario-based approach. Constraints, 11 (1), pp.53-80.

Thenail, C. and Baudry, J. 2004. Variation of farm spatial land use pattern according to the structure of the hedgerow network (bocage) landscape: a case study in northeast Brittany. Agriculture, Ecosystems and Environment, 101, pp.53-72.

Timmons, A.M., Charters, Y.M., Crawford, J.W., Burn, D., Scott, S.E., Dubbels, S.J., Wilson, N.J., Robertson, A., OBrien, E.T., Squire, G.R. and Wilkinson, M.J. 1996. Risks from transgenic crops. Nature, 380 (6574), pp.487-487.

Tulloch, D.L., Myers, J.R., Hasse, J.E., Parks, P.J. and Lathrop, R.G. 2003. Integrating GIS into farmland preservation policy and decision making. Landscape and Urban Planning, 63 (1), pp.33-48.

Turner, M.G. 1990. Spatial and temporal analysis of landscape patterns. Landscape Ecology, 4 (1), pp.21-30.

Veldkamp, A. and Lambin, E.F. 2001. Predicting land-use change. Agriculture Ecosystems and Environment, 85, pp.1-6.

White, R.G. 1929. Rotation of Crops. London: HMSO, Ministry of Agriculture and Food, 22 pages.

Wibberley, J. 1996. A brief history of rotations, economic considerations and future directions. Aspects of Applied Biology, 47, pp.1-10.

Wilkinson, M.K., Elliott, L.J., Allainguillaume, J., Norris, C., Welters, R., Alexander, M., Cuccato, G., Sweet, J.B., Shaw, M.W. and Mason, D.C. 2005. Compiling a spatially explicit profile for gene flow from GM oilseedrape in the UK. Aspects of Applied Biology, 74, pp.35-45.

Winder, N. 2000. Chapter Sixteen: Simulating crop choice on the Argolid Plain. in The Pathfinder as Historian: an assay of modernism, of the emerging science of cultural ecodynamics and the mathematics of history.eds. Kiruna, Sweden, SMC, pp.155-162.

Winder, N. 2003. Successes and problems when conducting interdisciplinary or transdisciplinary (integrative) research. in Interdisciplinary and Transdisciplinary Landscape studies: Potential and Limitations. B. Tress, A. van der Valk and G. Fry, eds. Wageningen.

Wünsch, J.A. 2004. Integration des contraintes du marche dans la conduite des cultures: effet de la differenciation des produits sur la conduite de la culture de pomme de terre de conservation dans les exploitations agricoles de Picardie. Ph.D Thesis, Institut National de la Recherche Agronomique. 149 pages.

Young, M.W., Brown, L., Dawson, G. and Squire, G.R. 2006. Field pattern and land use in the Fife study area: a GIS dataset. Agroecology Data Archive, SCRI Dundee DD2 5DA, UK.

Appendix A: Existing crop allocation models

Summary of the studies relating to crop allocation models.

	Authors	General aim	Scale of study	Base unit	Initial Inputs	Final outputs	Modelling processes	Mathematical processes
ARABLE based on SFARMOD	(Rounsevell *et al.*, 2003a; Rounsevell *et al.*, 1998)	Evaluate optimum agricultural land use	European scale (model at Farm scale)	Grid (5*5km) 1cell = 1farm	Soil, climate, farmer decision making very detailed (labour, timing…)	% agricultural land use / cells	Farmer decision process (from IMPEL)	Optimisation process Linear programming
Crop generator	(Klöcking *et al.*, 2003)	Create a virtual crop pattern	regional 7,500km^2	Grid 100* 100m; 1 parcel >= 4 cells	% of landuse per homogeneous areas	Crops / cells	Statistical Year / year	Random allocation Crop statistics respected
For STICS model	(Mignolet *et al.*, 2004)	Nitrate model	Water-shed	Homogeneous areas 425km^2	Expert knowledge Agricultural statistics	Main crop rotations	Expert knowledge	Temporal data mining
Crop succession	(Klein Haneveld and Stegeman, 2005)	Crop succession requirement	1 unit (no spatial model)	-	Crop sequences not allowed	Crop sequences	No agronomic / economic consideration	Mathematical approaches Generic multi-year linear programming model Linear constraints in the decision variables
ROTAT	(Dogliotti *et al.*, 2003)	Generate reproducibly crop rotation	Farm (no spatial hetero-geneity)		Not allowed successions, profitability, return period, dates	Possible crop rotations (& classify Bio & physical and properties)	Agronomic filters Farmer-specific constraints & objectives	Linear programming to optimise temporal interactions filters
CropSyst	(Donatelli *et al.*, 1997; Stöckle *et al.*, 2003)	Cropping systems simulation model (effect climate change)	Watersheds	Block of field (same environment + same management) = 1 polygon	Crop rotation template, Soil, crop specificity , crop management, climate biomass production, intakes.	Effects on environment.	Modelling crop growth, soil water, erosion… Daily time step	Deterministic model Event driven model

