

**A COMPARISON OF REMOTE SENSING
CHANGE DETECTION METHODS FOR URBAN
CREEP IDENTIFICATION IN NORWICH**

**Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester**

by

Andrew Phillip Tewkesbury BSc

Department of Geography

University of Leicester

2017

A COMPARISON OF REMOTE SENSING CHANGE DETECTION METHODS FOR URBAN CREEP IDENTIFICATION IN NORWICH

By

Andrew Phillip Tewkesbury

Abstract

Change detection is one of the most active research areas in remote sensing, driven by the desire to monitor the highly dynamic world around us. Modern very high resolution (VHR) images from satellite and aerial platforms present an opportunity to reveal landscape change in great detail. However, there is little research focusing on detailed change detection. This research addresses this problem, by investigating the utility of remote sensing change detection to detect fine urban changes called 'urban creep'. Urban creep is the addition of impermeable surface to an existing property, after its initial construction. This is problematic because cumulatively, urban creep significantly increases flood risk. Moreover, up-to-date urban creep statistics are not readily available. Therefore, urban creep identification is a challenge to change detection research because of its subtle expression, small extent, and complicated contextual setting. The investigation of urban creep is conducted over the city of Norwich using aerial images from 2006 and 2010. The research focuses on three methodological areas. Firstly, statistical sampling is employed to quantify a baseline rate of urban creep. Secondly, a direct classification of the remotely sensed data is undertaken to assess the utility of the state-of-the-art in this application. And lastly, as a counter-point, change vector analysis (CVA) is applied to explore the utility and relevance of differencing methods in a complex urban change detection setting. The results have shown that 1) The rate of urban creep in Norwich is increasing; 2) Quantitative urban creep evaluation remains beyond the current state-of-the-art, but qualitative detection is possible; 3) Simple spectral CVA has almost no utility in this application. These findings contribute to knowledge by clarifying the capabilities of change detection techniques in detailed urban applications; building upon, and extending existing CVA research; and finally by adding to urban creep knowledge.

Acknowledgements

Preparing this thesis has been the biggest challenge of my life. In fact, with a part time duration of seven years, the juggling of commitments, the late nights and weekends, it has felt like a life sentence. In completing this thesis I greatly appreciate the help and support that I have received, and take this opportunity to express my gratitude.

Firstly, I would like to thank my supervisors for guiding me along this long and eventful journey, and remaining patient and supportive. Throughout this time I have enjoyed the support and influence of no less than four supervisors at various times. Dr Nicholas Tate remains the only founding member of ‘the band’ to have lasted the distance, and I thank him especially for his attention to detail and commitment to high quality, well written research. For the final two years Professor Kevin Tansey has offered a fresh prospective, keen insight and clear advice which has really helped towards submission. A special mention must go to Professor Alexis Comber for leading the supervision team for the first five years. His enthusiasm, support and advice have proved invaluable for the entire duration. Lastly, and sadly, Professor Peter Fisher passed away during the preparation of this thesis. His guidance and sentiment still live on, as does his legacy, both as an accomplished and celebrated scientist, and as a thoroughly great guy!

I would like to thank my sponsor, Airbus Defence and Space Ltd. for supporting this research with finances, with data, and most importantly personnel support. Dr Anthony Denniss deserves a special mention – you got me in to this mess! But seriously, thank you for the opportunity, and the ongoing support. Thank you to Alistair Lamb for being my change detection sounding board: discussing papers, developing ideas and providing an endless conveyor belt of methods to evaluate. Lastly, thank you to Andrew Anstee for listening to me drone on about this research for seven years.

This thesis would not have been possible without the support of my friends and family. I thank them all for putting up with me during this long task and for faking interest so convincingly. In particular, my wife Caroline for coping with long stints as a ‘PhD widow’ and going along with me when I suggest that a week in Norwich mapping

pavement is an ideal summer holiday. Thank you to my son Thomas for providing the giggles and welcome distractions. My father Martyn, and sadly departed grandfather Denis for installing in me inquisitiveness and a passion and commitment to lifelong learning. Thanks must also go to my mother, brother, and in-laws for your kind support. My fellow PhD students have been a huge help throughout this process. The shared support, advice and discussion have been invaluable.

Several organisations have supported this research in various ways. Thank you to Mark Ogden from Norfolk County Council and Andrew Bird from Anglian Water for providing local insight and helping to refine the study area. Thanks also go to UKWIR for granting access to its publications, the Ordnance Survey of Great Britain for data provision, and Leica Geosystems, Geospatial Solutions Division for granting permission to reproduce the ADS40 illustrations.

Chapter 2 of this thesis is based on an article published in the journal *Remote Sensing of Environment*, and so I thank the editorial team and publisher Elsevier for permitting re-publication. The original article (Tewkesbury et al. 2015) was a collaborative effort between myself, Alexis Comber, Nicholas Tate, Peter Fisher and Alistair Lamb.

However, the co-authors contribution was largely editorial and I can confirm that the scientific content of the article did originate from the lead author, Andrew Tewkesbury.

List of Contents

Abstract	i
Acknowledgements	ii
List of Contents.....	iv
List of Tables	viii
List of Figures.....	ix
List of Equations	xii
List of Acronyms and Abbreviations.....	xiii
List of Publications and Awards.....	xv
List of Appendix Figures and Tables	xvi
1 Introduction.....	1
1.1 Overview	1
1.2 Problem definition	2
1.3 Research questions.....	7
1.4 Research objectives	8
1.5 Thesis structure	8
2 Change detection techniques.....	10
2.1 Introduction	10
2.2 Unit of Analysis	12
2.3 Comparison methods.....	21
2.4 Summary	31
3 Change detection challenges.....	33
3.1 Introduction	33
3.2 Scene illumination	34
3.3 Viewing geometry.....	37

3.4	Analysis scale	40
3.5	Segmentation inconsistency and object comparison	42
3.6	Object-based feature utilisation	43
3.7	Summary	44
4	Urban creep: a social concern and change detection challenge	46
4.1	Introduction	46
4.2	Urban creep as a challenge to change detection research	46
4.3	Why is urban creep important?	49
4.4	Summary	56
5	Study area and reference data collection	57
5.1	Introduction	57
5.2	The study area and monitoring timeframe	58
5.3	Nomenclature	60
5.4	Reference data objectives	61
5.5	Sample design	62
5.6	Sample size	65
5.7	PSU mapping.....	68
5.8	Field checking and final review.....	72
5.9	Population inference	73
5.10	Urban creep rate	75
5.11	Summary	75
6	Reference data collection results	77
6.1	Introduction	77
6.2	Change distribution	77
6.3	Change statistics	80
6.4	Urban creep rate.....	82

6.5	Resolving image interpretation confusion	83
6.6	Ongoing urban creep	87
6.7	Summary	88
7	Remotely sensed change detection	91
7.1	Introduction	91
7.2	Remotely sensed data and study area review	91
7.3	Data quality	97
7.4	Data organisation	99
7.5	Spatial co-registration	101
7.6	Radiometric co-registration	103
7.7	Modelling viewing geometry	107
7.8	Modelling scene illumination	110
7.9	Unit of analysis	113
7.10	Feature extraction	120
7.11	Unsupervised feature selection	124
7.12	Change vector analysis	125
7.13	Training data	130
7.14	Supervised classification	131
7.15	Accuracy assessment	133
7.16	Experiment structure	135
7.17	Summary	138
8	Change detection results	140
8.1	Introduction	140
8.2	Unsupervised feature selection	141
8.3	Classification results	147
8.4	Best case classification	167

8.5	Urban creep and CVA.....	175
8.6	Summary	185
9	Discussion	187
9.1	Introduction	187
9.2	Purpose	187
9.3	Results summary.....	188
9.4	Urban creep estimation	189
9.5	Remotely sensed change detection.....	196
9.6	Summary	209
10	Conclusion	211
10.1	Introduction	211
10.2	Research questions	212
10.3	Contribution to knowledge	213
10.4	Implications	215
10.5	Recommendations for future work	216
	Appendix.....	217
	References	230

List of Tables

Table 1: Analysis units commonly used in remote sensing change detection studies ..	14
Table 2: An overview of commonly used comparison methods	22
Table 3: Image resolution in change detection research	41
Table 4: A summary of urban creep monitoring research	55
Table 5: PSU sample size estimation table	67
Table 6: Remotely sensed reference data	69
Table 7: The reference data interpretation key	71
Table 8: Norwich change estimates, 2006 to 2010	80
Table 9: Norwich urban creep rate, 2006 to 2010	83
Table 10: The reference data results summary table	89
Table 11: Aerial image data layer summary	95
Table 12: 2006 image flight lines	96
Table 13: 2010 image flight lines	97
Table 14: Work block monitoring timeframe	101
Table 15: Data positional error summary	102
Table 16: Summary of the spatial calibration	102
Table 17: Flight line sun positioning	111
Table 18: The CEMRS parameters used to create IOL L2, and L3	118
Table 19: The twenty-five feature themes	122
Table 20: Classification group and classification scenario summary	136
Table 21: Overall map accuracy	150
Table 22: Urban creep classification results summary	154
Table 23: Other change and no change summary	161
Table 24: The error matrix of the most successful classification	168
Table 25: The change detection results summary table	185
Table 26: The subset of the results for discussion	188
Table 27: Urban creep rate comparison	190
Table 28: A brief literature survey of VHR change detection overall accuracy	197
Table 29: A comparison between spectral and feature selected C ² VA	205

List of Figures

Figure 1: Urban creep and urban growth visualised in bi-temporal aerial imagery	4
Figure 2. The four components of a change detection technique.....	12
Figure 3. Commonly used remote sensing change detection analysis units	15
Figure 4: Two dimensional geometry of three formulations of CVA	30
Figure 5: The Norwich study area.....	59
Figure 6: One-stage cluster sampling schematic.....	63
Figure 7: The distribution of the reference samples.....	68
Figure 8: A mapped PSU sample	72
Figure 9: Norwich urban creep and other change distribution.....	79
Figure 10: An example of urban creep, identified in the reference data	81
Figure 11: An example of urban creep, identified in the reference data	82
Figure 12: The visual similarity between some gravels and small block paving	85
Figure 13: Urban creep obscured by a car	86
Figure 14: An example of urban creep subjectivity.....	87
Figure 15: An illustration of evolving urban creep	88
Figure 16: Remote sensing data overview	93
Figure 17: Push broom tri-stereo sensor image capture	94
Figure 18: The aerial image flight lines.....	96
Figure 19: An illustration of the data resolution and quality	99
Figure 20: The extent of the fourteen work blocks.....	100
Figure 21: The implemented PIF identification algorithm	105
Figure 22: Example output from the PIF identification algorithm	105
Figure 23: PIF radiometry before and after the radiometric transformation.....	106
Figure 24: Hidden surface detection algorithm outline	108
Figure 25: Hidden surface identification results	109
Figure 26: The hill-shading illumination modelling	112
Figure 27: An illustration of a hierarchical, multi-scale segmentation	114
Figure 28: The results of the L1 multi-temporal MRS segmentation.....	115
Figure 29: A summary of the CEMRS algorithm	116
Figure 30: A cross comparison of MRS and CEMRS	117

Figure 31: The results of the L2 multi-temporal CEMRS	119
Figure 32: The results of the L3 multi-temporal CEMRS	120
Figure 33: C ² VA geometry.	126
Figure 34: CTA illustration	129
Figure 35: Direct classification workflow	132
Figure 36: CVA classification workflow	133
Figure 37: The error matrix preparation workflow	134
Figure 38: Feature selection by analysis scale	143
Figure 39: Feature selection by image interpretation principle	144
Figure 40: Feature selection by comparable state	144
Figure 41: Difference feature selection by analysis scale	145
Figure 42: Difference feature selection by image interpretation principle	146
Figure 43: Overall map accuracy	148
Figure 44: Norwich change map.....	151
Figure 45: Norwich change map.....	152
Figure 46: Norwich change map.....	153
Figure 47: Urban creep average accuracy comparison	155
Figure 48: Urban creep producer accuracy comparison	158
Figure 49: Urban creep user accuracy comparison.....	159
Figure 50: Change classification example	162
Figure 51: Other change and no change average accuracy comparison	164
Figure 52: Other change and no change producer accuracy comparison	165
Figure 53: Other change and no change user accuracy comparison	166
Figure 54: Urban creep classification example	170
Figure 55: Urban creep classification example	171
Figure 56: Urban creep classification example	172
Figure 57: Urban creep classification example discovered from the classification	173
Figure 58: Classification feature importance and semantic summary.....	174
Figure 59: WC ² VA magnitude histogram comparison by class	176
Figure 60: CTA polar visualisation of the Red and NIR bands	179
Figure 61: CTA polar visualisation of the four spectral bands.....	180
Figure 62: CTA visualisation.....	181

Figure 63: CTA polar visualisation of 39 features.....	182
Figure 64: CTA visualisation.....	183
Figure 65: CTA visualisation.....	184

List of Equations

Equation 1.....65

Equation 2.....65

Equation 3.....66

Equation 4.....73

Equation 5.....74

Equation 6.....74

Equation 7.....74

Equation 8.....74

Equation 9.....75

Equation 10.....75

Equation 11.....75

Equation 12.....75

Equation 13.....103

Equation 14.....103

Equation 15.....106

Equation 16.....106

Equation 17.....126

Equation 18.....126

Equation 19.....127

Equation 20.....127

Equation 21.....127

Equation 22.....127

Equation 23.....128

Equation 24.....128

Equation 25.....128

Equation 26.....130

Equation 27.....130

Equation 28.....135

List of Acronyms and Abbreviations

CDA	Critical Drainage Area
CEMRS	Compactness Enforcement Multi-Resolution Segmentation (defined in Chapter 7)
CTA	Change Trajectory Analysis (defined in Chapter 7)
CV	Change Vector
CVA	Change Vector Analysis
C ² VA	Compressed Change Vector Analysis
DSM	Digital Surface Model
FOV	Field of View
GDP	Gross Domestic Product
GIScience	Geographic Information Science
GLCM	Grey-Level Co-occurrence Matrix
IOL	Image Object Level
L1	Image object level one (large cartographic scale)
L2	Image object level two (moderate cartographic scale)
L3	Image object level three (small cartographic scale)
MRS	Multi-Resolution Segmentation
NIR	Near Infra-Red
NOAA	National Oceanic & Atmospheric Administration
OBIA	Object-Based Image Analysis
PIF	Pseudo Invariant Feature
PSU	Primary Sampling Unit
RMSE	Root Mean Square Error
SAM	Spectral Angle Mapper
SCM	Spectral Correlation Mapper
SSU	Secondary Sample Unit
SUDS	Sustainable Urban Drainage Systems
SWMP	Surface Water Management Plan
T0	Time period zero, equivalent to 2006 in this thesis
T1	Time period one, equivalent to 2010 in this thesis

List of Acronyms and Abbreviations

UK	United Kingdom
UKWIR	United Kingdom Water Industry Research
VHR	Very High Resolution
WC ² VA	Weighted Compressed Change Vector Analysis (defined in Chapter 7)

List of Publications and Awards

- Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A. & Fisher, P.F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, 160, pp.1–14.
- Comber, A., Umezaki, M., Zhou, R., Ding, Y., Li, Y., Fu, H., Jiang, H. & Tewkesbury, A.P., 2012. Using shadows in high-resolution imagery to determine building height. *Remote Sensing Letters*, 3(7), pp.551–556.
- Tewkesbury, A.P., Comber, A.J., Tate, N.J. & Fisher, P.F., 2012. Practical considerations for remote sensing urban creep classifications. In *Proceedings of RSPSoc Annual Conference*. p. 8.
- Tewkesbury, A.P., 2011. Mapping the extent of urban creep in Exeter using OBIA. In *Proceedings of RSPSoc Annual Conference*. p. 163.
- Tewkesbury, A.P. & Hinchliffe, G., 2011. Object-based Image Analysis for the Operational Production of High Resolution Land Cover: the Development of Landbase. In *Proceedings of the International Cartographic Conference 2011*.
- Tewkesbury, A.P. & Allitt, M., 2010. Urban creep mapping from Remote Sensing data. In *Proceedings of RSPSoc Annual Conference and Irish Earth Observation Symposium*. p. 34.
- Allitt, M. (Richard Allitt Associates Ltd.) & Tewkesbury, A.P. (Infoterra Ltd.), 2009. SUDS & Flood Management Initiative of the Year, Water Industry Achievement Award 2010. For: 'Impact of Urban Creep on Sewerage Systems'. <http://www.waterindustryachievementawards.info/2017/en/page/past-winners>
- Allitt, M. & Tewkesbury, A.P., 2009. Investigations into Urban Creep at 5 Cities. In *WaPUG Autumn Conference*. pp. 1–10.
- Tewkesbury, A.P. & Hinchliffe, G., 2009. LandBase: the development of a high resolution, generic land cover classification. In *Proceedings of RSPSoc Annual Conference*. pp. 746–755.

List of Appendix Figures and Tables

Figure A-1: Example field map, sample ID 7388, panel 1.....	217
Figure A-2: Example field map, sample ID 7388, panel 2.....	218
Figure A-3: PIF identification algorithm pseudo code.....	221
Figure A-4: Pseudo code outlining the hidden surface detection algorithm	222
Figure A-5: Pseudo code describing the CEMRS algorithm.....	222
Table A-1: Change proportions by stratum.....	219
Table A-2: Change proportions by population	220
Table A-3: Full feature listing	223
Table A-4: The top 15 representative features from all available features	228
Table A-5: The top 15 representative features from all difference features	229

1 Introduction

1.1 Overview

Change detection is one of the most active research areas in remote sensing, driven by the desire to monitor the highly dynamic world around us. Change detection research includes applications that monitor urban areas (Homer & Xian 2011; Lu et al. 2010; Doxani et al. 2011), forestry (Desclée et al. 2006; Bontemps et al. 2012; Cohen & Fiorella 1998), the natural environment (Linke & McDermid 2012; Stow et al. 2008), and natural disasters (Barazzetti 2016; Martha et al. 2016). Even though the huge body of work has been consolidated in a number of literature reviews (Tewkesbury et al. 2015; Hussain et al. 2013; Lu et al. 2004; Coppin et al. 2004), the science is still evolving rapidly. This is in part, due to the increasing uptake of very high resolution (VHR) images in recent years. VHR images, with ground resolutions below 1m -often referred to as 'VHR1' within the Copernicus program (European Space Agency 2015, p11)-, can image change in great detail. However, there is very little research focusing on detailed change detection, analysing the images at large cartographic scales. This thesis addresses this problem, by investigating the utility of remote sensing change detection to address fine urban changes called 'urban creep'.

Urban creep is the incremental, sub-property level, addition of impermeable surfaces, such as the construction of new driveways or patios (Richard Allitt Associates Ltd. 2008). Urban creep is difficult to identify (Trioulet 2012) and is not explicitly recorded in land use change statistics (Perry & Nawaz 2008). However, it is important because in the UK, it is proven to increase flood risk (UKWIR 2010; Ofwat 2011). Urban creep is typically small in extent and is defined by a complicated contextual setting.

Consequently, it poses a challenge to remote sensing change detection. Other researchers have sought to identify urban creep in multi-temporal aerial images (Duckworth 2005; Newcastle City Council 2008), but little research exists applying state-of-the-art remote sensing technology to the problem. This would prove challenging because documented issues with VHR image change detection such as coping with variations in viewing angle (Listner & Niemeyer 2011b) and shading patterns (Hussain et al. 2013) would likely confuse an analysis. Furthermore, the

application could prove so challenging, that the relevance of image differencing methods, such as change vector analysis (CVA) could be questioned. This thesis explores urban creep identification using remote sensing by means of a case study covering the city of Norwich, in the UK.

1.2 Problem definition

Urban creep is the addition of impermeable, sealed surfaces to a development after the initial construction (Wright et al. 2011). Examples include the paving over of a front garden, to create additional parking; a housing extension; or a paved, low maintenance garden. This type of impermeable surface addition is a change in land cover (not land use), typically from a vegetated to a non-vegetated state. This loss of permeable, often vegetated land is linked to a range of different impacts including increased flood risk, urban heat island effects and reduced biodiversity and carbon sequestration (Scalenghe & Marsan 2009).

Urban creep then is incremental change, made by the occupier to an existing property. This change is made based on occupier lifestyle choices such as the desire for a modern, minimalist garden or by addressing constraints such as a lack of parking space. The proliferation of such activity is complex and can be linked to factors such as increased car ownership, a limited availability of public transport and the popularity of low-maintenance gardens (Greater London Authority 2005).

This type of change may also be described as curtilage development (Department for Communities and Local Government 2008a) where curtilage is defined as 'A yard, courtyard, or piece of ground, included within the fence surrounding a dwelling house' (American Congress on Surveying and Mapping 1994, p133). However, curtilage is not purely a residential concept and relates to all land uses, with the definition based upon access and primary activity (Dickinson & Shaw 1977). In the same way that curtilage development can refer to a range of different land uses, so too does urban creep. Therefore parking space added to an existing industrial unit or retail outlet is also an example of urban creep.

While these examples may represent small individual changes; of little apparent significance, their cumulative effect is now beginning to be recognised (Wright et al.

2011). Recent research has indicated that across London, 30 km² of vegetated garden area was lost between 1998-99 and 2006-08 (London Wildlife Trust et al. 2011). Such a large area would no doubt have been subject to much scrutiny if concentrated in a few developments, yet when spread out over thousands of small modifications it is largely unchallenged. The combination of these small, incremental changes has a profound impact on local drainage and flooding. Specifically, there is a proven link between urban creep and increases in pluvial flooding through the sewer network (UKWIR 2010; Trioulet 2012; DEFRA et al. 2008). Furthermore, the magnitude of urban creep induced flooding far exceeds levels anticipated by small, apparently insignificant change (Ofwat 2011).

Urban creep is entirely different from urban growth or 'sprawl'. Urban growth relates to new developments either within the urban footprint on vacant or derelict land; or growing out at the fringes into agricultural, forested or unused land. Growth of this kind is a bold expression of change imprinting a significant visual impact on the landscape. Examples of this include new residential estates, retail complexes or newly established towns. These changes are usually associated with developmental control governance and a connection to the local highways, utilities and service infrastructure. Such governance and process act to limit development to those with significant societal value (Willis & Whitby 1985), while maintaining a sustainable support infrastructure. Therefore these changes, if legal and following process, are thoroughly documented and accounted. Urban creep, on the other hand, is a much more subtle change under limited developmental control. The combined effect of urban creep may be to overwhelm both infrastructure and ecosystem services beyond the initial, planned state of development. The concept of urban creep and urban growth is illustrated in Figure 1 with the use of a time series of aerial images.



Figure 1: Urban creep and urban growth visualised in bi-temporal aerial imagery. *The aerial images are captured over Exeter in 2003 (left) and 2006 (right). The South East of the scene is dominated by urban growth, highlighted in blue shading. The pre-urbanised region to the North West has undergone significant urban creep, highlighted in red. All imagery ©Airbus Defence and Space Ltd. 2017 and GeoPerspectives 2017.*

Urban creep identification is an under-explored avenue for remote sensing change detection. Early urban creep identification work by Cutting (2003) recognised that aerial images give very effective support to a field survey, helping to delineate the extent of impermeable additions. Later studies recognised the benefit that multi-temporal aerial imagery can bring, giving two or more specific snapshots in time. For example, Duckworth (2005) tracked urban creep in Keighley, Bradford over three epochs by sampling land cover in a set of aerial images between 1972 and 2002. Perry & Nawaz (2008) mapped urban creep for part of Leeds using aerial images from 1971

and 2004. Moreover, London Wildlife Trust et al. (2011) undertook a very ambitious project mapping the change in 1292 gardens throughout London. Each of these projects was the result of considerable mapping effort, prohibiting regular monitoring. Verbeeck et al. (2011) raise the point that remote sensing change detection, specifically with Object-based Image Analysis (OBIA) is an ideal technology for urban creep identification. However, there is very little work investigating this. A pioneering study by UKWIR (2010), applied a semi-automated OBIA workflow to map urban creep across five entire cities in the UK. This methodology is detailed further in the presentation by Allitt & Tewkesbury (2009), where pre-prepared OBIA land cover constrains a semi-automatic, layer arithmetic change detection. This work was later evolved by Tewkesbury (2011) to apply context to the changes with a rule-based OBIA. However, the full capabilities of remote sensing change detection have not been applied in this application. Therefore, there is a gap to build upon the work of Allitt & Tewkesbury (2009) and Tewkesbury (2011) with a thorough investigation of remotely sensed urban creep detection.

With urban creep detection, there is an opportunity to advance the capability of change detection techniques. In light of the challenges posed by VHR image change detection (G. Chen et al. 2012; Hussain et al. 2013) researchers have sought novel techniques to monitor change in the presence of view angle and shading differences. For example, Wu et al. (2016) and Wen et al. (2016) both compared aggregated land cover statistics as an indicator of urban change. Gueguen & Hamid (2016) compared aggregated statistics of feature shape to consistently report urban change, despite encountering a range of different imaging conditions. Moreover, neural networks have also been used to compare image patches directly to identify change (Pacifi & Frate 2010). These pieces of research effectively tackle difficult urban monitoring applications by modulating scale. Specifically, change is assessed by analysing the structural similarity of image patches, typically around 150x150 pixels in size (Wu et al. 2016). On the other hand, there is little research investigating large-scale change detection, aiming to identify fine land cover changes explicitly. Therefore, there is a clear gap to conduct more large-scale change detection. Urban creep identification is an excellent candidate application with which to focus. Unfortunately, this is a non-

trivial undertaking because the features of interest are on a comparable scale, and occur in close proximity, to VHR imaging problems such as building lean and scene shading. Furthermore, a high level of scene understanding is required to separate urban creep features from other changes effectively. Therefore, future research in this area would need to consider: 1) land cover change, imaging anomalies and scene context, and 2) a change detection comparison method capable of identifying complex relationships. Given this challenge, research questions would focus on the capability of change detection in this application. A specific question could focus on the capability of advanced comparison methods, capable of identifying complex change, such as a direct classification (Tewkesbury et al. 2015).

Given the potential complexity of urban creep identification, one could question the relevance of many classical change detection techniques when applied to VHR image change detection. Change detection techniques are rooted in a fundamental, and long-standing tenet that radiometric changes relate to real land cover changes (Singh 1989), with a large amount of work dedicated to identifying strong radiometric changes (Adar et al. 2014; Wang 2014a; Patra et al. 2011; Sinha & Kumar 2013; Bruzzone & Prieto 2000). Specifically, the comparison methods: layer arithmetic, transformation, CVA, and hybrid change detection (Tewkesbury et al. 2015), are all variations of this theme. However, the position of these methods in relation to VHR image urban change applications is not clear in the literature. For example, there have been calls for new methods (Lu et al. 2014), there is strong support for the use of OBIA instead of image pixels (G. Chen et al. 2012; Hussain et al. 2013; Boldt et al. 2012), and Bruzzone & Bovolo (2013) stress the importance of separating real land cover change from less interesting radiometric changes. Despite these advances, there is little experimental work investigating the utility and relevance of differencing methods in a complex urban change scenario. Therefore, it is an interesting counterpoint of this research to formally test the utility of differencing methods when applied to urban creep identification. Specifically, CVA would act as an ideal, representative differencing method because the breadth of the technique would facilitate simple Euclidean similarity measures (Homer & Xian 2011), and more complicated, and largely unexplored multi-dimensional change vector analysis (Bovolo et al. 2012).

Given the complexity involved in urban creep identification we can hypothesise that urban creep cannot be adequately identified with CVA. This thesis seeks to identify how true this hypothesis is, and in doing so push the boundaries of differencing methods and clarify their capability in VHR urban applications.

The city of Norwich is chosen as the case study for this research, aiming to monitor urban creep between 2006 and 2010. Norwich has a population of 132 500 (Office for National Statistics 2013) and is the primary settlement in the county of Norfolk, to the East of England. Norwich has a high risk of urban flooding (Wilson 2010; Pelling & Kelly 2010) and so is an ideal candidate for urban creep monitoring. High quality aerial images acquired in 2006 and 2010 were made available by Airbus Defence and Space Ltd. to support the research. Furthermore, previous work (UKWIR 2010) has estimated the rate of urban creep for Norwich between 1999 and 2006. Therefore, the outcome of this research can give insight into the changing nature of urban creep.

1.3 Research questions

To focus the research, three research questions are posed. The first question relates to the capability of remote sensing change detection in general. It seeks to find out if state-of-the-art comparison methods, such as a direct classification, can effectively identify urban creep. In answering this question, it is hoped to gain more insight into the capability of large-scale urban change detection. The second question is a counterpoint to the first question. Specifically, it seeks to assess the utility and relevance of differencing methods such as CVA, when applied to a complex urban change application. In the course of the research, the rate of urban creep in Norwich will be revised. Given the social importance of the subject, this revision is in itself an important outcome. Therefore, the third question seeks to assess the revised rate against previously published research.

1. Can a direct classification of multi-temporal VHR imagery adequately identify urban creep in Norwich?
2. Can a Change Vector Analysis of multi-temporal VHR imagery adequately identify urban creep in Norwich?

3. Has the rate of urban creep in Norwich changed since the last published estimate?

1.4 Research objectives

To structure the research, six objectives are outlined detailing the intermediate milestones to be achieved throughout the thesis.

1. Establish a baseline urban creep reference dataset with which to compare against the results of subsequent remote sensing classifications.
2. Establish a comprehensive set of object-based features contributing to the description of land cover change, context, and fluctuations in scene illumination and viewing geometry.
3. Identify the 'best case' direct classification to act as the remote sensing state-of-the-art, and as a point of reference for subsequent CVA classifications.
4. Exhaustively explore a range of different CVA configurations.
5. Use the reference data to revise the rate of urban creep in Norwich.
6. In light of the classification results, add to academic knowledge by clarifying the capability of change detection techniques in relation to urban creep identification and the wider discipline of remote sensing.

1.5 Thesis structure

Following on from this introduction, the thesis is structured in nine further chapters. Chapter 2 presents a change detection technique literature review, based on a previously published article (Tewkesbury et al. 2015). Chapter 3 discusses the challenges facing remote sensing change detection and proposes an agenda for future research. Chapter 4 reviews the social importance of urban creep and why it is an ideal application to challenge and advance change detection techniques. Chapter 5 presents the Norwich study area and describes the reference data collection methodology. Chapter 6 presents the results of the reference data collection to include quantitative change statistics and qualitative observations. Chapter 7 describes the remote sensing change detection methods, to include data preparation, feature and unit of analysis extraction, classification, and accuracy assessment. Chapter 8 presents the change detection results. Chapter 9 discusses all of the results,

towards impactful findings. Chapter 10 concludes the thesis, stating the contribution to knowledge and recommendations for future research.

2 Change detection techniques

2.1 Introduction

This chapter describes a concise, synoptic overview of change detection techniques, building upon, and advancing beyond existing reviews (Lu et al. 2004; G. Chen et al. 2012; Hussain et al. 2013; Singh 1989; Coppin et al. 2004; İlsever & Ünsalan 2012). The review is based on the presentation of change detection techniques formulated by Tewkesbury et al. (2015). This presentation focuses on the unit of analysis by which change is analysed and the method of comparison used to decide which elements have changed. Furthermore, this summary of change detection techniques is a welcome addition to the literature because it significantly reduces the conceptual overlap introduced by opposing pixel and object-based methods, promotes a more scientific foundation for experimentation and encourages more informed method selection (Tewkesbury et al. 2015).

This chapter contributes to the thesis by demonstrating a command and understanding of current change detection methods. The command of the discipline has steered this thesis to knowledge gaps and avenues for future research, which are presented in Chapter 3, and it also provides a foundation for rigorous change detection research conducted in Chapter 7.

Remote sensing change detection is a disparate, highly variable and ever-expanding area of research. There are many different methods in use, developed over several decades of satellite remote sensing. These approaches have been consolidated in several reviews (Coppin et al., 2004; Hussain et al., 2013; Lu et al., 2004; Radke et al., 2005; Warner et al., 2009) and even reviews of reviews (İlsever & Ünsalan 2012), each aiming to better inform applied research and steer future developments. However, most authors agree that a universal change detection technique does not yet exist (Ehlers et al., 2014), leaving end-users of the technology with an increasingly difficult task selecting a suitable approach. For instance, Lu et al. (2004) present seven categories divided into 31 techniques; making an overall assessment very difficult. Recent advances in Object-Based Image Analysis (OBIA) have also further complicated this picture by presenting two parallel streams of techniques (G. Chen et al., 2012;

Hussain et al., 2013) with significant conceptual overlaps. For instance, direct image comparison and direct object comparison (Hussain et al. 2013) could relate to identical operations applied to different analysis units. This chapter provides a clearer nomenclature with less conceptual overlap by providing a clear separation between the unit of analysis, be it the pixel or image object, and the comparison method used to highlight change.

Previous reviews (Lu et al. 2004; Hussain et al. 2013) have identified three broad stages in a remote sensing change detection project, namely pre-processing, change detection technique selection and accuracy assessment. This chapter focuses on the second stage, aiming to bring an improved clarity to a change detection technique selection. A change detection technique can be described in four components (Figure 2): the pre-processed input imagery, the unit of analysis, a comparison method and finally the derived change map ready for interpretation and accuracy assessment. To identify change(s), the input images are compared, and a decision is made as to the presence or degree of change. Prior to this, the geographical 'support' (Atkinson 2006) must be defined so that it is understood exactly which spatial analysis units are to be compared over time. At a fundamental level, this might be individual image pixels, but could also include: systematic groups of pixels, image-objects, vector polygons or a combination of these. With a comparison framework established, analysis units are then compared to highlight change. There are many different methods of achieving this, from simple arithmetic differencing, sequential classifications or statistical analysis. This comparison results in a 'change' map which may depict the apparent magnitude of change, the type of change or a combination of both.

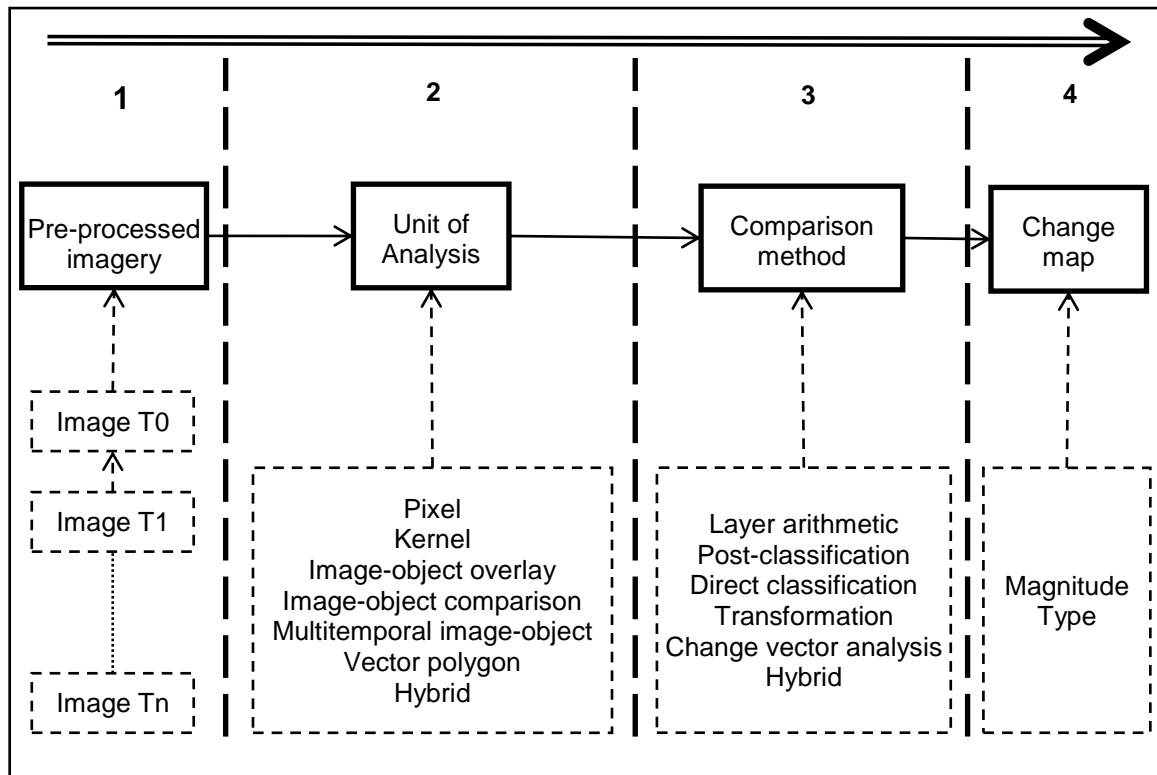


Figure 2. The four components of a change detection technique.

2.2 Unit of Analysis

Modern remote sensing and image processing facilitate the comparison of images under several different frameworks. In the broadest sense, image pixels and image-objects are the two main categories of analysis unit presented in the change detection literature (Hussain et al. 2013; G. Chen et al. 2012). When further exploring the possible interactions, there are in fact many more permutations by which a change comparison can be made. For instance, image pixels may be considered individual autonomous units or part of a systematic group such as a kernel filter or moving window. Listner and Niemeyer (2011a) outlined three different scenarios of image-object comparison; those generated independently, those generated from a multi-temporal data stack, and lastly a simple overlay operation. In addition to these, one could also consider mapping objects, typically vector polygons derived from field survey, or stereo or mono photogrammetry (Comber et al., 2004b; Sofina et al., 2012; Walter, 2004). Furthermore, a mixture of analysis units may be utilised, with this strategy sometimes referred to as a hybrid approach (G. Chen et al. 2012; Hussain et al. 2013). These elements are discussed in seven categories, namely pixel, kernel,

image-object overlay, image-object comparison, multi-temporal image-object, vector polygon and hybrid. These categories are summarised in Table 1 to include a brief description of each, advantages and disadvantages and some examples from the literature. To further clarify these definitions, illustrations are given in Figure 3, where the absolute change magnitude under each unit of analysis is depicted for a bi-temporal pair of images. The chapter then continues with a more detailed discussion of each unit of analysis.

Table 1: Analysis units commonly used in remote sensing change detection studies. The comparable features are based on Avery & Colwell's fundamental features of image interpretation; as cited by Campbell 1983, p43.

	Description	Comparable features	Advantages	Limitations	Example studies
Pixel	Single image pixels are compared.	Tone Shadow (limited)	Fast and suitable for larger pixels sizes. The unit does not generalise the data.	May be unsuitable for higher resolution imagery. Tone is the only comparable reference point.	Abd El-Kawy et al. (2011); Deng et al. (2008); Green et al., (1994); Hame et al., (1998); Jensen & Toll, (1982); Ochoa-Gaona & Gonzalez-Espinosa (2000); Peiman (2011); Rahman et al. (2011); Shalaby & Tateishi (2007); Torres-Vera et al. (2009)
Kernel	Groups of pixels are compared within a kernel filter or moving window.	Tone Texture Pattern (limited) Association (limited) Shadow (limited)	Enables measures of statistical correlation and texture. Facilitates basic contextual measures.	Generalises the data. The scale of the comparison is typically limited by a fixed kernel size. Adaptive kernels have been developed but multi-scale analysis remains a challenge. Contextual information is limited.	Bruzzone & Prieto (2000); He et al. (2011); Im & Jensen (2005); Klaric et al. (2013); Volpi et al. (2013)
Image-object overlay	Image-objects are generated by segmenting one of the images in the time series. A comparison against other images is then made by simple overlay.	Tone Texture Pattern (limited) Association (limited) Shadow (limited)	Segmentation may provide a more meaningful framework for texture measures and generalisation. Provides a suitable framework for modelling contextual features.	Generalises the data. Object size and shape cannot be compared. Sub-object change may remain undetectable.	Comber et al. (2004a); Listner & Niemeyer (2011a); Tewkesbury & Allitt (2010); Tewkesbury (2011)
Image-object comparison	Image-objects are generated by segmenting each image in the time series independently.	Tone Texture Size Shape Pattern Association Shadow	Shares the advantages of image-object overlay plus an independent spatial framework facilitates rigorous comparisons.	Generalises the data. Linking image-objects over time is a challenge. Inconsistent segmentation leads to object 'slivers'.	Boldt et al. (2012); Dingle Robertson & King (2011); Ehlers et al. (2006); Gamanya et al. (2009); Listner & Niemeyer (2011a); Lizarazo (2012)
Multi-temporal image-object	Image-objects are generated by segmenting the entire time series together.	Tone Texture Pattern Association Shadow	Shares the advantages of image-object overlay plus the segmentation can honour both static and dynamic boundaries while maintaining a consistent topology.	Generalises the data. Object size and shape cannot be compared.	Bontemps et al. (2012); Chehata et al. (2011); Desclée et al. (2006); Doxani et al. (2011); Teo & Shih (2013)
Vector polygon	Vector polygons extracted from digital mapping or cadastral datasets.	Tone Texture Association Shadow (limited)	Digital mapping databases often provide a cartographically 'clean' basis for analysis with the potential to focus the analysis using attributed thematic information.	Generalises the data. Object size and shape cannot be compared.	Comber et al. (2004b); Duro et al. (2013); Gerard et al. (2010); Sofina et al. (2012); Walter (2004)
Hybrid	Segmented image-objects generated from a pixel or kernel level comparison.	Tone Texture Pattern Association Shadow	The level of generalisation may be chosen with reference to the identified radiometric change. Although size and shape cannot be used in the comparison, it may be used in the interpretation of the radiometric change.	Object size and shape cannot be compared.	Aguirre-Gutiérrez et al. (2012); Bazi et al. (2010); Bruzzone & Bovolo (2013)

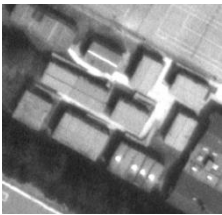
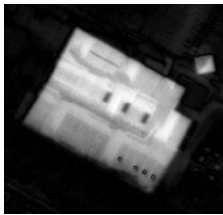
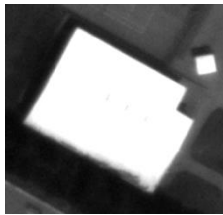
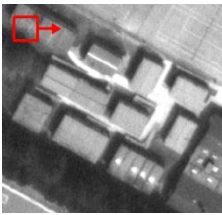

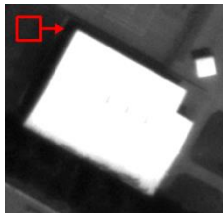
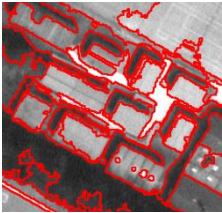

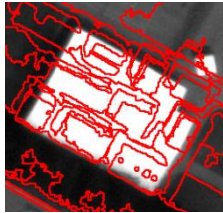
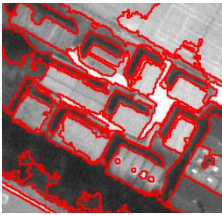
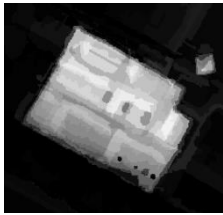
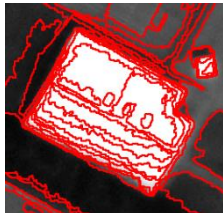
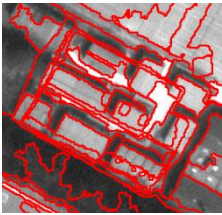
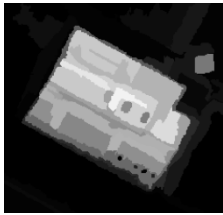
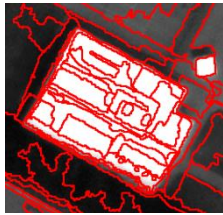
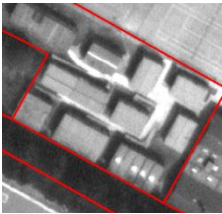
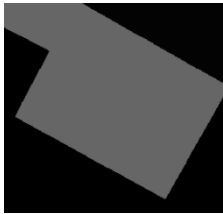
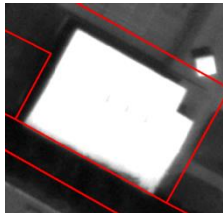
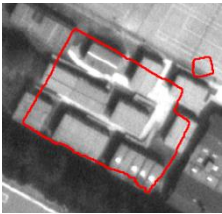

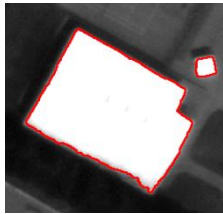
	Image 1	Change magnitude	Image 2
Pixel			
Kernel (moving window)			
Image-object overlay			
Image-object comparison			
Multi-temporal image-object			
Vector polygon			
Hybrid			

Figure 3. Commonly used remote sensing change detection analysis units. Image 1 is 25cm resolution aerial imagery over Norwich, UK from 2006. Image 2 is aerial imagery captured over the same area in 2010, also at 25cm resolution. The change magnitude is the absolute difference between Image 1 and Image 2 calculated over the respective unit of analysis. All imagery ©Airbus Defence and Space Ltd. 2017.

