Deconstructing the complexity of land use and cover classification and land change modelling

by

Marcelo de Castro Chaves Stabile

M.Sc. Agronomy (Texas A&M University)
B.Sc. Agro. Eng. (ESALQ-USP)

A thesis submitted to the University of Sydney
in fulfilment of the requirements for the
degree of Doctor of Philosophy

Faculty of Agriculture, Food and Natural Resources
The University of Sydney
New South Wales, Australia

MMXII
Certificate of originality

I hereby certify that the text of this thesis contains no material that has been accepted as part of the requirements for any other degree or diploma in any University, nor any material previously published or written unless the reference to this material is made.

Marcelo de C. C. Stabile

July, 2012
Abstract

Land use and land cover (LULC) dynamics are an integral component of global change. In this thesis, various approaches were developed to unravel the complexity of LULC classification and the subsequent application of the multi-temporal LULC data for land change modelling. This complexity is particularly relevant in this study, whereby the available multi-temporal remote Landsat images are noisy and of relatively low spatial resolution. First, a semi-automated object-based method using rulesets and supervised classification was developed. This method was applied to the multi-temporal Landsat images to produce LULC maps. As the outcomes of the classification were not sufficiently accurate for land change modelling, the LULC maps were subsequently augmented using expert knowledge and input from landowners. Second, since high-resolution aerial photos were available for portions of the study area for 1998 and 2004, a case study was done with image fusion. The case study compared LULC maps derived from the different levels of fusion to those from the non-fused images. The results indicated that the feature- and decision-level fusion produced LULC maps which could be used for land change modelling. Third, in order to develop a land change model, the augmented multi-temporal LULC maps were used for extracting transition probabilities for a Markov-chain land change model. However, the classical Markov-chain method does not consider the neighbourhood influence, whereas the cellular automata does. A flexible hybrid approach, combining the Markov-chain and cellular automata algorithms, was developed. This was done to model the LULC dynamic transition probabilities to drive the change. The model’s sensitivity was assessed and the hybrid approach was tested by simulations of contemporary and future LULC patterns in the lower Hunter Valley, NSW with transition probabilities derived from various methods.
Preface

“Um passo à frente e você não está mais no mesmo lugar”

Chico Science

Travelling half way around the world is not something you do every day, nor is a PhD thesis. Journeys like these require commitment, planning and an open mind willing to learn, not to mention a lot of support. Journeys like these are like taking your first steps; one at a time and your reality is changed, forever.

This journey began when I was finishing my M.Sc. and was asked if I would like to continue on to a PhD in Texas, to which I replied, “No, I’d like to do remote sensing, land change modelling and be exposed to a new culture, maybe in Australia”. Luckily, Steve Searcy put me in contact with Alex McBratney and, after a few bumps, including me managing an organic pecan farm in Texas and my wife starting her PhD in Seville, Spain, we eventually made it here in July of 2007. I left my “comfort zone” of Agriculture and the warm waters of Brazil (Recife, PE), to embark on a journey that I knew would change me.

My journey would not have been possible without the support of my supervisor, Associate Professor Inakwu Odeh and associate supervisor Professor Alex McBratney, both of whom contributed with ideas and suggestions that enhanced the significance of this work. I would also like to thank the Faculty of Agriculture, Food and Natural Resources and the University of Sydney for the financial support.

This PhD would have been less exciting without the chats, ideas, beers, coffees and meals that I’ve shared with staff and students of our Faculty. I have decided not to name all of them for two reasons: first, the list would be quite long; and secondly, I’d surely forget someone and thus would be unfair. Thanks to all, as you know who you are. However, I’d like to acknowledge Tom Bishop and Michael Nelson, for the words of ‘suppoRt’ and for pushing
me with learning R. As for Australia, we will miss the natural beauty and warmth of this country.

Finally, I’d like to dedicate this thesis to three people. To my mother, Carmen de Castro Chaves for educating and pushing me forward; to my wife Roberta Viegas e Silva for sharing her life with me and for the continuous support; and to my son, Pedro Viegas-Stabile, who even so little has already changed our perspective in life.

The journey has not ended yet, as it is a life-long task. Thus, the submission of this thesis is only a milestone. The thesis is a treatise of the work done since July 2007 and I hope that it will open more doors to other opportunities. Where it will take me, I do not yet know, but I am excited about another step forward and embarking in a new journey, with Roberta and Pedro.
# Table of contents

Certificate of originality ........................................................................................................... i
Abstract....................................................................................................................................... ii
Preface.......................................................................................................................................... iii
Table of contents.......................................................................................................................... v
List of figures ............................................................................................................................... ix
List of tables ................................................................................................................................... xiii

Chapter 1. Thesis introduction ................................................................................................. 1

Chapter 2. Literature review: land use and land cover classification and land change modelling ........................................................................................................................................ 5
  2.1 Introduction .......................................................................................................................... 6
  2.2 Land use and land cover classification ............................................................................... 9
     2.2.1 Definition and challenges of LULC classification ....................................................... 9
     2.2.2 Workflow of LULC classification ............................................................................... 12
     2.2.3 Satellite images for LULC classification .................................................................... 14
     2.2.4 Image classification algorithms ............................................................................... 18
     2.2.5 Classification accuracy assessment ......................................................................... 21
  2.3 Land change modelling ....................................................................................................... 22
     2.3.1 Definitions and challenges of land change modelling ............................................... 22
     2.3.2 Classification of land change models ........................................................................ 25
     2.3.3 Markov-chain models of land change ...................................................................... 29
     2.3.4 Cellular automata and land change modelling ......................................................... 32
     2.3.5 Hybrid, CA / Markov-chain land change models ..................................................... 34
     2.3.6 Remarks on land change models ............................................................................. 37
  2.4 Summary ............................................................................................................................. 40
  2.5 References .......................................................................................................................... 42

Chapter 3. Object-based classification of multi-temporal satellite images: a case study of the lower Hunter valley, NSW .......................................................................................... 51
  3.1 Introduction ......................................................................................................................... 52
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2 Data and methods</td>
<td>55</td>
</tr>
<tr>
<td>3.2.1 Study area</td>
<td>55</td>
</tr>
<tr>
<td>3.2.2 Acquisition of satellite images and aerial photographs</td>
<td>58</td>
</tr>
<tr>
<td>3.2.3 Image segmentation and classification scheme using Landsat images</td>
<td>60</td>
</tr>
<tr>
<td>3.2.4 LULC classification steps</td>
<td>63</td>
</tr>
<tr>
<td>3.2.5 Detailed explanation of the ruleset based classification and supervised nearest neighbour</td>
<td>66</td>
</tr>
<tr>
<td>3.2.6 Accuracy assessment</td>
<td>69</td>
</tr>
<tr>
<td>3.2.7 Spatial resolution issues</td>
<td>70</td>
</tr>
<tr>
<td>3.3 Results and discussion</td>
<td>71</td>
</tr>
<tr>
<td>3.3.1 Classification accuracy assessment: comparison of “stage 1 maps” with the reference maps</td>
<td>71</td>
</tr>
<tr>
<td>3.3.2 Classification accuracy assessment: comparison of “stage 2 maps” with the reference maps</td>
<td>75</td>
</tr>
<tr>
<td>3.3.3 Spatial resolution issues</td>
<td>78</td>
</tr>
<tr>
<td>3.3.4 On producing historical LULC maps</td>
<td>80</td>
</tr>
<tr>
<td>3.3.5 LULC change analysis</td>
<td>82</td>
</tr>
<tr>
<td>3.4 Conclusions</td>
<td>84</td>
</tr>
<tr>
<td>3.5 References</td>
<td>85</td>
</tr>
</tbody>
</table>

Chapter 4. Improving accuracy of historical LULC maps through supervised object-based classification of multi-sensor fused images.......................... 89

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Introduction</td>
<td>90</td>
</tr>
<tr>
<td>4.2 Data and methods</td>
<td>95</td>
</tr>
<tr>
<td>4.2.1 LULC classes and sub-areas of study</td>
<td>95</td>
</tr>
<tr>
<td>4.2.2 Image data</td>
<td>96</td>
</tr>
<tr>
<td>4.2.3 Image fusion and segmentation</td>
<td>97</td>
</tr>
<tr>
<td>4.2.4 Training data and classification</td>
<td>99</td>
</tr>
<tr>
<td>4.2.5 Classification accuracy assessment</td>
<td>100</td>
</tr>
<tr>
<td>4.3 Results and discussion</td>
<td>102</td>
</tr>
<tr>
<td>4.3.1 Class separability and areal proportions</td>
<td>102</td>
</tr>
<tr>
<td>4.3.2 Comparing the patterns of LULC maps</td>
<td>104</td>
</tr>
<tr>
<td>4.3.3 Classification accuracy assessment</td>
<td>107</td>
</tr>
</tbody>
</table>
Chapter 5. Development and testing of a hybrid, weighted Markov-chain and cellular automata, land change model ......................................................... 120

5.1 Introduction ............................................................................. 121
  5.1.1 An insight into Markov-chain models .................................. 122
5.2 Data and methods .................................................................. 125
  5.2.1 Annualising the transition matrices .................................... 125
  5.2.2 Model A: a first-order Markov-chain model ....................... 127
  5.2.3 An example for incorporating neighbourhood influence .......... 128
  5.2.4 Model B: hybrid, weighted Markov-chain and CA, land change model ... 130
  5.2.5 Data and computation .................................................. 134
  5.2.6 Model comparison and sensitivity analysis .............................. 135
5.3 Results and discussion ............................................................. 137
  5.3.1 Annual transition matrices .................................................. 137
  5.3.2 Model A: a first-order Markov-chain model ....................... 139
  5.3.3 Model B: hybrid, weighted Markov-chain and CA, land change model ... 142
  5.3.4 Computation efficiency .................................................. 146
  5.3.5 Sensitivity analysis ...................................................... 147
5.4 Conclusions ............................................................................ 155
5.5 References .............................................................................. 157

Chapter 6. On the transition matrix and its parameterization ......................... 160

6.1 Introduction ............................................................................. 161
6.2 Data and methods .................................................................. 163
  6.2.1 Study area and brief history .............................................. 163
  6.2.2 Model calibration and simulation ...................................... 169
6.3 Results and discussion ............................................................. 177
  6.3.1 Model calibration ........................................................... 177
  6.3.2 Simulation of 2020 landscape ............................................ 184
  6.3.3 Discussion ................................................................. 190
List of figures

Figure 2.1. LULC transitions after Foley (2005)..........................................................8
Figure 2.2. A “pixelised” view of the world (after Fisher 1997).....................................11
Figure 2.3. Point spread function of Landsat 5 TM sensor (after Bastin et al. 2002)...........15
Figure 2.4. Issues of mixed pixels as shown in Fisher (1997)..........................................17
Figure 2.5. Three dimensional framework for describing land change models (Agarwal et al. 2002).........................................................................................24
Figure 2.6. Interaction of multiple models. Extracted from Verburg et al. (2008).................38
Figure 3.1. Region of study projected to its location in Australia (true colour composite of 1998 orthophoto)....................................................................................57
Figure 3.2. a) False colour display of 1975 MSS image, where the red circles show clouds and shadows; b) true colour display of 1990 TM image where banding occurred ........................................................................................................59
Figure 3.3. Decomposition of the segmentation algorithm built in to Definiens Professional 5.0 (extracted from Definiens 2006).................................................................61
Figure 3.4. A flowchart showing the LULC map production workflow..............................64
Figure 3.5. a) 1995 ruleset based and supervised NN LULC map; b) 1995 reference LULC map .........................................................................................................................73
Figure 3.6. a) 2000 ruleset based LULC map; b) 2000 reference LULC map.......................74
Figure 3.7. a) 1995 stage 2 map; b) 1995 reference LULC map..........................................76
Figure 3.8. a) 2000 stage 2 map; b) 2000 reference LULC map..........................................77
Figure 3.9. Proportion of LULCs with different spatial resolutions ....................................79
Figure 3.10. Population of Cessnock and Singleton from 1976-2006 (Australian Bureau of Statistics 2008)..................................................................................84
Figure 4.1. Workflow of the different fusion approaches......................................................98
Figure 4.2. A view of sub-area 1: a) 1998 orthophoto; LULC maps by: b) pixel-level fusion; c) feature-level fusion; d) decision-level fusion; e) non-fusion Landsat; and f) non-fusion orthophoto.................................................. 105

Figure 4.3. A view of sub-area 2: a) 2004 orthophoto; LULC maps by: b) pixel-level fusion; c) feature-level fusion; d) decision-level fusion; e) non-fusion Landsat; and f) non-fusion orthophoto.......................................................... 106

Figure 5.1. Conceptual workflow of model A........................................................................ 127

Figure 5.2. Conceptual workflow of model B ...................................................................... 133

Figure 5.3. 1995 initial conditions (as produced from Chapter 3) ........................................ 135

Figure 5.4. LULC maps generated by model A: a) initial conditions; b) after 1 year; c) after 10 years; and d) modal map of 500 runs ................................................................................... 141

Figure 5.5. LULC maps generated by model B after 10 years: a) four neighbours, $W_{TM}=0.2$; b) eight neighbours, $W_{TM}=0.2$; c) four neighbours, $W_{TM}=0.8$; and d) eight neighbours, $W_{TM}=0.8$ ........................................................................ 145

Figure 5.6. Sensitivity of models A and B to varying $W_{TM}$, as measured by overall accuracy ........................................................................................................... 149

Figure 5.7. Sensitivity of models A and B to varying $W_{TM}$, as measured by location Kappa ........................................................................................................ 149

Figure 5.8. Sensitivity of models A and B to varying $W_{TM}$, as measured by histogram Kappa ........................................................................................................ 150

Figure 5.9. Sensitivity of models A and B to varying $W_{TM}$, as measured by fractal dimension ................................................................................................. 151

Figure 5.10. Sensitivity of models A and B to varying $W_{TM}$, as measured by Simpson's diversity index ......................................................................................... 152

Figure 5.11. Sensitivity of models A and B to varying $W_{TM}$, as measured by edge density... 152
Figure 5.12. Sensitivity of models A and B, as measured by a three-dimensional surface plot of overall accuracy, the number of runs and $W_{TM}$ (using eight neighbours).........153

Figure 5.13. Sensitivity of models A and B, as measured by a three-dimensional surface plot of edge density, the number of runs and $W_{TM}$ (using eight neighbours).............154

Figure 6.1. Tonnage of grapes grown and imported to the Hunter (Australian Wine and Brandy Corporation 2008)........................................................................................................166

Figure 6.2. Annual number of visitors to the Hunter region (Tourism New South Wales 2009)........................................................................................................................................166

Figure 6.3. Population trends and projections of two major statistical regions in the Hunter region (Australian Bureau of Statistics 2009, Hunter Valley Research Foundation 2009)..................................................................................................................167

Figure 6.4. Histogram Kappa, as function of experimental transition matrix, $W_{TM}$ and number of runs......................................................................................................................................178

Figure 6.5. Location Kappa, as function of experimental transition matrix, $W_{TM}$ and number of runs......................................................................................................................................179

Figure 6.6. Overall accuracy, as function of experimental transition matrix, $W_{TM}$ and number of runs......................................................................................................................................180

Figure 6.7. Deviation from reference edge density, as function of experimental transition matrix, $W_{TM}$ and number of runs......................................................................................................................................182

Figure 6.8. Deviation from reference fractal dimension, as function of experimental transition matrix, $W_{TM}$ and number of runs......................................................................................................................................183

Figure 6.9. Percentage (%) of study area covered by each LULC class in 2020, predicted by different models, as a function of $W_{TM}$ and number of runs.................................186

Figure 6.10. Percentage (%) of study area covered by “intensive uses” in 2020, predicted by different models, as a function of $W_{TM}$ and number of runs.................................187
Figure 6.11. Percentage (%) of study area covered by “irrigated Agriculture” in 2020, predicted by different models, as a function of $W_{TM}$ and number of runs..........................188

Figure 6.12. a) Reference 2005; b) simulated 2020, MNLR AllClim, $W_{TM}$=0.0, 20 runs;
   c) simulated 2020, Empirical, $W_{TM}$=0.5, 20 runs; d) simulated 2020, MNLR Pop, $W_{TM}$=1.0, 20 runs..........................................................190
List of tables

Table 2.1. Historic LULC datasets (after Goldewijk and Ramankutty 2004) ........................................... 7
Table 2.2. Steps for image classification (adapted from Jensen 2005) ..................................................... 13
Table 2.3. Description of classification methods (adapted from Lu and Weng 2007) ......................... 20
Table 2.4. Categories of land change models, according to Lambin et al. (2000) ......................... 26
Table 2.5. Some land change models and the methods employed by them (adapted from Koomen and Stillwell 2007) .................................................................................................................................................................................. 27
Table 2.6. Characteristics of nine land change models (after Pontius Jr et al. 2008) ......................... 28
Table 3.1. Number of objects, for each scale parameter, for each year ................................................. 61
Table 3.2. Accuracy assessment of “stage 1 maps” versus reference ................................................. 71
Table 3.3. Accuracy assessment of “stage 2 maps” versus reference ............................................... 75
Table 3.4. Proportions (%) of LULC at different resolutions .............................................................. 78
Table 3.5. Fragmentation statistics at different resolutions ............................................................... 80
Table 3.6. Composition of the landscape in terms of LULC classes per year ....................................... 83
Table 4.1. Number of objects for each sub-area .................................................................................. 99
Table 4.2. Number of training polygons per sub-area ...................................................................... 100
Table 4.3. Proportion of points for SA1 ......................................................................................... 103
Table 4.4. Proportion of points for SA2 ......................................................................................... 104
Table 4.5. Accuracies of SA1 ........................................................................................................... 107
Table 4.6. Accuracies of SA2 ........................................................................................................... 107
Table 5.1. Change matrix for 1995-2000 ...................................................................................... 137
Table 5.2. Annualised transition probabilities (1995-2000) .......................................................... 138
Table 5.3. Error matrix of model A (1 run, 10 years) versus reference ........................................ 139
Table 5.4. Error matrix of model A (500 runs, 10 years) versus reference .................................... 140
Table 5.5. Error matrix of model B (1 run, 10 years), four neighbours and $W_{TM}=0.2$

versus reference ..............................................................................................................143

Table 5.6. Error matrix of model B (1 run, 10 years), four neighbours and $W_{TM}=0.8$

versus reference ..............................................................................................................143

Table 5.7. Error matrix of model B (500 runs, 10 years), four neighbours and $W_{TM}=0.5$

versus reference ..............................................................................................................144

Table 5.8. Approximate run times for models A and B with various configurations ..........147

Table 6.1. Proportions of LULC categories ....................................................................165

Table 6.2. Experimental setup and naming of each approach .......................................170

Table 6.3. Empirical transition probabilities ..................................................................173
Chapter 1. Thesis introduction
Land use\(^1\) and land cover\(^2\) (LULC) changes are an inevitable consequence of human interference with nature that, if not properly managed, may lead to severe impacts on the functioning of the Earth system. In Australia, there is a lack of rigorous land change analysis, especially in the non-urban environment, where the complexity of LULC and their associated dynamisms are prevalent. Such an analysis is needed to provide baseline information and knowledge for environmental monitoring and management. There is also no consensus on what is the best technique for determining LULC patterns. While LULC classification has been based on remote sensing technologies dating back to the early 20\(^{th}\) century, there is currently no universal agreement in terms of the best classifications. Nor is it to be expected there will ever be such an agreement. However, it is pertinent to revisit some of the techniques, particularly supervised classification algorithms, the recently emerging object-based approach and the need to explore fusion of the medium-spatial resolution multispectral satellite images with high-spatial resolution multispectral aerial photos.

A related issue is land change modelling, which is important for monitoring and managing the global LULC change. Like the case of LULC classification and mapping, there are plethora of methods and approaches for land change modelling. Amongst these, is the classical Markov-chain algorithm which uses transition probabilities to model the change. Despite its wide use, it is recognised that the Markov-chain lacks the ability to take into account the spatial structure of LULC patterns, i.e. that neighbouring land parcels are more likely to be of the same LULC category than parcels farther apart. To account for the spatial structure, the use of cellular automata (CA) has become prevalent. However, the classical CA lacks the dynamism expected of a complex changing system such as LULC patterns. Moreover, it is known that the LULC changes are not solely reliant on the neighbourhood configuration.

---
\(^1\) Land use is the consequence of the anthropogenic change to a part of earth’s surface.
\(^2\) Land cover is the natural state of earth’s cover before anthropogenic modification
Therefore, the main objective of this thesis is to explore the complexity of LULC classification and land change modelling, as well as to develop techniques that facilitate LULC mapping and land change modelling. Specifically, the thesis is therefore aimed at:

1. exploring in details the mainstream LULC classification algorithms, especially in relation to the use of the medium-resolution satellite (Landsat MSS – TM) images;
2. rigorously examining the fusion of the multispectral Landsat images with the high resolution aerial photos to enhance automated LULC mapping;
3. investigating the hybridisation of the classical Markov-chain algorithm with CA to take advantages of the two techniques of land change modelling;
4. assessing the hybrid techniques using a regional case study by its comparison with the generic techniques.

The organisation of this thesis is such that the chapters can be treated individually. However, by reading the whole body of work, it shows the reader how to create LULC maps, derive change matrices and project future LULC patterns caused by biophysical and socio-economic factors, using the methods and models that were developed here. The structure also allows the reader to focus on specific areas of interest and avoiding oneself to read the whole body of work in search of information regarding the topic of interest.

Thus Chapter 2 deals with the review of literature, covering such issues as the assumptions, advantages and limitations of a number of algorithms for LULC classification and land change modelling. It highlights the necessity and importance of accuracy assessment of the maps resulting from LULC classification. In Chapter 3, a multi-temporal series of LULC maps is developed through the use of an object-based approach associated to ruleset and supervised classification methods, augmented by expert knowledge. These LULC maps are later used to derive transition matrices for land change modelling, covered in Chapter 5. Chapter 4 presents a case study illustrating how to utilise multi-source data through image fusion, to enhance the accuracy of LULC maps produced by an automated object-based and
supervised approach. Thus in Chapter 4, the three different levels of image fusion are tested and assessed. The assessment is based on accuracy metrics and on benchmarking the LULC maps produced from the different levels of fusion against the LULC maps produced from the original (non-fused) images. In Chapter 5, two land change models are developed. The first model is based on first-order Markov-chains, while the second is a hybrid of the weighted Markov-chains and cellular automata models. In order to use the concept of first-order Markov-chains and to enable the running of the models in yearly time steps, it is necessary to annualise the transition matrices, which is also done in this chapter. Further to the development of these two models, the sensitivity to change in the parameters is also tested. In Chapter 6, a number of methods for deriving transition matrices are used to run the hybrid land change model, in a case study covering the lower Hunter Valley of NSW, Australia. Finally, Chapter 7 presents the salient findings of the thesis and synapses for future studies, while the Appendix includes the source code of the models developed here.
Chapter 2. Literature review: land use and land cover classification and land change modelling
2.1 Introduction

Earth’s surface has been thoroughly modified since the settlement of humans in society. According to Flannery (1973) and Harlan (1971), villages appeared and grew significantly, leading to the modification of land. In the early history of humankind, humans were hunter-gatherers, extracting their resources from the natural environment for survival. However, when land cultivation began about 10,000 years ago (Harlan 1971), humans became dependent on Agriculture\(^1\). This led to a system that supported more people living in a single place (Byrd 2005).

The emergence of Agriculture, which is generally regarded as the domestication of species, contributed to increased productivity (Glover et al. 2007) and allowed the development of villages. Consequently, larger areas were required to provide food and material for shelter and clothing for the ever growing population. The industrial revolution in the 18\(^{th}\) and 19\(^{th}\) centuries (Falkowski et al. 2000, Pittock 2003) that led to mechanisation (Lambin et al. 2001, Foley et al. 2005), marked the culmination of human development and provided the turning point of significant changes to the land surface.

It is now well known that since the industrial revolution, humans have contributed to the enhanced greenhouse gas emissions (Fischer et al. 1999), sped up deforestation (Laurance 1998), expanded the amount and size of our cities (Kalnay and Cai 2003), as well as increased our population significantly (Australian Bureau of Statistics 2008). Thus, all of these changes have consequences in terms of the dynamics of the Earth as a system. These consequences include the widely known climate change (IPCC 2001) and the increased need of land for food and fibre production (Tubiello and Fischer 2007). However, as Alley et al. (2003) have shown other consequences are yet unknown.

\(^1\) Agriculture will be capitalised throughout the thesis. This is due to its importance, in enabling our society to exist.
At global and national scales, some understanding of the effect of anthropogenic land use\(^2\) and land cover\(^3\) (LULC) changes\(^4\) (Lambin 2006) can be gauged, mainly through the use of historical LULC datasets created by modelling (Goldewijk and Ramankutty 2004). A summary of related works in this regard is presented in Table 2.1 below.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Spatial characteristics</th>
<th>Temporal characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Local/National level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White and Mladenoff (1994)</td>
<td>Northern Wisconsin, WI, USA</td>
<td>1860s, 1931, 1989</td>
</tr>
<tr>
<td>Bork et al. (1998)</td>
<td>Germany</td>
<td>7th century–present</td>
</tr>
<tr>
<td>Crumley (2000)</td>
<td>Burgundy (France)</td>
<td>Iron Age–present</td>
</tr>
<tr>
<td>Himiyama (1992)</td>
<td>Japan</td>
<td>1850, 1900, 1980</td>
</tr>
<tr>
<td>Larsson and Frisk (2000)</td>
<td>Sweden</td>
<td>ca. 1700–present</td>
</tr>
<tr>
<td>Manies &amp; Mladenoff (2000)</td>
<td>Sylvania Wilderness Area, MI, USA</td>
<td>pre-settlement</td>
</tr>
<tr>
<td>Ogaard &amp; Rasmussen (2000)</td>
<td>Denmark</td>
<td>past 2 millennia</td>
</tr>
<tr>
<td>Petit and Lambin (2002)</td>
<td>Belgium Ardennes</td>
<td>1700–present</td>
</tr>
<tr>
<td><strong>Continental level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Darby (1956)</td>
<td>Central Europe</td>
<td>900, 1900</td>
</tr>
<tr>
<td>Williams (2000)</td>
<td>Western Europe, reconstructed map</td>
<td>11th–13th century</td>
</tr>
<tr>
<td>AUSLIG (1989)</td>
<td>Australia, scale 1:20,000,000</td>
<td>pre-settlement (1780), 1980s</td>
</tr>
<tr>
<td><strong>Global level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramankutty and Foley (1999)</td>
<td>5 min. resolution</td>
<td>1700–1992</td>
</tr>
<tr>
<td>Goldewijk (2001)</td>
<td>0.5 x 0.5 degree grid</td>
<td>1700–1990</td>
</tr>
</tbody>
</table>

There is some understanding of regional and local anthropogenic LULC changes, as elucidated in the table above, however this is dependent on mapping LULC over time (Verburg et al. 2006).

Mapping of historical LULC patterns allows measuring and understanding the evolution of landscapes (DeFries and Bounoua 2004). This can provide information for better land management practices and alternative LULCs at the global, national, regional and local scales.

---

\(^2\) Land use, as defined by Lambin et al. (2001), is the human purpose or intent given to a specific part of earth’s surface.

\(^3\) Land cover, as also defined by Lambin et al. (2001), is the biophysical attribute of earth’s cover (e.g. a natural forest is a land-cover, while a planted forest is a land-use)

\(^4\) Land use and land cover change is also referred to as land dynamics throughout the thesis.
(Dale 1997). A general scheme of the evolution of landscapes is illustrated in Figure 2.1. While this shows the intensification of the land uses from the natural ecosystems in a pre-settlement scenario, to a mix of intensive uses as seen today, it has become increasingly apparent that the latter might not be sustainable in the long run. Thus, the future composition of the landscape is not known, as illustrated by the question mark in Figure 2.1.

![Figure 2.1. LULC transitions after Foley (2005)](image)

Understanding the underlying causes of land dynamics based on multi-temporal classification of LULC has been the theme of numerous studies. LULC classification is ultimately aimed at producing LULC maps that describe the features of a specific area (Jensen 2005). Multi-temporal LULC maps have been used for land resource management, through the retrieval of biophysical parameters of cover (Ju et al. 2005), for land change detection and modelling (Groot et al. 2009), amongst other uses.

LULC classification is not as simple as one would imagine. Different approaches exist, most of which rely on remote sensing images (Mather 2004). The limitations posed by these images are often related to their spatial, spectral and temporal resolutions, but are also related to image availability and spectral noise. Because of these problems, there is no consensus on a
global solution to LULC classification (Lu and Weng 2007). The choice of a classification strategy greatly depends on the issues mentioned above, but also on the availability of reference data, accessibility and spatial extent of the study area. In the next section (Section 2.2), some issues of LULC classification are reviewed and methods used to enhance LULC classification utilising multi-source data are highlighted.

As the case with LULC classification, Pontius Jr et al. (2008) points out that there is no consensus on the best land change model. This review also covers some of the widely used land change modelling methods/approaches, highlighting their strengths and weaknesses (Section 2.3).

2.2 Land use and land cover classification

2.2.1 Definition and challenges of LULC classification

LULC classification, based on remote sensing data, involves the identification and separation of land surface cover categories with the objective of producing a detailed LULC map describing land surface features of a specific area. Thus, LULC classification constitutes the base for many environmental and socio-economic applications (Lu and Weng 2007). For instance, the European Environment Agency recognises the importance of LULC classification in providing information on land status. Such classification can be used for: i) the assessment of environmental impacts of LULC change, ii) the assessment of land use potential and alternative uses. More importantly, the classification itself equips scientists and land planners to compare and analyse data from different sources and dates. Naturally, as mentioned above, LULC classification combined with land change modelling has only become possible in the last few decades, due to the development of personal computers.

---

5 Land change model is a model used to describe or understand the changes occurring in LULC patterns through time. These models may or may not be spatial and serve to answer different research questions as further elucidated in Table 2.4
Ju *et al.* (2005) emphasized the importance of classification as fundamental to land resource management and retrieval of biophysical parameters of land covers, such as leaf area index, albedo and surface roughness. At about the same time, Jensen (2005) affirmed that classification is a transformation process which converts data (pixel values) to information. In this sense, LULC classification requires the user to consider many factors such as: image pre-processing, classification system to be used, training sites (if required), feature extraction, post-classification processing and most importantly classification accuracy assessment.

The essential element of LULC classification is to organise land surface features categorically, spatially and temporally, therefore enabling the comparison of LULC maps from different areas and over a period of time (Anderson *et al.* 1976). For this reason, LULC classification is a pre-requisite for LULC change detection. LULC change detection as a pre-requisite for land change modelling, leading to the understanding of environmental impacts of anthropogenic interference (Lambin 2006) and providing information for policy decision-making regarding land use planning (Batisani and Yarnal 2009).

LULC classification depends on the availability of data and on the selection of appropriate classification methods. While some scholars emphasize the strengths of certain classification algorithms (Kahya *et al.* 2008, Yuan *et al.* 2008), others (e.g. Lu and Weng 2007) affirm that there is no superior classification algorithm to be used. This is because the choice of a given algorithm needs to consider data availability, data quality criteria (spatial, spectral, temporal and radiometric resolutions), accessibility to the study area and the reference data required for quality assurance.

In considering image resources for LULC classification, there are many sources of error intrinsically linked to the classification process, as summarised in the seminal article by Fisher (1997). The problem of mixed pixels, as coined by Fisher, negatively affects classification accuracy as the spatial resolutions of images limits the identification of land features (Verbeiren *et al.* 2008). For instance, a hard classification process will assign a single
class for any pixel. However, depending on the spatial resolution of the sensor, each pixel may comprise mixed LULC classes. Figure 2.2 below shows an ideal landscape for classification, where objects are square, as most image pixels, each one representing exactly one LULC class. Unfortunately, images acquired from older sensors often comprise pixels of mixed features.

![Figure 2.2. A “pixelised” view of the world (after Fisher 1997)](image)

Nowadays with the evolution of sensors and increase in the spatial resolution, it is possible to have spectrally pure pixels. However, earlier images such as those acquired from Landsat multispectral scanner (MSS) and Landsat thematic mapper (TM) were characterised by mixed pixels, due to their coarse spatial resolution. While mixed pixels are not the only
source of classification error, they pose a challenge when analysing both low spatial resolution (mixed pixels) and high spatial resolution images (spectral noise).

Lu and Weng (2007) categorised other sources of errors as being: classification errors (from mixed pixels or lack of ground data), positional errors (from registration), interpretation errors (inexperienced analyst) and poor quality of training data (for supervised classification approaches).

Nevertheless, the problems with classification do not end with the various error sources, image quality, or lack of it. One issue that has haunted the environmental science community is related to data availability. This is a direct result of several factors. For example, images acquired by optical sensors are often problematic, as they are limited by cloud cover (Mather 2004). This often leads to limited availability of remote sensing data for selecting anniversary dates for LULC change detection and for monitoring agricultural areas within the crop-growing season. Furthermore, if the research requires historical images, such as those acquired by Landsat MSS, the research is limited by the sensor’s coarse spatial resolution. Aerial photography, which pre-dates satellite remote sensing, would resolve the issue of low spatial resolution, but it lacks the temporal resolution of satellite remote sensing. Additionally, most of the early aerial photos had a low spectral resolution (B&W or RGB). This double-sided challenge demands a special combination of technologies and methodologies to extract historical LULC information in the quest to model and monitor global changes. This requires a well defined workflow of LULC classification, the topic of the following section.

2.2.2 Workflow of LULC classification

The evolution of LULC classification has been intrinsically linked to the resources available for classification. In the pre-remote sensing era, LULC mapping projects involved field-based land survey of small areas (Petit and Lambin 2002). Beginning from the early 20th century, aerial photography and later satellite remote sensing, provided data for detailed classification of larger areas. Such remote sensing classification schemes have equipped
scientists to assess the anthropogenic impacts on our planet, based on monitoring past changes using multi-temporal LULC surrogate data provided by the remote sensing technologies (Hurtt et al. 2006).

Table 2.2. Steps for image classification (adapted from Jensen 2005)

<table>
<thead>
<tr>
<th>1. Nature of classification problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Specify region of interest;</td>
</tr>
<tr>
<td>b. Define classes of interest;</td>
</tr>
<tr>
<td>c. Determine classification to be hard or soft; and</td>
</tr>
<tr>
<td>d. Determine classification to be per-pixel or object oriented.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Acquire appropriate remote sensing and ground reference data</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Select data based on:</td>
</tr>
<tr>
<td>i. R.S. considerations: spatial, spectral, temporal and radiometric resolutions;</td>
</tr>
<tr>
<td>ii. Environmental: atmospheric, soil moisture, phenological cycle, etc; and</td>
</tr>
<tr>
<td>iii. Availability of resources and cost of data.</td>
</tr>
<tr>
<td>b. Obtain initial ground reference data from a priori knowledge of study area.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Process remote sensor data</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Radiometric correction;</td>
</tr>
<tr>
<td>b. Geometric correction;</td>
</tr>
<tr>
<td>c. Select proper image classification logic:</td>
</tr>
<tr>
<td>i. Parametric;</td>
</tr>
<tr>
<td>ii. Non-parametric;</td>
</tr>
<tr>
<td>iii. Non-metric (e.g., ruleset based, decision-tree classifier).</td>
</tr>
<tr>
<td>d. Select appropriate image classification algorithm:</td>
</tr>
<tr>
<td>i. Supervised;</td>
</tr>
<tr>
<td>ii. Unsupervised;</td>
</tr>
<tr>
<td>iii. Hybrid, involving artificial intelligence (ANN).</td>
</tr>
<tr>
<td>e. Extract data from initial training sites;</td>
</tr>
<tr>
<td>f. Select most appropriate bands using feature selection criteria:</td>
</tr>
<tr>
<td>i. Graphical;</td>
</tr>
<tr>
<td>ii. Statistical.</td>
</tr>
<tr>
<td>g. Extract training statistics and rules based on:</td>
</tr>
<tr>
<td>i. Final band selection;</td>
</tr>
<tr>
<td>h. Extract thematic information</td>
</tr>
<tr>
<td>i. For each pixel or for each segmented image object (supervised);</td>
</tr>
<tr>
<td>ii. Label pixel or image objects (unsupervised).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Perform accuracy assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Select method:</td>
</tr>
<tr>
<td>i. Qualitative</td>
</tr>
<tr>
<td>ii. Statistical measurement</td>
</tr>
<tr>
<td>b. Determine number of samples required by class</td>
</tr>
<tr>
<td>c. Select sampling scheme</td>
</tr>
<tr>
<td>d. Obtain ground reference test information</td>
</tr>
<tr>
<td>e. Create and analyze error matrix</td>
</tr>
<tr>
<td>i. Univariate and multivariate statistical analysis</td>
</tr>
</tbody>
</table>

| 5. Accept or reject hypothesis                          |

| 6. Distribute results if accuracy is acceptable         |

Table 2.2 above presents the general steps necessary to perform remote sensing image classification. This table highlights the complexity of a classification scheme where many processes are involved, each of which is a possible source of error. From the processes
illustrated in Table 2.2, if two scientists were to produce independently a LULC map of a given area, the outcome might be different, even if they were to use the same data and processing methods. Moreover, the table above also points to the fact that there is no standard classification procedure/method, as adduced by Lu and Weng (2007).

As discussed above, a key limitation of LULC classification in many situations is limited data availability. While data availability has increased significantly, augmented by advances in remote sensing technology and sensible increase in remote sensing launches in various countries within the last decade or so, integration of the resulting multi-source data for meaningful multi-temporal monitoring of land surface has become a major challenge.

2.2.3 Satellite images for LULC classification

The possibility of observing the world from above has enabled scientists to understand and predict natural phenomena (Boyd and Danson 2005), through monitoring and modelling changes taking place at the global, regional and local scales (Lambin 2006). In fact, the development of satellites for civilian remote sensing was initially limited by the cold war (Aplin et al. 1997). The first of a series of sensors available for civilian use was the MSS onboard of Landsat, launched in 1972 by NASA. It had a spatial resolution of 80 m and four spectral bands. Since then, other remote sensing satellites have been launched and the spatial resolution has improved from 80-m pixels in the Landsat MSS, to sub-meter resolution in the panchromatic bands of Quickbird and IKONOS. This evolution has also led to a great leap in the spectral resolution, for example NASA’s Hyperion images comprise 220 spectral bands.