	Authors	General aim	Scale	Base unit	Initial Inputs	Final outputs	Mod. process	Mathematical processes
Simulated crop rotation	(Joannon, 2004)	Analyse farmer's flexibility to change crop systems	Regional	field	Rotation rules Return period Previous crop	Crop allocation to fields	Spatial & temporal constraints	Not specified Model not formalised
Crop choice	(Oxley *et al.*, 2002; Oxley *et al.*, 2004; Winder, 2000)	Farmer decision making, crop choice model	Regional	Grid, cell = 1ha parcel	Crops; Physical, socio-economic & institution conditions	Crops type Water requirement, water sources…	Farmer decision making Yearly	Decision trees Rule based model
	(Baudry *et al.*, 2003)	Compare connectivity on a real landscape, 7 years crop successions.	Water-shed ~ 5*5km	Polygons	Real landscape, farming systems (topology, major driving forces, classes of landuse, soil types).	Land use / cover classes in fields Hedgerows between fields	Within-farm rules of land allocation Empirical rules	Rules based model.
Crop rotation model	(Detlefesen, 2004)	Network model used for crop rotation	-	-	Crops Hectares in rotation	Average production plan		Network / transportation models
	(Le Ber *et al.*, 1998)	Simulate spatial organisation of agricultural approaches	Regional	1.homogeneous areas; 2. parcels of homogeneous soil, slope, distance	DEM Soil map 2. slope, distance constraints 3. spatial constraints	Crop allocation Number of cows fed.	1. Expert knowledge 3. closer to farmers thinking than 2.	2. multi-agent systems 3. simulated annealing
	(Parker and Meretsky, 2004)	Explore impacts of urban-rural conflicts on land-use patterns	Fictive grid	grid	Urban negative influence on agricultural production Repulsion of urban against urban land.	Location of urban and agricultural land	Macro & micro-scale feedbacks Externalities present Cost/profit	Agent-based model

Appendix B: Digital Appendices (CD)

The digital appendices (CD) provide the programs and example datasets developed within the thesis. For the structure, refer to Figure B.1.

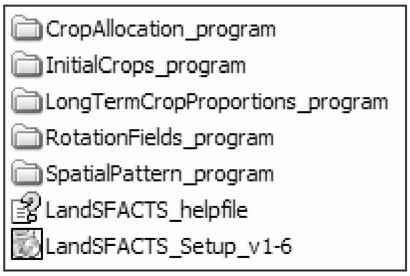


Figure B.1: Screenshot of CD content

- LandSFACTS_Setup_v1-6

This file is the installer of the LandSFACTS software. By double-clicking it, the LandSFACTS software with its adds-on will be installed. The software is provided with a comprehensive helpfile (LandSFACTS_helpfile, directly accessible from LandSFACTS interface), example datasets and projects, and a tutorial. The simulation programs (RotationFields, InitialCrops, and CropAllocation) embedded within the software are not the latest versions.