2.2.1 Pixel

The pixel is the most fundamental element of an image (Fisher 1997) and forms a convenient and well-used means of comparison. Since the launch of the first Landsat satellite in 1972 remotely sensed images have been analysed digitally, by comparing pixel intensities for changes in a range of applications. These include: urban development (Deng et al., 2008; Jensen & Toll, 1982; Torres-Vera et al., 2009), land cover and land use changes (Green et al., 1994; Ochoa-Gaona & Gonzalez-Espinosa, 2000; Peiman, 2011; Shalaby & Tateishi, 2007) and forestry (Coops et al., 2010; Hame et al., 1998; Wulder et al., 2008). The concept of comparing images is very simple, with arithmetic operations such as subtraction or division applied to continuous band radiance or reflectance (Jensen & Toll 1982; Green et al. 1994), or integer class labels (Abd El-Kawy et al., 2011; Rahman et al., 2011). These examples show that when the pixel spatially represents the anticipated change relatively well it can be a simple and effective focus by which to make change decisions, especially when there is a strong relationship between pixel intensity and the land cover transitions under investigation.

The pixel as a unit for change comparison does have many critics and is not seen as a suitable approach when considering modern Very High Resolution (VHR) imagery. For instance, G. Chen et al. (2012) argue that pixels have limited comparable classification features, typically just tone or radiance and so do not provide an adequate framework to model contextual information. Whereas Hussain et al. (2013) highlight that the pixel may be a source of geometric error, especially when integrating different data types. The overriding criticism of the pixel as an analysis unit for change detection is the susceptibility of producing spurious, noisy change pixels as a result of within class spectral variability and image registration issues. This issue, commonly referred to as classification ‘salt and pepper’ is widely discussed in the change detection (G. Chen et al. 2012; Hussain et al. 2013; Radke et al. 2005) and general remote sensing literature (Blaschke 2010; Baraldi & Boschetti 2012) as a prominent feature of pixel-based classifications, especially when dealing with VHR imagery. In light of these limitations, other means of comparison have been developed and implemented with a focus on groups of pixels.

2.2.2 Kernel

The use of a pixel kernel filter or moving window is a systematic way of generalising change results and introducing contextual information. By considering a local neighbourhood of image pixels, change can be interpreted statistically, aiming to filter noise and identify 'true' change. A neighbourhood of pixels is also a means of modelling local texture and contextual relationships by statistical and knowledge-based means. For instance, Im & Jensen (2005) used a neighbourhood correlation analysis to improve the identification of change information in VHR imagery by considering linear regression parameters instead of pixel radiance alone. The use of kernel-based texture measures have also proved to be a complementary addition to the change detection problem in several studies, including those by He et al. (2011) & Klaric et al. (2013). Furthermore, the use of contextual information is an effective method of filtering spurious change pixels (Bruzzone & Prieto, 2000; Volpi et al., 2013). These examples highlight the benefit of kernel filters; as a means of reducing spurious change and as a mechanism for allowing change decisions to be made beyond basic tonal differences. Unfortunately, kernel filters are often operated at a fixed scale, and the determination of optimum window sizes is not clearly defined (Warner 2011). Consequently, their use can lead to blurred boundaries and the removal of smaller features.

2.2.3 Image-object overlay

Objects segmented from one image may simply be overlaid on another to form the spatial framework for comparison (Listner & Niemeyer 2011a). Figure 3 illustrates this concept. These objects then form the basis of an arithmetic or statistical comparison of the underlying image pixels. Image-objects have been found to make the modelling of contextual information more accessible. For example, Tewkesbury & Allitt (2010) segmented aerial imagery and used mean image ratio differences to assist in the identification of impermeable surface change. In further work, a spatial knowledge base was applied to separate the identified change into those associated with existing properties and those that are part of a new development (Tewkesbury 2011). Research by Listner & Niemeyer (2011a; 2011b) segmented one image and then used a measure of object heterogeneity calculated on bi-temporal imagery to highlight

change. Comber et al. (2004a) overlaid classified image-objects on a pixel-based classification and then used expert knowledge to assist in the identification of true change from classification error. Overlaying existing objects onto new images can form a simple basis for change detection while benefiting from object-based contextual measures. The main disadvantage of this approach is that the geometry of the image-objects reflects only one of the images; with change in the opposing image not necessarily conforming to the imposed spatial framework.

2.2.4 Image-object comparison

The premise of image-object comparison is that two images are segmented independently so that the image-objects and their respective properties may be compared. The theoretical construct being that corresponding image-objects are 'linked' across space and time allowing a comparison to be made without the constraint of a geometric union. The distinct advantage being that all object properties can be compared including size and shape (Listner & Niemeyer 2011a) or class label (G. Chen et al. 2012). However, due to variations in factors such as illumination, viewing angle, phenology and atmospheric conditions, segmentations may be highly variable even under stable land cover and perfect co-registration.

The process of comparing one object with another is, therefore, complicated and non-trivial. Listner & Niemeyer (2011a) propose two approaches to comparison namely, *directed object correspondence* whereby an object is given a weighted sum of all overlapping objects and *correspondence via intersection* where object attributes are compared directly, but only over the spatial intersection created between the two time periods. The majority of the literature in this area uses the latter method, especially when applied to post-classification change (Boldt et al., 2012; Dingle Robertson & King, 2011; Gamanya et al., 2009). Image-object comparison by intersection is also illustrated in Figure 3. The main limitation of a spatial intersection of segmentations, also referred to as *correspondence via intersection*, is that it introduces a widely reported problem of 'sliver' objects under inconsistent segmentations (McDermid et al. 2008; G. Chen et al. 2012). Sliver objects can result in false change being detected and impacts the utility of updated land cover maps (Linke et al., 2009a). One method of minimising sliver objects is to simply remove smaller

change objects, as demonstrated by Boldt et al. (2012). However, this approach equates to a systematic reduction in the cartographic scale of the change analysis and information loss. Linke et al. (2009b) tackled this problem by using object width to highlight slivers before elimination. However, this approach remains insensitive to narrow change objects below the specified width threshold. While the work of Linke et al. (2009b) provides a robust strategy to suppress sliver objects more work is required on the rigorous matching of image objects so that their full properties may be used in a change comparison (Listner & Niemeyer 2011a; Hussain et al. 2013).

2.2.5 Multi-temporal image-object

Multi-temporal objects may be created by simply segmenting all available images together in a single data stack. Multi-temporal image-objects are illustrated in Figure 3. This approach has the distinct advantage of considering all images during object formation, therefore, minimising sliver errors and potentially honouring key multi-temporal boundaries. For example, Doxani et al. (2011) used this approach to detect urban change, an application that would be prone to widespread sliver errors due to differences in viewing geometry and shading. Teo & Shih (2013) also used multi-temporal image-objects as the basis for urban change detection, this time utilising LiDAR data, where it was found to perform well even in the presence of high magnitude spatial registration noise found at the edge of buildings. This approach has also proved successful in forest change applications at large (Chehata et al., 2011), moderate (Desclée et al., 2006) and small (Bontemps et al., 2012) cartographic scales. These examples show how multi-temporal image-objects are an elegant way of representing an image time-series, especially in applications involving elevated features where extensive viewing geometry differences are expected. However, this analysis unit is limited because object size and shape cannot be easily compared and smaller or indistinct changes may be generalised out during the segmentation process.

2.2.6 Vector polygon

Vector polygons originating from existing mapping databases can be overlaid over imagery and used as a basis to group image pixels in a change analysis. Groups of pixels across a temporal sequence may then be analysed statistically, the result of which may indicate changes within the corresponding polygons. This approach is often

linked to map updating, where remotely sensed images are used to automatically identify broad-scale change in polygons and regions, thereby reducing the manual review process. For instance, Walter (2004) calculated spectral means, variances and corresponding pixel class area for a set of land parcel polygons. These features were then used within a supervised classification to identify changed parcels. In a simpler workflow Gerard et al. (2010) overlaid recent CORINE land cover parcels against aerial images to visually assess historical changes over 50 years. These demonstrate how vector polygons can be used to spatially guide a change assessment. However, since the polygons often form part of a land informational database, this information may also be used to help inform the change detection process. For example, Comber et al. (2004b) used soil properties, rainfall and terrain to supplement the satellite spectral information when updating land cover mapping in Scotland.

Existing class labels can provide useful information in change detection workflows, allowing efforts to be focused and acting as a thematic guide for classification algorithms. For instance, Bouziani et al. (2010) & Sofina et al. (2012) used a 'map guided' approach to train a supervised classification algorithm to identify new buildings. While Duro et al. (2013) used cross-correlation analysis to statistically identify change candidates based on existing land cover map class labels. The use of vector polygons as a framework for change detection has great potential especially in cases where existing, high-quality attribution is used to inform the classification process. However, an assumption of this approach is that the scale of the vector polygons matches the scale of the change of interest. If this is not the case then a strategy will need to be considered to represent the change adequately; for instance, pixels may be used to delineate smaller change features within a vector polygon.

2.2.7 Hybrid

A hybrid approach refers to a combination of analysis units to highlight change in a stepwise way. In its most basic form, this relates to a change comparison of pixels which are then filtered or segmented as a mechanism to interpret what the change image is showing. For example, Bazi et al. (2010) first derived a pixel-based change image and then used multi-resolution segmentation (MRS) to group the results logically. Their approach proved successful when experimentally applied to Landsat

and Ikonos imagery. Figure 3 replicates the method employed by Bazi et al. (2010), first calculating the absolute difference between image pixels, then applying MRS to the difference image before finally calculating the mean absolute difference of the original images by image-object. Research by Linke et al. (2009b) found that MRS applied to pixel-based Landsat wetness difference images proved an effective method of identifying montane land cover change in Alberta, Canada. Aguirre-Gutiérrez et al. (2012) combined pixel and object-based classifications in a post-classification workflow that sought to retain the most accurate elements of each. Bruzzone & Bovolo (2013) modelled different elements of change at the pixel level to include shadows, registration noise and change magnitude. These pixel-based change indicators were then used to inform a change classification based on overriding multi-temporal image-objects. These examples show that using a hybrid of analysis units may be an intuitive approach whereby change in pixel intensity is logically grouped towards identifying features of interest.

2.3 Comparison methods

Previous reviews (Coppin et al. 2004; Hussain et al. 2013; Lu et al. 2004) have presented exhaustive lists of change detection techniques containing many comparison methods. Here six broad comparison methods are identified capturing the key features of previous research in a concise and accessible manner. These categories are summarised in Table 2 to include a brief description of each, advantages and disadvantages and some examples from the literature. This is followed by a more detailed discussion of each comparison method.

Table 2: An overview of commonly used comparison methods.

	Description	Advantages	Limitations	Example studies
Layer arithmetic	Image radiance or derivative features are numerically compared to identify change.	Can be simple to implement.	Usually gives little insight into the type of change.	Coulter et al. (2011); Dams et al. (2013); Desclée et al. (2006); Falco et al. (2013); Green et al. (1994); Homer & Xian (2011); Im et al. (2008); Im & Jensen (2005); Jensen & Toll (1982); Klaric et al. (2013); Lu et al. (2010); Tewkesbury & Allitt (2010)
Post-classification change	The comparison of multiple maps to identify class transitions.	Produces a labelled change map. Prior radiometric calibration may not be required.	Errors in any of the input maps are directly translated to the change map.	Abd El-Kawy et al. (2011); Boldt et al. (2012); Chou et al. (2005); Comber et al. (2004a); Dingle Robertson & King (2011); Gamanya et al. (2009); Hester et al. (2010); Li et al. (2012); Teo & Shih (2013); Torres-Vera et al. (2009); X. Chen et al. (2012)
Direct classification	A multi-temporal data stack is classified directly identifying both static and dynamic land cover.	Only one classification stage is required. Provides an effective framework to mine a complicated time series. Produces a labelled change map.	Classification training datasets can be difficult to construct, especially for a time series of images.	Chehata et al. (2011); Gao et al. (2012); Ghosh et al. (2014); Hame et al. (1998); Hayes & Sader (2001); Schneider (2012)
Transformation	A mathematical transformation to highlight variance between images.	Provides an elegant way to handle high dimensional data.	There is no defined thematic meaning to the results. Change may be difficult to locate and interpret.	Deng et al. (2008); Doxani et al. (2011); Listner & Niemeyer (2011a)
CVA	The computation of difference vectors between analysis units giving both the magnitude and direction of change.	Gives insight into the type of change occurring.	In its raw form, the change direction and magnitude may be ambiguous.	Bovolo et al. (2012); Bovolo & Bruzzone (2007); Bruzzone & Prieto (2000); Carvalho Júnior et al. (2011); Cohen & Fiorella (1998); Johnson & Kasischke (1998)
Hybrid change detection	The use of multiple comparison methods within a workflow. The most commonly used strategy is a combination of layer arithmetic to identify change and direct classification to label it.	Training data does not have to be collected over radiometrically stable areas.	No specific limitations.	Bruzzone & Bovolo (2013); Doxani et al. (2011); Seto et al. (2002); Xian & Homer (2010)

2.3.1 Layer arithmetic

Arithmetic operations such as subtraction or division applied to bi-temporal imagery are simple methods of change detection. These operations give an image depicting radiance differences, which is hoped reflect the magnitude of change on the ground (Singh 1989). This technique has long been used to highlight areas of image change quickly with minimal supervision (Jensen & Toll 1982; Green et al. 1994) and is still in use today, typically applied to image-objects (Desclée et al. 2006; Tewkesbury & Allitt 2010). To add thematic meaning to a difference image, image radiance may be transformed into a vegetation index or fractional cover image prior to the layer arithmetic. For example, Coulter et al. (2011) differenced regionally normalised measures of NDVI to identify vegetative land cover change while Tewkesbury & Allitt

(2010) used image ratios to identify vegetation removal in aerial imagery. It is also common to monitor urban expansion by subtracting multi-temporal impermeable surface fractional cover images obtained by sub-pixel analysis (Dams et al., 2013; Gangkofner et al., 2010; Lu et al., 2010). A highly evolved system of layer differencing is presented by Jin et al. (2013), whereby change is assessed by combining difference images of image spectral indices and biophysical transformations. These examples demonstrate how simple arithmetic operations of image radiance, or derivative features can be used to highlight changed areas, target specific features based upon an expected spectral response or quantify fractional, sub-pixel changes.

Layer arithmetic comparisons may go beyond simple radiometric differencing by leveraging different units of analysis. This empowers the comparison by considering texture, context and morphology; therefore reducing the dependency on a target's spectral characteristics as an indicator of change. For instance, Im & Jensen (2005) found that measures of kernel similarity –namely correlation coefficient, slope and offset- proved to be more effective indicators of change than simple pixel differencing. Further work showed that this same comparison method may also be applied to multi-temporal image-objects (Im et al., 2008); although no significant improvement was found when compared to the kernel based approach. When working with VHR imagery, several researchers have incorporated measures of texture and morphology into the arithmetic comparison as a means of reducing the dependence on image tone. For instance, Klaric et al. (2013) present a change detection system based on a weighted combination of neighbourhood spectral, textural and morphological features. The authors argue that this approach is not entirely dependent on spectral change and applies to multi-spectral and panchromatic imagery. The idea of reducing the dependence on spectral information is further developed by Falco et al. (2013) in research using Quickbird panchromatic imagery alone, as a basis for change detection, by comparing measures of morphology and spatial autocorrelation. Image change is not necessarily associated with a strong spectral difference, and these examples have shown how researchers have tackled this problem by using contextual information. However, there is still much research to be done in this area to improve classification accuracies over complex targets.

2.3.2 Post-classification change

Post-classification change (or map-to-map change detection) is the process of overlaying coincident thematic maps from different time periods to identify changes between them. The distinct advantage of this technique is that the baseline classification and the change transitions are explicitly known. Furthermore, since the maps may be produced independently, a radiometric normalisation is not necessary (Coppin et al. 2004; Warner et al. 2009). The direct comparison of satellite-derived land cover maps is one of the most established and widely used change detection methods. It is commonly applied to Landsat class imagery (Abd El-Kawy et al. 2011; Dingle Robertson & King 2011; Gamanya et al. 2009; Torres-Vera et al. 2009) and VHR imagery (Boldt et al., 2012; Demir et al., 2013; Hester et al., 2010). The approach may also be used to locate changes in a specific thematic target. For instance, Boldt et al. (2012) and Teo & Shih (2013) both used post-classification change to identify building changes uniquely. These examples show that post-classification change is a thematically rich technique able to answer specific change questions, making it suitable for a range of different applications.

Post-classification change is limited by map production issues and compounded errors, making it a costly and difficult method to adopt. The comparison method requires the production of two entire maps which may be an expensive (Lu et al. 2004) and an operationally complex task. Furthermore, input maps may be produced using differing data and algorithms. In this case, a distinction must be made between classification inconsistencies and real change, as explored by Comber et al. (2004a). The biggest issue with post-classification change is that it is entirely dependent on the quality of the input maps (Coppin et al. 2004; Lu et al. 2004). Specifically, individual map errors compound in the change map (Serra et al., 2003). Therefore, it is difficult and expensive to produce a time series of maps with sufficient quality to obtain meaningful change results.

There have been significant efforts to improve post-classification change results by accounting for classification uncertainty and by modelling anticipated change scenarios. Classification uncertainty may be spatial, thematic or a combination of both and accounted for by assigning confidences to these criteria. For instance, X. Chen et

al. (2012) compared fuzzy class probability, rather than crisp labels, to highlight uncertain land cover transitions. Hester et al. (2010) used spatial and thematic fuzziness in the classification of urban change using Quickbird imagery accounting for increased pixel level misregistration in VHR imagery. Specific change scenarios can also be modelled in an attempt to identify and remove unlikely land cover transitions. For instance, Chou et al. (2005) developed a spatial knowledge base, implemented as pixel kernel filters to remove change pixels that do not conform to predetermined change scenarios. This approach has also been extended to include full urban simulations as a means of identifying unlikely transitions (Li et al., 2012). These examples demonstrate that post-classification change has been extended from a simple map label arithmetic operation to one that considers the confidence of a particular label and the likelihood of its indicated change.

2.3.3 Direct classification

A multi-temporal stack of images can be directly classified to give a land cover inventory over stable areas and land cover transitions where change has occurred. The data stack consists of multiple sets of n -band images which may be treated by a classifier as one set of classification features. This is then classified with a supervised or unsupervised technique, aiming to give a set of stable land cover classes and changed land cover transitions. The technique is advantageous since only one classification stage is required, and identified changes are thematically labelled. Several researchers investigating forest change have used this approach as a means of directly identifying their target of interest. For instance, Hayes & Sader (2001), Hame et al. (1998) and Chehata et al. (2011) all implemented forest change detection systems based on an unsupervised classification of multi-temporal imagery, facilitated by a good understanding of the nature of the change. These examples from forestry applications show how the direct classification technique can be used to solve a relatively well-constrained problem. However, direct classification is a powerful tool in the context of a data mining problem such as the interpretation of a dense time series of images. Such a scenario is very difficult to conceptualise or model with expert knowledge, and is an ideal scenario for machine learning algorithms. For example, Schneider (2012) was able to successfully mine a time series of 50 Landsat images

from 1988 to 2010 for changes in urban extent using supervised support vector machine (SVM) and decision tree classifiers. The dense time series and machine learning approach allowed the extraction of meaningful change under complicated phenological patterns without explicitly modelling them. Gao et al. (2012) also used this strategy, applying a supervised decision tree classifier to extract impermeable surface change over 33 years using nine Landsat images. These examples demonstrate that direct classification of a time series of images can be an effective way of deciphering change that may be buried within complex patterns. However, deriving training datasets for such a classification can be very challenging (Lu et al. 2004) and unsupervised approaches can prove unresponsive to small magnitude change patterns (Warner et al. 2009). In light of these limitations, recent work by Ghosh et al. (2014) into semi-supervised change classification is extremely interesting with more research needed in this area.

2.3.4 Transformation

Data transformations such as principle component analysis (PCA) and multivariate alteration detection (MAD) are methods of data reduction by suppressing correlated information and highlighting variance. When applied to a multi-temporal stack of remotely sensed images there is the potential to highlight image change, since it should be uncorrelated between the respective datasets. For instance, Deng et al. (2008) applied PCA to a multi-temporal data stack of Landsat and SPOT 5 imagery to identify changed areas for a subsequent supervised change classification. The PCA image was classified into 'change' and 'no change' domains by labelling unsupervised clusters. In this case, 60 clusters were required to identify the change, indicating that the change signal was relatively well 'hidden' within the principle components. Doxani et al. (2011) found that applying the MAD transformation to image-objects was an effective method of highlighting change objects in VHR imagery. Listner & Niemeyer (2011a) also applied a MAD transformation to image-objects to highlight change. However, they highlighted that the MAD transformation might become mathematically unstable when applied to highly correlated features. This is particularly relevant when considering the large number of classification features available under OBIA. To ensure a robust change detection strategy, they proposed a

prior PCA, with the first three principle components acting as the inputs to the MAD transformation. Although this strategy worked in their application, it does highlight an issue with transformations, namely that the first 2 or 3 components may not necessarily contain the desired change information (Bovolo et al., 2012). Therefore, change features may either be missed or buried within a high number of transformation components. Furthermore, PCA and MAD transformations are scene dependent and may prove difficult to interpret (Lu et al. 2004; Warner et al. 2009; Carvalho Júnior et al. 2013). Transformations can be a useful way of assessing change within a complex time series of images. However, they usually only serve to highlight change and therefore should form part of a hybrid change detection workflow to provide change labels. Lastly, due to scene dependence, it may prove a difficult task to locate change within the multiple components, if the change is represented at all.

2.3.5 Change Vector Analysis (CVA)

Change vector analysis is a method of interpreting change based on its magnitude and direction. To facilitate this, bi-temporal datasets are described in three components; namely the feature vector at time 1, the feature vector at time 2 and an interconnecting vector. The interconnecting vector is called the change vector (CV), and its magnitude and direction can give us an insight into the type of change occurring. The geometry of a CVA is given in Figure 4a (in 2D for simplicity). Calculating the magnitude is very simple (see Cohen & Fiorella, 1998, p 91), easily extended to high dimensional feature space. For instance, the change magnitude of all six Landsat spectral bands (excluding the thermal) is often calculated to assess the apparent extent of change (Bruzzone & Prieto 2000; Xian & Homer 2010). In theory, the magnitude gives the degree to which the image radiance has changed, containing limited thematic content, while the direction indicates the type of change. Therefore, the combination of magnitude and direction can be a means of labelling change and minimising false positives (Bovolo & Bruzzone 2007). In the standard formulation of CVA (Figure 4a), the direction is described by a directional cosine for each axis of the feature space. Therefore, $n-1$ directional cosines are required to describe the change direction in n -dimensional feature space, leading to a complicated output data array which may be difficult to interpret (Carvalho Júnior et al., 2011). In light of this, many

researchers simplify the input feature space to two bands only. For example, Bovolo et al. (2007) defined a 2D feature space based on Landsat bands 3 and 4, allowing burnt area change to be uniquely identified from magnitude, and a single angular direction. Transformations can also be used to reduce multi-dimensional data down to two components, ready for CVA. Cohen & Fiorella (1998) and Johnson & Kasischke (1998) used this approach, transforming the six available Landsat bands into tasselled cap components as input into a 2D CVA. These examples highlight how CVA has the potential to be used as both a change identification and labelling tool. However, a complicated description of n -dimensional change limits its application. This point is discussed in detail by Bovolo et al. (2012), who note that limiting CVA to 2-dimensional features space requires prior knowledge of the nature of the change occurring and may lead to a poor analysis through an ill-informed band selection. This highlights a clear need to more elegantly describe change direction in n -dimensional feature space.

More recently, there has been some interesting research describing how n -dimensional change directional information can be conveyed in a CVA. These have sought to use several image channels while retaining a simple description of the change direction. For instance, Carvalho Júnior et al. (2011) proposed the use of the spectral angle mapper (SAM) and its statistically normalised derivative, spectral correlation mapper (SCM), both well-established techniques, common in hyperspectral remote sensing. Such techniques are used to describe how similar any two n -dimensional vectors are to each other, and so has clear applicability to change detection. SAM, mathematically based on the inner product of two vectors (Yuan et al., 1998) is the single angle between two n -dimensional vectors (Figure 4b). It is worth re-iterating that SAM and SCM are both measures of *similarity* and do not give change direction or type per-se. However, they can be highly informative and complementary to a change vector analysis (Carvalho Júnior et al. 2011).

The principle behind SAM was further explored by Bovolo et al. (2012) to relate the single angle back to change direction. This work used the same theoretical basis as Carvalho Júnior et al. (2011) but instead evaluated the angle between the change vector and an arbitrary reference vector (Figure 4c), and Bovolo et al. (2012)

normalised the reference vector by setting all elements equal to $1/\sqrt{n}$. The rationale for this approach is that the use of an arbitrary reference vector gives a consistent baseline for the change direction, allowing thematic changes to be consistently grouped throughout a scene. Bovolo et al. (2012) argue with reference to experimental examples, that this new formulation, Compressed CVA (C²VA) does not require any prior knowledge of the anticipated change or its remote sensing response. Moreover, the technique can identify more types of change since all of the available information is considered. These developments could go some way towards establishing CVA as a universal framework for change detection, as suggested by Johnson & Kasischke (1998). Considering future super spectral satellite missions and the wide variety of object-based features available, C²VA has great potential. At the time of writing, there is no published research applying and evaluating the theoretical work of Bovolo et al. (2012), therefore it is recommended that the application of n-dimensional CVA be explored further.

A little-reported limitation of CVA is that both the magnitude and direction can be ambiguous (Johnson & Kasischke 1998). Consider the three identified formulations of CVA shown in Figure 4a, b & c. It is evident that the change vector itself can be translated within the feature space, while retaining the same measures of magnitude and direction. There is the possibility that multiple thematic changes may be described by identical measures of magnitude and direction, limiting the power of CVA as a change labelling tool. In appraising this limitation, Cohen & Fiorella (1998) concluded that a baseline reference vector, typically from the first time period, should be used when attempting to further classify CVA results. This limitation of CVA is easily surmountable but clearly increases the burden of the interpretation task, especially in the case of high dimensional datasets.

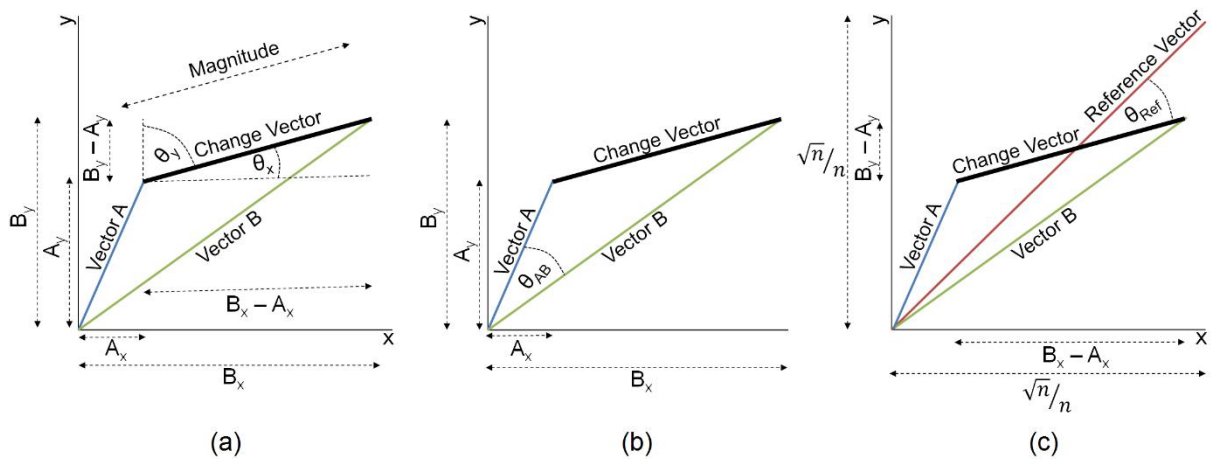


Figure 4: Two dimensional geometry of three formulations of CVA. For each, the x and y-axis represent the input features under analysis, typically spectral bands. Vector A and Vector B represent the value of a given analysis unit for a bi-temporal pair of images. (a) The 'standard' formulation of CVA describing the change vector by magnitude and a series of angular directions relative to each axis. (b) Spectral Angle Mapper (SAM) for CVA, after Carvalho Júnior et al. (2011). (c) n-dimensional, C²VA, after Bovolo et al. (2012).

2.3.6 Hybrid Change Detection

A hybrid approach uses more than one comparison method to increase the understanding of identified change. At an elementary level it could be thought of in two stages: Locating change and identifying change. This approach identifies change candidates, minimising reference data collection (Lu et al. 2004). Hybrid change detection is often expressed as a layer arithmetic operation to identify changed elements, followed by a supervised or unsupervised direct classification of the changed features giving them meaning (Lu et al. 2004). For example, Seto et al. (2002) first established a CVA depicting the radiometric change magnitude and direction, and then used a supervised classification to label change into specific land cover transitions. Whereas Doxani et al. (2011) tackled urban change detection in VHR imagery by first applying a MAD transform to highlight changed areas, and then applied a knowledge-based classification to filter and classify the results. An interesting formulation of hybrid change detection was presented by Bruzzone & Bovolo (2013). They argued that functional change detection must distinguish semantic change, relating to specific features from radiometric, or image change. This theory was experimentally implemented by combining pixel-based measures of shadow, radiometric change and noise within an object-based classification. These examples highlight a trend amongst research that seeks to use multiple stages of

change comparison to solve particular problems, a trend which is likely to continue as workflows become ever more complex.

2.4 Summary

This chapter has presented the current state-of-the-art in optical image change detection techniques to a clear, succinct nomenclature based on the unit of analysis and the comparison methodology. This presentation of change detection techniques is a welcome addition to the literature because it significantly reduces the conceptual overlap introduced by opposing pixel and object-based methods, promotes a more scientific foundation for experimentation and encourages more informed method selection (Tewkesbury et al. 2015).

The summary of analysis units shows that while the pixel remains the most common approach, object-based units are becoming more common. While image-object comparison is theoretically the most powerful unit, in light of inconsistent segmentations, matching image-objects over space and time requires far more sophisticated map conflation technology. Therefore, multi-temporal image-objects or a hybrid approach are likely the most robust analysis units, while the pixel is still suitable for many applications.

Post-classification change is the most popular comparison method due to the descriptive nature of the results allowing specific thematic questions to be answered. A direct classification of a complicated data stack is also an effective method of identifying semantic changes. However, the required training data is extremely difficult to obtain since the location of change is usually not known before an analysis. As highlighted by Lu et al. (2004) a hybrid approach may inherit the benefits of a direct classification while simplifying training data collection. Recent developments in CVA provide a powerful framework to compare multi-dimensional data but remain largely untested in the literature. Therefore, more research is required exploring recent formulations of CVA, in particularly the effect of integrating object-based features and other contextual measures.

As high level summary, change detection is rooted in a fundamental, and long standing tenet that radiometric changes relate to real land cover changes (Singh

1989). Specifically, the comparison methods: layer arithmetic, transformation, CVA, and hybrid change detection, are all variations of this theme. Furthermore, post-classification can also fall within this theme, if for example, a rule-based classification is applied. However, where VHR imagery is concerned, it is recognised that this is not always the case because of shading and mis-registration effects (Bruzzone & Bovolo 2013). Although these kind of VHR image effects are recognised and regularly cited (Hussain et al. 2013; Listner & Niemeyer 2011b), they are rarely discussed in detail. In light of this, the next chapter reviews the specific challenges faced by modern change detection and outlines a future research agenda.

3 Change detection challenges

3.1 Introduction

This chapter continues on from the review of change detection techniques presented in Chapter 2 and reviews the specific challenges of modern change detection. The content discussed builds upon, and goes beyond existing reviews (G. Chen et al. 2012; Hussain et al. 2013) by explicitly describing the underlying causes and mechanisms and establishes a research agenda towards solving these challenges. In particular, it describes:

- The large knowledge gap in the effect of 3D modelling of scene illumination and viewing geometry in change detection.
- The observation VHR change detection at large cartographic scales is under explored, and not well supported by current methods.
- The link between inconsistent segmentations and map conflation. Specifically, GIScience research could help to guide research in object-based change detection.
- The need to continually explore a variety of object-based features in change detection to bolster our knowledge of scene understanding.

This chapter contributes to the thesis by highlighting the specific challenges that prohibit the identification of complex urban change patterns such as urban creep. This thesis seeks to categorically illustrate that classical differencing methods cannot adequately describe complex urban change features, and that a more holistic approach is required. In particular, the highlighted change detection challenges must be addressed, to minimise their impact.

Remote sensing change detection is a vast subject that has evolved significantly in the last 30 years, but more research is required to tackle persistent problems. These include: scene illumination effects (Hussain et al. 2013; Singh 1989), changes in viewing geometry (Listner & Niemeyer 2011a; Lu et al. 2004), scale and the identification of small, ‘sub-area’ change (G. Chen et al. 2012), object-based feature utilisation (G. Chen et al. 2012; Hussain et al. 2013) and segmentation inconsistency

and comparison (Hussain et al. 2013; Listner & Niemeyer 2011a). This chapter expands these issues in more detail than previous discussions and crucially highlights potential solutions, establishing a research agenda.

3.2 Scene illumination

Scene illumination is a significant problem to change detection but it is rarely addressed in sufficient detail. Varying solar illumination angles over time give changes in shading patterns which can lead to false readings of change. Furthermore, deep shadows may mask the underlying land cover entirely leading to coverage gaps. The problem is exaggerated in very high-resolution (VHR) imagery because of varying local imaging times, and the scale of the shaded features relative to the image resolution. Moreover, urban change and forestry applications are particularly prone to errors of illumination because both landscapes have an inherently heterogeneous canopy with complicated shading and shadow patterns. Although the overall problem is widely reported in the literature (Lu et al. 2004; G. Chen et al. 2012; Listner & Niemeyer 2011a) the issue is rarely treated in enough detail to focus future research.

Scene illumination is expressed in two different ways, namely bi-directional reflectance and shadows. Bi-directional reflectance is the result of variations in image brightness from direct solar radiation, while shadows are regions occluded from direct solar radiation. Bi-directional reflectance is due to solar and sensor angles, atmospheric scattering and surface characteristics (Schaepman-Strub et al. 2006) and is a result of both direct (beam) and diffuse radiation (Chen 2011, p112). In practical terms, bi-directional reflectance equates to a decrease in image brightness as the solar incidence angle increases, i.e. more light is reflected away from the sensor. On the other hand, shadows are regions where the direct radiation is blocked. Therefore, the signal received at the sensor is a result of diffuse solar radiation only. Correcting for bi-directional effects is a well-established feature of satellite image pre-processing, and can be achieved by physical modelling (Richter 1998) or semi-empirical methods (Reese & Olsson 2011). More importantly, bi-directional reflectance correction methods have specifically been shown to improve change detection results (Tan et al. 2013). On the basis that bi-directional reflectance differences are an established and addressable part of the literature, they shall not be discussed further in this chapter.

However, the effect of shadows on the analysis of remotely sensed images is still an open research question and will be addressed in more detail. The remainder of the discussion focuses on suppression at source, illumination invariant metrics and shadow classification and masking techniques.

Suppressing illumination effects at source can be achieved by capturing all images with a similar sun position, ideally, as close to solar noon as possible. This concept has been a long-standing recommendation of the change detection literature (Singh 1989; Lu et al. 2004; G. Chen et al. 2012; Hussain et al. 2013). The strategy is feasible with sun-synchronous, fixed imaging geometry sensors such as Sentinel-2 and those in the Landsat program. However, this is extremely difficult with modern high and VHR satellite sensors because of variations in local time and viewing geometry. The situation is even more challenging for airborne sensors. In light of these challenges, and the rapid increase in the availability and use of VHR images, suppressing at source is a very limited and impractical approach. Therefore more emphasis must be placed on other strategies.

Illumination invariant methods and change metrics can be used to suppress the effect of scene illumination by minimising intensity changes in the reported change maps. Post-classification change is the most common method of achieving illumination invariance (Hester et al. 2010; Wen et al. 2016). However, the production of high quality multi-temporal maps can be a non-trivial task, especially for large-scale land cover applications. Therefore, image-to-image change detection methods invariant to illumination would be very valuable. In particular, Carvalho Júnior et al. (2011) propose the use of spectral similarity measures, such as Spectral Angle Mapper (SAM) and Spectral Correlation Mapper (SCM) which in theory remain insensitive to image intensity. Research testing these illumination invariant metrics is starting to emerge with encouraging results (Moughal & Yu 2014; Zhuang et al. 2016). However, while these methods should be effective in suppressing bi-directional reflectance effects and brighter shadows, in theory, they would not be insensitive to deep shadow changes. Furthermore, these methods would not provide effective illumination invariance to measures of texture or context. On this basis, illumination invariant metrics could be a

useful component of a change detection workflow but would not solve shading problems in isolation.

Image shadows can be masked to either constrain image analysis or attempt to reconstruct the underlying signature. Constraining delineates the shadow extent, excluding them from further analysis (Carleer & Wolff 2006; Van de Voorde et al. 2004; Hamedianfar & Shafri 2013). For change detection this would remove shadow induced false alarms. On the other hand, reconstruction will attempt to recover the data under the shadow. Shadow reconstruction can be applied via contrast enhancement (Guo et al. 2010; Kasetkasem & Varshney 2011; Liu & Yamazaki 2012), contextual classification (Salehi et al. 2012) or substituting with pixels from other, unshaded images (Zhou et al. 2009). Shadow correction for change detection is less common. However, demonstrations of shadow constraining are beginning to emerge (Bruzzone & Bovolo 2013; Volpi et al. 2013; Wang et al. 2015). These examples demonstrate that shadow masking is becoming increasingly important in change detection research due to the increased uptake of VHR imagery.

There are currently two main methods of shadow classification present in the literature, namely *spectral* and *contextual* classification. Spectral methods use image intensity, ratios and transformations to delineate dark areas. Contextual methods use neighbouring relationships to determine shadow position. Bruzzone & Bovolo (2013) propose the use of spectral methods of shadow masking in change detection studies. However, spectral methods can be prone to confusion with other dark areas (Shahtahmassebi et al. 2013). Therefore context is crucial in effective shadow classification. G. Chen et al. (2012) comments that OBIA is a suitable framework to model shaded features utilising contextual elements. For instance, Comber et al. (2012) used the sun's position and image object orientation to more effectively identify shadows, while Shao et al. (2011) used a model of expected land cover adjacency to highlight building shadows. Lastly, even when using a contextual analysis Huang & Bu (2015) report confusion with other dark objects, highlighting that effective shadow detection is a non-trivial task, especially in complex urban scenes.

Modelling is a third and underexplored method of shadow masking, which should be investigated in change detection research. Modelling uses the 3D geometry of the

scene and the Sun's position to simulate illumination. Only two conference papers (Nakajima et al. 2002; Zhan et al. 2005) were identified following this methodology. Shahtahmassebi et al. (2013) highlight that modelling is limited by the availability of 3D information, a point which is likely to explain the lack of research. The question of data availability is even more important for change detection because a rigorous analysis would require multi-temporal 3D data. However, modern digital photogrammetry and the latest sensors make the production of suitable Digital Surface Models (DSMs) technically achievable from both aerial (Dazhao, 2007; Gehrke, Uebbing, Downey, & Morin, 2011) and satellite sensors (Gehrke et al. 2011). Furthermore, high-resolution DSMs have long been available from commercial companies for both aerial (Bluesky 2016; Remote Aerial Surveys 2016) and satellite sensors (Airbus Defence & Space 2016; DigitalGlobe 2016). Crucially, multi-temporal DSM change detection is beginning to emerge (Tian et al. 2014; Qin et al. 2015). Therefore as the availability of 3D information increases, shadow modelling becomes more feasible and research is required to demonstrate this within change detection workflows.

3.3 Viewing geometry

Scene viewing geometry is a problem in remote sensing change detection sharing many of the key issues linked with scene illumination. Elevated features will show differing degrees of lean (or parallax) depending on the position of the sensor. Therefore even stable, unchanged buildings and trees can be imaged very differently leading to false readings of change. In a similar manner to illumination, viewing geometry becomes increasingly important when considering VHR data sources and urban and forestry applications. This is due to VHR sensor pointing agility, the scale of the analysis relative to parallax features, and the inherent canopy heterogeneity of forested and urban surfaces. The issue of viewing geometry has been highlighted by some authors as a challenge in change detection studies (Lu et al. 2004; G. Chen et al. 2012; Listner & Niemeyer 2011a) but is rarely discussed in detail.

There are several strategies available to deal with viewing geometry; namely: suppressing at source, removal by generalisation, comparing structural similarity, modelling or true ortho-rectification. In a similar vein to scene illumination, it has

been a long-standing recommendation of the literature to use images with similar or identical viewing geometry (Lu et al. 2004; G. Chen et al. 2012; Hussain et al. 2013). This strategy is feasible when applied to satellite systems with a fixed nadir view such as the Landsat program. However, it is very difficult to obtain geometrically consistent VHR satellites images, and near impossible from unsystematic aerial surveys. Therefore, other strategies are required to account for this problem.

Generalisation is often used to suppress viewing geometry and other misregistration effects present in VHR imagery by simply analysing change at a reduced scale. The rationale of this approach is that false alarms introduced by misregistration will be relatively small in extent, compared to the features of interest. Therefore, targeting larger change features will suppress the apparent noise. For example, when identifying urban change with Quickbird imagery, Boldt et al. (2012) excluded all change with an extent of less than 300 pixels. While Pacifici & Frate (2010) reported Quickbird image change in 100x100 pixel units. Marchesi et al. (2010) presented experimental work identifying registration errors with CVA. The results showed that at full resolution, the noise was interspersed with the change and no change domain, whereas when the CVA analysis was performed at a lower spatial scale the noise was heavily suppressed, leaving a more pronounced changed domain. Recent work by Wen et al. (2016) showed that under large viewing angle differences, urban change detection quality was related to the size of the unit of analysis, where larger units generalised opposing image content changes. These examples show that current methods of dealing with differences in viewing geometry and other misregistration effects are very limited, with simple scale filtering the primary strategy. This is a point of concern if the change features or interest are small and would themselves be liable to removal.

The use of structural similarity measures is a new and exciting development in change detection research. In theory, measures of structural similarity will highlight differences in the nature of structures present in the image independent of their position. In particular, methods of change detection employing the bag-of-visual-words classifiers are beginning to emerge. For example, Wu et al. (2016) and Wen et al. (2016) used counts of different land cover types aggregated at the patch level as a

means of comparison. More interestingly Gueguen & Hamid (2016) used feature shapes as the image ‘words’ and demonstrated that the approach performed well with VHR imagery across a range of different viewing geometries and illumination conditions. This research direction shows great promise. However, Wen et al.'s (2016) results highlight that the bag-of-visual-words technique is prone to spatial generalisation. Specifically, larger units of analysis can compensate for differences in viewing geometry at the expense of detail. Therefore there is a challenge to evolve this technique further to remain sensitive to fine detailed changes.

Modelling the 3D properties of a scene is another method of viewing geometry compensation, but is largely absent from the remote sensing literature. Contrary to some descriptions in the literature (Marchesi et al. 2010; Bovolo & Bruzzone 2007; Bruzzone & Bovolo 2013), viewing geometry changes are not ‘noise’. They are in fact systematic and can be modelled accordingly. For instance, if the geometry of a building’s roof is known then its projection can be modelled at different viewing angles and compensated for in a change analysis. No examples in the remote sensing literature were found replicating this hypothesis, but some related research was found. For instance, Liu et al. (2010) projected 3D building models onto newly acquired imagery as a means of identifying changes. While in the computer vision literature, van de Wouw et al. (2016) demonstrated a novel method of projecting recent imagery onto baseline imagery using its 3D properties. No such research exists in the remote sensing literature. Therefore, modelling scene viewing geometry for change detection is a huge gap in the literature, with a clear avenue for future research. This is especially true considering multi-temporal DSMs are now being used for change detection (Tian et al. 2014; Qin et al. 2015).

Lastly, the production of true ortho-rectified images prior to change detection is a method that would effectively eliminate viewing geometry problems. A true ortho-rectified image is processed to reconstruct orthogonality at feature level rather than ground level as in standard ortho-rectification. Practically this results in images with building and tree ‘lean’ removed. This level of pre-processing is ideally suited to urban and forestry change detection using VHR imagery because the data stack will remain co-registered over complex canopies. The technique can apply to aerial (Sheng 2007)

and to a lesser extent satellite images (Bang et al. 2007). However, no research was identified using true ortho-rectified images for change detection. This omission is likely because of the data volumes involved. Specifically, true ortho-rectified images can only be produced where significant stereo overlap exists, and the scene elevation is modelled precisely at the image pixel level. While this would be an interesting area for future research at present, the volume of stereo data required would prohibit the use of this approach in most applications.

3.4 Analysis scale

The majority of the change detection literature is focused on the identification of spatially broad features, either a result of limited image resolution or an imposed analysis scale. A non-systematic survey of change detection articles reviewed for this thesis (see Table 3) found a significant majority (64%) utilised imagery with a resolution of 20m or more. Incidentally, this figure is likely to be much higher, given that this thesis and the Author's interest is focused on VHR image analysis.

Nonetheless, this gives a clear indication that VHR imagery is still under-utilised. While image resolution influences analysis scale, it is not a diagnostic indicator due to multi-scale object-based analysis, and conversely, sub-pixel analysis. In other words, when whole pixels are chosen as the unit of analysis, scale is proportional to image resolution. When pixels are aggregated to image objects of varying sizes, in the case of OBIA, or when fractional spectral mixtures are considered in the case of sub-pixel analysis (Dams et al. 2013; Lu et al. 2010) the scale of the analysis is not clear.