Satellite remote sensing images have the advantage of fine spatial-resolution, high geometric precision, short revisit intervals and rapid data supply (Aplin et al. 1997), when compared with aerial photographs and manual data collection methods. Thus, the availability of a variety of sensors has enabled the use of multi-source, multi-temporal data for LULC monitoring and land change modelling.
In order to produce accurate LULC maps, it is beneficial to use more than one source of information. This could be multi-source data, comprised of images from different sources or images and a spectral data library, for instance. Integration of multi-source, multi-temporal data has been approached differently by scientists. A comprehensive review of the challenges posed by data integration is provided by Pohl and van Genderen (1998). These challenges are related mainly to the purpose of data integration. These depend on a number of features: i) availability of different sources of data; ii) the selection of sensors and their specific characteristics; and iii) the geometric correction of the raw data, among others. The choice of an appropriate technique for integrating multi-source, multi-temporal data also depends on the criteria to be evaluated. For instance, if the objective is to preserve visual integrity of the data, the methods may be quite different from the requirements for producing LULC maps.

If the purpose of using multi-source images is LULC classification, it may be necessary to consider the differences in acquisition dates, spatial and spectral resolutions, sensor calibration parameters and the point spread function for each sensor. Figure 2.3 illustrates the point spread function of Landsat TM, whereby any given pixel value is composed mainly of the reflected light captured at the centre of the pixel. The practical application of this is that if multi-sensor images were to be used, the effective pixel size of the resulting classification would be partially determined by the point spread function of each sensor.

Figure 2.3. Point spread function of Landsat 5 TM sensor (after Bastin et al. 2002)
Turning to the problem of LULC classification, Addink *et al.* (2006) proposed a method for assessing classification performance using images of different spatial resolutions, without actually classifying the images for a specific area. By using different settings of a spectral/spatial clustering (segmentation) algorithm on images with variable spatial resolution, it would be possible to determine which best matched the “reference” high-spatial resolution image. In this sense, while this method enables the user to select the appropriate scale for classification, it requires more data than is often available for LULC classification. In other words, in a hypothetical situation where a variety of images of different spatial resolutions are available, the method proposed by Addink *et al.* (2006) would enable selection of the image most suitable for classification. Unfortunately, this multitude of images is not commonly available.

If a study’s objective is the creation of historical LULC maps for change detection and land change modelling, then the method chosen would unavoidably rely on archive data such as Landsat (Alvarez *et al.* 2003, Peterson *et al.* 2004, Zhang *et al.* 2007) and aerial photos (Erdogan *et al.* 2008). Since the level of classification accuracy depends on the spatial resolution of the images and the spectral characteristics of the objects to be classified, the readily available spatial ancillary data and historical aerial photos can be used to enhance the classification. The non-geographical ancillary information such as census data and information obtained by interviews of the locals, as well as spectral libraries can be incorporated to the classification, increasing its accuracy.

As noted above, images characterised by low spatial resolution are faced with the problems of mixed pixels (Fisher 1997). When generating LULC maps from such images, the mixed pixel problem, as shown in Figure 2.4 below, may contribute to classification error due to misleading proportions or even exclusion of classes. There are two solutions to this

![Sub-pixel and Boundary pixel](image1)

![Intergrade and Linear features](image2)

Figure 2.4. Issues of mixed pixels as shown in Fisher (1997)

The first work on sub-pixel classification was by Foody and Cox (1994). In their approach the authors used a linear mixture model, regression and fuzzy memberships associated with knowledge of the different spectral signatures of the ground features, to determine multiple class memberships within each pixel. A similar approach was used by Verbeiren et al. (2008), who compared the linear mixture model with a neural network approach for determining class memberships within the pixels. In a critical analysis of the sub-pixel classification, Tatem et al. (2003) pointed out the limitation of the sub-pixel classification approach in failing to predict the location of each class within the pixel and proposed a solution using neural networks. Another known limitation of sub-pixel classification is that it requires detailed information of the pure spectral responses of each cover (for the mixture model) and/or
cadastral data to aid in the classification procedure, limiting its application when either of these data is lacking.

Image fusion, which is another solution to the problem of mixed pixels, enables the combination of multi-temporal, multi-sensor images. Welch and Ehlers (1987) were among the first to use image fusion, combining a 10-m resolution SPOT-1 panchromatic band with the 28.5-m resolution Landsat TM images, over an urban area, to increase the quality of the image for visual interpretation. The procedure of merging the digital numbers of two different sensors was labelled as pan-sharpening. This procedure has become commonplace for the panchromatic and multispectral bands of sensors such as ETM+, SPOT, Quickbird, amongst others.

Pohl and van Genderen (1998) differentiated three levels of image fusion: i) pixel-level fusion, where the digital numbers of two sensors were merged, increasing the spatial resolution while preserving the spatial resolution; ii) feature-level fusion, where multi-source data would be used for delineating objects, which would then be classified; and iii) decision-level fusion, where the higher-resolution data would be used for delineating objects, but multi-source data would be used in the classification process.

In terms of LULC classification, the three levels of image fusion provide a flexible tool for enhancing image classification. Image fusion enables the integration of multi-source, multi-date data, but since there is no general agreement on benchmark algorithms (Colditz et al. 2006), different fusion approaches need to be tested.

### 2.2.4 Image classification algorithms

The selection of a suitable classification approach, as mentioned earlier, depends on data availability (or lack of it), access to the study area and often the spatial extent to be mapped. As such, the first decision is related to the smallest unit of classification, if it will be done on sub-pixels, pixels or objects. Following this, it is necessary to decide whether to use supervised or unsupervised methods, using parametric or non-parametric statistics and finally
if the output classified map will be raster or vector, with a single class (hard classification) or if they can have membership values to many classes (soft classification).

A summary of these different distinctions was compiled by Lu and Weng (2007), as shown in Table 2.3 below. Thus the different criteria for classification are compared and each category within the criteria on the left of the Table 2.3 can be combined. The implication of such combinations is that it gives rise to a multitude of classification approaches.

While the selection of an appropriate classification approach is often a task left to the end user (Wang et al. 2007), it is important to understand some of their characteristics. As a general rule, the first decision is whether classification would be done on pixels or objects. While the object-based approach is more intuitive, as it groups pixels into objects, the data processing requirements for this approach have only become available in the last decade or so (Blaschke and Hay 2001). On the other hand, pixel-based approaches are more common, but are faced with the “salt and pepper” effect (Rogan et al. 2002). This is due to the fact that spectral discrimination does not consider the neighbourhood influence, as single pixels are classified independently of their neighbours. While some studies demonstrate the superiority of the pixel-based approach (Wang et al. 2008), or of the object-based approach (Mitri and Gitas 2006), other studies find the performance of both approaches to be similar (Wang et al. 2007).

Segmentation procedures (Benz et al. 2004) employed on the object-based approaches are no more than a weighted spatial/spectral clustering technique. This can be interpreted as an unsupervised classification algorithm, whereby the spatial and spectral characteristics of the pixels are used for clustering similar pixels that will form objects (Jobin et al. 2008). This makes the boundary of pixel-based versus object-based approaches similar to that of unsupervised versus supervised algorithms.
<table>
<thead>
<tr>
<th>Criteria</th>
<th>Categories</th>
<th>Characteristics</th>
<th>Example of classifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of training data or not</td>
<td>Supervised</td>
<td>LULC classes are defined. Reference data is used as training samples. Spectral signatures of training samples are used as class definitions and a thematic map is derived</td>
<td>Nearest neighbour, max. likelihood, min. distance, artificial neural network (ANN), decision-tree (DT), expert system (ES)</td>
</tr>
<tr>
<td></td>
<td>Unsupervised</td>
<td>Clustering into a number of spectral classes based on image statistics. Class definitions are determined by the algorithm. Clusters are labelled and merged by use of a posteriori</td>
<td>ISODATA, K-means clustering algorithm</td>
</tr>
<tr>
<td>Use of statistical parameters or not</td>
<td>Parametric classifiers</td>
<td>Normal distribution is assumed and parameters are extracted from the image. Limitations include the difficulty of integrating ancillary data, spatial attributes and non-statistical information</td>
<td>Maximum likelihood, linear discriminant analysis</td>
</tr>
<tr>
<td></td>
<td>Non-parametric classifiers</td>
<td>No distribution of the data is assumed. It is not based on statistical measures and it is suitable for incorporating ancillary data and non remote sensing data in classification</td>
<td>ANN, decision-tree, support vector machine (SVM), expert system</td>
</tr>
<tr>
<td>What is the minimum classification unit</td>
<td>Per-pixel classifiers</td>
<td>Pixels are individually classified, thus results rely greatly on the quality of training data. Mixed pixels are ignored and put in to some class. Spectral signature of training data determines class properties</td>
<td>Most classifiers: max. likelihood, min. distance, ANN, decision-tree, SVM</td>
</tr>
<tr>
<td></td>
<td>Sub-pixel classifiers</td>
<td>Each pixel may represent a pure class or a mix of classes. This is determined by a linear or non-linear combination of different endmembers. Each pixel may have proportion of classes</td>
<td>Fuzzy-set classifiers, sub-pixel classifier, spectral mixture analysis</td>
</tr>
<tr>
<td></td>
<td>Object-oriented classifiers</td>
<td>Images are clustered spatially/spectrally producing segments. These should represent features on the ground. Limitation is associated to quality of segmentation, which can incorporate GIS layers</td>
<td>eCognition, Erdas, ENVI, watershed segmentation, tessellation</td>
</tr>
<tr>
<td></td>
<td>Per-field classifiers</td>
<td>GIS and remote sensing data are integrated. Parcel boundaries are used to define class boundaries and thus spectral variations within parcels are ignored</td>
<td>GIS-based classification approaches</td>
</tr>
<tr>
<td>Output has definitive or membership to classes</td>
<td>Hard classification</td>
<td>The smallest unit of analysis (pixel or object) is given a single class. Easier to interpret, but coarse data often has mixed pixels, causing misclassification</td>
<td>Max. likelihood, min. distance, ANN, decision-tree, SVM</td>
</tr>
<tr>
<td></td>
<td>Soft (fuzzy) classification</td>
<td>Each pixel has a membership to one or more classes. Potentially more accurate results with coarse spatial resolution images, but spatial allocation of classes is an issue</td>
<td>Fuzzy-set classifiers, sub-pixel classifier, spectral mixture analysis</td>
</tr>
<tr>
<td>Classification incorporates spatial information</td>
<td>Spectral classifiers</td>
<td>Spectral information is used for class determination, but may result in “salt and pepper” effect, as classes are determined purely based on spectra</td>
<td>Usually pixel-based: max. likelihood, min. distance, ANN, decision-tree, SVM</td>
</tr>
<tr>
<td></td>
<td>Contextual classifiers</td>
<td>Neighbouring pixel information is used for classification</td>
<td>Iterated conditional modes, point-to-point contextual correction</td>
</tr>
<tr>
<td></td>
<td>Spectral-contextual classifiers</td>
<td>Spatial and spectral information is used for classification. This can be done through segmentation or using parametric or non-parametric classifiers to generate initial classification images and then contextual classifiers are implemented in the classified images</td>
<td>ECHO, combination of parametric or non-parametric and contextual algorithms</td>
</tr>
</tbody>
</table>
As pointed out above, no single classification method is definitive (Gao 2008), with each having its strengths and weaknesses. While some authors claim the superiority of their own method, it entirely depends on the availability of resources for classification. For instance, a supervised classification approach is not recommended if the training data were not readily available. On the other hand, if expert knowledge and high-resolution images are readily available, then the supervised classification generally yields better results than the unsupervised approach. Ultimately, the choice of the classification method will largely depend on data availability, associated to the knowledge of the area of study (Mather 2004, Berberoglu and Akin 2009).

In summary, classification of LULC classes depends not only on a combination of methods, but also on the data and expertise available, as expounded by Lu and Weng (2007). It is noteworthy to keep in mind the importance of data selection for classification, to minimise spectral redundancy provided by bands with similar spectral characteristics. In doing so, the efficiency of image processing can be increased without affecting the quality of the end product— the LULC maps.

2.2.5 Classification accuracy assessment

Over the years, the methods for accuracy assessment of LULC classification maps have increased in complexity. Foody (2002) discussed the evolution of classification accuracy assessment through its various stages. Accordingly, the first stage was characterised by the visual assessment, whereby the question was: “does the map look right?” The visual approach was commonly used in the pre-digital age and pre-remote sensing era.

The second stage of classification accuracy assessment was slightly more objective than the visual assessment. It involved the comparison of areal proportions of the map with the ground measurements, however disregarding the spatial distribution of the ground and mapped classes. In this stage, a high agreement in terms of class proportions would indicate
good classification accuracy. The problem with this approach was that the spatial allocation of the classes would not matter, only the class proportions.

After the comparison of areal proportions, the third stage involved spatial metrics which used site-specific data and label agreement. Reference or ground data were used to spatially build overall accuracies based on the percentage of cases correctly classified.

The final and most recent stage of accuracy assessment, as detailed in Foody (2002), is the refinement of the spatial metrics. It uses the widely popular error matrices, which show class proportions and agreements; it summarises the producer and user’s accuracy (per class) and gives an indication of overall accuracies, as well as overall and class Kappa statistic. The latter can be disaggregated into location and histogram components, as proposed by Hagen (2002). The above mentioned methods, from the most recent stage of accuracy assessment, are commonly employed for estimating the accuracy of classified LULC maps (Kahya et al. 2008, Yuan et al. 2008, Dewan and Yamaguchi 2009) and for comparing different mapping methods (Gao 2008, Berberoglu and Akin 2009, Zhou et al. 2009).

2.3 Land change modelling

2.3.1 Definitions and challenges of land change modelling

Land change modelling is an important tool to understand and predict future land patterns. It is most relevant if done in a spatially-explicit, integrated and multi-scale manner (Veldkamp and Lambin 2001), as it enables experimentation in gaining the understanding of processes, causes and consequences of land change.

The environmental applications of land change models range from monitoring biodiversity (Verburg et al. 2008) and vegetation losses (Echeverria et al. 2008), to modelling the environmental impacts of climatic change (Lasch et al. 2002). Land change models are
also used as learning tools (Lambin 2006, Schaldach and Priess 2008) and for policy development and decision-making (Beurden et al. 2007), amongst many other uses.

In their review, Verburg et al. (2004) defined priority issues in land change modelling, which include: i) level of analysis; ii) cross-scale dynamics; iii) driving forces; iv) spatial interaction and neighbourhood effects; v) temporal dynamics; and vi) level of integration. While these authors pointed out the increased complexity of the models, they called for new ideas in their area that would better address the multi-scale characteristics, quantify the neighbourhood effects, deal with the temporal effects and integrate different disciplinary modelling approaches. This, they asserted, would lead to better understanding of the land dynamics and associated land management decisions and policy formulation.

A key problem with land dynamics is that they are a consequence of multiple processes, occurring at different scales (Ju et al. 2005). Therefore, understanding all of the interactions between processes at different scales requires large (and often unavailable) amounts of input data. Furthermore, the driving factors of LULC change should be accounted for at each of these scales, which is difficult to attain at the right balance. Consequently, most models either over-simplify these interactions or require significant resources (financial, human, computational, amongst others) to attain plausible outcome (Koomen et al. 2008).

A number of the earlier approaches to land change modelling were somewhat generic, involving simple linear relationships, whereby any change in a variable would directly affect another. With the evolution of computing power, there have been considerable advances from these simple models to more complex ones, using statistical approaches, cellular automata, fuzzy logic and interconnectedness amongst different modelling methods.
In any land change model, the predictive power is determined by the amount and significance of the input data, while the outcomes are limited by the model’s complexity. To illustrate this, Agarwal et al. (2002) compared a series of nineteen land change models in terms of their spatial, temporal and decision making complexity. Figure 2.5 above, from Agarwal et al. (2002), illustrates this three-dimensional space of overall model complexity.

In explaining Figure 2.5, an ideal model would need to be: i) spatially complex, addressing the drivers acting at different spatial scales; ii) temporally complex, allowing for events to be predicted at multiple scales; and iii) able to handle complex decision-making strategies, encompassing deterministic and non-deterministic events. This “utopian” model, which has not yet been constructed, would be capable of testing our understanding of processes and specific variables that influence land change (Veldkamp and Lambin 2001).

Syphard et al. (2005) opined that, while most land changes (urban or agricultural) were a top-down phenomenon influenced by policies and environmental constraints, the change itself would emerge from the bottom-up, whereby patterns could appear from the local characteristics. Based on this premise, different modelling approaches have been proposed.
For instance, Irwin and Geoghegan (2001) described the evolution of models that integrate spatially-explicit land change models with economic processes. Dietzel et al. (2005) and others used multi-agent systems to describe and model land change based on an ecological or an urban perspective, while Hietel (2007) used statistics to model land changes based on socio-economic indicators.


Earlier, Thornton and Jones (1998) suggested models should be characterised by the minimum complexity possible. Such simple models would serve as an exploratory tool and to generate temporary hypotheses. They further emphasised that a simple model is generally preferred, as too much time is often spent on developing and calibrating models which are only applicable to a small region.

### 2.3.2 Classification of land change models

As discussed above, land change modelling has been undertaken by various research groups with different foci; they have been developed for variable uses, modelled at a variety of scales and with diverse disciplinary backgrounds (Castella et al. 2007). To understand these different land change models, their assumptions and their purposes, there are many reviews which discuss the strengths and weaknesses of various models and provide insight to modelling methods/approaches (Briassoulis 2000, Agarwal et al. 2002, Verburg et al. 2004, Koomen and Stillwell 2007, Pontius Jr et al. 2008, Schaldach and Priess 2008). However, a simple classification of land change models by Lambin et al. (2000), reproduced in Table 2.4 below, is more relevant here.
In Table 2.4, it can be deduced that the knowledge on the past patterns of LULC and its causes ("why") would allow for different models to be constructed. When prior knowledge is limited to the past patterns of LULC ("where and when"), questions related to the short-term change can be answered by stochastic, empirical and statistical methods of land change modelling. However, if prior knowledge associated to the causes of LULC change is available, then long-term change scenarios can be simulated. This can be achieved using mechanistic, economic, as well as process- and agent-based methods.

The distinctions above are focused on what was known and what needed to be known about land change, providing an indication of suitable methods/approaches to land change modelling. However, as there are many models, Koomen and Stillwell (2007) classified them based on their temporal characteristics (static or dynamic), simulation process (transformation or allocation) and simulation approach (deterministic or probabilistic). In explaining further, while static models are used to extrapolate past trends into the future, dynamic models are capable of modifying the trends during simulations. While transformation models use transition probabilities and/or neighbourhood effects to determine transitions, allocation models employ suitability, such as soil quality, for agricultural allocation and/or economic characteristics to determine LULC preferences. In contrast, deterministic models rely on rules
to gauge changes, whereas probabilistic models combine transition matrices and random numbers to determine change.

The two classifications of land change models described above are complementary. While Lambin et al. (2000) elucidated what questions could be answered based on the prior knowledge and which modelling method was suitable for answering those questions, Koomen and Stillwell (2007) decomposed the different approaches to modelling based on the models’ characteristics. Table 2.5 illustrates some available land change models and their modelling approaches.

Table 2.5. Some land change models and the methods employed by them (adapted from Koomen and Stillwell 2007)

<table>
<thead>
<tr>
<th>Model name</th>
<th>Economic principles</th>
<th>Spatial interaction</th>
<th>Cellular automata</th>
<th>Statistical analysis</th>
<th>Optimisation</th>
<th>Rule-based</th>
<th>Multi-agent</th>
<th>Micro-simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markov model</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial interaction</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetic algorithm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear programming</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GeneticLand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UrbanSim</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Multi-agent simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PUMA</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>DSSM</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LUMOS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOLAND</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLUE-s</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELES environment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land Use Scanner</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ProLand and UPAL</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As shown in Table 2.5 above, different models employ a variety of methods. Some models are limited to one method, such as Markov and spatial interaction, while others combine two or more methods (e.g. CLUE-s, LUMOS). The combination of two of more methods has produced many land change modelling approaches, however these different models are built with different objectives, thus a global solution is not possible (Veldkamp and Verburg 2004).
A recent collaborative research lead by Pontius Jr et al. (2008) compared nine commonly used, peer-reviewed, land change models. The characteristics of these models are summarised in Table 2.6 above.

The results of the work by Pontius Jr et al. (2008), indicate that despite much effort spent in building, calibrating and validating the land change models, all nine tested had poor predictive power, failing to outperform a null (no change) model. While the outcomes of this work (Pontius Jr et al. 2008) were surprising, it was not meant to discredit the land change modelling community, but concluded that models should be used with caution, as already pointed out by Verburg et al. (2006).

Furthermore, Table 2.5 and Table 2.6 elicit how model complexity is quite variable. Some land change models are based on one method (CLUE, Land Use Scanner, Logistic Regression and SAMBA), while others are more complex combining multiple methods (CLUE-s, Environment Explorer, Geomod, LTM and SLEUTH) for land change modelling.

From the discussion above, it can be seen that model complexity does not necessarily lead to higher predictive power (Pontius Jr et al. 2008). Therefore, parsimonious models which are flexible, allowing increasing complexity are preferred (Thornton and Jones 1998). Moreover, simpler models require less input data and are generally more flexible, serving as learning tools for students (Lambin 2006) and for unravelling the interrelations of land dynamics (Verburg et al. 2006).
Given the importance of LULC classification associated to the information which can be extracted from multi-temporal LULC maps; this limits the knowledge to “when and where”, as defined by Lambin et al. (2000). For this reason, in the remainder of this chapter, the focus will be on simpler models, such as Markov-chain and cellular automata land change models. Other models, such as those focused on economic theory and multi-agent models are not covered here, as they are outside the scope of this thesis.

### 2.3.3 Markov-chain models of land change

A popular tool used for land change modelling is the transition matrix based on the first order Markov-chains. The transition matrix, extracted from multi-temporal LULC maps, constitutes probabilities of change from one class to another (Urban and Wallin 2002). While a brief review of this model is provided here, more detailed treatment of the model is given in Chapter 5.

Lambin et al. (2000) classified Markov-chain models as stochastic, emphasising their usefulness for simulation of land change into the near future (Table 2.4). Baca (2002) integrated Markov models with other models to serve as a tool for understanding land dynamics, not only in the short term as suggested by Lambin et al. (2000), but also in the medium to long term.

Even though the concept of Markov-chains was developed in the early 19th century, an early application of Markov-chain models to understand land dynamics was carried out by Muller and Middleton (1994). In their study, they compiled five LULC maps at different times, from which they extracted four transition matrices for different time intervals. The transition matrices were then used, in conjunction with the LULC maps, to derive hypothetical “equilibrium” distributions of LULC classes for each of the four intervals. They found that urbanisation was the main driver of change in the first two time steps. This study is important in showing the proportional distribution of LULC classes over time, but made no prediction of the spatial distribution of the LULC classes.
As mentioned earlier, Thornton and Jones (1998) proposed a simple agricultural LULC model based on transition probabilities extracted from a transition matrix. They defend the simplistic approach based on the fact that land change models are usually tailored for specific areas or situations. Accordingly, time spent optimising a particular model of limited broad use is not justifiable. Furthermore, they opined that more complex models require large amounts of input data and time spent on calibration. For efficiency and cost effectiveness, the authors proposed a simplistic, empirical model, specific to a given situation, with complexity being added as needed. However, their proposed model was merely conceptual.

A more applied use of Markov-chains was developed by Lopez et al. (2001). They used multi-temporal transition matrices coupled to regression models to understand the causes of land change in a region of Mexico. The authors found that urbanisation was one of the main drivers of land change, followed by Agriculture. However, they concluded that the Markov-chain model performed poorly in predicting the location of land change. In spite of this, they pointed out that the Markov-chain approach was good as a descriptive tool rather than a predictive tool. This was a result of the transition matrix being static, therefore the transitions were random.

Weng (2002) reached a similar conclusion modelling land dynamics in a coastal region of China. This study used multi-temporal transition matrices to determine an “equilibrium” state, as was done by Muller and Middleton (1994). Again, the focus was on producing a simulation of the LULC class distributions without taking into consideration the spatial structure of the results. Weng (2002) is notable for recognising importance of integrating remote sensing images with land change modelling, which can be done effectively with Markov-chain models.

More recently, Wu et al. (2008) and Cabral and Zamyatin (2009) used the first-order Markov-chain method for land change modelling in a region of China and Portugal, respectively. In both studies, transition matrices were derived from multi-temporal LULC
maps and then used to forecast LULC patterns. Once again, the authors found that the predicted maps lacked spatial structure.

From what has been seen so far, a major limitation of Markov-chain models is that while they often correctly simulate the amounts of change, the spatial structure is lost (Muller and Middleton 1994, López et al. 2001, Wu et al. 2008). To overcome the lack of spatial structure found in pure Markov-chain models, Baca (2002) incorporated an algorithm which he called “border effect algorithm” to simulate LULC patterns while preserving spatial structure. The author’s approach stipulates the amount of change by the Markov-chain chain method (Muller and Middleton 1994), but the spatial allocation of the change depends on the pixel’s immediate neighbourhood. This was achieved by computing the position (in the landscape) of every neighbourhood combination and then allocating the changes randomly only at border pixels, so as to preserve the spatial structure. A limitation of this approach though, is that every time a single pixel changes, the neighbourhood combination tables has to be reconstructed. In effect, the method maintains the proportion of change as predicted by the Markov-chain model, but it can be very time consuming. Furthermore, the method only considers the immediate neighbourhood.

While the Markov-chain models are incapable of reproducing the spatial structure of LULC patterns, they have been used as tools for describing land dynamics and for land change modelling, in many studies (Muller and Middleton 1994, Baca 2002, Weng 2002, Wu et al. 2008, Cabral and Zamyatin 2009). Additionally, Markov-chain models have also been used in conjunction or as part of the well-established, integrated land change models, such as Environment Explorer (Engelen et al. 2003), DINAMICA (Soares-Filho et al. 2002) and CLUE-s (Verburg et al. 2002).

The main limitation of Markov-chain models, as previously described, is related to the lack of spatial structure. Baca (2002) proposed a solution to this problem; however some of its limitations were shown above. Another method used for land change modelling which
requires limited inputs, but is focused on neighbourhood interactions is based on the cellular automaton, the topic of the following Section 2.3.4.

2.3.4 Cellular automata and land change modelling

The idea of cellular automata (CA) began in the 1940s through the work of von Neumann and Ulam (Voorhees 1996). Von Neumann’s idea was one of making a self-replicating “robot” while Ulam suggested the use of mathematical abstraction (Kemp 2008). The first formal appearance of CA and its rules were presented in a series of papers by Wolfram (1983, 1984), describing the functioning of the system.

CA models are pattern-based and mechanistic. They require a landscape composed of a series of pixels covering a finite space, with values of each pixel and a series of rules determining the change. The rules should be local so that the change in each pixel depends only on their neighbouring cells, which characterises the deterministic CA method. Furthermore, the search neighbourhood needs to be the same for every pixel but should not be limited to any specific size (Voorhees 1996). Given an initial landscape and simple transition rules, where the state at time \( t+1 \) is dependent on state at time \( t \) and nothing else, complex patterns appear after a series of iterations, which justifies the use of this approach by the biological science community (Irwin and Geoghegan 2001). However, Balzter et al. (1998) noted that the CA rules do not need to be deterministic. They can also be stochastic, extracted from a transition matrix. In this case, there is a probability of change and changes do not rely solely on neighbourhood configuration.

From the land change modelling point of view, the deterministic CA method is not commonly used, since the change is not exclusively attributed to the neighbourhood alone. There are many applications of the stochastic CA in the urban environments. An example was given by Clarke et al. (1997). To understand the impact of urbanization in San Francisco’s climate, they proposed the use of CA with self-modifying rules and Monte Carlo methods for accumulating probabilistic rules of change. The expansion of this early model is the SLEUTH
(slope, land use, exclusion, urban extent, transportation and hill shade). Later, Silva and Clarke (2002) applied the SLEUTH model to Lisbon and Porto, in Portugal. Their calibration results showed that the patterns of change were well represented. SLEUTH has since been widely used for a variety of applications, mostly related to urban LULC change (Arthur-Hartranft et al. 2003, Goldstein et al. 2004, Jantz et al. 2004, Xian and Crane 2005, Xian et al. 2005, Al-Awadhi 2007, Mahiny and Gholamalifard 2007, Lin et al. 2008). SLEUTH has also been used to simulate the run-off and water quality in a watershed (Carlson 2004), as well as to model the possible effects of climate change in a multi-county study in the USA (Solecki and Oliveri 2004).

A land change model which uses a stochastic CA is the monitoring land use dynamics (MOLAND), developed in the late 1990s by a company called RIKS. Initially the model was developed to identify trends of urban expansion in the European context. However, since the mid 2000s it has been used to evaluate the occurrence of extreme climatic events and as a spatial planning tool (Engelen et al. 2007). MOLAND utilises input from LULC maps, suitability, zoning and accessibility maps, as well as socio-economic data. Through the use of CA, spatial interaction and transition rules, it forecasts future LULC patterns. Hoffman et al. (2002) used Ikonos images to develop LULC maps which were then input to MOLAND to identify trends of urban sprawl in Shanghai and Belgrade.

Another application of CA, to model agricultural expansion and deforestation, was reported by Walsh et al. (2008). Their results include scenario based patterns of LULC caused by the change in rural agricultural income, access to roads and off-farm employment. Thus the combined use of LULC maps and associated information led to the development of rules describing the neighbourhood and the socio-economic interactions. Although they also computed and used suitability layers which involved competition among the different LULC classes, they found that the model predictions were “weak” and emphasised the importance of
the influence of scale-effect, whereby changes would occur at the farm-level and not necessarily at the pixel level.

In the above mentioned models, emphasis was put on the importance of selecting the stochastic transition rules for CA. This could be done by directly extracting these rules from multi-temporal LULC maps (Geertman et al. 2007). In this approach CA was employed with generic, regional and time-specific rules. They (Geertman et al. 2007) found that the regional rules performed best in predicting land change, however they emphasised some limitations of the regional rules. The creation of such region-specific rules demanded a big effort, requiring a balance between the predictive accuracy and the time spent in developing them.

Another attempt to automate the selection of CA rules was undertaken by Liu and Li (2007) whereby a combination of Fisher discriminant and discrete selection models was used to automatically establish CA transition rules. The authors further compared their automated model to a logistic regression model for selecting CA rules. Their model was found to be superior because the transition rules had some “physical meaning”. The issue concerning rule generation was further elucidated by Hagoort et al. (2008), who suggested that neighbourhood rules should be derived from a series of complementary methods, such as understanding of land dynamics from the literature search, expert interviews and known spatial metrics.

An alternative and emerging hybrid approach to land change modelling combines the stochastic CA and the aforementioned Markov-chains to take advantage of both modelling methods. While Markov-chains can be used for determining the amounts of change, the CA can preserve the spatial integrity of the output maps. A number of authors have combined these two methods of the hybrid approach, which is covered in the following Section 2.3.5.

2.3.5 Hybrid, CA / Markov-chain land change models

An early approach that combined CA and Markov-chains was developed by Turner (1987). In their study, three different models were developed; the first was a first-order Markov-chain model (which was already covered in Section 2.3.3), while the other two, also
discussed here, incorporate CA using four neighbours and eight neighbours, respectively. Turner’s (1987) hybrid model uses a transition matrix extracted from LULC maps, calculates a transition index by multiplying the frequency of each neighbour by its transition probability. Furthermore, the amount of change can be determined by multiplying the transition probabilities by the frequency of each LULC class. In an iterative process, the cells with highest transition indices are changed until the total amount of change per class is exhausted.

Jenerette and Wu (2001) applied a similar method, although by generating the CA transition rules using a modified genetic algorithm. It was found that their hybrid model performed “reasonably” in reproducing LULC changes. In both cases (Turner 1987, Jenerette and Wu 2001), the issue of scale was highlighted and the output maps were spatially coherent, but did not take into account the scales at which the transition processes occurred.

Another model which integrates CA and Markov-chains is the CA_Markov built into IDRISI software. CA_Markov uses Markov-chains, CA and a multi-criteria allocation to determine where the land changes are likely to occur (Eastman 2006). Pontius Jr and Malanson (2005) used CA_Markov and Geomod (Hall et al. 1995) in a comparative study to assess both models’ performance. The authors used a study area in Massachusetts, USA and data from two time periods for calibration, then the trends extracted from this data was used to project LULC changes. Both models’ outcomes were poor and did not outperform a null-model at the resolution in which the simulations were run. However, they pointed out that the main advantage of the CA_Markov approach was that it was capable of dealing with multiple two-way transitions, while Geomod could only handle one-way transitions between two classes. CA_Markov was also used by Peña et al. (2005) to understand land dynamics in a region of Spain and to construct future scenarios. However, they did not detail the calibration accuracy, nor thoroughly discuss the scenarios. Modifications to the original CA_Markov model have been proposed by Lee et al. (2008), who indicated that the original model was limited as it only considered LULC maps from two points in time. They constructed a
modified CA_Markov that accounts for the LULC maps from four points in time, thus allowing for trends in the land dynamics to appear.

Heldens (2006) also developed a land change model which combines Markov-chain with CA. In the author’s approach, different neighbourhood combinations were grouped and empirical assumptions were made in relation to the transition process so that all neighbourhood configurations could be computed in an empirically derived transition matrix. Effectively, the author generated a spatially coherent, stochastic CA transition matrix. The outcome was the improved visual aspect of the output maps, but did not forecast much change in the landscape.

From the discussion above, two different hybrid approaches can be identified. The first approach uses the Markov-chain to model the amount of change, followed by the CA component used for allocating the change based on the spatial configuration of the landscape and/or suitability layers (Turner 1987, Jenerette and Wu 2001, Peña et al. 2005, Pontius Jr and Malanson 2005, Lee et al. 2008). The second approach, proposed by Heldens (2006), used an empirical transition matrix whereby the different neighbourhood combinations were empirically summarised. In both cases, the spatial coherence of the output maps was improved even though they applied different methods for determining the transitions.

To maintain the spatial coherence of the output maps, Turner (1987) used the transition index to identify where changes could occur, while in CA_Markov (Eastman 2006) the transition was derived from LULC maps (through suitability) and in Heldens (2006) it was derived empirically. However, it is well known that the transition rules should perform better if the neighbourhood interactions were modelled and combined with expert knowledge (Liu and Li 2007, Hagoort et al. 2008).

Unfortunately, there is no model in which the user is allowed to determine the influence of the Markov-chain and CA components independently. This would be important because, as shown in Section 2.3.2, the two modelling methods (Markov-chain and CA) perform better
for answering multiple research questions. The Markov-chain models are more descriptive tools and indicate future LULC class distributions (Section 2.3.3), while CA models are better at reproducing spatial patterns of change in a deterministic or stochastic way (Section 2.3.4). Thus it would be beneficial to have a more flexible land change model that could be used in Markov-chain mode or in CA mode or a combination of these, as set by the user.

2.3.6 Remarks on land change models

As shown in the sections above, there is an abundance of models and modelling methods. Some methods have been around for a long time, while others are yet in their development stage. The world is becoming more complex, in terms of economics (e.g. global markets) or biophysical conditions (e.g. climatic variability) and this complexity is being unravelled using an ever increasing, complex web of knowledge. Furthermore, with the increasing storage and computing power, it would be naïve to believe that a single model or modelling method/approach could suffice for explaining the interconnectedness of biophysical systems with human decision-making, to understand and predict LULC change.

In fact, models are mere representations of reality of the land system, serving as learning tools to understand land dynamics, human decision-making, as well as impacts of certain socio-economic policies and to aid in land planning. Verburg et al. (2006) agree with this premise, emphasizing the importance of utilising computers as virtual laboratories, where our understanding of the system could be tested and visualised.

The process of land change occurs not only from the bottom-up (emergence of patterns), but also from the top-down, where policies directly affect the land patterns. Therefore, research in land change modelling is becoming a global exercise with researchers from different fields interacting and integrating their models to make better predictions (Castella et al. 2007, Verburg et al. 2008). Such integrated models are now required to answer questions regarding social, economic and biophysical dimensions. An illustration of a multi-model approach is shown in Figure 2.6 below.
Verburg et al. (2008) applied a combination of two models (GTAP and IMAGE) to predict changes in land demand for countries of the European Union. In the same study, they also applied CLUE-s to transform these demands of LULC to patterns of LULC at 1-km resolution, at a national level. GTAP - a global economic model - and IMAGE - an integrated assessment model, were combined to provide estimates of demand for agricultural land by countries of Europe. The integrated model takes advantages of both: GTAP, being an equilibrium economic model, unable to deal with different demands, but that maximises a profit function for each country and IMAGE, which was used to estimate effects of LULC and climatic change on crop yield levels, as well as to calculate global environmental indicators. The output was then used by CLUE-s to allocate the LULCs at the national level with a relatively good (1-km) spatial resolution.

Other multi-model approaches include neural networks and CA used by Li and Yeh (2002) to model urban expansion; the LUMOS (Beurden et al. 2007), which uses Environment
Explorer (Engelen et al. 2003) as a CA spatial allocation model integrated with Land Use
Scanner (Hilferink and Rietveld 1999), which is an econometric model that determines land
demand; SAMBA by Castella et al. (2007), who integrated LUPAS (Roetter et al. 2005) and
CLUE (Veldkamp and Fresco 1996) to address LULC changes in the Bac Kan province of
Vietnam.

In spite of these multiple land change models, there is no consensus as to which is the best
model (Koomen et al. 2007) nor will there be one, as each model has its strengths and
weaknesses. This was better demonstrated by Pontius Jr et al. (2008) who compared the
outcomes of a number of well established land change models (See Table 2.6 on page 28).
Their study was based on the premise that no data post-calibration were used in the validation
phase for any of the models. Thus, calibration and validation would be done independently.
Furthermore, another requirement in their study was that the quantity of change for the
validation period should be extracted from the calibration data. The results of this multi-
model comparison are intriguing. The authors found that most models did not perform better
than a null model (no change). Nevertheless, a few models did outperform a null model, but
“cheated”, by incorporating the quantity of change from the validation period.

The findings of Pontius Jr et al. (2008) show that while much has been done in creating
and calibrating land change models, our understanding of these models is far from ideal. It
should be remembered that models are tools and as such, subject to errors and limitations.
This is why models’ source code should be published for transparency. This should be so that
the science community could collaborate in improving them.

A comment about land change models, relative to data availability and calibration, is
necessary here. As model complexity increases, so does the need for more (detailed) data,
resulting in the calibration task becoming more difficult. As an example, SAMBA, a multi-
agent model, has been used to predict and elucidate land change dynamics in Vietnam, but it
relies significantly on field data. Therefore applying this same model to other areas would
equally depend on field data collection. More general models such as ones based on CA and/or Markov-chains, on the other hand, rely simply on transition rules or probabilities, thus serving as adequate learning and exploratory tools (Verburg et al. 2006) which can be used globally.