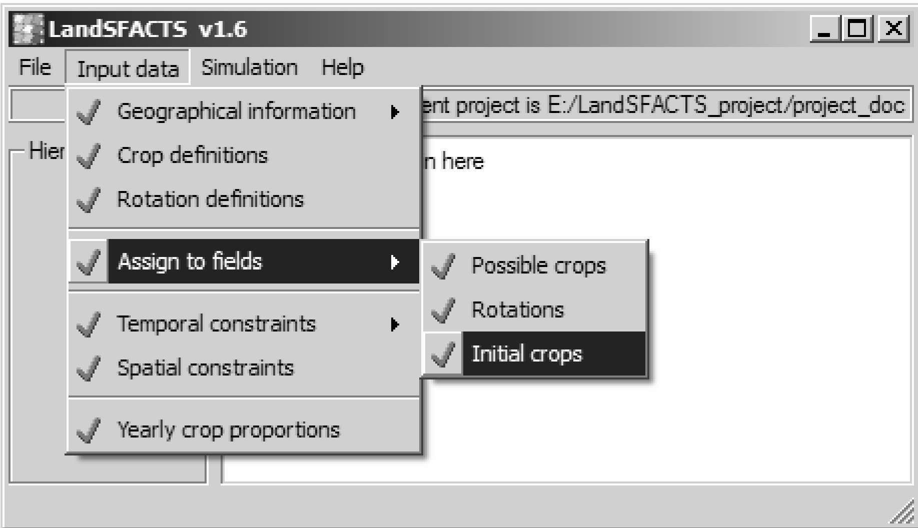


Figure B.2: Screenshot of LandSFACTS software interface

- HelpFile for LandSFACTS software

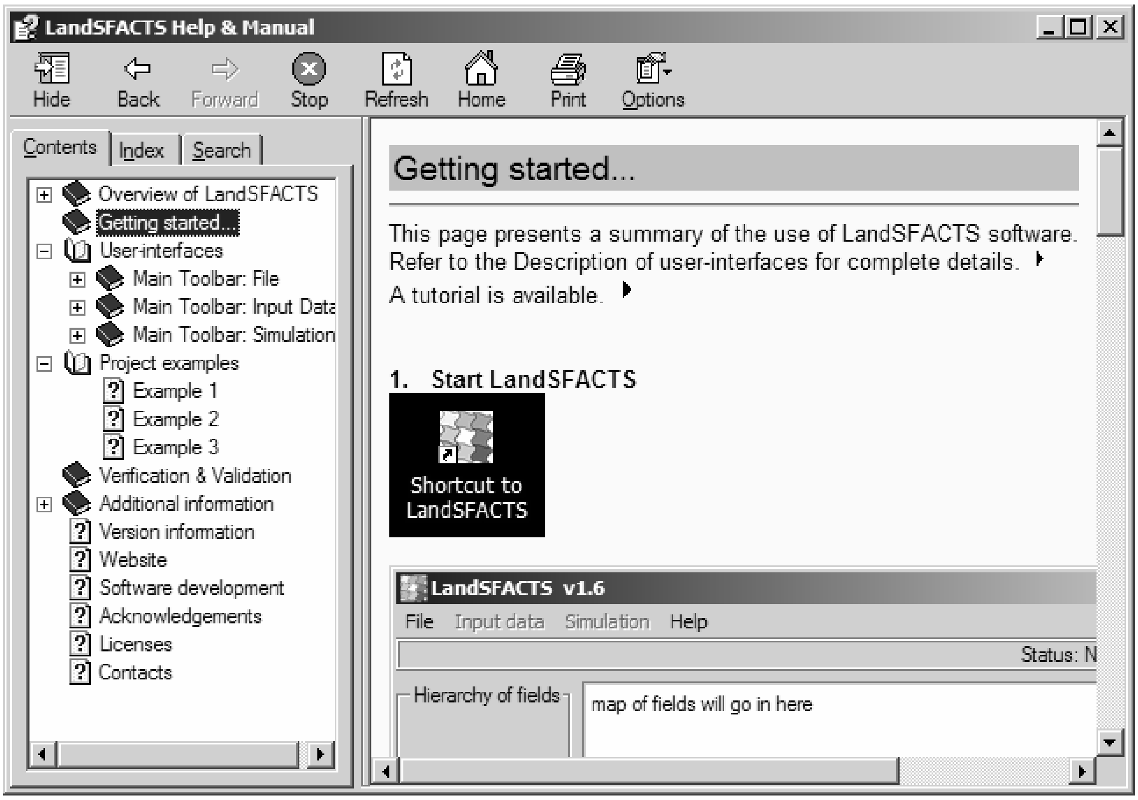


Figure B.3: Screenshot of helpfile

- LandSFACTS C++ stand-alone programs.