Therefore a meaningful breakdown of analysis scale is only possible if the change detection scale is explicitly reported. Unfortunately scale is rarely reported; only six of the articles identified here that did so. However, examples indicating the application of relatively large units of analysis, and aggregated change metrics, are common when analysing VHR imagery (Wu et al. 2016; Wen et al. 2016). This high-level summary highlights that change detection is normally conducted at relatively coarse scales, even when the source imagery contains more detail.

Table 3: Image resolution in change detection research. Based on a non-systematic survey of 83 articles.

Image resolution	Frequency	Percentage
<=1m	23	28%
1-20m	7	8%
20m+	53	64%

VHR image change detection is commonly reported at a reduced scale, when compared to the image resolution. Noisy, ‘salt n pepper’ elements in high-resolution change maps (and classified maps in general) are often targeted for removal by spatial filtering. Moreover, using a large unit of analysis will generalise the complex signal of VHR imagery. Both of these strategies relate to a reduction in the scale of the analysis. For example, Klaric et al. (2013) effectively demonstrate ‘salt n pepper’ removal using pixel-based filtering. Moreover, object-based methods are routinely used to suppress the high-frequency content in change detection (Cao et al. 2014; Hao, Shi, Deng, Zhang, et al. 2016; Tang et al. 2015; Boldt et al. 2012). In fact, the suppression of ‘salt n pepper’ effects is often cited as a key motivation for using object-based methods of change detection (G. Chen et al. 2012; Parmentier & Eastman 2014). Furthermore, it is common to use large units of analysis to report aggregated change over relatively large areas. For instance, Wen et al. (2016) reported change results aggregated over large (the size was unreported) square cells. Wu et al. (2016) tested a similar technique over 150 x 150 pixel blocks, while Pacifici & Frate (2010) applied a neural network analysis of 100 x 100 pixels blocks. These examples illustrate that even though VHR imagery has the potential to detect very fine changes, this is very rarely done with the tendency to report on a much-reduced scale.

The suppression of false change detection due to illumination and viewing geometry effects is the primary reason for reducing the scale of change analysis. Rigorously modelling and removing illumination and viewing geometry issues is a non-trivial task. Therefore, spatial filtering and generalisation are one strategy employed to suppress them. Furthermore, the thematic accuracy of VHR change detection increases at coarser scales (Wen et al. 2016; Mishra & Crews 2014). While it is clear that not all

applications require large-scale, sub-parcel analysis, for those that do this is a rather unsatisfactory state. While it is the assumption of many change detection methods, that change is more likely if adjacent to other change (Bovolo & Bruzzone 2015), we should challenge this assumption if the change is not associated with random noise processes. In other words, smaller change objects imaged with VHR sensors can be made up of many pixels, and their formation is not a random process. Therefore, in VHR imagery smaller, multi-pixel objects contain valid data and should not be removed by size and isolation alone. Furthermore, some applications such as urban creep analysis (Tewkesbury & Allitt 2010; Tewkesbury 2011) require the identification of small change features. In light of this, it is inevitable that applications requiring detailed change analysis must consider illumination and viewing conditions and therefore there is a clear case to specifically model and remove these effects to facilitate larger scale analysis.

3.5 Segmentation inconsistency and object comparison

Independent segmentations over time are highly likely to be inconsistent, primarily due to issues of illumination and viewing geometry. Image-object comparison is theoretically the most powerful framework for change detection (Tewkesbury et al. 2015) but is a non-trivial task due to inconsistent segmentations over stable land cover conditions. Tewkesbury et al. (2015) and chapter 2.2 of this thesis discuss current methods of tackling this issue, in particular, the identification and removal of sliver objects using spatial and morphological filtering and removal. While it is highlighted that linking of image objects over time is an issue (Listner & Niemeyer 2011a) and one that requires more research (Hussain et al. 2013), little has been discussed as to the nature of this work.

There are clear parallels between inconsistent segmentations in remote sensing change detection and map conflation in GIScience which should be further explored. G. Chen et al. (2012) highlights that the problem of linking image objects is most likely to be solved under a 'GIS framework'. This is, of course, indicating map conflation in GIScience. Ruiz et al. (2011) identify three domains of map conflation matching techniques, namely geometric (to include spatial), semantic and topological. For instance, recent research by Barazzetti (2016) demonstrated a geometric conflation

(affine transformation) for object-based post-classification change before sliver object removal. However, a rigorous conflation between segmentations would need to move beyond this and utilise two, or all three (geometric, semantic and topological) matching techniques. For example, Jones et al. (1999) demonstrated conflation using combined geometric and semantic matching criteria to identify change between multi-temporal vector polygons. While graph theory can help to explain object-based classification structure (Comber, Brunson, et al. 2012), it could be an interesting avenue of research for topological matching. These examples show that there are unexplored avenues in GIScience conflation that could be utilised to provide an improved framework for image-object matching over time.

3.6 Object-based feature utilisation

Object-based Image Analysis (OBIA) is still a relatively immature resource when compared to its conceptual potential. The conceptual framework of OBIA states that image objects are analysed based on their tone, texture, morphology and spatial context in a manner approximating image interpretation. For instance, Hölbling et al. (2015) used scene morphology, context and texture to improve the identification of landslides in Taiwan. While Toure et al. (2016) used multiple analysis scales to improve urban land use change analysis. However, OBIA is still far from delivering the level of image understanding or ‘geographic-based intelligence’ (Blaschke 2010) that was conceptualised and in fact, has gathered some criticism (Baraldi & Boschetti 2012). For instance, it is common for object-based change classifications to be dominated by the influence of spectral features (Yu et al. 2016; Hao, Shi, Deng, Zhang, et al. 2016) or exclusively use spectral features (Walter 2004; Bontemps et al. 2012; Hamada et al. 2013) to depict change. Recent commentary (Bovolo & Bruzzone 2015; Jian et al. 2016; Lu et al. 2014) highlights that simple image differencing is still the most widely used indicator of change from remote sensing. This is despite the consensus that object-based methods are the more advanced and preferred solution (G. Chen et al. 2012; Hussain et al. 2013). The cautious use of object features for change analysis is likely due to difficulties modelling complex features, as reported by Barazzetti et al. (2015). While effective for certain change targets, focusing on spectral features does little to bolster our knowledge of scene understanding. For object-based change

detection to develop towards its potential more research is required utilising a range of object-based features.

Future difference images and CVA should consider a range of object-based features to visualise a range of changes, both spectral and non-spectral. Visualising a selection of informative object-based features using CVA could be an interesting avenue for research, simplifying change detection workflows. Recent research by Bovolo et al. (2012) and Carvalho Júnior et al. (2011) both explore n-dimensional CVA and hypothesise that their use should be extended to include contextual and spatial information. Research by He et al. (2011) showed that image texture is a useful addition to a CVA analysis, while work by Tian et al. (2013) cleverly include feature height. A logical direction for future research would be to apply multiple object-based classification features to recent formulations of CVA as defined by Bovolo et al. (2012) and Carvalho Júnior et al. (2011). This type of research would effectively create pseudo difference images which may help to visualise and identify changes outside of the spectral domain.

3.7 Summary

There is a clear case to investigate and evaluate 3D change detection methods explicitly modelling scene illumination and view angles. As discussed in section 3.2 and 3.3 scene illumination and viewing geometry are critical issues in VHR change detection especially in urban and forestry applications. Furthermore, these two issues are intrinsically linked because they can both be physically modelled and corrected with suitable 3D information. Given that DSMs are feasibly generated from both aerial and satellite imagery and change detection studies have begun to emerge using multi-temporal DSMs there is a clear case to investigate complex 3D change detection. Specifically, scene illumination and viewing geometry should be modelled as part of the change detection process.

Automated change detection is more often conducted at small cartographic scales, with little research demonstrating the identification of small features at large cartographic scales. There is an increasing amount of good research utilising VHR imagery for change detection, many of which are tackling complex illumination and

viewing geometry challenges. However, many of these are achieved at moderate scales, heavily aggregating the image content. Applications identifying high-resolution urban changes such as urban creep analysis are not well served by these techniques. Therefore further research is required investigating larger-scale change analysis.

Inconsistent segmentation is a hugely challenging problem and should be investigated with reference to map conflation research in GIScience. In particular, the three matching criteria of geometry, semantics and topology should form the framework for future work. Recent research is already utilising geometric matching to some degree, and graph theory could be an interesting method of topological matching.

Compared to its conceptual potential object-based change detection and OBIA, in general, remains an immature technology. As highlighted in previous reviews (G. Chen et al. 2012; Hussain et al. 2013; Blaschke 2010), the identification and utilisation of pertinent object-based features towards a high level of image understanding is a huge challenge; one which is not likely to be solved in the short to medium term. Therefore a concerted effort must be made towards utilising the contextual power of OBIA, exploring a range of different object-based features for change detection to bolster our knowledge of scene understanding. In particular, the use of object-based features in CVA provides an opportunity to readily combine features and generate pseudo difference images illustrating non-spectral changes.

Urban creep identification provides an ideal, benchmark application with which to test modern change detection techniques. Firstly, urban creep features are typically small which challenges the default analysis scale for change detection. Secondly, observing small urban features is prone to confusion with illumination and viewing geometry changes. Lastly, urban creep features are often subtle with minimal spectral differences and effective identification will likely require textural and contextual support. These points along with the wider need for urban creep identification are discussed in Chapter 4.

4 Urban creep: a social concern and change detection challenge

4.1 Introduction

This chapter continues on from Chapter 3 by describing how the application of urban creep detection is an ideal platform with which to advance remote sensing change detection techniques. Furthermore, the context of urban creep is described, and it is explained why small, apparently insignificant changes are actually of great importance to society. Key highlights include:

- Despite the apparent insignificance of small changes, the cumulative effect is hugely important and has proven links to significant increases in urban flooding.
- As an example target for remote sensing change detection, urban creep provides an excellent ‘fit’ to the challenges identified in Chapter 3.
- Urban creep influences government policy, has resulted in new legislation and is an established fixture of urban planning in the UK. On this basis, continued monitoring to verify and update existing understanding is advised.

This chapter contributes to the thesis by introducing a real world case study that may expose change detection technique weaknesses and push strongly towards solving key challenges in the discipline. Specifically, preserving an analysis scale appropriate to VHR imagery, tackling the effect of scene illumination and viewing geometry and discovering object-based features that facilitate increased scene understanding.

4.2 Urban creep as a challenge to change detection research

Urban creep is an ideal application with which to challenge the state-of-the-art of change detection techniques because it is sensitive to the primary issues faced by the discipline. Chapter 3 emphasised that scene illumination, viewing geometry, analysis scale, object-based feature utilisation and inconsistent segmentation all pose significant challenges to change detection. The identification of urban creep features, will in most cases, be directly affected by these challenges. For example, a new

extension or driveway adjacent to an existing building will be prone to masking and confusion with scene shading and viewing angle effects. Features such as a conservatory or garden hard standing are far smaller than the vast majority of remotely sensed change detection targets. The detailed textual, contextual and multi-scale interpretation required to interpret urban creep features is an ideal test of the ability of object-based features to model high-level image interpretation processes. Lastly, since urban creep features usually reside in highly heterogeneous landscapes, segmentations will be highly inconsistent and prone to sliver objects that would be very challenging to separate from real change at this scale. These examples demonstrate that it is impossible to tackle this application without considering the key challenges in change detection. Therefore, research in this area is likely to push the state-of-the-art towards new breakthroughs and progress in the discipline.

The identification of urban creep features in remotely sensed imagery is directly affected by illumination conditions and viewing geometry. Typical examples of urban creep are small in extent and reside in close proximity to elevated features. Consequently, real changes are readily confused with feature lean or masked by shadows. For example, when interpreting multi-temporal aerial images to identify garden changes across London, the London Wildlife Trust et al. (2011) found that occlusions by feature lean and deep shadows were common and needed to be removed from the analysis. Perry & Nawaz (2008) also identified difficulty interpreting the state of change in the presence of shadows and feature lean. It was reported that a high level of contextual evidence, -such as a pathway running into and then out of a shadow- was utilised to assign change in these circumstances. Furthermore, Duckworth (2005) and Newcastle City Council (2008) both reported interpretation difficulties as a result of shadow and feature lean. These examples show that urban creep identification is a change detection application heavily influenced by illumination conditions and sensor viewing geometry.

Effective remotely sensed urban creep monitoring requires VHR change detection at large cartographic scales. Changes to existing properties are usually small, and their depiction requires a large-scale analysis of VHR imagery. Aerial imagery is the principal data source for this type of research. For example, Verbeeck et al. (2011) utilised

25cm imagery, Perry & Nawaz (2008) 60cm imagery, while several others have used aerial imagery at undisclosed resolutions (London Wildlife Trust et al. 2011; Newcastle City Council 2008; Duckworth 2005; Richard Allitt Associates Ltd. 2008; Cutting 2003). While it is clear VHR imagery is the primary choice of data, the imagery itself is also interpreted and classified at large-scales to facilitate the demands of the application. For instance, Verbeeck et al. (2011) identified sub-parcel changes in baseline mapping at scales ranging from 1:200 to 1:1250. Richard Allitt Associates Ltd. (2008) augmented 1:1250 scale mapping with small changes identified from field survey and aerial imagery. While Duckworth (2005) sampled multi-temporal aerial imagery 2-5 times per square metre, giving a very detailed analysis. Clearly then, this application demands spatially detailed results at the limit of most VHR imagery. From a remote sensing perspective, urban creep monitoring presents a scale of analysis far beyond most change detection applications, and in some circumstances urban creep features may be characterised as classification 'salt n pepper' or 'noise'. Therefore, change detection research focusing on urban creep or sub-parcel change would push the state-of-the-art.

Urban creep interpretation is a complicated task relying on a range of criteria and OBIA is a prime candidate to model these criteria towards automated classification. Effective identification of urban creep requires a complicated assessment of permeability indicators –such as colour or texture-, and context to include the association to surrounding developments. For example, Perry & Nawaz (2008) describe how they used fine-scale patterns and texture, and contrast between neighbours in their interpretation. Modelling by UKWIR (2010) found that the rate of creep was correlated to both building density and type, while Verbeeck et al. (2011) found that change probability increased near to buildings and roads. Furthermore, several studies have used historical panchromatic aerial images (Duckworth 2005; Perry & Nawaz 2008) meaning that the interpretation was entirely reliant on albedo, texture and contextual indicators. Verbeeck et al. (2011) comment that OBIA is an ideal framework with which to identify urban creep, with object-based features the mechanism to model contextual relationships. In fact, the modelling work of UKWIR (2010) was based on semi-automated change detection utilising OBIA (Allitt &

Tewkesbury 2009), while further work conducted over Exeter demonstrated the use of object-based features to separate urban creep from other impermeable surface change (Tewkesbury 2011). These examples demonstrate that urban creep identification is a truly multi-criteria problem requiring a range of image features and one that certainly challenges modern change detection techniques.

Urban creep identification would be further complicated by inconsistent segmentations where true change could be easily confused with sliver objects. A new building extension or garden pathway adjacent the established house will be prone to viewing angle and shading effects, which in turn would lead to inconsistent segmentations and sliver objects. Since sliver objects are typically removed using size and morphological filtering (Linke, McDermid, Pape, et al. 2009; Barazzetti 2016) small changes are at risk of removal. No research was found directly investigating the consequence of inconsistent segmentations on urban creep or sub-parcel change. However, Verbeeck & Van Orshoven (2012) found that large-scale garden classifications were improved when segmentations were simplified using auxiliary mapping. The impact of inconsistent segmentations and sliver objects on large scale change analysis is largely unknown, but urban creep would form an ideal case study with which to benchmark opposing methods.

4.3 Why is urban creep important?

4.3.1 Flooding

There is a proven link between urban creep and increases in pluvial flooding through the sewer network. Pluvial flooding occurs when the rate of runoff exceeds the capacity of the local drainage system, resulting in surface pooling of water (Han 2011). In the wake of the widespread UK floods in 2007 and recommendations made in the 'Pit Review' (London Cabinet Office 2008, p71), there has been a raft of work proving the link between urban creep and flooding, and assessing its impact. For example, results of an extensive modelling program across 97 catchments UK wide, indicated that urban creep could contribute a 12% increase in sewer flooding by 2040 (Ofwat 2011). While UKWIR (2010) found increases in flood locations of up to 23% within 20 years. Research by Trioulet (2012) found that urban creep in Cranbrook, East London

had the potential to increase the number of flood locations in excess of 20cm depth from 28 in 2010 to 45 by 2040. A pilot study in Bradford concluded that the combination of climate change and urban creep could increase surface water volumes by 77% by 2085 (Gill 2008). Hydraulic modelling in Birmingham indicated that adding 25% more paved area to existing properties would have the effect of increasing the volume of sewer flooding by 18% (DEFRA et al. 2008). These examples have proven the hypothesis in the 'Pit Review' that urban creep was likely a significant contributing factor to the UK wide floods in 2007. Furthermore, they describe significant problems with flooding in the future if the issue is not addressed.

The observed flooding impact of urban creep far exceeds the anticipated impact, considering the relatively small changes observed. One could easily conclude that urban creep would equate to relatively small, proportional contributions to pluvial flooding. However, research has shown that the impact is disproportionately large. For example, UKWIR (2010) estimated an increase of flood volume of up to 20% within 20 years. While Trioulet (2012) estimated between 9% and 22% increases in flood volume by 2030 for Cranbrook, East London. Moreover, Ofwat (2011) calculate that by 2040 urban creep is likely to contribute more to sewer flooding volumes than new housing and UKWIR (2010) state that it will often overshadow the effects of climate change. The disproportionate response is because of inappropriate drainage to waste and combined sewer networks. Urban creep and other uncontrolled development is often drained to the dedicated wastewater sewer or combined systems. As much as 50% of urban creep may drain into the waste water system (Ofwat 2011). This type of inappropriate connection overwhelms the system well beyond its initial design leading to sewer overflow (Butler & Davies 2011). This body of work shows that despite its relatively small extent, urban creep can disproportionately contribute to significant flooding.

4.3.2 Legislation and policy

UK government policy is influenced by urban creep, and legislation has been passed in an attempt to control it. Historically, many changes to existing properties -especially within gardens- were not reported in national mapping or land use statistics and therefore fall outside of planning laws (Perry & Nawaz 2008). However, on 1st October

2008, new legislation came into force requiring planning permission for any additional impermeable area greater than 5m² within the front or sides of a dwelling (Department for Communities and Local Government 2008b). This legislation comes as a result of a public enquiry into the widespread UK floods in 2007, the 'Pitt Review' (after Sir Michael Pitt the inquiry chair), specifically recommendation 9:

'Householders should no longer be able to lay impermeable surfaces as of right on front gardens and the Government should consult on extending this policy to back gardens and business premises.'

(London Cabinet Office 2008, p15)

However, even with this legislation established urban creep is still a consideration for the UK government because of its irregular occurrences and difficulties of control and monitoring as discussed in the 'Foresight Land Use Futures Project':

'even where there is control over urbanisation, 'urban creep' adds hard surfaces in an uncontrolled and unpredictable manner.' (The Government Office for Science 2010, p163)

Therefore, while the UK government has taken action and legislated to mitigate the effects of urban creep, it remains on the agenda. Specifically, urban creep monitoring will remain important as a mechanism to gauge the effectiveness of legislation and the ongoing risk that it poses.

Sub-parcel change also influences the legislation and policy of other governments beyond the UK. While the term 'urban creep' is rarely used outside of the UK, sub-parcel change -the underlying process of urban creep- is of interest universally. For example, urbanisation control legislation similar to the UK exists in Belgium (Verbeeck et al. 2011). While properties in the City of Austin, Texas are allowed a maximum extent of impermeable surface depending on a defined zoning scheme (Wu et al. 2007). Although sub-parcel change control laws exist in other nations, the nature of the legislation is highly variable (Wright et al. 2011).

4.3.3 Stream condition and biodiversity

Urban creep increases sewer discharge to local water courses, which in turn will elevate pollution levels and threaten biodiversity. The increased surface runoff caused by urban creep makes the prevalence of sewer overflows into local water courses more frequent, degrading stream condition. Instances, where the sewer overflows into the local watercourse, are called Combined Sewer Overflow (CSO) spills and are a deliberate design feature of the sewer system to alleviate surface water flood risk in storm conditions (Ofwat 2011). CSO spills are significant because the overflow contains a mixture of storm and waste water. The resulting cross contamination can diminish fish stocks through oxygen depletion as well as raise concerns over public health and water course aesthetics (Butler & Davies 2011). Urban creep has been shown to increase the frequency of these spills. For example, hydraulic modelling by UKWIR (2010) yielded a 15% increase in the frequency of CSO spills and a 29% increase in CSO spill volume over a 14-year timeframe. Modelling by Bradford Metropolitan District Council et al. (2008) also indicated a link between urban creep and increased CSO spills, while Anglian Water (2010) have recognised the potential environmental risk of creep induced CSO spills. This interesting set of results coming out of the UK water industry show a clear link between urban creep and increased sewer discharge into local water courses.

Urban creep may further contribute to reduced stream condition and biodiversity through increased runoff to local waterways. Most modern sewer designs have a separate storm water system which diverts surface water to local water courses with no treatment and little filtration (Butler & Davies 2011). This is a different type and mechanism of discharge to CSO spills discussed earlier. While not contaminated with waste water, storm water runoff can be contaminated by surface pollution. The types of pollutants expected include hydrocarbons, fertilisers and heavy metals (Wright et al. 2011). Impermeable surface runoff has been shown to damage stream condition and biodiversity by increasing pollution and water temperature and altering stream profiles (Slonecker et al. 2001; Stone 2004). Furthermore, surface pollutants are deemed the most significant threat to water resources by the U.S. Environment Protection Agency (Stone 2004). There is no empirical research specifically linking

urban creep to a degradation in stream condition and biodiversity. However, it is a reasonable assumption that urban creep in part, contributes to the widely recognised effect of urbanisation on stream condition and biodiversity.

4.3.4 Urban environmental quality

Urban creep may contribute to degrading urban environmental quality by impacting biodiversity, regional climate and public health and well-being. While urban creep always results in an increase in impermeable surface, it also results in a loss of vegetation and a reduction in urban green space. The removal of vegetation will always result in a loss of biodiversity (Scalenghe & Marsan 2009). For example, there is a proven link between urban intensification and reduced bird species diversity and abundance (Carbó-Ramírez & Zuria 2011; Blair 1996). Incremental impermeable creep is also likely to exasperate the urban heat island effect, where urban areas are linked to raised surface temperature (Pan et al. 2011; Weng et al. 2011; Huang et al. 2011). Lastly, the threat to public health may be elevated. Primarily because sewer overflow poses a significant pathogen and bacterial contamination risk (Butler & Davies 2011), but also due to the holistic link between urban green space and public health and well-being (James et al. 2009; Schipperijn et al. 2010). Even though there is no research explicitly linking urban creep to the degradation of urban environmental quality, there are strong indications that it would contribute to the proven traits of urbanisation.

4.3.5 Planning

In the UK, urban creep is an established consideration in urban planning. Impermeable creep is considered in new development sewer design and the continued maintenance and upgrade of existing infrastructure. For example, a set impermeable addition of between 5m² and 10m² per property might be considered (Wessex Water 2009) or a 10% increase as used by Scottish Water (Wright et al. 2010). More recently empirical evidence has resulted in the development of models, evolving creep over time. For instance UKWIR (2010) propose a mean rate of urban creep of between 0.4 to 1.1 m² per property, per year. Interestingly, Verbeeck et al. (2011) observed an impermeable creep rate of 1.3m² per property, per year in Leuven, Belgium. More sophisticated models proposed by UKWIR (2010) integrate building density and property type to refine estimates of creep rate and have been applied to samples nationwide in

research by Ofwat (2011). These examples show that estimates of the rate of urban creep are an important fixture of urban planning in the UK. Further monitoring will be required to verify and improve these models for future planning purposes.

4.3.6 Existing monitoring

Urban creep has been deemed worthy of monitoring by several researchers, with work dating back nearly two decades. Adopting urban creep as the thesis case study may seem like an obscure choice. However, many researchers have deemed it worthy of investigation with work conducted using a range of different methodologies. The body of work in this area is succinctly catalogued in Table 4. Attempts have been made to gain insight from existing use land use statistics (Whitehand & Carr 1999), but it is widely acknowledged that these statistics and large-scale mapping layers do not comprehensively quantify parcel permeability (Perry & Nawaz 2008; Richard Allitt Associates Ltd. 2008; UKWIR 2010). Therefore field survey and resident questionnaires (Wright et al. 2011) have been adopted, but with limited extent. Several studies (Cutting 2003; Richard Allitt Associates Ltd. 2008; Verbeeck et al. 2011) have recognised that aerial imagery may assist with field survey by extending visibility into back gardens. However, the use of a single image does not present any direct evidence of land cover change. Therefore, multi-temporal aerial imagery with a defined monitoring period is the common and most adopted approach (Duckworth 2005; Perry & Nawaz 2008; Newcastle City Council 2008; London Wildlife Trust et al. 2011; Allitt & Tewkesbury 2009; Tewkesbury 2011). Several authors (Perry & Nawaz 2008; Verbeeck et al. 2011) have noted the time-consuming nature of urban creep interpretation and the limited extent that can be covered. With this in mind, pioneering work by Tewkesbury & Allitt (2010) facilitated the analysis of over 533 000 properties using semi-automated OBIA for a UK wide investigation of urban creep (UKWIR 2010). These examples show that there is considerable interest in urban creep and researchers have gone to great lengths to monitor and begin to understand its behaviour.

Chapter 4 – Urban creep: a social concern and change detection challenge

Table 4: A summary of urban creep monitoring research.

Methodology	Study	Description
Land use statistics and large-scale mapping	The Changing Fabrics of Ordinary Residential Areas (Whitehand & Carr 1999)	Analysed planning applications between 1948 and 1992 in pursuit of small scale development trends. The authors noted the limitation to permitted house extensions only.
Field survey and questionnaire	Urban creep in Scotland: stakeholder perceptions, quantification and cost implications of permeable solutions (Wright et al. 2011)	Involved the visual inspection of property curtilage from the street to estimate the type and extent of apparently impermeable additions. This was supplemented by a resident's questionnaire regarding property evolution. The study yielded deep insight but over a very limited extent.
Image-assisted field survey	Property Creep: A Case Study (Cutting 2003)	Field surveyed ten sample sites in Oakwood; Derby supported by aerial imagery.
	Wessex Water final business plan 2010 - 2015: Appendix C6.ii Urban Creep Study (Richard Allitt Associates Ltd. 2008)	Used recent aerial imagery and mapping, historical mapping and field survey to identify the extent of urban creep over 30 samples throughout Bournemouth, Bristol and Weston-Super-Mare.
	Measuring extent, location and change of imperviousness in urban domestic gardens in collective housing projects (Verbeeck et al. 2011)	Field surveyed regions of Leuven; Belgium supported by the original building plans and recent 25cm aerial imagery.
Multi-temporal image interpretation	Assessment of urban creep rates for house types in Keighley and the capacity for future urban creep (Duckworth 2005)	Prepared detailed multi-temporal land cover maps from aerial images at 1972, 1989, 1997 and 2002 to track urban creep in Keighley, Bradford.
	An investigation into the extent and impacts of hard surfacing of domestic gardens in an area of Leeds, United Kingdom (Perry & Nawaz 2008)	Prepared detailed land cover maps from 1971 and 2004 using multi-temporal aerial imagery.
	Urban flood risk and integrated drainage - Ouseburn and North Gosforth Pilot Project (Newcastle City Council 2008)	Mapped impermeable surface in 1996 and 2005 for 100x100m sample sites across Newcastle using multi-temporal aerial imagery.
	London: Garden City? (London Wildlife Trust et al. 2011)	Used multi-temporal aerial imagery to estimate land cover change within 1292 garden plots across London.
Semi-automated multi-temporal image analysis	Impact of Urban Creep on Sewerage Systems (UKWIR 2010), further methodology in 'Investigations into Urban Creep at 5 Cities' (Allitt & Tewkesbury 2009) and Urban creep mapping from Remote Sensing data (Tewkesbury & Allitt 2010)	Semi-automatic land cover maps were prepared from recent imagery using OBIA. The image objects were then backdated against historical aerial imagery to identify changes. Changes were identified using a combination of object-based image ratio differences, recent land cover class and manual re-classification. Impermeable changes were then manually assigned as either 'creep' or 'growth'.
	Mapping the extent of urban creep in Exeter using OBIA (Tewkesbury 2011)	Semi-automatic object-based urban creep classification of multi-temporal imagery covering the whole of Exeter. Utilised and extended the UKWIR (2010) methodology to include automatic sub-division of creep and growth components.

4.4 Summary

This chapter has described why urban creep is an important application for change detection research. This is because the key challenges in change detection are deep-rooted in the effective identification of urban creep from remotely sensed images. Specifically, urban creep was argued to be an ideal application with which to explore scene illumination, viewing geometry, analysis scale, object-based features and inconsistent segmentations.

The importance of urban creep was summarised, demonstrating why it is a worthwhile monitoring activity. Primarily, urban creep has influenced government policy and resulted in legislation change, chiefly because of the proven link to urban flooding. Also, the wider impacts on stream condition and biodiversity and urban environmental quality were discussed. The body of research monitoring urban creep was summarised showing that there is interest in the subject and that significant work has been undertaken. Given that urban creep is now an established consideration in UK urban planning, interest in the subject will continue and further monitoring is recommended.

5 Study area and reference data collection

5.1 Introduction

The literature reviews in Chapters 2, 3 and 4 have described the current status of remote sensing change detection, the key challenges with which to focus future research and described how urban creep identification is a meaningful case study with which to challenge change detection techniques and contribute to effective land use policy. This chapter establishes the city of Norwich as the research study area, and describes the methods used to capture a reference data set quantifying urban creep within it. The reference dataset, derived in this chapter, will act as a benchmark to objectively measure the success of subsequent change detection results derived in Chapter 7. This chapter contributes to the thesis by introducing a meaningful case study area and demonstrating rigorous and accurate reference data collection methods. The reference data will also act to answer the third research question, concerning the current status of urban creep in Norwich.

Due to data currency constraints, it was not possible to apply stratified random sampling based on the change detection maps derived later in this thesis (Chapter 7). This presented a challenge because stratified random sampling is well supported in the literature (Congalton & Green 2009; Olofsson et al. 2014) and is universally recommended (Stehman 2009). Consequently, a novel cluster sampling method was designed to capture rare change classes in the absence of a stratification framework. An exhaustive and rigorous mapping procedure was employed, combining multiple sources of remotely sensed imagery and field observations, all contributing to a high-quality reference data set.

The chapter is structured in ten sections describing the study area, reference data collection, and analysis methods. Sections 5.2, 5.3, and 5.4 establish the study area and monitoring timeframe and describe the thematic information captured. Sections 5.5, 5.6, 5.7 and 5.8 detail the sampling design and mapping procedure. Sections 5.9 and 5.10 outline the methods employed to obtain aggregated, population level statistics from the samples. Lastly, Section 5.11 summarises the chapter.

5.2 The study area and monitoring timeframe

The city of Norwich in the UK was chosen as the case study area to monitor urban creep. Norwich is the primary settlement in the county of Norfolk in the East of England. Initially a Saxon town, Norwich is a historical city containing medieval, Georgian and modern architecture (Natural England 2010, p14). Norwich is relatively flat, with a maximum elevation difference of approximately 50m (Ogilvie et al. 2011, Figure A2). The surrounding countryside is also low-lying, consisting mainly of arable agriculture (Natural England 2010c; Natural England 2010a) bordered by the pasture rich Norfolk Broads (Natural England 2010b) to the South East, and woodland to the North. The district of Norwich has a population of 132 500 and a density of 34 people per hectare (Office for National Statistics 2013). Norwich is a moderately dense city when compared to 4.1 people per hectare for England as a whole, and 52 for Greater London.

Urban creep in Norwich is important because of the risk of local surface water flooding. Regional hydraulic modelling has identified eleven key areas across the city at risk of surface water flooding (Wilson 2010, p1). In fact, Norwich is ranked as having one of the highest numbers of properties at risk of surface water flooding in the UK (Pelling & Kelly 2010, p46), and the highest flood risk in Norfolk (Norfolk County Council 2011, p16). Between 2000 and 2010, Anglian Water records report that 65 properties have been affected by sewer flooding (Ogilvie et al. 2011, p39), while hydraulic modelling indicated rain events as low as 1 in 5 could cause sewer overflow (Ogilvie et al. 2011, p47 & 54). For extreme forecasts, such as a 1 in 200 rainfall event, 6500 properties may be at risk (Ogilvie et al. 2011, p19). The risk is such that there have been warnings against any further surface water runoff to combined and foul sewer systems (Pelling & Kelly 2010, p97). This warning could be interpreted as an indirect reference to urban creep because it often drains to foul or combined sewer systems (Ofwat 2011). Going further, the Norwich Strategic Flood Risk Assessment recommends that urban creep should be managed in gardens and curtilages, and be part of development control planning policy (Wilson 2010, p50-51). There is then a clear case to monitor urban creep over Norwich, for local drainage planning purposes, and to provide evidence for future planning policy formulation.

The Norwich study area was carefully considered to include areas important to local surface water flooding. In the UK, it is the responsibility of local authorities to draft a Surface Water Management Plan (SWMP) and coordinate with relevant stakeholders to action its outcomes. Therefore, the Surface Water Project Officer at Norfolk County Council, and a representative from Anglian Water were consulted to discuss the research. These joint discussions highlighted the three Critical Drainage Areas (CDAs) Drayton, Catton Grove and Sewell, and Nelson and Town close for inclusion in the study. In addition to the CDAs, the suburb of Sprowston, and Broadland Business Park were also included as extensions because Sprowston is an important hydraulic catchment, and Broadland is a hotspot of new development.

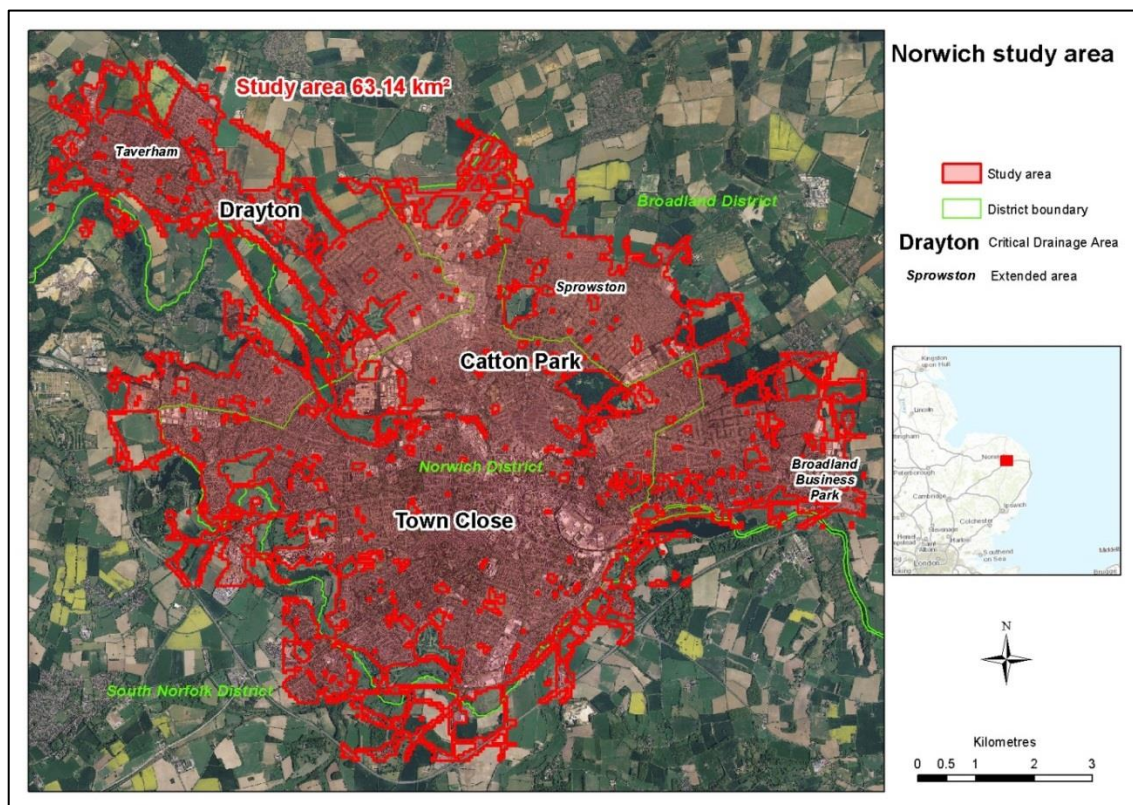


Figure 5: The Norwich study area. The area of interest is shaded in red. Backing imagery ©Airbus Defence and Space Ltd. (2017). Overview base mapping courtesy of ESRI and its data contributors 2013.

The study area focused on urban land to maximise population inference precision. For the purpose of this research, urban land was defined as having more than 1% artificial surfaces. Limiting to urban land excluded large areas of parkland, surrounding agricultural and wooded areas, all of which would not benefit from urban creep monitoring. Restricting the size of the population would also increase the precision of

sample with a fixed size (Stehman 2001). Artificial surface was determined by querying the 'MAKE' attribute of Ordnance Survey MasterMap™ (accessed May 2012). Unit percentages were evaluated against a 50x50m tessellation grid. Only urban grid squares intersecting Norwich district, a CDA, or one of the extended areas were retained. This resulted in a 63.14 km² study area, as illustrated in Figure 5.

The period 2006 to 2010 was chosen as the monitoring time frame because of data availability. High quality 25cm aerial imagery from July 2006 and May to July 2010 was made available by Airbus Defence and Space Ltd. This bi-temporal imagery is the primary source for reference data collection and the subsequent change detection described in Chapter 7. Previous urban creep research (UKWIR 2010), including evaluations of Norwich (Allitt & Tewkesbury 2009) indicate that it is highly likely that the change observed within this timeframe will be sufficient to undertake meaningful research.

A large, city-wide study area was deliberately chosen to replicate the scope of VHR image analysis often demanded by end users and to provide a sample pool large enough for meaningful change analysis. Modern VHR imagery can give us detailed information on ground conditions and their change. This type of imagery is now available globally from sensors such as Pleiades and Worldview-3, while VHR image availability is driving demand for increasing levels of information extraction. The demand for detailed spatial information, over wide areas, exists now and will likely rise rapidly in line with the huge increase in VHR imagery availability over the next 5-10 years. Observing a large study area, even with VHR imagery, is indicative of current monitoring requirements. Furthermore, a larger study area will increase the probability that relatively rare change classes can be sampled frequently enough to provide meaningful statistics.

5.3 Nomenclature

The reference data will honour a three class nomenclature: urban creep, *other change*, and *no change*. The use of three classes, as opposed to a binary urban creep/no urban creep scheme, will give a more informative description of errors

present in remotely sensed urban creep maps. The three classes are now briefly described.

5.3.1 Urban creep

Urban creep is the addition of impermeable surface to a development after the initial construction (Wright et al. 2011). Examples include the paving over of a front garden to create additional parking; a housing extension; or a paved, low maintenance garden. Urban creep is a change in land cover and context. Therefore, a particular land cover transition can be considered urban creep or *other change* depending on context. For example, a land cover change of herbaceous vegetation to tarmac would be labelled as urban creep if part of a driveway extension, but would not be labelled urban creep if part of a new road.

5.3.2 Other change

Other change encompasses all other land cover changes that are not determined to be urban creep. *Other change* includes the development of new properties (urban growth), construction and site clearance, tree removal and tillage and ground disturbance.

5.3.3 No change

This class is reserved for areas where no land cover change is evident. *No change* includes 'radiometric change' (Bruzzone & Bovolo 2013) such as shadow and building lean where no change on the ground has occurred.

5.4 Reference data objectives

The reference data sample design and data collection was specified according to the following objectives:

5.4.1 Objective 1

Provide an accurate estimate of urban creep and *other changes* to the urban area of Norwich between 2006 and 2010. Therefore a probabilistic sample of the city is required with enough change samples to accurately estimate population proportions. Given that previous research (Allitt & Tewkesbury 2009; UKWIR 2010) estimated

urban creep in Norwich between 1999 and 2006 at approximately 1% of the city, the maximum acceptable standard error is set at 0.25%.

5.4.2 Objective 2

The reference data must be collected before the test data creation to minimise the time difference between the image acquisition and field observation. In this thesis, as in most change detection studies (Khorram et al. 1999, p11), simultaneous field work and remote sensing image capture were not possible. Given that this thesis was prepared part-time between 2010 and 2017, it was crucial to obtain field validation as soon as possible, to minimise currency issues. Minimising currency issues equates to capturing the reference data, as soon as possible, and before preparing change detection maps.

5.5 Sample design

The sampling design was limited by Objective 2, requiring data collection before map production. The majority of the change detection literature apply stratified random sampling after map production (Im et al. 2008; Frohn 2006; Weng et al. 2011). This is with good reason. Stratified random sampling is well supported in the literature (Congalton & Green 2009; Olofsson et al. 2014) and universally recommended (Stehman 2009). Crucially, it facilitates accurate population inference from rare class samples (Padilla et al. 2015) by increasing their sampling rate. Consequently, precisely quantifying urban creep in the absence of a stratification framework presents a significant challenge.

One-stage cluster sampling was the starting point of the design. Cluster sampling selects samples from constrained sub-regions of the population, named Primary Sampling Units (PSU). Secondary Sampling Units (SSU) are the assessment units nested within each PSU. A one-stage cluster sampling design was chosen, sampling all SSUs. Figure 6 illustrates the cluster sampling implemented. While cluster sampling is usually employed to reduce costs (Stehman 2009), the primary motivation here is that each PSU acts as a container for undefined assessment units. In other words, at the time of the sample design, the assessment units did not exist, and the PSU acts as a container for mapped units when available.

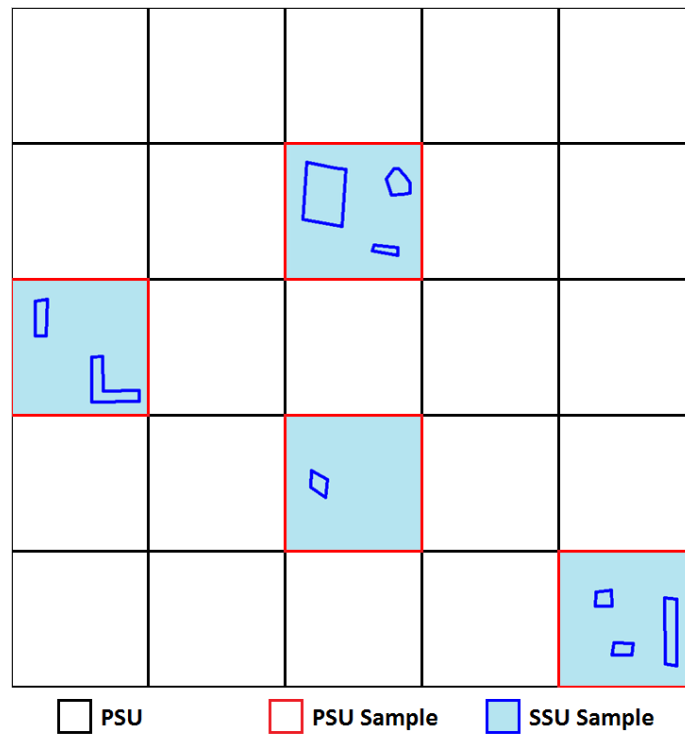


Figure 6: One-stage cluster sampling schematic. Sampled PSUs are depicted as red grid squares, while SSUs are shown as blue outlined polygons.

A qualitative assessment of different PSU sizes was undertaken, and a 50x50m grid square was chosen. 50x50m being large enough to encapsulate most instances of urban change while being sufficiently repeatable to sample the population meaningfully. The PSU size was chosen as a balance between meaningful feature representation, and estimation accuracy. A higher number of small PSUs will increase the accuracy of population inference (Stehman 2009), but prove more costly, and will more likely partially capture instances of change. Conversely, larger PSUs would be more cost effective to map and more likely encapsulate instances of change, while reducing inference accuracy. Intersecting the study area with a 50x50m tessellation grid gives a population of 25256 PSUs.

Polygon areas were selected as the SSU and mapped at a scale of 1:500 to minimise abstraction. The large cartographic scale ensures small urban creep features are represented in sufficient detail, with minimal abstraction from the reality on the ground. Furthermore, the reference data will likely reside at a level of detail superior to the classification maps; a feature recommended by Olofsson et al. (2014).

Continuous mapping of each PSU will form irregular polygons of different sizes. Since

the number of SSUs is not known at the sample design stage, it is estimated here to aid sample size and precision calculations. At a scale of 1:500, a Minimum Mapping Unit (MMU) of 4m² is adopted. We can then assume that each mapped polygon is an aggregation of one or more MMU sized polygons. On this basis, there are 625 SSUs per PSU.

The PSUs were stratified to increase population inference accuracy. While a simple random sample is statistically strong and unbiased (Czaplewski 2000), change is usually a rare occurrence (Foody 2002; Congalton & Green 2009), and simple random sampling may struggle to sample such classes with adequate accuracy unless very large sample sizes are employed (Stehman et al. 2005). Therefore, the population was broadly stratified to increase the representation of the rare change classes (Biging et al. 1998), in this case, urban creep and *other change*. Since the reference data was collected before map creation, no stratification framework existed. Therefore a method was formulated to estimate three PSU strata *No change candidate*, *Urban creep candidate* and *Other change candidate*. The purpose of each stratum being to increase the probability that a selected PSU will contain instances of its candidate class.