2.4 Summary

In this review, the aim was to highlight the recent advances in the science of LULC mapping and land change modelling, but also to identify research gaps which should be addressed further. As pointed out, there is no consensus about what method is “best” for LULC classification and mapping or is there a single best land change model. Different approaches have their strengths and weaknesses.

In spite of this, historical LULC maps are necessary for monitoring and detecting change, depending on data availability (or lack of it). While pixel-based methods are limited by the “salt and pepper effect”, unsupervised methods are inadequate if expert knowledge is available. An object-based approach has the advantage of spatially/spectrally clustering similar pixels into objects which can be classified. Moreover, supervised methods associated to rulesets enable the use of contextual information in the classification process, increasing the quality of the resulting LULC maps. Additionally, expert knowledge can be used to augment these LULC maps, which in turn can be used for deriving transition matrices for land change modelling. Few studies have outlined a framework for producing multi-temporal LULC maps from medium-spatial resolution satellite images, which could be used for monitoring and land change modelling. This is one of the gaps identified in this review that needs further research and is covered in Chapter 3.

When multi-source images are available for a single year, LULC classification can be done using the concepts of image fusion and supervised classification. The combination of
these two methods would enable automation of the LULC classification while increasing accuracy in comparison with LULC maps derived from single-source images. Moreover, no studies have been identified which compare the different levels of image fusion for LULC mapping, while benchmarking the resulting (fused) maps against the single-source maps (no fusion), which is the topic of Chapter 4.

In terms of land change models, a multitude of methods have been employed (Table 2.4 and Table 2.5) and a number of models have been developed (Figure 2.5 and Table 2.6). Significant time and effort has been dedicated to creating, calibrating and validating these models, but it has recently been shown that the outcomes of these generally do not outperform a null model (Pontius Jr et al. 2008). For this reason and for parsimony, simpler models are preferred.

Markov-chain models of land change have shown to serve as adequate descriptive tools and capable of predicting LULC frequency class distributions, but they fail at preserving spatial patterns as each pixel is treated individually. CA models, on the other hand, can successfully produce spatially coherent output, but the transition is often only linked to the neighbourhood distribution. In any case, these two methods require little input data, namely an initial condition map and a transition matrix (for Markov-chain models) or a set of transition rules (for CA).

Most importantly, integrating Markov-chain and CA models, called the hybrid approach, allows answering multiple research questions. A few hybrid approaches have been developed, but they are limited by the amount of change based on the Markov-chain and spatially allocate the change based on suitability factors or over-simplified neighbourhood interactions through empirical rules. A model which is flexible enough and can be run in Markov-chain or CA mode or any combination of both has not yet been developed. Such a model would be more flexible for answering different research questions, related to frequency and spatial distribution of LULC classes and this is done in Chapter 5 and Chapter 6.
2.5 References


Lambin, E.F., 2006. Land use and land cover change local processes and global impacts Berlin; Heidelberg; New York: Springer.


Chapter 3. Object-based classification of multi-temporal satellite images: a case study of the lower Hunter valley, NSW

* Part of this work was presented at the 14th Australasian Remote Sensing and Photogrammetry Conference (14ARSPC) held in Darwin, NT, from the 29th of September to the 3rd of October 2008
3.1 Introduction

It is well known that local and regional land use and land cover (LULC) changes are a congruent part of global environmental change (Riebsame et al. 1994). According to Lambin (2001), LULC changes are so pervasive that they constitute the most important aspect of the Earth system functioning. Aggregation of local and regional environmental changes has contributed immensely to current global climate change, which makes it necessary to investigate environmental impacts of LULC changes and to determine potential feedbacks of the system through time (Dale 1997, Wentz et al. 2006). This is particularly important at a regional scale, once land utilisation itself is the major driver of local (Barson et al. 2000) and global changes affecting the Earth (Betts et al. 2007). It has, thus, become increasingly important that accurate LULC classification be achieved, in order to model LULC change characterised by high accuracy and low uncertainty.

In Australia, the post-European impacts of LULC change have led to environmental degradation in both rural and urban settings. The case study presented here covers a portion of the Hunter valley in east-central New South Wales, Australia, which has experienced LULC change more rapidly than most other parts of the country, especially in the last 40 years. This portion of the Hunter valley is characterised by rolling hills with grasslands, wineries and infrastructure for tourism. In fact tourism has become one of the main attractors to the area due to its proximity to urban centres and natural beauty of the Hunter valley. The region is very dynamic due to recent expansion of an already established wine producing area (since late 1800s), mining and increased infrastructure for tourism, attracting people from Sydney, Newcastle and overseas (McManus et al. 2000, Beer et al. 2003). Due to its dynamism, the area was selected for the study, as well as good relations with landowners in the area and an image database which would serve the study.
LULC classification methods have increasingly relied on remote sensing data, as recent advances in satellite technology enabled the acquisition of data on land surface repeatedly over time. Such satellite images have been available since the 1970s (Agarwal et al. 2002), with aerial photographs preceding them. While these data are an excellent sources for LULC classification, a major challenge lies within the transformation process that converts the raw image data into LULC information (Jensen 2005). In fact, in the beginning of the last decade Watson et al. (2000) pointed out that LULC classes are poorly enumerated, in spite of recent improvement achieved with Earth observation satellites. This statement is still largely true, especially when considered in the regional context.

This poor performance in LULC classification is due to two main factors: first, the lack of a generalised algorithm suitable for all situations and different LULC patterns; second and more importantly, due to lack of quality remote sensing data, limited by atmospheric conditions and LULC dynamics, especially when data is acquired off-season from actively crop-growing periods. The latter case often results in high noise-to-signal ratios, which makes it difficult to apply conventional classification algorithms.

Traditionally, classification has focused on pixels, where each pixel would be classified by any of the available algorithms (Barson et al. 2000, Lu and Weng 2007). The pixel-based approach is still in use today (Wang et al. 2008). In fact, while it has advantages, it also has limitations. The “salt and pepper effect”, for example, is a well known problem of pixel-based methods (Rogan et al. 2002, Zhou et al. 2009). It occurs when, one or a few pixels within an image have significantly different spectral signature than its neighbours, causing assignment of a different class to single pixels within the image.

As an alternative to classifying individual pixels, the object-based approach aims at clustering pixels into meaningful objects, which can then be classified. This spatial/spectral clustering of pixels is often referred to as “segmentation” (Benz et al. 2004, Lu and Weng 2007) and clusters of pixels referred to as “objects”. Even though the concept of object-based
classification was idealised a few decades ago (Mason 1988), it has only become mainstream in the last decade or so, due to the increase in computing power (Aplin and Smith 2008).

One advantage of the object-based approach is that it is not affected by the “salt and pepper” effect, as segmentation is responsible for delineating homogeneous and contiguous objects. If the objects are properly extracted, boundaries of fields, as well as dams and edges of forests, will be identified. The objects can then be classified into proper and meaningful classes (Schiewe et al. 2001, Jobin et al. 2008).

Classification of objects can be achieved utilising the same algorithms as in the pixel-based approach. This may employ an unsupervised approach, ruleset based algorithms (thresholding) (Barlow et al. 2003, Lucas et al. 2007), as well as supervised methods, including maximum likelihood or nearest neighbour classifiers (Qian et al. 2007, Gao 2008), with an advantage that the smallest unit of analysis is composed of contiguous objects and not of individual pixels.

Another advantage of the object-based approach is that it enables multi-scale analysis, where LULCs which occur at different spatial scales are more easily identified. For instance, segmentation can be done at various hierarchical levels, allowing the identification of forests on a broad scale, or of roads on a smaller scale (Lucas et al. 2007).

The delineation of objects through segmentation, associated with the multi-scale analysis aids in minimising errors typically caused in the traditional, pixel-based approach. Furthermore, segmentation at various levels enables the identification of large and small features, irrespective of high noise-to-signal ratio in the data.

Various studies have shown superiority of the object-based approach. Mitri and Gitas (2006), for instance, compared the object-based approach with a pixel-based method for burned areas. The above mentioned authors found that classification applied to image objects improved the overall accuracy by over 18%. Aplin and Smith (2008), showed how object-based classification aided in enhancing the accuracy of the UK’s national LULC map, while
Jobin et al. (2008) utilised object-based classification to identify potential habitat areas for an endangered species.

With this in mind, it was decided to utilise the object-based paradigm to produce a time-series of LULC maps, which would be useful for characterising change. Moreover, transition matrices extracted from these maps will be used for subsequent LULC change modelling (Chapters 5 and 6).

The main objective of this chapter is to develop an object-based classification model, in order to create a time-series (1972, 1985, 1995, 2000 and 2005) of classified LULC maps for the study area. The accuracies of these maps will be assessed and other support information used in order to enhance their overall quality. This is done to ensure that the final LULC maps are highly representative of past patterns, thus enabling LULC change modelling that is undertaken subsequently in chapters 5 and 6.

In summary, the aims are to:

i) determine the spatial LULC patterns in the lower Hunter valley at selected time intervals since 1972; and

ii) assess the accuracy of object-based classification of multi-temporal Landsat images (using Definiens Professional 5.0), as well as subsequent knowledge-based map enhancements.

3.2 Data and methods

3.2.1 Study area

Throughout this thesis the focus will be on one study region: the lower Hunter valley of NSW, Australia. The methods proposed here are of much broader use and the choice of this region was related to the importance it has for the state as a wine producer but more importantly as a tourist destination, which has caused rapid changes to the landscape in recent
decades. Furthermore the area was chosen due to data availability, as well as accessibility for visitation and a well established contact with stakeholders.

The lower Hunter valley is about 160 km north of Sydney. It is Australia's oldest wine producing region, inextricably linked with tourism. Since the mid 1960s, tourism has greatly developed, due to wineries, touristic infrastructure such as golf courses and parks, as well as the construction of a freeway which has facilitated the access to Sydney and Newcastle.

According to Geoscience Australia (2004), the pre-European vegetation of the region was composed predominantly of tufted grasslands, with about 10-30% natural cover of medium size (10-30 m) eucalyptus. From the 1800s onwards, the region has developed with the establishment of wineries and mines. In fact, significant human impact in the last decades has characterised the Hunter as one of the fastest growing regions of Australia (Beer et al. 2003).

Cessnock City Council released a report (Cessnock City Council 2002), in which they estimated the number of overnight tourists to be 1.5 million for the year of 1998-1999. More recent figures by the Hunter Valley Wine Country (2008), however indicate that this has increased to over 2.5 million overnight tourists in the year of 2007. Daily visitation to the Hunter has increased steadily over the last few years, from 4.4 million day visitors in 2001-2002 to over 6.5 million in 2007-2008 (Hunter Valley Research Foundation 2009).

In this sense, recent developments aimed at accommodating such flux of people have brought major changes in LULC, resulting in conversion of traditional grazing and natural landscapes to vineyards, urban development, tourist resorts and golf courses. LULC changes vary from single (e.g. grazing) to multiple LULCs (e.g. viticulture and tourism). Field observations and interviews with landowners have shown some LULC back-transformation, e.g., vineyards being reversed to pasture, especially so in the mid 1980s fuelled by government subsidies.

Figure 3.1 below, illustrates the study area: the south-western portion as being characterised by a mountain range and the north-western area comprising a military training
facility set up many decades ago. Most of the area is comprised of rolling hills from north to south. The study area is delimited by Cessnock on the south-eastern edge and encompasses around 287 km$^2$. It is characterised by a sub-temperate climate, with annual average temperature range of 11º C to 24.6º C and annual mean rainfall of 800 mm.

Figure 3.1. Region of study projected to its location in Australia (true colour composite of 1998 orthophoto)
As the region has undergone fast-paced LULC change within the last 40 years, it is important to understand these LULC dynamics. For this purpose, determining past LULC patterns is necessary.

3.2.2 Acquisition of satellite images and aerial photographs

A time-series of Landsat images was provided by the Australian Greenhouse Office (AGO) in Canberra, ACT. The necessary pre-processing steps, such as, orthorectification, radiometric normalisation and spatial resampling (as described in Jensen 2005) were done by the AGO. The detailed procedures are explained in Furby (2002), but the spatial resampling was done so that pixels from AGOs time-series of images fell within the same grid. Radiometric resampling was done to match the baseline histogram of the 2000 images that the AGO used for their carbon accounting studies. The Landsat Thematic Mapper (TM) images were resampled to 25 m spatial resolution and provided with six bands, while the Landsat MultiSpectral Scanner (MSS) images were characterised by four bands and resampled to 50 m spatial resolution. Images from Landsat TM and MSS were provided with 8-bit radiometric depth and no thermal band.

From the available time-series, selection was based on image quality, which fitted the time sequence roughly at decadal interval: 1972, 1985, 1995, 2000 and 2005. Images for 1972 and 1985 were acquired by MSS sensor on board the earliest of the Landsat series. The other images, namely, 1995, 2000 and 2005 were acquired from Landsat TM sensor and two scenes from different seasons were made available.

Initially, it was intended to have images at five year time intervals, starting from 1975. However this proved to be very difficult. The 1975 image, for instance, had cloud cover, while the 1990 images had some strips that were result of the pre-processing done by AGO. These issues could not be solved and therefore images had to be excluded from analysis. An illustration of these problems can be seen in Figure 3.2 a) where the red circles show clouds and shadows. Figure 3.2 b) at its turn, illustrates some banding that occurred.
Further to the Landsat images, a series of aerial photos were acquired from the NSW Department of Lands. Hard copy colour aerial photos, at a 1:25,000 scale, from 1976 and 1990 were available; however they did not cover the whole study area. Black and white aerial photos for 1984, at a 1:40,000 scale, were also available. Digital colour orthophotos were available for 1998 and 2004, although the latter did not cover the whole study area. These digital orthophotos were provided by the NSW Department of Lands, scanned to 1 m spatial resolution. They comprised a 3-band true colour composite, taken with a Wild RC30 camera and scanned using a Vexcel Ultrascan 5000 photogrammetric scanner and orthorectified by the NSW Department of Lands using a 25 m DEM (personal communication, NSW Lands, 2009).

The hard copy aerial photos (1976, 1984, 1990) were scanned and georeferenced (RMS error < 2 pixels) to match the location of digital orthorectified aerial photographs for 1998 and 2004.
3.2.3 Image segmentation and classification scheme using Landsat images

The classification was done in Definiens Professional 5.0 (Definiens 2006) and individual projects were created for each year of the study. The 1972 and 1985 projects were comprised of a single Landsat MSS image each, while the 1995, 2000 and 2005 projects had two images Landsat TM, from different seasons which covered the whole study area.

The first step for object-based image analysis is segmentation (Baatz and Schäpe 2000). Segmentation was based on a bottom-up merging algorithm, starting with single seed pixels and later merging them with similar neighbours, while minimising heterogeneity of resulting objects. Heterogeneity, as defined by Benz et al. (2004) is a function \( f \) of:

\[
Heterogeneity = f \left( w_{\text{colour}} \Delta h_{\text{colour}} + w_{\text{shape}} \Delta h_{\text{shape}} \right)
\]

\[
w_{\text{colour}} \in [0,1], w_{\text{shape}} \in [0,1], w_{\text{colour}} + w_{\text{shape}} = 1
\]

Equation 3.1

where the spectral characteristics of the objects are represented by “colour” and the spatial characteristics are represented by “shape”. Moreover, the user can define the weight given to the spectral and spatial components by changing “w”. A threshold for heterogeneity is an abstract measure defined by the scale parameter. Segments are allowed to grow until they reach the heterogeneity threshold, when “\( \Delta h \)” is larger than the allowable threshold. An in depth explanation of the algorithm can be found in Benz et al. (2004) and Figure 3.3 illustrates a decomposition of the multi-resolution segmentation algorithm.
For all projects, *shape* and *compactness* parameters were kept constant; however segmentation was done with various scale parameters, to enable identification of features at different scales. This was only possible due to the nested structure of segmentation (Zhou and Troy 2008), where boundaries of smaller objects were limited by boundaries of larger objects. A variety of scale parameters were used for each year, Table 3.1 shows the resulting number of objects according to the scale parameter and year.

### Table 3.1. Number of objects, for each scale parameter, for each year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>-</td>
<td>-</td>
<td>46</td>
<td>83</td>
<td>88</td>
</tr>
<tr>
<td>70</td>
<td>-</td>
<td>-</td>
<td>79</td>
<td>145</td>
<td>134</td>
</tr>
<tr>
<td>50</td>
<td>46</td>
<td>35</td>
<td>169</td>
<td>306</td>
<td>256</td>
</tr>
<tr>
<td>40</td>
<td>67</td>
<td>47</td>
<td>269</td>
<td>447</td>
<td>428</td>
</tr>
<tr>
<td>30</td>
<td>118</td>
<td>96</td>
<td>481</td>
<td>757</td>
<td>715</td>
</tr>
<tr>
<td>20</td>
<td>243</td>
<td>225</td>
<td>1039</td>
<td>1663</td>
<td>1467</td>
</tr>
<tr>
<td>10</td>
<td>903</td>
<td>825</td>
<td>3619</td>
<td>5938</td>
<td>5292</td>
</tr>
<tr>
<td>7</td>
<td>1805</td>
<td>1620</td>
<td>6841</td>
<td>11287</td>
<td>10079</td>
</tr>
<tr>
<td>5</td>
<td>3308</td>
<td>2936</td>
<td>12847</td>
<td>21004</td>
<td>18656</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>19567</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>21773</td>
<td>18098</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
As stated in Definiens (2006), scale parameter defines the maximum standard deviation of the homogeneity criterion in relation to the weighted layers used for segmentation. The direct implications of this were that higher scale parameters resulted in higher tolerance for heterogeneity, producing larger objects than when using lower scale parameters, as illustrated by Zhou and Troy (2008).

Considering the images from Landsat MSS (50 m pixels) and TM (25 m pixels), the lower-resolution Landsat MSS images of 1972 and 1985 should have fewer objects at the same scale parameter in comparison to higher spatial resolution Landsat TM images. In this sense, while the 1995 image had 46 objects with a scale parameter of 90, the segmentation of the 1972 image had the same number of objects with a smaller parameter of 50 (Table 3.1). This multi-scale approach was used to identify land features at broader and finer scales; however the final maps were produced utilising a small scale parameter as to accommodate the small features within the images.

Following segmentation, images were then classified utilising either a ruleset based approach (Li and Chen 2005, Civco et al. 2006) for 1985, 2000 and 2005 or a combination of a ruleset based approach with supervised nearest neighbour classification (Hubert-Moy et al. 2006, Laliberte et al. 2006) for the 1972 and 1995 images. The combined method is discussed in detail in the following section.

It is important to point out that the classification scheme was based on the Australian collaborative land use mapping programme or ACLUMP (Bureau of Rural Sciences 2006). The choice of the ACLUMP was made in order to allow integration with LULC maps produced by government agencies, as suggested by Anderson et al. (1976).

Five broad LULC classes were used, pertaining the first hierarchical level of the ACLUMP as: “woodland”, “dryland Agriculture” (including pasture), “irrigated Agriculture” (including vines), “intensive uses” (including urban areas and roads) and “water bodies”. The
classification was set at the first level of the ACLUMP, since Landsat images were only adequate for classification at a broader level, as stated by Anderson et al. (1976).

3.2.4 LULC classification steps

As mentioned previously, image classification was done exclusively based on rulesets, or on a combined approach, utilising a supervised nearest neighbour algorithm associated to the rulesets. The combined approach was used for the 1972 and 1995 images, while the rulesets were applied for 1985, 2000 and 2005.

The image classification and correction procedures are demonstrated by the overall workflow represented in Figure 3.4. The first part of the workflow resulted in “stage 1 maps”, produced with no manual intervention. The “stage 1 maps” illustrate the accuracy achieved by a generalised approach, which can be used in other areas and can be automated to produce LULC maps for large regions. However, since the objective of this chapter was to generate these LULC maps as accurately as possible, aerial photo interpretation (API) was used to edit the “stage 1 maps”. This manual editing was very time consuming, but aided in successfully increasing the accuracy of the LULC maps and the final result of API was the “stage 2 maps”. The “stage 2 maps” were subsequently taken to the field for verification and editing by landowners, resulting in the “stage 3 maps”, which were used as the reference for land change modelling done in Chapter 5 and Chapter 6. The “stage 3 maps” were also used to assess the accuracies of the maps produced by the fully automated approach (stage 1) and with API (stage 2).
In order to produce the “stage 2 maps”, the “stage 1 maps” were overlain on the orthophotos. These overlain orthophotos were subsequently used with API for manual refinement, resulting in the “stage 2 maps” (Figure 3.4). This was done to minimise errors introduced by the rulesets or the combined methods. Based on this procedure, the following aerial photos were used: the 1976 aerial photos, as the nearest available photos in the 1970s, served as reference for the 1972 LULC map; the 1984 aerial photos, for 1985 LULC map; 1990 photo, for 1995 LULC map; 1998 orthophoto, for the 2000 map; and finally the 2004 orthophoto was used for the 2005 map. While it was desirable that the Landsat and aerial photos images were acquired at periods as close as possible to compensate for changes in LULC (Jensen 2005), as described earlier, some Landsat images had pre-processing issues and could not be used. Furthermore, the aerial photos were the closest representation of reality that cover the whole extent of the study area and that could be used for validating the LULC maps.
Even though the “stage 2 maps” were more accurate than “stage 1 maps”, visits to the area and consultation with landowners were carried out in order to further edit and verify them. Hardcopies of the “stage 2 maps” and those of the aerial photos were produced on A0-size format. The hardcopy maps were then used in the field, in consultation with the landowners, to determine veracity of the LULC maps. This process was initiated with a conversation involving two landowners that agreed to look at the maps and suggest corrections, these two landowners also agreed to consult with other farmers in the area and through phone conversations a date was set for a meeting with the whole group.

In this meeting, four landowners were present and looked through each of the maps as a group, discussing property boundaries and clarifying specific doubts of LULC classification. In this meeting, a number of suggestions were made to alter the maps, however the landowners’ knowledge was limited to about 70% of the geographic extent of the study area. In subsequent visits to the study area (about nine), a number of conversations were undertaken with locals. In these conversations informal questions were made and annotated which also impacted the final maps.

The input from all of the people that were consulted (about 10 in total) included, but was not limited to reshaping objects, correcting classifications and identifying small dams and areas of development. The changes were annotated on A0 and digitised later.

The four landowners that participated in the initial meeting, who lived in the area since the early 1950s and 1960s, were consulted because of their extensive knowledge. They were also very committed to research and willing to share knowledge. As for the other locals who were consulted, these meetings were more informal, but they had knowledge that was more spread though the area, which covered about 80% of the study’s extent.

The “stage 3 maps” (Figure 3.4) were based on this verification approach and as mentioned earlier, were used as our reference maps for accuracy assessment and for land change modelling.
3.2.5 Detailed explanation of the ruleset based classification and supervised nearest neighbour

Whilst some of the “stage 1 maps” (1985, 2000 and 2005) were produced based exclusively on the ruleset based approach (Li and Chen 2005, Civco et al. 2006), the 1972 and 1995 maps were produced using the combined ruleset based and supervised nearest neighbour classification approach (Hubert-Moy et al. 2006, Laliberte et al. 2006).

The ruleset based approach utilised different parameters for each year, however the same sequential workflow was maintained. For instance, “woodland” recognition was done by thresholding the normalized difference vegetation index (NDVI) values or brightness at a relatively coarse scale, then synchronising the classification at the finer scale.

As discussed previously (Section 3.2.3, Table 3.1), the final 1972 map was produced with a scale parameter of “2”. In the specific case of 1972, objects with NDVI > 0.5 at level “30” were classified as “woodland” and then synchronised with objects of level “2”, utilising the class related features within the software. The practical implication of this rule was that any object at level “2”, which had a super-object classified as “woodland”, would inherit the class. This procedure of classifying at a higher level then synchronising with lower level objects was important for classification of “woodland”, as well as “dryland Agriculture”, “intensive uses” and “irrigated Agriculture”, even though each of the LULC classes used different features and thresholds.

For “woodland” classification, this procedure identified large patches of forest. However, it failed to delineate smaller patches. In the cases of 1972 and 1995 outputs, this was solved through the nearest neighbour classification. For the other years, NDVI thresholding (NDVI > 0.35) at a lower level, assisted in “woodland” identification.

Since the spatial resolution of the 1995, 2000 and 2005 images was higher, it was possible to identify smaller patches of “woodland”, such as riparian vegetation. The identification of the latter was done by identifying the thin, long objects (compactness > 4) and assigning them
to a temporary class. The objects of the temporary class with low NDVI (< 0.2) were identified to be roads and classified as “intensive uses”, while objects with higher NDVI were identified as riparian vegetation, a component of “woodland”.

Identification of “water bodies” was done by NDVI thresholding at the final map level (illustrated in Table 3.1, ~19000 objects), where objects with NDVI < 0 and little texture were classified as “water bodies”.

The LULC “irrigated Agriculture”, mainly composed of vines, was one of the most challenging areas to identify. For the lower spatial resolution images, such as 1972, classification was done through the nearest neighbour classifier, while in 1985 they were classified utilising object texture. For the other projects (1995, 2000 and 2005), with higher spatial resolution data and two images per year, it was possible to calculate the change in NDVI between two scenes. NDVI was defined as:

\[ NDVI = \frac{NIR - Red}{NIR + Red} \]  

and the difference as:

\[ \Delta NDVI = NDVI_{t2} - NDVI_{t1} \]

Since the multi-temporal images were taken within the same year, one during actively growing season of vines and the other during their period of dormancy, “irrigated Agriculture” areas could be identified where the difference in NDVI was high (\( \Delta NDVI > 0.3 \)), also using the synchronisation with higher level objects and band thresholding.

Due to its mixed spectral nature, “intensive uses” were quite difficult to extract, as they included golf courses, small houses throughout the study area, the city of Cessnock and its surroundings. Spectral signatures of the different LULCs were obviously not homogeneous and thus difficult to classify. Identification of large patches of “intensive uses” was done by NDVI and band thresholding at higher levels, followed by synchronisation at lower levels, as described previously.
Smaller patches of “intensive uses” were also identified by NDVI and band thresholding, but only at the final map level. Nevertheless, thresholds varied from year to year. As mentioned previously, long and thin objects with low NDVI were classified as “intensive uses”. For 1985, 2000 and 2005 projects, objects left unclassified after the procedures described above were then assigned to “dryland Agriculture”.

The use of NDVI was justified as it assisted in differentiating LULC classes, so that low NDVI represented water, bare soil, roads and buildings, while higher, but invariant NDVIs could be used to identify woodland areas and NDVIs that varied within season could show where crops were being grown (e.g. winter with low NDVI and summer with high NDVI would imply an annual or deciduous crop).

Subsequent to classification of objects by a ruleset based approach, for 1972 and 1995 projects, the nearest neighbour (NN) classifier (Singh et al. 2001), built-in the software, was used. Training samples were collected for each class utilising field based knowledge and the aerial photographs to represent each LULC as accurately as possible. The NN algorithm utilised a user-defined feature-space where it calculated the distance of each unclassified object to the centroids of the class samples. Following this, the algorithm then assigned fuzzy membership values to each of these objects, where higher probabilities were assigned to objects “closer” to class centroids (Definiens 2004). This allowed for post-classification enhancements.

With ruleset based and combination of ruleset and supervised NN approaches, LULC maps were generated for 1972, 1985, 1995, 2000 and 2005. These comprised the “stage 1 maps”, which would be enhanced by API and knowledge based editing generating “stage 2 maps”. Both of these were compared to the reference “stage 3 maps” and their accuracies assessed.
3.2.6 Accuracy assessment

Accuracy assessment is an essential part of LULC mapping, as it is the final stage of quality assurance. In this case, accuracy assessment was done by producing an error matrix (Congalton 1991) with standard measures of accuracy, utilising the freely available Map Comparison Kit (MCK) (Visser and De Nijs 2006).

As mentioned earlier, “stage 3 maps” were used as reference, as they were field validated and edited according to the landowners’ input.

The reference maps were compared to the “stage 1 maps”. This was done to assess the accuracy of the two automated approaches proposed here: firstly, the ruleset based approach used for 1985, 2000 and 2005; and secondly, for the combined ruleset and supervised NN used for 1972 and 1995. These comparisons would give an indication of the accuracies of the automated approaches and could also indicate whether the ruleset based or the combined approach was more appropriate. This comparison would provide a quality estimate of the “stage 1 maps”.

Subsequent to this comparison, the API and knowledge based enhanced maps, namely “stage 2 maps” and were also compared to the reference. These comparisons helped in understanding errors introduced by the map producer, as editing was based on the overlay of geo-referenced aerial photos or orthophotos with the “stage 1 maps”.

The accuracy statistics that were calculated were: overall, producer’s and user accuracies, as described in Jensen (2005). Overall Kappa statistic as described by Cohen (1960) and its decomposition into location and histogram Kappa. Location kappa is used as an indicator of spatial agreement (spatial structure), while histogram kappa is an indicator of class distribution agreement (frequency/proportion of each class) (Hagen 2002).
3.2.7 Spatial resolution issues

The Landsat MSS images were originally acquired with a spatial resolution of 80 m, while Landsat TM images were acquired at 30 m spatial resolution (USGS and NASA 2005). The images utilised here had a 50 m pixel for Landsat MSS and 25 m pixels for Landsat TM data. As described in Section 3.2.2, the resampling, as well as other pre-processing steps were undertaken by the AGO and the procedures were detailed in Furby (2002).

If our focus were to look at individual LULC maps and no inference made from them, there would be no spatial resolution issues. However the maps produced here were to be used in a land change model (Chapter 5 and Chapter 6). Downscaling the 50 m maps to 25 m was not recommended, as no information would be gained (Turner et al. 1989) and certain LULCs would be overestimated (Aplin 2006). The preferred alternative was to upscale the 25 m maps to 50 m, for comparisons.

Upscaling was done by a majority rule algorithm and proportions of each LULC computed. Even though it was known that downscaling would not add any information to the process, it was also done. The proportions of LULCs at 25 m and at 50 m were compared to verify the effect of upscaling and downscaling on the proportions of each LULC class.

Borrowing from the domain of landscape ecology, fragmentation statistics were calculated utilising Fragstats 3.3 (McGarigal et al. 2002). Mean fractal dimension and its standard deviation were used as indicators of shape complexity and so was perimeter-area fractal dimension (Turner and Ruscher 1988). Simpson’s diversity index indicated the probability of any two random pixels being part of different classes (Lasch et al. 2002). These statistics were calculated for all years with 25 m and 50 m pixels. It was assumed that if the difference between these metrics were small, spatial resolution would not be the main issue of concern.

Traditional, pixel-based, approaches have pixels as the smallest unit of analysis (Atkinson and Curran 1997). A single pixel could be assigned to any class and, therefore, a difference in image resolution would directly affect analysis. When single pixels are classified differently
from their neighbours, this characterises the previously mentioned “salt and pepper effect” (Rogan et al. 2002, Zhou et al. 2009).

On the other hand, the object-based approach has individual objects as the smallest unit of analysis and classification. When segmentation is done, if the scale parameter is adjusted, images of different resolutions can yield similar number of objects, which should minimise any spatial resolution issues. The approach used here aimed at maintaining similar numbers of objects, regardless of the image’s spatial resolution, with the purpose of minimising errors caused by differences in spatial resolution.

3.3 Results and discussion

3.3.1 Classification accuracy assessment: comparison of “stage 1 maps” with the reference maps

Since the “stage 1 maps” and the reference covered the same area, a direct map comparison was made. The results are illustrated in Table 3.2.

| Table 3.2. Accuracy assessment of “stage 1 maps” versus reference |
|------------------|-----|-----|-----|-----|-----|
| Overall accuracy (%) |
| Kappa          | 0.53| 0.51| 0.57| 0.50| 0.56|
| Kappa location | 0.56| 0.60| 0.70| 0.66| 0.68|
| Kappa histogram| 0.95| 0.85| 0.82| 0.75| 0.82|

User’s accuracy (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
<td>74.92</td>
<td>78.52</td>
<td>72.34</td>
<td>85.30</td>
<td>82.15</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>76.23</td>
<td>78.99</td>
<td>79.76</td>
<td>68.09</td>
<td>68.07</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>47.18</td>
<td>25.76</td>
<td>62.82</td>
<td>40.54</td>
<td>70.21</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>18.29</td>
<td>84.28</td>
<td>39.35</td>
<td>69.77</td>
<td>64.23</td>
</tr>
<tr>
<td>Water bodies</td>
<td>70.09</td>
<td>50.90</td>
<td>47.88</td>
<td>96.98</td>
<td>63.70</td>
</tr>
</tbody>
</table>

Producer’s accuracy (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
<td>80.66</td>
<td>84.51</td>
<td>94.66</td>
<td>64.86</td>
<td>81.33</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>71.41</td>
<td>66.99</td>
<td>73.78</td>
<td>87.59</td>
<td>84.80</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>48.84</td>
<td>44.70</td>
<td>21.49</td>
<td>26.43</td>
<td>29.51</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>19.05</td>
<td>48.61</td>
<td>38.75</td>
<td>41.96</td>
<td>33.41</td>
</tr>
<tr>
<td>Water bodies</td>
<td>31.65</td>
<td>48.46</td>
<td>34.70</td>
<td>19.36</td>
<td>26.36</td>
</tr>
</tbody>
</table>
As shown, overall accuracies vary from 70% to about 74%. Literature suggests that these results are in the normal range. As an example, while Gao (2008) reports an overall accuracy of 70.6% for a pixel-based classification of ASTER images, Corbane et al. (2008) report accuracy of between 63% and 84% for a multi-scale object-based approach utilising aerial photos. Laliberte et al. (2006), compared a decision-tree method to a nearest neighbour classifier with Quickbird images and found overall accuracies of 80% and 77%, respectively. Higher accuracies for object-based approach have been reported (Laliberte et al. 2006, Zhou et al. 2009), but are usually produced from images of higher spatial resolution images than those of Landsat used in this study.

Overall Kappa statistics indicate moderate agreement (Viera and Garrett 2005), varying from 0.50 to 0.57. These values are not high, but when decomposed into location Kappa and histogram Kappa values, the histogram Kappa value was found to be high. While the location Kappa indicates an agreement of spatial allocation of LULCs, the histogram Kappa is an indicator of quantitative similarity (Hagen 2002). This shows that, while the maps produced by the automated procedure, namely “stage 1 maps”, showed very high agreement in terms of class quantities (histogram Kappa of 0.75 to 0.95), they had only a moderate agreement in terms of their spatial allocations (location Kappa of 0.56 to 0.70).

The less than good performance of the ruleset and combined approaches, in terms of their spatial allocations of LULC classes, is due to the relatively low spatial resolution of the Landsat images used, as pointed out by Currit (2005). It is also due to inadequate sample selection, especially for the 1972 and 1995 images, caused by the mixed pixels (Fisher 1997, Cracknell 1998, Bastin et al. 2002). Nevertheless, the producer and user’s accuracies for the most common LULC categories (“woodland” and “dryland Agriculture”) varied from 65% to 95%. This means that the ruleset classifier effectively identified these two categories, but had difficulties in identifying the less frequent categories, such as “water” and “intensive uses”.
Another cause of relatively low Kappa statistics is due to the dates of image acquisition. While the aim was to select images acquired on anniversary dates as suggested in Mather (2004) and Jensen (2005), in the present study this was limited by the non-availability of images to coincide with the actively growing season of grapes in the Hunter valley.

An alternative to overcome the limited spatial resolution of Landsat MSS and Landsat TM, while increasing the overall accuracies, would be to fuse these images with higher resolution orthophotos (1998, Colditz et al. 2006, Wentz et al. 2006). This would possibly increase the overall accuracy and Kappa statistics of the resulting maps, while minimising manual labour. This approach was applied on two different subset areas, for two different years and its procedures and results are the topic of the following chapter (Chapter 4).

![Figure 3.5. a) 1995 ruleset based and supervised NN LULC map; b) 1995 reference LULC map](image)

The 1995 LULC map produced by the combined ruleset/NN classifier (Figure 3.5 a) and the reference map (Figure 3.5 b) are visually compared in Figure 3.5. While the combined approach correctly identified much of the predominant LULCs (“woodland” and “dryland...
Agriculture”), it incorrectly identified many small patches of “irrigated Agriculture”. It should also be noted that the extent of “irrigated Agriculture” areas were very much underestimated by the combined approach, in comparison to the reference map.

Figure 3.6 illustrates an example of the ruleset approach and reference for 2000. Figure 3.6 a) shows that the ruleset based approach correctly identified most of the “woodland” areas, much of “dryland Agriculture” and “intensive uses”. However, in the latter case, it failed to correctly identify the airport and golf courses just north of Cessnock (as shown in Figure 3.6 b). The location of vine plantations was generally correctly identified, but the amount was largely under-identified.

![Figure 3.6. a) 2000 ruleset based LULC map; b) 2000 reference LULC map](image)

An attempt to solve this problem of underestimating the area of “irrigated Agriculture” utilised a ruleset based algorithm (not shown here). All of the “irrigated Agriculture” areas were merged and small patches were reclassified (area < 10 ha) to “dryland Agriculture”. Then, what was left of the “irrigated Agriculture” was iteratively “grown” utilising
neighbourhood functions (i.e. if the shared border was above a certain threshold, then it was reclassified to “irrigated Agriculture”). This attempt failed, as some of the incorrectly identified areas “grew” too much, especially those located in the north-western part of the image. For this reason, these results were not included in the present study.

3.3.2 Classification accuracy assessment: comparison of “stage 2 maps” with the reference maps

The “stage 1 maps” were manually edited utilising API and expert knowledge, producing the “stage 2 maps”. This process was laborious and time-consuming, but aimed at producing the best possible LULC maps so to minimise field verification and consultation with landowners.