Each program folder is composed of a source code folder, a Files folder (inputs and outputs examples), the executable of the program, cf. Figure B.4, 5. Individual help files are provided within each executable, further advices are available in the helpfile for LandSFACTS software.

- RotationFields program
- InitialCrops program
- CropAllocation program
- Long-term crop proportions
- Statistical test of spatial patterns (E analysis)

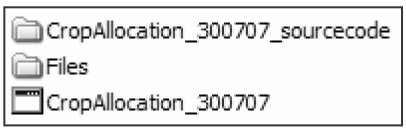


Figure B.4: Screenshot of contents of CropAllocation_program folder

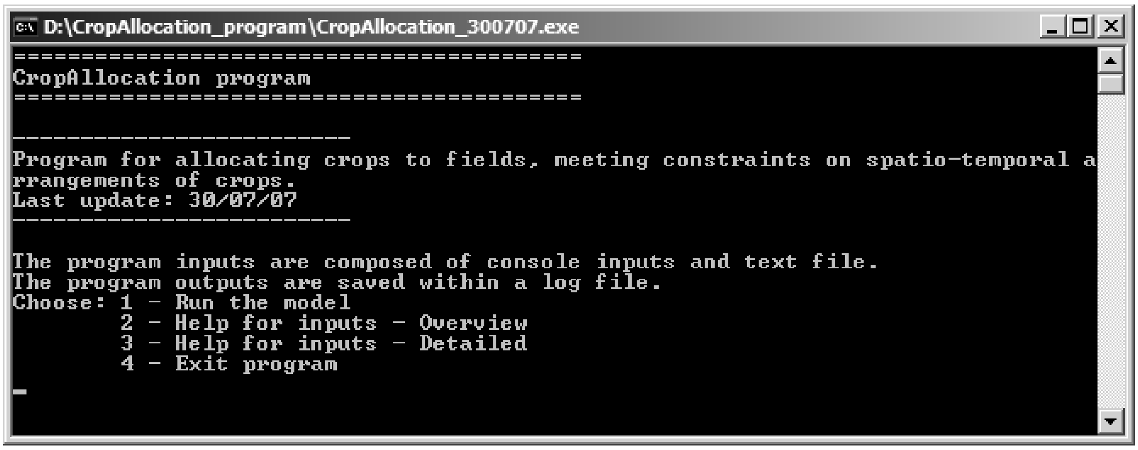
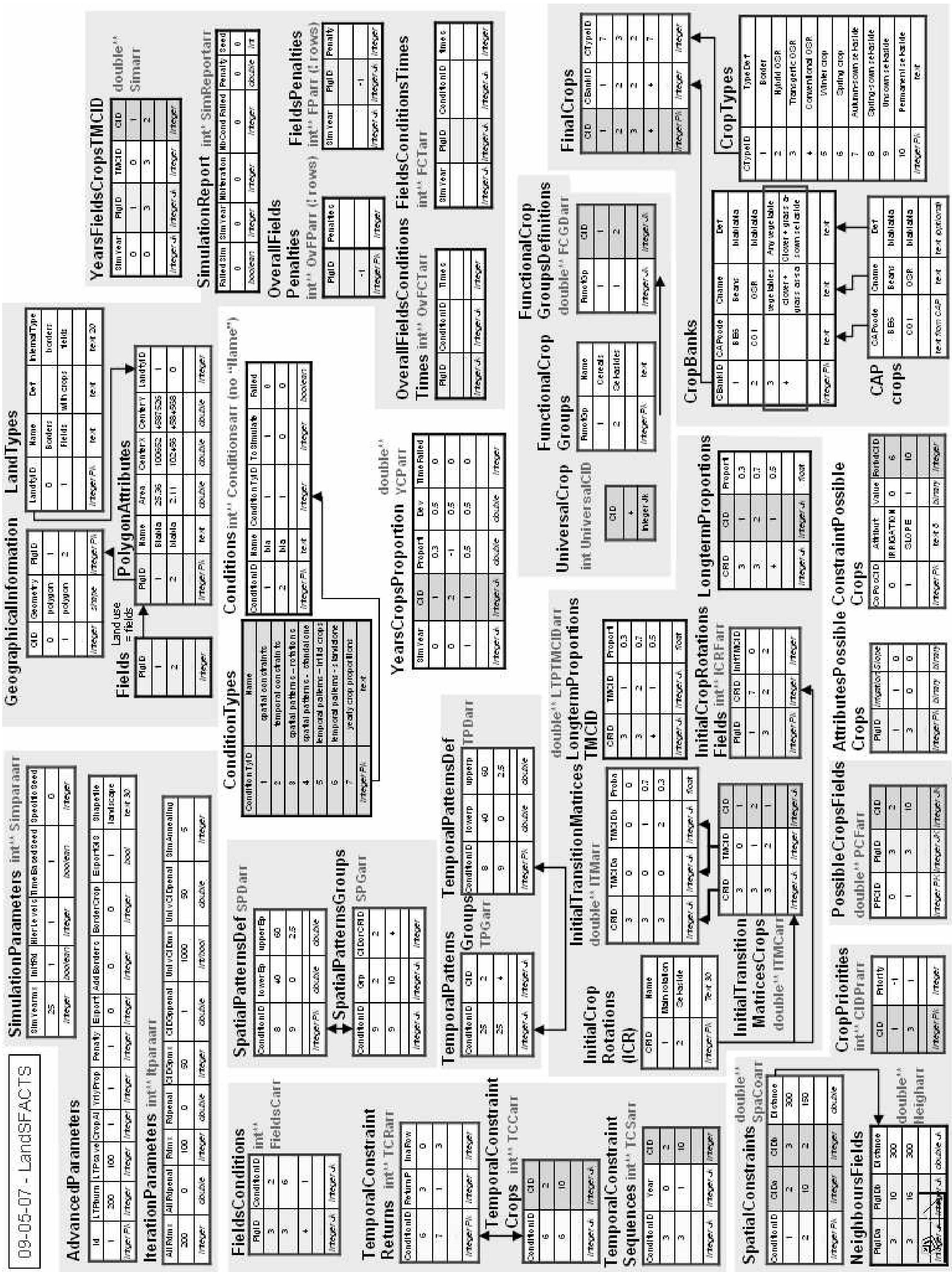