The strata were approximated using records of topographic mapping change and rapid visual interpretation. Firstly, The Ordnance Survey MasterMap™ change record was used to approximate the *Other change candidate* stratum by including all PSU's with >10% new buildings since 2006. It is assumed that urban creep would, by and large, occur outside of the *Other change candidate* stratum because it is not formally recorded as a change in MasterMap™. *Urban creep candidate* PSUs were identified by reviewing the remaining 'No change' area at a scale of approximately 1:2000. *Urban creep candidates* were highlighted with a simple Red-Green difference composite image prepared from the 2006 and 2010 imagery. Candidate identification was very time-consuming, and some of these candidates would be disproved under more rigorous inspection. Conversely, a detailed review of PSUs outside of the *Urban creep candidate* stratum may also display urban creep. PSUs not selected as *Other change* or *urban creep candidates* were labelled as the class *No change candidate*.

A random sample selection protocol was strictly employed to provide a scalable and unbiased estimate of the population. As stressed by Stehman (2009), a true probability sample is crucial for unbiased population inference. Therefore each sample within each stratum was selected entirely randomly. Random selection was chosen over a systematic scheme because of a more flexible sample size determination.

5.6 Sample size

Sample size estimation for single stage clustered sampling is a non-trivial task because of the difficulty in estimating class proportions and intra-cluster correlation. Here, sample sizes are calculated using reasonable approximations, but it was observed that small variations in these estimates could dramatically change sample size estimates. Therefore the final sample size allocation was determined by consulting the calculations and guideline sample sizes published in the literature.

Bennett et al. (1991) outline a cluster sample size estimation method based on a weighted variant of simple random sample size estimation. The number of PSU clusters required c , is estimated by the target standard error S , class proportion ρ , number of SSUs per PSU, M , and the design effect D . D takes the anticipated reduction in precision inherent in a cluster sampling design into account by inflating the sample size estimation.

$$c = \frac{\rho(1 - \rho)D}{S^2 M} \quad \text{Equation 1}$$

The design effect D is related to M and the rate of homogeneity of the clusters, r :

$$D = 1 + (M - 1)r \quad \text{Equation 2}$$

Bennett et al. (1991) highlights the fact that r estimates ‘the variability between clusters as compared to the variation within clusters’. Its derivation and calculation are detailed by Cochran (1977, p242), Where S_b^2 is the variance between clusters and S^2 is the population variance:

$$r = \frac{S_b^2 - S^2}{(M - 1)S^2}$$

Equation 3

For a stratified cluster sample design, c must be estimated separately for each class of each stratum. The maximum estimate for each stratum is used as the strata sample size allocation (Khorram et al. 1999, p43). Table 5 presents class and stratum level cluster sample size estimates. The class level estimates range from 10 to 92 clusters, based on a target standard error of 0.25%. The most clusters are required to estimate *No change* and *Other change* within the *Other change candidate* stratum. However, since the stratum proportions of these classes are likely to be much greater than urban creep, the defined target standard error may be over-specified for these classes. For example, relaxing the target standard error to 0.5% reduces the 'Other change candidate' stratum cluster requirements from 92 to 23.

Sample sizes are dependent on estimates of class proportion, PSU variance and strata variance. While class proportions are readily approximated from experience and qualitative observations, class variance is far more difficult to estimate and are the primary source of uncertainty in these estimations. Here, variances are approximated based on a constant PSU standard deviation of 5% and a constant stratum standard deviation of 2.5%. Due to this uncertainty, further guidance on sample size is required.

Table 5: PSU sample size estimation table

Stratum	No change candidate strata			Urban creep candidate strata			Other change candidate strata		
Class	No change	Urban creep	Other change	No change	Urban creep	Other change	No change	Urban creep	Other change
Number of SSU per PSU, M	625								
Target standard error, s	0.0025								
Target precision \pm (95% confidence)	0.0049								
PSU variance estimate	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025
Strata variance estimate	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006
Rate of Homogeneity, r	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048	0.0048
Design effect, D	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000
Estimated proportion, p	0.9800	0.0100	0.0100	0.9700	0.0200	0.0100	0.9000	0.0100	0.0900
Estimated number of clusters, c	20.0704	10.1376	10.1376	29.7984	20.0704	10.1376	92.1600	10.1376	83.8656
Estimated number of SSU samples per stratum	12293	63	63	18065	251	63	51840	63	4717
Maximum required clusters per stratum	20			30			92		
Maximum required clusters over the population	142								

There are several guidelines for sample size and stratum allocation in the literature. For example Congalton & Green (2009, p75) propose a ‘rule of thumb’ guideline of a minimum of 50 samples for each class while Stehman (2001) suggests a sample size of 100 will ensure a standard error no greater than 0.05. Olofsson et al. (2014) recommend that stratified class samples should be allocated approximately proportional to population proportions while increasing sampling of rare classes. In consideration of the results obtained in Table 5 and the guidelines in the literature a total sample size of a 150 PSU’s, with 50 PSU’s in each stratum is initially chosen as the sample size. Under this allocation, the target precision is maintained for the rare classes, while the relative proportions are approximately maintained by the variety of SSU samples within each PSU. During the data collection phase, it was possible to collect four more PSUs. Given the results obtained in Table 5, these additional samples were allocated to the ‘Other change candidate’ strata to increase its count to 54 and the total PSU’s to 154. Figure 7 shows the distribution of the final sample selection.

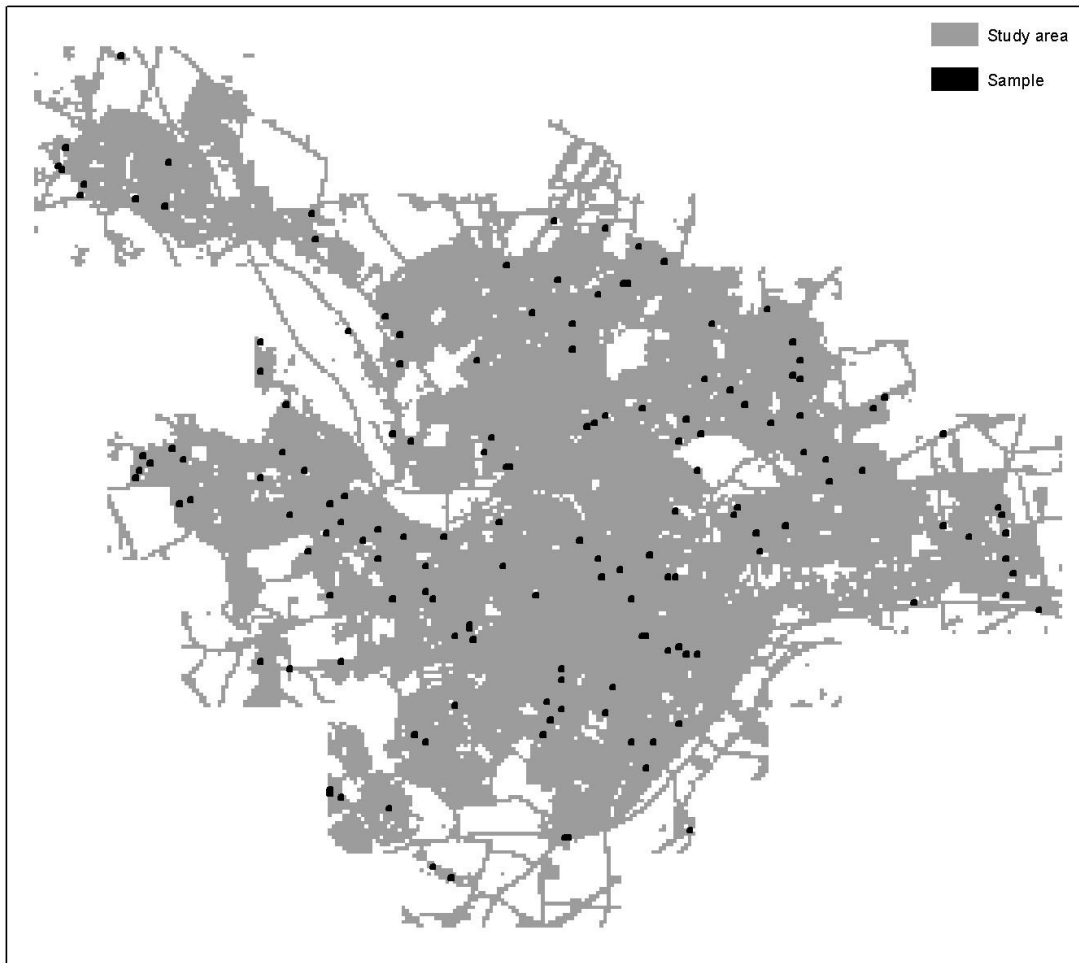


Figure 7: The distribution of the reference samples. The grey shaded area is the study area, and the black shaded squares are the 154 reference PSU samples.

5.7 PSU mapping

Each PSU sample was mapped at a scale of 1:500 using multi-temporal, multi-source and multi-scale remotely sensed imagery. The reference data was collected with as high spatial and thematic accuracy as possible, exceeding the anticipated quality of the classification with which it will be compared against (Stehman 2009; Olofsson et al. 2014). In this case, maximising spatial and thematic accuracy was the result of careful, methodical mapping, based on multiple sources of imagery and field observations. Features were mapped as polygons in ArcGIS 10.0 at a set scale of 1:500 and a defined MMU of 4m². The large cartographic scale minimised mapping abstraction. To reduce mapping effort, Ordnance Survey MasterMap™ was used as the initial spatial framework and was modified as necessary to capture the change features.

The multi-temporal imagery was inspected to assess the presence of changes between 2006 and 2010. The primary, baseline reference dataset was the 25cm Airbus Defence and Space Ltd. imagery. There are two main problems involved in large-scale image mapping of this kind. Firstly, the presence of occlusions from shadows and feature parallax and secondly, interpretation ambiguity; in particular difficulties associated with impermeable surface identification. Perry & Nawaz (2008) provide a good account of impermeable surface interpretation ambiguities at this scale. These issues are addressed by supplementing the Airbus imagery with remotely sensed data from Bing and Google maps. All of the sources used in the reference data collection are detailed in Table 6. The additional nadir and oblique aerial imagery significantly reduced occlusions and land cover ambiguity. Google street view proved extremely useful in reducing interpretation ambiguity, in particular by resolving gravel surfaces which may otherwise be interpreted as impermeable.

Table 6: Remotely sensed reference data.

Data type	Source	Capture year	Image resolution
Nadir aerial imagery	Airbus Defence and Space Ltd.	2010	25cm
Nadir aerial imagery	Airbus Defence and Space Ltd.	2006	25cm
Nadir aerial imagery	Bing maps	2011	10cm
Oblique aerial imagery	Bing maps	Various	Variable, estimated 15cm at frame centre
Nadir aerial imagery	Google maps	2006	10cm
Street view	Google maps	May-June 2012	Variable

An image interpretation key was defined to maintain mapping consistency and introduce a degree of confidence to the interpretation. Core to this was a four-point classification scheme describing: apparent permeability change and confidence, and

apparent contextual change and confidence. The interpretation key is shown in Table 7, and an example of its application can be seen in Figure 8. The key was designed to collect a wealth of thematic information, beyond the simple three-class nomenclature described in section 5.3. This added thematic redundancy, which was important for a long and evolving research project. The interpretation key was applied in four steps. Firstly, the apparent permeability was recorded, as either *No change*, *Change to impermeable* or *Change to permeable*. Next, the confidence of this change was graded from 1 (Very low) to 5 (Very high). The confidence classes were based on a range of land cover and transitions and validation scenarios. For example, field verified permeability changes are labelled with very high confidence. Thirdly, the context of the change is assigned one of seven contextual classes. The seven classes were defined to describe a range of common changes observed over the study area. Lastly, the confidence of the contextual class was also graded from 1 (Very low) to 5 (Very high). The resulting four-point class description gave a comprehensive, flexible account of the change. For the simple three-class nomenclature, the results were aggregated by the apparent permeability and context. For example, urban creep has a permeability change class of 1 and contextual class of 1. Ultimately, the interpretation key aided consistent classification across multiple sites and data collection sessions, a point corroborated by Khorram et al. (1999, p38).

Table 7: The reference data interpretation key. The top table gives the permeability change classes. The middle table outlines the context of change while the bottom table is a confidence scoring system for the two preceding criteria. Please note, specific data and context scenarios have been defined for permeability class confidence. On the other hand, contextual class confidence is more ambiguous. Using this criterion a score of 1 5 1 5 would relate to new tarmac, which has been field verified and is clearly a post construction addition to an existing residential property (urban creep with high confidence).

Apparent permeability change class	Description
0	No change
1	Change to impermeable
2	Change to permeable

Apparent contextual class	Description	Example
0	No change	No relevant land cover change identified.
1	Urban creep	Change to an established property.
2	Urban growth	Change associated with a new property. Construction of a property may be considered growth when the foundations are established.
3	Construction and site clearance	Site clearance, demolition and construction work without the establishment of a new building on the specific site.
4	Tillage and other ground disturbance	The exposure of bare earth through agriculture, horticulture or other unspecified activities.
5	Change to permeable	Either through re-development or specific conversions back to permeable.
6	Felled tree	Complete removal of a tree canopy.

Confidence class	Description	Permeability example	Contextual example
1	Very low	Obscured in aerial imagery. Changes associated with construction.	Complex re-build or on-going construction. ↓ New parking to front garden or a new build on green field site.
2	Low	Apparent new gravel in aerial imagery	
3	Moderate	Apparent new hard standing in aerial imagery.	
4	High	Clearly still vegetated or new building in aerial imagery. New hard standing in street level imagery.	
5	Very high	Field verified	

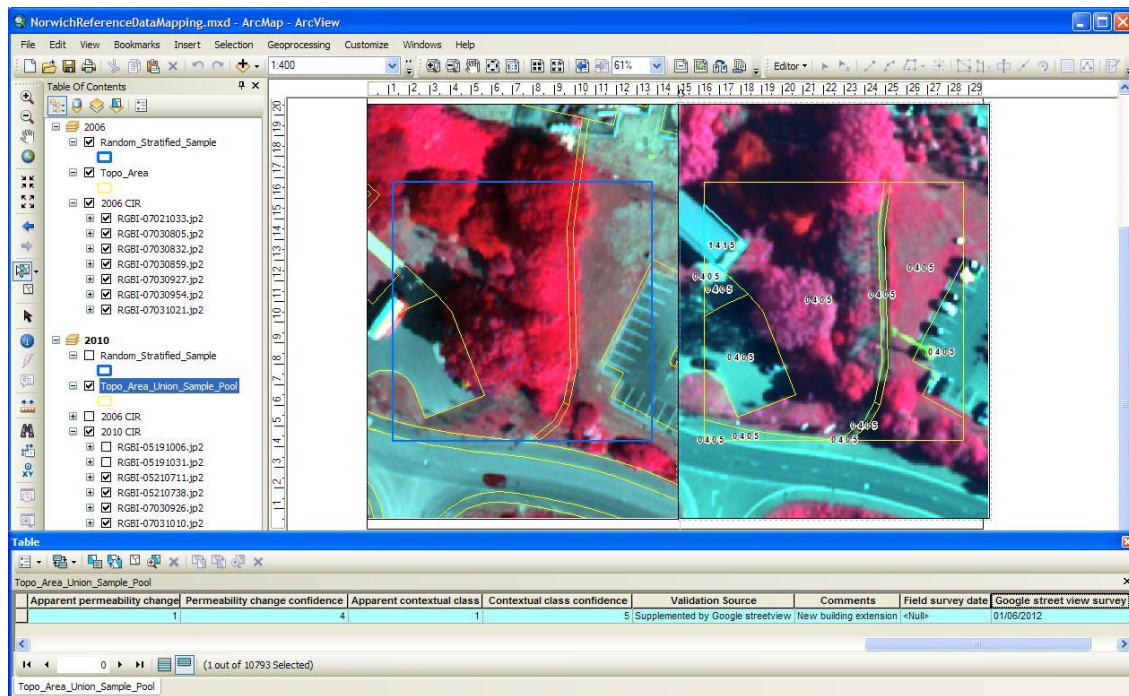


Figure 8: A mapped PSU sample. The example shows a new building extension highlighted, which was identified in both the aerial imagery and Google Street View. The map panel on the left contains the 2006 imagery overlaid with MasterMap™. The map panel on the right contains the 2010 imagery overlaid with the MasterMap™ derived reference mapping. Backing imagery ©Airbus Defence and Space Ltd. (2017).

5.8 Field checking and final review

Due to the resources involved, it was not possible to field validate all 154 PSUs. Therefore a random subsample of 34 (22%) PSU samples were field validated. The fieldwork was carried out between 15th and 22nd September 2012. Supporting image interpretation with partial field validation is a method recommended by Czaplewski (2000) and is employed by Aguirre-Gutiérrez et al. (2012). The field check samples were sequentially drawn at random from each PSU strata to maintain a strict probability sampling design.

Hardcopy maps were annotated in the field to record thematic labels, modified boundaries, access, photograph references and comments. Comments included any supplementary supporting evidence for the change interpretation, to include verbal verification from residents. An example annotated field map is included in the Appendix, Figure A-1 and Figure A-2. A considerable 27.5% of the polygons selected for field validation were not accessible and labelled as inaccessible. Inaccessible polygons were labelled from the remotely sensed data (Table 6) only.

All PSUs outside of the field validation were then reviewed to ensure consistency. The familiarity with the study area can have a significant impact on the quality of image interpretation (Lu et al. 2004) and the experience here was that the partial field validation did increase the quality and confidence of the whole sample. In particular, there were many examples where the aerial imagery indicated an impermeable surface increase, but field validation and Google street view review confirmed a permeable gravel surface. Therefore, combining multiple sources of aerial imagery, terrestrial imagery and field survey, increased the overall thematic accuracy of the reference data.

5.9 Population inference

Population class proportions and confidence intervals were inferred from PSU class proportion mean and variance. Class proportions were calculated for each PSU, by aggregating class areas of the underlying SSUs. This aggregation effectively collapses the stratified random cluster sample to a stratified random sample. Stratum level estimates of class proportion mean and variance were weighted and combined to give population-level estimates. Population level mean class proportions and standard errors facilitate the calculation of change areas with confidence limits.

Starting with the cluster sample, if we consider the i th PSU, is composed of M SSUs, with a_i SSUs classified to particular class, the PSU class proportion is calculated by (Cochran 1977, p246):

$$p_i = \frac{a_i}{M} \quad \text{Equation 4}$$

In other words, the SSUs are aggregated to a single property of the PSU, p_i . In this case, a random cluster sample collapses to a random sample. Therefore, the population proportion P is estimated by calculating the mean, \bar{p} , of the PSU proportions, p_i (Cochran 1977, p246). For a stratified random sample, stratum class proportion is equivalent to the stratum property y_h . Therefore, the stratum class proportion mean is given by (Cochran 1977, p90):

$$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$$
Equation 5

Where n_h are the number of PSU samples in the stratum h , and y_{hi} is the class proportion of the i th PSU. Note, that when inferring class proportion, y_{hi} is equivalent to Equation 4.

To move from a set of stratum class proportion estimates to a single population class proportion estimate, we must consider each stratum's influence or weight. Stratum weight is the simple ratio of stratum units, N_h , to population units, N (Cochran 1977, p90):

$$W_h = \frac{N_h}{N}$$
Equation 6

The population proportion is estimated by calculating the stratum weighted mean, \bar{y}_{st} , for L strata (Cochran 1977, p91):

$$\bar{y}_{st} = \sum_{h=1}^L W_h \bar{y}_h$$
Equation 7

To add confidence limits to the proportion estimates, we must calculate the standard error of the class proportion mean. If we repeated the same sampling procedure several times, and calculated class proportion means, the standard error is the standard deviation of these means. In other words, the standard error estimates the stability and precision of the sample. Although we do not measure the standard error from multiple samples, it can be estimated from the variance of the class proportions.

The class proportion variance of the h th stratum and i th PSU is given by (Cochran 1977, p95):

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$$
Equation 8

We can then estimate the variance of the stratum class proportion mean, \bar{y}_{st} (Cochran 1977, p95):

$$s^2(\bar{y}_{st}) = \sum_{h=1}^L \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^L \frac{W_h s_h^2}{N} \quad \text{Equation 9}$$

The standard error of the stratum class proportion mean, $s(\bar{y}_{st})$ is simply the square root of the variance, $s^2(\bar{y}_{st})$. Assuming a normal distribution, the confidence interval is then given by multiplying the standard error by a z-score, t , (Cochran 1977, p95).

Therefore, class proportion (\bar{y}_{st}) confidence limits are described by:

$$\bar{y}_{st} \pm ts(\bar{y}_{st}) \quad \text{Equation 10}$$

Given the Norwich study area covers 6314 Ha, class extents and 95% (a z score of 1.96) confidence limits, are estimated by:

$$\text{Class extent(Ha)} = (6314 \times \bar{y}_{st}) \pm (1.96 \times s(\bar{y}_{st}) \times 6314) \quad \text{Equation 11}$$

5.10 Urban creep rate

Urban creep is often reported in the units percent per year (Cutting 2003; Newcastle City Council 2008) to normalise results across different time periods and extents.

Therefore the area of urban creep, A_{creep} , is divided by the extent of the study area, A_{all} , and the monitoring time period in years, T_{years} , specifically:

$$\text{Urban creep rate, \% per year} = \frac{A_{creep}}{A_{all} T_{years}} \quad \text{Equation 12}$$

5.11 Summary

This chapter has established the case study, identifying urban creep in Norwich between 2006 and 2010. Environmental and water industry evidence show that Norwich has a significant problem with surface water flooding. Therefore interest may extend beyond the remote sensing and GIS scientific community, to environmental scientists and urban planners. A data search yielded remotely sensed imagery suitable for reference data collection and image change detection.

A stratified random cluster sampling scheme was designed to overcome the lack of a sample level stratification framework. PSU level strata were approximated using a combination of topographic mapping and rapid image interpretation. The approximate stratum facilitated oversampling of rare change classes to increase estimation accuracy. 154 PSU samples were mapped at a scale of 1:500 giving an equivalent of 96 250 SSU samples. The samples were exhaustively mapped, referring to multiple sources of remotely sensed data and field observations.

The class proportions of each of the 154 PSUs were aggregated to give population level change statistics. This was achieved by calculating the stratum weighed mean, while confidence limits were derived from the class proportion variance amongst PSUs. Lastly, the change statistics were converted to an urban creep rate, by considering the change proportions and the monitoring time frame.

6 Reference data collection results

6.1 Introduction

This chapter presents the results obtained from the reference data collection and analysis described in Chapter 5. Primarily, the results describe urban change statistics in Norwich between 2006 and 2010, extrapolated from 154 cluster samples. Each sample recorded change as two separate classes, urban creep and *other change*, to facilitate urban creep rate calculations. Furthermore, observations captured during the data collection are presented to add insight into the classification confusion and ambiguities encountered. In all, eight noteworthy results were recorded, including:

- Urban creep accounts for 0.95% of the study area, equivalent to a rate of 0.24% per year
- Accurate determination of small, sub-parcel impermeable features was achieved by consulting multiple sources, and could not be achieved with a single set of VHR images alone.

The results described in this chapter contribute to the thesis by establishing a reference data baseline, quantifying urban creep in Norwich, and gaining insight into the nature of likely classification errors and confusion. The primary objective of the reference data was to act as a precise, unbiased baseline with which to benchmark the change detection results presented in Chapter 8. The derived change statistics will specifically address the third research question concerning the changing rate of urban creep in Norwich. Lastly, presenting the difficulties, confusion and ambiguities observed are highly valuable in understanding the change detection results, and provide important insight for VHR image monitoring in general.

The results are organised in six sections detailing the change distribution and statistics, urban creep rate, image interpretation confusion, and ongoing urban creep. Lastly, the results are collated and summarised.

6.2 Change distribution

Figure 9 shows the distribution of urban creep and *other change* observed across the urban extent of Norwich, between 2006 and 2010. As expected, *other change* (Figure

9a) exhibits greater levels of change than urban creep (Figure 9b). The mean PSU percentage of *other change* is 19.27% compared to 1.86% for urban creep. This is because *other change* is normally expressed as new building developments of considerable extent. Furthermore, although *other change* occurs throughout the city, there are clear clusters of change. For instance, there are clusters of *other change* occurring near Bowthorpe industrial estate, to the West; Taverham, to the Northwest; the Southeast of the city centre; and near Broadlands business park, to the East. In contrast, urban creep is more evenly distributed across the PSU samples. However, there are clusters of urban creep visible. In particular, to the south of New Costessey, to the West; and to a lesser extent, Sprowston, to the Northeast. While urban creep is usually limited in extent, there are examples where an appreciable creep is observed. For example, there is a large occurrence to the South of New Costessey; another near Eaton, to the South; one near Broadlands business park, to the East; and one to the Northwest in the outskirts of Taverham. These examples, where urban creep covers a relatively large extent, are in most cases, due to the extension of an industrial or commercial building.

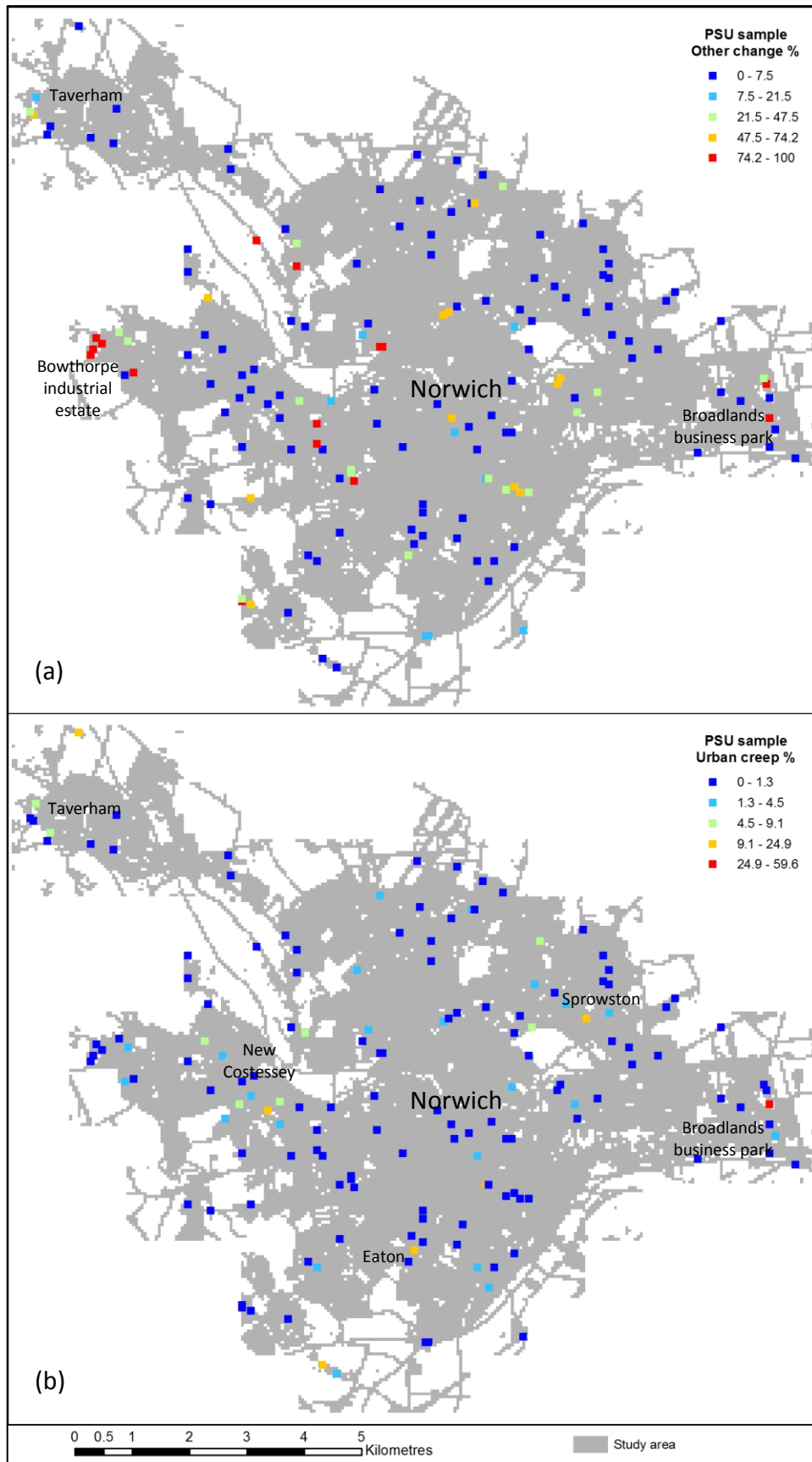


Figure 9: Norwich urban creep and other change distribution. (a) Other change distribution. (b) Urban creep distribution. Note, the scale varies between plots (a) and (b), with urban creep exhibiting less fractional change than other change.

6.3 Change statistics

Table 8 shows the summary change statistics calculated using the methods described in Chapter 5.9. The results of the intermediate calculations are recorded in the thesis Appendix. Specifically, Table A-1 shows intermediate calculations of the stratum class proportion mean, and variance, while Table A-2 shows the population class proportion mean, standard error and confidence limits.

Table 8: Norwich change estimates, 2006 to 2010.

	Standard error %	Class proportion estimate, 95% confidence limits	Class area (Ha) estimate, 95% confidence limits
No change	1.48%	94.11% \pm 2.90%	5942.07 \pm 183.14
Urban creep	0.13%	0.95% \pm 0.26%	59.71 \pm 16.36
Other change	1.48%	4.94% \pm 2.90%	312.22 \pm 183.32

The results show that urban creep accounts for 0.95% of the urban area in Norwich, which equates to an area of nearly 60 hectares. Two examples of urban creep, identified in the reference data, are shown in Figure 10 and Figure 11. *Other change* is estimated at 4.94% of the urban area, equating to 312 hectares. The combination of urban creep and *other change* combines, giving an equivalent rate of change approaching 1.5% per year.

Estimation precision is high relative to the population, but relatively low when compared to the change proportions. For example, urban creep has a standard error of 0.13%, which equates to a confidence limit (95%) of 0.26% (Table 8). At first glance, this appears to be an extremely precise estimate. However, since urban creep accounts for a relatively small fraction of Norwich, the error is high relative to the fraction of urban creep. Specifically, these confidence limits equate to 27% of the urban creep class proportion. The relative precision of *other change* is even less, with the confidence limits accounting for nearly 59% of the class proportion.

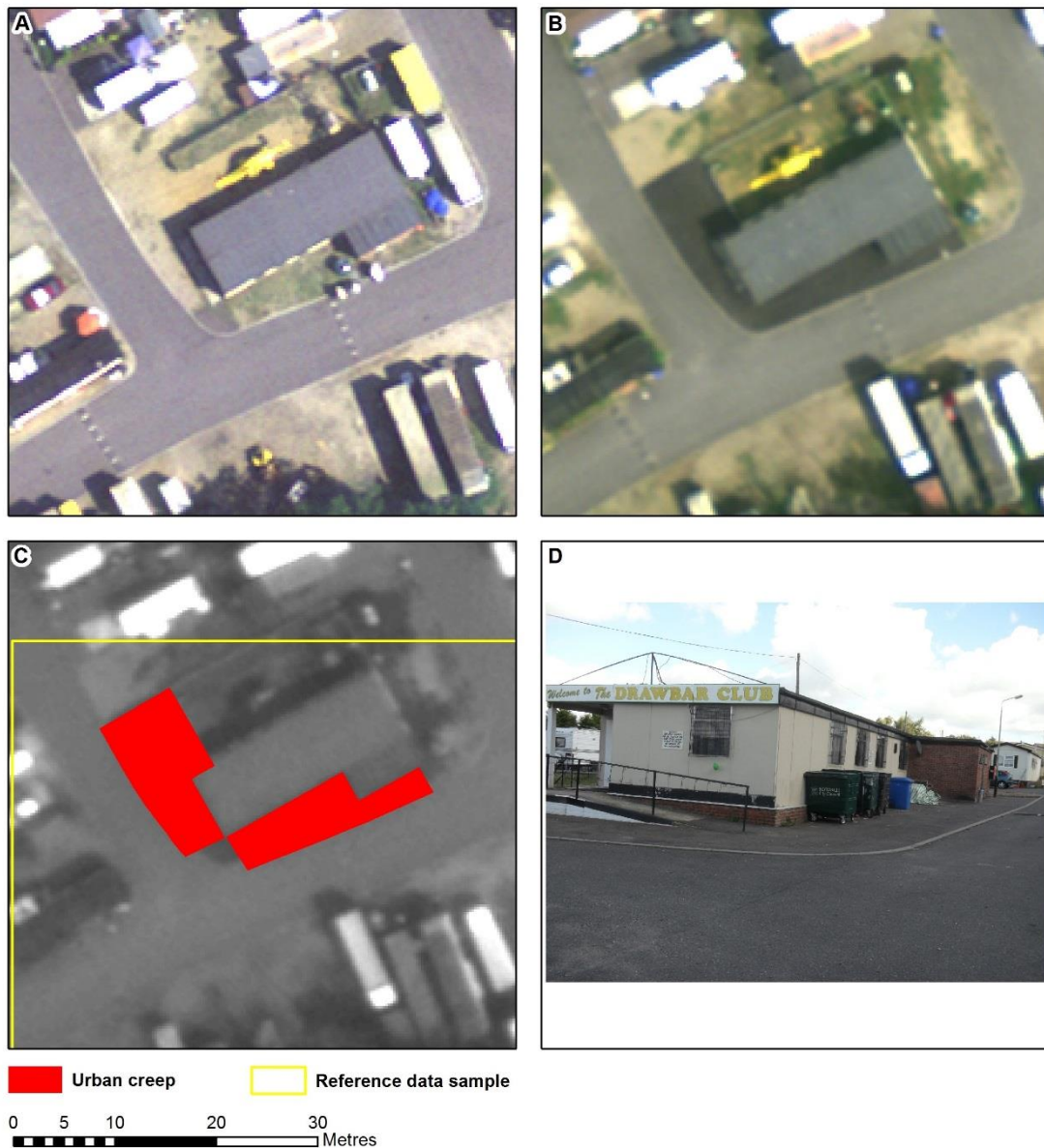


Figure 10: An example of urban creep, identified in the reference data. Much of the grass and bare earth surrounding the property has been paved over with tarmac. Panel A & B show the 2006 and 2010 aerial imagery respectively. Panel C shows the reference data overlaid over the 2010 imagery. Panel D shows the corresponding field survey photograph taken on 18th September 2012.

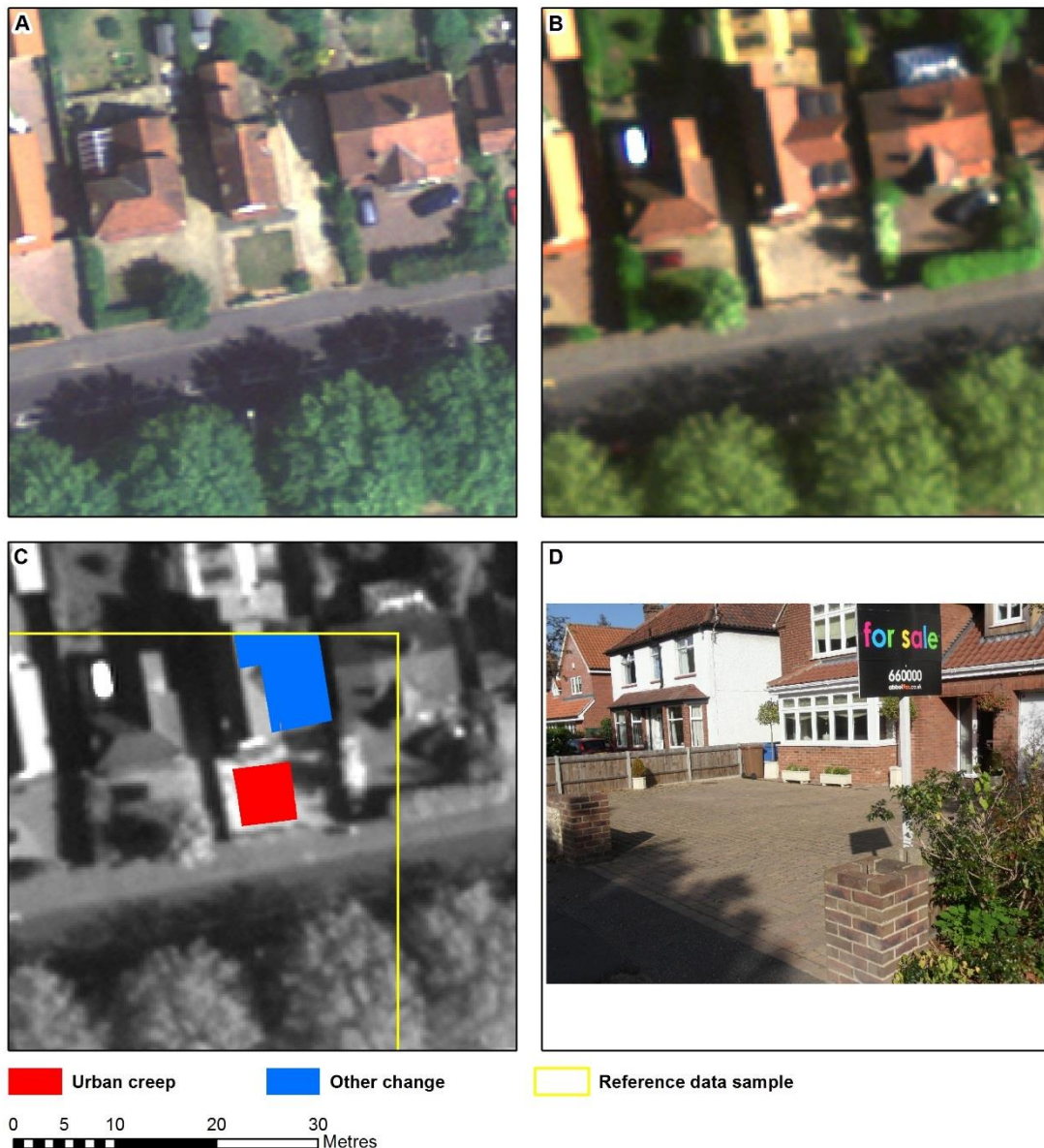


Figure 11: An example of urban creep, identified in the reference data. A house has been extended and the driveway re-modelled. Part of the front driveway has changed from grass to permeable, small block paving, and so is designated as urban creep. The house extension appears to have been built over previously impermeable paving. Consequently, the house extension, in this case, is not labelled as urban creep. Panel A & B show the 2006 and 2010 aerial imagery respectively. Panel C shows the reference data overlaid over the 2010 imagery. Panel D shows the corresponding field survey photograph taken on 17th September 2012.

6.4 Urban creep rate

The rate of urban creep was calculated to normalise the observed change by time and area for a more meaningful cross comparison against other research. Table 9 summarises the observed rate of urban creep between 2006 and 2010. Rate enumeration was made per hectare and as a fraction of the study area. Of particular

note, urban creep rate was estimated at 24.11 m²/Ha/year, equivalent to 0.24% per year.

Table 9: Norwich urban creep rate, 2006 to 2010.

Study Area (Ha)	6314
Time elapsed (years)	3.92
Urban creep rate, m² per hectare, 2006 to 2010, with 95% confidence limits	94.56 ± 25.91
Urban creep rate, m² per hectare, per year, with 95% confidence limits	24.11 ± 6.61
Urban creep rate, %, 2006 to 2010, with 95% confidence limits	0.95 ± 0.26
Urban creep rate, % per year, with 95% confidence limits	0.24 ± 0.065

6.5 Resolving image interpretation confusion

Urban creep interpretation is a complex task with scope for considerable ambiguity and uncertainty. The reference data collection was carefully undertaken, consulting multiple data sources, and where possible, clarifying ambiguities in the field. In some circumstances, accurate urban creep classification was reliant on multiple sources of evidence and was not resolvable from the aerial imagery alone. Specifically, classification confusion may occur due to tonal and textural ambiguity, occlusion, or context.

In some circumstances, contrasting surface types have indistinguishable image tone and texture. In other words, an impermeable surface change may appear identical to a permeable surface change, making an urban creep assignment impossible. For example, Figure 12 shows two new driveways observed during the fieldwork. The driveway depicted on the left of the figure is paved with impermeable block paving and is labelled as urban creep. The driveway depicted on the right of the figure is paved with permeable gravel and is *not* labelled as urban creep. However, both

driveways have nearly identical image tone and texture. This alarming similarity is an extreme case. Thankfully, field and image analysis experience facilitated accurate labelling in most cases. Tarmac (see Figure 10) and concrete is usually distinct, displaying very dark and bright albedo respectively. Although, some bright gravels may be confused with concrete. As illustrated in Figure 12, the majority of confusion occurs between small block paving and gravel because both materials are available in a variety of different materials, colours and granularity.

The rigorous methodology employed highlights that urban creep could be readily overestimated with VHR image interpretation alone. Chapter 5 discusses the rigorous reference data collection methodology applied. Specifically, the final labelling was a result of several stages of multi-source image interpretation and field survey. In most cases, the corrections applied during these iterations downgraded urban creep to *other change* because it was constructed with permeable gravel. Importantly, the use of permeable gravel as a construction material for driveways, pathways and garden features was widespread. Of the area initially labelled as urban creep, 78.62% was retained after detailed analysis. In this case, VHR image interpretation alone gave urban creep commission errors of 21.38% compared to a rigorous, multi-source approach. This result is important as it has significant implications for large-scale impermeable surface maps ‘validated’ with VHR image interpretation alone.

Chapter 4 discussed how urban creep features are susceptible to occlusion by shadow and feature lean. The reference data collection confirmed that this was indeed the case. The use of multi-angular VHR imagery, Google Street View and field survey was necessary to resolve many features. Surface clutter, such as vehicles also proved a challenge. For example, Figure 13 shows that building lean and a car masks much of the extent of a new paved parking space.

The contextual assignment of urban creep is complicated and somewhat subjective. In some circumstances, there is uncertainty assigning impermeable additions to either an existing property, for urban creep, or a new development, for *other change*. For example, Figure 14 illustrates an example where a new development has resulted in impermeable surface additions to an existing property. Urban creep was assigned in this case because the change belongs to an existing property. However, one could

argue that the change occurs outside of the initial property extent, and should be labelled as *other change*, assigned to the new development.

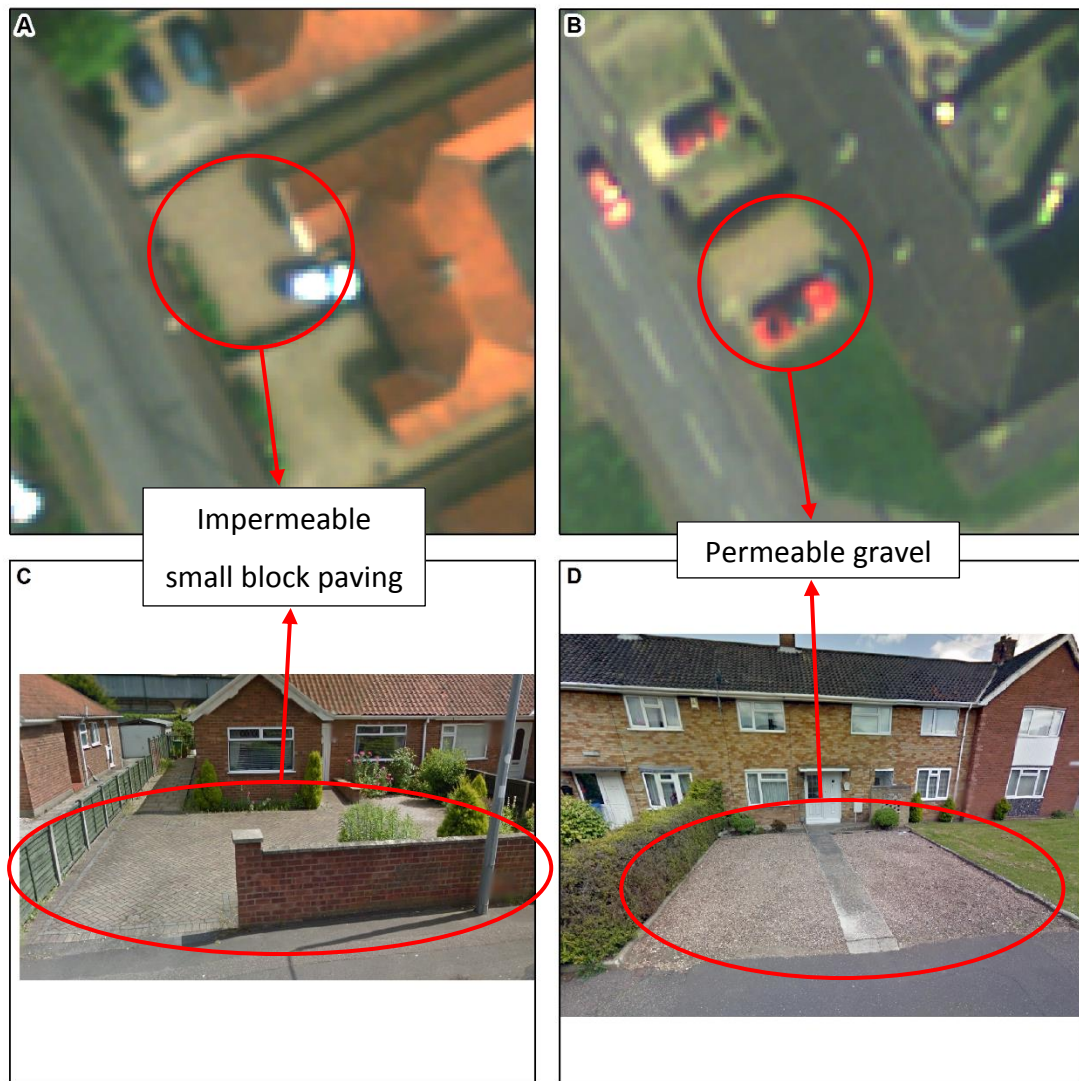


Figure 12: The visual similarity between some gravels and small block paving. Panel A & B are subsets of the aerial imagery from 2010. The two locations were field surveyed on 21st, and 18th September 2012 respectively. Panel C & D shows the corresponding Google Street view images (©Google 2016) taken in June 2012.