Table 3.3. Accuracy assessment of “stage 2 maps” versus reference

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall accuracy (%)</td>
<td>92.64</td>
<td>96.23</td>
<td>89.24</td>
<td>91.22</td>
<td>90.46</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.88</td>
<td>0.93</td>
<td>0.83</td>
<td>0.86</td>
<td>0.85</td>
</tr>
<tr>
<td>Kappa location</td>
<td>0.94</td>
<td>0.95</td>
<td>0.89</td>
<td>0.89</td>
<td>0.87</td>
</tr>
<tr>
<td>Kappa histogram</td>
<td>0.94</td>
<td>0.99</td>
<td>0.93</td>
<td>0.97</td>
<td>0.98</td>
</tr>
</tbody>
</table>

User’s accuracy (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
<td>93.92</td>
<td>98.99</td>
<td>93.94</td>
<td>93.16</td>
<td>92.55</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>96.21</td>
<td>97.12</td>
<td>93.90</td>
<td>93.86</td>
<td>90.24</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>76.85</td>
<td>82.91</td>
<td>65.13</td>
<td>78.01</td>
<td>86.95</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>89.71</td>
<td>96.27</td>
<td>78.05</td>
<td>87.65</td>
<td>88.94</td>
</tr>
<tr>
<td>Water bodies</td>
<td>100.00</td>
<td>93.01</td>
<td>78.25</td>
<td>73.31</td>
<td>74.46</td>
</tr>
</tbody>
</table>

Producer’s accuracy (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
<td>96.02</td>
<td>98.49</td>
<td>94.47</td>
<td>94.11</td>
<td>93.84</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>89.86</td>
<td>96.44</td>
<td>86.39</td>
<td>90.84</td>
<td>91.89</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>97.22</td>
<td>88.63</td>
<td>89.92</td>
<td>89.56</td>
<td>80.82</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>73.95</td>
<td>95.92</td>
<td>76.35</td>
<td>83.10</td>
<td>82.98</td>
</tr>
<tr>
<td>Water bodies</td>
<td>82.28</td>
<td>77.89</td>
<td>71.32</td>
<td>70.57</td>
<td>70.45</td>
</tr>
</tbody>
</table>

Table 3.3 above, shows the accuracy of the “stage 2 maps” versus the reference. The overall accuracy varies from 89% to 96%, with Kappa statistics indicating almost perfect agreement (> 0.81) (Viera and Garrett 2005). Additionally, the decomposition of Kappa into histogram and location Kappa statistics shows their values to be above 0.87, corroborating a very good agreement. These results are a very large improvement from the previous stage.
These large values are also in line with literature (Kahya et al. 2008). However, since the “stage 2 maps” were a result of API and knowledge based editing to enhance the outcomes, these large accuracy values are not surprising. Further to this, the laborious task of API, to produce the “stage 2 maps”, proved itself worthy, as the changes proposed by the landowners to generate the reference maps were minimal.

The 1995 “stage 2 map” is shown in Figure 3.7 a), while the reference is shown in Figure 3.7 b). It can be noticed that areas of “irrigated Agriculture” were overestimated and that the golf course north of Cessnock was almost omitted in the “stage 2 map”.

Also shown is the comparison between the “stage 2 map” (Figure 3.8 a) and the reference map for 2000 (Figure 3.8 b). Note that the shapes and sizes of the “irrigated Agriculture” areas were modified, as well as the boundaries of Cessnock and some of the areas of “woodland”. In this sense, the “stage 2 map” identified the golf course north of Cessnock, but its boundary was reshaped in the reference.
In summary, the “stage 1 maps” produced by a fully automated procedure through either a ruleset based or combined approach yielded overall accuracies in the order of 70% (Table 3.2). There were no differences between the overall accuracies of the two approaches (ruleset and combined, ruleset and NN). These results are in line with what was found in literature, whereby moderately high Kappa statistics were limited by low spatial resolution of the images used for the classification, due to the problem of mixed pixels (Fisher 1997).

While it would be sensible to adopt the image fusion techniques to enhance the spatial resolution of the Landsat images used for classification, this was not feasible for all years due to non-availability of quality aerial photos taken close to the dates of Landsat image acquisition. In Chapter 4, however, the benefit of such an approach was demonstrated for two sub-sections of the study area.

The API and knowledge based “stage 2 maps” involved manual enhancement of “stage 1 maps”, which was found to be very time-consuming. It was, nevertheless, the only alternative
to improving accuracy. The “stage 2 maps” were shown and discussed with a group of landowners who contributed to produce the “stage 3 maps”. These were used as reference and were an accurate representation of historical LULC patterns, from which transition probability matrices could be derived, as required for land change modelling. The land change modelling is a subject covered in Chapters 5 and 6. However, the problem of difference in spatial resolution of the Landsat images needed attention.

3.3.3 Spatial resolution issues

In order to assess the effect of the difference in spatial resolution of Landsat MSS and Landsat TM images, the maps produced from TM images (25 m) were up-scaled to 50 m by a majority rule algorithm. Also, the maps produced from MSS (50 m) were subsequently down-scaled to 25 m. The proportions of each LULC category were computed, as was suggested by Turner et al. (1989) and Aplin (2006), to investigate the effect of “grain” in regards to landscape metrics (Table 3.4).

As shown in Table 3.4 below, down-scaling the 50 m resolution maps to 25 m resolution did not affect the proportions of any of the categories, for both 1972 and 1985 LULC maps. This was expected, as the algorithm simply split the 50 m pixels into four 25 m pixels. The problem with this approach was that since down-scaling the maps did not add information it would give a false idea that if a 25 m image were classified, the results would be exactly the same as classifying a 50 m image (oversimplification).

<table>
<thead>
<tr>
<th>Table 3.4. Proportions (%) of LULC at different resolutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Woodland</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
</tr>
<tr>
<td>Intensive uses</td>
</tr>
<tr>
<td>Water bodies</td>
</tr>
</tbody>
</table>

78
Upscaling, on the other hand, did change the proportions for 1995, 2000 and 2005. In fact, it moderately increased the proportion of “woodland”, while decreased the proportions of the less dominant LULC categories, such as “water bodies” and “intensive uses”. For 2005, “woodland” increased from 35.8% of the area to 36.8%, meanwhile “irrigated Agriculture”, “intensive uses” and “water bodies” had a systemic decrease of 0.4% and 0.3% respectively. While 0.3% of the study area is not a large amount of the area, it represented one third of the cover of “water bodies”.

![Figure 3.9. Proportion of LULCs with different spatial resolutions](image)

This reinforced the need for upscaling the later LULC maps from 25 m to 50 m resolution, rather than downscaling, for subsequent land change modelling. Figure 3.9 above, illustrates these differences, whereby the main effect occurs on the less dominant LULC categories.

Fragmentation statistics, as described in Section 3.2.7, were computed. For simplicity, only Perimeter-area fractal dimension and Simpson’s diversity index are shown (Table 3.5). As was the case with the proportions of the different LULC categories for 1972 and 1985 maps, downscaling had no effect on any of the fragmentation parameters; meaning that the
spatial distribution patterns remained the same. However, the effect of upscaling from 25 m resolution to 50 m resolution confirms the effect on the proportions of LULC categories, as discussed above. This outcome is expected, since upscaling tends to reduce the complexity of shapes and sizes of LULC patches (i.e. if you upscale increasingly, you will end up with one shape of fractal dimension close to one). This is illustrated by the decrease of the perimeter-area fractal dimension for 1995, 2000 and 2005. It also reduced slightly the Simpson’s diversity index.

$$\text{Table 3.5. Fragmentation statistics at different resolutions}$$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>25 m</td>
<td>1.36</td>
<td>1.36</td>
<td>1.36</td>
<td>1.42</td>
<td>1.43</td>
</tr>
<tr>
<td>50 m</td>
<td>1.36</td>
<td>1.36</td>
<td>1.36</td>
<td>1.32</td>
<td>1.35</td>
</tr>
<tr>
<td>25 m</td>
<td></td>
<td></td>
<td></td>
<td>1.43</td>
<td>1.44</td>
</tr>
<tr>
<td>50 m</td>
<td></td>
<td></td>
<td></td>
<td>1.35</td>
<td>1.37</td>
</tr>
</tbody>
</table>

The above-mentioned exercises, of computing and comparing the proportions and landscape metrics of LULC categories, as consequence of upscaling (1995, 2000 and 2005) or downscaling (1972 and 1985) the maps, indicated that 50 m resolution LULC maps should be used. In other words, the subsequent LULC change modelling should be based on the combined multi-temporal 50 m resolution LULC maps.

### 3.3.4 On producing historical LULC maps

The objective of this chapter was to determine the spatial patterns of LULC for 1972, 1985, 1995, 2000 and 2005 in the lower Hunter valley. This was achieved using an object-based classification approach, whereby the images were first segmented and then classified. The classification procedure was based on the ruleset approach (1985, 2000 and 2005) and a combination of ruleset approach and supervised nearest neighbour (1972 and 1995).

The first stage of the process was fully automated and produced maps with overall accuracies from 70% to 74% with overall Kappa’s from 0.5 to 0.57 (detailed in Section 3.3.1). These results had a limited accuracy and were caused by the low signal to noise ration within
the images, where: 1) the classes were spectrally close; and 2) the images are in a limited spectral range. Given the limited availability of data, these results were acceptable and agreed with references in literature (Corbane et al. 2008, Gao 2008); however they were not considered accurate enough to extract transition probabilities for land change modelling.

The low spatial resolution, as well as the high noise-to-signal ratio of the Landsat images, contributed to misclassification. The spatial resolution, as confirmed by Blaschke and Strobl (2001) and Cracknell (1998), was a problem of mixed pixels, where more often than not the pixel was bigger than the minimum mapping unit (e.g. creeks). However, the object-based approach, be it through a ruleset based approach or a combination with supervised classification, assisted in identifying LULCs, such as “woodland”, “dryland Agriculture”, “irrigated Agriculture”, “intensive uses” and “water bodies”.

To enhance the “stage 1 maps” and improve their accuracies, they were overlain with aerial photographs (characterised by 1 m spatial resolution). Subsequently, through API, the misclassified patches were manually corrected. The corrected maps comprised the “stage 2” and were much more accurate, as detailed in Section 3.3.2.

Furthermore, since the objective was to produce these maps as accurately as possible, hard copies of the latter maps were taken to a number of field meetings with landowners, whose knowledge of the study area contributed to modifications of the “stage 2 maps”, resulting in the final reference maps. The field verification and validation through farmer interviews produced the “stage 3 maps”, which are used as reference in subsequent Chapter 5 and Chapter 6. The “stage 3 maps” were also used to assess the accuracies of the automated (stage 1 maps) and API enhanced maps (stage 2).

The reference maps produced here were accurate representations of LULC from the 1970s and transition matrices extracted from these maps will serve as input for a land change model, detailed in Chapter 5 and Chapter 6. One of the limitations identified here was the limited accuracy associated to the coarse spatial resolution of the Landsat images. In order to further
enhance accuracies, fusion of the Landsat images with higher spatial-resolution images was suggested and is the subject of the following chapter (Chapter 4).

Lessons learnt were two-fold, object-based approaches are useful, but their potential is not fully realised with medium resolution data, as mixed pixels limit the identification of objects. In fact, supervised classification was also hampered by the spectral signatures of the objects and therefore manual editing was necessary. However, the time-series of LULC maps, validated in the field, along with socio-economic data, will allow some understanding of the reasons of change and enable prediction of how the region might look in the future.

3.3.5 LULC change analysis

The Hunter region (lower and upper) has historically been a place of coal mines, wines and horse breeding (McManus 2008). Here the focus is on a portion of the lower Hunter Valley, which is the region characterized by wine grape growing. The region was established as a wine producing region in the 1800s (Beer et al. 2003) and up to the mid 1970s the Hunter was composed mainly of large commercial vineyards, some of which are still producing wines today. In the early 1980s, however, there was a policy from the Australian government for “pulling out the vines”, which affected production in the region (Australian Bureau of Statistics 2011).

Later, namely in the 1990s and 2000s various new wineries were established, either with new plantings or using some of the old/abandoned vines. This was regarded as the consequence of the proximity of major urban centres, such as Newcastle and more importantly Sydney, as well as the construction of an expressway between these centres and Cessnock (Halliday 2005, McManus 2008). Furthermore, the lifestyle of the countryside associated to the ease of access (through the F3) to urban centres impacted the expansion of Cessnock and of Singleton.

Table 3.6 summarises the paragraphs above, where it can be seen that the areas under Irrigated Agriculture decreased from 10.7% in 1972 to 9.5% of the area in 1985. From 1985,
the areas under wine grapes increased and reached about 13.5% of the area in 2005. Furthermore it illustrates that while the “intensive uses” only occupied 1.8% of the landscape in 1972, it more than doubled to compose 4.1% of the landscape in 2005.

Table 3.6. Composition of the landscape in terms of LULC classes per year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
<td>37.1</td>
<td>31.8</td>
<td>36.0</td>
<td>34.5</td>
<td>36.8</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>50.2</td>
<td>56.0</td>
<td>50.8</td>
<td>50.8</td>
<td>45.0</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>10.7</td>
<td>9.5</td>
<td>10.2</td>
<td>10.1</td>
<td>13.5</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>1.8</td>
<td>2.1</td>
<td>2.5</td>
<td>3.8</td>
<td>4.1</td>
</tr>
<tr>
<td>Water bodies</td>
<td>0.2</td>
<td>0.6</td>
<td>0.5</td>
<td>0.8</td>
<td>0.6</td>
</tr>
</tbody>
</table>

As stated by Halliday (2005), the conditions for wine grape growing in the area are not optimal, however its location close to urban centres attract tourists from Australia and internationally. Tourism statistics were only available from the early 2000s but show that the visitation has grown from six million tourist visits per year in 2001 to over eight and a half million in 2007 (Tourism New South Wales 2009). This was only possible due to the infrastructure build for tourism (roads, hotels, restaurants) (O’Neill 2000), furthermore, the majority of the wine in the region is not sold/produced in bulk, but usually sold in small boutique wineries (McManus 2008). Therefore, the landscape patterns (as illustrated by fragmentation statistics in Table 3.5) also changed, where the wineries of the 1970s were fewer and bigger and in 2005 they were smaller, but more numerous, explaining the increase in fractal dimension.

The increase in the areas of “intensive uses” is corroborated by census of the population for the cities that fall within the area of study, Cessnock and Singleton, as shown in Figure 3.10.
The other LULC categories have followed other patterns. While there was a loss of “woodland” to “dryland agriculture” from 1972-1985, this was reversed in the subsequent periods (1985-2005). The causes of this change in reduction of “woodland” are directly linked to the establishment of the Hunter as an area of tourism, and lifestyle consumption. As the wineries became the main attraction to the lower Hunter as well as its natural beauty and countryside lifestyle, then the woodlands also became necessary to support the tourism business.

It can be noted here, however, that the main focus of this chapter was only to establish the LULC patterns of the region, while the discussion of its historical importance and future consequences are the main theme of Chapter 6.

3.4 Conclusions

The main objective of this chapter was to create a time series of accurate LULC maps of the lower Hunter valley. The second objective was to assess their accuracies and improve them, so that they could be used in a land change model.
• The “stage 1 maps” were produced utilising the object-based paradigm, with either rulesets or a combination of rulesets and supervised classification. The advantage was that the procedures were automated, but the overall accuracies around 72% implied that these maps should be improved;

• API and knowledge based augmentation were used to generate the “stage 2 maps”, which had higher accuracies (~92%) but demanded large amount of manual input;

• Input from landowners was used to further enhance the “stage 2 maps”, resulting in the final reference maps (stage 3); and

• Improvements to the automated procedure were suggested, utilising image fusion. Image fusion and automated LULC map production is the topic of the subsequent research chapter (Chapter 4).

3.5 References


Definiens, 2004. eCognition version 4 object oriented image analysis user guide Munich: Definiens Imaging GmbH.


Chapter 4. Improving accuracy of historical LULC maps through supervised object-based classification of multi-sensor fused images

* Part of this work was presented at the 30th Asian Conference on Remote Sensing (30th ACRS), held in Beijing, China, from the 19th to the 23rd of October 2009
4.1 Introduction

As discussed previously in Chapter 3, monitoring land use and land cover (LULC) is important not only for the classification of rural and urban areas allowing for change detection (Coppin et al. 2004), but also important for a variety of other uses such as habitat mapping (Cousins et al. 2003, Gottschalk et al. 2005, Lucas et al. 2007), land management (Dale 1997, Priess et al. 2001, Yuan et al. 2005), urban planning (Xu et al. 2006, Sheng et al. 2008) and as a tool for climate change science (Feddema et al. 2005), amongst other uses. Recent developments in the science and technologies of remote sensing, providing data with increasing spatial, spectral and temporal resolutions have facilitated these efforts (Aplin 2004).

To take advantage of the advances in remote sensing, many studies on LULC dynamics and its interactions with the environment often rely on historical data. For instance, Betts et al. (2007) reconstructed LULC maps from 1750 onwards until early 2000s, using databases of historical cropland inventories and LULC modelling, to demonstrate the effect of deforestation on surface temperature based on a general circulation model. Similarly, Hurtt et al. (2006) created a global gridded dataset of deforestation, to enable others to assess the consequence of anthropogenic changes. Historical LULC data have also been used to derive the drivers of LULC change, as they enable researchers to examine the links between urban and rural areas (Currit 2005). Thus, the legacy LULC maps are an important component of land change modelling (López et al. 2001, Heldens 2006), hydrological modelling (Lin et al. 2008) and are used for detecting spatial-temporal patterns of deforestation (Turner et al. 2003) amongst many other uses. All of these studies are geared towards the medium- to long-term monitoring of the environment, vis-à-vis human influence.

As such, a time series of legacy LULC maps are needed to quantify and detect the spatial patterns of LULC and environmental changes (Riebsame et al. 1994, Brown et al. 2000). This
is especially important when the spatial extent of a given study area is too small for global datasets to be useful; however at such a spatial extent that manual aerial photo interpretation is unfeasible or may be impracticable.

The creation of historical LULC data is often limited by the problem of mixed pixels (Cracknell 1998), limited by the low spatial resolution of satellite images acquired in the 1970s and/or low spectral resolution of historical aerial photographs, as elucidated in Chapter 2. In spite of these limitations, the Landsat series of satellites has been used to monitor Earth’s surface since the early 1970s, providing images acquired at a relatively short revisit time (USGS and NASA 2005). Comparatively, aerial photos which are usually characterised by high spatial resolution have also been used, but are limited by their infrequent, irregular revisit time and their low spectral (often B&W or RGB) resolution. The challenge then is how to maintain a regular monitoring of the environment at reasonable accuracy, backed by historical data that have limited spatial resolutions.

One of the possible solutions to this challenge is sub-pixel classification, as proposed by Foody and Cox (1994). This approach has improved the accuracy of LULC classification (Aplin and Atkinson 2001, Verbeiren et al. 2008). However, it requires an in-depth understanding of the underlying structure within the pixel. Moreover, this approach may be problematic if used for large areas that are characterised by different LULC categories with variable internal structures.

An alternative solution is image fusion, where Landsat data could be fused with higher spatial resolution imagery, such as orthophotos, in order to incorporate more detailed information about the LULC patterns, prior to a detailed automated classification scheme (Wald 2000). This is particularly so, if there is the need to enhance the quality of historic LULC maps, generated from multispectral Landsat images fused with the high resolution orthophotos (Geneletti and Gorte 2003).
Image fusion is only a small part of a larger framework named data fusion, defined by Wald (2000), as a series of tools and methods which enable data from different sources to be combined, which ultimately would improve the “quality” of the output. Despite data fusion being widely used in a variety of fields (from industrial processes to LULC mapping), in this study the focus is on multi-date, multi-sensor image fusion for automated LULC classification.

The significance of fusion lies on the fact that it is a flexible tool, which enables the integration of disparate multiple data, with different spatial, temporal and spectral resolutions (Pal et al. 2007). The value of this tool is related to the capacity of: increasing the spatial resolution of imagery (Svab and Ostir 2006), improving image registration (Ehlers 1991), increasing classification accuracy (Colditz et al. 2006); and enabling accurate LULC change detection (Zhang et al. 2005), amongst other uses.

Image fusion could be classified into three distinct levels: a) pixel-level, in which pixel values from different sensors are merged (i.e. pan-sharpening); b) feature-level, whereby images from different sensors are fused during the process of feature extraction (i.e. segmentation) and c) decision-level, which involved feature extraction from one image, followed by fusion of the information contained in both, during the process of classification (i.e. post-segmentation fusion) (Pohl and van Genderen 1998, Ehlers et al. 2008).

While the three levels of fusion described above have their uses, much work has focused on developing algorithms for pixel-level fusion for the purpose of enhancing visualisation and to increase the spatial resolution without the loss of spectral information (Ehlers et al. 2008). For instance, Welch and Ehlers (1987) utilised the intensity-hue-saturation (IHS) colour transform to merge the panchromatic band of SPOT HVR with Landsat. Similarly, Ranchin and Wald (2000) proposed the use of ARSIS concept, which employs a wavelet transformation. Other approaches included PCA fusion, Brovey and Gram-Schmidt transformations, as well as Ehlers fusion algorithm. In spite of these multiple approaches and
different algorithms, there is no general agreement on which performs best (Karathanassi et al. 2007).

The issue of the influence of different pixel-level fusion approaches on LULC classification accuracy was discussed by Colditz et al. (2006), who found that adaptive image fusion, wavelet and PCA transformations outperformed Brovey and IHS. However while many pixel-level fusion approaches, such as the ones cited above, have been developed and compared (Svab and Ostir 2006), detailed study and comparison of other levels of fusion have been neglected for the purpose of LULC mapping.

References to feature- and decision-level fusion for LULC classifications are sparse in the applied literature. Some work has been done for land-mine detection (Gunatilaka and Baertlein 2001), where authors compared feature- and decision-level fusion, however using different data for each. A more comprehensive comparison, for instance, was undertaken by Yuan et al. (2008) who compared various pixel-level fusion algorithms to one approach of feature- and decision-level fusion for LULC classification, finding that decision-level fusion performed best, but also not using the same data for the three levels of fusion. In particular little work, if any, has been done using the same datasets for LULC classification, comparing all the three levels of fusion and benchmarking them with classification of the non-fused images.

In terms of LULC mapping, a variety of algorithms exists and were summarised in Chapter 2. A review by Lu and Weng (2007) pointed out that successful classification depended on: the quality of available data, the design of a classification procedure and the analyst’s skills. A number of studies have focused on comparing a variety of classification algorithms (Qian et al. 2007, Wang et al. 2007, Zhou et al. 2009), however Lu and Weng (2007) also pointed out that there was no standardisation or agreement on choice of classification algorithms.
Choice of the classification algorithm starts with the decision if individual pixels or groups of pixels (segments/objects) will be classified. In some studies, the pixel-based approach was shown to be superior (Wang et al. 2008), whereas in others, the object-based approach presented better results (Gao 2008) (also discussed in Chapter 2 and Chapter 3).

According to Geneletti and Gorte (2003), object-based classification had the advantage of using contextual information to delineate objects. In a two-stage process, pixels were grouped according to their similarity, forming segments (Baatz and Schäpe 2000), which were then classified. The segmentation process produced a set of geographical entities such as LULC classes, the type of information required for LULC mapping (Aplin and Smith 2008). This approach eliminated the “salt and pepper effect” of pixel-based classification (Yu et al. 2006), but was limited by the quality of the segmentation procedure (Koch et al. 2003).

Image classification would, however, be affected by the image’s spatial resolution. Medium resolution data, such as Landsat, were often faced with mixed pixels that: either were on the boundary of two covers; had small objects within a pixel; were composed of linear features, such as roads that comprised part of the pixel or finally pixels which represented an intergrade between covers (Fisher 1997) (Figure 2.5 in Chapter 2).

The different levels of image fusion associated to an object-based classification procedure provided an opportunity to enhance the quality of resulting LULC maps (Ehlers 2006), as it would help in minimising the problem of mixed pixels.

Further to the choice of the object-based approach, a classification algorithm had to be chosen. As mentioned previously, there is no consensus on a best algorithm (Lu and Weng 2007), however a supervised approach is deemed to be appropriate. While there were a variety of algorithms, such as maximum likelihood (Rogan et al. 2002, Gao 2008), Warfield (1996) suggested a nearest neighbour algorithm for classification of remote sensing images. Atkinson (2004) found that this algorithm had better accuracies than other traditional supervised classifiers, however it would work better on segments, rather than on individual
pixels. Furthermore, the nearest neighbour algorithm has been reported to perform well in image classification by other authors (Singh et al. 2001, Laliberte et al. 2006, Yan et al. 2006) and was readily available within Definiens software.

A supervised classification within Definiens Professional 5.0, utilising the object-based paradigm with a nearest neighbour classifier, would enable an unbiased comparison of different levels of fusion. Consequently, the main aims of this chapter were to:

i) illustrate the effects of pixel-, feature- and decision-level fusion of Landsat and digital orthophotos on the accuracy of historical LULC maps; and

ii) benchmark the performance of the three fusion levels against classification of non-fused images. Landsat and orthophotos were individually classified, without fusion.

4.2 Data and methods

4.2.1 LULC classes and sub-areas of study

As it has been suggested that the LULC classes selected in a given study may significantly influence the classification outcomes (Jensen 2005), LULC classes that were discernible from the images that matched the broadest level of the Australian standards (ACLUMP, Bureau of Rural Sciences 2006) were selected and followed those of Chapter 3.

The selected LULC classes include: “woodland”, characterised by native vegetation and remnants around creeks; “dryland Agriculture” areas, mainly composed of grazing areas and natural grasslands; “irrigated Agriculture”, mostly composed of wine grapes; “intensive uses”, such as roads, hotels, built-up areas; and finally “water bodies”, such as dams and creeks.

As described in Chapter 3, the study area was located in the lower Hunter valley of NSW, Australia. Here, however, due to data constraints two different sub-areas of 8 by 8 km for two different dates were selected, to test the robustness of the procedures proposed here. Field
validation of the selected sub-areas and consultation with landowners combined with visual interpretation of the aerial photos enabled the accurate characterisation of each sub-area.

The first sub-area (SA1) was mainly composed of “dryland Agriculture” areas, with patches of “woodland”, “irrigated Agriculture” and “water bodies” of variable size. There were also limited number of “intensive uses” patches, characterised mainly by roads and rural properties.

The second sub-area (SA2) was characterised by a large part of Cessnock city, including its roads and airport, some contiguous blocks of “woodland”, including riparian vegetation, as well as “dryland Agriculture” and some “irrigated Agriculture” areas. A few “water bodies” were identified as well in SA2.

4.2.2 Image data

Two subsets of Landsat TM scenes were used; both were obtained courtesy of the Australian Greenhouse Office in Canberra. They were provided in an 8-bit radiometric depth, without the thermal band, but were already orthorectified and resampled to a 25-m DEM (see Furby 2002 for processing details). While the first image (path 90/row 83) was acquired in February 2000, the second (path 89/row 83) was acquired in December 2003. The selection of these images was based on their quality and matching the peak growing season of grapes.

To match the above-mentioned Landsat scenes, orthophotos from two different years were used. The first orthophoto was a 1 m spatial resolution digital orthophoto, obtained from NSW Department of Lands. The image was a 3 band true-colour composite, taken with a Wild RC30 camera in October 1998. The hard-copy aerial photos were scanned using a Vexcel Ultrascan 5000 photogrammetric scanner and orthorectified by the NSW Department of Lands using a 25-m DEM (personal communication, NSW Lands, 2009). The second orthophoto, taken in September 2004, followed the same procedures as above.

In this study, the orthophotos served two different purposes: first, at their original resolution of 1 m, as reference for manual classification of points within the two sub-areas;
second, both images were degraded to a 5-m spatial resolution using a cubic convolution algorithm (van Der Meer 1997, Jensen 2005). The 5-m orthophotos were then used for the different fusion approaches. This was done to keep the ratio of fusion to 5:1, as suggested by Ehlers and Klonus (2008).

While SA1 was composed of an 8 by 8 km tile utilising the 1998 orthophoto and the 2000 Landsat scene, SA2 utilised the 2004 orthophoto and the 2003 Landsat scene. Data availability constraints did not allow for the gaps between the orthophotos and the aerial photos to be smaller.

Image fusion is more commonly done with images from the same sensor, such as Landsat ETM+ which has its multispectral bands (natively with 30 m pixels) and one panchromatic band (15 m pixel). The difference is that while the multispectral bands cover specific portions of the spectra (R, G, B, near IR, mid-IR), the panchromatic band covers the whole visible spectrum. The implication of this is that fusion within a sensor eliminates various problems. For instance, registration of one image to another is much simpler and since the multispectral images are taken on the same day, there is no issue with differences in vegetation, sun angle and even LULC changes.

Using images for fusion which were taken on different dates adds the complexity of differences in sun-angles, thus the shadows will appear differently, but more importantly when the images are taken in different seasons, then the vegetation changes (i.e. imagine a tree in the winter and again in the summer). In the specific case of using fusion to create historical LULC maps, the issues mentioned above have to be considered in the data analysis, however as mentioned earlier, it was not possible to find images with smaller date gaps.

### 4.2.3 Image fusion and segmentation

As pointed out above, Pohl and van Genderen (1998) categorized image fusion techniques into three levels; pixel-level fusion (PLF), feature-level fusion (FLF) and decision-level fusion (DLF), all of which were tested in this study. Here PLF was done using the PCA
resolution merge technique, as suggested by Erdogan et al. (2008) and Colditz et al. (2006), while noting that other algorithms could have been used as well. Following PLF, the image was then segmented and classified. Secondly, FLF used both datasets (Landsat and orthophoto) for segmentation and subsequent classification. For DLF, the higher spatial-resolution orthophotos were used for segmentation, but classification of the objects used information contained in the orthophoto and Landsat. In order to verify whether the use of fusion did improve classification accuracy, Landsat images (NFL) and orthophotos (NFO), were classified independently, without fusion. The schematic of the methodology is illustrated in Figure 4.1.

![Figure 4.1. Workflow of the different fusion approaches](image)

Whilst there are various segmentation algorithms, such as the ones built into proprietary software such as: Definiens, Erdas Imagine and ENVI, to mention a few; there are various approaches to segmentation, including k-means clustering, region growing, edge detection, tessellation, histogram thresholding amongst other approaches (Pantofaru and Hebert 2005, Ma et al. 2010).

The objective of the present study was not to assess the accuracy of different segmentation algorithms. Therefore, one standard segmentation algorithm, that is native to Definiens Professional 5.0, was applied to all of the different approaches (Benz et al. 2004).

In applying the segmentation algorithm, the spectral and textural parameters were kept constant, while maintaining different numbers of spectral bands for the different approaches.
The knowledge acquired from this exercise was used to modify the variability threshold required for generating objects to ensure all images yielded around 2500 objects (Table 4.1).

<table>
<thead>
<tr>
<th>Fusion approach</th>
<th>SA1</th>
<th>SA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLF</td>
<td>2496</td>
<td>2503</td>
</tr>
<tr>
<td>FLF</td>
<td>2500</td>
<td>2496</td>
</tr>
<tr>
<td>DLF</td>
<td>2533</td>
<td>2552</td>
</tr>
<tr>
<td>NFL</td>
<td>2412</td>
<td>2337</td>
</tr>
<tr>
<td>NFO</td>
<td>2542</td>
<td>2552</td>
</tr>
</tbody>
</table>

One of the key factors affecting classification accuracy are over- and under-segmentation (Moller et al. 2007). To avoid both these situations, 2500 was chosen as the number of objects used for fusion. By visual inspection, combined with the knowledge of the area and consultation with landowners, it was possible to judge that the objects generated followed the boundaries seen in the images.

4.2.4 Training data and classification

The selection of appropriate training data was essential for any supervised classification, particularly because the sample size and locations would greatly impact the classification results (Jensen 2005). The training areas were selected based on the knowledge acquired during field visitation, combined with aerial photo interpretation (Table 4.2 below). The total area of the selected training sites varied with the different approaches (fusion and no fusion), since the object boundaries from segmentation were not the same.

In all cases though, the total number of pixels for training data was larger than the minimum proposed by Mather (2004), who suggested an optimal sample size of 30 pixels x p x c, where p is the number of spectral bands and c is the number of classes used in classification. This agreed with numbers suggested by Park and Stenstorm (2008) and van Genderen et al. (1978). As image fusion occurred at three levels, the number of spectral bands required for classification would vary from 6 (for PLF) to 9 (for FLF and DLF). Given that
five LULC categories were considered here, the samples size for training should be no less than 1350 pixels (9 bands, 5 classes, 30 pixels per band per class).

Table 4.2. Number of training polygons per sub-area

<table>
<thead>
<tr>
<th>Category</th>
<th>SA1</th>
<th>SA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Water bodies</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

Once the training polygons were selected, they were used for supervised classification based on the nearest neighbour algorithm, as suggested by Warfield (1996) and described in Jensen (2005). This algorithm was chosen because of its efficiency and produced comparable, if not better, accuracies than other supervised methods (Atkinson 2004).

The feature-space used for each method was automatically selected in such a way to maximise class separability, but included no more than the mean values of each layer and their standard deviations. Important indices, such as NDVI, were not used because they could not be computed for all approaches (e.g. NFO).

4.2.5 Classification accuracy assessment

The selection of an appropriate sample size for accuracy assessment was no trivial task. While this issue has been topic of discussion and research for decades now, the scientific community has not reached a consensus so far. For example, Congalton (1991) and Mather (2004) suggested, as a rule of thumb, a minimum of 50 samples per class, which for this study meant 250 samples. However, Jensen (2005) and Foody (2009) suggested the binomial distribution function for determining sample size. In this, the number of samples ($N$) is defined by the binomial probability function:

\[
N = \frac{Z^2(p)(q)}{E^2}
\]

Equation 4.1
where Z is the normal deviate from a 95% two-sided confidence interval; p is the expected accuracy of the map; q=100-p; and E is the allowable error. In our case, Z=2, p=85% and E=5%, and hence N=204 samples. This parameter is counter intuitive because it is based on the estimated accuracy of the map (p), however the product (p)*(q) would reach a maximum when p was set to 50%. It also did not consider the number of classes within the map.

A similar approach to the binomial function for sampling selection sampling is the multinomial distribution (Tortora 1978, Grenier et al. 2008). One of its advantages was that it considered the number of classes and the estimated proportion of the most frequent class (Π). The multinomial function can be derived as:

\[ N = \frac{B\Pi(1-\Pi)}{b^2} \]  

where B is the upper 1-(α/k) percentile of the \( \chi^2 \) distribution with 1 degree of freedom; α is the desired accuracy; k is the number of classes; b is the desired precision; and Π was the proportion of the most frequent class. Tortora (1978) suggested setting Π to 50% to maximize the number of samples. In this case, with k=5, α and b of 5%, and Π of 50%, the estimated number of samples was 664.

In order to increase the confidence and precision, 1000 points were randomly allocated within each sub-area. Since it was not possible to fully validate the historical LULC data in the field, orthophotos at their native resolutions of 1 m were used to manually classify the randomly allocated points within each sub-area, through aerial photo interpretation. More importantly, field visitation and consultation with the landowners enabled the validation of some of the sample points in which there was doubt. While this approach was not ideal, it was the best means of populating the error matrix. In order to keep the comparisons unbiased, the
randomly allocated sample points were forced to be outside of the training areas, as suggested by Foody (2009).

An error matrix was derived comparing the reference points to each of the fusion methods, for each sub-area. A number of indicators of map quality were computed, including: overall accuracy and Kappa statistic (Mather 2004); location and histogram Kappa (Hagen 2002); and finally disagreement due to allocation and quantity, as proposed by Pontius (2002). The output of the different approaches was also visually compared and the proportions of points in each LULC class, for each approach and the reference were computed.

4.3 Results and discussion

In this chapter the effect of different approaches of image fusion in producing LULC maps from multi-temporal, multi-sensor data were compared. The different fusion (and no fusion) approaches were tested on two different sub-areas. The selection of these areas was based on data availability and on the fact that their LULC distribution was different. SA1 was characterized by a predominant rural landscape, while SA2 encompassed the city of Cessnock.

4.3.1 Class separability and areal proportions

The first step of quality assessment of the LULC maps produced by the three levels of fusion and no fusion was to compute the areal proportions with the reference. After assignment of the feature-space, the minimum Euclidian distance between the class centroids (class separability) was computed, generating a 5x5 matrix for each of the five approaches (PFL, FLF, DLF, NFL and NFO) and two sub-areas. Ten matrices were computed, but are not shown here.

The separability values from PLF and NFL were quite similar. The PLF fusion algorithms’ primary aim is to increase the spatial resolution, while preserving the spectral characteristics
(Klonus and Ehlers 2007), thus these similarities were expected as the spectral information for PLF and NFL were similar. The FLF and DLF produced the best class separability. These approaches (FLF and DLF) utilized information from both Landsat and orthophotos images in their feature-space for classification it would be natural that the separability values were the highest. However, the separability of the NFO was similar to those of NFL and PLF, which was a surprise. Initially it was thought that the multispectral information from the multiple bands of Landsat, used in NFL and PLF, would provide more spectral separability than the three visible bands of the orthophotos (NFO), but that was not the case. The reason for NFO, NFL and PLF to have similar separability values indicated that the quality of spectral information in the orthophoto and Landsat image was similar.

The lowest separability values in all cases were between “dryland Agriculture” and “irrigated Agriculture”. “Water bodies”, “intensive uses” and “woodland” were easily separable as their spectral signatures were very distinguishable. These separability values were expected and similar outcomes have been reported consistently in the literature (Colditz et al. 2006).

Once the classification was done for all of the approaches and in both sub-areas, the proportion of points classified in each LULC class was computed. This is shown in Table 4.3 for SA1 and in Table 4.4 for SA2. The PLF underestimated “dryland Agriculture” at the expense of other classes. FLF and DLF produced a better class distribution than PLF, but overestimated “irrigated Agriculture”. All approaches (PLF, FLF, DLF and NFL), but NFO, overestimated “irrigated Agriculture”. NFL greatly overestimated “irrigated Agriculture” in SA1 and SA2, while NFO performed quite well in terms of class proportions.

<table>
<thead>
<tr>
<th>Table 4.3. Proportion of points for SA1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
</tr>
<tr>
<td>Woodland</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
</tr>
<tr>
<td>Intensive uses</td>
</tr>
<tr>
<td>Water bodies</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Woodland</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
</tr>
<tr>
<td>Intensive uses</td>
</tr>
<tr>
<td>Water bodies</td>
</tr>
</tbody>
</table>

### 4.3.2 Comparing the patterns of LULC maps

A visual comparison of the output maps, produced for the different sub-areas and fusion approaches, was the second step of map quality assessment. The comparison between the original orthophotos and the final classified maps for each of the approaches is shown in Figure 4.2 for SA1 and in Figure 4.3 for SA2.

It is important to note that, while there was no reference LULC map, it could be seen that PLF greatly overestimated the amount of “woodland” in SA1 (Table 4.3 and Figure 4.2 b). The maps produced from NFL alone appear to be rougher than others (Figure 4.2 e and Figure 4.3 e).
Figure 4.2. A view of sub-area 1: a) 1998 orthophoto; LULC maps by: b) pixel-level fusion; c) feature-level fusion; d) decision-level fusion; e) non-fusion Landsat; and f) non-fusion orthophoto.
Figure 4.3. A view of sub-area 2: a) 2004 orthophoto; LULC maps by: b) pixel-level fusion; c) feature-level fusion; d) decision-level fusion; e) non-fusion Landsat; and f) non-fusion orthophoto
4.3.3 Classification accuracy assessment

Table 4.5 presents the accuracy statistics for SA1. The overall accuracy and Kappa were highest for FLF, while location and histogram Kappa were highest for PLF and NFO, respectively. The degree of disagreement due to allocation and to quantity was minimal for PLF and NFO, respectively.