Figure B.5: Screenshot of CropAllocation program

Appendix C: Database structure



Appendix D: Code verification tests

List of general code verification tests:

ID	Verification	Status
1	Crop allocation with a known fixed rotation and initial crop for each field	Ok
2	Crop allocation with a known flexible rotation and initial crop for each field	Ok
3	Detecting spatial constraints	Ok
4	Detecting temporal constraints - return period of crops	Ok
5	Detecting temporal constraints - maximum consecutive years of crops	Ok
6	Detecting temporal constraints - forbidden crop sequences	Ok
7	Checking many temporal constraints at the same time	Ok
8	Reach target yearly crop proportions	Ok
9	Check all types of conditions (temporal, spatial, crop proportions) at the same time	Ok

Appendix E: Assessment of stochastic processes

Random initial crops

A simulation is set up to test the true randomness of random choice of initial crops (crops for first year of simulation in every field), by giving an equal probability between 10 potential initial crops for every field. The simulation is run 1,000 times for only one year. For every simulation, the number of fields with each crop is recorded.

By determining the relative standard deviation of the number of fields with each crop for all the simulations, the randomness of the process is evaluated.

Dataset:

- Beauce dataset
- 1 crop rotation for all fields: equal probabilities of transition between 10 crops
- random initial crops in fields
- no constraints and iteration parameters
- simulation years = 1
- simulation batch = 1000

Results:

The summary of the number of fields for each crop between the 1,000 simulations is reported in Table E.1. As for each field there is an equal probability between the 10 crops (probability of 0.1), and as there is 1,993 fields, for every simulation, each crop should tend towards 199.3 fields. For each crop, the average of number of fields over the 1,000 simulations is close to 199.3. The relative standard deviations between the numbers of fields for each crop over all the simulations are all under 0.007. Therefore the random choice of initial crops between all the crops in the rotation is concluded to be adequate, as it provides an equal probability of obtaining any of the crops for identical probability within the transition matrix. Further tests could be useful, such as by giving non-identical probabilities between the crops in the transition matrix.