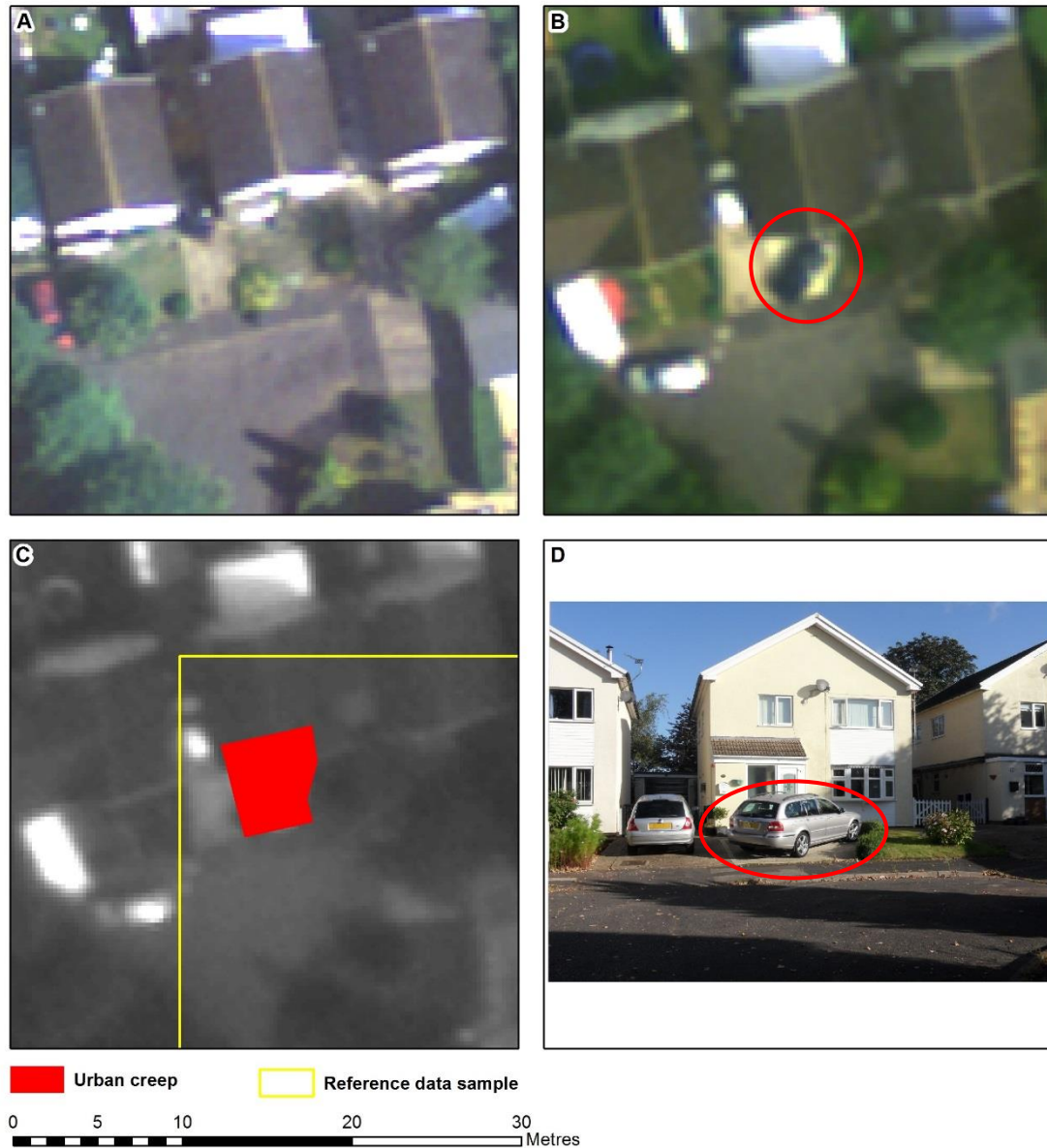


Figure 13: Urban creep obscured by a car. The lawn is paved over to create a new parking space (circled in red). The urban creep is almost entirely obscured by a car, shading and building lean. Panel A & B show the 2006 and 2010 aerial imagery respectively. Panel C shows the reference data overlaid over the 2010 imagery. Panel D shows the corresponding field survey photograph taken on 18th September 2012.

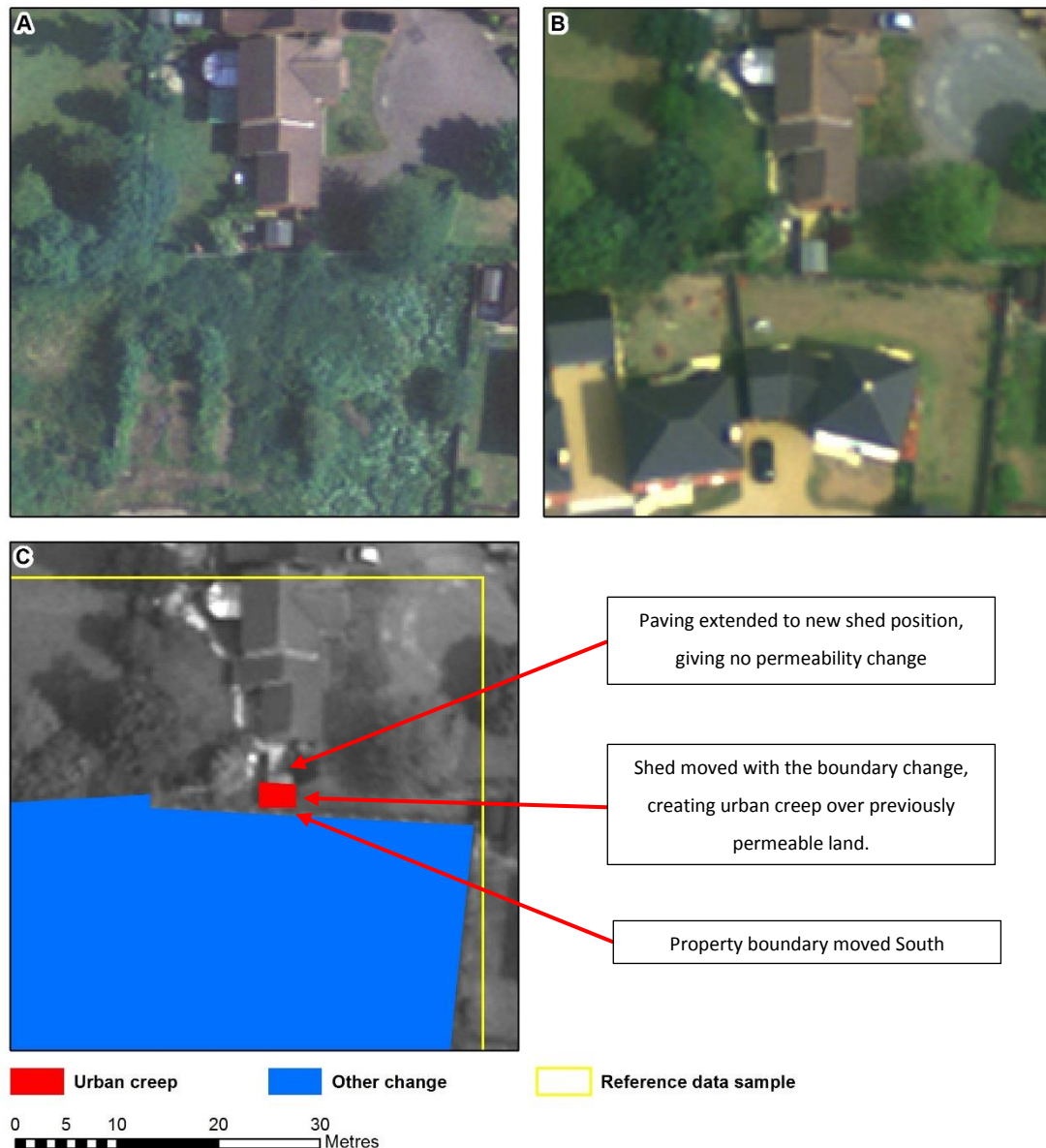


Figure 14: An example of urban creep subjectivity. A repositioned shed and additional paving create urban creep adjacent to a new development. Panel A & B show the 2006 and 2010 aerial imagery respectively. Panel C shows the reference data overlaid over the 2010 imagery.

6.6 Ongoing urban creep

Ongoing urban creep was observed during the reference data collection. This thesis is concerned with quantifying urban creep in Norwich between 2006 and 2010.

However, the 2012 field survey identified several examples where urban creep had occurred between 2010 and 2012. For instance, Figure 15 shows how urban creep observed between 2006 and 2010 is continuing, with the driveway extended between the latest 2010 image and the field survey in 2012. This circumstantial byproduct of

the field survey does indicate that, at least in parts, urban creep in Norwich is continuing, as expected.

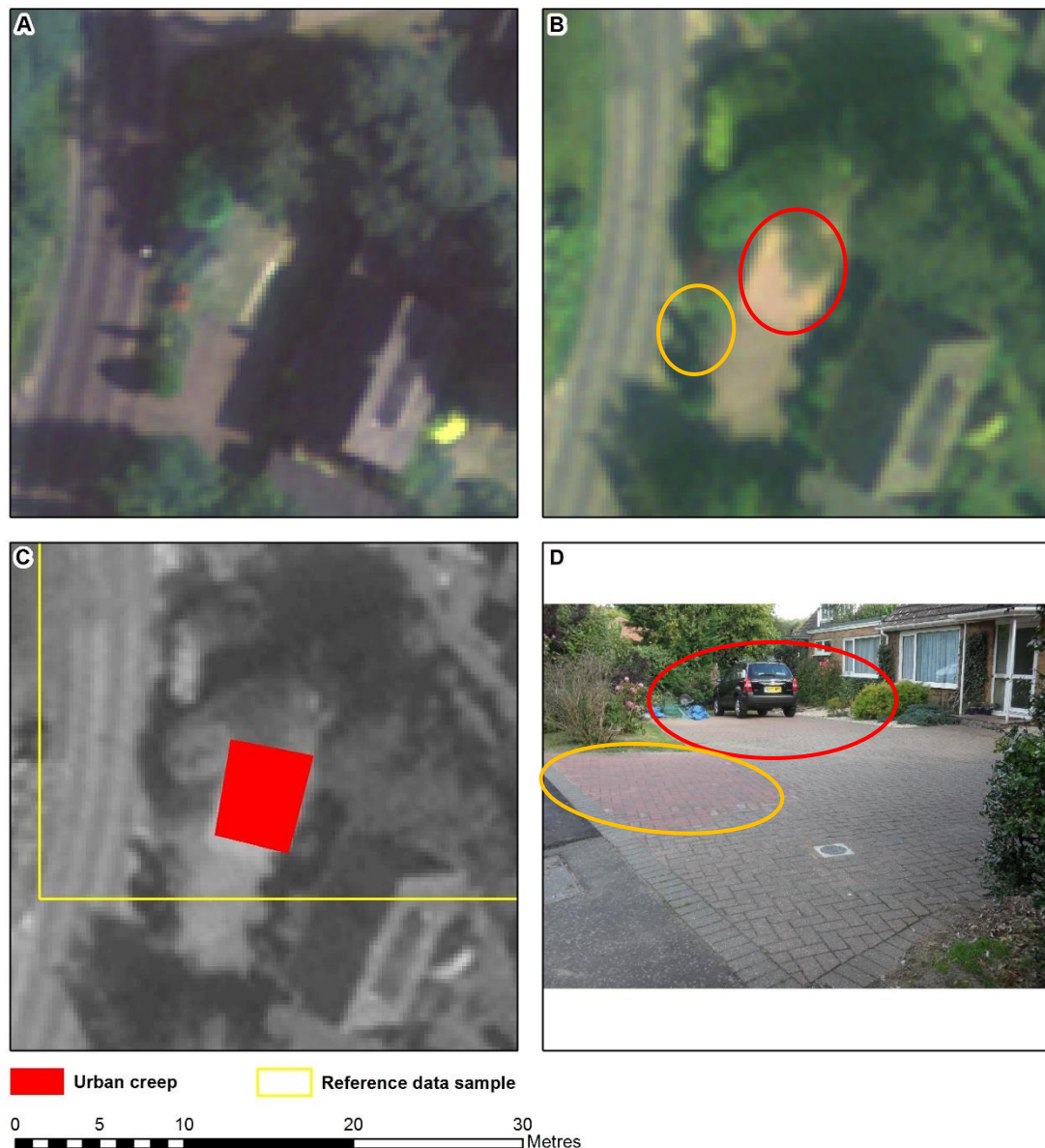


Figure 15: An illustration of evolving urban creep. The lawn is paved over to create a new parking space (circled in red). The parking space is extended (circled in orange), between 2010 and 2012. Panel A & B show the 2006 and 2010 aerial imagery respectively. Panel C shows the reference data overlaid over the 2010 imagery. Panel D shows the corresponding field survey photograph taken on 19th September 2012.

6.7 Summary

This chapter has presented the results of the reference data collection, to include land change distribution and statistics, and urban creep identification considerations. The results quantified urban creep in Norwich between 2006 and 2010 to establish a baseline to compare against subsequent remotely sensed change detection maps and

assess the changing nature of urban creep in Norwich. Also, urban creep identification difficulties were reported to include permeability confusion and contextual ambiguity.

The reference data results are summarised in Table 10. Result 1 highlights that in general urban creep is well distributed, while *other change* is more localised. Although some urban creep clustering was observed (Result 2). Results 3 & 4 indicate that Norwich is a highly dynamic city, but *other change*, dominated by large, new developments is highly variable, giving relatively low estimation precision. The rate of urban creep reported in Result 5 is the most important result from this chapter. Urban creep in Norwich has covered a considerable extent, and the reported metric forms a point of reference against other research. Results 6, 7 & 8, summarise additional points of interest identified during the data collection. In particular, permeability was easily confused in the VHR imagery and contextual assignment of urban creep was ambiguous in some cases. The importance and consequence of these results are discussed in Chapter 9.

Table 10: The reference data results summary table.

ID	Description
Result 1	In general, other change is localised, whereas urban creep is distributed more evenly throughout the city.
Result 2	Urban creep clustering is visible in New Costessey, close to Dereham Road and to a lesser extent, Sprowston.
Result 3	Urban creep accounts for 0.95%, or 59.71 Ha of the study area, while Other change accounts for 4.94%, or 312.22 Ha.
Result 4	Reference data class standard errors were calculated as 1.48%, 0.13% and 1.48% for No change, Urban creep and Other change, respectively.
Result 5	An urban creep rate of 24.11 m ² /Ha/year was observed, equivalent to 0.24% per year.
Result 6	The addition of new permeable gravel surfaces was widespread and easily confused in VHR imagery as impermeable surface creep. A more rigorous, multi-source interpretation yielded urban creep commission error rates of 21%.

ID	Description
Result 7	The contextual assignment of urban creep was complicated and subjective to a certain degree.
Result 8	Circumstantial field evidence indicates that as expected, urban creep in Norwich is continuing beyond 2010.

7 Remotely sensed change detection

7.1 Introduction

This chapter describes the change detection methods tested in respect of urban creep identification in Norwich. The methods presented include data preparation and preprocessing, scene modelling, change detection unit of analysis definition and classification and accuracy assessment. This work follows on from Chapter 6 and 7, by using the reference data in the change detection accuracy assessment.

This chapter contributes to the thesis by defining methods and experimentation which may reveal answers to the first two research questions posed in Chapter 1. These two questions concern the identification of urban creep with a direct classification and CVA classification. Therefore, the experimentation is organised into two broad groups generating direct classification results and CVA classification results. Each broad group is exhaustively explored to include many different classification permutations adding rigour to the method. Furthermore, this chapter contributes to the thesis by introducing several novel methods which give more comprehensive results and add to the change detection literature.

This chapter is presented as sixteen further sections. Firstly, the remotely sensed data and pre-processing methods are described. Then a novel scene modelling approach is presented to include viewing geometry and scene illumination simulation. Next, the change detection unit of analysis is introduced, to include image object formation and feature attribute extraction ready for change detection. An unsupervised feature selection routine is described to aid classification and result interpretation. Then, the CVA methods are presented, followed by the supervised classification approach and accuracy assessment. Lastly, the chapter is summarised to highlight the key stages and methods.

7.2 Remotely sensed data and study area review

Change detection was conducted over Norwich using multi-temporal VHR aerial imagery and DSMs from 2006 and 2010. Chapter 5 introduced the Norwich study area and monitoring timeframe. In review, the 63.14 km² study area depicts the built-up

regions of the city of Norwich in the United Kingdom, to include catchments deemed important by Norfolk County Council. The 2006 to 2010 monitoring timeframe is linked to the VHR imagery available to support this thesis. Two image datasets from 2006 and 2010, captured by the Leica ADS40 sensor at a ground resolution of 25cm, came courtesy of Airbus Defence and Space Ltd. Figure 16 displays overviews of the imagery and DSM layers with respect to the Norwich study area.

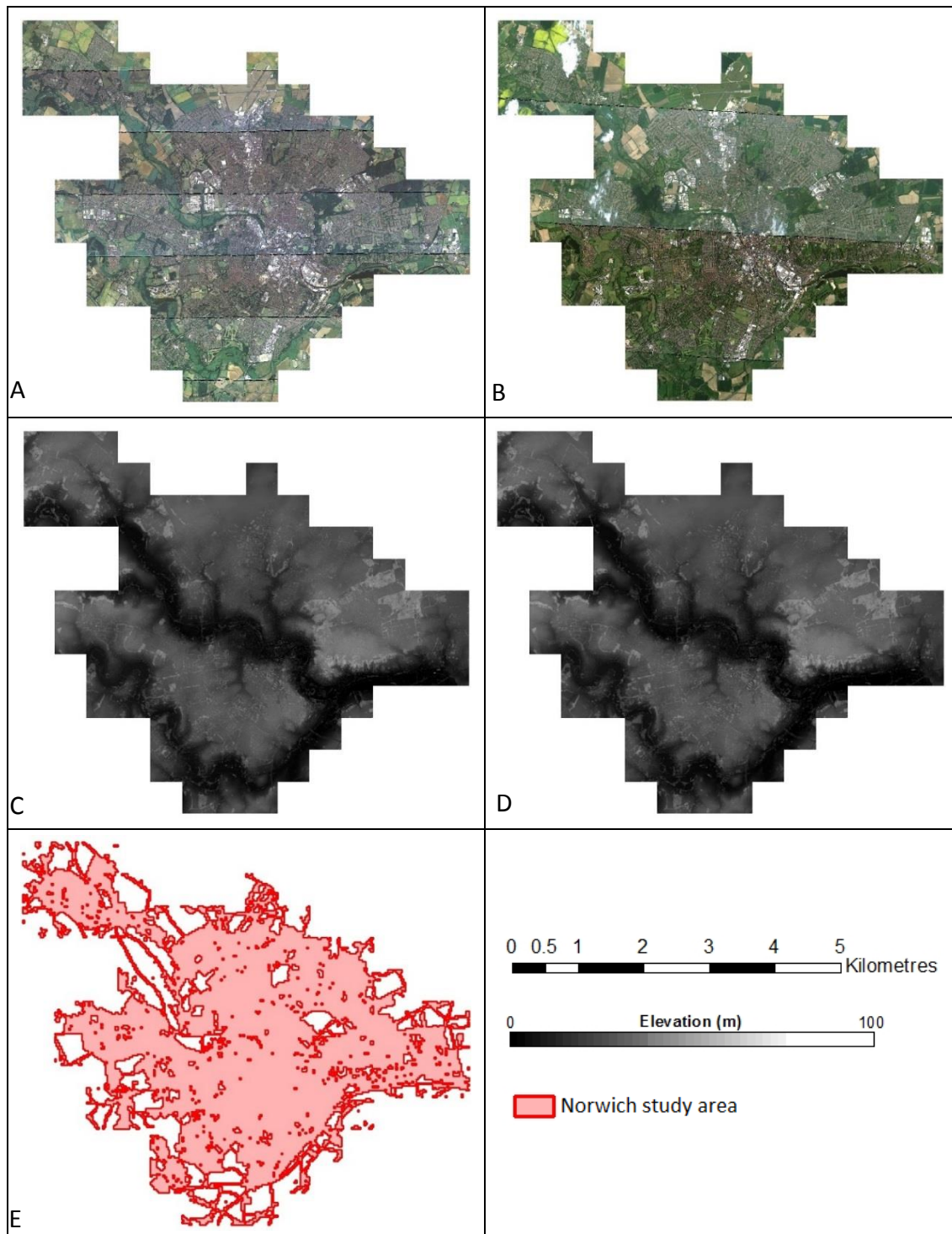


Figure 16: Remote sensing data overview. Panels A & B show the 2006 and 2010 image flight lines respectively. Panels C & D show the 2006 and 2010 DSMs respectively. Panel E shows the Norwich study area, defined by the built-up extent.

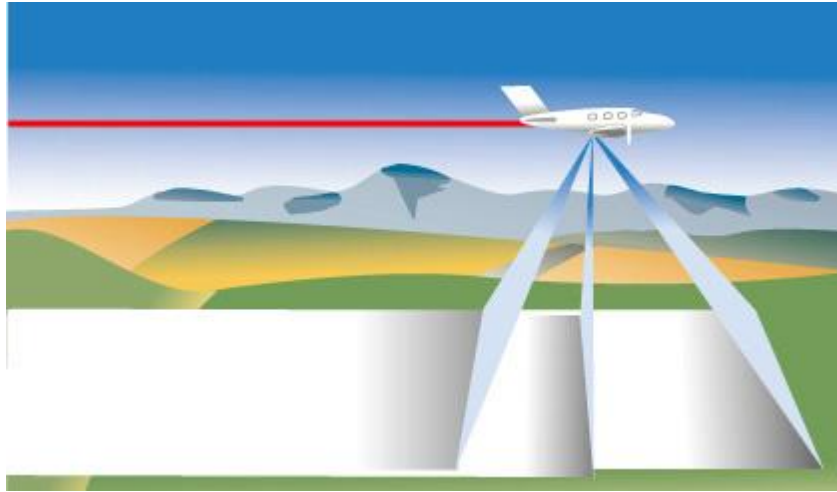


Figure 17: Push broom tri-stereo sensor image capture. In the case of the ADS40 sensor, four-band multispectral imagery is captured with a nadir view, while the panchromatic imagery is captured with forward, nadir, and backwards views. This figure is reproduced courtesy of Leica Geosystems, Geospatial Solutions Division.

The provision of simultaneous image and DSM layers was a result of the sensor configuration and data processing. The Leica ADS40 is a push broom sensor which simultaneously captures nadir and tri-stereo imagery suitable for wide-area image acquisition, and photogrammetric DSM generation (Figure 17). Push broom sensors capture 1D lines of data perpendicular to the direction of travel, and the 2D image is constructed by the forward procession of the platform. The term *flight line* is used in this chapter to describe a line of image data captured by the sensor. The ADS40 houses a nadir viewing multispectral scanner and a tri-stereo panchromatic scanner. The multi-spectral scanner acquires red, green, blue and near infra-red (NIR) images. The panchromatic scanner simultaneously acquires along-track tri-stereo panchromatic imagery, in a similar manner to imaging satellites such as ALOS PRISM or SPOT 5 (Poli & Toutin 2012). The nadir viewing multispectral scanner was the source of the 25cm ortho-rectified imagery used in this research. The tri-stereo panchromatic scanner imagery was used to produce a photogrammetric DSM with a resolution of 2m. Airbus Defence and Space Ltd. conducted all of the aerial data pre-processing, achieving planar and vertical RMSE accuracies of $\leq 1.5\text{m}$. Table 11 summarises the aerial image data layers produced for each flight line.

Table 11: Aerial image data layer summary.

Data layer	Resolution	Planar accuracy	Vertical accuracy
Ortho-rectified multispectral imagery	0.25m	≤1.5m RMSE	N/A
DSM	2m	≤1.5m RMSE	≤1.5m RMSE

The imagery was acquired using thirteen flight lines to achieve two cloud-free coverages of the city. The 2006 imagery is made up of seven flight lines captured on 2nd and 3rd July 2006, at an altitude of approximately 2400m. All of the 2006 imagery is usable because there was no cloud cover. Figure 18A and Table 12 respectively show the extent and characteristics of the 2006 flight lines. The 2010 imagery is made up of six flight lines captured on 19th & 21st May and 3rd July 2010, at an altitude of approximately 4800m. Figure 18B, C, and Table 13 show the extent and characteristics of the 2010 flight lines. The 2010 imagery was acquired in two batches because of cloud cover. In 2010 the ADS40 super-resolution imaging mode was engaged to reduce image acquisition costs. The super-resolution imaging mode reconstructs high-resolution images from a pair of images at half resolution, staggered by half a pixel. This facilitates higher flying heights and wider swath widths. The super-resolution methodology employed by Airbus Defence and Space Ltd. is described in detail by Reulke et al. (2006).

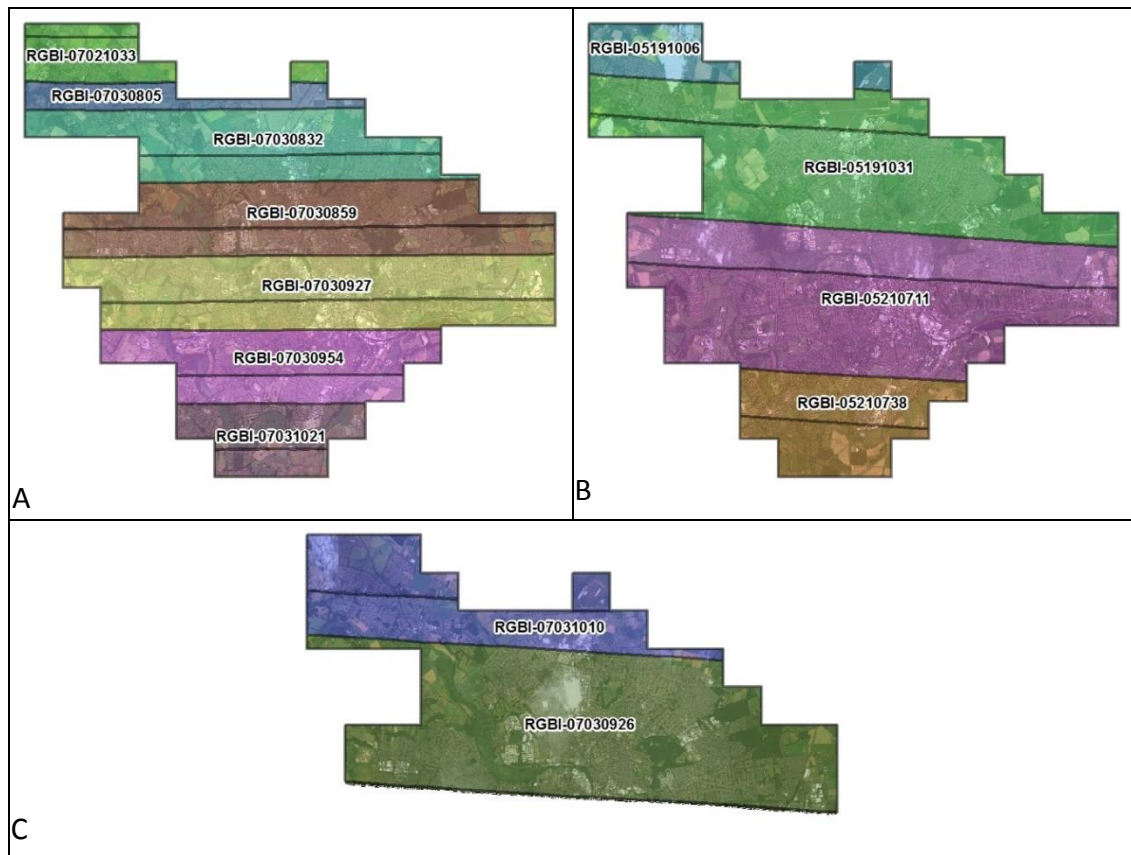


Figure 18: The aerial image flight lines. In total, thirteen aerial image flight lines were required to achieve complete multi-temporal image coverage. Panel A shows the extent of the seven flight lines from July 2006. Panel B shows the four initial flight lines from May 2010, noting some cloud cover is present. Panel C shows an additional two flight lines from July 2010 captured to infill cloudy regions from May 2010. The July 2010 flight lines also contain cloud. Therefore cloud free coverage was achieved with a mixture of May and June. In all panels, the flight line boundary extents are shown as a black outline, where we can see that significant lateral overlap exists.

Table 12: 2006 image flight lines.

Flight line	Approximate flying height, metres	Imaging mode	Ground resolution, metres	Acquisition date	Acquisition time	Area km ²	Image size, megapixels	Image size, Gigabytes
RGBI-07021033	2400	Tri-stereo panchromatic, four band at nadir	0.25	02/07/2006	10:33:00	5.73	92	0.68
RGBI-07030805				03/07/2006	08:05:00	19.62	314	2.34
RGBI-07030832					08:32:00	28.63	458	3.41
RGBI-07030859					08:59:00	37.35	598	4.45
RGBI-07030927					09:27:00	33.92	543	4.04
RGBI-07030954					09:54:00	19.80	317	2.36
RGBI-07031021					10:21:00	7.67	123	0.91

Table 13: 2010 image flight lines.

Flight line	Approximate flying height, metres	Imaging mode	Ground resolution, metres	Acquisition date	Acquisition time	Area km ²	Image size, megapixels	Image size, Gigabytes
RGBI-05191006	4800	Super-resolution mode, tri-stereo panchromatic, four band at nadir	0.25	19/05/2010	10:06:00	14.02	224	1.67
RGBI-05191031					10:31:00	50.09	801	5.97
RGBI-05210711				21/05/2010	07:11:00	49.47	791	5.90
RGBI-05210738					07:38:00	12.09	193	1.44
RGBI-07030926				03/07/2010	09:26:00	52.22	836	6.23
RGBI-07031010					10:10:00	16.51	264	1.97

7.3 Data quality

The remotely sensed data quality was evaluated prior to the change detection to assess its suitability. A qualitative assessment of the data was conducted to ascertain its suitability to the application and to support the interpretation of the change detection results. Overall, it was concluded that the data is suitable for this research and urban creep features are meaningfully resolved. Figure 19 shows an example of urban creep in the multi-temporal data stack. However, the data is far from perfect, and the following observations must be considered when interpreting the change detection results:

2010 image contrast. While both datasets are imaged and processed to a ground resolution of 25cm, the 2010 imagery is not as sharp as the 2006 imagery. This is interpreted to be largely due to the effect of the super-resolution imaging mode processing. However, the atmospheric conditions may also be a contributing factor because of the cloud present in all 2010 imaging campaigns.

Overall DSM quality. It is important that the resolution and accuracy of the available 2m photogrammetric DSM is not confused with the resolution and accuracy of a Lidar DSM. Airbus Defence and Space Ltd. designed the processing and editing with the objective of producing a DSM with an RMSE $\leq 1.5\text{m}$, suitable for viewing and analysis at cartographic scales no larger than 1:5000. Generally, the DSM quality is good and within the defined accuracy tolerance. However it is clearly not as accurate or detailed as a typical airborne lidar DSM.

2010 DSM quality. Both the 2006 and 2010 DSM are within the defined accuracy tolerance of $\leq 1.5\text{m}$ RMSE. However, a qualitative evaluation of both datasets indicates that the 2010 DSM is less accurate than the 2006 DSM. The lower two panels of Figure 19 show that when compared to the 2006 DSM, the 2010 DSM appears less accurate, heavily generalising some smaller buildings. It is assumed that the difference is because of the super-resolution imaging mode and, to a lesser extent, increased atmospheric haze.

Sensor and sun geometry. Sensor and illumination geometry varies between all flight line image pairs. This is to be expected with VHR image change detection. While, extreme differences are not observed, the data contains a mixture of differences and serves as a good testbed for illumination and viewing geometry invariance methods testing.

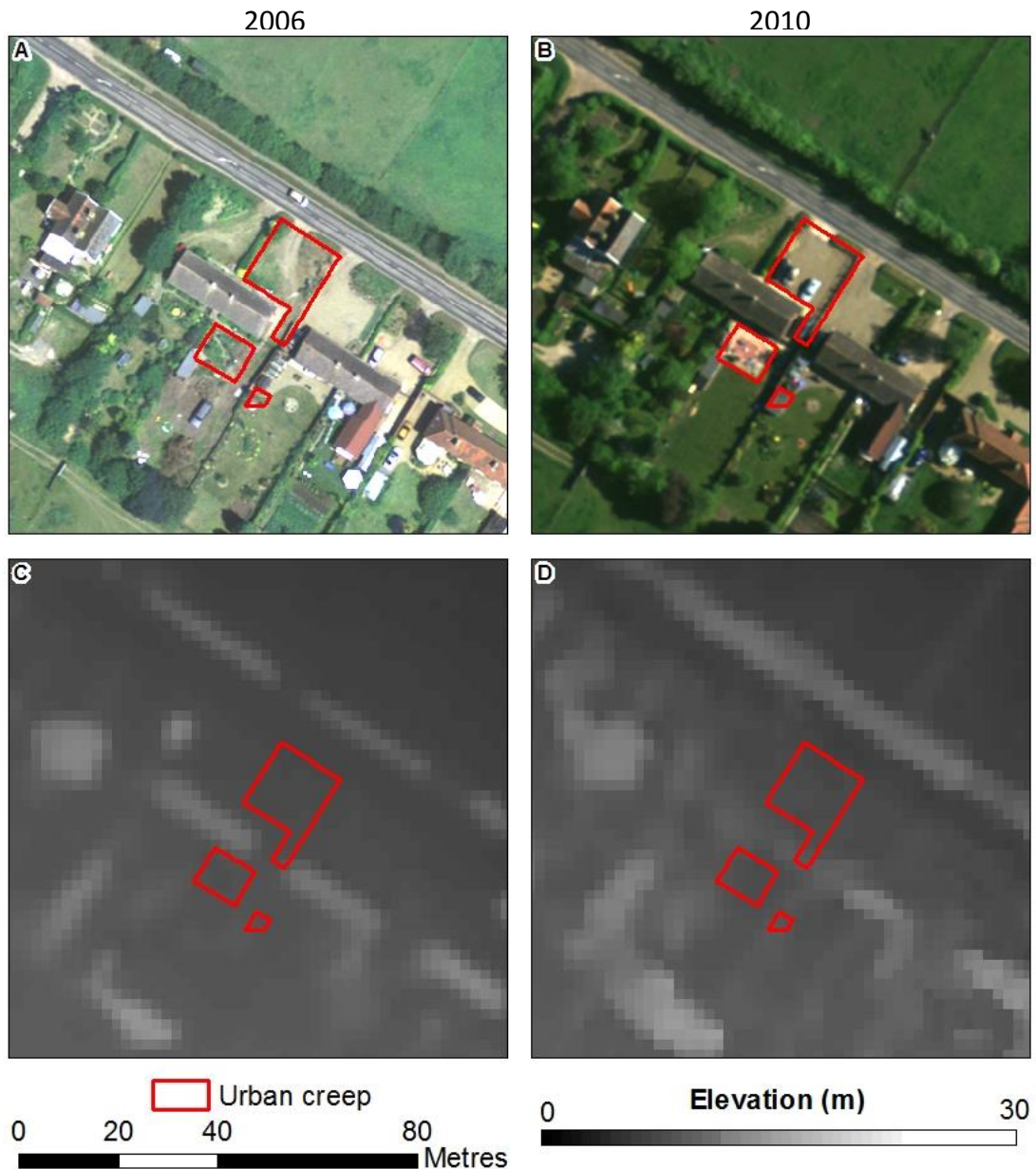


Figure 19: An illustration of the data resolution and quality. Panel A & B show the 25cm imagery for 2006 and 2010 respectively. We can see that although the image quality in both is fair to good, 2010 is not as sharp as 2006. Panel C & D show the 2m DSM. We can see that at this scale, elevation accuracy is variable, especially in the 2010 data, with some of the buildings and other details heavily generalised.

7.4 Data organisation

The data was organised into fourteen work blocks to maintain temporal consistency and facilitate image processing. The available images were organised into work blocks containing bi-temporal image pairs with no cloud cover. Therefore, each work block observes a sub-set of the Norwich study area over a specific monitoring period, with consistent acquisition conditions and image radiometry. Furthermore, because of the

size of the images, image and GIS processing was only possible by work block, not the entire, mosaiced dataset. Figure 20 shows the extent of the fourteen work blocks, while Table 14 summarises the monitoring timeframe for each. The time difference between work blocks ranges from 3.8795 to 4.0055 years. For simplicity, the area weighted mean time difference is calculated as 3.92 years.

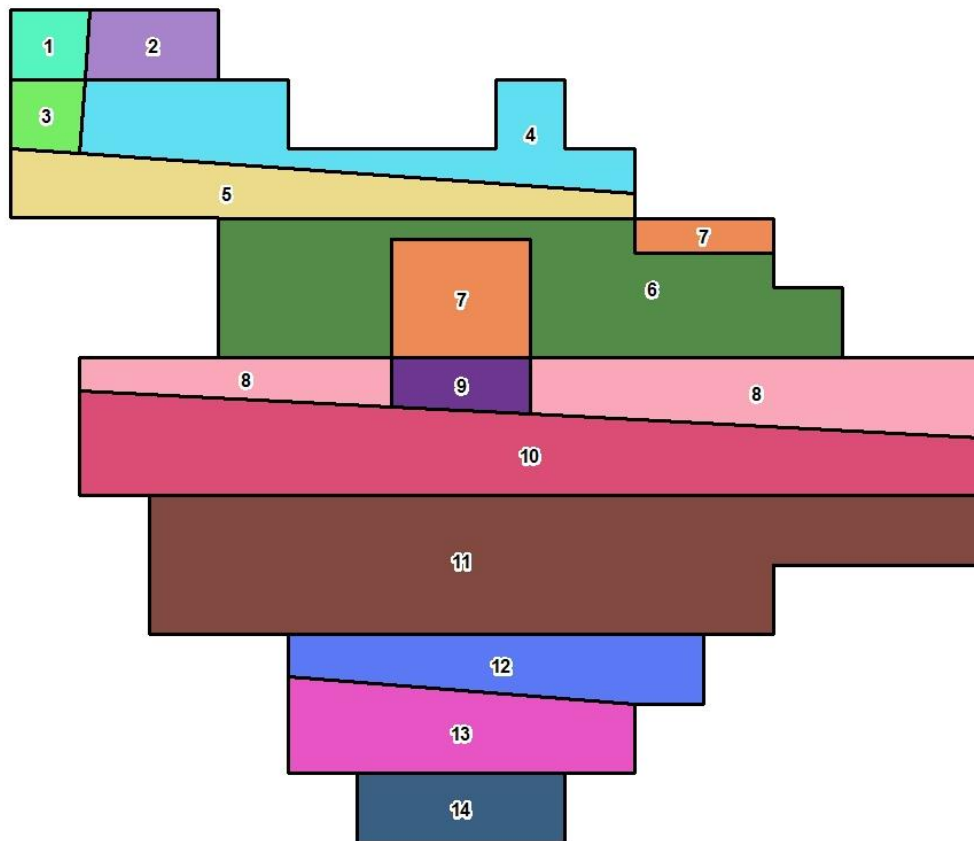


Figure 20: The extent of the fourteen work blocks.

Table 14: Work block monitoring timeframe. Note, the fractional time differences in years respect the difference in image acquisition down to the minute.

Work block number	2006		2010		Time difference, years
	Flight line	Date	Flight line	Date	
1	RGBI-07021033	02/07/2006	RGBI-05191006	19/05/2010	3.8822
2			RGBI-07031010	03/07/2010	4.0055
3	RGBI-07030805	03/07/2006	RGBI-05191006	19/05/2010	3.8795
4			RGBI-07030926	03/07/2010	4.0027
5			RGBI-07031010		
6	RGBI-07030832		RGBI-05191031	19/05/2010	3.8795
7			RGBI-07030926	03/07/2010	4.0027
8	RGBI-07030859		RGBI-05191031	19/05/2010	3.8795
9			RGBI-05210711	21/05/2010	3.8849
10			RGBI-07030926	03/07/2010	4.0027
11	RGBI-07030927		RGBI-05210711	21/05/2010	3.8849
12	RGBI-07030954				
13					
14	RGBI-07031021		RGBI-05210738		

7.5 Spatial co-registration

The 2006 imagery was spatially co-registered to the 2010 imagery to facilitate meaningful change detection. While all of the available images were already ortho-rectified, a qualitative evaluation indicated that a further stage of spatial registration was required. Klaric et al. (2013) employed a similar two-stage methodology in their

research, specifically, ortho-rectification followed by spatial co-registration. A further stage of image registrations was almost inevitable, given that literature recommendations of co-registration accuracy range from sub-pixel (Kuntz et al. 2011) to as little as 0.2 pixels (Dai & Khorram 1998). Table 15 summarises the errors observed between the images sets and sixteen ground control points provided by Airbus Defence and Space Ltd. These results indicate, the positional accuracy of both image sets is within the defined ortho-rectification tolerance (Table 11), and that the 2010 imagery was the most accurate. However, it confirms the spatial registration is not suitable for change detection. Therefore, the 2006 imagery was registered to the 2010 imagery using a 1st order polynomial transformation. Table 16 summarises the co-registration of the 2006 flight lines. Between eight and twelve tie points were used for the co-registration, yielding errors ranging from 0.519 to 0.855 pixels.

Table 15: Data positional error summary. Reports the error observed between the multi-temporal aerial data and sixteen ground control points.

Image set	RMSE X	RMSE Y	RMSE
2006	0.35m	0.91m	0.93m
2010	0.32m	0.26m	0.62m

Table 16: Summary of the spatial calibration.

Flight line	Number of tie points	RMSE (pixels)	RMSE (metres)	Registration type	Resample type
RGBI-07021033	9	0.79	0.1975	1st order polynomial	Bilinear
RGBI-07030805	10	0.855	0.21375		
RGBI-07030832	10	0.546	0.1365		
RGBI-07030859	12	0.519	0.12975		
RGBI-07030927	11	0.736	0.184		
RGBI-07030954	9	0.835	0.20875		
RGBI-07031021	8	0.726	0.1815		

7.6 Radiometric co-registration

A relative radiometric calibration was undertaken using pseudo-invariant features (PIFs) to ‘match’ the 2010 imagery to the 2006 imagery. PIFs are stable areas of the image that have not changed, and are commonly a means of achieving relative radiometric calibrations (Davis 2011; Schroeder et al. 2006; Canty et al. 2004).

Unchanged pixels are used to calculate a radiometric transformation between images, which is applied to invoke a relative calibration. Therefore, after a PIF calibration, unchanged pixels should have near-identical pixel intensities, and meaningful differences can be observed.

A subset of the PIF identification method employed by Carvalho Júnior et al. (2013) is used for the calibration. Carvalho Júnior et al.'s (2013) method identifies PIFs based on the sequential application of three spectral similarity metrics and two statistical filtering techniques. In summary, their method selects PIFs for those pixels that satisfy all five approaches. For simplicity, the method employed here selects only two spectral similarity metrics, SAM and SCM as the means of PIF identification. Spectral Angle Mapper (SAM) and Spectral Correlation Mapper (SCM) are remote sensing terms for cosine similarity, and Pearson's r respectively. They are both gain invariant similarity metrics useful for matching spectral signatures under varying illumination conditions.

To simplify the notation, 2006 is referred to as T0 and 2010 as T1. This notation relates to T0 as the ‘baseline’ image and T1 as the next image in a given time series up to Tn.

For the i^{th} band in the image at T0 and T1, SAM and SCM are given by:

$$SAM = \cos^{-1} \left[\frac{\sum_{i=1}^n T0_i T1_i}{\sqrt{\sum_{i=1}^n T0_i^2 \sum_{i=1}^n T1_i^2}} \right], \quad SAM \in [0, 180] \quad \text{Equation 13}$$

$$SCM = \frac{\sum_{i=1}^n (T0_i - \bar{T0})(T1_i - \bar{T1})}{\sqrt{\sum_{i=1}^n (T0_i - \bar{T0})^2 \sum_{i=1}^n (T1_i - \bar{T1})^2}}, \quad SCM \in [-1, 1] \quad \text{Equation 14}$$

Therefore, PIF pixels are designated as those with SAM values close to 0, where there is little or no angle between the vectors; and SCM values close to 1, where the vectors have a strong positive correlation.