Table 4.5. Accuracies of SA1

<table>
<thead>
<tr>
<th></th>
<th>PLF</th>
<th>FLF</th>
<th>DLF</th>
<th>NFL</th>
<th>NFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall accuracy (%)</td>
<td>75.7</td>
<td>87.8</td>
<td>83.3</td>
<td>75.3</td>
<td>83.1</td>
</tr>
<tr>
<td>Overall Kappa</td>
<td>0.62</td>
<td>0.81</td>
<td>0.74</td>
<td>0.63</td>
<td>0.72</td>
</tr>
<tr>
<td>Location Kappa</td>
<td>0.89</td>
<td>0.86</td>
<td>0.81</td>
<td>0.82</td>
<td>0.76</td>
</tr>
<tr>
<td>Histogram Kappa</td>
<td>0.70</td>
<td>0.93</td>
<td>0.91</td>
<td>0.77</td>
<td>0.95</td>
</tr>
<tr>
<td>Disagreement due to allocation (%)</td>
<td>5.2</td>
<td>8.1</td>
<td>10.9</td>
<td>9.2</td>
<td>13.7</td>
</tr>
<tr>
<td>Disagreement due to quantity (%)</td>
<td>19.1</td>
<td>4.1</td>
<td>5.8</td>
<td>15.5</td>
<td>3.2</td>
</tr>
</tbody>
</table>

However for SA2, the overall accuracy and Kappa, as well as location and histogram Kappa were highest for FLF (In Table 4.6), with location Kappa being the same for FLF and FLF. Like the previous results, the degree of disagreement due to allocation and quantity were minimal for PLF and NFO, respectively.

Table 4.6. Accuracies of SA2

<table>
<thead>
<tr>
<th></th>
<th>PLF</th>
<th>FLF</th>
<th>DLF</th>
<th>NFL</th>
<th>NFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall accuracy (%)</td>
<td>78.5</td>
<td>86.3</td>
<td>83.0</td>
<td>79.3</td>
<td>81.9</td>
</tr>
<tr>
<td>Overall Kappa</td>
<td>0.70</td>
<td>0.80</td>
<td>0.75</td>
<td>0.71</td>
<td>0.73</td>
</tr>
<tr>
<td>Location Kappa</td>
<td>0.87</td>
<td>0.87</td>
<td>0.83</td>
<td>0.85</td>
<td>0.79</td>
</tr>
<tr>
<td>Histogram Kappa</td>
<td>0.80</td>
<td>0.92</td>
<td>0.91</td>
<td>0.67</td>
<td>0.92</td>
</tr>
<tr>
<td>Disagreement due to allocation (%)</td>
<td>7.6</td>
<td>8.1</td>
<td>11.1</td>
<td>9.1</td>
<td>12.9</td>
</tr>
<tr>
<td>Disagreement due to quantity (%)</td>
<td>13.9</td>
<td>5.6</td>
<td>5.9</td>
<td>11.6</td>
<td>5.2</td>
</tr>
</tbody>
</table>

As pointed out by Mather (2004), the classical accuracy assessment relies on the error matrix, overall accuracy and Kappa. Gao (2008) reported 74.2% accuracy for an object-based classification. In this study, the largest overall accuracy was roughly 88% (FLF in SA1, Table 4.5), which is considered high, as the benchmark for a very good accuracy assessment is 85% (Mather 2004, Jensen 2007). This is particularly very good because, indices, such as NDVI which could have further enhanced the classification, were not included in the feature-space;
furthermore, classification was fully automated. For example, Geneletti and Gorte (2003) obtained an overall accuracy of 86% using object-based classification of Landsat and orthophotos, but their approach incorporated both the NDVI and empirical classification rules.

Other studies have reported higher classification accuracies, however here it has to be considered that: 1. training sites were selected directly on the images, so errors by the producer could have affected classification; 2. the images were taken on different dates and even though every effort was made to select areas which wouldn’t be affected by seasonal changes, some error could have been introduced; and 3. While the multispectral information from Landsat was helpful in discriminating LULC classes, its relatively low spatial resolution was an issue that could only be partially solved using image fusion.

As shown in Table 4.5 and Table 4.6, the overall Kappa values varied from 0.62 to 0.81. Other studies suggested that Kappa greater than 0.6 is indicative of substantial agreement and values above 0.8 represented almost perfect agreement (Viera and Garrett 2005). A previous study by Yuan et al. (2008), comparing different fusion approaches for LULC classification using the Kappa statistic, found that Kappa indices of PLF may vary between 0.64 and 0.74, while FLF and DLF may vary between 0.77 and 0.80. These results are similar to the findings of this study in which the FLD and DLF generally performed better than PLF, however the authors (Yuan et al. 2008), used vegetation indices for FLF and an ensemble of classifiers for DLF.

The overall accuracies and Kappa followed a pattern for both sub-areas whereby the highest overall accuracy and Kappa was achieved by FLF, followed by DLF, NFO and then similar values for both PLF and NFL.

In addition to the use of overall Kappa, Hagen (2002) proposed its decomposition into components of histogram (relative to the frequency of the class) and location (relative to the spatial allocation). For each of these indices, the values can vary from 0.70 to 0.95 (Table 4.5 and Table 4.6). Here, the location Kappa was highest for the PLF in SA1, with the values for
PLF and FLF in SA2, tied. In SA1, the histogram Kappa was highest for NFO, while for SA2 it was highest for FLF and NFO.

Turning to the location Kappa, its values were found to be generally above 0.80 for all approaches except for the NFO. This high level of accuracy is attributable to the effectiveness of the supervised classification. While there have been a few studies on the location Kappa comparing different fusion approaches, it seems that the use of the spectral data from the Landsat images contributed to the increase on the location Kappa, as its lowest value was produced by NFO.

Comparatively, the only large values of histogram Kappa seemed to be produced by those approaches involved with the high resolution orthophotos. Once again, to our knowledge, no study has been reported comparing the different fusion approaches using histogram Kappa. Coincidentally or not, all approaches which used the orthophotos for segmentation/classification had histogram Kappa’s greater than 0.9. This is probably the case because the spectral signatures of the orthophotos contained fewer mixed pixels (Cracknell 1998) and aided in a better segmentation.

While the application of Kappa and its decompositions are widespread for accuracy assessment of categorical maps (Qian et al. 2007, Gao 2008), other authors, such as Pontius Jr and Millones (2008), have pointed out one of Kappa’s flaws. Kappa and its variations are indicative of an overall agreement relative to an expected agreement with random allocation. In other words, how much more accurate is the map produced by different approaches than one, whereby, the same proportions of each class were randomly assigned? The flaw is due to the fact that the landscape patterns are known to be non-random (Turner 1987).

Pontius Jr and Millones (2008) proposed a solution which takes into consideration the disagreement due to quantity and allocation. These indices were computed and are shown in Table 4.5 and Table 4.6. The results indicate that the PLF had the lowest allocation disagreement; however it had the highest disagreements in terms of quantity, in both sub-
areas. NFO had the lowest quantity disagreement in both areas; however had a quite high
disagreement in terms of allocation. The sum of these measures would equal 100% minus the
overall accuracy, therefore indicating how much of the inaccuracies of the map were
associated to over/under representation of a class and how much was associated to the spatial
allocation. Thus, the FLF had the highest overall accuracies and consequently the smallest
combined disagreement measures.

The accuracy statistics shown in Table 4.5 and Table 4.6, discussed above show
interesting trends. First, that classification based only on the orthophoto (NFO) provided some
meaningful results, as the overall accuracy was similar to that of the DLF. In comparison, the
NFL produced the lowest accuracy and the orthophotos did improve the classification
accuracy through the FLF and DLF.

In terms of comparing the different fusion approaches, some studies showed that most
approaches were based on PLF (Pohl and van Genderen 1998). This is understandable, as
newer sensors usually are capable of simultaneously capturing a panchromatic band of high
spatial-resolution plus a number of lower spatial resolution multispectral bands. The previous
scenario allows for the fusion of the lower spatial-resolution bands with the panchromatic
band. Such development is too late for the purpose of producing good historical LULC maps,
as the older sensors, such as MSS and TM on board Landsat, were not capable of capturing
the high-resolution panchromatic band.

The main objective of the PLF is to increase the spatial resolution without interfering with
spectral characteristics (Ehlers 2004, Klonus and Ehlers 2007), thus a successful PLF would
not increase the quantity or the quality of the spectral information. This implies that the
results of classification of such fused images (PLF) should be similar to the results of
classifying the original non-fused image (NFL), as seen here.

Generally, data required for historical LULC classification are often sourced from multi-
sensor, multi-date images, it would be more appropriate to utilize the entirety of these data. Of
the approaches compared here, the FLF incorporated all of the nine spectral bands (six from Landsat and three from the orthophoto) for segmentation and classification. This provided the best data available to resolve the spatial structures identified during segmentation and classification, thus FLF outperformed all other approaches.

Colditz et al. (2006) compared six different algorithms of PLF, with varying results in terms of overall accuracy (which varied between 62% and 81%) and Kappa (0.56 and 0.79). The best performing PLF algorithm was the PCA resolution merge, as chosen to be used in the present work. The authors (Colditz et al. 2006), however, did not benchmark their different PLF algorithms against a classification of the original Landsat scene (NFL) as done here. Here, PLF and NFL had similar accuracies, leading to question the usefulness of PLF for historical LULC mapping.

An example of comparison of the three levels of fusion for LULC classification was done by Yuan et al. (2008). They utilized four different PLF algorithms, but when considering FLF they employed textural measures and NDVI for classification, while for DLF they utilized a classifier ensemble. In their case, PLF and FLF had comparable outcomes, with accuracies between 74% and 81%, marginally lower than the accuracies reported here. DLF at its turn, had accuracies between 83% and 85%, which are in line with the results shown in Table 4.5 and Table 4.6. The limitation of what was proposed by Yuan et al. (2008) was, once again, that data utilized for classification varied with the fusion approach and no benchmark, such as NFL and NFO, was established.

In the example shown here, the DLF performed second best overall. It was hypothesized that the segmentation of the high resolution orthophoto would delineate the “best” objects and that the spectral information from Landsat would then contribute to higher classification accuracy than otherwise. However, the results indicated that a combination of both the Landsat and orthophoto (FLF) provided the best data for classification, but DLF performed second best.
Based on the work reported here vis-à-vis similar studies in the literature, there are a few issues which need to be addressed. First is that, while Wald (2000) proposed a conceptual approach for differentiating the levels of fusion, this has not been adopted by the scientific community. Consequently, the definitions of FLF and DLF vary greatly (see for instance: Ehlers et al. 2008, Yuan et al. 2008). Second, while a number of studies compared different PLF algorithms (as discussed above); many were focused visual comparison of the fusion results, without utilising the fused data for LULC classification. The study reported here is an attempt to correct this and to demonstrate the need and benefit of using the fused images for LULC classification.

Finally, there is a lack of studies benchmarking of the results of LULC classification of the fused data (PLF, FLF and DLF), with those obtained from the original images (NFO and NFL). In this study, this benchmarking was achieved based on the accuracy assessment. Thus, a combination of visual inspection, with accuracy assessments, indicated improved classification accuracy as a result of image fusion. Overall accuracy of classification increased from about 77% for NFL to 87% for FLF. Additionally, Kappa indices and minimisation of the disagreement due to quantity and to allocation were improved as a result of fusion. Overall, the FLF and DLF performed better than PLF. Classification using NFO also produced reasonable results, however the disagreements due to allocation were quite high (Table 4.5). This was probably due to the limited spectral information contained in the three bands of the orthophotos.

While this study indicated that FLF and DLF performed better than PLF in terms of the accuracy measures (Table 4.5 and Table 4.6), other segmentation, classification or PLF algorithms could yield different results.

For the specific case shown in this study, with segmentation parameters kept constant and using a nearest neighbour supervised classification algorithm, FLF and DLF outperformed the more commonly used PLF. PLF was also outperformed by NFO, but produced results similar
to NFL. Furthermore, the procedures proposed here could be fully automated, enabling rapid mapping of large areas.

4.3.4 Further work

In a time-constraint work like this, it is difficult to cover all aspects of the research. Therefore it is important to point out a few pertinent gaps that need more research. In order to further validate the results reported here, the different fusion approaches should be tested with independent data set from other areas.

Another issue is the need for comparison using other algorithms for PLF, as there is no consensus on the best algorithm (Karathanassi et al. 2007). Thus, other PLF algorithms, such as Brovey transform, HPF, Ehlers fusion, subtractive, wavelet, amongst others, could be compared to FLF, DLF, NFL and NFO.

4.4 Conclusions

The main objective of this chapter was to compare the accuracy of historical LULC maps generated by the three different fusion levels of multi-source, multi-date images using object-based supervised classification methods. In this case, Landsat images were fused with orthophotos. It was shown that:

- Classification based on FLF and DLF produced the best results. This was because they had more spectral information for segmentation and classification. As such the overall accuracy was above 83%, while minimizing the disagreements due to quantity and allocation;

- Surprisingly, the third best classification outcome was produced from NFO, also characterised by high overall accuracy, although by high disagreement due to allocation;
• The PLF and NFL had the poorest performance. As the objective of PLF was to enhance spatial resolution without affecting spectral resolution, the similarities of results were expected. However, testing other PLF algorithms may be necessary;

• Based on literature, it was expected that the PLF approach would match if not outperform FLF and DLF approaches. Therefore, further tests with other datasets and PLF algorithms would be necessary to corroborate the results produced here;

• The methods proposed here could be fully automated, enabling LULC mapping of large areas in a time-effective manner;

• The standardisation of the definitions of FLF and DLF would be beneficial to the profession; and

• This study has also demonstrated the need for comparing the results from different methods of fusion (PLF, FLF and DLF) through benchmarking with the results of classification of the non-fused data (NFL and NFO).

4.5 References


Chapter 5. Development and testing of a hybrid, weighted Markov-chain and cellular automata, land change model*

* Part of this work was presented at the 15th Australasian Remote Sensing and Photogrammetry Conference (15ARSPC) held in Alice Springs, NT, from the 13th to the 17th of September 2010
5.1 Introduction

Land use and land cover (LULC) change is one of the main drivers of global environmental change, as discussed previously. Since the industrial revolution, humankind has altered a significant proportion of Earth’s ecosystems to foster economic and capital development and to improve living standards to an unprecedented level. Arguably, one of the most important factors of change is population pressure caused by its rapid increase in the last two centuries: from under a billion people in the 1800s to over 6 billion in early 2000s. This increase has led to skyrocketed demand for food and fibres, which in turn has led to devastating deforestation required for expansion of agricultural production (Foley et al. 2005). Additionally, even though cities do not cover a large proportion of land, rapid urbanisation since the 19th century has significantly changed many natural ecosystems directly and indirectly. Other primary production systems, such as mining and forestry, have also contributed to the rapid change of the natural ecosystems.

Modelling these changes would enable the understanding of LULC dynamics caused by specified factors and thus provide tools for analysing and monitoring LULC patterns (Verburg et al. 2004), as discussed in Chapter 2. In fact, recent developments in LULC change modelling have produced a number of modelling approaches, ranging from simple regression models to more complex ones, such as Markov-chain and cellular automata (CA). Indeed, the complexity of some of the models is limited by the researcher’s needs and available computing power. In other words, the exponential increase in computing power, in accordance with Moore’s law (Moore 1965), has enabled the development of more integrated and complex models.

The application of these LULC change models is wide-ranging from monitoring biodiversity (Verburg et al. 2008) and vegetation (Echeverria et al. 2008) losses, to modelling the impacts of climatic change (Lasch et al. 2002), as well as tools for policy decision-making
(Beurden et al. 2007), models can and should be used as learning tools (Verburg et al. 2006). As Syphard et al. (2005) put it, while LULC change (may it be change to urban or Agriculture) is a top-down phenomenon, as influenced by policies, environmental constraints and trends of our society, the change itself can also spread from the bottom-up, where patterns emerge from the local characteristics.

There is no consensus on a best approach for land change modelling (Koomen et al. 2007), as elucidated in Chapter 2. Thornton and Jones (1998) proposed a relatively simple agricultural LULC change model based on transition probabilities. This was justified by the fact that models should be useful to answer a broad range of questions, without region specific constraints. For efficiency and cost effectiveness, they proposed an empirical, simplistic top-down model. This model could be used for a broad range of situations and complexity added as needed. In effect, these models would be “temporary hypotheses”, which, if not sufficient, could be substituted by other more complex ones. These models would be based on Markov-chain theory and governed by transition probabilities, which themselves could be extracted from real data (e.g. LULC maps of different years) or estimated by other statistical methods.

5.1.1 An insight in to Markov-chain models

Transition matrix models are based on Markov-chain theory, which was introduced by Andrei Markov in early 1900s. Since then, a number of LULC change models have used the underlying principle of Markov-chains, such as the models proposed by Turner (1987), Thornton and Jones (1998), Baca (2002) and Heldens (2006). There are also more elaborate application tools using Markov-chains, such as Environment Explorer (Engelen et al. 2003) and MOLAND (Barredo and Demicheli 2003).

In describing Markov-chains, the assumption is that, for any given set (A), the conditional probability (P) of occurrence of a state (X_t) is solely dependent upon the previous state (X_{t-1}) for a first-order Markov-chain or states (X_{t-1}+X_{t-2}+...+X_{t-n}) for n^{th}-order Markov-Chains.
(Roberts 1998). Here, the focus is on first-order Markov-chains, in LULC modelling terms, where LULC at time $t$ only depended on LULC at time $t-1$

$$P[X_t \in A \mid X_{t-1}]$$

Equation 5.1

Each state of the matrix could be characterised by a LULC map (comprising of image pixels). Once the two states were known, it would be possible to compute the transition probabilities. An example can be given as a simple array, whereby a set “A” of LULC categories transit from $t-1$ to $t$:

$$A_{t-1} = \begin{bmatrix} 3 & 4 \\ 3 & 3 \end{bmatrix} \rightarrow A_t = \begin{bmatrix} 3 & 3 \\ 4 & 3 \end{bmatrix}$$

Equation 5.2

this example may yield something like the following transition matrix ($TM$) with rows normalised to 1:

$$To: \begin{bmatrix} 3 & 4 \\ 3 & 2 & 1 \\ 4 & 1 & 0 \end{bmatrix} \rightarrow TM = \begin{bmatrix} 0.667 & 0.333 \\ 1 & 0 \end{bmatrix}$$

Equation 5.3

The $TM$ (Equation 5.3) describes the transition probabilities between the different states. Utilising the $TM$ (Equation 5.3) and the set $A_{t(i)}$, it would be possible to determine the future state of $A$, at $t+1$. Every cell that takes the state of “3” is characterised by 0.667 probability of maintaining its state, while every cell in a state of “4” would change to “3”. However, one of the problems of this approach is that this ignores neighbourhood influence of the surrounding cells.
The reason why the neighbourhood influence is ignored in classical Markov-chains is because they were not specifically designed for spatial problems (Balzter 2000). Due to this reason, their application to modelling LULC change posed some challenges. For example, it has been noted that LULC changes are not strictly Markovian (Turner 1987). Even though the probability of change may be Markovian, it is not independent of its neighbours. For instance, a patch of woodland surrounded by cropping areas would be more likely to change, than if that same patch of woodland were surrounded by woodland pixels.

Therefore, in running simple Markov-chain models, which are typically limited by their lack of spatial structure (Urban and Wallin 2002), the first iteration would greatly alter the inherent spatial distribution of pixels and subsequent iterations further disorganise the distribution. To overcome this limitation, a bootstrapping approach could be used, whereby a model could run in a number of years for a number of times, enabling the estimation of a probability distribution for each pixel and computation of a modal map (Baca 2002).

Another challenge posed by classical transition matrix models is the interval at which they run. If the LULC maps, from which the transition probabilities are extracted, have an interval between them of $y$ years, each run of a first-order Markov-chain model is equivalent to $y$ years. However, it is often necessary to run these models on a yearly basis. The problem is not one of scalar mathematics as proposed by Urban and Wallin (2002), but one which is mainly based on matrix algebra, eigenvalues and eigenvectors as seen in Heldens (2006) and more recently in Takada et al. (2010).

The final challenge posed by this approach is determining the $TM$ itself. This could be done by a variety of approaches, from $TMs$ established from LULC data (as shown in Equation 5.2 and Equation 5.3), to utilising other approaches, such as logistic regression and empirical knowledge (covered in Chapter 6). The transition probabilities should be capable of predicting change scenarios. However, as they are static, relating them to dynamic factors, such as future climatic patterns, political, social and economic development can be
problematic. There is, therefore, the need for their modification to make them dynamic, which would enable the incorporation of the dynamic factors listed above. The modification can be achieved through tweaking of regression models and the related correlations that would allow an insight into the future landscape patterns (Hill et al. 2002). This is useful for scenario evaluation (e.g. if the temperature changed by 5°C, how that will affect a certain agricultural region), but also for planning purposes (e.g. given this might happen, where shall investment in infrastructure go to). Detailed elaboration on dynamic transition matrices and their relationship to climate and socio-economic factors will be covered in Chapter 6.

The focus of this chapter is to build a spatially coherent transition matrix model, which could be modified with relative ease and served as a temporary hypothesis. This model would allow aggregation of complexity (Thornton and Jones 1998) such as externally driven transition matrices, detailed in Chapter 6.

In this chapter, two different LULC change models, based on Markov-chains were developed in “R”. The specific aims of this chapter were to:

i) annualise the transition matrices, so that models could be run on a yearly time-step;

ii) build a first-order Markov-chain model without neighbourhood influence;

iii) incorporate neighbourhood influence into the first-order Markov-chain model, through the use of CA; and

iv) test both models’ sensitivity to change in parameters.

5.2 Data and methods

5.2.1 Annualising the transition matrices

The reference LULC maps created in Chapter 3 were used for establishing transition matrices, as described in Wright (2001). The transition matrices were computed by cross-tabulating the results of LULC maps from two different points in time. Here, since the
intervals between the maps were not constant (from 5 to 13 years) and it was deemed necessary to run the model on a yearly time-step, the transition matrices had to be annualised, as suggested in Heldens (2006).

Initially, the procedures described by Urban and Wallin (2002) were used to compute annual transition probabilities. Their approach used simple scalar mathematics to annualise the transition matrices. The process consisted in dividing the off-diagonal cells by the interval (years) and then computing the diagonals by differencing, as each row had to add to one. It was found, however, that these procedures were not robust and did not produce expected results. Annual probabilities produced by these methods, when multiplied to achieve the desired time step (5, 10 or 13 years), did not yield same proportions as using the original transition matrices and therefore it was necessary to find another methodology.

A more robust method of matrix manipulation was found in Whitelegg and Oliver (1975) and more recently in Takada et al. (2010). This methodology of annualising the transition matrices produced the same proportions and quantities as the original transition matrices. The method shown in Whitelegg and Oliver (1975) uses eigenvalues, eigenvectors and matrix multiplication to find a solution, which is summarised by:

\[ TM_{(t)} = TM_{(t)}^t \]  \hspace{1cm} \text{Equation 5.4}

where \( TM \) is transition matrix and \( t \) is the interval in years. Solving this would result in:

\[ TM_{(t)} = \sqrt[1/t]{TM_{(t)}} \]  \hspace{1cm} \text{Equation 5.5}

since \( TM \) are matrices, the solution depends on eigenvalues and eigenvectors that satisfy:

\[ TM_{(t)} = V.A.V^{-1} \]  \hspace{1cm} \text{Equation 5.6}

and thus:

\[ TM_{(t)} = V.A^{(t/t)}.V^{-1} \]  \hspace{1cm} \text{Equation 5.7}

where \( V \) is a matrix of eigenvectors; and \( A \) a matrix with eigenvalues on the diagonal. The solution of Equation 5.7 followed the approaches seen in Takada et al. (2010), Heldens (2006).
and Whitelegg and Oliver (1975). This was implemented using the package “panel” (Gentleman 2006) in R (R development core team 2009). The annualised TMs were used subsequently.

### 5.2.2 Model A: a first-order Markov-chain model

The first model, model A, aimed at mimicking what was done by Turner (1987) and Heldens (2006). It considered every single pixel in the landscape and, according to the transition probability for a given LULC and a random number, it changed or not to a different state. The workflow of model A is illustrated in Figure 5.1.

![Figure 5.1. Conceptual workflow of model A](image-url)
The model shown above could be applied not only to images but also to other data analyses that require first-order Markov-chain algorithms. Although a single run of this model should not produce realistic LULC maps, it should maintain correct proportions of each of the LULC throughout the landscape. In the present case, the model was run for 10 years once. A modal map of the final state of 500 runs of this model was computed, as presented in the results (Section 5.3.2).

It is worth noting that, for the outputs of this models to be more realistic, it was necessary to incorporate some neighbourhood influence in the transition process, as suggested by Baca (2002).

5.2.3 An example for incorporating neighbourhood influence

Turner (1987) proposed three different transition models: i) a first-order Markov-chain model (as in Section 5.2.2), in which changes in LULC were governed solely by transition probabilities and without neighbourhood effects; ii) a first-order Markov-chain model which utilised a von Neumann neighbourhood (four closest neighbours) for determining transitions; and iii) a first-order Markov-chain model using a Moore neighbourhood (eight closest neighbours) for determining transitions. The approach for using the neighbourhood is described below.

Turner’s approach utilised a LULC frequency table and TM to quantify the amount of change in LULCs at time t to LULC categories at time t+1. These could be “stored” in a pool of possible changes, as described by the author. Subsequently, a transition index was calculated for every pixel (using four or eight neighbours) and changes were made to pixels with the highest indices.

Below is an example adapted from Turner (1988), who utilised a LULC map converted to a numerical matrix (A(t)):
\[
A_{(t)} = \begin{bmatrix}
4 & 4 & 4 \\
5 & 3 & 3 \\
3 & 5 & 5
\end{bmatrix}
\]

Equation 5.8

and the transition matrix \((TM)\) for each LULC

\[
\begin{array}{c}
To: \\
3 & 4 & 5 \\
3 & 0.80 & 0.15 & 0.05 \\
From: 4 & 0.00 & 0.95 & 0.05 \\
5 & 0.10 & 0.30 & 0.60
\end{array}
\]

Equation 5.9

Considering the centre pixel \((a_{2,2})\), which has LULC “3” and a Moore neighbourhood, a frequency of each type of neighbour was established. Subsequently a transition index based on probabilities from the matrix above was calculated as:

\[
t_{3,3}N_3 = 0.80 \times 2 = 1.6 \\
t_{3,4}N_4 = 0.15 \times 3 = 0.45 \\
t_{3,5}N_5 = 0.05 \times 3 = 0.15
\]

Equation 5.10

where \(t_{i,j}\) indicates the transition probability from \(i\) to \(j\) extracted from \(TM\) and \(N_i\) indicates the number of neighbours of each LULC class. In this case, although the pixel \((a_{2,2})\) would not change its state, a search algorithm would modify pixels with highest transition indices and subtract these changes from the “pool”. The transition indices would then be recalculated and random pixels selected to be changed. This procedure would then be repeated until the “pool” was empty. A neighbourhood approach very similar to this was used by Baca (2002).

In critically analysing this method, two limitations are identified. The first is that the transitions solely relied upon a predefined amount of change (LULC frequency table multiplied by the \(TM\)), meaning that the change is completely driven by the transition index. Because of this, if the procedure were to be applied to a pre-established landscape, it would yield similar results every time. This indicates that, while Turner’s method is focused on utilising the neighbourhood influence, it ignored emergence of different LULCs in the middle of a homogeneous landscape.
The second limitation is related to the time span of each of the model runs. If the “pool” of change were exhausted after the first iteration, this would equate to one year. A different situation is when the “pool” is exhausted after \( n \) iterations; \( n \) being the number of years, there is the need to allow for evolving patterns, as transition indices were re-calculated at every iteration. If the transition matrix is annualised (Section 5.2.1) and if it is necessary to go through a few iterations to achieve the required amount of change determined by the “pool”, this would mean a multi-year transition, as transition indices would change between iterations.

Alternatively, Heldens (2006) implemented a model based upon Jenerette and Wu (2001) that tried to deal with every possibility of neighbourhood transitions. The author sought to determine transition probabilities for every possible combination of a von Neumann neighbourhood. After deciding this was not practical, since she did not have enough data, the dimensionality of the transition matrix was very much reduced, based on a few assumptions. Such assumptions included, but were not limited to: that intensive use and water pixels would not change; that the only factor affecting the \( TM \) was the presence or absence of a LULC type and not its frequency; and also that certain classes could only appear if there were a neighbour in that class. This approach was considered limited, as some of the assumptions were deterministic.

In summary, while Turner’s approach (Turner 1987) is considered sensible, it does not consider the transition probabilities directly, but only the transition index. Heldens’ approach (Heldens 2006), on the other hand, utilises the transition probabilities. However, some of the transition rules are deterministic or based on empirically derived assumptions, which may not be true for different locations or situations.

5.2.4 Model B: hybrid, weighted Markov-chain and CA, land change model

Considering the limitations above, an alternative method is deemed necessary. In this new proposition, both the transition matrix (\( TM \)) and neighbourhood can be used to determine the
transition probability, such that a final transition matrix ($FinalTM$) is the sum of weighted components:

$$FinalTM = TM \ast W_{TM} + NB\_TM \ast W_{NB}$$

Equation 5.11

where weights should be such that

$$W_{TM} + W_{NB} = 1$$

Equation 5.12

and $W_{TM}$ is the weight of the Markov-chain component, $W_{NB}$ is the weight attributed to the neighbourhood influence (the CA component), $TM$ is the transition matrix and $NB\_TM$ is the neighbourhood transition matrix.

In order for Equation 5.11 to work properly, it would be necessary for the rows of the $TM$ and $NB\_TM$ each sum to one. $TM$ multiplied by $W_{TM}$ is a simple scalar operation, as would be the multiplication of $NB\_TM$ by $W_{NB}$.

The challenge would then be to determine $NB\_TM$. This could be done using CA rules, as described by Balzter et al. (1998) and discussed in Section 2.3.4 of Chapter 2. The transition indices rely on a function similar to Turner (1988), but would need to be normalised to one. In summary:

$$NB\_TM = f(a_{i,j}, n(a_{i,j}), TM_{Row}(a_{i,j}))$$

Equation 5.13

In explaining further, $NB\_TM$ as a function of: $a_{i,j}$ is the pixel value, $n(a_{i,j})$ is its neighbourhood, the $TM_{Row}(a_{i,j})$ would be the appropriate row of $TM$ according to $a_{i,j}$’s value. For each $a_{i,j}$, its value and neighbourhood are computed, $NB\_TM$ component becomes dynamic. A “woodland” pixel surrounded by “woodland” pixels would be characterised by a $NB\_TM$ that will ensure that it stays as “woodland”. However, if it is surrounded by “dryland Agriculture” and “woodland”, $NB\_TM$ would have “woodland” and “dryland Agriculture” components which would purposely modify the $FinalTM$. 

131
Considering the example used for Turner’s method above (Section 5.2.3, page 129) and considering that $TM$ and $NB\_TM$ as the same, then:

\[
W_{TM} = 0.4 \quad \therefore \quad W_{NB} = 0.6
\]

\[
TM * W_{TM} = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.32 & 0.06 & 0.02 \end{bmatrix}
\]

\[
NB\_TM = \begin{bmatrix} 1.6 & 0.45 & 0.15 \\ 2.2 & 2.2 & 2.2 \end{bmatrix}
\]

Equation 5.14

\[
NB\_TM * W_{NB} = \begin{bmatrix} 0.73 & 0.2 & 0.07 \end{bmatrix} * 0.6 = \begin{bmatrix} 0.44 & 0.12 & 0.04 \end{bmatrix}
\]

\[
FinalTM = TM * W_{TM} + NB\_TM * W_{NB}
\]

\[
FinalTM = \begin{bmatrix} 0.76 & 0.18 & 0.06 \end{bmatrix}
\]

In this example, model A (Section 5.2.2) computes that the pixel would have 80% chance of not changing. Turner’s model would also keep the pixel in its state (Equation 5.10, Section 5.2.3), due to the transition index. However, the introduction of the neighbourhood effect (model B) to the same example marginally increased the probability of change from 20% to 24%. This marginal increase was a result of two factors: i) the initially high probability of not changing; and ii) the $W_{NB}$, which if increased would further affect the probability of change.

The hybrid, model B, is therefore advantageous in a number of ways. For instance, it allows the investigation of different settings from a simplistic stochastic model when $W_{TM}$ is set to one, to a completely CA driven transition, when $W_{NB}$ set to one. The hybrid model is also more flexible in that different transition matrices for each component ($TM$ and $NB\_TM$) could be applied. Finally, the current implementation allows the use of four or eight neighbours to influence each pixel, although modifications to the code will enable the influence of a larger neighbourhood.

As already explained, the structure of the hybrid, model B, is such that it allows for improved flexibility, extending from mimicking model A to testing different settings of neighbourhood weights by varying $W_{NB}$ from 0 to 1. The $FinalTM$ was based on a combination of weights from the $TM$ and $NB\_TM$, each set by the user, permitting a von
Neumann or Moore neighbourhood to be tested. The workflow of model B is illustrated in Figure 5.2 below.

Model B was run with different settings of $W_{TM}$ from 0.0 to 1.0 in 0.1 increments and with four and eight neighbours, totalling 22 different setups. This was designed to analyse the effect of different weights and neighbourhood configurations. Moreover, this would validate the premise that a single run of the model using neighbourhood functions would produce outputs that were spatially coherent. In this study, each model setting was run once, 20, 40, 60, 80, 100 and 500 times for 10 years and the modal outputs computed and mapped. Finally,
a sensitivity analysis for different numbers of runs was carried out to discern the model behaviour and stability. This is described in the results Section 5.3.5.

5.2.5 Data and computation

As described above, both models were written in R 2.10.0 (R development core team 2009), using the packages “rgdal” (Keitt et al. 2009) for input/output of georeferenced images and “spdep” (Bivand 2009) to compute neighbourhood of each pixel. As mentioned previously (item 5.2.1), the annualisation of the transition matrices was also implemented in R using the package “panel” (Gentleman 2006). Simulations were done on either an Intel Core 2 6300 with 2 GB of RAM or an Intel Core 2 Duo E8400 and 4 GB of RAM, running Windows XP.

Since the models were characterised as stochastic simulation models, they were run for: 1, 20, 40, 60, 80, 100 and 500 times for each setting, enabling elaboration of modal maps, as well as establishing the distribution of pixel values and understanding of model stabilisation.

Initial conditions for both models were the 1995 LULC map (produced in Chapter 3) and transition probabilities extracted from the 1995-2000 LULC maps. Probabilities were then converted to annual time-steps, as described previously (Section 5.2.1).

While some widely used land change models typically only deal with two classes simultaneously, such as Geomod (Pontius Jr and Malanson 2005), the initial landscape, which is the 1995 LULC map (Figure 5.3 below) was composed of five different classes simultaneously modelled for change.

As mentioned previously, the main objective of this Chapter was to develop and test model A and model B with different settings. Here, the 1995 LULC map (from Chapter 3) was used as the initial condition and the TM that was extracted from 1995 to 2000 LULC maps. This TM was then annualised, as described in Section 5.2.1. As for modifying the TM, this is the subject of the following, Chapter 6. In the present Chapter, the differences between
models and their settings are illustrated by comparing model outputs to the 2005 reference (produced in Chapter 3).

5.2.6 Model comparison and sensitivity analysis

While simulation models A and B are both stochastic, it was deemed necessary to evaluate their sensitivity in order to understand the outputs resulting from changing model parameters. Kocabas and Dragicevic (2006) stated that, while there was much research done in building LULC change models, there is a lack of understanding of the effect of different model parameters on the outputs. Consequently, sensitivity analysis using quantitative and
qualitative components was needed. Furthermore, Hagen-Zanker and Martens (2008) commented that researchers should not limit themselves to simple, single criteria approaches when calibrating and validating models, as this might lead to false conclusions.

For this reason, a multi-criteria approach, based on landscape metrics, such as fractal dimension, shape index, Simpson’s diversity index and edge density (described in McGarigal et al. 2002), could be used. Additionally, standard measures of agreement, such as overall accuracy and Kappa statistics, as recently used by Hagen-Zanker and Martens (2008) and Kocabas and Dragicevic (2006), could also be applied.

In this study, a number of metrics were calculated using the Map Comparison Kit (MCK) (Visser and De Nijs 2006) version 3.2.0. As each model was run with a number of settings, their output comprised of predicted LULC maps from single runs and modal maps from various iterations. In terms of comparison, it should be noted that model A is not amenable to flexible parameter settings and therefore its output was limited to the modal maps of different combinations of iterations. Model B, on the other hand, enables changes in $W_{TM}$ and the number of neighbours, resulting in 22 output maps (four or eight neighbours and $W_{TM}$ of 0.0 to 1.0 in 0.1 increments) for each set of iterations. These maps were each compared to the 2005 reference LULC map from Chapter 3.

The error matrix from MCK provided us with overall accuracies (described in Jensen 2005) along with overall Kappa (Cohen 1960) and its decomposition to location and histogram Kappa (Hagen 2002). Furthermore, MCK also provided the landscape metrics, of which fractal dimension, Simpson’s diversity index and edge density were used. These landscape metrics were chosen, as they illustrate the spatial coherence of the output. Such metrics indicated shape complexity (fractal dimension and edge density) and diversity of LULC patterns (Simpson’s diversity index). Furthermore they have been used by a number of authors to evaluate LULC map patterns (Jenerette and Wu 2001, Parker and Meretsky 2004, Yanyan et al. 2008).
For the sake of brevity, only a few error matrices will be shown. However, plots that compare the different model setups to the reference will be produced. It is acknowledged that changing the \( TM \) would have a large impact on the outcomes. Due to its importance, modification of the \( TMs \) is the main focus of the following Chapter 6.

### 5.3 Results and discussion

#### 5.3.1 Annual transition matrices

The multi-temporal 50 m resolution LULC maps from Chapter 3 were used to compute the \( TMs \) using R program (R development core team 2009). These \( TMs \) are essential for land change modelling. As described in Chapter 3, the maps had variable intervals ranging from 5 to 13 years. It was therefore necessary to transform the computed \( TMs \) to annual ones. Here the results of annualising the 1995-2000 matrix will be shown, as it will be used in model testing. However, the same procedure was applied to the other \( TMs \).