Table E.1. Results on random initial crops

crop	sum	average	min	max	stdev	% stdev
1	200344	200.344	156	245	13.50741	0.006742
2	200101	200.101	159	248	13.72017	0.006857
3	199143	199.143	159	246	13.26365	0.00666
4	198537	198.537	162	238	13.07117	0.006584

5	199520	199.52	165	255	13.26188	0.006647
6	199528	199.528	152	245	13.85327	0.006943
7	198845	198.845	158	241	13.41406	0.006746
8	199013	199.013	154	240	13.17081	0.006618
9	198944	198.944	160	244	13.55	0.006811
10	199025	199.025	163	248	13.5164	0.006791
sum	1993000					

The above results were obtained from 1,993,000 values, derived from 1,000 simulations over 1 year crop allocation, on 1,993 fields. The number of initial random choice tested are 1,000 simulation multiplied by 1,993 fields.

Random next crop

A simulation is set up to test the true randomness of random choice of the next crop when using a transition matrix (crops for any year but the first year of simulation), by giving an equal probability in the transition matrix between 10 potential crops for every field. The simulation is run 50 times for 21 years, the first year (initial year) is discarded as it uses the random initial crop instead of the random next crop. For every simulation, the number of fields with each crop is recorded.

By determining the relative standard deviation of the number of fields with each crop for all the simulations, the randomness of the process is evaluated.

Dataset:

- Beauce dataset
- 1 crop rotation for all fields: equal probabilities of transition between 10 crops
- random initial crops in fields
- no constraints and iteration parameters
- simulation years = 21
- simulation batch = 50

Results:

The summary of the number of fields for each crop between the 50 simulations is reported in Table E.2. As for each field there is an equal probability between the 10 crops, and as there are 1,993 fields, the number of each crop for every year of simulation should tend towards 199.3. The table reports that the average value of number of fields for each crop are very close to the 199.3 value. The relative standard deviation between all the simulations of the number of fields for each crop are all lower than 0.007, indicating a low variation from the target value (199.3). Therefore the random choice of the next crop is considered as adequately random for the purpose of LandSFACTS. Further random tests could be carried out to test the properties of the random choice, such as by setting up non-equal probabilities within the transition matrix.

Table E.2. : Results on random next crop

crops	sum of all values	average value	minimum value	maximum value	standard deviation	relative standard deviation
1	199596	199.60	161	248	13.218	0.0066
2	199360	199.36	157	236	13.381	0.0067
3	199514	199.51	158	241	13.572	0.0068
4	199658	199.66	154	242	13.728	0.0069
5	199152	199.15	158	246	13.994	0.0070
6	199957	199.96	144	242	13.046	0.0065
7	198953	198.95	151	245	13.474	0.0068
8	199221	199.22	163	243	13.600	0.0068
9	198700	198.70	152	234	13.131	0.0066
10	198889	198.89	163	238	12.973	0.0065

The above results were derived from 50 simulations over 20 year crop allocation (initial year was not considered), on 1,993 fields.

Random field to alter

A simulation is set up to test the true randomness of the fields chosen to be altered when several fields must have their crop altered. Every field of the dataset is linked with a crop rotation featuring continuous wheat, at the same time the crop sequence wheat-wheat is forbidden. In consequence, one year over two, all the fields, one at a time, must have their wheat crop altered for the universal crop (only option authorised). Iteration by iteration, the simulation records how many times a field does not meet a constraint, therefore the field, which has its crop altered first will have the value 1; the field, which has its crop altered last will have the highest value (388 for this example as it is the total number of fields). Those values are recorded in the table called FieldsConditionsTimes, in the column "Times". By investigating the variations between years and simulation of those "Times" values; the randomness of the fields to alter is evaluated.