The PIF calibration method employed in this thesis goes beyond the literature by enforcing spatial, and radiometric distribution. This modification ensures that PIFs are well distributed across the image extent and radiometric scale, which will, in turn, facilitate the computation of a robust radiometric transformation. The enforcement was deemed necessary after practical experience yielded PIFs that tended to cluster in bright and dark saturated regions only, such as deep water and bright roofs, yielding an unsatisfactory regression model. Radiometric enforcement was applied by binning image albedo and searching for at least 50 PIFs in each bin. Spatial enforcement was achieved by splitting the input images into tiles, applying the radiometric enforcement PIF identification to each tile and then aggregating the results. To minimise radiometric noise, only PIFs of more than two pixels in size were retained. The PIF identification algorithm was implemented in eCognition 9.0. Figure 21 depicts a summary flow diagram of the algorithm, while more detailed pseudo code of the algorithm is provided in the Appendix, Figure A-3. Figure 22 shows an illustration of the PIF distribution for a subset of work block 1. Note that the PIFs primarily congregate over stable urban surfaces but also exist over other stable land covers.

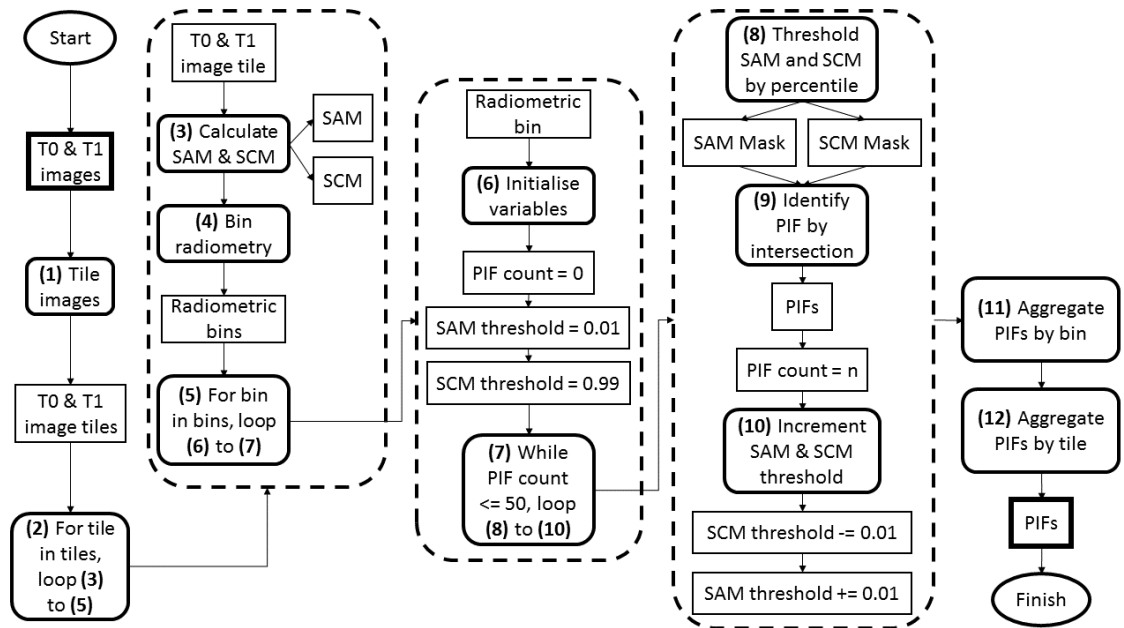


Figure 21: The implemented PIF identification algorithm.

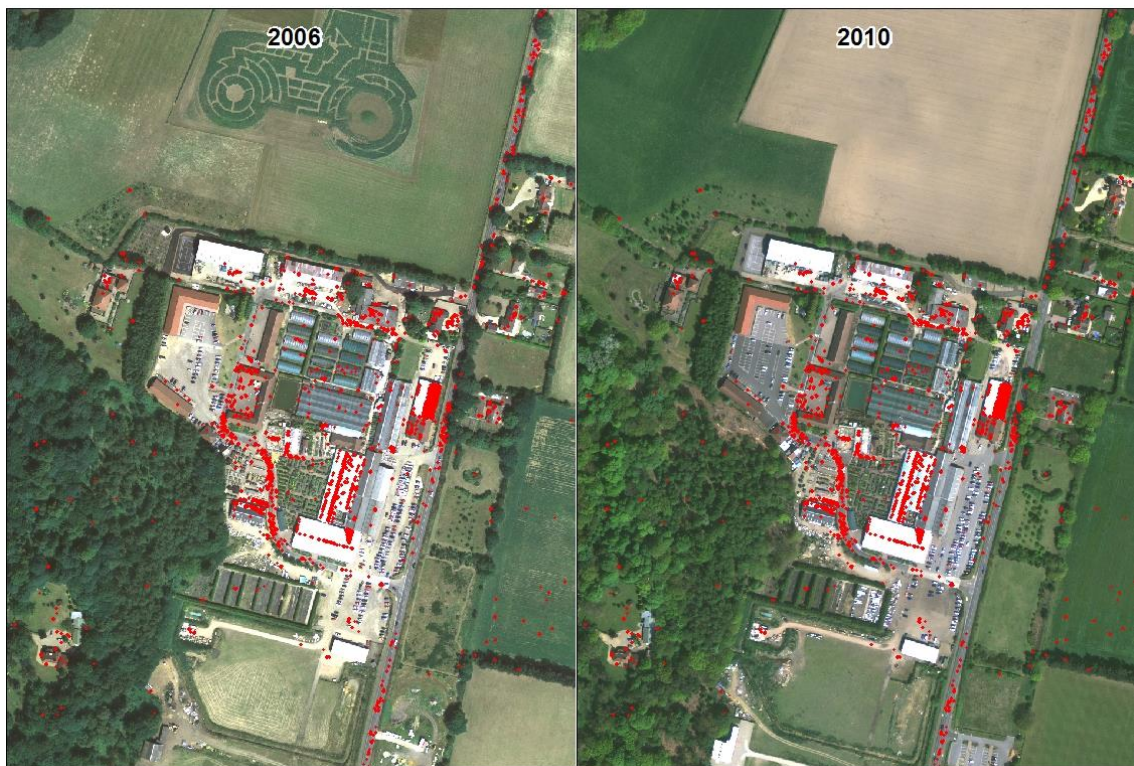


Figure 22: Example output from the PIF identification algorithm. PIFs are highlighted in red. The left-hand panel contains the 2006 imagery, while the 2010 imagery is on the right.

The PIFs were used to radiometrically co-register the imagery, by work block and by image band. Specifically, a linear regression model was calculated between T0 and T1, for each of the four image bands, and for each of the fourteen work blocks. In each case the linear regression model is defined by:

$$T1 = mT0 + c \quad \text{Equation 15}$$

$$T0 = \frac{T1 - c}{m} \quad \text{Equation 16}$$

Where m is the regression slope gradient, and c is the intercept. Therefore, applying Equation 16 removes the radiometric offset for the particular image band pair. Figure 23 shows example scatter plots before and after the PIF calibration. Before the calibration, there is an intercept offset and slope difference between T0 and T1. After these offsets are removed a perfect 1:1 radiometric relationship is achieved. In other words, pixel intensity remains near identical over PIF features.

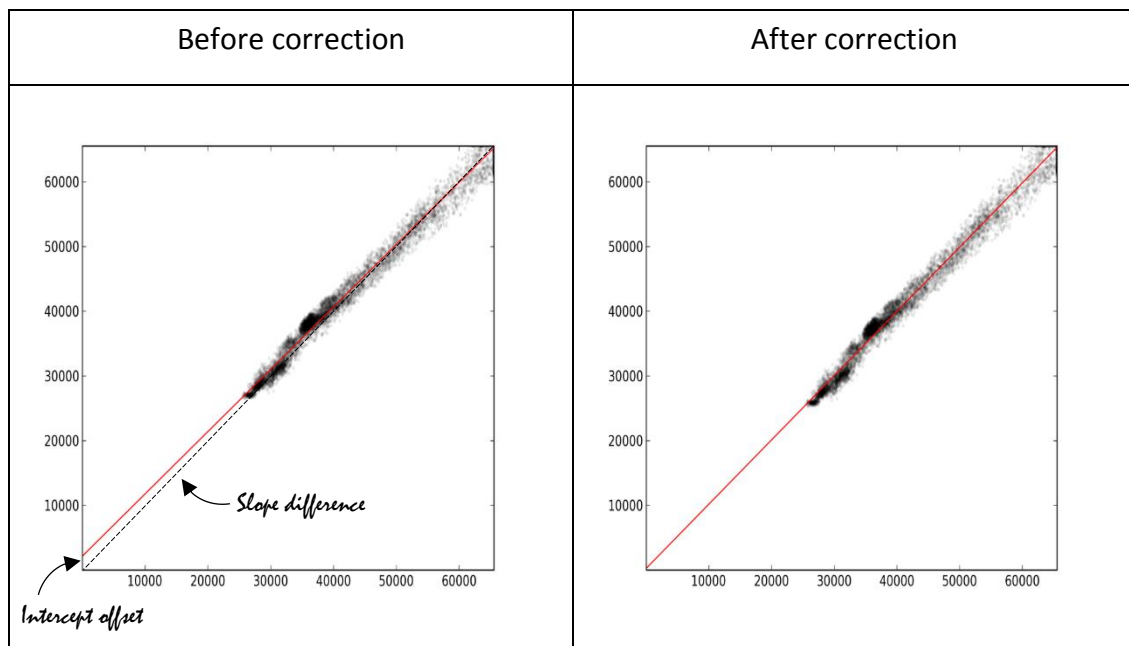


Figure 23: PIF radiometry before and after the radiometric transformation. Both plots show the results of the blue band in work block 11. The PIFs are plotted as black dots, and the linear regression line is plotted in red. The left-hand panel shows the 2010 blue band plotted against the 2006 blue band before the correction is applied. The right-hand panel shows the 2010 blue band plotted against the 2006 blue band after the correction has been applied. Note, that after the correction a 1:1 radiometric relationship is maintained.

7.7 Modelling viewing geometry

This research models hidden (occluded) surfaces in the available imagery, as a means of suppressing change detection false alarms induced by differences in viewing geometry. Hidden surfaces are areas on the ground not visible to the sensor due to occlusions. In the remotely sensed imagery, they are visible as elevated features leaning over and hiding the ground surface. In practical terms, this modelling aims to highlight the parts of building roofs and tree canopies that hide the ground surface; and may otherwise be confused with real land cover change. When treated rigorously, the identification of hidden surfaces is a photogrammetric problem, solved by precise sensor modelling and stereo image measurement. However, this photogrammetric approach can be approximated with simple desktop GIS tools.

A novel GIS-based method of hidden surface identification method is implemented by approximately simulating a push broom sensor. A GIS viewshed analysis and a concurrent DSM is used to simulate the optics of a push broom sensor and depict its ground visibility along a flight line. The ADS40, like other push broom sensors, works by imaging a procession of pixel strips perpendicular to the flight direction. Therefore, each row of the acquired image strip inherits different imaging geometry. If we consider each row a field of view (FOV) from an observation point, we can simulate each with a GIS viewshed analysis. The viewshed created for each observation will depict the visibility of the sensor in that position, which conversely, can map the hidden surfaces. If we then mosaic the viewsheds for each observation together, we obtain a map of hidden surfaces for the entire flight line. The algorithm was implemented in Python 2.7.8 using the ArcGIS 10.1 viewshed tool. Figure 24 outlines the hidden surface detection algorithm as a flow diagram, while more detailed pseudo code is provided in the thesis Appendix, Figure A-4. The 3D flight paths of the sensor were estimated from the extent of each flight line, and the approximate flying heights described in Table 12 and Table 13. Figure 25 shows example hidden surface maps from 2006 and 2010 over Norwich city centre. In this example, we can see that the hidden surface map does correlate with the feature lean present in the source imagery. However, the scale of hidden surface maps is strictly limited by the scale of the source DSM. Consequently, only large, prominent hidden surfaces are depicted.

The hidden surface maps were used directly in the change detection classifications as a means of reducing false alarms. The method described here represents a novel and simplified simulation of viewing geometry for change detection.

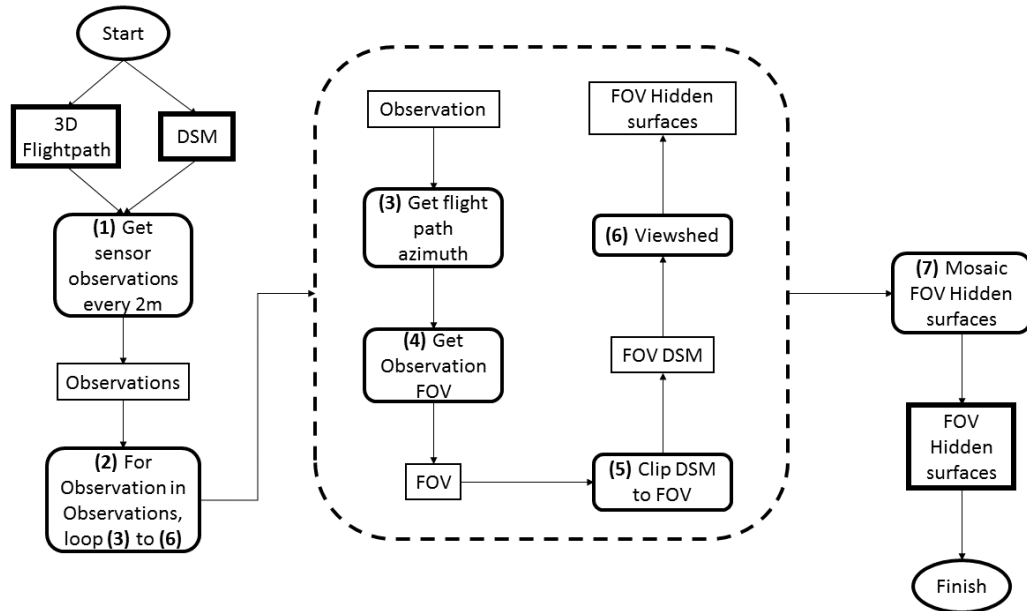


Figure 24: Hidden surface detection algorithm outline.

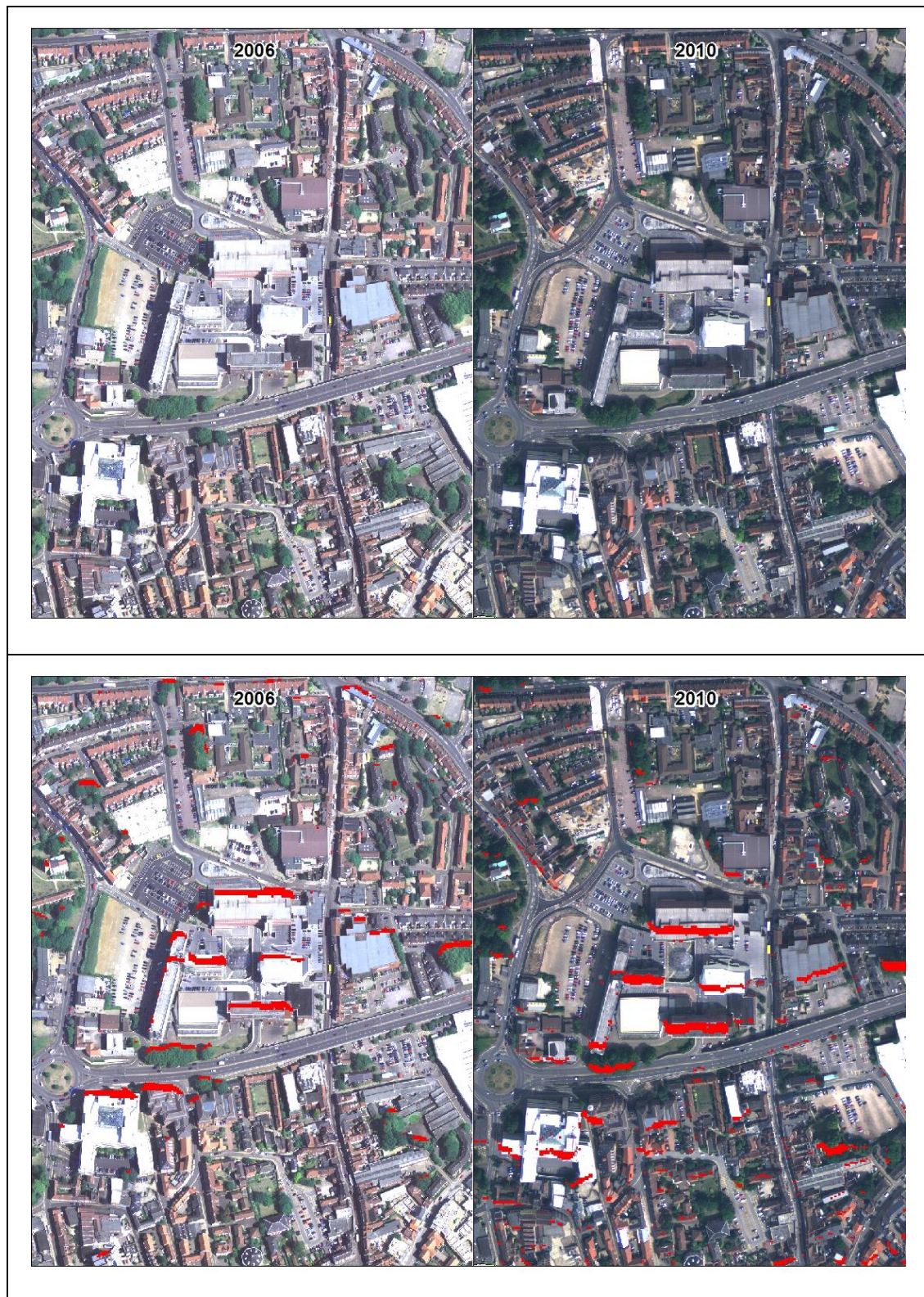


Figure 25: Hidden surface identification results. The top row shows the multi-temporal imagery over Norwich city centre. The bottom row shows the same imagery with the hidden surfaces highlighted in red.

7.8 Modelling scene illumination

The lighting conditions at the time of image acquisition were modelled as a novel addition to the change detection process. Scene illumination was simulated to create shading maps depicting the extent of shadows and bi-directional reflectance induced shading. The shading maps were used as a direct input to the change detection process. Specifically, the supervised classifier would have a clear opportunity to learn the relationship between shading patterns and false alarms.

Scene illumination change was simulated using a multi-temporal GIS hill shading procedure. Since each image is accompanied by a DSM and the sun position can be reasonably estimated, shading patterns are readily simulated with common GIS hill-shading tools. Solar azimuth and elevation were estimated for each image based on the acquisition time, date and location. The NOAA online solar calculator (National Oceanic & Atmospheric Administration 2016) was used to calculate the sun position. Table 17 catalogues the sun position for each of the fourteen flight lines. The hill-shading tool in ArcGIS 10.1 was used to create shading maps for all of the flight lines. Crucially, the hill shading tool considers bi-directional reflectance shading and cast shadows. The output 8-bit raster images depict shadows with a digital number of 0, while shaded areas range from 1 to 255 as the solar incidence angle decreases. Figure 26 illustrates the results of the hill-shading over Norwich city centre. The main shading patterns are honoured. However, the scale of the shadow detection is limited by the resolution and accuracy of the source DSM. Although, hill-shading is a well-established and readily available technique, the application to change detection workflows is a novel application.

Table 17: Flight line sun positioning.

Image	Year	Month	Day	Image centre (lat long)	Approx. time	Solar azimuth	Solar elevation
RGBI-07021033	2006	July	2	52.690 1.242	11:30	141.63	55.85
RGBI-07030805	2006	July	3	52.679 1.262	09:30	106.19	40.69
RGBI-07030832	2006	July	3	52.661 1.284	09:30	106.19	40.71
RGBI-07030859	2006	July	3	52.644 1.293	10:15	117.59	47.04
RGBI-07030927	2006	July	3	52.627 1.291	10:45	126.3	50.91
RGBI-07030954	2006	July	3	52.610 1.275	11:15	136.15	54.34
RGBI-07031021	2006	July	3	52.598 1.272	11:30	141.58	55.84
RGBI-05191006	2010	May	19	52.684 1.246	11:30	146.58	53.53
RGBI-05191031	2010	May	19	52.658 1.285	11:45	152.3	54.72
RGBI-05210711	2010	May	19	52.627 1.293	08:15	94.03	28.24
RGBI-05210738	2010	May	19	52.602 1.278	08:45	100.29	32.75
RGBI-07030926	2010	July	3	52.655 1.293	10:45	126.33	50.89
RGBI-07031010	2010	July	3	52.682 1.256	11:45	147.4	57.08

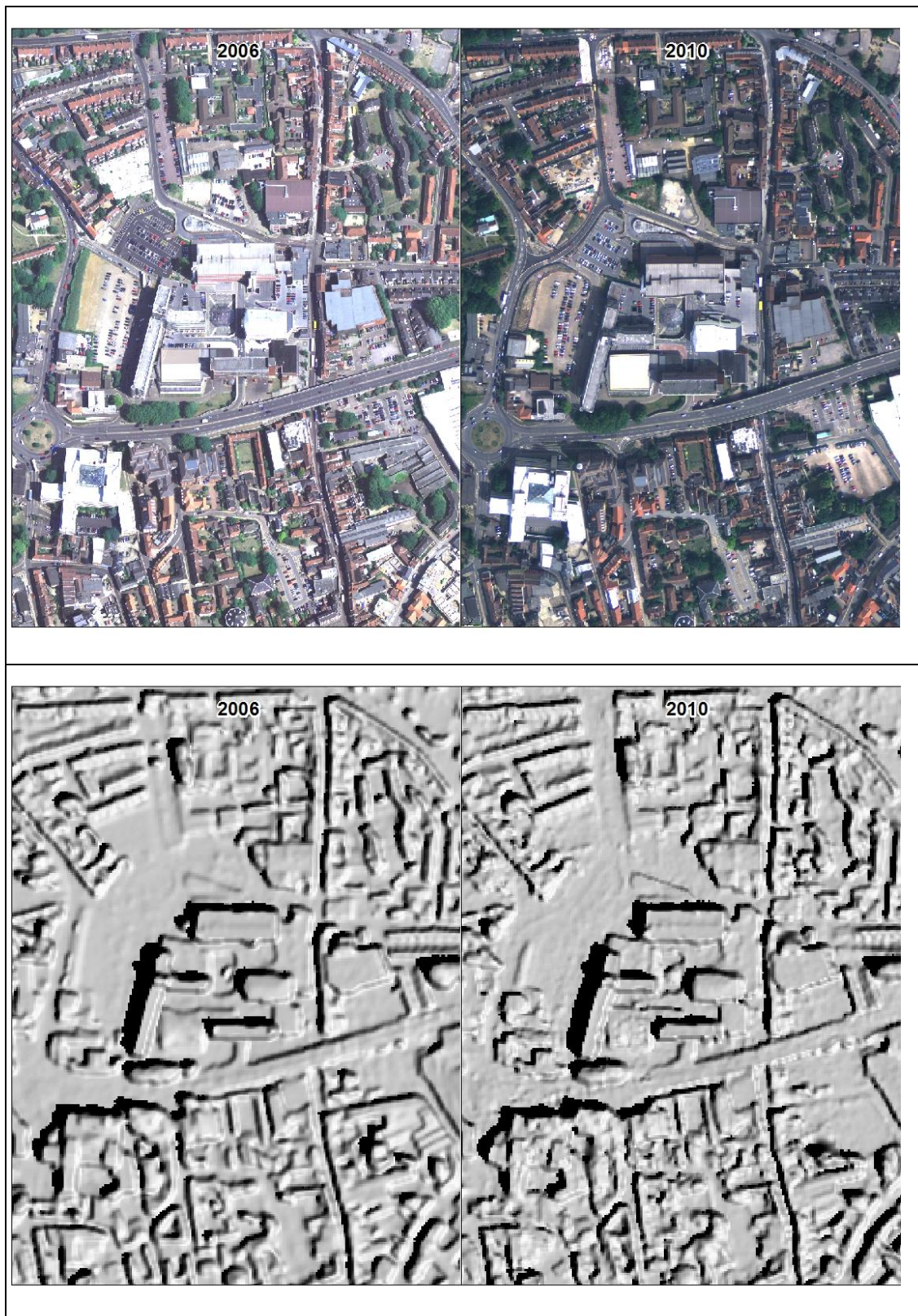


Figure 26: The hill-shading illumination modelling. The top row shows the multi-temporal imagery over Norwich city centre. The bottom row shows the hill-shaded DSM with shadows applied for the solar geometry at the time of acquisition.

7.9 Unit of analysis

Multi-temporal image objects were selected as the unit of analysis for change detection. Therefore, image objects were formed by segmenting the entire multi-temporal image stack. Chapter 2.2.5 explored multi-temporal segmentation in detail, and it was found that under this framework, image tone, texture, pattern, association and shadow can be monitored over time. On the other hand, because the spatial framework is fixed, size and shape cannot be monitored over time. However, size and shape of the shared geometry can be calculated and used as a static, non-comparable input to classifications.

A hierarchical, multi-scale segmentation was employed to add context to the change detection. A hierarchical, multi-scale segmentation is a nested set of image objects at different scales. At the bottom of the hierarchy reside the smallest objects, representing a large cartographic scale. The bottom of the hierarchy, in this thesis, is labelled L1. The next level in the hierarchy is labelled L2 and contains larger, parent objects. As more levels are added, the child objects are aggregated further. Crucially, image objects at one level inherit the properties of their parent objects. Therefore, L1 objects inherit the properties of L2, L3, through to Ln. Inheriting parent object information adds context to L1 by quantitating the surrounding area. Figure 27 illustrates the structure and inheritance of a multi-scale segmentation with three levels.

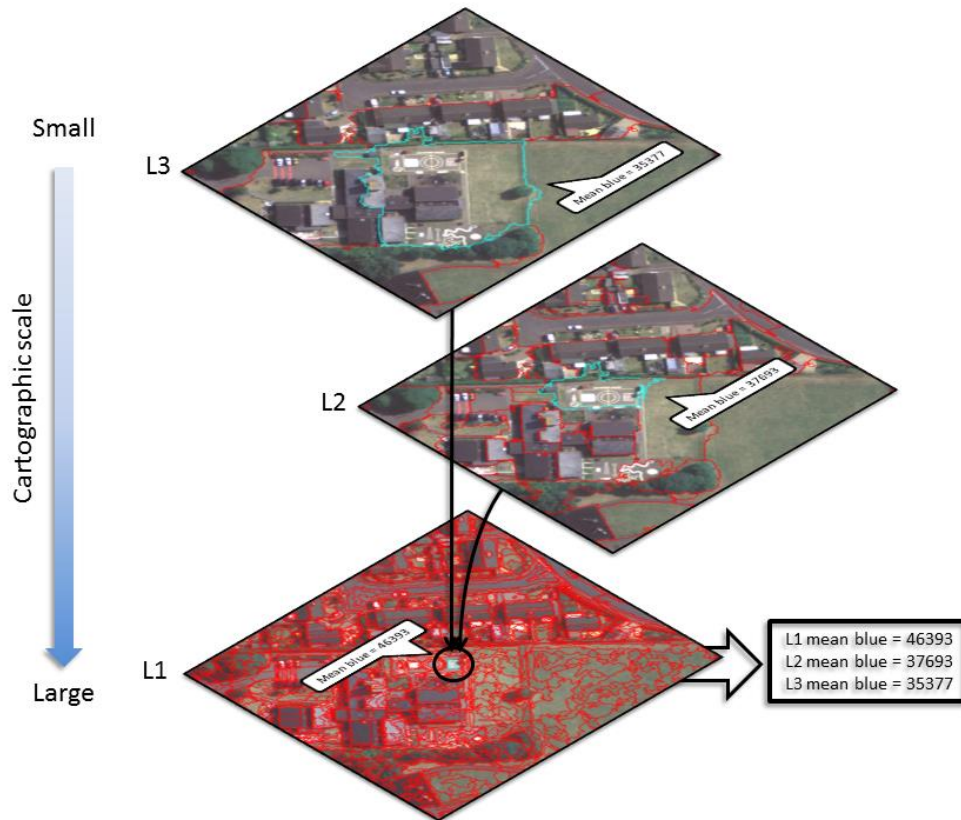


Figure 27: An illustration of a hierarchical, multi-scale segmentation. The illustration shows a nested segmentation at three different scales. The highlighted object at L1 is a child object and grandchild object to highlighted objects in levels 2 and 3 respectively. An example is given, where the highlighted L1 object inherits the mean image blue value of L2 and L3.

L1 objects are formed using the multi-resolution segmentation (MRS) algorithm found in eCognition 9.0 (Trimble 2015). The algorithm parameters were calibrated qualitatively to form small objects with little, or no land cover mixing, and where marginal over-segmentation is preferred over under-segmentation. The results of this type of segmentation are often referred to as *image object primitives* (Batz et al. 2008; Tabib Mahmoudi et al. 2013; Diesing et al. 2016; Martha et al. 2016). The L1 objects were created with the MRS parameters: scale, shape and compactness set to 250, 0.1 and 0.5 respectively. To control over-segmentation and stabilise the scale, objects of less than 10 pixels were merged to the adjacent object with the best colour match.

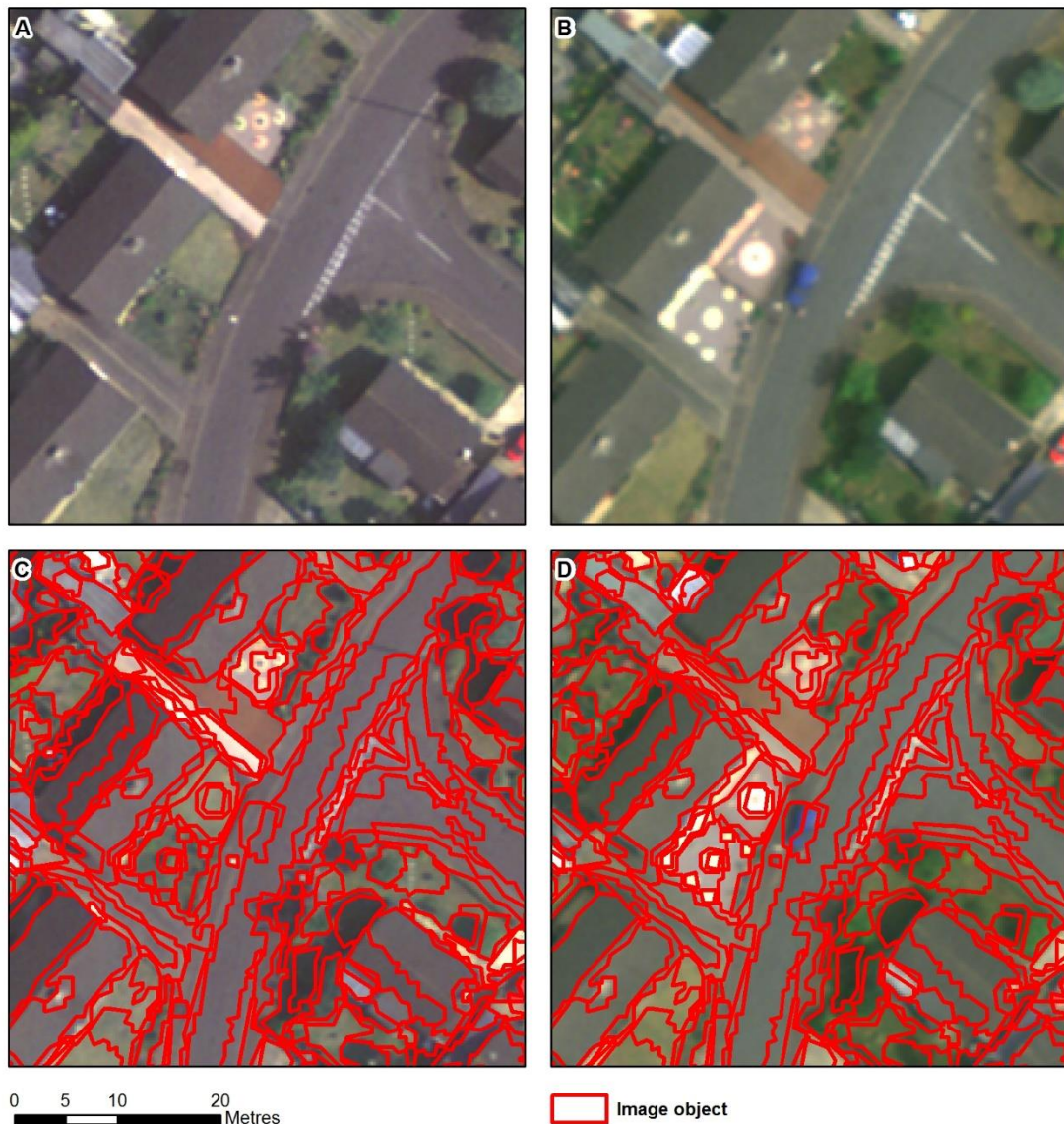


Figure 28: The results of the L1 multi-temporal MRS segmentation. Panels A and B show the 2006 and 2010 imagery respectively. Panels C and D show the 2006 and 2010 imagery respectively, overlaid with the L1 image objects.

The MRS algorithm was improved to give more meaningful image objects at smaller scales. Deriving image objects that tend towards meaningful geographic objects is a fundamental part of OBIA (Castilla & Hay 2008; Blaschke 2010; Hussain et al. 2013). Meaningful image object primitives, L1 in this thesis, are easily achievable with MRS. However, deriving meaningful small-scale objects over urban areas, with increasing levels of abstraction, is a non-trivial task. This is because the spectral and textural patterns in the image do not necessarily relate to urban structures. Qualitative experimentation identified that even when the MRS shape and compactness parameters are set to maximum, the resultant image objects favour image colour over

form. This results in overly complex objects, that often aggregate roof structures and shading patterns, but do not adequately depict urban structures. To improve urban structure representation, a modification of MRS was derived, Compactness Enforcement Multi-resolution Segmentation (CEMRS). CEMRS gives segmentations that tend towards compact objects; that better describe urban structures. It works by targeting complex objects, unlikely to outline urban forms, for re-segmentation with a more abstract approach. Figure 29 gives an overview of the CEMRS algorithm, while a more detailed description is given in pseudo code in the thesis Appendix, Figure A-5. To summarise the algorithm; a seed image object level (IOL), L_{n-1} , is aggregated using the standard MRS to give a parent IOL, L_n . Then L_n complex objects are identified and split back to their child, sub-objects (L_{n-1}). The child objects are then dissolved into the surrounding, compact objects with the greatest shared border. Then, surrounded objects, ring objects, and partial ring objects are removed. The algorithm iterates until no more complex objects remain. Figure 30 shows a side by side comparison of MRS and CEMRS when applied to the Norwich imagery. It highlights that MRS image objects at relatively small scales do not adequately describe urban form. The CEMRS algorithm shows clear improvement.

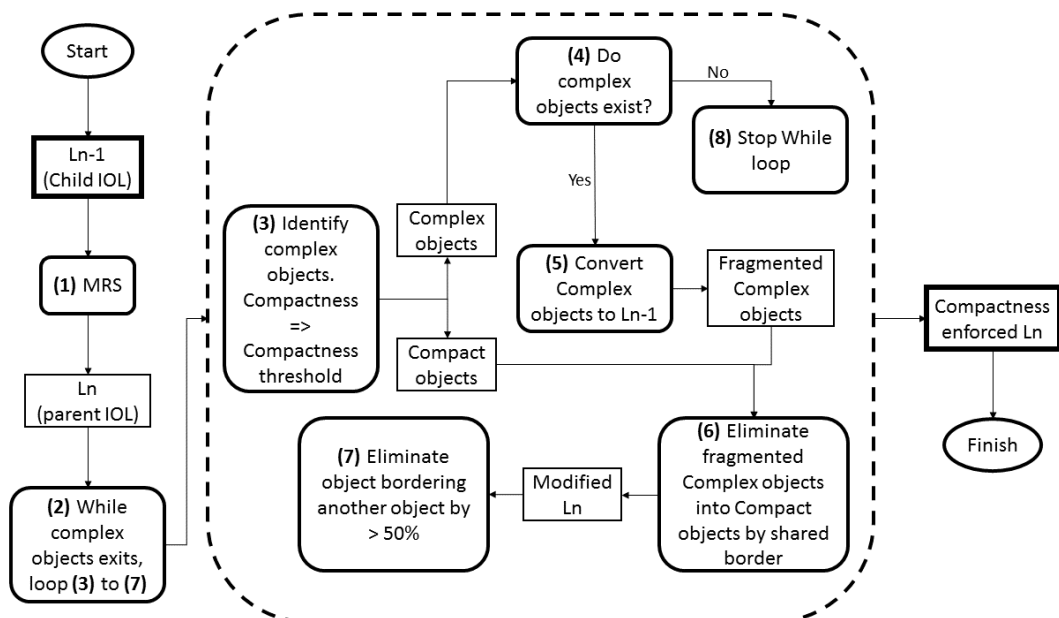


Figure 29: A summary of the CEMRS algorithm.

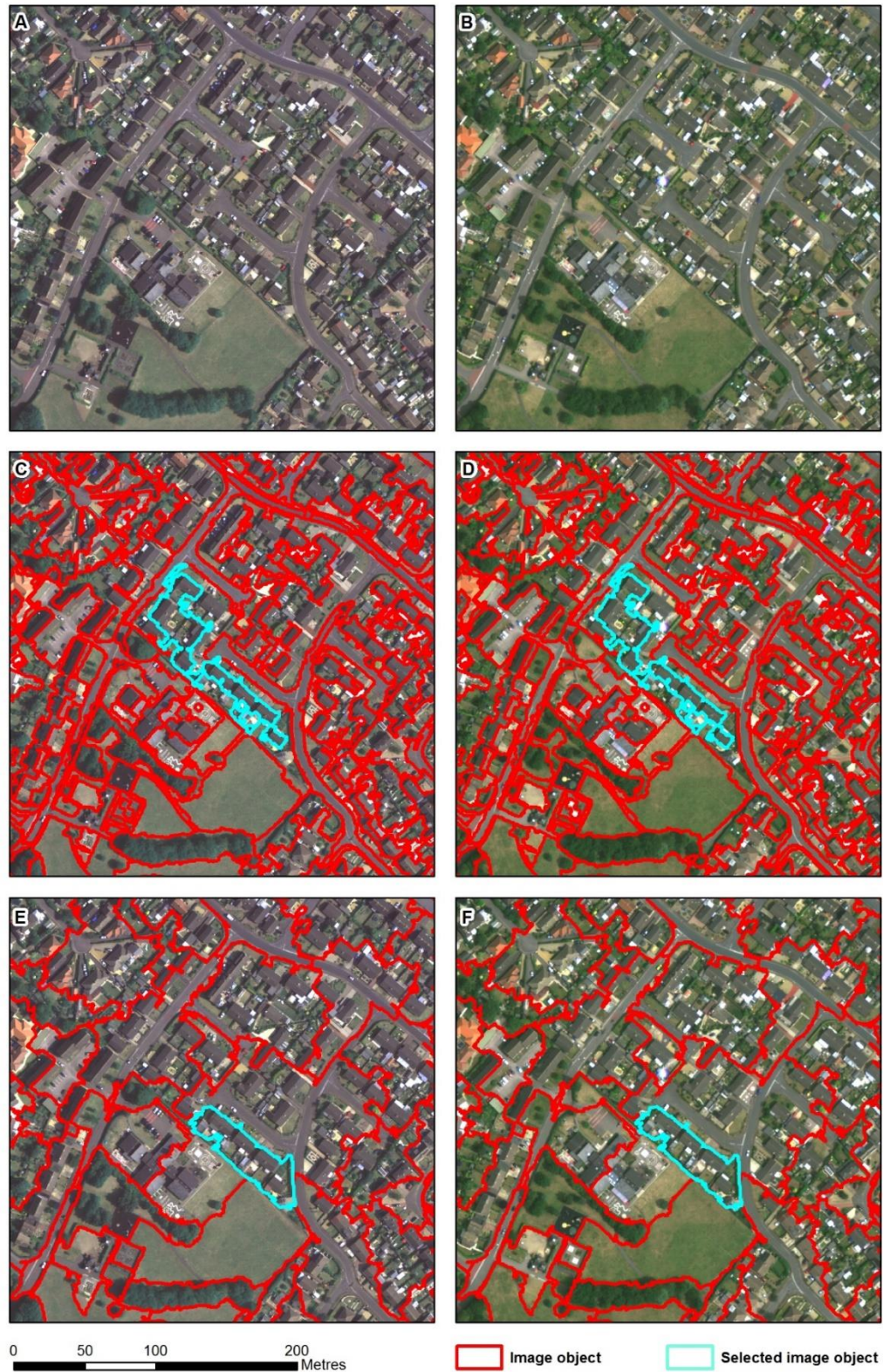


Figure 30: A cross comparison of MRS and CEMRS. Panels A and B show the 2006 and 2010 imagery respectively. Panels C and D show the 2006 and 2010 imagery respectively, overlaid with the MRS image objects. Panels E and F show the 2006 and 2010 imagery respectively, overlaid with the CEMRS image objects. The MRS shape and compactness parameters are set to maximum. The Scale parameters of both algorithms are calibrated to deliver the same number of image objects within the image chip. MRS creates objects that tend towards image colour, and do not adequately depict urban structures. CEMRS delivers compact abstracted objects that more closely match urban forms.

The CEMRS algorithm was calibrated qualitatively to derive L2 and L3 image object levels that approximate residential property and residential block respectively. Since this research targets urban creep, property extent and urban block are meaningful, complimentary analysis scales to the image object primitives described in L1. Table 18 shows the CEMRS parameters used to achieve these abstractions. Figure 31 and Figure 32 show the results of the L2 and L3 segmentations applied to the Norwich imagery. In total 5499378 objects were formed for L1, 193014 for L2, and 47696 for L3.

Table 18: The CEMRS parameters used to create IOL L2, and L3.

Ln	L2	L3
Ln-1	L1	L2
Scale	400	800
Shape	0.9	0.9
Compactness	0.9	0.9
Compactness threshold	3	3

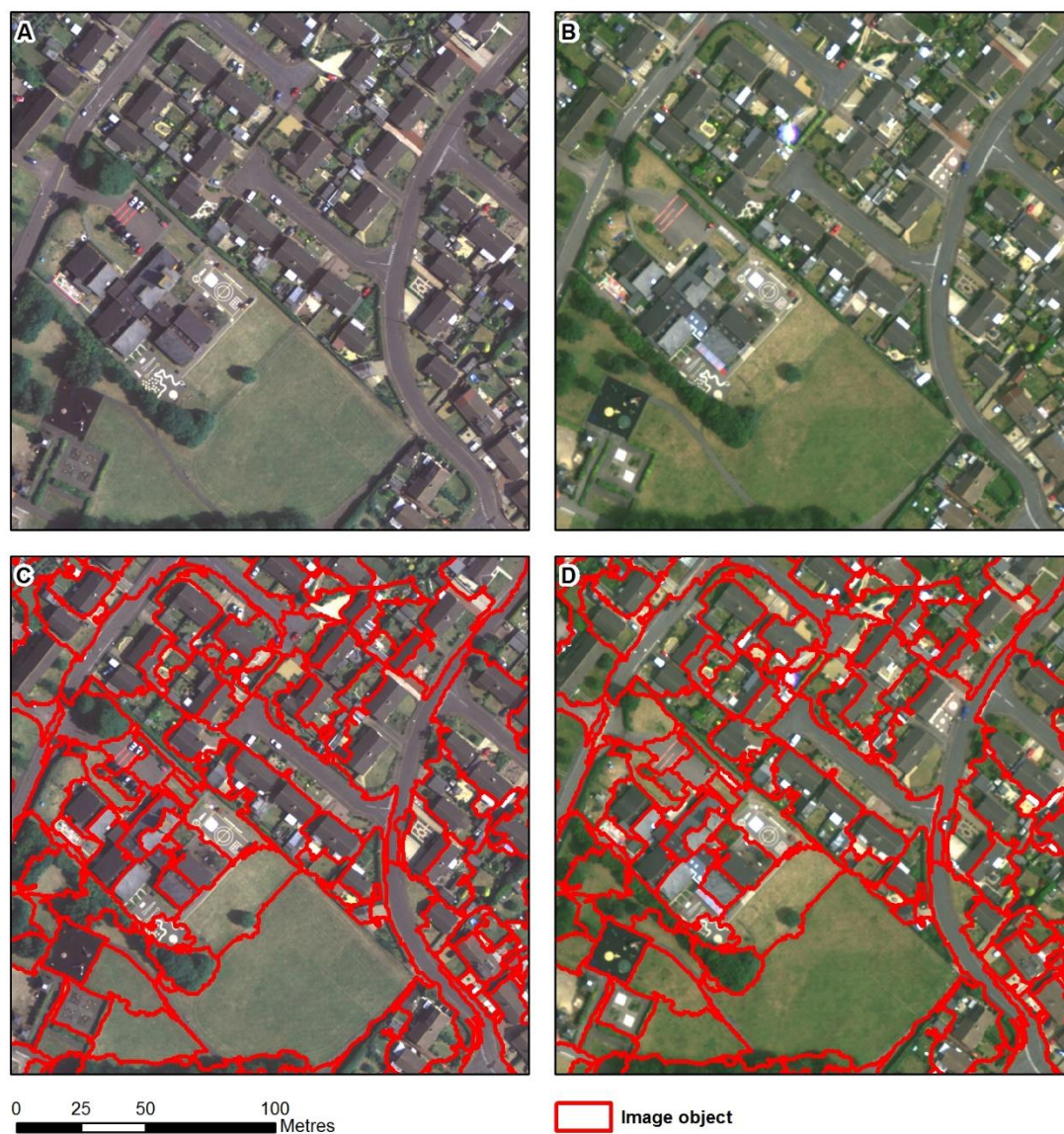


Figure 31: The results of the L2 multi-temporal CEMRS. Panels A and B show the 2006 and 2010 imagery respectively. Panels C and D show the 2006 and 2010 imagery respectively, overlaid with the L1 image objects.