<table>
<thead>
<tr>
<th>in %</th>
<th>Woodland</th>
<th>Dryland Agriculture</th>
<th>Irrigated Agriculture</th>
<th>Intensive uses</th>
<th>Water bodies</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
<td>29.69</td>
<td>5.63</td>
<td>0.29</td>
<td>0.33</td>
<td>0.04</td>
<td>35.98</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>4.19</td>
<td>41.85</td>
<td>3.07</td>
<td>1.39</td>
<td>0.32</td>
<td>50.81</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>0.36</td>
<td>2.88</td>
<td>6.69</td>
<td>0.14</td>
<td>0.17</td>
<td>10.25</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>0.19</td>
<td>0.32</td>
<td>0.08</td>
<td>1.89</td>
<td>0.01</td>
<td>2.49</td>
</tr>
<tr>
<td>Water bodies</td>
<td>0.05</td>
<td>0.11</td>
<td>0.02</td>
<td>0.02</td>
<td>0.28</td>
<td>0.47</td>
</tr>
<tr>
<td>Total</td>
<td>34.47</td>
<td>50.80</td>
<td>10.15</td>
<td>3.77</td>
<td>0.81</td>
<td>100.00</td>
</tr>
</tbody>
</table>

The change from 1995 to 2000 is summarised in Table 5.1. It can be seen that there was some swap between “dryland Agriculture” and “woodland”, as well as between “dryland Agriculture” and “irrigated Agriculture”. “Intensive uses”, which comprised roughly 2.5% of the landscape in 1995, increased in area to about 3.8% of the landscape in 2000. This was likely caused by the establishment of a new golf course, as well as the expansion of the city of
Cessnock. “Water bodies” also increased from about 0.5% of the landscape in 1995 to 0.8% in 2000. This was caused by the development of new dams by farm owners as well as the addition of artificial lakes in or near golf courses.

Even though this matrix allows the understanding of changes that had occurred from 1995 to 2000, it was based on a 5 year gap. Iterations in any model derived directly from this matrix would run at 5 year time interval. A logical solution was therefore to “annualise” this matrix.

As mentioned earlier, the first attempt to annualise the TMs followed the procedures described by Urban and Wallin (2002). While their approach was a simple one, the resulting annualised TM did not produce robust results, as described in Section 5.2.1.

The second approach used here was based on the methodology of Whitelegg and Oliver (1975) and Takada et al. (2010). This approach was described in detail in Section 5.2.1 and summarised by Equation 5.7 (Page 126). This methodology utilised eigenvectors and eigenvalues, addressing the problem that matrix calculations were not scalar. This solution was implemented in R (R development core team 2009) using the “panel” package (Gentleman 2006). The resulting annual TM for 1995-2000 is presented in Table 5.2. The R code used to annualise these matrices is shown in Appendix A.

This TM (Table 5.2), with minor adjustments, was utilised for model testing as described in the following sections.

<table>
<thead>
<tr>
<th>1995</th>
<th>Woodland</th>
<th>Dryland Agriculture</th>
<th>Irrigated Agriculture</th>
<th>Intensive uses</th>
<th>Water bodies</th>
</tr>
</thead>
<tbody>
<tr>
<td>95.86</td>
<td>3.80</td>
<td>0.00</td>
<td>0.00</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>3.45</td>
<td>93.14</td>
<td>2.92</td>
<td>0.49</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>2.47</td>
<td>6.52</td>
<td>90.77</td>
<td>0.24</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>3.25</td>
<td>2.37</td>
<td>2.12</td>
<td>92.26</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>4.07</td>
<td>5.06</td>
<td>2.27</td>
<td>0.68</td>
<td>87.92</td>
<td></td>
</tr>
</tbody>
</table>
5.3.2 Model A: a first-order Markov-chain model

As stated in the method section (5.2.2), model A was limited by the lack of spatial structure. This is evident in Table 5.3, which shows the error matrix comparing the output of model A for 2005 to the 2005 reference. It could be seen that overall accuracy was around 51%, while overall Kappa had a very poor agreement (Viera and Garrett 2005) of 0.24. These low values can be related to inherent characteristics of the model being completely stochastic (Baca 2002) and indeed non-spatial. This low value of location Kappa can be explained by the fact that the chance of a “woodland” pixel changing to any other LULC was independent of its location and driven exclusively by the TM.

Poor agreement of first-order Markov-chain models without neighbourhood influence have been reported previously in Turner (1987) and also by Heldens (2006). On the other hand, histogram Kappa showed that there was almost perfect agreement based on the frequency of each LULC class. This is expected, since transition probabilities should mimic the amount of change, but were known not to perform well when predicting its location (Lambin et al. 2000).

| Table 5.3. Error matrix of model A (1 run, 10 years) versus reference |
|---|---|---|---|---|---|---|
| in hectares | Reference 2005 LULC map |
| model A: 1 run for 10 years | Woodland | Dryland Agriculture | Irrigated Agriculture | Intensive uses | Water bodies | Total |
| Woodland | 6553.3 | 3655.5 | 927.3 | 355.5 | 46.8 | 11538.3 |
| Dryland Agriculture | 3101.5 | 6823.0 | 1732.3 | 429.3 | 74.0 | 12160.0 |
| Irrigated Agriculture | 584.0 | 1951.5 | 1096.0 | 149.8 | 32.5 | 3813.8 |
| Intensive uses | 138.3 | 368.5 | 104.0 | 231.5 | 5.3 | 847.5 |
| Water bodies | 151.5 | 71.5 | 11.3 | 7.8 | 21.3 | 263.3 |
| Total | 10528.5 | 12870.0 | 3870.8 | 1173.8 | 179.8 | 28622.8 |
| Location Kappa | 0.26 | | | | | |
| Histogram Kappa | 0.94 | | | | | |

For the purpose of comparison, model A was run for 10 years, 500 times and a modal map created. The outcome of the modal maps was compared to the reference, resulting in Table 5.4. In this (Table 5.4), the overall accuracy and Kappa are much better: 76.99% and 0.63.
respectively. However, histogram Kappa is slightly lower than in the case for the model run once (Table 5.3), but still in the range of an almost perfect agreement, according to Viera and Garrett (2005). The location Kappa, on the other hand, is much higher than that of the single run model, jumping from 0.24 to 0.70. This means that the modal map of 500 runs is more representative of the reality than of a single run.

<table>
<thead>
<tr>
<th>Table 5.4. Error matrix of model A (500 runs, 10 years) versus reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>model A: 500 runs for 10 years</td>
</tr>
<tr>
<td>Woodland</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
</tr>
<tr>
<td>Intensive uses</td>
</tr>
<tr>
<td>Water bodies</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

To illustrate the outcome for multiple-runs of the model versus the case for single-runs, Figure 5.4 below provides the comparative LULC maps. While Figure 5.4 a) shows initial conditions (t), Figure 5.4 b) represents the LULC after a single iteration for a single year (t+1). Due to the lack of influence of neighbouring cells, the landscape appears to be noisy, meaning that random pixels were converted to other LULCs. This is even more evident in Figure 5.4 c), which represents the structure of a single run of model A for 10 years (t+10). Most of the landscape’s inherent spatial structure was suppressed by the stochastic nature of model A. On the other hand, Figure 5.4 d) shows the modal outcome of 500 iterations. It can be seen that some changes to the landscape patterns are evident in Figure 5.4 a), compared to Figure 5.4 d). Since the latter is an amalgamation of most frequent outcomes of 500 iterations, such changes should be expected.

However, it is important to point out that there is some loss of information in Figure 5.4 d) as it is not representative of the frequency distribution of individual pixels. In other words, as
the most frequent outcome is shown, Figure 5.4 d) does not indicate whether the occurrence is in 90% of the cases or 51%.

Figure 5.4. LULC maps generated by model A: a) initial conditions; b) after 1 year; c) after 10 years; and d) modal map of 500 runs
As pointed out above, the performance of model A was limited by its stochastic nature, as well as it being a transformation model (Koomen and Stillwell 2007). Moreover, it is wholly driven by static transition probabilities, which were known not to be completely independent (Turner 1988). In this sense, it was deemed necessary to incorporate the neighbourhood influence in the transition process. Therefore the emergence of different LULCs over time was possible, but not plausible, as there was no spatial structure associated driving the process and the likelihood of this happening independently was very low. By introducing neighbourhood functions, through the use of CA, more realistic results would be expected. These would allow the emergence of different LULCs while maintaining an organised spatial structure. This led to the development of the hybrid, model B, the focus of the next section.

5.3.3 Model B: hybrid, weighted Markov-chain and CA, land change model

In this section, only a few examples of the outcomes of model B are shown, however a detailed account of the different model setups is shown in Section 5.3.5 Sensitivity analysis.

Model B compensates for the lack of neighbourhood effect in model A, as it is required for model outputs that are spatially coherent (Balzter et al. 1998). In doing so, some flexibility was introduced to the model in terms of weight associated with the neighbourhood pixels. This flexibility meant that model B produced variable results.

A change in $W_{TM}$ would greatly affect not only the overall accuracy, but also location and overall Kappa. This is demonstrated in Table 5.5 where $W_{TM}=0.2$ and Table 5.6, where $W_{TM}=0.8$. Note the decrease in overall accuracy and location Kappa from $W_{TM}=0.2$ to $W_{TM}=0.8$. 
Table 5.5. Error matrix of model B (1 run, 10 years), four neighbours and $W_{TM}=0.2$ versus reference

<table>
<thead>
<tr>
<th>in hectares</th>
<th>Reference 2005 LULC map</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Woodland</td>
</tr>
<tr>
<td>Woodland</td>
<td>8528.8</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>1796.5</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>129.8</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>65.0</td>
</tr>
<tr>
<td>Water bodies</td>
<td>8.5</td>
</tr>
<tr>
<td>Total</td>
<td>10528.5</td>
</tr>
</tbody>
</table>

Location Kappa 0.66  Kappa 0.58
Histogram Kappa 0.89 Overall accuracy (%) 74.33

Increasing the weight of $W_{TM}$ resulted in increasing histogram Kappa, which is expected from a stochastic model. However, when $W_{TM}$ is decreased (more neighbourhood influence), the location Kappa increased substantially from 0.41 ($W_{TM}=0.8$) to 0.66 ($W_{TM}=0.2$). This indicates that the neighbourhood, as a factor of change, enabled spatial coherence in the model (Balzter et al. 1998, Soares-Filho et al. 2002).

Table 5.6. Error matrix of model B (1 run, 10 years), four neighbours and $W_{TM}=0.8$ versus reference

<table>
<thead>
<tr>
<th>in hectares</th>
<th>Reference 2005 LULC map</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Woodland</td>
</tr>
<tr>
<td>Woodland</td>
<td>7459.0</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>2591.5</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>327.5</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>80.5</td>
</tr>
<tr>
<td>Water bodies</td>
<td>70.0</td>
</tr>
<tr>
<td>Total</td>
<td>10528.5</td>
</tr>
</tbody>
</table>

Location Kappa 0.41  Kappa 0.37
Histogram Kappa 0.92 Overall accuracy (%) 60.74

The modal outcome of model B with different parameter settings had overall accuracies which did not change much, from 77.0% to 77.2% (shown in Figure 5.6). Overall Kappa was around 0.63 and, when split to histogram and location Kappa, had values of ~0.89 and ~0.70 respectively. A sample error matrix will be shown here, where $W_{TM}=0.5$ and four neighbours were used (Table 5.7).
Table 5.7. Error matrix of model B (500 runs, 10 years), four neighbours and $W_{TM}=0.5$ versus reference

<table>
<thead>
<tr>
<th>in hectares</th>
<th>Reference 2005 LULC map</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Woodland</td>
</tr>
<tr>
<td>Woodland</td>
<td>8782.5</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>1554.5</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>118.3</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>72.3</td>
</tr>
<tr>
<td>Water bodies</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>10528.5</strong></td>
</tr>
</tbody>
</table>

Location Kappa: 0.71
Histogram Kappa: 0.89
Location Kappa: 0.63
Overall accuracy (%): 77.11

Utilising four or eight neighbours did not produce large differences in model outcome, as illustrated in Figure 5.5 below. There is little noticeable difference between the output from four neighbours (Figure 5.5 a) and Figure 5.5 c)) or eight neighbours (Figure 5.5 b) and Figure 5.5 d)). Spatial structure was greatly affected when varying $W_{TM}$. Whereas Figure 5.5 a) and Figure 5.5 b), which were produced from $W_{TM}=0.2$ had a coherent spatial configuration, with little noise, Figure 5.5 c) and Figure 5.5 d) have a lot of noise (speckle), even though their histogram Kappa are higher.

This indicated that different model setups would be more appropriate in answering different research questions. If the objective were to establish the frequency of each LULC class, then a model with a high $W_{TM}$ would be appropriate. If the main concern were to reproduce the spatial structure of the landscape, then a model with a low $W_{TM}$ should perform better. This is better illustrated in Section 5.3.5.

The hybrid, model B, represents a new development in what had previously been proposed by Turner (1987), Baca (2002) and Heldens (2006). The aforementioned approaches either were limited to the first-order Markov-chain concept, relying exclusively on the $TM$, or were driven by the neighbourhood through a transition index. Here, however the user is allowed to set the model to be completely stochastic, neighbour-driven or any combination in between, through changing $W_{TM}$.  

144
Figure 5.5. LULC maps generated by model B after 10 years: a) four neighbours, \( W_{TM} = 0.2 \); b) eight neighbours, \( W_{TM} = 0.2 \); c) four neighbours, \( W_{TM} = 0.8 \); and d) eight neighbours, \( W_{TM} = 0.8 \)

This flexibility associated to the hybrid, model B, allowed for the establishment of more realistic LULC predictions. Model B is therefore equivalent to a hybrid, Markov-chain
(Balzter 2000) with CA rules (Kocabas and Dragicevic 2006). The advantages of the hybrid model include, but are not limited to:

- the amount of transition was semi-stochastic and not pre-determined, as proposed in Turner (1987) and Baca (2002). Weights allowed for the model to be completely stochastic ($W_{TM}=1.0$) to completely CA driven ($W_{TM}=0.0$);
- the influence of neighbourhood could be set by the user and was not pre-determined (Heldens 2006). This was an advantage, but also required experimentation. Changing weights of neighbourhood and transition probabilities was, however, very straightforward, as model flexibility allowed for experimentation and optimisation of settings; and
- the current implementation of the model was set up to run with four or eight nearest neighbours, however it could be modified to include more neighbouring pixels if needed.

5.3.4 Computation efficiency

The study area was composed roughly of 170,000 pixels. As each pixel needed to be processed individually, the complexity of each model determined the run time. Model A generated a random number and would change (or not) pixel value, depending exclusively on $TM$ (Section 5.2.2). It would then proceed to the following pixel, doing so for the entire image and reiterating for however many years and runs.

Model B, on its turn, considered neighbouring pixels to affect transitions (Section 5.2.4). Therefore, its first step was to calculate index location of each pixels neighbours, which was done using the package “spdep” (Bivand 2009). This process took about 2.4 hours to run on a Core 2 Duo E8400.

Subsequently, $FinalTM$ (Equation 5.14) was computed according to user defined weights and according to neighbourhood configuration of each individual pixel. A random number was then generated and compared to the modified transition probabilities, with neighbourhood
influence. Both models ran for a user-defined amount of years, along with a user-defined amount of iterations. The final step was to compute the most frequent outcome, for each individual pixel and output a simulated LULC map, as well as a table with frequency distribution of individual pixels.

Table 5.8 illustrates approximate run times of the models with different configurations. It shows that model A was much quicker to run than B, since it was simpler. Utilising four or eight neighbours did not alter running times much, as determining the index location of the neighbours was done previous to the model run.

<table>
<thead>
<tr>
<th>model</th>
<th>Neighbourhood</th>
<th>Years</th>
<th>Iterations</th>
<th>CPU</th>
<th>Approximate run time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>n/a</td>
<td>1</td>
<td>1</td>
<td>Intel Core 2 6300</td>
<td>0.43</td>
</tr>
<tr>
<td>A</td>
<td>n/a</td>
<td>10</td>
<td>1</td>
<td>Intel Core 2 6300</td>
<td>1.34</td>
</tr>
<tr>
<td>A</td>
<td>n/a</td>
<td>10</td>
<td>500</td>
<td>Intel Core 2 6300</td>
<td>622.80</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>Intel Core 2 Duo E8400</td>
<td>1.44</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>10</td>
<td>1</td>
<td>Intel Core 2 Duo E8400</td>
<td>14.40</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>10</td>
<td>500</td>
<td>Intel Core 2 Duo E8400</td>
<td>7416.00</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>Intel Core 2 Duo E8400</td>
<td>1.47</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>Intel Core 2 Duo E8400</td>
<td>14.70</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>10</td>
<td>500</td>
<td>Intel Core 2 Duo E8400</td>
<td>7546.60</td>
</tr>
</tbody>
</table>

5.3.5 Sensitivity analysis

It was further necessary to understand the effects of varying models parameters as pointed out by Kocabas and Dragecivic (2006). For this reason, a multi-criteria sensitivity analysis was carried out as suggested by Hagen-Zanker and Martens (2008). It should be noted that the analysis, as undertaken here, should not be viewed as model accuracy assessment. The accuracy assessment and modification of the TMs is treated in the following, Chapter 6. However, some of the classical map accuracy criteria were used here to test the sensitivity of the models.

It should be noted that model B with a $W_{TM} = 1$ gives the same results of model A. So the parameters that can be set in model B are: $W_{TM}$, which follows that $W_{TM} + W_{NB} = 1$; number of neighbours (4 or 8); and the number of runs. Adjusting the $W_{TM}$ is dependent on the research
question which needs addressing. If the research question is related to predicting the proportions of different LULC classes, then $W_{TM}$ should be set as a high value, however when doing the spatial simulation this will incur in the loss of spatial patterns, so the resulting map is very speckled. If the research question is more related to predicting the spatial patterns of change, then $W_{TM}$ should be low. The consequence of the latter is that the proportions of LULC classes might not follow the expected pattern. The number of neighbours is set to determine the size of the neighbourhood to be considered and this is empirically determined. As for the number of runs, this will determine the most frequent outcome. If the intent of the scientist is to look at a single snapshot in time, then the number of runs should be 1. If, however, the intent is to see the distribution of possible outcomes, then the number can be set higher.

The results of the sensitivity analysis are shown in Figure 5.6 through Figure 5.13. As shown in Figure 5.6, the modal outcomes of models A and B have similar accuracies (>76%), irrespective of changing $W_{TM}$ in model B. This could be anticipated, since both model outputs came from multiple iterations. However, when this outcome is compared with the results of a single run, the overall accuracy of model B based on both four and eight neighbours decreased considerably from around 76% to 52%, as $W_{TM}$ increased. Model B based on $W_{TM}=1$ mimicked model A, indicating that both were functioning properly. When the model B was driven mainly by the neighbourhood (low $W_{TM}$), edges were allowed to grow, consequently producing more realistic LULC patterns and higher overall accuracies.
Pontius Jr and Millones (2008) pointed out that Kappa based indices of agreement should not be taken blindly as sole indicators of model performance. Here, location and histogram Kappa were used to test models’ A and B sensitivity. For the latter, $W_{TM}$ varied from 0.0 to 1.0. The results indicate that increasing values of $W_{TM}$ decreased the location Kappa (Figure 5.7) for single runs. The reverse is the case when the test criterion is histogram Kappa (Figure 5.8) for single runs.

In regards to the modal outcomes, a CA driven model (low $W_{TM}$) should mimic locations better, while a completely Markov-chain model ($W_{TM}=1$) is more representative of amounts of
change. But, because Markov-chain models performed poorly when predicting location (Balzter et al. 1998), a correct balance between the weights and proper calibration of transition matrices was needed to produce acceptable results.

Figure 5.8. Sensitivity of models A and B to varying $W_{TM}$, as measured by histogram Kappa

Other criteria used to test the sensitivity of changing model parameters were a number of landscape metrics. The first metric used was the fractal dimension (FD), which describes the complexity of shapes in the landscape patterns (Turner and Ruscher 1988). It has been used for analysing LULC maps by a number of authors (Jenerette and Wu 2001, Hagen-Zanker and Martens 2008) and its values can vary between 1 and 2. Figure 5.9 shows the FD values for the output maps of model A and B with varying $W_{TM}$. It also shows the reference FD, extracted from the reference 2005 LULC map, as being around 1.4. Models A and B were considered to perform better when the deviation from the reference FD was minimal. For a single run of model B, the deviation from the reference FD increased with $W_{TM}$. However, the models driven by eight neighbours had slightly higher deviations. When considering the modal outcomes of model B, their FDs were lower than that of the reference FD ($W_{TM}<0.8$), with an increasing trend that matched the reference FD at $W_{TM}$ of about 0.9.
In a previous study, Yanyan et al. (2008) used Simpson’s diversity index to study changes in LULC in China. This index represents the probability of two random pixels within the landscape being assigned to different LULC classes (Lasch et al. 2002, McGarigal et al. 2002). In this study, a number of single runs of model B with increasing $W_{TM}$ approached the reference value (Figure 5.10), leading to a false conclusion that the output of model B done with high $W_{TM}$ lacks spatial structure. The behaviour of modal outcomes shows a similar trend to those of the single runs, but not as pronounced. The reference value was slightly over 0.64 and the outcomes of model A and B varied from 0.585 to 0.64, thus these differences were within a small (10%) range.

Figure 5.9. Sensitivity of models A and B to varying $W_{TM}$, as measured by fractal dimension
Figure 5.10. Sensitivity of models A and B to varying $W_{TM}$, as measured by Simpson’s diversity index.

Edge density, representing the number of edges of different LULCs within a given area, has been used by various authors (Jenerette and Wu 2001, Parker and Meretsky 2004) for LULC model assessment. Figure 5.11 shows that the reference edge density was around 0.22. Single runs of model B with a low $W_{TM}$ (~0.1) had the same edge density as the reference map. This was also an indicator that increasing $W_{TM}$ generated more edges than necessary, which implied a general lack of spatial structure. Modal outcomes of models A and B had lower values than the reference value.

Figure 5.11. Sensitivity of models A and B to varying $W_{TM}$, as measured by edge density.
In addition to the figures shown above (Figure 5.6 to Figure 5.11), surface plots of $W_{TM}$ and modal maps produced by different number of runs were created for overall accuracy (Figure 5.12) and edge density (Figure 5.13) to illustrate the behaviour of these metrics. This allowed a further understanding of the relationship of $W_{TM}$ and the modal outcomes, varying the number of runs. The surfaces shown in Figure 5.12 and Figure 5.13 are those of the models run with eight neighbours. The surface plots for four neighbours (not shown) were also computed and support the results shown here.

![Surface plot of overall accuracy, number of runs, and $W_{TM}$](image)

**Figure 5.12.** Sensitivity of models A and B, as measured by a three-dimensional surface plot of overall accuracy, the number of runs and $W_{TM}$ (using eight neighbours).

The surface plot (Figure 5.12 above) indicates that increasing the number of runs greatly altered overall accuracy. It also shows that, by decreasing $W_{TM}$, the accuracy increased. Moreover, it is also worth noting that the model run with a low $W_{TM}$ achieved a semi-steady state after just 20 runs. In contrast, models with more than 80 runs were necessary to achieve a similar state when $W_{TM}=1$. This led us to accept the premise that, when considering a single
run of the model, a low $W_{TM}$ was desirable and indeed it should produce more realistic results, preserving spatial structure.

In terms of edge density, the reference value was around 0.22 (already mentioned in Figure 5.11). The surface plot represented by Figure 5.13 below shows consistent decreasing values with increase in the number of runs and decreasing values of $W_{TM}$. For a single run, $W_{TM}=0.1$ closely matched the reference, as well as when $W_{TM}=0.8$ with 20 runs and $W_{TM}=0.9$ with 40 runs. This means that the lower values of $W_{TM}$ associated with increased number of runs underestimated edge density.

Based on these accuracy and landscape metrics, a few noteworthy findings could be deduced. Model B with $W_{TM}=1$ showed the same results as model A. Not only was this expected, but, more importantly, confirmed that both models were functioning correctly.

Figure 5.13. Sensitivity of models A and B, as measured by a three-dimensional surface plot of edge density, the number of runs and $W_{TM}$ (using eight neighbours)

By increasing the weight attributed to the CA component, the models preserved or recovered the spatial structure (fractal dimension, location Kappa, edge density and overall
accuracy), while utilising a higher $W_{TM}$, this reproduced the frequency of LULCs better (histogram Kappa, Simpson’s diversity index). Turner (1987) showed the effect of neighbourhood in a transition model, but as described previously, the amounts of change were pre-determined. A study by Balzter et al. (1998) also showed preservation of spatial structure when utilising CA rules for neighbourhood. Here, a hybrid model, capable of incorporating weight and CA components to the transition matrices was developed. The hybrid model is also amenable to fine tuning, which could further improve its performance.

Although it is well known that transition matrices should affect model’s outcomes (Baca 2002, Heldens 2006), this was not attempted in the present chapter, but will be the subject of Chapter 6. This chapter was specifically aimed at developing these models and illustrating their sensitivity to change in parameters.

### 5.4 Conclusions

Two different transition matrix models were developed. The first was a first-order Markov-chain model with no neighbourhood effects (model A). The second model incorporated the influence of neighbourhood, through a weighted CA and had flexible settings (hybrid, model B).

- While transition matrices ($TM$ and $NB_{TM}$) were kept constant, $W_{TM}$ and number of neighbours were changed in order to test the model’s sensitivity. $W_{TM}=1$ indicated a completely Markov-chain driven model, with no neighbourhood influence and in which even though quantities of change were better represented, spatial structure was lost.
- When $W_{TM}=0$, transitions were exclusively driven by neighbourhood configuration and $NB_{TM}$. This situation better simulated locations of change, but did not perform so well in representing the amount of change.
• Utilising four or eight neighbours did not alter outcomes much, although small changes could be identified.

• It would be possible to incorporate the effect of more neighbours, but this should only be explored if, after correct parameterization of $NB\_TM$, it was found that the number of neighbours had a large influence in the outcome.

• A number of different metrics were computed and shown, both as a function of changing $W_{TM}$ and as surface plots, showing modal outcomes of different number of iterations. These metrics indicated that there should be an optimal model setting, which should be experimented with.

• Most importantly, correct parameterization of transition matrices should be done, in order to assess model performance under a different set of conditions. This exercise of modifying transition matrices is reserved to the next chapter, where model B will be calibrated, for optimal prediction of future LULC patterns.

• Since these models were written in R and the code is published in the appendices (appendix B and appendix C), these models can be used for a number of different objectives:
  a) as a teaching/learning tool, as the code is public and change of parameters is relatively simple;
  b) to further understand LULC change dynamics by using the model with different $TMs$ and $NB\_TMs$; and
  c) to generate scenarios of future LULC patterns.
5.5 References


Bivand, R., 2009. Package "spdep". 0.4-56 ed.: R foundation for statistical computing, spdep: Spatial dependence: weighting schemes, statistics and models.


Gentleman, R., 2006. Package "panel". 1.0.6 ed.: R foundation for statistical computing, Panel: functions and datasets for fitting models to panel data.


Chapter 6. On the transition matrix and its parameterization

* Part of this work was presented at the Advance Institute in Ecosystem Services Valuation and Modelling, held by the GLP, at the University of Hokkaido in Sapporo, Japan, from the 9th to the 13th of August 2010
6.1 Introduction

As covered in Chapter 5, it is widely accepted that land change models are useful tools not only for scientific investigations, but for environmental planning and monitoring by community and government agencies (Parker et al. 2002, Verburg et al. 2004, Schaldach and Priess 2008). In Chapter 5, two land change models were developed and their sensitivity to change in parameters was shown; however, parameterization of the transition matrix, due to its importance, was assigned a chapter to itself. Here a variety of different methods to parameterize the transition matrix were used, within R, to predict landscape patterns.

Osaragi and Aoki (2006) discussed a number of limitations of Markov-chain models including their lack of spatial structure and the static nature of the transition probabilities, often extracted from simply computing the transitions between two points in time. The hybrid model proposed in Chapter 5 addressed the first limitation. Thus, the lack of spatial structure in terms of influence of neighbourhood was addressed by a novel approach, incorporating a weighted cellular automaton driving the transition matrix. In this way, the transition matrix changes in accordance with the nature of the central pixel and configuration of its immediate neighbourhood. Furthermore, the weighting enabled the user to pre-define the level of influence of each component, namely, the Markov-chain and cellular automata.

In the case of the second limitation, this chapter explores various approaches to deal with the static transition probabilities and how the dynamic transition matrices, incorporating biophysical and socio-economic factors, are implemented. Here, transition matrices are not only extracted from two points in time, but are extracted using a variety of methods proposed in the literature, including the incorporation of multinomial logistic regression (MNLR) to create transition matrices related to specific variables.

It is widely known that land change models vary greatly in terms of their assumptions and objectives (Brown et al. 2004, Pontius Jr and Malanson 2005), mainly because there are a
great number of spatially-explicit land change models with different purposes (Agarwal et al. 2002, Koomen and Stillwell 2007). Nevertheless, while this multitude of models is beneficial to the community, as they address a variety of problems (Brown et al. 2004), there is no consensus on how to properly drive changes (Veldkamp and Lambin 2001).

For the specific case of Markov-chain models, where changes are driven by transition matrices, selecting and/or producing appropriate and useful transition probabilities is the main challenge. There are different approaches to achieve this (Parker et al. 2002), such as: static trends (Houet and Hubert-Moy 2006, Huang and Cai 2007, Xiongwei 2008), averaging (Baca 2002), logistic regression (Mertens and Lambin 2000, Geoghegan et al. 2001), empirical modelling (Turner 1988, Heldens 2006, Huang and Cai 2007), genetic algorithms (Tang et al. 2007), amongst other approaches, all of which have been used for forecasting LULC. Each of the above may fit a specific purpose and, while it is not the intention of the work presented in this chapter to discuss all of them, some will be highlighted and compared.

Another challenge for a successful implementation of a land change model is the latter’s calibration and validation. There is no consensual approach for validation and calibration of a given land change model, although they vary from face value inspection, accuracy statistics and spatial metrics. However, Verburg et al. (2006) point out that scientists should state clearly to which degree a model is valid, so that improvements could be made to it.

While calibration, a term used for parameterization of a model, allows the model to reproduce patterns seen in the data used to create it (Verburg et al. 2006), validation can be applied to an independent dataset to assess model performance. It is now well understood that both are necessary (Kok and Veldkamp 2001, Pontius Jr and Schneider 2001). While calibration can only reproduce the land use and land cover (LULC) patterns previously identified, validation of future LULC is impossible (Kok et al. 2001).

From the foregoing discussion, this chapter deals with the calibration of the hybrid model, in which various approaches for generating transition matrices are assessed. The chapter is
focused on predicting future scenarios of future LULC patterns, based on biophysical and socio-economic drivers of change. As such, LULC change at the so called macro-scale (Brown et al. 2004) are linked to the dynamic nature of transition matrices. The chapter also explores the patterns that emerge from the interaction and influence of neighbours. This enables the simulation of future spatial patterns of LULC as driven by the biophysical and socio-economic factors.

The specific aims of this chapter are therefore to:

i) correlate socio-economic and biophysical changes to the transition probabilities;

ii) use different approaches found in the literature for parameterizing the transition matrix within R;

iii) run the model with different transition probabilities and $W_{TMS}$ to simulate the 2005 LULC, therefore enabling model performance assessment; and

iv) simulate future LULC patterns based on different forecasts of biophysical and socio-economic indicators.

6.2 Data and methods

6.2.1 Study area and brief history

The study area was composed of a portion of the lower Hunter Valley of NSW, Australia. The general features of the study area were described in Chapter 3, but a brief history of the area is appropriate here.

Interviews conducted with the landowners in the area (Tinkler et al. 2008) indicated that first plantations of wine grape were established in the 1800s. Until the 1960s the region was characterised by large land holdings. By the 1970s, some of the wineries became insolvent, while large companies continued to thrive. The early 1980s were characterised by removal of
grape vines with government subsidies (O'Neill 2000, Tinkler et al. 2008), however in the 1990s smaller plots were planted along with the increase of the tourism industry.

These facts, as revealed by the interviews of Tinkler et al. (2008), were corroborated by McManus (2008), who pointed out that even though the climate and soils of the Hunter were not optimal for wine making, the chief reason for the region’s winery success was related to its proximity to major urban centres, such as Sydney and Newcastle. Moreover, the ease of access, due to development of the major Sydney-Newcastle expressway and other road expansion in the 1990s, also contributed to the area’s development.

O'Neill (2000) also reported that vine plantations which were first established in the region in the 1800s, declined towards the end of that century due to economic depression. However, in the mid 1900s, plantations were re-established and big vineyards resumed supplies of grapes to the newly built, medium to large wineries. More recently in the 1990s, a number of small boutique wineries have been established, leading to the development and rapid expansion of accommodation and leisure activities to meet the tourist needs. As a result of this expansion, in the 1990s, the district was generally referred to as the Wine Country (O'Neill 2000), having been transformed from a traditional place of wine production to a place of “lifestyle consumption”. Both Australian and international visitors trooped to the region on holidays, for leisure, indulging in wine, but also in fine food, as well as boutique accommodation and services.

Specific to the work covered in this thesis, in order to understand how this shift has transformed the Hunter, the characterisation of the landscape of the Wine country was covered in Chapter 3. As a result, LULC maps were created at selected time intervals, following the broadest level of the ACLUMP (Bureau of Rural Sciences 2006). These maps were then used for extracting transition probabilities as described in Chapter 5. As stated above, this chapter is focused on modelling the past LULC change and showing the effect of driving the hybrid
model with different transition matrices. The modelling process could enable the simulation of future LULC needed for different purposes.

<table>
<thead>
<tr>
<th>Table 6.1. Proportions of LULC categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
</tr>
<tr>
<td>Intensive uses</td>
</tr>
<tr>
<td>Water bodies</td>
</tr>
</tbody>
</table>

Based on the results of multi-temporal LULC classification in Chapter 3, Table 6.1 above summarises the composition of the Hunter’s landscape between 1972 and 2005 at a decadal interval. Evidently there is increase in area under “irrigated Agriculture” (dominated by wine grapes) from the mid 1990s. This increase was accompanied by concomitant expansion of “intensive uses”, characterised by the construction of parks, golf courses and hotels. Comparably, while in the 1970s the combined area of “intensive uses” and “irrigated Agriculture” corresponded to roughly 12% of the landscape; in 2005 both represented over 17.5% of the area.

While the percent increase in terms of area is relatively small (from 12.5% to 17.6%), this represents major changes and reshaping of the landscape. These large changes are related to infrastructure (roads, hotels, entertainment, etc) as well as to the surrounding environment. The areas under “irrigated Agriculture” more than doubled and as explained in section 3.3.5, this also meant a large increase in revenue. In the 1970s, a large portion of the wine was sold in bulk, not associated to the “atmosphere of the Hunter”, as a place of “lifestyle consumption” (O’Neill 2000) as it was from the 1990s. The value of the Hunter was closely interlinked with its proximity to Sydney, being a destination for national and foreign visitors.

The large increase in “irrigated Agriculture” from the mid-1990s is intriguing, since the terrior of the Wine Country is not ideal for wine making. Halliday (2005) stated that “the Lower Hunter surely has to be one of the most capricious and vexatious wine regions in the
world”. In fact, Australian law allows grapes to be brought in from other regions to produce consistent wines (O’Neill 2000).

Evidence of this can be seen in Figure 6.1, where the tonnage of grape purchased from other regions has been almost equal to that grown in the area. While the number of crushed tonnes of grapes has steadily declined from 1999 until 2008, the number of visitors to the region has increased steadily (Figure 6.2). This corroborates the hypothesis that the development of the Hunter is intrinsically linked to the tourism industry.

![Figure 6.1. Tonnage of grapes grown and imported to the Hunter (Australian Wine and Brandy Corporation 2008)](image1)

![Figure 6.2. Annual number of visitors to the Hunter region (Tourism New South Wales 2009)](image2)
Due to increase in tourism, the population of Cessnock and Singleton have increased quasi-linearly from the mid 1970s and are forecast to follow the trend (Figure 6.3), demanding more infrastructure and services. It is important to note that the forecast of population growth of these two statistical local areas are highly correlated ($r > 0.95$) to the forecast of the state’s population.

![Figure 6.3. Population trends and projections of two major statistical regions in the Hunter region (Australian Bureau of Statistics 2009, Hunter Valley Research Foundation 2009)](image)

This pre-data analysis discourse is important because historical data and predicted information related to population, GDP and climate are used to establish a cause-effect relationship with LULC patterns in the area. As an illustration, these relations are then used to simulate LULC in 2020.

Utilising spatial data for predicting LULC change can also be done. This is useful if the interest is to answer specific research questions as posed in Table 2.4. The first question that can be answered with spatial data, specifically a digital elevation model (DEM) and a detailed soil map is to predict the most suitable location for new wineries or to assess whether some existing wineries might disappear due to limited soil conditions. Some of this work has already been done by James Taylor in his PhD thesis (Taylor 2004).

A second question that can be answered with spatial data is the effect of climate change on wine grape production, for this it is necessary to have a DEM, soil maps, as well as historical,
current, and predicted spatial climatic data. The increase in average temperature and occurrence of extreme climatic events can impact wine production in the Hunter Valley and some of this work has been published Webb (2006), Jones et al. (2005), Hood et al. (2006).

Thirdly, it can be argued that the value of the production requires wineries to be located closer to roads, as they need transportation to the consumption market and to ports, for this it is be necessary to have road networks at different points in time and the projected roads for scenario establishment. In the above examples, socio-economic spatial data can also be used to add a dimension in the complexity of winery allocation.

However it was decided to use only a limited number of variables to modify the transition matrices. This decision was based on a few facts which will be summarised below. Firstly, there is little novelty in allocating future wineries to the “optimal” location; secondly, it can be argued that the main product of the Hunter valley is not the wine itself, but the experience it provides through the beautiful scenery, restaurants, country-side atmosphere and proximity to Sydney.

The assumption then becomes that the tourism industry is the main factor for the prosperity of the area. However, tourism statistics only exist for the past decade (Cessnock City Council 2002, Hunter Valley Wine Country 2008, Tourism New South Wales 2009). The roads are needed, but not necessarily so that the wines can be exported in bulk, by truck, but that the visitors take a piece of the experience with them. The study area only spans two census districts, which is not sufficient to further elaborate on these differences.

As for the climate variables, it is out of the scope of the present work to extend the research. First, because as stated by a few authors (Jones et al. 2005, Webb 2006), the area is already on the boundary climate for grape production and it is not known for the quality (Halliday 2005) nor the productivity of its grapes, with a large proportion yearly being imported in to the region (O’Neill 2000). Furthermore, while out of the scope of the present
work, downscaling climatic models to the regional scale, due to the limited size of the study area, will not yield results that could change the course of implementation of new areas.