Dataset:

- Fife dataset with 388 fields
- 1 crop rotation: continuous wheat
- random initial crops in fields
- 1 forbidden crop sequence for all fields: wheat after wheat
- iteration parameters: only option 2.3. Universal crop, which is “fallow”, maximum iterations: 1,000, penalty to fields = 1.
- simulation years = 200 (as only one year in two is tested for wheat-wheat, there is 100 sample year)
- simulation batch = 50

Results:

The table E.3 reports the statistic summary of the number of “Times” for each field (i) over all years of simulation, (ii) over one year of simulation, (iii) over all years if the order of the fields altered was constant (maximum skewness). The relative standard deviation of all simulations, is two order of magnitude lower than the one from maximum skewness (0.002 instead of 0.15). This difference is also reflected in the maximum and minimum values, which are both closer to the average. The evaluation indicates that the order of the fields to alter seems random, i.e. even distribution between the fields of the “Times” values. This test could be further complemented with an evaluation of the lowest possible skewness, by simulating an even distribution of the order of altered fields.

Table E.3. Results on random field to alter

	all runs	1 year	maximum skewness
number of values	5000	1	5000
sum	377330000	75466	377330000
average	972500	194.5	972500
median	972508	194.5	972500
standard deviation	7607	112	560750
relative standard deviation	0.0020	0.1486	0.1486
max	992682	388	1940000
min	952572	1	5000

Appendix F: Crop rotations for sensitivity analyses

The following tables reports the crop rotation matrices used for the sensitivity analyses in Section 8.6: Sensitivity analyses, p.120.

Table F.1: Transition matrix for permanent grassland rotation (Fife dataset)

		Year n + 1
		Permanent grassland
Year n	Permanent grassland	1

Table F.2: Transition matrix for "all crops" rotation (Fife dataset)

		Year n + 1										
		Ww	Wb	Sb	Wo	So	WOSRc	WOSRgm	SOSRc	P	o	Sa
Year n	Winter wheat	0	0.1	0.2	0	0.2	0	0	0.2	0.1	0.1	0.1
	Winter barley	0	0.1	0	0.4	0	0.4	0	0	0	0.1	0
	Spring barley	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.1	0.09
	Winter oats	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.1	0.09
	Spring oats	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.1	0.09
	Winter OSR conv	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.1	0.09
	Winter OSR GM	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.1	0.09
	Spring OSR conv	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.1	0.09
	Potatoes	0.4	0.25	0.25	0	0	0	0	0	0	0.1	0
	Other	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.1	0.09
	Set-aside	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.1	0.09

Ww = winter wheat; Wb = winter barley; Sb = spring barley; Wo: winter oats; So = sunflower; WOSRc = winter oilseed rape conventional variety; WOSRgm = winter oilseed rape GM variety; SOSRc = spring oilseed rape conventional; P = potatoes; o = other; Sa= set-aside.

Table F.3: Transition matrix for "all crops" rotation (Beauce dataset)

		Year n + 1									
		OSR conv	OSR GM	W	S	M	Sa	P	Wb	o	Sb
Year n	OSR conv	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	OSR GM	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	Wheat	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	Sunflowers	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	Maize	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	Set-aside	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	Peas	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	Winter barley	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	Other cereals	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07
	Spring barley	0.15	0.05	0.3	0.07	0.07	0.07	0.07	0.07	0.08	0.07

OSRconv = oilseed rape conventional variety; OSRGM = oilseed rape GM; W = wheat; S = sunflower; Sa = set-aside; P = peas; Wb = winter barley; o = other cereals; Sb = spring barley.

Appendix G: Results from scenarios

The following tables reports the crop rotation matrices used for the sensitivity analyses in Section 8.7: Scenario testing for a real landscape, p.131.