Figure 32: The results of the L3 multi-temporal CEMRS. Panels A and B show the 2006 and 2010 imagery respectively. Panels C and D show the 2006 and 2010 imagery respectively, overlaid with the L1 image objects.

7.10 Feature extraction

In total, 117 features were calculated for each L1 image object, resulting in over 640 million data items across the study area. The features were chosen to represent a range of image interpretation principles and analysis scales, across two different time periods. All features were calculated using eCognition 9.0.

The complete 117 feature set is based on twenty-five feature themes, duplicated by scale and time. Table 19 catalogues the twenty-five feature themes, along with their brief description, image interpretation principle and comparability. Nine image interpretation principles were defined, namely, *tone*, *texture*, *height*, *shadow*, *view*

angle, pattern, association, shape and *size*. The nine principles are based on the set identified by Bianchetti & Maceachren (2015), with the specific addition of *view angle* to describe the hidden surfaces. Each principle is represented by one or more feature themes. For example, *tone* refers to all of the available spectral bands, in this case, red, green, blue and NIR. The feature themes belonging to *tone, height, shadow* and *view angle* all report the mean value of the pixels intersecting the image object. The texture metrics are calculated from an albedo image, specifically, the mean of the red, green and blue bands. Texture measures range from a simple standard deviation, through to GLCM measures. The choice of GLCM texture metrics was informed by recommendations identified by Warner (2011). Specifically, one contrast measure (homogeneity), one orderliness measure (entropy), and two descriptive statistics (mean and standard deviation). The *pattern* features refer to crude measures of spatial autocorrelation. The mean difference to neighbours is the mean albedo difference between an image object and its neighbours. In other words, how locally bright the object is. The standard deviation to neighbouring objects is the standard deviation of the albedo differences between an image object and its neighbours. In other words, how variable the local pattern is. *Association, shape* and *size* features all describe image object geometry.

Table 19: The twenty-five feature themes.

Feature theme	Description	Image interpretation principle	Analysis unit comparability
Blue	Object mean blue	Tone	Comparable
Green	Object mean green		
Red	Object mean red		
NIR	Object mean near infra-red		
Standard deviation	Object albedo standard deviation	Texture	
GLCM Entropy	Object albedo GLCM entropy		
GLCM Homogeneity	Object albedo GLCM homogeneity		
GLCM Mean	Object albedo GLCM mean		
GLCM Standard deviation	Object albedo GLCM standard deviation		
DSM	Object mean DSM	Height	
Hill shaded DSM	Object mean hill shaded relief and shadows	Shadow	
Hidden surfaces	Object mean visibility and feature parallax	View angle	
Mean difference to neighbours	Object albedo, mean difference to neighbouring objects	Pattern	
Standard deviation to neighbours	Object albedo, standard deviation to neighbouring objects		
Direction	Object direction, or azimuth	Association	Non-comparable
Asymmetry	Object asymmetry, a measure of elongation	Shape	
Compactness	Object compactness, the similarity to the most compact shape a circle		
Density	Object density, the ratio of area divided by radius		
Elliptical fit	Object elliptical fit, similarity to an equivalent ellipse		
Length/Width	Object length/Width		
Rectangular fit	Object rectangular fit		
Roundness	Object roundness, radius difference between the enclosing and enclosed ellipse		
Area	Object area	Size	
Length	Object length		
Width	Object width		

All of the feature themes described are calculated at three different scales. The three different scales are the image object levels, L1, L2 and L3, which relate to relatively large, moderate and small cartographic scales respectively. The only exception to this rule are the *pattern* features, which are defined by a neighbourhood of image objects. All of the *pattern* feature scale variants are calculated at L1, with increasingly large neighbourhood sizes. Specifically, large-scale neighbourhoods are all objects adjacent to the target. Moderate scale neighbourhoods are all objects within 100 pixels (25m) of the target, while small-scale neighbourhoods are within 200 pixels (50m) of the target.

It is important to make the distinction between *comparable* and *non-comparable* features. Not all the properties of multi-temporal image objects can be described over time because of the fixed spatial framework. Specifically, the *association*, *shape* and *size* feature themes can only be calculated once for both time periods, and so are deemed *non-comparable*. Comparable feature themes are calculated twice, one for each period, whereas non-comparable feature themes are calculated once and shared between the time periods. For the features to remain readily compatible with CVA, comparable features are stored as T0, for the 2006 data and Δ as the difference between 2006 and 2010. Non-comparable features are labelled as T0 & T1 to denote shared properties. In summary, the source data state of the feature themes was calculated and recorded as either T0, Δ , or T0 & T1.

The complete set of 117 features is constructed by a multiplication of the feature themes at different scales and time periods. There are fourteen comparable feature themes. Each of these is calculated at three different scales and two different source data states. Therefore, each feature theme multiplies to six features. For example, the feature theme *Blue* multiplies to the features: *T0 L1 Blue*, *T0 L2 Blue*, *T0 L3 Blue*, *Δ L1 Blue*, *Δ L2 Blue* and *Δ L3 Blue*. Conversely, each of the eleven non-comparable feature themes is calculated at three different scales, but only one source data state. For example, feature theme *Area* multiplies to the features: *L1 Area*, *L2 Area* and *L3 Area*. The final feature set is a result of the multiplication:

$$117 = (14 \text{ Feature themes} \times 3 \text{ Scales} \times 2 \text{ Source data states}) \\ + (11 \text{ Feature themes} \times 3 \text{ Scales} \times 1 \text{ Source data state})$$

For reference, the full feature listing can be found in Table A-3 of the Appendix.

7.11 Unsupervised feature selection

The unsupervised feature selection technique agglomerative hierarchical clustering was applied to identify meaningful feature sets, with minimal data redundancy.

Agglomerative hierarchical clustering is a bottom-up algorithm that recursively aggregates data points or features into a nested structure using a dissimilarity matrix (Murphy 2012, p893). In other words, each feature is initialised as a cluster, and a dissimilarity score is computed for all cluster pairs. The cluster pair with the lowest dissimilarity are merged, and the dissimilarity matrix is re-computed. The cluster merging and dissimilarity computation continue until all features are aggregated to one cluster. The agglomerative hierarchical clustering was implemented using the function *FeatureAgglomeration*, from the Python module, Scikit Learn (Scikit Learn 2016b). The dissimilarity criterion specified was Ward's error sum of squares (Ward 1963), in Euclidean feature space. Ward's method minimises cluster variance by considering the 'loss' (Ward 1963, p237) of information incurred by each cluster pair.

The cluster representation method proposed by Park (2013) was used to select individual features. Agglomerative hierarchical clustering is an intermediate stage in the feature selection process because it does not indicate which feature from each cluster should be chosen. Park (2013) demonstrated a method utilising feature-to-feature correlation. Specifically, a correlation matrix is computed for the features within a cluster, and the correlation scores are summed by feature. The feature with the highest correlation sum is the one selected. In other words, from each cluster, the feature selected is the most representative of the entire cluster.

The cluster representation method was iterated 117 times to reveal all feature selection scenarios. Starting at the root of the hierarchy tree, all features are contained in one cluster. The single most representative feature is then identified. We then move one position up the hierarchy, giving two clusters, where the two most representative features are revealed. This process continues until the hierarchy tree leaves are reached, and each cluster contains only one feature. Classifying all of the feature selection scenarios will reveal the optimum feature set.

All data was standardised before feature selection to allow meaningful clustering among features with different data ranges and distributions. Many of the 117 features reside on different scales, which would invalidate clustering operating in Euclidean space. For example, image spectral features, such as *T0 L1 Blue*, have a range of 0 to 65535, while the GLCM texture measures, such as *T0 L3 GLCM Homogeneity*, have a range of 0 to 90. Furthermore, differing data distributions may adversely affect the clustering. Dy & Brodley (2004) show how standardisation improves feature selection performance because clustering tends towards low variance distributions. Therefore all of the data was standardised, to have a mean of 0 and a standard deviation of 1.

7.12 Change vector analysis

7.12.1 Established formulation (C²VA)

Firstly, the established state-of-the-art formulation of CVA is described graphically and mathematically. The state-of-the-art formulation, compressed CVA (C²VA), is presented in detail by Bovolo et al. (2012) and describes CVA in n -dimensions, where the CV direction is evaluated against an arbitrary reference vector. Figure 33 illustrates the geometry of a C²VA system, where Vector A represents T0, Vector B represents T1, and the reference vector is a unit vector in the centre of the feature space. As a unit vector, the magnitude of the reference vector in each of the n dimensions is $1/\sqrt{n}$. The illustration depicts the system in two dimensions, x and y , while C²VA is scalable up to n dimensions, i.e. $x, y, z \dots \dots n$.

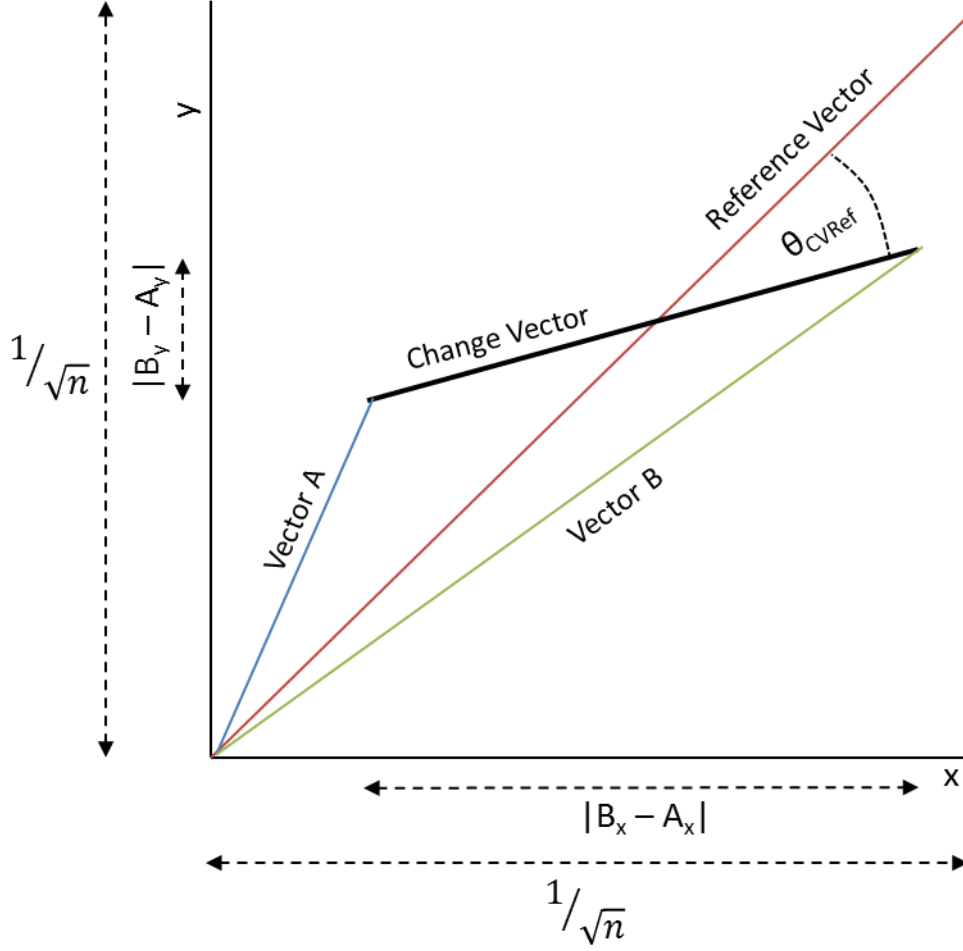


Figure 33: C^2VA geometry. Vector A represents $T0$ and Vector B $T0$.

Given a feature vector at $T0$, A and a feature vector at $T1$, B the difference from $T0$ to $T1$, Δ , is clearly:

$$\Delta = B - A$$

If we consider n features, the Euclidean norm, or CV magnitude is:

$$CV \text{ magnitude} = \sqrt{\sum_{i=1}^n \Delta_i^2} \quad \text{Equation 17}$$

While the angle between the change vector and the reference vector, namely the CV direction is given by (Bovolo et al. 2012):

$$\theta_{CVRef} = \cos^{-1} \left[\frac{1}{\sqrt{n}} \left(\frac{\sum_{i=1}^n \Delta_i}{\sqrt{\sum_{i=1}^n \Delta_i^2}} \right) \right] \quad \text{Equation 18}$$

7.12.2 Weighted C²VA

This thesis presents Weighted C²VA (WC²VA) as a novel extension to the established state-of-the-art. WC²VA is advantageous because it normalises CV magnitude under variations in the number of input features, and facilitates the application of feature weightings to determine importance. Without weightings, CV magnitude is proportional to \sqrt{n} . In other words, as more features are added the magnitude increases, making cross comparisons difficult. We can normalise CV magnitude by applying weights to each feature:

$$\text{Weighted CV magnitude} = \sqrt{\sum_{i=1}^n w_i (A_i - B_i)^2} \quad \text{Equation 19}$$

Where, $0 \leq w_i \leq 1$ and $\sum_{i=1}^n w_i = 1$

Next, we apply weights to the direction, θ_{CVRef} . The unweighted angle between vectors is related to the dot product of the two vectors and their magnitudes (Olver & Shakiban 2006, p139):

$$\cos \theta_{AB} = \frac{A \cdot B}{|A||B|} \quad \text{Equation 20}$$

Where $|A|$ and $|B|$ are the vector magnitudes, and the dot product $A \cdot B$ is calculated by multiplying the row vector of A by the column vector of B (Olver & Shakiban 2006, p131):

$$A \cdot B = \sum_{i=1}^n A_i B_i \quad \text{Equation 21}$$

Multiplying each axis by a weight, w , we derive expressions of the weighted dot product, and weighted magnitude (Olver & Shakiban 2006, p133):

$$A \cdot B = \sum_{i=1}^n \omega_i A_i B_i \quad \text{Equation 22}$$

$$|A| = \sqrt{\sum_{i=1}^n \omega_i A_i^2} \quad \text{Equation 23}$$

Substituting Equation 22 and Equation 23 into Equation 20, we have the weighted angle, $\theta_{AB\omega}$:

$$\cos \theta_{AB\omega} = \frac{\sum_{i=1}^n \omega_i A_i B_i}{\sqrt{\sum_{i=1}^n \omega_i A_i^2 \sum_{i=1}^n \omega_i B_i^2}} \quad \text{Equation 24}$$

Considering the change vector and reference vector, R_i , Equation 24 becomes:

$$\cos \theta_{CVRef\omega} = \frac{\sum_{i=1}^n \omega_i \Delta_i R_i}{\sqrt{\sum_{i=1}^n \omega_i \Delta_i^2 \sum_{i=1}^n \omega_i R_i^2}}$$

Finally, we invert the cosine and simplify the reference vector to the per dimension magnitude, $1/\sqrt{n}$, to give the weighted direction, $\theta_{CVRef\omega}$:

$$\theta_{CVRef\omega} = \cos^{-1} \left[\frac{\sum_{i=1}^n \omega_i \Delta_i \left(\frac{1}{\sqrt{n}} \right)}{\sqrt{\sum_{i=1}^n \omega_i \Delta_i^2 \sum_{i=1}^n \omega_i \left(\frac{1}{\sqrt{n}} \right)^2}} \right] \quad \text{Equation 25}$$

Comparing C^2VA direction (Equation 18) to WC^2VA direction (Equation 25), it is noted that in the weighted variant, the denominator no longer collapses down to a single vector magnitude. This is because under unequal weightings, or if the weightings do not sum to 1, the weighted magnitude of the reference vector is no longer 1, so it cannot be ignored.

In summary, WC^2VA has the following properties:

- Weighted CV magnitude remains on a comparable scale (the magnitude does not escalate proportionally to \sqrt{n}) to the input features where the weights sum to 1, specifically: $0 \leq w_i \leq 1$ and $\sum_{i=1}^n w_i = 1$

- Under equal weights, reported direction is identical to unweighted, C^2VA , i.e. the outcomes of Equation 18 and Equation 25 are identical if $\omega_i = \omega_n$ for all axes.
- Unequal weights increase or decrease the relative importance of the particular axes.

7.12.3 Change trajectory analysis

This thesis presents a further novel extension of CVA, namely Change Trajectory Analysis (CTA). CTA summarises the magnitude and direction of the change vector, and the baseline vector $T0$. Considering the CV and $T0$ reflect the complete trajectory, from origin to $T1$. Furthermore, adding the baseline position, $T0$, removes CV ambiguity. CV ambiguity is established in the literature (Johnson & Kasischke 1998), as is its removal by adding a baseline vector (Cohen & Fiorella 1998). However, this research goes further by formalising these concepts under the WC^2VA framework.

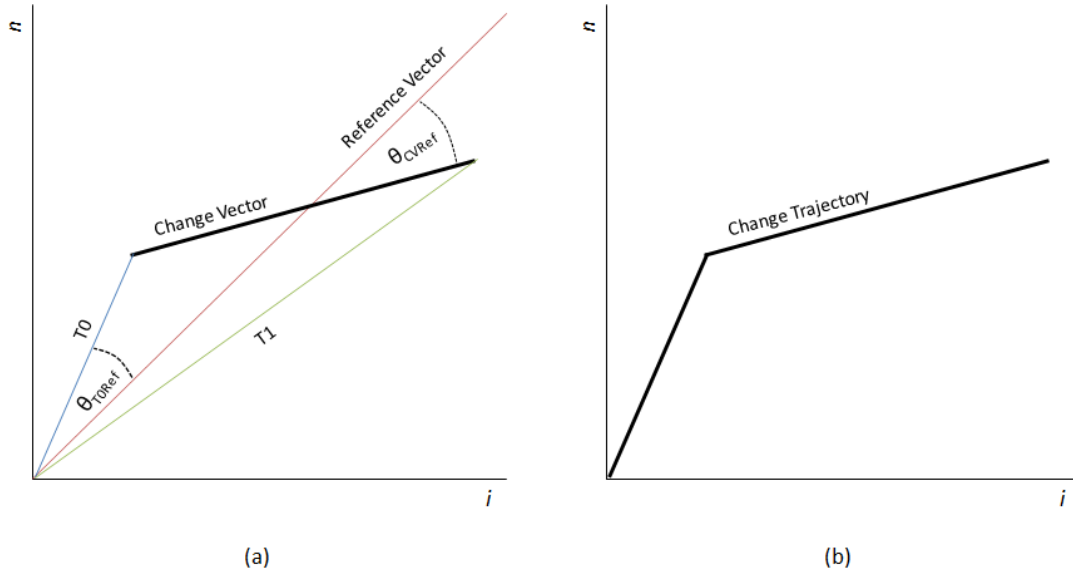


Figure 34: CTA illustration. (a) Illustrates the two reference vector angles. (b) Highlights the change trajectory, running from the origin, to $T0$, to $T1$.

Figure 34 illustrates CTA, to include the direction, θ_{T0Ref} and the trajectory geometry. CTA extends WC^2VA , by including the magnitude and direction of $T0$. Specifically:

$$\text{Weighted } T0 \text{ magnitude} = \sqrt{\sum_{i=1}^n w_i T0_i^2} \quad \text{Equation 26}$$

$$\theta_{T0Ref\omega} = \cos^{-1} \left[\frac{\sum_{i=1}^n \omega_i T0_i \left(\frac{1}{\sqrt{n}} \right)}{\sqrt{\sum_{i=1}^n \omega_i T0_i^2 \sum_{i=1}^n \omega_i \left(\frac{1}{\sqrt{n}} \right)^2}} \right] \quad \text{Equation 27}$$

Therefore, CTA transforms any number of inputs into the four components:

1. Weighted CV magnitude (Equation 19)
2. Weighted CV direction, $\theta_{CVRef\omega}$ (Equation 25)
3. Weighted T0 magnitude (Equation 26)
4. Weighted T0 direction (Equation 27)

7.13 Training data

The training data were collected in a similar manner to the reference data collection method described in Chapter 5. Chiefly, candidate reference samples were reviewed against multiple data sources to minimise labelling error. The training data selection started with the approximately stratified PSUs. In review, the 50x50m PSUs were assigned approximate strata: *urban creep*, *other change* and *no change* candidates. Recalling from Chapter 5, a total of 154 random stratified samples were selected for the reference dataset. The training samples were drawn from the remainder of the pool. Therefore the reference and training datasets were fully independent.

Training data samples were assigned by linking image objects to thematic classes identified in the multi-source data stack. Firstly, a PSU is drawn at random from one of the three strata, *urban creep candidate*, *other change candidate* or *no change candidate*. The PSU was then reviewed against the remotely sensed reference data, detailed in Table 6 of Chapter 5. If, after a rigorous review, there is confidence in the change class, the pattern on the ground is linked to a representative L1 image object. The image object attributes and the assigned training data label are the source data

for the supervised classification. This process is repeated until a suitable training data sample pool is achieved. A total of 300 training samples were collected, evenly split between the three thematic classes: urban creep, *other change* and *no change*.

7.14 Supervised classification

This section details the supervised classification method applied to the feature extracted image objects. This stage trains a classifier to apply thematic labels to the unlabeled image objects. Overall, two workflows are followed: direct classification, and CVA classification. The direct classification utilises the features directly. The CVA classification utilises CVA features, derived from the 'raw' features. For example, in the CVA classification image objects are trained and labelled by CV magnitude and direction, instead of the features directly. These two broad workflows relate to the thesis research questions, concerning the urban creep classification performance of direct, and CVA classifications.

All supervised classifications were based on a random forest classifier, with identical configurations. Therefore, the supervised classifier is fixed, and not confused with the experiments independent variables (the input features and classification type).

Random forest is an ensemble classifier based on a 'forest' of classification and regression trees, where each tree is a discrete classifier in their own right. Each tree is created from a random subset of the training data and feature space. The final classification is aggregated from the individual trees. The random forest classifier was chosen for three reasons:

1. Random forest is the 'Classifier of choice' for many applications (Kuncheva 2014, p191).
2. Tree aggregation produces a more generalised classification model that suppresses overfitting (Cichosz 2014, p443).
3. Random forest can handle a large number of features (Rokach 2009, p55).

The Python Scikit Learn implementation of random forest (Scikit Learn 2016a) was used in all experimentation. The classifier was configured to construct forests with 128 trees. The forest size was based upon Oshiro et al.'s (2012) recommendation of between 64 and 128 trees, where fewer trees often result in a compromised model,

and, more trees yields little tangible benefit. Each tree was generated using the standard subset of $\sqrt{n \text{ features}}$ (Cichosz 2014, p443), with no pruning applied (Kuncheva 2014, p191).

7.14.1 Direct classification

The supervised classification workflow applied to each direct classification is outlined in Figure 35. The workflow starts with the unlabelled, feature extracted, image objects containing N features. The image objects then participate in the unsupervised feature selection, described in section 7.11, to derive a subset of n features. A subset of n features is then also created from the training data, which is passed to train the classifier. The derived classifier is then applied to the image objects, to add thematic labels. The process is repeated for all feature selection scenarios.

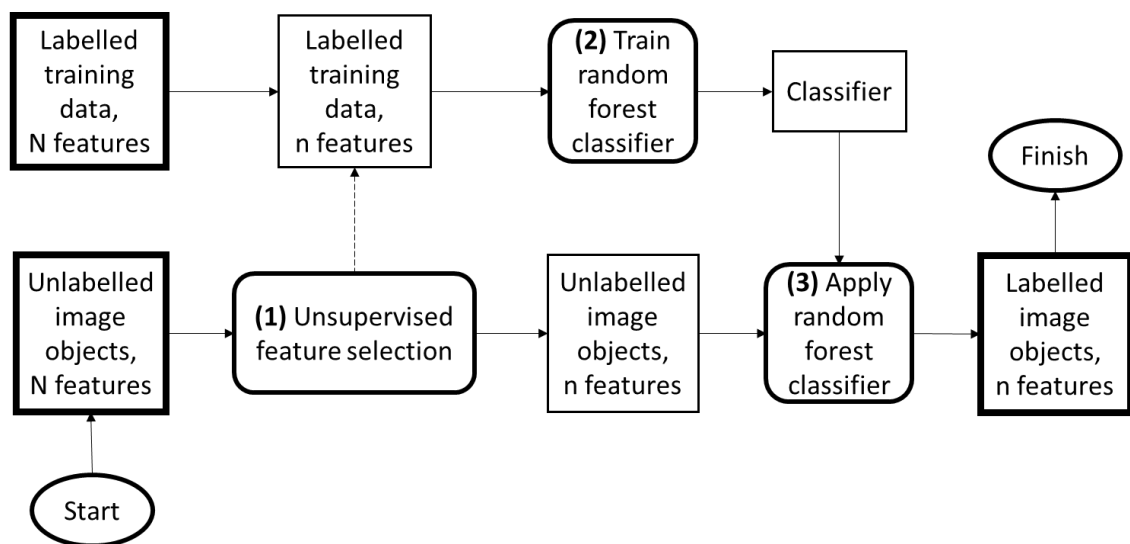


Figure 35: Direct classification workflow. The flow chart describes the main stages of the direct classification. Rectangles represent data or parameters. Rounded rectangles represent operations, numbered by their sequence.

7.14.2 CVA Classification

Figure 36 outlines the supervised classification workflow applied for each CVA classification. The workflow mirrors the direct classification workflow, with the addition of two CVA stages. Prior to classifier training or classifier application, the set of n -features was transformed to a CVA feature set. Firstly, the features are standardised so that the data has a consistent range and distribution. Then, the standardised data undergoes either W^2 CVA, or CTA to derive a set of CVA features.

The CVA features are either: CV magnitude, CV Magnitude and direction, or T0 and CV Magnitude and direction (For CTA).

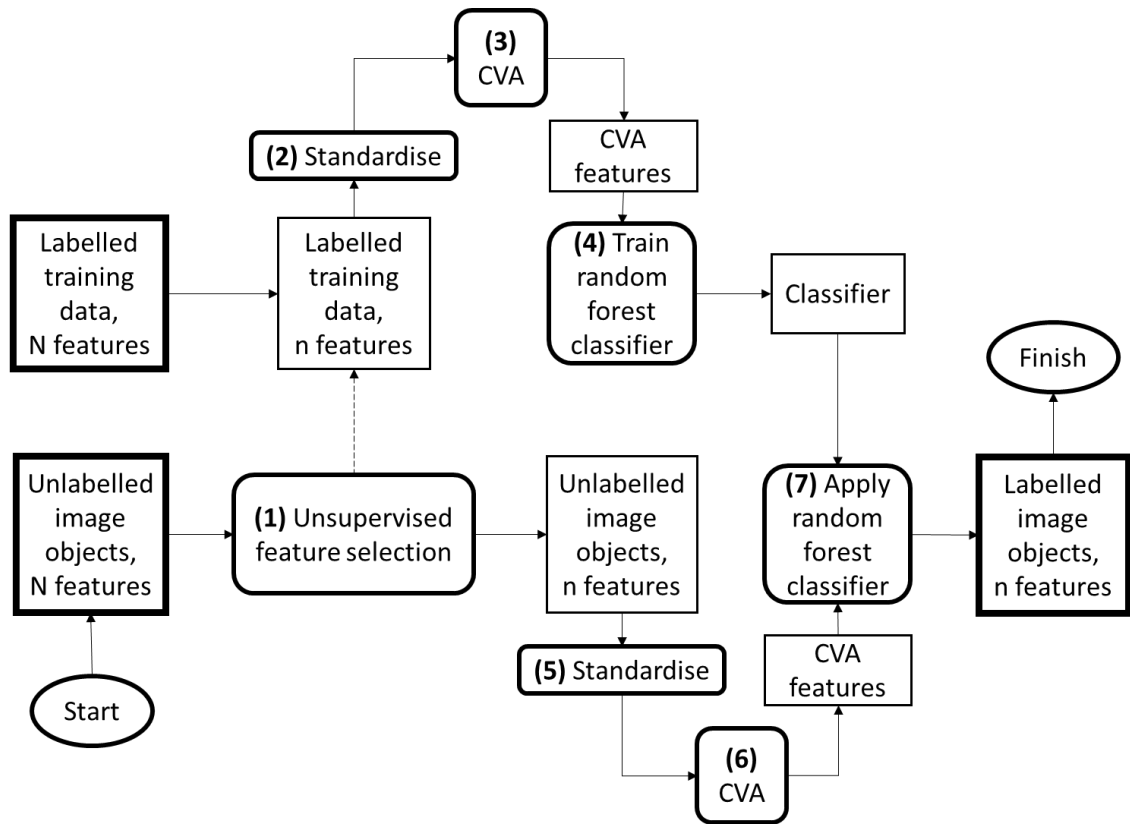


Figure 36: CVA classification workflow. The flow chart details the main stages of the CVA classification. Rectangles represent data or parameters. Rounded rectangles represent operations, numbered by their sequence.

7.15 Accuracy assessment

The performance of the supervised classifications was assessed with a site specific area weighted error matrix. The error matrix is a widely accepted and demonstrated representation of thematic accuracy (Congalton & Green 2009, p55). The method has the advantage of revealing individual class confusion and deriving properties that summarise thematic accuracy. The accuracy assessment was site specific. Therefore the location and extent of classified features are as important as the thematic label. In GIScience terms, a site-specific accuracy assessment is an overlay procedure, and thematic agreement is only achieved where the classified and reference data labels correspond at a particular location. By considering the agreement at all locations, the assessment is area weighted, since classified image objects may be partially correct, or indeed, partially incorrect. Therefore, the agreement is not binary. For example, an image object classified as urban creep could overlap a reference data geo-object by

80%, with the remaining 20% overlapping *no change*. The weighted error matrix directly reflects the degree of correspondence. Recalling the reference data in Chapter 5, reference polygons are mapped at a scale of 1:500. Therefore, the site specific area weighted error matrix essentially describes the level of cartographic similarity. By cartographic similarity, the Author means that both the thematic label and spatial configuration are tested. While other research (Adelabu et al. 2015) has shown that it is possible to implement an accuracy assessment without an independent dataset, using random forest out-of-bag samples only, this approach was not chosen here. This is because of the reference data objectives described in Chapter 5.4. Specifically, due to the duration of the research the reference data needed to be collected prior to the remote sensing change detection work.

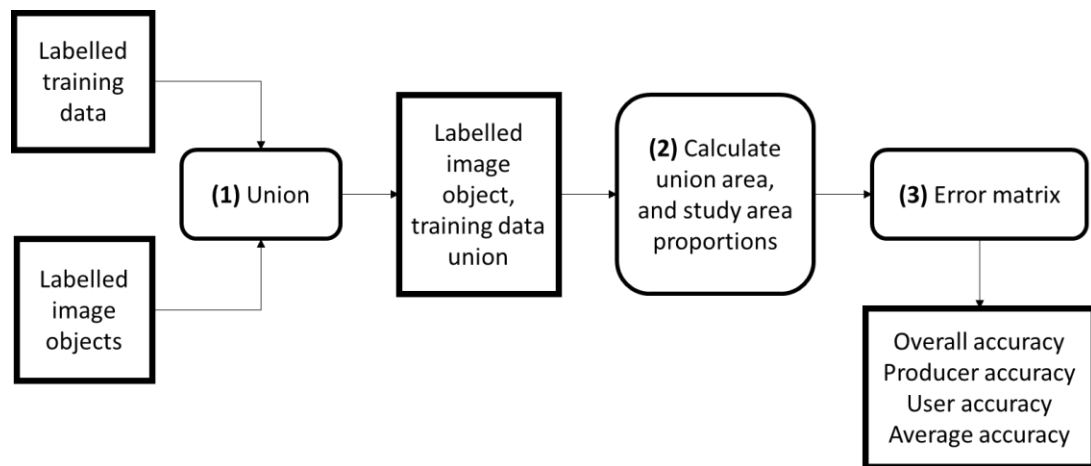


Figure 37: The error matrix preparation workflow. Rectangles represent data or results. Rounded rectangles represent operations, numbered by their sequence.

A geometric union of the labelled image objects and reference data mapping is executed to facilitate the area weighted accuracy assessment. Figure 37 shows the error matrix preparation workflow. Firstly, a geometric union of the labelled image objects and the reference data mapping is created. The union combines boundaries, disaggregating objects while retaining classified and reference class labels. The area and proportional area, relative to the study area, is calculated for each disaggregated object. Upon thematic agreement or disagreement, the proportional area is reported in the error matrix. Therefore, the error matrix directly reflects partial agreement and in turn, cartographic similarity.

Urban creep average accuracy was chosen as the primary measure of supervised classification success. The error matrix properties: overall accuracy, producer accuracy, and user accuracy, are commonly recommended summaries (Liu et al. 2007). While overall accuracy is the most commonly reported measure of success (Congalton & Green 2009, p59), urban creep average accuracy is a more suitable measure in this thesis. Average thematic accuracy is simply the mean of the user accuracy, ua , and producer accuracy, pa , for the i^{th} thematic class (Liu et al. 2007), as formalised in Equation 28. Judging supervised classifications by average urban creep accuracy will explicitly reveal the success of the classification in determining complex urban change. On the other hand, overall accuracy may mask poor urban creep classification performance, in favour of other classes. The Kappa statistic was not reported, because of its reported problems (Olofsson et al. 2014).

$$aup_i = \frac{ua_i + pa_i}{2}$$

Equation 28

7.16 Experiment structure

The experiment consists of 375 classifications, split across two groups, and twelve scenarios. Table 20 summarises the classification groups and scenarios. The two classification groups, direct classification and CVA classification relate to the first two thesis research questions. The first research question questions the ability of a direct classification to identify urban creep, where the direct classification represents the state-of-the-art in change detection. The second question questions the ability of CVA to identify urban creep, where CVA represents change detection differencing methods. The classification scenarios further explore the parent classification group to cover a range of different input scenarios and facilitate more meaningful comparison between the two classification groups.

Table 20: Classification group and classification scenario summary.

Classification group	Classification scenario	Individual classifications
Direct classification	Direct classification, all features	117
	Direct classification, Δ features	42
	Direct classification, T0 & Δ features	84
CVA classification	WC ² VA magnitude	42
	WC ² VA magnitude & direction	42
	CTA	42
	Red & NIR WC ² VA magnitude	1
	Red & NIR WC ² VA magnitude & direction	1
	Red & NIR CTA	1
	4 Band WC ² VA magnitude	1
	4 Band WC ² VA magnitude & direction	1
	4 Band CTA	1
	Total	375

The three direct classification scenarios explore the impact of considering all features, or two subsets of the features mirroring the CVA inputs. The first, ‘direct classification, all features’, considers all 117 features, over 117 classifications. The high number of classifications equate to an exhaustive exploration of all feature selection scenarios, from 1 to 117 features. Starting with all features is interesting to assess the most comprehensive classification approach, but it does not enable a fair comparison against CVA. This is because WC²VA is limited to the 42 difference features, while CTA also considers the T0 vector, effectively doubling the features to 84. In light of this the scenarios ‘Direct classification, Δ features’ and ‘Direct classification, T0 & Δ features’ operate on subsets of the full feature set, facilitating meaningful comparisons against the WC²VA and CTA classifications respectively.

The nine CVA classification scenarios exhaustively explore a range of experimental permutations, while also testing some common input variations. The scenarios

‘WC²VA magnitude’ and ‘WC²VA magnitude & direction’ both classify all of the 42 feature selections using WC²VA magnitude, and WC²VA magnitude and direction respectively. In doing so, the experiment exhaustively investigates the utility of a diverse range of object-based features in CVA. For example, in contrast to typically spectral CVA, the experiment yields quantitative results describing the merit of CVA using features such as multi-scale texture or pattern. The scenario ‘CTA’ extends ‘WC²VA magnitude & direction’ to include the T0 vector of each feature set. Note that CTA reports 42 classifications, i.e. the classifications do not double to 84. This is because the feature selection is decoupled from the T0 vectors; the selections are made on the 42 difference features, and the T0 vectors are added by default.

Four scenarios were added to represent simple, spectral CVA commonly applied in the literature (Bruzzone & Prieto 2000; Xian & Homer 2010; Bovolo & Bruzzone 2007), where image ‘colour’ change is the key discriminator. These four scenarios are:

1. Red & NIR WC²VA magnitude
2. Red & NIR WC²VA magnitude & direction
3. 4 Band WC²VA magnitude
4. 4 Band WC²VA magnitude & direction

These are referred to collectively as ‘Spectral CVA’ and serve as a crucial point of reference against the experimental CVA in this research. Furthermore, a comparison between the two and four band spectral CVA scenarios further test Bovolo et al.’s (2012) hypothesis that utilising all available spectral bands improves CVA class discrimination. WC²VA has been applied in these cases, not C²VA, to remain consistent with the other experiments. In other words, the input features are the independent variables under test and changing the CVA method for these scenarios would add a confusing variable to the experiment. Recalling the properties of WC²VA (section 7.12.2), under equal features weights (as applied here), CV direction is identical under C²VA and WC²VA. Conversely, CV magnitude range inflates proportionally to \sqrt{n} for C²VA but remains static for WC²VA. Under the random forest classification used in these experiments, the magnitude differences are unlikely to affect the results.

However, in the interest of good science, it is deemed prudent to keep the CV magnitude ranges static.

Lastly, the scenarios: Red & NIR CTA and 4 Band CTA, were added to explore the effect of adding T0 to the spectral CVA scenarios.

To summarise the form of the experiments, the dependent variable is the classification accuracy determined by the error matrix. Urban creep average accuracy will be the key measure of success. However, overall accuracy, the average accuracy of the other classes, and specific user and producer accuracies will all be assessed. The classification accuracy will be modulated against the independent variables: input features, direct classification, WC²VA classification and CTA classification.

7.17 Summary

This chapter documented the methods applied to the Norwich study area to quantify urban creep classification performance of two approaches: direct classification and CVA classification. Firstly, the remotely sensed data was outlined describing a multi-temporal 25cm, four-band imagery with accompanying 2m DSM. It was noted that, although the data is of good quality, the 2010 imagery has poorer contrast and DSM accuracy when compared to 2006. The data was spatially and radiometrically co-registered to facilitate meaningful change detection. The radiometric co-registration involved a novel PIF identification method, ensuring spatial and radiometric distribution of the PIF targets.

The methods presented contribute to the literature by presenting novel approaches to tackle key challenges such as scene illumination, viewing geometry, and object-based features. In particular, simplified GIS methods were described to identify hidden surfaces and simulate scene illumination from DSM layers. While, the chosen unit of analysis describes the theoretical merit of simultaneously reporting features at multiple viewing scales. When evolving this approach, the most commonly used segmentation algorithm MRS, proved inadequate at smaller scales. Therefore, CEMRS was developed to improve the delineation of urban form in VHR imagery.

The feature extraction yielded 117 features representing two time periods, three cartographic scales, and twenty-five feature themes. The feature themes covered a range of image interpretation principles such as tone, texture, pattern and association. The features were then subject to an unsupervised feature selection process to reduce feature redundancy and facilitate semantic interpretations of the feature sets.

The urban creep application is used to explore CVA exhaustively. A huge variety of inputs were tested to include object-based features describing tone, texture and pattern, at three different scales. This complicated CVA application required method extensions be developed. In particular, this work follows on from Bovolo et al. (2012) to include WC²VA, and the baseline vector aware variant, CTA.

The experimentation was organised into two groups, *Direct classification* and *CVA classification* to address the two research questions. Each group was subdivided into several classification scenarios exploring different inputs, such as the difference features alone, or all available features. Each classification scenario was classified for all feature selections, exhaustively identifying the optimum configurations. Lastly, the accuracy of each classification was assessed against a site specific, area weighted error matrix.

8 Change detection results

8.1 Introduction

This chapter presents the remote sensing change detection experimental results. In total, 375 remote sensing change detection classifications were executed to explore the relative performance of a direct classification and CVA, when applied to urban creep identification. In all, fifteen notable results were recorded. The highlights of these results include:

- The most successful classification mapped urban creep to an accuracy of 52.15%.
- In most cases, more than half of urban creep is not identified, and increasing classification complexity does not improve identification rates.
- Urban creep is a special case of change, with properties closer to unchanged areas.
- The identification of other changes is far more accurate.
- Direct classifications of urban creep significantly outperform equivalent CVA classifications.
- The novel CVA methods defined in this thesis improve overall change detection performance when compared to established CVA methods.

This chapter contributes to the thesis by providing the evidence by which the first two research questions will be answered and contextualised within the wider discipline. The first question refers to a direct classification as the benchmark, state-of-the-art, method, and whether or not it can satisfy the demands of this challenging application. Therefore, this chapter presents results to show the most successful direct classifications, the contributing factors to that success, the nature of errors and the sources of those errors. The second question asks if classical differencing methods, specifically CVA, are applicable in complex urban change detection? Thus, this chapter will present comparative results to the direct classifications and highlight the extent of their differences.

This chapter presents the results in five sections detailing the unsupervised feature selection, classification results, best case classification and the relationship between urban creep and CVA. Lastly, the results are collated and summarised.

8.2 Unsupervised feature selection

This section presents the results of the unsupervised feature selection. Specifically, the results highlight the classification features which are most representative of the entire dataset. Selecting the most representative features will reduce the redundancy in the classification models and simplify feature set interpretation.

In total, 117 classification features were considered, aiming to reduce these to their fundamental components. The full 117 features contain 42 baseline (T0) features, 42 difference ($\Delta T0$ to T1) features, and 33 shared (T0 and T1) features. All features are available to participate in the direct classifications, whereas shared features are not available to CVA change detection. In light of this, two separate feature selections are presented, one including all features and one containing the differences only.

The raw, tabular feature selection results are cumbersome and difficult to interpret. They take the form of $n \times n$ tables. For example, for the single, most representative feature, one feature is returned. For the two most representative features, two features are returned, and so on. Clearly, directly displaying and interpreting the selection from 117 features is not an informative approach. For reference, the first fifteen feature selections from all 117 features, and the 42 difference features are retained in Table A-4 and Table A-5 of the Appendix. Incidentally, the baseline, moderate scale object mean blue value, *T0 L2 Blue*, is the single most representative feature of all features. The moderate scale texture difference, *$\Delta L2 GLCM Mean$* , best represents all of the available difference features.

The results are organised into semantic groups and plotted as relative gains with which to compare against other semantic groups. All features were assigned semantic labels describing the analysis scale, image interpretation principle, and source state. For each feature selection, the proportional gain of a particular semantic group is plotted to show its contribution to the selected feature set. Positive gains mean that more features are selected from a semantic group than a proportional allocation. In

other words, the semantic group is over-represented in the selection. Negative gains mean that fewer features are selected from a semantic group than a proportional allocation. In other words, *negative gains* mean that a semantic group is under-represented in the selection. When all features are selected, all semantic groups have a neutral gain because the selection is forced into proportionality.

8.2.1 All features

The relative importance of analysis scale varies as increasing numbers of features are drawn. Figure 38 shows the analysis scale selection gain for all of the feature selection scenarios. Initially, with one feature selected, as expected a moderate-scale feature is most representative. Then from 2 to 4 features, large-scale features are over-represented in the selection, indicating that a combination of large-scale features best describes the fundamental elements of the data. From 4 to 14 features, large-scale suddenly becomes under-represented, replaced by small-scale features, while moderate-scale feature allocation remains approximately neutral. From 14 to 88 features moderate-scale features become over-represented in the selection, with large-scale remaining roughly neutral, while small-scale features are under-represented. This pattern indicates that there could be redundancy between small and moderate scales, with small-scale ‘context’ being selected when moderate and large-scale features are exhausted. Finally, from 88 to 117 features, the selections tend towards neutrality. This last selection pattern indicates new features are added by default, not because they contain unique information.

Figure 39 shows the proportional gain, of the nine image interpretation principles, for all feature selection scenarios. Generally, from 1 to 20 features, tonal features are the most over-represented, indicating that initially, tonal features hold the most information. However, from 25 to 94 features, image tone is under-represented. Below 20 features, texture, has a noisy, under-represented response before settling at marginal under-representation. This trend for under-representation indicates redundancy within the texture measures. Size features are largely over-represented in the selections. Aside from a positive anomaly from 4 to 9 features, pattern features are, in general, under-represented until neutrality is achieved from 47 features. From 8 to 40 features, shadow is over-represented before reaching neutrality.

Figure 40 shows the proportional gain, of the three source states, for all feature selection scenarios. The source state displays only subtle anomalies. In general, up to 20 features, the response is noisy and difficult to interpret. From approximately 20 features, the shared T0 and T1 features are marginally over-represented, largely at the expense of the T0 features.