Finally, the thesis was also intended to be a manual in which inexperienced readers could understand and replicate the whole process of LULC map creation; address map quality issues through other methods such as image fusion; create/understand and experiment with a relatively simple land change model; and lastly to understand how, in literature, that the transition matrices are created/modified and the results of doing so, which is the main topic of the present chapter.

6.2.2 Model calibration and simulation

While Markov-chain models rely on information contained in the transition matrix to accurately reproduce the amount of change, cellular automata models utilise neighbourhood functions to determine change. In the present work, the amount and location of change was simulated using the newly developed, hybrid model.

Calibration of the transition matrix was carried out using a number of methods, varying from simple, static trends to more complex multinominal logistic regression models, all of which were built in R. The calibration runs were seeded by the initial conditions of 2000 and LULC data from 1972 to 2000 for modelling changes to 2005, ensuring that no data post the calibration period were used to simulate 2005, as suggested by Pontius Jr and Malanson (2005). The reference data is the LULC map of 2005 produced and validated in Chapter 3. The work here, therefore, compared the simulated LULC maps for 2005, created by different transition matrices, to the 2005 reference map, which enabled the assessment of model performance. Subsequently, the 2005 map was incorporated into the models, which were then used to simulate the 2020 landscape.

In Chapter 5, it was demonstrated that the plain Markov-chain model utilising transition probabilities extracted from the data was quite successful in reproducing the amounts of change, but did not perform so well in predicting the location of these changes. In contrast,
the model with a higher weight given to the cellular automata component was better at predicting location of change, even though it was limited in reproducing the amount of change. This behaviour could be controlled by the user when modifying \( W_{TM} \) as it was attempted and assessed here.

Since the hybrid model is flexible in terms of weighting values, \( W_{TM} \), it was run with three pre-determined values of \( W_{TM} \) (0, 0.5 and 1). For each possible value of the three \( W_{TM} \), a single run and multiple runs (20 iterations) were carried out. In the case of the multiple runs, only the most frequent outcome of 20 iterations were used further. For all of the above-mentioned cases, the cellular automata employed eight nearest neighbours.

Model performance was assessed by comparing simulated maps to the 2005 reference map. To achieve this overall accuracy, Kappa and its decomposition to location and histogram Kappa (Pontius Jr and Schneider 2001, Veldkamp and Lambin 2001) were computed. Additionally, selected spatial metrics were calculated such as edge density, Simpson’s diversity index and fractal dimension as explained in McGarigal et al. (2002). These metrics were used by Jenerette and Wu (2001), Herold et al. (2005) and Hagen-Zanker and Martens (2008) for comparing outcomes of land change models.

<table>
<thead>
<tr>
<th>Name</th>
<th>Method for deriving trans. matrix</th>
<th>Num of runs</th>
<th>( W_{TM} )</th>
<th>Calibration IC</th>
<th>Transition probability</th>
<th>Simulation IC</th>
<th>Transition probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static</td>
<td>Static trend</td>
<td>1, 20</td>
<td>0, 0.5, 1</td>
<td>2000</td>
<td>1995-2000</td>
<td>2005</td>
<td>2000-2005</td>
</tr>
<tr>
<td>A2TP</td>
<td>Average</td>
<td>1, 20</td>
<td>0, 0.5, 1</td>
<td>2000</td>
<td>85-95 &amp; 95-00</td>
<td>2005</td>
<td>95-00 &amp; 00-05</td>
</tr>
<tr>
<td>A3TP</td>
<td>Average</td>
<td>1, 20</td>
<td>0, 0.5, 1</td>
<td>2000</td>
<td>72-85 &amp; 85-95 &amp; 95-00</td>
<td>2005</td>
<td>85-95 &amp; 95-00 &amp; 00-05</td>
</tr>
<tr>
<td>Empirical</td>
<td>Empirical</td>
<td>1, 20</td>
<td>0, 0.5, 1</td>
<td>2000</td>
<td>Empirical</td>
<td>2005</td>
<td>Empirical</td>
</tr>
</tbody>
</table>

*Multinomial logistic regression (MNLR)*

<table>
<thead>
<tr>
<th>Name</th>
<th>Method for deriving trans. matrix</th>
<th>Num of runs</th>
<th>( W_{TM} )</th>
<th>Calibration IC</th>
<th>Transition probability</th>
<th>Simulation IC</th>
<th>Transition probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>AllClim</td>
<td>MNLR climate</td>
<td>1, 20</td>
<td>0, 0.5, 1</td>
<td>2000</td>
<td>MNLR 1972-2000</td>
<td>2005</td>
<td>MNLR 1972-2005</td>
</tr>
<tr>
<td>GDP</td>
<td>MNLR GDP</td>
<td>1, 20</td>
<td>0, 0.5, 1</td>
<td>2000</td>
<td>MNLR 1972-2000</td>
<td>2005</td>
<td>MNLR 1972-2005</td>
</tr>
<tr>
<td>Pop</td>
<td>MNLR population</td>
<td>1, 20</td>
<td>0, 0.5, 1</td>
<td>2000</td>
<td>MNLR 1972-2000</td>
<td>2005</td>
<td>MNLR 1972-2005</td>
</tr>
</tbody>
</table>
As stated earlier, the validation of the 2020 (predicted) LULC map was not possible, therefore the LULC categories were compared in terms of their class proportions and spatial distribution. Table 6.2 above, presents a summary of the experimental setup in terms of: the name given to different setups, the method for determining the transition matrix, the number of runs, the weighting and finally the model calibration and validation conditions.

6.2.2.1 Static trend

The static trend (Static) experiment only requires the transition probabilities to be extracted from two different LULC maps. These trends are then used for prediction. Nevertheless, a few authors have successfully used trends extracted from two points in time to predict future LULC maps (Houet and Hubert-Moy 2006, Huang and Cai 2007, Xiongwei 2008). One advantage of the Static approach is its simplicity. However, this is also its major limitation, since it assumes that the transition probabilities do not change over time. This contradicts the findings of many land change studies (Lambin et al. 2000, Jenerette and Wu 2001, Turner et al. 2001), which reported that land change is generally dynamic, even though it is often treated as static (Schaldach and Priess 2008).

Furthermore, static transition probabilities imply that a state of equilibrium can be reached, which is incompatible with the findings in land change literature. As shown by Urban and Wallin (2002), a stochastic projection of static trends into the future can be achieved by multiplying the proportions of LULC classes by the static transition probabilities using matrix multiplication. In this study however, the neighbourhood configuration (when $W_{TM}$ was 0 or 0.5) was used to dynamically change the transition probability, as an improvement to the simple static trend proposition.

The static trend calibration was done by annualising the 1995-2000 transition matrix, as described in Chapter 5 and by subsequently using the resulting transition probability matrix to predict the LULC in 2005. Furthermore, simulation involving the annualised transition matrix of 2000-2005 was used to forecast the LULC in 2020 and to test the performance of the

6.2.2.2 Averaging

The reasoning behind averaging two or more transition matrices is that it would incorporate more information to the process. It is a simple scalar mathematical operation, where transition matrices for different periods in time are averaged. Its limitation is related that it is also static and eventually assumes a state of “equilibrium”. This procedure transformed known probabilities of change between time steps into simple average transition values between independent periods.

Baca (2002) utilised transition matrices of 1972-1984 and 1984-1996 to derive an average matrix for prediction of future LULC for an area in Rio de Janeiro state in Brazil. This approach could have originated from Bierzychudek (1999), who used averaging for projecting population of specific plant species. This approach has rarely been used in the literature.


6.2.2.3 Empirical

Empirical transition probabilities are obtained based on expert knowledge of the nature of the transitions (Heldens 2006, Huang and Cai 2007), which makes it possible to use computer programming to generate matrices which preserved such specific knowledge (Lambin et al. 2000, Walton and Poore 2000). They are more appropriate for scenarios required to be tested from pre-conceived assumptions (Lambin 1997, Turner et al. 2001, Verburg et al. 2006,
Huang and Cai (2007). This approach is therefore centred on the framing of the transition matrix using the empirical knowledge of the process generating the change (Heldens 2006). For example, water body is a LULC category which is very unlikely to change. When a set of transition probabilities extracted from such LULC data do not corroborate this situation, then this empirical knowledge of the transition can be incorporated into the matrix.

In the work reported in this chapter, an empirical transition matrix was built from the knowledge gained during field visits (Tinkler et al. 2008), the observed trends in the multi-temporal LULC data generated in Chapter 3 and from the literature (McManus et al. 2000, O'Neill 2000, Beer et al. 2003, McManus 2008). Therefore the transition matrix was based on the following assumptions: a) “woodland” was mostly preserved, with a small chance of their conversion to “dryland Agriculture”; b) “dryland Agriculture” had moderate to high probabilities of changing to a variety of other LULC categories. The latter assumption was based on the fact that the study area was experiencing expansion of “irrigated Agriculture”, in addition to “dryland Agriculture” transiting to “woodland”. Moreover the socio-economic factors were more amiable to these transitions, as both irrigated vineyards and aesthetic landscape were contributing to rapid growth in tourism. Also the rapid population growth in the study area is a major driver of “dryland Agriculture” transiting to “intensive uses” such as housing and recreational development; c) “irrigated Agriculture”, which consisted mostly of wine grapes, were expanding at the expense of “dryland Agriculture”; and d) “intensive uses” and “water bodies” were stable. The transition probabilities generated from these assumptions are presented in Table 6.3; they were used for both calibration and simulation.

<table>
<thead>
<tr>
<th></th>
<th>Woodland</th>
<th>Dryland Agriculture</th>
<th>Irrigated Agriculture</th>
<th>Intensive uses</th>
<th>Water bodies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodland</td>
<td>0.98</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dryland Agriculture</td>
<td>0.035</td>
<td>0.903</td>
<td>0.035</td>
<td>0.025</td>
<td>0.002</td>
</tr>
<tr>
<td>Irrigated Agriculture</td>
<td>0</td>
<td>0</td>
<td>0.98</td>
<td>0.018</td>
<td>0.002</td>
</tr>
<tr>
<td>Intensive uses</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Water bodies</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
6.2.2.4 Incorporation of drivers of land change into MNLR models

A statistical approach for relating land change to ancillary variables is MNLR. Whereas regular logistic regression only considers two categories, MNLR can handle multiple categories. Considering \( j \) possible categories, MNLR predicts the log odds of \( j-1 \) categories compared to a baseline category \((j)\) (Agresti 2002) as:

\[
\text{Pr}(y_i=k) = \frac{e^{x_i \beta_k}}{1 + \sum_{k=1}^{j-1} e^{x_i \beta_k}}
\]

where \( y_i \) is the observed outcome, \( k \) is 1 to \( j-1 \) categories and \( x_i \) is a vector of explanatory variables. The unknown \( \beta_k \) are estimated by maximum likelihood. \( \text{Pr}(y_i=k) \) can be calculated as:

\[
\text{Pr}(y_i=1) = \frac{1}{1 + \sum_{k=1}^{j-1} e^{x_i \beta_k}}
\]

for the baseline category \((j), \beta = 0\), therefore \( e^{x_i \beta} = 1 \), and

\[
\text{Pr}(y_i=k) = \frac{e^{x_i \beta_k}}{1 + \sum_{k=1}^{j-1} e^{x_i \beta_k}}
\]

MNLR has been used by a number of authors (Mertens and Lambin 2000, Geoghegan et al. 2001, Schneider and Pontius Jr 2001) to model changes in LULC. A major limitation of this, as with other regression models, is that explanatory variables should not be correlated.

In this study, MNLR was utilised to create transition matrices based on current and projected changes to ancillary variables. The MNLR models were built using the R programming language (R development core team 2010) and the multinomial function was fitted using the package ‘nnet’ (Ripley 2009). Subsequently, these models were used to estimate the transition probabilities and were then annualised using the package ‘panel’ (Gentleman 2006) (details in section 5.2.1 of Chapter 5). The ultimate aim was to use the transition matrices for model calibration and simulation. Thus the modelling process established a cause-effect relationship between transitions and explanatory variables, which
are capable of generating legacy, current and future LULC scenarios (Brown et al. 2002, Verburg et al. 2006).

Three ancillary variables were used in this study: a) population growth, which has been shown to correlate with urbanisation and economic development (Jenerette and Wu 2001); b) annual per capita GDP, as an indicator of economic growth (Alcamo et al. 2005); and c) climatic projections, as a known paramount for successful wine production (Jones et al. 2005).

While the objective of this chapter was to forecast LULC for 2020 for the Wine Country of the Hunter valley, it was by no means exhaustive, as other variables could contribute to the model.

6.2.2.4.1 Population

It is well known that population is closely related to development and urbanisation (Verburg and Chen 2000, Jenerette and Wu 2001, Agarwal et al. 2002). In the case of the study area used here, by the 1970s, “intensive uses”, including urban, transport and recreational, occupied approximately 1.7% of the landscape. By 2005, they represented over 4.0% of the land area (Table 6.1). Much of this increase was due to the expansion of Cessnock and to growth in tourism.

Evidently, the overall growth of NSW state population (Barson et al. 2001) had also contributed indirectly to the economic growth of the study area through increased tourism, leading to a steady increase in the number of visitors (Figure 6.2) and thus demanding more infrastructure and services. The evidence of this can be seen in the steady increase in population in the two statistical local districts- Cessnock and Singleton- that constitute the study area (Figure 6.3) (Hunter Valley Research Foundation 2009). The population trends of the two districts were highly correlated with the trend of NSW population (r > 0.95). Subsequently, the projected population of NSW (series B) (Pink 2009) was used as an ancillary modifier of the transition matrix for the prediction of LULC in 2020.
6.2.2.4.2 GDP

Alcamo et al. (2005) pointed out that GDP is often used as a measure of economic development. Recently it was used as an ancillary variable in land change modelling of continental Europe by Verburg et al. (2008). In the case of Australia, one of the issues is the declining effect of GDP for rural Australia (Pittock 2003), as the farm sector’s contribution to the national GDP has decreased (26.1% of GDP in 1950 compared to 4.0% in the 2000s). Nevertheless, GDP is still an important determinant or surrogate of overall economic growth of a given region or nation.

A key assumption of using GDP as an ancillary variable in this study, was more related to economic significance of the study area as a region of lifestyle consumption (O’Neill 2000), than to the agricultural value of the economic production system. This assumption was corroborated by McManus (2008), who stated that the success of wine-making in the Hunter was related to its proximity to Sydney and Newcastle and not necessarily on the suitability of the area for wine-making. The substantial increase in number of tourists from the 1990s onwards, facilitated by ease of access, has made the region a major touristic destination. This increased in domestic tourism has been accelerated by the rapid growth of Australian GDP.

The GDP data (Productivity Comission 2005a, Productivity Comission 2005b) was used in this study to develop a multinomial logit function. Then, the projected values of GDP were used to modify the transition matrix to predict LULC in 2020.

6.2.2.4.3 Climate (temperature & rain)

Previous studies have demonstrated that a changing climate would affect grape production and have a major impact on grapes, hence on wine quality (Schultz 2000, Keller 2010). The Hunter Valley is located in a region already considered hot for wine grape production and the projected increase in temperature could hamper production (Jones et al. 2005). Relating wine grape production to projected climates generated by OzClim (CSIRO 2005), Webb (2006)
demonstrated that increased temperatures (0.6 to 2.6°C in 2050) would lead to declining the window of optimum harvest time and hence would compromise grape quality. Webb’s work, involving three regional climatic models combined with three different scenarios (B1, A1B and A1F1) (IPCC 2001), found that grape quality would decrease partially due to increased temperature and partially due to increase in disease pressure. The latter was predicted to affect overall decline in returns of up to 30% due to 50% price drop per tonne of Cabernet.

By zooming to the Hunter region, the scenarios by Webb (2006) showed that grape production in the region could be weather-limited, as the region was already in a boundary climate (Jones et al. 2005). The author (Webb 2006) identified the most significant climatic factors affecting grape production, which were: average annual temperature, average annual rainfall and average temperature of the harvest months. All these factors were considered here to build the MNLR model. The national annual climatic data were obtained courtesy of the AGO (Kesteven et al. 2004). They were sub-sampled at every pixel within the study area. The dataset was then used in the MNLR model and projected changes by Webb (2006) were used to generate a modified transition matrix.

### 6.3 Results and discussion

#### 6.3.1 Model calibration

Calibration involved performance assessment of each model based on accuracy statistics and spatial metrics (Hagen-Zanker and Martens 2008). Here the maps produced by each of the approaches detailed in Section 6.2.2 are compared to the reference map of 2005 to assess the accuracy of the model. Such comparison is made by direct map comparison and not by sampling of points in each map.

One of the most popular accuracy statistic used here is Kappa (Cohen 1960), which could be decomposed into location (Figure 6.5) and histogram (Figure 6.4) components, but the
overall accuracy (Figure 6.6) results are also shown. Location Kappa indicates the spatial component of agreement and histogram kappa quantifies the similarity between different maps (Hagen 2002). It is generally accepted that Kappa values above 0.8 indicate almost perfect agreement between maps (Viera and Garrett 2005).

![Experimental transition matrix and histogram Kappa](image)

**Figure 6.4. Histogram Kappa, as function of experimental transition matrix, $W_{TM}$ and number of runs**

The histogram Kappa showed that all models performed similarly and well when $W_{TM}=0$ (Figure 6.4), independent of the number of runs. However, AllClim (as defined in Table 6.2) had the lowest values despite changing $W_{TM}$. When $W_{TM}$ was adjusted to 0.5 or 1.0, histogram Kappa decreased for 20 runs in comparison to a single run. Such results were expected and have been discussed previously (Chapter 5 and the present). When considering
$W_{TM}=1$ and the Static, A2TP and A3TP it can be seen that histogram kappa increases for single runs when compared to $W_{TM}=0$. Averaging (A2TP and A3TP) smoothes the transition matrix and makes it more similar to Static, these three indicated that there was little change thus justifying the high accuracies. Literature suggests that Markov-chain models usually perform well when reproducing the amount of change, but are limited in reproducing the spatial structure of change (Turner 1987).

![Figure 6.5. Location Kappa, as function of experimental transition matrix, $W_{TM}$ and number of runs](image)

In Figure 6.5, it is demonstrated that both single and 20 runs with $W_{TM}=0$ result in location Kappa’s around 0.7. Increasing the weight of the Markov component to $W_{TM}=0.5$ and $W_{TM}=1$ decreased the location Kappa to about 0.5 when the model is run only once, but did
not affect the results of 20 runs. Thus, the poor performance of single runs and $W_{TM}=1$, can be overcome by multiple iterations of the models. It is also interesting to note that the static transition matrix, which produced consistent histogram Kappa independent of both $W_{TM}$ and the number of runs (Figure 6.4), produced the lowest location Kappa with $W_{TM}=1$ and a single run. When considering single runs of a model with $W_{TM}=1$, it will allow for much noise to be imprinted in the final map as the spatial structure is lost, with the accuracies being higher when $W_{TM}=0$ due to the fact that the cellular automata will change neighbouring cells.

Figure 6.6. Overall accuracy, as function of experimental transition matrix, $W_{TM}$ and number of runs

The overall accuracy statistics (Figure 6.6) corroborates that the best accuracies are associated with small $W_{TM}$ or increased number of runs when $W_{TM}$ is higher. The performance
of all models was similar when $W_{TM}=0$, as in Figure 6.4 and Figure 6.5. Since the time span for the calibration is only five years, the amount of projected change is limited and these are spatially coherent, thus models with low values of $W_{TM}$ perform best.

In addition to the accuracy measures, landscape metrics, such as: fractal dimension, edge density and Simpson’s diversity index, as suggested by Jenerette and Wu (2001), Hagen-Zanker and Martens (2008), were calculated for each of the model’s output maps and the reference map (from Chapter 3), to assess model performance.

The edge density is a measure that determines the amount of edges per unit area. Thus, in a map with little spatial structure, there are a considerably higher number of edges than in a map where the spatial configuration is preserved. Figure 6.7 illustrates the deviation from the reference edge density, indicating that $W_{TM}=0$ or multiple runs with larger $W_{TMS}$ have low deviation from the reference value. Once again this is due to the fact that the cellular automata ($W_{TM}=0$) performs well at maintaining spatial structure, still preserving the edge density. On the other hand, the single runs of the models run with $W_{TM}=1$ had such a high deviation because the spatial structure was lost. A balanced, hybrid model (i.e. $W_{TM}=0.5$), combining the cellular automata and Markov-chain components has lower deviations than the purely Markov-chain model ($W_{TM}=1$), but much higher than those of when the models are run in cellular automata mode ($W_{TM}=0$).
Fractal dimension determines the complexity of shapes, however here the difference from the reference fractal dimension is computed. The reference was calculated for the 2005 map and fractal dimension of the map was calculated for each model and $W_{TM}$. As the maps with preserved spatial structure would have fractal dimensions similar to the reference, then the deviation would be close to zero, while maps with a lack of spatial structure would have higher values due to increased shape complexity. Figure 6.8 illustrates the deviation of each models fractal dimension to that of the 2005 reference. The results of Figure 6.8 and Simpson’s diversity index (not shown), are similar to those of edge density. In summary, the
deviations from the reference fractal dimension and Simpson’s diversity index increased with the increase of $W_{TM}$ and when a single run of the model was computed.

![Figure 6.8. Deviation from reference fractal dimension, as function of experimental transition matrix, $W_{TM}$ and number of runs](image)

In terms of calibration of the models, the three accuracy metrics shown here (Figure 6.4, Figure 6.5, Figure 6.6) and the two landscape metrics shown here (Figure 6.7, Figure 6.8) seem to indicate similar results. Generally speaking, the best performance of all models was achieved when $W_{TM}=0$ or when the modal maps of 20 runs was computed, irrespective of $W_{TM}$.

As for the performance of specific models, their performances were quite similar when considering both the accuracy and landscape metrics. Since the initial conditions were the
same for all models, the variability in outcome would be associated to changes in the transition matrix and the time for which the models ran.

Recalling Table 6.2, it could be noted that for “A2TP”, “A3TP” and Static, the transition matrices for these approaches would be very similar and should yield similar results. The empirical transition matrix was derived from knowledge of the area and was similar to Static, thus the results should also be similar.

The same could not be said about the matrices generated from MNLR. These were a result of relating specific variables (“AllClim”, “GDP” and “Pop”) to LULC changes and generating modified transition matrices. The limitation of this approach however, was associated to the data. GDP and population data were represented by single values for the whole region. Historical climatic variables were available in space/time, however the projections extracted from Webb (2006) also represented a single value for the whole region. The cause-effect relationship was summarised in a single transition matrix, which might not be an appropriate solution. Future research would aim at developing locally adaptable transition matrices. Since MNLR also considered the historical data, the dynamics of the above mentioned variables resulted in little changes to the transition matrix, thus yielding similar results.

Furthermore, in the calibration phase the models were only run for five years. It has been discussed above that the transition matrices for the different models were quite similar, thus the differences in these matrices would be elucidated better if the model were run for a longer period of time, as in the simulation results below.

6.3.2 Simulation of 2020 landscape

Figure 6.9 below, shows some interesting results of model simulations of 2020 LULC. All model simulations with $W_{TM}=0$, produced very similar LULC patterns except for the models “AllClim” and “GDP”. These two models indicating that “dryland Agriculture” would lose area to “woodland”. Furthermore, increasing the value of $W_{TM}$ accentuated this difference between two models- “AllClim” and “GDP”- compared with all the other models. The two
models also produced outcomes indicating areas of “intensive uses” were seriously compromised. In support of this, the most frequent outcome of 20 runs of the “GDP” model with $W_{TM}=1$ produced LULC patterns composed mostly of “woodland” (>95%). It is unrealistic that built infrastructure areas would be converted back to a semi-natural landscape in such a short period.

Simulations using the “Empirical” model, on the other hand, produced outcomes with large increase in “irrigated Agriculture” and “intensive use” at the expense of “dryland Agriculture”.

Other studies also use the methods described here for simulating LULC patterns (Turner 1987, Soares-Filho et al. 2001, Baca 2002, Heldens 2006, Zhang et al. 2007, Cabral and Zamyatin 2009), but rarely on a comparative manner as shown in Figure 6.9. Given the same initial conditions, a LULC change model and transition matrices derived from different approaches, the outcomes can be of little change (“A2TP”, “A3TP”, ”Pop” and “Static”), some change (“Empirical”) and huge (though unrealistic) change (“AllClim” and “GDP”). This indicates that the variability of the results is given by the selection not only of the $W_{TM}$ parameter in the model, but also on the method for determining transition matrices, for which there is much research but no consensus. It would seem reasonable the transition matrices derived from more complex models which accounted for spatial variables would perform best, but this would only be the case if these spatial variables were the drivers of change. In a scenario where the spatial changes are very limited a model that predicts no change at all would perform very well, which seems to be the case with “Static”, “A2TP” and “A3TP”.

185
Given that the evolution of the landscape composition from 1972 to 2005 presented in Table 6.1, indicating “intensive uses” and “irrigated Agriculture” had expanded considerably over time, it would be interesting to know the proportions of these two LULC categories, in 2020, based on various model predictions. With $W_{TM}=0$, the models’ outputs indicated little variation (Figure 6.10 below). Given that in 2005 “intensive uses” represented about 4.1% of the area, all models except “Empirical” showed a slight decrease in this LULC category by 2020. Interestingly though, the outcome of models with increased $W_{TM}$ showed very little difference in the proportions of “intensive use” among the models: “A2TP”, “A3TP” and “Static”. In contrast, while “AllClim” and “GDP” produced outcomes that showed “intensive
uses” would disappear, the “Empirical” model produced an increase of “intensive uses”, regardless of the number of runs.

![Experimental transition matrix](image)

**Figure 6.10.** Percentage (%) of study area covered by “intensive uses” in 2020, predicted by different models, as a function of $W_{TM}$ and number of runs

Once again the results derived from transition matrices with little changes (“A2TP”, “A3TP” and “Static”) presented results which were very similar to the initial maps, promoting little change while the “empirical” model signalled a large increase in the area of “intensive uses”, while “Pop” and “GDP” virtually suppressed its cover. These changes became more accentuated with the use of a high $W_{TM}$, when doing so the estimated proportions of the classes are more accurate; however the map lacked spatial structure.
Figure 6.11 illustrates the effect of number of runs and $W_{TM}$ on the final composition of “irrigated Agriculture” predicted by different models in 2020. In 2005 “irrigated Agriculture” comprised only 13.52% of the landscape (Table 6.1). For 2020, considering $W_{TM}=0$, all models but “AllClim” and “Empirical” indicated a slight decrease in the areas of “irrigated Agriculture”. However in 2020, the aforementioned two models produced increase in the proportion of “irrigated Agriculture”. This increase was accentuated with $W_{TM}=0.5$, while the other five models indicated further decrease of “irrigated Agriculture”.

As $W_{TM}$ was increased to 1, there was more discrepancy of the results of the different combinations of the single and multiple runs. Focusing on the single runs, all models except
“GDP” and “Pop” produced outcomes with increased “irrigated Agriculture”, ranging from 14.3% (A3TP) to about 22.5% (Empirical). In contrast, the multiple runs produced different outcomes; all models except “AllClim” and “Empirical” produced lower proportion of “irrigated Agriculture” area in comparison to 2005. In particular the “GDP” model predicted “irrigated Agriculture” area of less than 1% of the landscape and the “Pop” model predicted a proportion of about 6% in 2020.

As discussed in Chapter 5, the single runs of the models with $W_{TM}>0$ resulted in loss of the spatial structure. In contrast, running the models exclusively in CA mode ($W_{TM}=0$), increased spatial coherence; however this compromised the accuracy, in terms of histogram Kappa. It was indicated that this shortcoming could be partially overcome by using the most frequent outcome of a number of iterations, regardless of the weight given to $W_{TM}$.

Visual comparison of the outcomes of different models is useful in highlighting some of the practical outcomes of the simulation results. Figure 6.12 a) shows the reference 2005 LULC map. Figure 6.12 b) shows the most frequent of 20 runs of the MNLR model from the climatic prediction, with $W_{TM}=0.0$. It suggests an increase in the areas of “woodland”, mostly at the expense of “dryland Agriculture”, without any large change in the proportion of “irrigated Agriculture” or “intensive uses”. Figure 6.12 c) which is the product of most frequent outcome of 20 runs, with $W_{TM}=0.5$ and utilising the empirical transition matrix, shows some increase in “irrigated Agriculture” areas. However, the 20 runs were not sufficient to remove the graining patterns present in patches of “dryland Agriculture”. Figure 6.12 d) shows the outcome of the most frequent outcome of 20 runs of MNLR “Pop” model with $W_{TM}=1.0$. This simulation indicated decrease of “irrigated Agriculture” area although the 20 runs were not sufficient to eliminate the graining effect. Thus, increasing the value of the Markov-chain component effectively compromised the spatial structure of the simulations.
6.3.3 Discussion

In comparing the methods of determining transition matrices, there was little difference between using the averages (“A2TP”, “A3TP”) and “Static”. Moreover, the averages and
static simulated little change in landscape patterns even when the model was run for 15 years. This could be seen in Figure 6.9, Figure 6.10 and Figure 6.11, where average and static transition matrices did not forecast much change in relation to the proportions of the 2005 reference (Table 6.1).

Using static trends extracted from recent LULC maps or from averaging of different transition matrices have been used in many studies to forecast LULC patterns (Baca 2002, Houet and Hubert-Moy 2006, Xiongwei 2008), but it is faced with limitations. The first issue related to these trends, is due to the assumption that future changes will follow the same patterns as those of the most recent transition matrix or average of the most recent ones. The second issue with this approach is that if there is some error associated to the LULC map from which the transition matrix is extracted, the matrix itself will be flawed (Pontius Jr and Li 2010).

It has been widely discussed in literature that future changes do not necessarily follow the recent past trends (Engelen et al. 2007, Koomen and Stillwell 2007). The recent past trends (static) are inadequate, especially so when the intention of the study is to evaluate scenarios (Castella and Verburg 2007), hence the need for more elaborate approaches, such as ones derived from empirical knowledge and MNLR.

The transition matrix built from empirical knowledge predicted large changes to “irrigated Agriculture”, “intensive uses” and consequently to “dryland Agriculture”, especially with $W_{TM} > 0$. This was expected, as assumptions were made, which would increase the areas of “irrigated Agriculture” and “intensive uses”. This sort of transition matrix was suggested by Heldens (2006) and Huang and Cai (2007) and while it did forecast reasonable proportions of LULC, there was no formal driving factor to support it, thus the results could be questioned.

To increase the objectivity of deriving the transition matrices, MNLR associated to climate (AlClim), GDP (GDP) and population (Pop) individually, was used. Simulation results of the different MNLR approaches differed. With $W_{TM} = 0$, the “Pop” produced results
similar to “Static”, “Empirical” and average models (A2TP, A3TP). The other two MNLR transition matrices, namely “AllClim” and “GDP” indicated an increase in “woodland” at the expense of “dryland Agriculture”. When increasing $W_{TM}$ though, the latter two models indicated a large increase in “woodland”, with it representing over 99% of the landscape for “GDP”, with $W_{TM}=1$ and the modal outcomes of 20 runs.

Results shown by the MNLR derived transition matrices would then raise the doubt of whether they represented any improvement over using the static, recent past trends, derived directly from the LULC maps. A method for comparing model fitting is the Akaike information criterion (AIC), as described in Hengl (2009) and Hastie et al. (2008). The AIC penalises for including variables in a model and gives a final score to the model, where lower scores indicate better fitting. The AIC was computed for the three MNLR models and for the past trends (Static and Average), finding that MNLR performed better in all three instances.

Another limitation of the MNLR approach undertaken here was that the prediction by GDP, population and the climatic variables were aspatial. Furthermore, as the explanatory variables were highly correlated, a model considering two or all three of them would be inappropriate, due to multicollinearity. While fitting the “Pop” and “GDP” models, it was thought that they would yield similar results, as both were represented by a single number for the whole area (for model fitting and prediction), however the final results were very different (Figure 6.9).

The outcomes of “Pop” were more conservative, similar to those of “A2TP”, “A3TP” and “Static”. This could be due to population representing a linear increase, thus replicating the recent trends. In the case of models based on GDP, two different issues were identified. The first is that the increase in Australian GDP may be invariant with LULC change in the Hunter. The point made previously (6.2.2.4.2) by Pittock (2003) is reiterated here; that the importance of the rural sector GDP had diminished in Australia. The initial and naïve assumption was that the increase of GDP would indicate an increase of wealth, which in its turn, would reflect in
the “lifestyle consumption” in the Hunter. The second point is that the GDP projections incorporate complexities in the domain of economics (such as adjusting values), which are not relevant to this work.

The model built from climatic variables (AllClim), also had its limitations. A key one, among others, is that while the model was fitted with spatial data, the predictions were made using aspatial predictions for the whole area. Webb (2006) provided estimated average changes to each climatic variable for the whole Hunter valley. While it would be valuable to have spatial predictions, even with a downscaled climatic model for the region, the resolution, as currently forecast, is too coarse for regional-scale study such as done here. In summary, the importance of building representative transition matrices is emphasised, as to produce scenarios of future change, as suggested by Verburg et al. (2006).

In regards to the predictive accuracy of the methods for deriving the transition matrices, the calibration results showed similar results for all seven approaches, while the simulation results showed a different picture. The simulation results showed either, little change (for 4 models) or unrealistic predictions for 2020 (for 2 models), while ‘Empirical’ showed more consistent results.

The poor predictive accuracy shown in the calibration has also been demonstrated in Pontius Jr et al. (2008), when comparing a number of land change models. In that study, the authors found that none of the models performed better than a null model (no change), if the validation map only used data from the calibration period. In that sense, persistence (no change), would yield high accuracies, but did not represent reality. The different methods for parameterizing the transition matrix, while limited in their predictive power, have to their advantage their simplicity.

Breiman (2001), defends that complex problems usually require complex solutions, better solved by algorithmic modelling such as: artificial neural networks, random forests, support vector machines, amongst other approaches. The algorithmic modelling would have higher
predictive power, but understanding the inner working of these models would be a “Herculean task”, as stated by the author.

Thus, once more, the main strength of what has been developed relies on its simplicity and flexibility. The novelty of the work brought forward in Chapter 5 and this one, is the development of a hybrid, Markov-chain and CA land change models, where the code is open-source and the models is amendable to change and addition of complexity.

This hybrid model has a number of strengths:

1. It is relatively simple, capable of forecasting LULC change from limited inputs, namely: an initial map and a transition probability matrix;
2. The user has the option of how to parameterise the transition matrix and the work shown here, illustrates how this has been done by other authors;
3. It allows the user to select the number of neighbours used in the cellular automata component, although in the current implementation only allows the choice of four or eight; future enhancement of the algorithm will allow for more flexible number of neighbours;
4. The weight given to the cellular automata and Markov-chain components is set by the user, allowing for model testing;
5. The standard output of the model includes a GeoTIFF, which can easily be imported to any GIS software, as well as a text file containing a pixel by pixel frequency distribution table, which can also be used for further model testing;
6. The model also outputs a GeoTIFF for each year of the simulation, thus allowing the user to visually compare the landscape’s evolution;
7. Since the model has a cellular automata component, the outcomes in terms of quantity of each LULC type cannot be pre-determined. Different spatial configurations with the same proportions of each LULC type will yield different results; and
8. Last but not least, the model is built using open-source software (R) and associated packages. The source code can be found in Appendix C.

Since the source code is open and will be made available through the public domain, the model can be modified and complexity added to it as the user wishes. Some examples of complexity that can be added to the model include (but are not limited to): masking out areas of no change; using other layers to influence the transition probabilities; creating layers of suitability for specific uses which may influence change and others.

While the hybrid model had a number of strengths, it also had some limitations. First, since the model is capable of being run from limited data, its outcome is only reasonable if the initial condition and transition probability matrix are both representative. Second, while the CA component is effective in increasing the spatial coherence of the model, here, it only considers a limited neighbourhood, so influence such as proximity to an urban centre, for instance, needs to be incorporated.

Third, the current implementation of the model is based on a single transition probability matrix for both components (Markov and CA). The neighbourhood configuration consequently becomes the main factor in dynamically changing transition probabilities. In a subsequent implementation of the model, it is intended to disaggregate the transition probability matrices, allocating one for the Markov-chain component, with a second one specific to the neighbourhood.

Finally, the model is based on raster data. Urban land models often use blocks of land as the minimum mapping unit. In a rural setting though, land ownership seldom follows a pre-determined structure, as it does in the city. Since it is not possible, in this scenario, to use pre-determined blocks of land, changes were modelled on a pixel by pixel basis.

Further to what has been shown here, a spatial MNLR associated to self-modifying transition matrix through a feedback loop could be integrated into the model, but has not yet been attempted. The idea would be that at every iteration of the loop, a transition matrix were
built and the spatial MNLR model were used with ancillary variables to create a modified transition matrix, which could be used in the following iteration. This approach would allow for the dynamics of the socio-economic variable to be dynamically linked to the process of change; however it would also imply the need for updating of the spatial socio-economic variable at every time step.

Here, it was shown how projecting future trends of change have been undertaken in the literature (static, average, empirical and MNLR), but these changes have not been linked to the monetary value of each LULC class. Thus, it would be sensible to value the predominant LULC classes and to use this knowledge combined with an econometric model to understand the inter-relations between the developments of the tourism industry in the region. This would require a valuation of the wine industry as a producer of agricultural product, but would also have to determine the impact of tourism in the industry. This is left for future work.

6.4 Conclusions

This chapter demonstrated the effect of utilising a variety of transition probability matrices to calibrate our hybrid model of LULC change. It also showcases the effect of using these matrices to forecast LULC patterns in 2020. The selection of methods for calibration and simulation varied from simple, static trends, to more complex ones derived from MNLR. This chapter illustrated the effect of running the models with different settings, including seven sets of transition probabilities; one or twenty runs and with $W_{TM}$ of 0.0, 0.5 and 1.0.

- The calibration runs showed similar results for most transition probability models, but with large effects due to changing the weight of the Markov component ($W_{TM}$) and the number of runs.
The simulation runs also showed that transition probability models from static and average trends performed similarly, while MNLR had variable results. “Pop” showed results similar to “Static”, “A2TP” and “A3TP”.

“GDP” and “AllClim” simulated that most of the landscape would be covered by “woodland”.

Empirical transition probabilities indicated an increase of “irrigated Agriculture” and “intensive uses”, mostly at the expense of “dryland Agriculture”.