Table G.1: Unprocessed results of each simulation

	ID	FailedSim	Year	NbIteration	NbCondFailed	Penalty	Seed
A0	1	0	0	3218	4842329	337	13667
	2	0	0	3431	5230842	324	5376
	3	0	0	3130	4759391	350	19291
	4	0	0	4367	6841779	390	10145
	5	0	0	2242	3084814	292	22522
	6	0	0	4445	7023564	376	6055
	7	0	0	4611	7196019	352	26405
	8	**0**	**0**	**4194**	**6508242**	**348**	**20840**
	9	0	0	5394	8464827	407	10777
	10	0	0	4693	7401647	376	13450
B0	1	0	0	5695	9542636	405	9466
	2	0	0	3461	5640678	307	20370
	3	0	0	4706	8139041	388	7732
	4	0	0	3547	5738110	344	3010
	5	0	0	4674	8084697	386	10413
	6	0	0	4087	6909869	354	14464
	7	**0**	**0**	**4093**	**6713520**	**336**	**5160**
	8	0	0	4052	6823957	354	6684
	9	0	0	3964	6684982	341	26025
	10	0	0	3106	4949373	312	12849
A1-50m	1	0	0	11977	22465910	769	15343
	2	0	0	11301	20715924	728	3197
	3	0	0	9014	16004598	641	6583
	4	0	0	11338	20719267	660	23437
	5	0	0	8438	15348210	607	26348
	6	0	0	10798	19551917	712	10023
	7	0	0	9069	16667877	642	23975
	8	0	0	10514	19391363	724	1730
	9	0	0	8678	15835887	622	10192
	10	0	0	7905	14193776	600	4879
B1-50m	1	1	2	18158	35245606	1793	27115
	2	0	0	14666	28096724	675	16455
	3	1	2	21960	42980004	1893	16889
	4	1	2	20102	39426562	1810	16911
	5	1	2	20149	38629261	1837	31101
	6	1	2	19926	38872849	1810	6867
	7	1	2	16918	33045120	1797	31528
	8	0	0	12551	24219992	676	28660
	9	1	2	16918	33045120	1797	31528
	10	0	0	19548	38041588	787	11509

FailedSim = 1 if simulation failed; Year = if simulation has failed, year that failed; NbIteration = number of iteration used for the simulation; NbCondFailed = number of failed condition during the iteration processes; Penalty = overall penalty to field of the simulation; Seed = random seed, unique for each simulation.

Table G.2: Statistical summary of results:

A0

	FailedSim	Year	NbIteration	NbCondFailed	Penalty
average	0	0	3973	6135345	355
median	0	0	4281	6675011	351
min	0	0	2242	3084814	292
max	0	0	5394	8464827	407
stdev	0	0	939	1608003	34

B0

	FailedSim	Year	NbIteration	NbCondFailed	Penalty
average	0	0	4139	6922686	353
median	0	0	4070	6768739	349
min	0	0	3106	4949373	307
max	0	0	5695	9542636	405
stdev	0	0	742	1364015	32

A1-50m

	FailedSim	Year	NbIteration	NbCondFailed	Penalty
average	0	0	9903	18089473	671
median	0	0	9792	18029620	651
min	0	0	7905	14193776	600
max	0	0	11977	22465910	769
stdev	0	0	1438	2807494	58

B1-50m

	FailedSim	Year	NbIteration	NbCondFailed	Penalty
average	0.7	1.4	18090	35160283	1488
median	1	2	18853	36643597	1797
min	0	0	12551	24219992	675
max	1	2	21960	42980004	1893
stdev	0	1	2866	5696071	536

FailedSim = 1 if simulation failed; Year = if simulation has failed, year that failed; NbIteration = number of iteration used for the simulation; NbCondFailed = number of failed condition during the iteration processes; Penalty = overall penalty to field of the simulation; stdev = standard deviation.

Table G.3: Student's t-test

	Tests	NbIteration	NbCondFailed	Penalty
t-value	A0-B0	0.4386	1.1808	0.1699
	A1-B1	7.8010	8.2104	5.0778
	A0-A1	10.9180	11.6840	14.7927
	B0-B1	14.3179	14.6389	7.0856
	B0-B1s	9.8810	10.1058	15.8145
	A1-B1s	4.6503	4.9649	1.5336
Significantly different at p = 0.01?	A0-B0	no	no	no
	A1-B1	yes	yes	yes
	A0-A1	yes	yes	yes
	B0-B1	yes	yes	yes
	B0-B1s	yes	yes	yes
	A1-B1s	yes	yes	no

NbIteration = number of iteration used for the simulation; NbCondFailed = number of failed condition during the iteration processes; Penalty = overall penalty to field of the simulation.

Degree of freedom for Student's t-test: 18, except for B1s (only 11)

Threshold value from t-test table, with p = 0.01: 2.88, except for B1s (3.10)