Here it was shown how the parameterization of the transition matrix is commonly used, noting that it is important to objectively select a method for doing so.

Better fitting and prediction would be achieved using spatial MNLR, however the ancillary variables at such spatial-resolution (50 m) are not available.

Strengths and limitations of the model were also discussed, but an emphasis is put on its simplicity and flexibility. As the code for the hybrid model is published in Appendix C, users can modify the model to include other functions and outputs.

6.5 References


Gentleman, R., 2006. Package "panel". 1.0.6 ed.: R foundation for statistical computing, Panel: functions and datasets for fitting models to panel data.


Parker, D.C., Berger, T. & Manson, S.M., 2002. Agent-based models of land-use and land-cover change: report and review of an international workshop. Irvine, California, USA.


Chapter 7. Summary, conclusions and future research
7.1 Summary and conclusions

The importance of land monitoring has been highlighted throughout the thesis. Monitoring is a result of land use and land cover (LULC) mapping associated to land change modelling. This thesis aimed at deconstructing the complexity of LULC classification and land change modelling. This was done by breaking down the components into chapters, which are summarised below.

The thesis began with Chapter 1, where the main topics of each chapter were presented. In Chapter 2 a review of the literature concerning LULC was undertaken, highlighting the importance of what one could define as the most biophysically diverse and dynamic components of the Earth system, the land and its change. The review also explored the issues of LULC classification and land change modelling for the monitoring of the land surface. It identified some limitations of various LULC classification algorithms and concluded, as in many studies, that there was no single best method for LULC classification, nor was there a global solution to land change modelling. The advantages of an object-based based approach were discussed, where it was seen that segmentation was responsible for delineating representative image objects which in turn could be classified. Furthermore, a choice of classification algorithm was greatly dependant on data availability, previous knowledge and access (or not) to the study area. Considering the lower Hunter valley, the chosen study area for this thesis, the availability of medium spatial resolution multi-temporal satellite images and the access to landowners in the area, it was pointed out that an object-based approach coupled to ruleset based and supervised classification methods would be appropriate.

In terms of land change modelling, a number of methods/approaches were discussed, along with their strengths and limitations. While some of the land change modelling methods used economic theory, others were focused mainly on stochastic simulation, multi-agents or statistical analysis. It was, however, noted that a more parsimonious approach would be
appropriate, specifically a flexible approach that would allow variable complexity as the need may be.

In Chapter 3, historical LULC maps of the lower Hunter valley study area were created using an object-based approach combined with rulesets and supervised nearest neighbour classification. The classification was done using historical Landsat images as corroborated by the interpretation of near-time aerial photo (API), landowners’ information and expert knowledge. The classification process was broken into: i) a fully automated approach; ii) enhancements using API and expert knowledge; and iii) field validation. The fully automated maps had overall classification accuracy in the order of 70-75%. This low level of accuracy was due to the problem of mixed pixels and coarse spatial resolution of the images.

Improvements to the fully automated approach were achieved by API and expert knowledge whereby the aerial photos were overlain on the LULC maps as a post-classification correction strategy. There was an increase in overall accuracy to nearly 90% and other accuracy metrics such as the Kappa statistic, but the time spent on elaborating the LULC maps would not be practical if the intention were to map large areas. Field validation and consultation with landowners allowed for the LULC maps to be further corrected, which in turn served as input for a land change model, both as the initial conditions for the calibration and for extracting transition matrices from multiple time steps.

In Chapter 4 the focus was on elucidating and unravelling LULC classification. It applied the principles of image fusion to enhance the LULC maps of two sub-sections of the study area. Through an automated object based approach, three levels of fusion were explored by experimentation, whereby Landsat TM images and orthophotos were fused at the pixel-, feature- and decision-levels to generate LULC maps using a supervised nearest neighbour algorithm. Furthermore, for comparison purposes, the LULC maps were also created from the non-fused images. The results indicate that fusion at the feature- and decision-levels produced the most accurate LULC maps, based on Kappa statistics, overall accuracy and level of
disagreement due to quantity and allocation, compared to those maps produced with pixel-level fusion and non-fusion methods. This was attributed to the fact that feature-level fusion used more spectral information for segmentation and classification than decision-level fusion, which used the higher spatial-resolution data for segmentation and incorporated the Landsat image for the classification.

The overall accuracy of feature- and decision-level fusion was above 83%, allowing for rapid mapping of larger areas. The process was fully automated, requiring only appropriate samples of training data for the supervised classification algorithm. However, since the orthophotos did not cover the whole study area and were not available for five different time-periods for land change modelling, the procedures could not be used for deriving the transition matrices for land change modelling. In order to fully validate the results of fusion experiments, though, it was necessary to test the procedures in another area and comparing with other algorithms of pixel-level fusion.

In Chapter 5, the focus was to develop two land change models. The first model was based on a first-order Markov-chain which relied exclusively on the transition matrix to determine change. The problem with this approach though, was that the neighbouring cells were not influential in the transition process. Thus, when running the first-order Markov-chain model, the spatial structure of the landscape was lost. To solve this problem, a second, hybrid model was created using a combination of a weighted Markov-chain ($W_{\text{TM}}$) and cellular automata (CA) ($W_{\text{NB}}$) to drive change. The transition of pixels, in this case, was dependent on the weight given to the Markov-chain component and the neighbourhood configuration. Therefore the transition matrices under the hybrid model were dynamic.

Both models were developed in the R programming language. A sensitivity analysis of both models was also undertaken. The results of the hybrid with a Markov-chain weight of 1.0 were similar to those of the pure Markov-chain model, as expected. In terms of varying the weighting components in the hybrid model, the results showed that a higher weight in the
Markov-chain component performed better at preserving the proportions of LULC classes, while a higher weight given to the CA component preserve the spatial structure of the simulation.

The hybrid model has a number of strengths:

- it can be used to predict LULC change from limited inputs;
- the amount of change is not pre-determined as in pure Markov-chain models;
- the weight of the neighbourhood, by varying $W_{NB}$, is set by the user;
- it allows four or eight neighbours to influence transition;
- the standard output of the hybrid model includes GeoTIFF files, which can be used in any GIS software;
- the output also includes a frequency distribution table for each pixel, as the model can be run multiple times;
- the hybrid model can and shall be used as a learning tool as the code is published here; and
- complexity can be added by the user, if needed.

Since the drivers of change in the hybrid model were associated with the transition matrix, Chapter 6 illustrated various ways of its parameterization. The methods explored ranged from simple, static transition trends, which were extracted from two LULC maps (from Chapter 3), to more complex approaches, including averages of multiple transition matrices, an empirical transition matrix and ones relying on multinomial logistic regression (MNLR) with socio-economic variables. The models were run with various settings: with $W_{TM}=0.0$, 0.5 and 1.0; single and most frequent outcome of 20 runs were also computed.

Model calibration runs on the different transition matrices were used to predict LULC in 2005, using data up until 2000. The results indicated little variability in regards to the overall accuracies of the different parameterization methods. This was due to the similarity among the transition matrices and the short period (5 years) for the model run.
Simulation runs, which used data up to 2005 to predict 2020, showed that the transition matrices extracted from static trends and averages produced similar results. The same could not be said about the models run with the empirical transition matrices and those derived from MNLR. Some of the outcomes were flawed as they predicted a back-transformation mainly to “woodland”. This was attributable to the fact that the socio-economic variables used were not spatial and therefore further enhancements to the parameterization of the transition matrices were suggested. There was no obvious conclusion reached from the experimentation with different methods for deriving transition matrices, as the outcomes showed little change for some models and unrealistic change for other models.

Further work in parameterizing the transition matrices should be done, including approaches with spatial predictors of land change. Moreover, the hybrid model should be coupled to an econometric model with economic value ascribed to different land uses. One other feature that could be added would be a feedback mechanism to self-modify the transition matrix at every time-step, relating the changes in LULC to socio-economic variables through MNLR or other statistical methods.

In the scientific community there are scholars who defended the Markov-chain and cellular automata approaches to land change modelling. There are others who object these approaches. From this study it can be said that the hybrid approach developed here is a relatively simple solution, to which complexity can be added. Moreover, this hybrid approach would serve as an exceptional learning tool, as the processes are transparent.

A discussion of the significance of the thesis is appropriate here, even though these have been mentioned in the respective research chapters.

The importance of LULC monitoring has been discussed widely throughout the thesis. The workflow approach, presented in Chapter 3, associated to the Landsat images did not produce maps that were directly suitable for LULC monitoring and land change modelling. However, if the workflow approach were combined with the principles of image fusion it
presented a very suitable candidate for automating the elaboration of current and historical LULC maps for large areas. The novelty of Chapters 3 and 4 was related to formalising the automated object-based workflow approach, testing and assessing the different levels of fusion and suggesting benchmarking LULC maps produced from fusion against the non-fused maps.

The work carried forward in Chapters 3 and 4 would facilitate the efforts of monitoring and land change modelling, as the change could be characterised (“when and where”) at relatively short time intervals. Furthermore, having a better understanding of LULC patterns over time would enable the understanding of the causes of change (“why”), however this could only be done using a land change model.

The hybrid land change model developed here would be a suitable candidate for understanding the causes of change, but also for forecasting future LULC patterns. In its current state the model is relatively simple, only needing a set of initial conditions, characterised by a LULC map and two transition matrices.

If the research question was one of determining the frequency of the different LULC classes, then the model should be used in Markov-chain mode ($W_{TM}$ closer to 1.0), whereas if the research question were more related to the spatial distribution of the LULC classes, then the model should have a higher weight given to the CA component ($W_{TM}$ closer to 0.0).

Further enhancements to the hybrid model could include suitability layers and feedback-loops to improve model output. It was also shown that commonly used methods for deriving the transition matrices were not necessarily representative of future LULC patterns, as they often indicated too little change or large areas transforming back to natural land covers, which was highly unlikely.

Regardless, the hybrid model is a novel approach to combining CA and Markov-chains, as compared to what has been done previously. It is further emphasised that it is an
exploratory tool, which can and should be modified, not only for LULC mapping, but for other spatio-temporal problems.

Finally, as the objective of a PhD is research training, in this work much knowledge and insight was gained. As mentioned previously, my background was in Agriculture, thus learning land classification and change modelling has been very rewarding. A few topics which have become part of my repertoire, include, but are not limited to: object-based classification, supervised classification algorithms, image fusion techniques, R, Markov-chains, transition matrices, cellular automata, accuracy assessment through different metrics, as well as programming the different models and methods.

7.2 Future research

As stated in the Preface, research is a life-long task. From the chapters of this thesis, while much has been achieved, the knowledge gained also helped in identifying areas where I could further elaborate on specific topics. These are summarised below:

- further enhancements to the object-based approach, using ancillary data for classification could aid in increasing the accuracy of the historical LULC maps. This could be achieved by expanding the rulesets;

- furthermore, the image fusion techniques proved to produce accurate LULC maps. This approach should be further explored in different areas, as it enabled an automated approach for historical monitoring of LULC patterns;

- in regards to the hybrid model, various opportunities exist. The first enhancement to the model would allow it to work with more than eight neighbours;

- second, it would be beneficial to include in the model, exclusion zones (where change could not occur) and hotspots (where LULC changes were more likely to occur). This could be done through the use of specific suitability layers;
- third, it would be beneficial to develop a statistically sound approach for deriving the neighbourhood transition matrix. A few ideas have been thought for that, but none were yet implemented; and

- in regards to the transition matrices themselves, a feedback loop could be used where the MNLR model fitting could occur at every time-step, thus generating a new matrix that would dynamically change with the change occurred in the previous time-step and change in the ancillary variable.

The topics mentioned above are likely to keep me busy for a while and that is the next step.
Appendices
Appendix A – Annualising transition matrices

The code shown here is useful for annualising transition matrices in R. It was used with R 2.10 and requires the package panel.

# Created by Marcelo Stabile on 20/08/2009
# Thanks to Michael Nelson for help with coding
# Using transition matrices for n years, this program annualises these matrices
# using eigenvalues and eigenvectors as described by Heldens (2006),
# Whitelegg and Oliver (1975), Takada et al. (2010)

# Load library
library(panel)

#### Insert transition matrix below
trans_XX_YY = rbind(
c(0, 0, 0, 0, 0),
c(0, 0, 0, 0, 0),
c(0, 0, 0, 0, 0),
c(0, 0, 0, 0, 0),
c(0, 0, 0, 0, 0))

# Look at transition matrix
trans_XX_YY

# Calculate eigenvectors and eigenvalues
eigen_XX_YY = eddcmp(trans_XX_YY)

# Look at eigenvectors and eigenvalues
eigen_XX_YY

# Calculate transition matrix for 1 year
eigen_XX_YY_1yr = eigen_XX_YY$evectors %*% ((diag(eigen_XX_YY$evalues))^(1/(YY-XX))) %*% solve(eigen_XX_YY$evectors)

# Inspect
eigen_XX_YY_1yr

setwd("E:/TEMP")

# Export text file containing the annualised transition matrix
write.table(eigen_XX_YY_1yr, file="50m_XX_YY_ann_TM.txt", row.names=FALSE, col.names=TRUE, sep=".", )
Appendix B – Model A

The code shown here represents Model A as in Chapter 5. It was used with R 2.10 and requires the packages maptools and rgdal

# This program reads a GeoTIFF LULC map and uses defined transition probabilities and random numbers to change the pixel values, It reiterates for however many years the user wants and outputs GeoTIFF files

#Cleaning Up
rm(list=ls())
gc()
setwd("E:/TEMP")

# Loading libraries
library(maptools)
library(rgdal)

# Setting Conditions
NumOfYear = 10  # How many years do we want it to run for
NumRuns = 500  # How many Runs do you want (for bootstrapping)
Sel_IC <- "95"  # What is the Year of the Initial Condition LULC map
Sel_TP <- "95-00"  # What Transition Probabilities are we using?
fileName <- as.character(paste("IC",Sel_IC,"_TP",Sel_TP,"_Runs",NumRuns,"_Years",NumOfYear,sep=""))

#Read GeoTIFF data in
InitCond = readGDAL("E:/TEMP/LULC_MAP.tif")

# Inspecting InitCond
# Plotting InitCond
#spplot(InitCond)

# Year 0 is = to the selected Initial Conditions
Year0 <- InitCond

# Transition probabilities from...
# Insert transition probability matrix
TM_5x5 = rbind(
  c(1,0,0,0,0),
  c(0,1,0,0,0),
  c(0,0,1,0,0),
  c(0,0,0,1,0),
  c(0,0,0,0,1))
as.array(TM_5x5)

#### Beginning of loop and functions
# Get Start Time
IniTime= Sys.time()
IniTime

# Creating function where:
# Random number is generated
# Looks at individual pixel value, and along with Transition rule, evaluates change or no change
EvalTrans = function (pixij,TM_5x5)
{
  RandArrij = runif (1,0,1)
  if (pixij==0) pixij=0 else
  (if (pixij==100)
    (if(RandArrij<=TM_5x5[1,1]) pixij=100 else
    (if(RandArrij<=(sum(TM_5x5[1,1:2]))) pixij=320 else
    (if(RandArrij<=(sum(TM_5x5[1,1:3]))) pixij=444 else

213
(if(RandArrij<=(sum(TM_5x5[1,1:4]))) pixij=500 else 
  if(RandArrij<=(sum(TM_5x5[1,1:5]))) pixij=600 )))  else 
  if (pixij==320) 
    if(RandArrij<=TM_5x5[2,1]) pixij=100 else 
      if(RandArrij<=(sum(TM_5x5[2,1:2]))) pixij=320 else 
        if(RandArrij<=(sum(TM_5x5[2,1:3]))) pixij=444 else 
          if(RandArrij<=(sum(TM_5x5[2,1:4]))) pixij=500 else 
            if(RandArrij<=(sum(TM_5x5[2,1:5]))) pixij=600 ))))  else 
  if (pixij==444) 
    if(RandArrij<=TM_5x5[3,1]) pixij=100 else 
      if(RandArrij<=(sum(TM_5x5[3,1:2]))) pixij=320 else 
        if(RandArrij<=(sum(TM_5x5[3,1:3]))) pixij=444 else 
          if(RandArrij<=(sum(TM_5x5[3,1:4]))) pixij=500 else 
            if(RandArrij<=(sum(TM_5x5[3,1:5]))) pixij=600 ))))  else 
  if (pixij==500) 
    if(RandArrij<=TM_5x5[4,1]) pixij=100 else 
      if(RandArrij<=(sum(TM_5x5[4,1:2]))) pixij=320 else 
        if(RandArrij<=(sum(TM_5x5[4,1:3]))) pixij=444 else 
          if(RandArrij<=(sum(TM_5x5[4,1:4]))) pixij=500 else 
            if(RandArrij<=(sum(TM_5x5[4,1:5]))) pixij=600 ))))  else 
  if (pixij==600) 
    if(RandArrij<=TM_5x5[5,1]) pixij=100 else 
      if(RandArrij<=(sum(TM_5x5[5,1:2]))) pixij=320 else 
        if(RandArrij<=(sum(TM_5x5[5,1:3]))) pixij=444 else 
          if(RandArrij<=(sum(TM_5x5[5,1:4]))) pixij=500 else 
            if(RandArrij<=(sum(TM_5x5[5,1:5]))) pixij=600 ))))

} 

# Function to evaluate pixel value
PixelEval <- function (pix, evalu) 
  { 
    ret <- 0 
    if (pix == evalu) ret <- 1 
    return (ret) 
  } 

# Initialises Year 0 to be IC then iterates for however many years and for however many runs
runs<- file(paste(fileName,".txt",sep=""),’wt’)  # Creates TXT file, easier to read
for (j in 1:NumRuns) 
  { 
    Year0 <- InitCond 
    # Have to reinit Year0, as the function modifies it
    for (i in 1:NumOfYear) 
      { 
        Year0@data <- data.frame(unlist(apply(as.matrix(Year0@data),1,EvalTrans,TM_5x5=TM_5x5))) 
        write(as.vector(Year0@data[,1]),runs,ncolumns=1) 
      } 
    close(runs)
  } 

# Comparing InitCond and output
#spplot(InitCond)
#spplot(Year0)
# Read text file
runs.read <- file(paste(fileName,".txt",sep=""),’rt’)
# Creates matrix of the same size as Year0
dist.matrix <- matrix(0,nrow=length(Year0@data[,1]),ncol=6)
#Populates the matrix with the occurrence of each LULC
for (i in 1:NumRuns)
temp.100 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=100))
temp.320 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=320))
temp.444 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=444))
temp.500 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=500))
temp.600 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=600))
temp.000 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=0))
dist.matrix[,1] <- dist.matrix[,1]+(temp.100)
dist.matrix[,2] <- dist.matrix[,2]+(temp.320)
dist.matrix[,3] <- dist.matrix[,3]+(temp.444)
dist.matrix[,5] <- dist.matrix[,5]+(temp.600)
dist.matrix[,6] <- dist.matrix[,6]+(temp.000)
}
close(runs.read)

# Function to evaluate pixel value and convert map to LULC
# 1=100, 2=320, 3=444, 4=500, 5=600, 6=0
Replacing <- function (pix)
{
  if (pix == 1) ret <-100
  if (pix == 2) ret <-320
  if (pix == 3) ret <-444
  if (pix == 4) ret <-500
  if (pix == 5) ret <-600
  if (pix == 6) ret <-0
  return (ret)
}

# Get end time and Calculates and outputs the time it took to run
EndTime=Sys.time()
RunTime=EndTime-IniTime

# 500 runs for 10 years took 10.38 hours on Core 2 CPU 6300
# 50 runs for 10 years took about 66 minutes Core 2 CPU 6300
# 1 Run for 10 years took 1.34 min Core 2 CPU 6300
# 1 Run For 1 year took 26 seconds Core 2 CPU 6300

# Creates dataset similar to InitCond, populates pixel values with the most common occurrence of # LULC class
MaxProbMap <- InitCond
GetMax <- as.numeric(max.col(dist.matrix))
GetMax_C <- apply(as.matrix(GetMax),1,Replacing)
MaxProbMap@data[,1] <- as.numeric(GetMax_C)

# Inspecting outcomes
# spplot(MaxProbMap)
# summary(MaxProbMap)

# Creates TXT File with distribution of pixel values
write(t(dist.matrix), file= paste("DistMat_","fileName",".txt",sep=""),ncolumns=6)

# Creates GeoTiff, with most probable occurrence
writeGDAL(MaxProbMap, paste(fileName,".tif",sep=""),drivername="GTiff", type = "UInt16")

#Outputs GeoTiffs for every year
#for(i in 1:NumOfYear)
# { 
#  yearly <- byYear
#  yearly@data <- data.frame(byYear@data[,i+1])
# writeGDAL(yearly,paste("E:/TEMP/IC","Sel_IC","_TP","Sel_TP","_year","i",".tif"),drivername = "GTiff", type = "UInt16")
# }
Appendix C – Model B

The code shown here represents Model B as in Chapter 5. It was used with R 2.10 and requires the packages maptools and rgdal

# This program reads a GeoTIFF LULC map, it then converts to a matrix and pads edges with 0's.
# It gives weights to the static Trans. Prob and uses neighbourhood functions to dynamically allocate
# a transition probability according to neighbours (works with 4 and 8), all neighbours have same weight.
# Weighted Trans prob + Weight Proportional Neighb Trans probs are compared to a random number.
# Pixel transitions are then re-evaluated. The number of years and runs can be specified by the user.
# in the end there is a function to output the most frequent map as a GeoTIFF file
# This file also produces a time series of images for each TP, but only for 1 run.
# Modified By Marcelo Stabile 20/05/2010

#Cleaning up & loading libraries
rm(list=ls())
gc()
setwd("C:/TEMP")
library(rgdal)
library(spdep)

# Defining variables
Sel_TP <- "test"  # What Transition Probabilities (TP) are we using?
TP_weight.vec <- c(0,0.5,1)  # What TP_weights you want to test?
NumRuns.vec <- c(1,20)  # Defining how many runs you want (loop)
NumNeighb <- 8  # Can be 4 or 8
Sel_IC <- "00"  # What is the Year of the Initial Condition LULC map
IniYear <- XXXX
IniCond = readGDAL("C:/TEMP/XXXX.tif")

# Reading transition probabilities from file according to TP name
TM_5x5 <- as.matrix(read.csv(paste("TM_","Sel_TP","_ann.txt", sep=""), header=F))

# Numeric classes of LULC
dimnames(TM_5x5) <- list(c("100","320","444","500","600"))

# To convert to SpatialPixelsDataFrame
fullgrid(IniCond) <- FALSE
# Checking IniCond, frequency and plot
summary(IniCond)
spplot(IniCond)

# Creating a matrix with 0's around it (so that we can extract index of neighbours)
pad_mat <- matrix(0, nrow= (IniCond@grid@cells.dim[2]+2), ncol= (IniCond@grid@cells.dim[1]+2))
coordinx.pad <- pad_mat[2:(IniCond@grid@cells.dim[2]+1),2:(IniCond@grid@cells.dim[1]+1)] <- matrix(IniCond@coords[,1],nrow= IniCond@grid@cells.dim[2], ncol= IniCond@grid@cells.dim[1], byrow= TRUE)
coordiny.pad <- pad_mat[2:(IniCond@grid@cells.dim[2]+1),2:(IniCond@grid@cells.dim[1]+1)] <- matrix(IniCond@coords[,2],nrow= IniCond@grid@cells.dim[2], ncol= IniCond@grid@cells.dim[1], byrow= TRUE)

# assigning x coordinates
coordinx.pad[,1] <- coordinx.pad[,2] <- IniCond@grid@cells.size[1]
coordinx.pad[,2](IniCond@grid@cells.dim[1]+2] <- coordinx.pad[,1](IniCond@grid@cells.size[1]+1)] + IniCond@grid@cells.size[1]
coordinx.pad[1,] <- coordinx.pad[,2]
coordinx.pad[1,1] <- coordinx.pad[2,2] <- IniCond@grid@cells.dim[2]+2.

216
# Assigning y coordinates
coordy.pad[,1] <- coordy.pad[2,] + InitCond@grid@cellsize[2]
coordy.pad[(InitCond@grid@cells.dim[2]+2),] <- coordy.pad[(InitCond@grid@cells.dim[2]+1),] -
InitCond@grid@cellsize[2]
coordy.pad[,1] <- coordy.pad[2]
coordy.pad[(InitCond@grid@cells.dim[1]+2),] <- coordy.pad[(InitCond@grid@cells.dim[1]+1),]
Data.pad <- pad_mat Data.pad[2:(InitCond@grid@cells.dim[2]+1),2:(InitCond@grid@cells.dim[1]+1)] <-
matrix(InitCond@data[,1], nrow= InitCond@grid@cells.dim[2], ncol= InitCond@grid@cells.dim[1], byrow= TRUE)
# Making the matrix a single line
x.col <- as.numeric(t(coordx.pad)) # The X coordinates
y.col <- as.numeric(t(coordy.pad)) # The Y coordinates
coords.mat <- cbind(x.col, y.col) # Combining coordinates
rm(x.col, y.col, coordy.pad, coordx.pad) # Cleaning up, removing unused bits
data.vec.pad <- as.numeric(t(Data.pad)) # The data

############################################################# END PHASE II #############################################################

############################################################# START PHASE III #############################################################
IniTime_NN <- Sys.time() # Initialising timer of NN
PixNN_ID <- knearneigh(coords.mat, k=NumNeighb) # Getting index of nearest neighbours
EndTime_NN <- Sys.time() # End timer of NN
Time_NN <- EndTime_NN - IniTime_NN # Computing time difference
write(t(PixNN_ID$nn), file = "NN.txt", ncolumns = 8) # Creating text file of Neighbour ID’s
Year0 <- data.vec.pad # Year 0 is the selected Initial Conditions

############################################################# END PHASE III #############################################################

############################################################# START PHASE IV #############################################################
# Function that returns the pixel values for the corresponding neighbourhood index
NN_PixVal <- function (PixNN_ID_nni, Year_nni)
{
  return(Year_nni[PixNN_ID_nni])
}

# Function to compute frequencies. For every row it puts the frequency of each LULC class
Year_Freq <- function (Year_NN_nni)
{
  tab_nni <- table(Year_NN_nni)
  out.tab <- rep(0,6)
  names(out.tab) <- c("100","320","444","500","600","0")
  out.tab[names(tab_nni)] <- tab_nni
  return (out.tab[1:5])
}

# Function to extract correct row of TM according to pixel value
Extract_TM <- function(pixij, TM_5x5)
{
  if (pixij==0) out.tm <- rep(0,5)
  else
    {
      out.tm <- TM_5x5[as.character(pixij),]
    }
  return(out.tm)
}

# Function where Random number is generated and compared to TM
# Looks at individual pixel value and along with Transition rule evaluates change or no change
EvalTrans_NN = function (YearDataRow)
{
  RandNum <- runif (1,0,1)
pixij <- YearDataRow[1]
  TM <- YearDataRow[2:6]
  if (pixij==0) pixij <- 0 else

if(RandNum<=TM[1]) pixij=100 else if(RandNum<=(sum(TM[1:2]))) pixij=320 else if(RandNum<=(sum(TM[1:3]))) pixij=444 else if(RandNum<=(sum(TM[1:4]))) pixij=500 else if(RandNum<=(sum(TM[1:5]))) pixij=600 )
}

return(pixij)

# Function to evaluate pixel value
PixelEval <- function (pix, evalu)
{
  ret <- 0
  if (pix == evalu) ret <-1
  return (ret)
}

# Function to replace pixel value
Replacing <- function (pix)
{
  if (pix == 1) ret <-100
  if (pix == 2) ret <-320
  if (pix == 3) ret <-444
  if (pix == 4) ret <-500
  if (pix == 5) ret <-600
  if (pix == 6) ret <-0
  return (ret)
}

# Creating a loop to use the vector of NUMBER OF YEARS
for (w in 1:length(NumOfYear.vec))
{
  NumOfYear <- NumOfYear.vec[w] # Defining Number of Years
  # Creating a loop to use the vector of TP_WEIGHTS
  for (z in 1:length(TP_weight.vec))
  {
    TP_weight <- TP_weight.vec[z] # Defining TP WEIGHT
    NB_weight <- 1-TP_weight # NB WEIGHT is 1 - TP_WEIGHT
    # Creating a loop to use the vector of NUMBER OF RUNS
    for (y in 1:length(NumRuns.vec))
    {
      NumRuns <- NumRuns.vec[y] # Defining Num RUNS
      # Establishing the Standard FILENAME
      fileName <- as.character(paste("IC",Sel_IC,"_TP",Sel_TP,"_Runs",NumRuns,"_Years",NumOfYear,"_TPw",TP_weight,"_NB",NumNeighb, sep=""))

      #### Begining of loop
      IniTime_Loop= Sys.time() # Get Start Time
      IniTime_Loop
      # Initialises Year 0 to be IC then reiterates for however many years and for however many runs
      runs<- file(paste(fileName,".txt",sep=""),"wt") # Creates TXT file, easier to read
      for (j in 1:NumRuns)
      {
        Year0 <- data.vec.pad # Reinit. Year0, as function modifies it
        for (i in 1:NumOfYear)
        {
          # Matrix with NN pixel val
          Year_NN <- t(apply(PixNN_ID$nn,1,NN_PixVal, Year_nni=Year0))
        }
      }
    }
  }
}
# Matrix with prop of LULC
FreqMat <- t(apply(Year_NN,1, Year_Freq))

# Matrix with correct row from original TM according to pixel value
TM_byRow <- t(apply(as.matrix(Year0),1, Extract_TM, TM_5x5=TM_5x5))

### Final TM, weighted by TM and by Neighbourhood TM
# Multiply the neighbourhood prop. by row
NB_TM <- FreqMat*TM_byRow

# Determining the sum of the rows
NB_Sum <- unlist(apply(NB_TM,1,sum))

# Calculating temp Normalised matrix to 1
Norm_NB_TM_T <- NB_TM/NB_Sum

# Replacing NaN's for 0's
Norm_NB_TM <- replace(Norm_NB_TM_T, is.nan(Norm_NB_TM_T)==T,0)

# Gets Final TM
Final_TM <- (TP_weight*TM_byRow) + (NB_weight*Norm_NB_TM)

# Combining Year i data with appropriate TM for evaluation
YearData <- cbind(Year0,Final_TM)

# Evaluating the Change or no change
Year0 <- unlist(apply (YearData,1, EvalTrans_NN))

# Removing unnecessary stuff
remove(Year_NN, FreqMat, TM_byRow, NB_TM, NB_Sum, Norm_NB_TM_T, Norm_NB_TM, Final_TM, YearData)

gc()
#
# Writes TXT file, easier to read
write(as.vector(Year0),runs,ncolumns=1)
#
# Closing file that has final state of each run
close(runs)

# Get end time and Calculates and outputs the time it took to run
EndTime_Loop=Sys.time()
RunTime_Loop=EndTime_Loop-IniTime_Loop

#################### END PHASE V ###############################################

#################### START PHASE VI ###################################################################

# Reads TXT file, easier to read
runs.read <- file(paste(fileName,".txt",sep=""),"rt")

# Creates matrix of the same size as Year0
dist.matrix <- matrix(0,nrow=length(Year0),ncol=6)

#Populates the matrix with the occurrence of each LULC
for (l in 1:NumRuns)
{
  temp <- scan(runs.read, integer(), length(Year0))
  temp.100 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=100))
  temp.320 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=320))
  temp.444 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=444))
  temp.500 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=500))
  temp.600 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=600))
  temp.000 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=0))

dist.matrix[,1] <- dist.matrix[,1]+(temp.100)
  dist.matrix[,2] <- dist.matrix[,2]+(temp.320)
  dist.matrix[,3] <- dist.matrix[,3]+(temp.444)
  dist.matrix[,5] <- dist.matrix[,5]+(temp.600)
  dist.matrix[,6] <- dist.matrix[,6]+(temp.000)
}

close(runs.read)

# Deletes the file with dist
unlink(paste(fileName,".txt",sep=""))
# Creates a DataFrame with maximum value and coordinates, which is converted to SGDF
MaxProbMap <- as.numeric(max.col(dist.matrix))
MaxProbMap_C <- apply(as.matrix(MaxProbMap),1,Replacing)
MaxProbMap_C_XY <- cbind(coords.mat,MaxProbMap_C)
MaxProbMap.df <- data.frame(MaxProbMap_C_XY)
coordinates(MaxProbMap.df) <- ~x+y
gridded(MaxProbMap.df) <- T
fullgrid(MaxProbMap.df) <- T

# Comparing InitCond with output
#par(mfrow = c(1,2))
#image(InitCond)
#image(MaxProbMap.df)

########################################
#### START PHASE VII
########################################
# Creates TXT File with distribution of pixel values
#write(t(dist.matrix), file= paste("DistMat_",fileName,".txt",sep=""), ncolumns=6)
# Outputs GeoTiff of the most probable occurrence
writeGDAL(MaxProbMap.df,paste(fileName,".tif",sep=""), drivername = "GTiff", type= "UInt16")

PREPARING FOR YEARLY EVOLUTION

# Variables have been defined
# Functions have been defined
# Evolution only occurs with 1 run
NumRuns <- 1

YEARLY EVOLUTION

# Creating a loop to use the vector of NUMBER OF YEARS
for (w in 1:length(NumOfYear.vec))
{
  NumOfYear <- NumOfYear.vec[w]       # Defining Number of Years
  # Defining TP WEIGHT
  TP_weight <- TP_weight.vec[z]
  NB_weight <- 1-TP_weight
  # Creates Matrix of Num Years + 1
  Year0i<-matrix(nrow=length(data.vec.pad),ncol=(NumOfYear+1))
  # First column of matrix is Initial Condition
  Year0i[,1] <- data.vec.pad
  # Establishing the Standard FILENAME
  fileName <- as.character(paste("IC","_Sel_IC","_TP","_Sel_TP","_Runs","_NumRuns","_Years","_NumOfYear","_TPw","_NB","_NumNeighb", sep=""))

  #### Beginning of loop
  IniTime_Loop=Sys.time() # Get Start Time
  IniTime_Loop
  # Initialises Year 0 to be IC then reiterates for however many years and for however many runs
  # Loops for a number of runs
  runs<- file(paste(fileName,".txt",sep=""),"wt") # Creates TXT file, easier to read
  for (j in 1:NumRuns)
  {
    Year0i[,1] <- data.vec.pad # Have to reinit Year0, as the function modifies it
  }
for (i in 1:NumOfYear)  # Loops For a Number of Years
{
  # Matrix with NN pixel values
  Year_NN <- t(apply(PixNN_ID$nn,1,NN_PixVal, Year_nni=Year0i[,i]))
  # Matrix with proportions of LULC for each LULC class
  FreqMat <- t(apply(Year_NN,1,Year_Freq))
  # Matrix with correct row from original TM according to pixel value
  TM_byRow <- t(apply(as.matrix(Year0i[,i]),1,Extract_TM, TM_5x5=TM_5x5))
  ## Final TM, weighted by TM and by Neighbourhood TM
  # Mult. the neighbourhood prop. by row
  NB_TM <- FreqMat*TM_byRow
  # Determining the sum of the rows
  NB_Sum <- unlist(apply(NB_TM,1,sum))
  # Calculating temporary Normalised matrix to 1
  Norm_NB_TM_T <- NB_TM/NB_Sum
  # Replacing NaN's for 0's
  Norm_NB_TM <- replace(Norm_NB_TM_T, is.nan(Norm_NB_TM_T)==T,0)
  # Gets Final TM
  Final_TM <- (TP_weight*TM_byRow) + (NB_weight*Norm_NB_TM)
  # Evaluating the Change or no change
  Year0i[,i+1] <- unlist(apply (YearData,1,EvalTrans_NN))
  # Removing unnecessary stuff
  remove(Year_NN, FreqMat, TM_byRow, NB_TM, NB_Sum, Norm_NB_TM_T, Norm_NB_TM, Final_TM, YearData)
  # Clearing up some memory
  gc()
}
# Writes TXT file, easier to read
write(as.vector(Year0i[, (NumOfYear+1)]),runs,ncolumns=1)
# Closing file that has final state of each run
close(runs)

runs.read <- file(paste(fileName,".txt",sep=""), 'rt')  # Reads TXT file, easier to read
# Creates matrix of the same size as Year0
dist.matrix <- matrix(0,nrow= length(Year0i[,1]),ncol=6)
#Populates the matrix with the occurrence of each LULC
for (l in 1:NumRuns)
{
  temp <- scan(runs.read, integer(), length(Year0i[,1]))
  temp.100 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=100))
  temp.320 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=320))
  temp.444 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=444))
  temp.500 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=500))
  temp.600 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=600))
  temp.000 <- unlist(apply(as.matrix(temp),1,PixelEval, evalu=0))
  dist.matrix[,1] <- dist.matrix[,1]+(temp.100)
  dist.matrix[,2] <- dist.matrix[,2]+(temp.320)
  dist.matrix[,3] <- dist.matrix[,3]+(temp.444)
  dist.matrix[,5] <- dist.matrix[,5]+(temp.600)
  dist.matrix[,6] <- dist.matrix[,6]+(temp.000)
}
close(runs.read)
unlink(paste(fileName,".txt",sep=""))  # Deletes the file with dist
# Creates a DataFrame with maximum value and coordinates, which is converted to
# Spatial Grid Data Frame
MaxProbMap <- as.numeric(max.col(dist.matrix))
MaxProbMap_C <- apply(as.matrix(MaxProbMap),1,Replacing)
MaxProbMap_C_XY <- cbind(coords.mat,MaxProbMap_C)
MaxProbMap.df <- data.frame(MaxProbMap_C_XY)
names(MaxProbMap.df) <- c("x","y","maxProb")
coordinates(MaxProbMap.df) <- ~x+y
grided(MaxProbMap.df) <- T
fullgrid(MaxProbMap.df) <- T

# Comparing InitCond with output
#par(mfrow = c(1,2))
#image(InitCond)
#image(MaxProbMap.df)

# Creates TXT File with distribution of pixel values
#write(t(dist.matrix), file= paste("DistMat_".fileName,".txt",sep=""),ncolumns=6)

# Creates an image for each year of the simulation
for (k in 1:length(Year0i[1,])) {
  YearEvol <- cbind(coords.mat,Year0i[,k])
  YearEvol.df <- data.frame(YearEvol)
  names(YearEvol.df) <- c("x","y","LU LC")
  coordinates(YearEvol.df) <- ~x+y
  gridded(YearEvol.df) <- ~x+y
  fullgrid(YearEvol.df) <- T
  writeGDAL(YearEvol.df,paste(fileName,"_Yr",IniYear+k,".tif",sep=""),drivername = "GTiff", type= "UInt16")
}

************************************************************************ END OF YEARLY EVOLUTION **************************************************************************