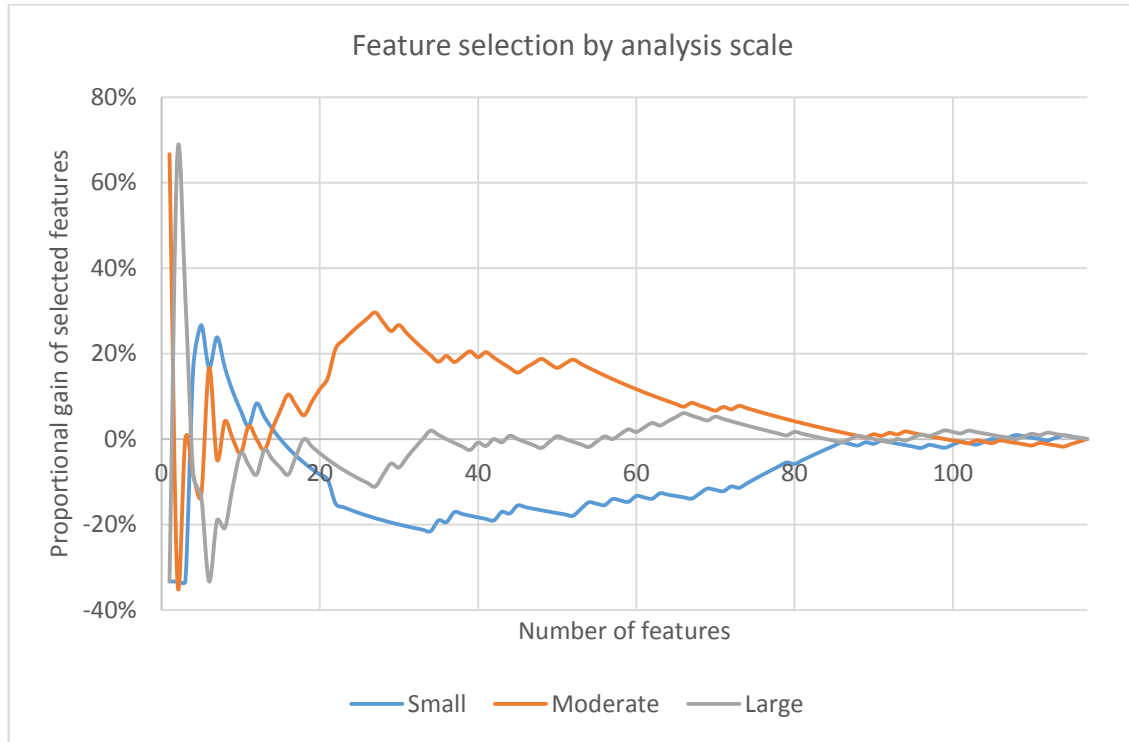


Figure 38: Feature selection by analysis scale. The selection is made on all available features.

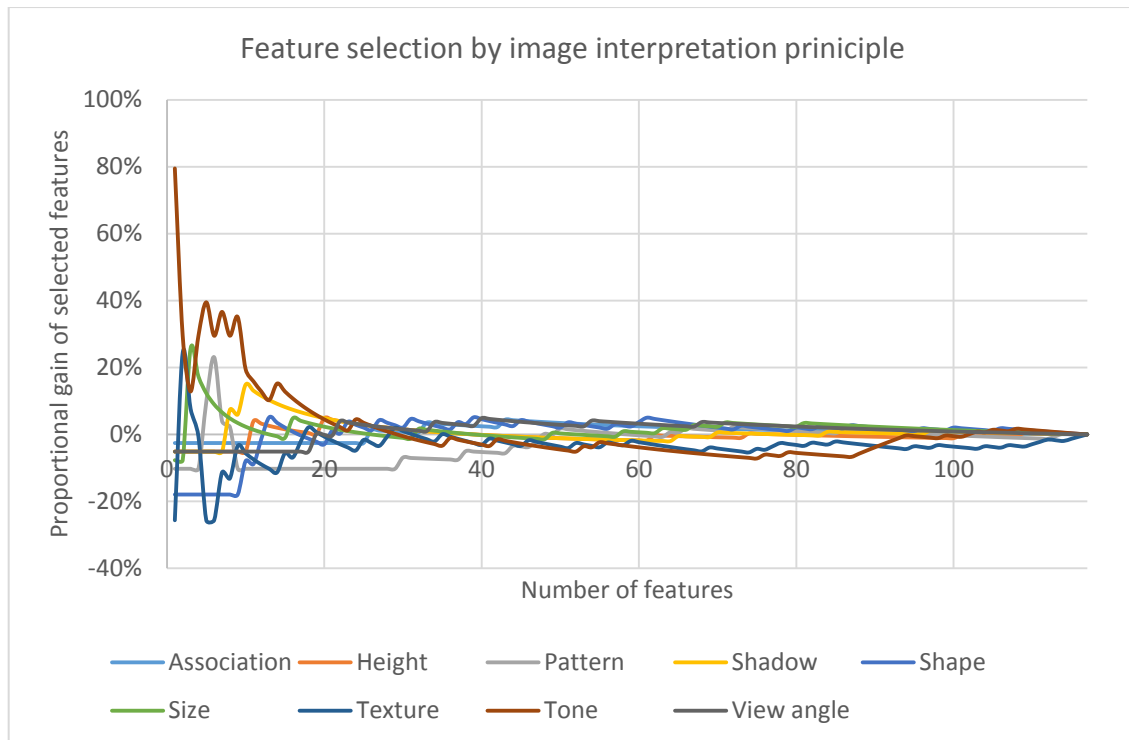


Figure 39: Feature selection by image interpretation principle. The selection is made on all available features.

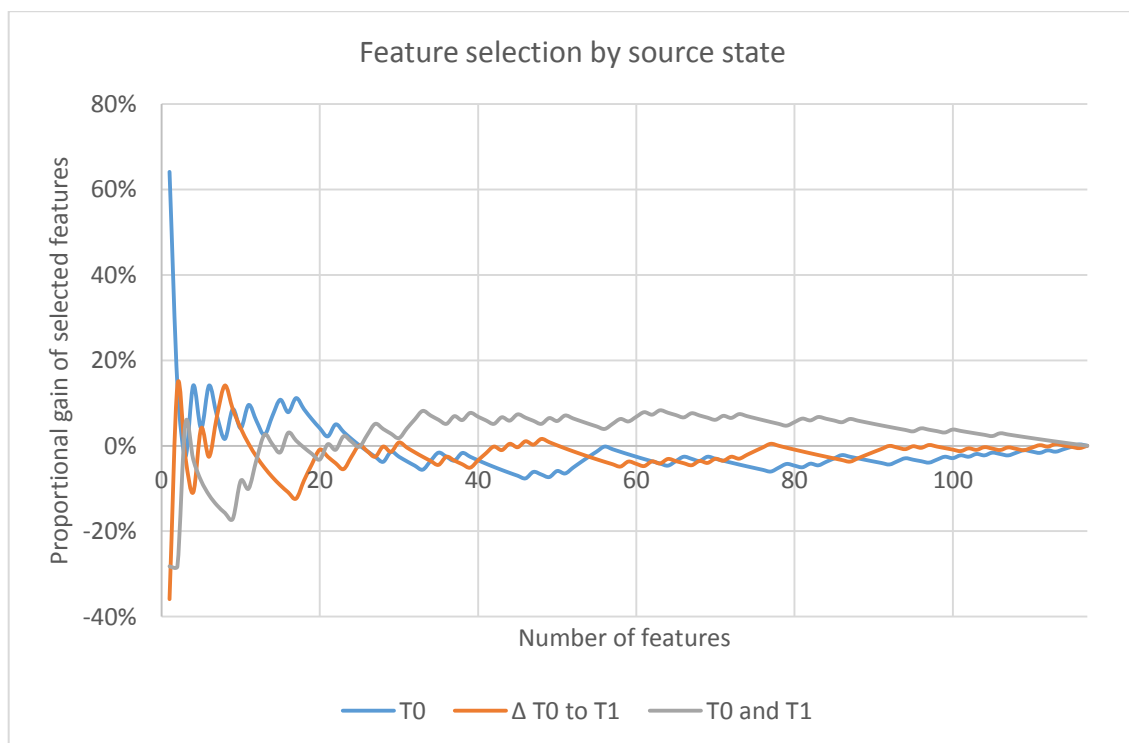


Figure 40: Feature selection by comparable state. The selection is made on all available features.

8.2.2 Difference features

Figure 41 shows the proportional gains, of the scale features, for all difference feature selections. Between 1 and 5 features, moderate-scale dominates the feature selection,

with large and small-scale under-represented. From 5 to 30 features, moderate-scale remains significantly over-represented, while small-scale features are significantly under-represented. On the other hand, the proportion of large-scale features increases in the 5 to 30 range. After 30 features, the selection is approximately proportional indicating reduced data uniqueness.

Figure 42 shows proportional gains, of the six interpretation principles, for all difference feature selection scenarios. Initially, texture dominates the difference feature selection, before oscillating between positive and negative gains, and tending towards under-representation from 20 features. Interestingly, height features are strongly over-represented from 2 to 6 features, before tending towards proportionality. Strikingly, tone remains strongly under-represented up until 5 features, before showing broad swings between positive and negative gain.

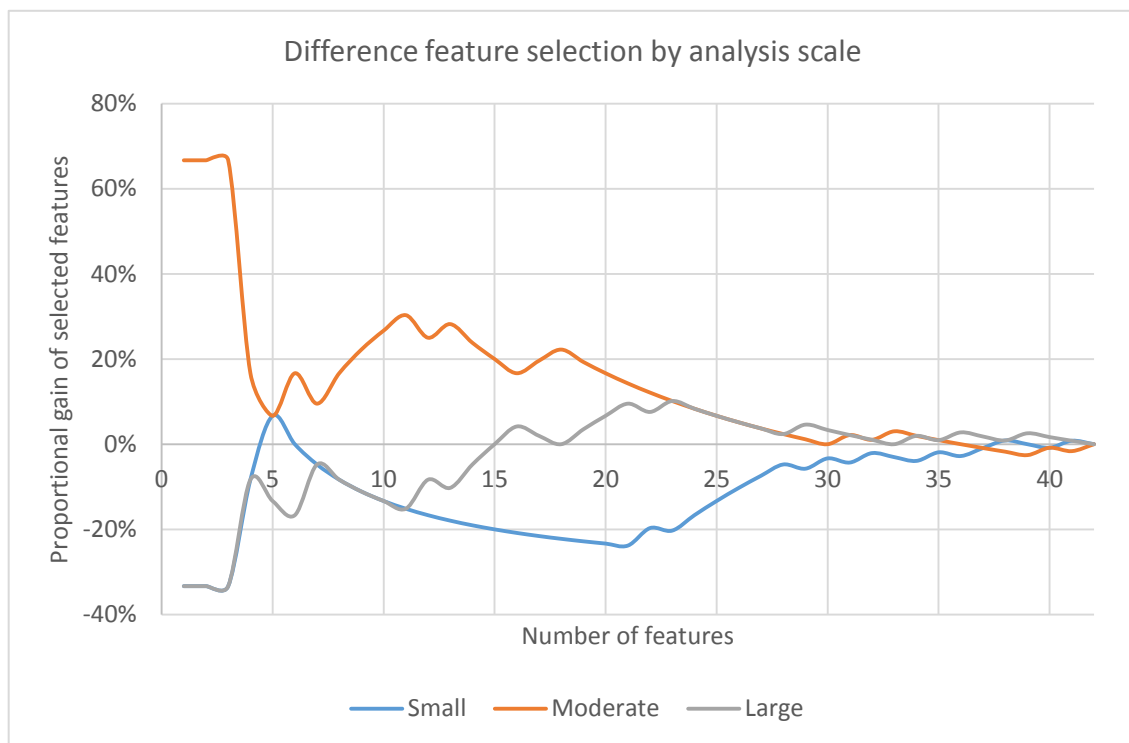


Figure 41: Difference feature selection by analysis scale.

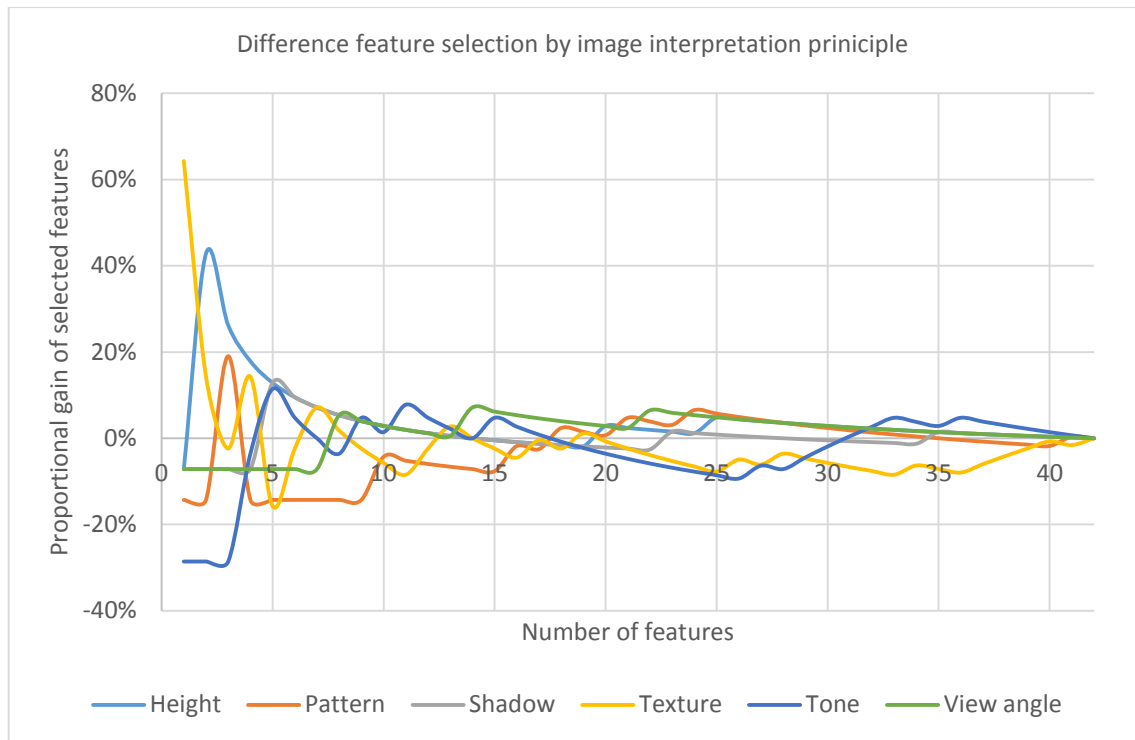


Figure 42: Difference feature selection by image interpretation principle.

8.3 Classification results

This section presents the change detection results, benchmarked against the reference data. The results are presented in three broad categories, overall map accuracy, urban creep accuracy, and a comparison between urban creep and the thematic classes: *other change* and *no change*. The overall map and urban creep accuracy sections highlight the most successful approaches and how CVA performs against a direct classification. The thematic comparison highlights the key differences in classification behaviour among the thematic classes.

8.3.1 Overall map accuracy

The overall map accuracy of all 375 classifications is plotted in Figure 43. Plotted lines depict the accuracy achieved by varying numbers of features, seeded by the unsupervised feature selection. Each line represents the results from of a particular classification, applied to a specific pool of data. Note, that the *direct classification of all features* has a feature pool of 117, compared to 41 difference features. The plotted points show the performance of selected CVA scenarios, such as a simple Red & NIR WC²VA. We can see that there is a general trend of increasing classification accuracy with the number of features. This Observation is particularly pertinent with the classification scenario *direct classification, all features*, which increases rapidly from one to eight features before plateauing. Most of the classification scenarios plateau at approximately eight input features, but gradual improvement is seen beyond eight features. For example, the highest overall map accuracy is seen with 99 features (*direct classification, all features*). We can also see that the spectral CVA scenarios, plotted as points, achieve some of the lowest accuracies.

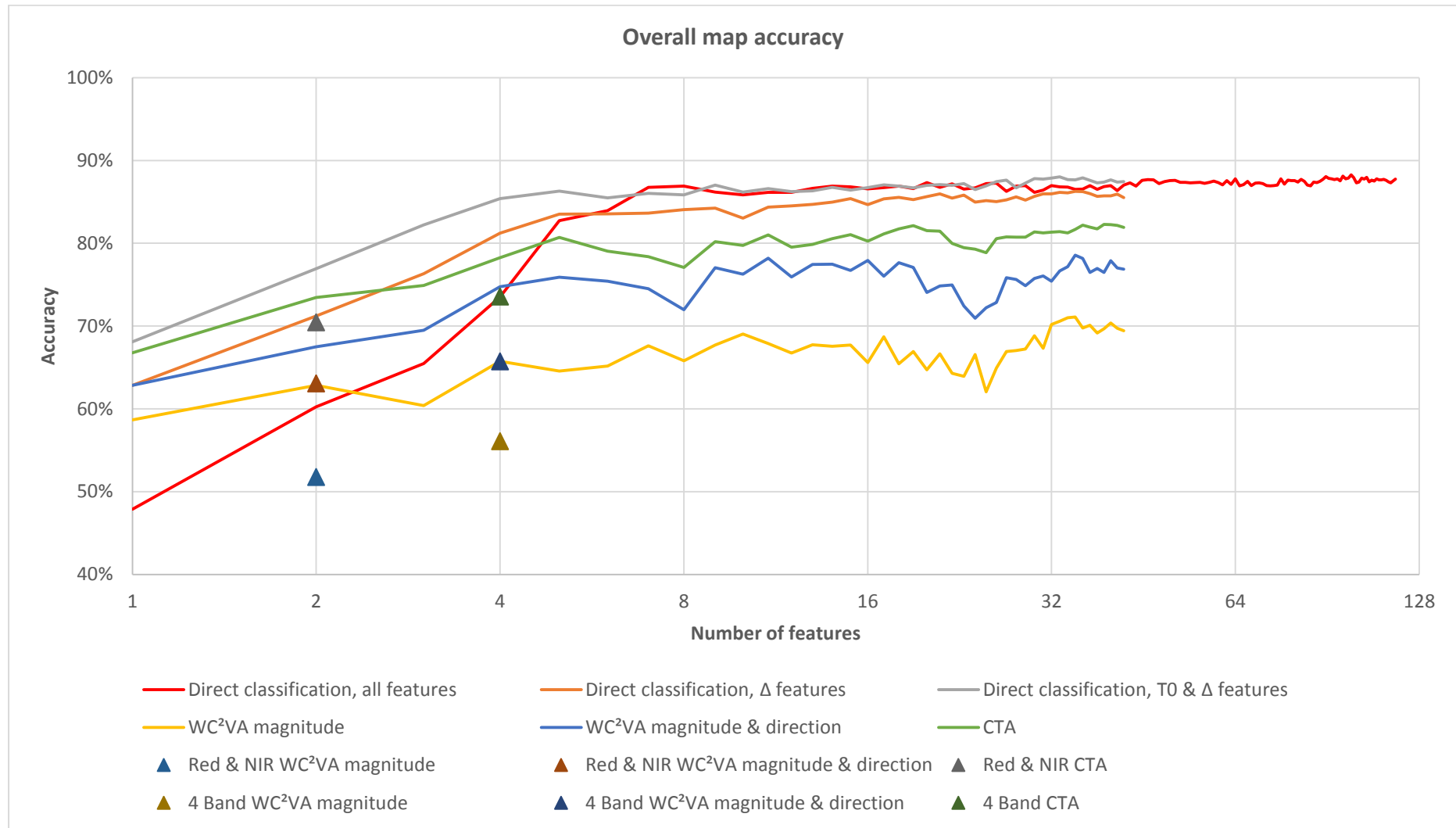


Figure 43: Overall map accuracy.

The maximum overall map accuracy achieved in each of the twelve classification scenarios is summarised in Table 21. The twelve classification scenarios give maximum overall map accuracies ranging from 88.26% (Direct classification, all features), shown in Figure 44, down to 51.78% (Red & NIR WC²VA magnitude), shown in Figure 45. The two classification groups: *Direct classification* and *CVA classification* also show two distinct groups in the results. All of the direct classifications outperform the CVA classifications, while there is only a small difference between the different direct classification scenarios. Conversely, the CVA classifications are the lowest performers and exhibit a large range of different accuracies within the group. Of the CVA results there are three notable features:

1. The simple, spectral approaches gave the four lowest accuracies, with spectral magnitude performing very poorly, giving results as low as 51.78%
2. The use of feature selected object-based features improves CVA performance compared to spectral approaches.
3. CTA (shown in Figure 46) shows a marked improvement over WC²VA in all cases. For example, there is a gain of nearly 8% between the 4 band WC²VA and the 4 band CTA. Furthermore, CTA seeded with object-based features shows a maximum accuracy of 82.29%, which is a very marked improvement over the simple spectral approaches.

Table 21: Overall map accuracy.

Classification group	Classification scenario	Maximum overall map accuracy	
		Accuracy	Number of features
Direct classification	Direct classification, all features	88.26%	99
	Direct classification, T0 & Δ features	88.03%	33
	Direct classification, Δ features	86.27%	35
CVA classification	CTA	82.29%	39
	WC ² VA magnitude & direction	78.55%	35
	4 Band CTA	73.56%	4
	WC ² VA magnitude	71.08%	35
	Red & NIR CTA	70.46%	2
	4 Band WC ² VA magnitude & direction	65.75%	4
	Red & NIR WC ² VA magnitude & direction	63.09%	2
	4 Band WC ² VA magnitude	56.11%	4
	Red & NIR WC ² VA magnitude	51.78%	2

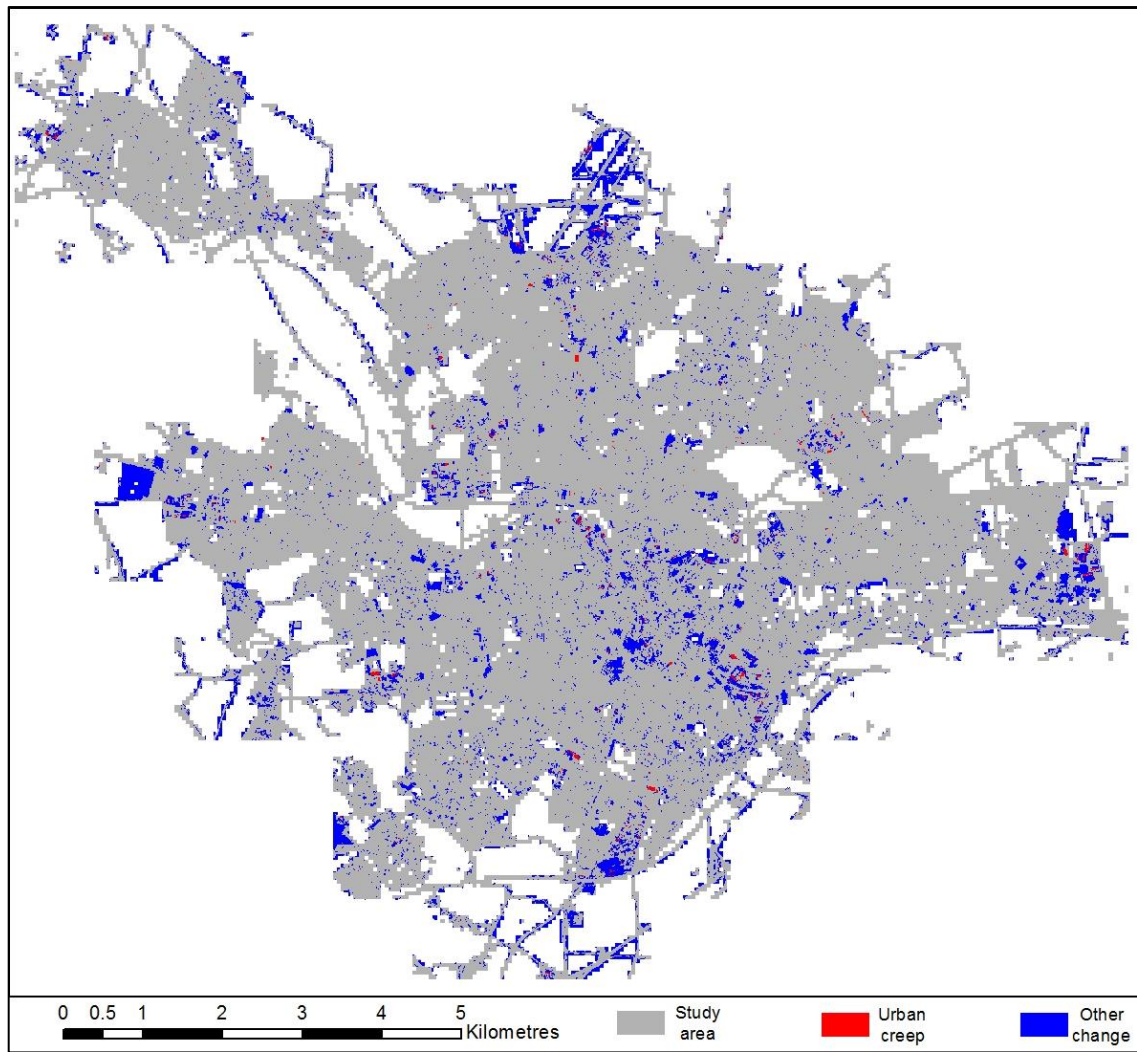


Figure 44: Norwich change map. Based on the direct classification of 99 features.

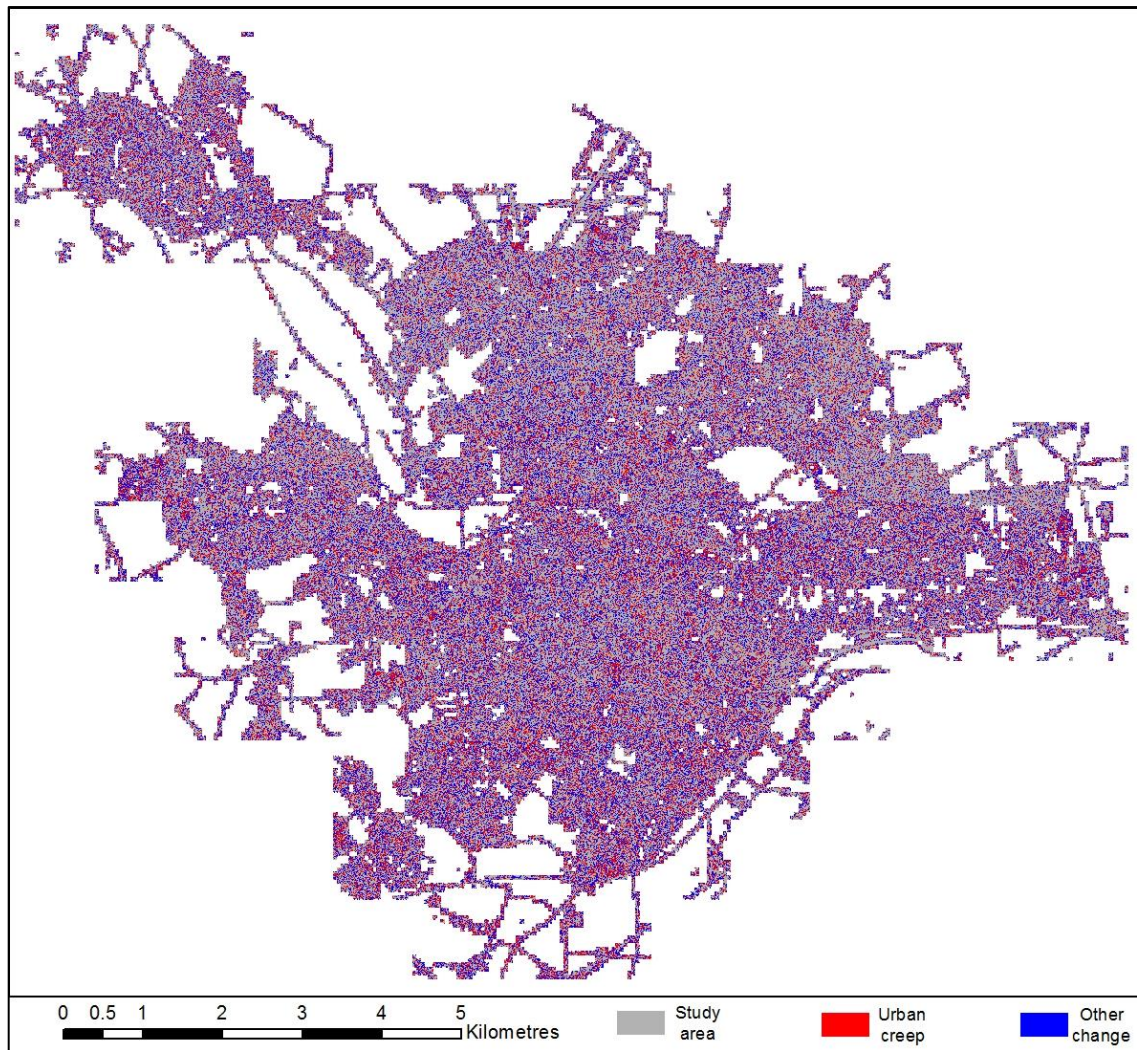


Figure 45: Norwich change map. Based on the CVA classification of Red & NIR WC^2VA .

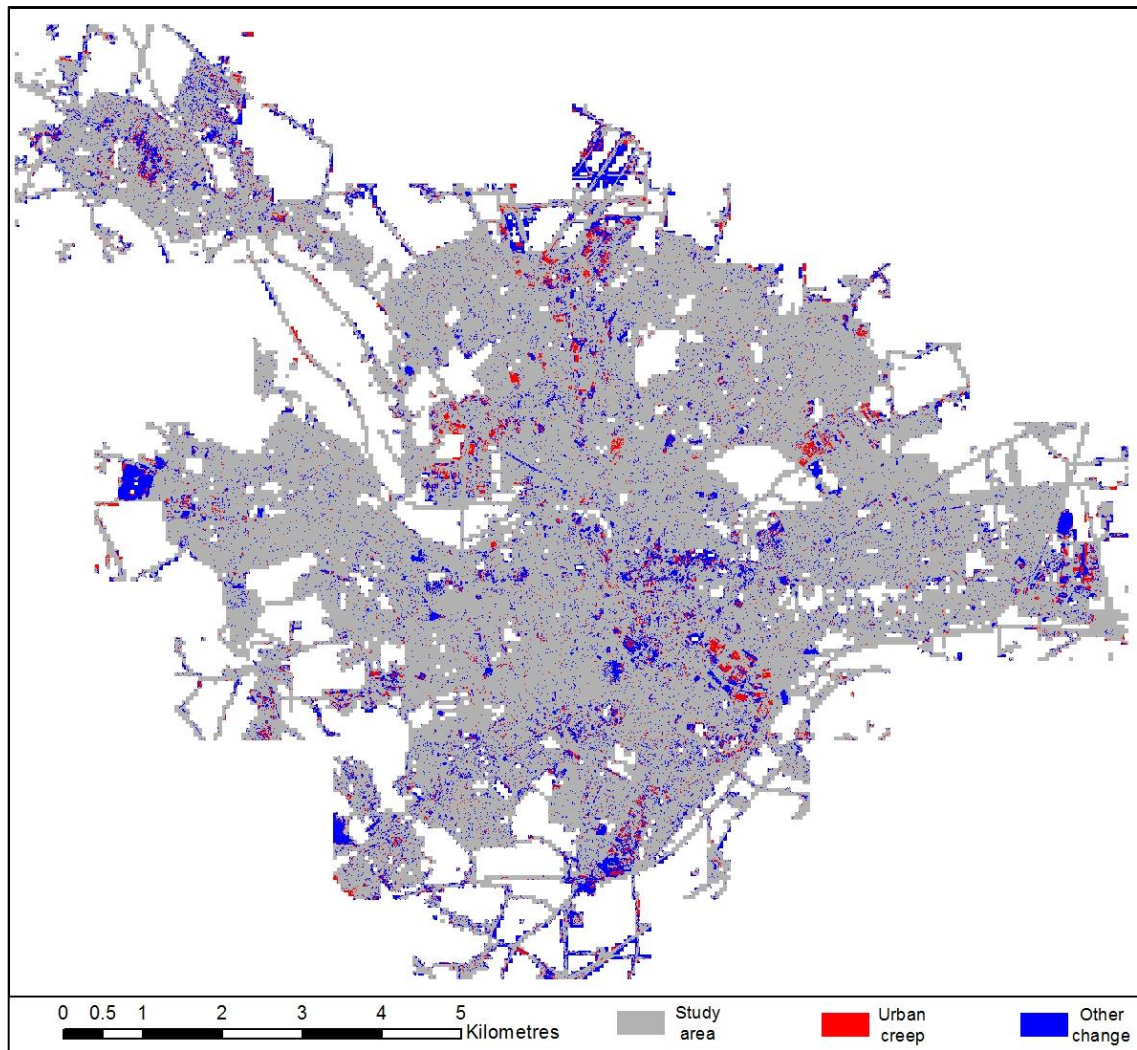


Figure 46: Norwich change map. Based on the CTA classification of 39 features

8.3.2 Urban creep accuracy

When all results are considered, 52.15% is the highest urban creep average accuracy achieved. Figure 47 plots urban creep average accuracies for all classifications conducted. Table 22 ranks and summarises the maximum accuracy, for each classification scenario. We can see that the most successful classification, at 52.15%, is achieved with a direct classification of 62 features, selected from all available features. The least successful classification, at 18.46%, was achieved using the WC²VA magnitude of the red and NIR and bands.

Table 22: Urban creep classification results summary.

Classification group	Classification scenario	Maximum urban creep average accuracy	Number of features
Direct classification	Direct classification, all features	52.15%	62
	Direct classification, T0 & Δ features	49.12%	11
	Direct classification, Δ features	39.92%	41
CVA classification	CTA	34.07%	2
	WC ² VA magnitude & direction	29.47%	1
	WC ² VA magnitude	29.28%	1
	Red & NIR WC ² VA magnitude & direction	28.09%	2
	4 Band WC ² VA magnitude & direction	26.47%	4
	4 Band CTA	25.89%	4
	Red & NIR CTA	24.91%	2
	4 Band WC ² VA magnitude	20.49%	4
	Red & NIR WC ² VA magnitude	18.46%	2

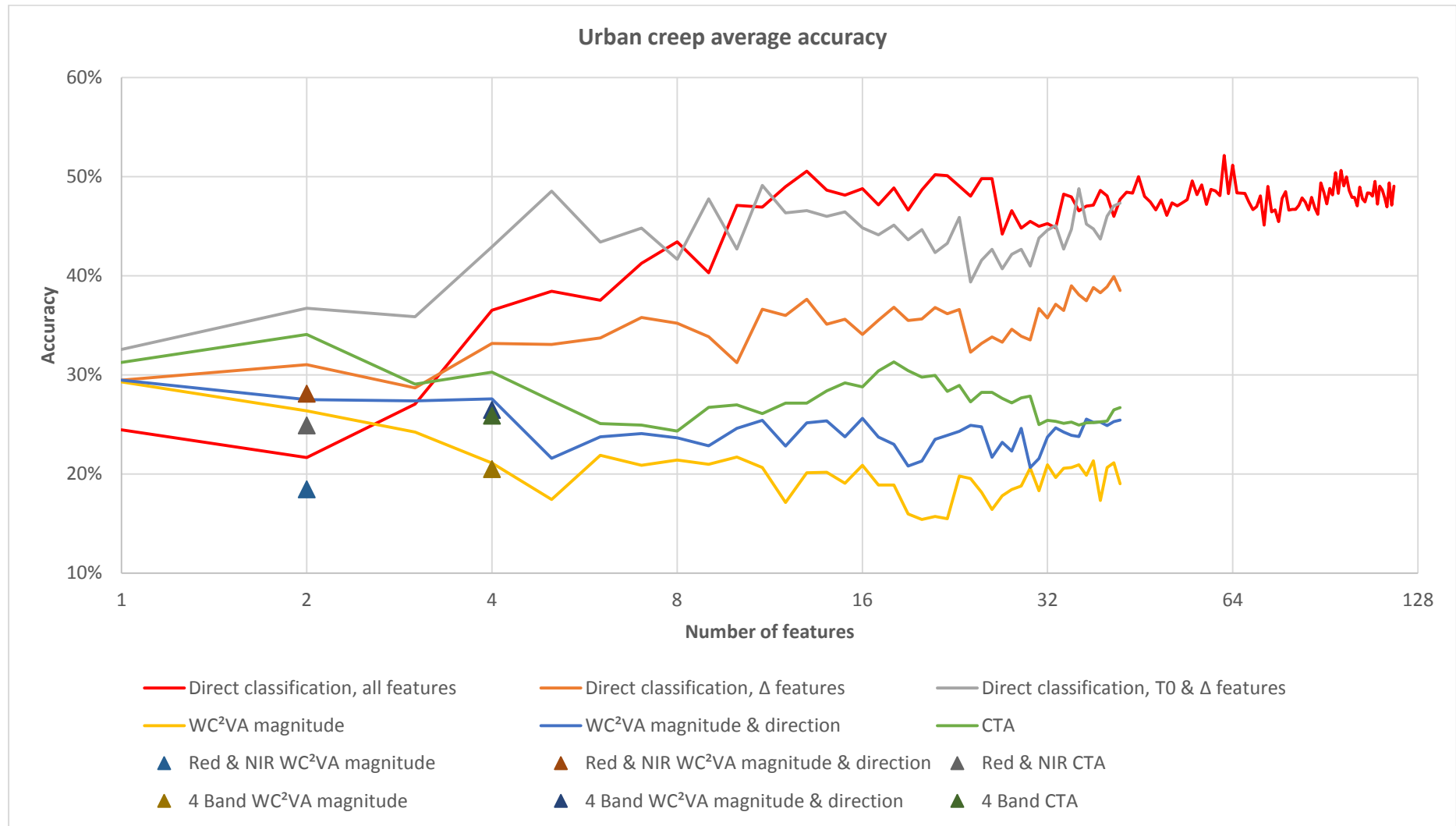


Figure 47: Urban creep average accuracy comparison.

There is a clear divide between the performance of the direct classifications and the performance of the CVA classifications. The direct classifications account for the top three most accurate results (Table 22), outperforming all of the CVA classifications. Similarly, if we consider the two scenario pairs, 'Direct classification Δ features' and 'CVA magnitude & direction', and 'Direct classification T0 & Δ features' and 'CTA', we can see that the direct classification always outperforms the CVA variant, for all feature inputs. Beyond four input features, this difference is typically 10% to 15%. Furthermore, there is almost 34% standing between the most successful direct classification and the Red & NIR WC²VA classification. Furthermore, Figure 47 shows that in general, direct classification accuracy increases with the number of features. Although, the direct classification of all features, displays a pronounced plateau at 13 features. On the other hand, the CVA classifications display a general decline in accuracy as the number of features increases.

The results show that the baseline state, T0, improves urban creep classification accuracy. In other words combining T0 and the Δ features significantly improves urban creep classification accuracy, when compared to the Δ features alone. For instance, Figure 47 shows that including T0 in a direct classification boosts accuracy by as much as 15%. While adding T0 to the best direct difference classification boosts accuracy by nearly 10% (Table 22). We can also see that CTA, integrating T0, consistently outperforms a WC²VA magnitude and direction, albeit with a more moderate gain. The best case CTA classification accuracy is 4.6% higher than the best case WC²VA magnitude and direction classification. There are two notable exceptions. The Red & NIR, and 4 Band CTA classifications are marginally less accurate than the comparable WC²VA classifications. These results show that feature differences do not comprehensively characterise urban creep while the initial conditions are clearly an important part of the description.

The novel CVA methods conceived and implemented in this thesis generally achieve higher urban creep classification accuracy when compared to spectral CVA. Table 22 shows WC²VA and CTA seeded with the unsupervised selection of object-based features outperform the spectral WC²VA variants. Furthermore, Figure 47 shows that the CTA, largely outperforms WC²VA, except in the case of the simple spectral inputs

were WC²VA shows a marginal gain. This selection of results shows that a variety of object-based features generally improves urban creep characterisation over spectral measures.

The results show that in general, gains in urban creep accuracy are made by false positive reduction only. In other words, for a given selection of urban creep candidates, there is the capacity to reduce classification false alarms, but there is limited scope to identify new, previously omitted candidates. Observing Figure 48, we can see that for all cases, the producer accuracy response under varying numbers of features and classification scenarios remains generally flat; fixed close to 40%. With the exception of the Red & NIR WC²VA at 50.91%, producer accuracy peaks at approximately 46%. These producer accuracy results tell us two things: 1) Over half of urban creep is missed from the classification, 2) The missed instances of urban creep remain elusive, no matter how sophisticated the classification. Interestingly, the simple, spectral CVA classifications tend to produce a more inclusive classification, with higher producer accuracy, but more false positives.

User accuracy results show that in general, direct classifications have the capacity to reduce urban creep false positives, while CVA classifications do not. Figure 49 shows that in general, direct classification urban creep user accuracy increases as more features participate in the classification. Therefore, urban creep false positives are reduced as the classification becomes more complex. This effect is particularly strong for the direct classification of all features, which improves rapidly up to thirteen features, before plateauing (Figure 49). On the other hand, the CVA classifications behave differently; with a general decline in user accuracy as more features are considered. Furthermore, we can also observe extremely low user accuracies for the CVA classifications, especially for the simple spectral scenarios. The end user of these CVA classifications would, of course, be overwhelmed by false positives, making them all but unusable. Notably, this situation is improved by applying CTA, which in some cases outperforms a direct classification.

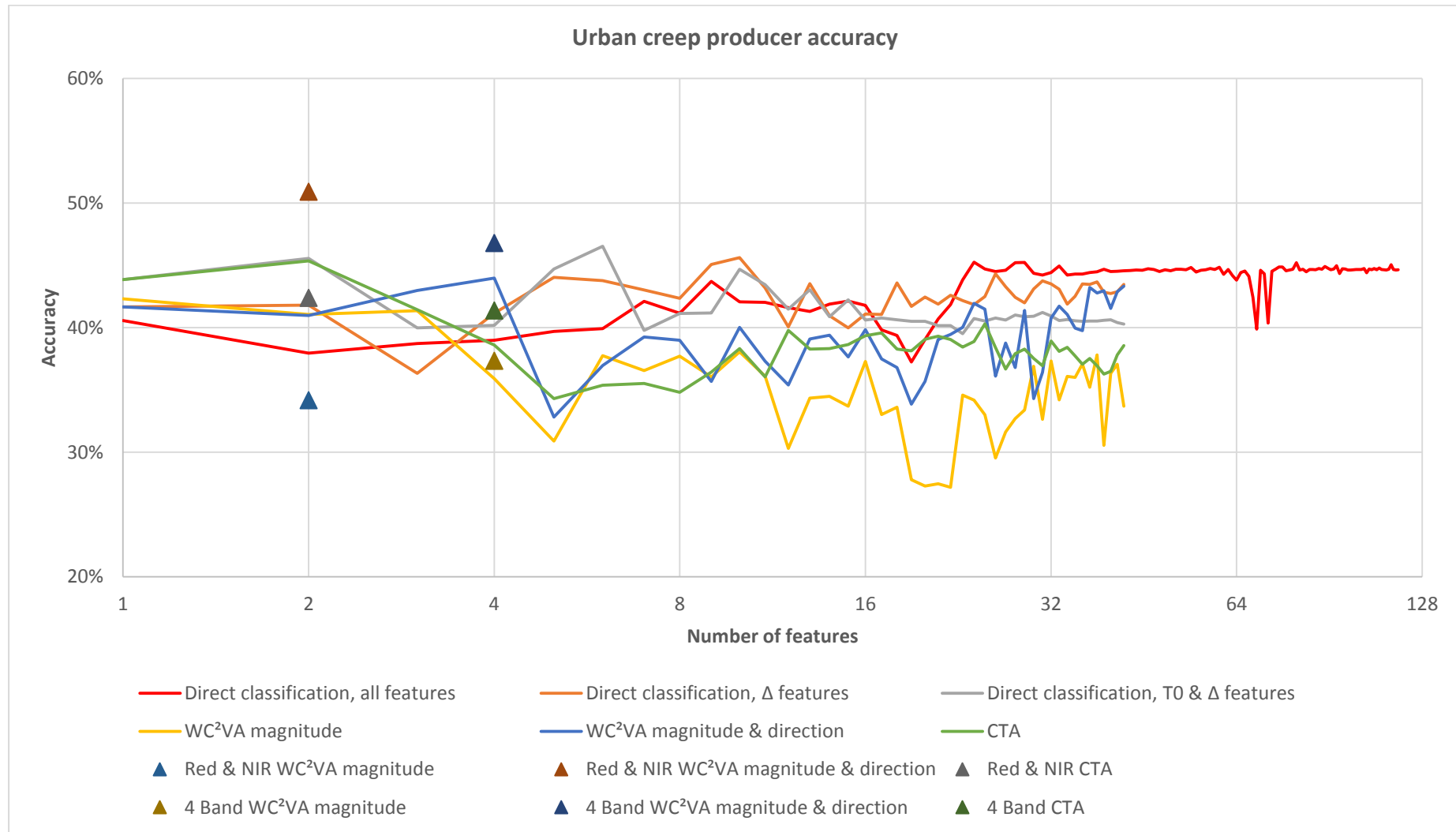


Figure 48: Urban creep producer accuracy comparison.

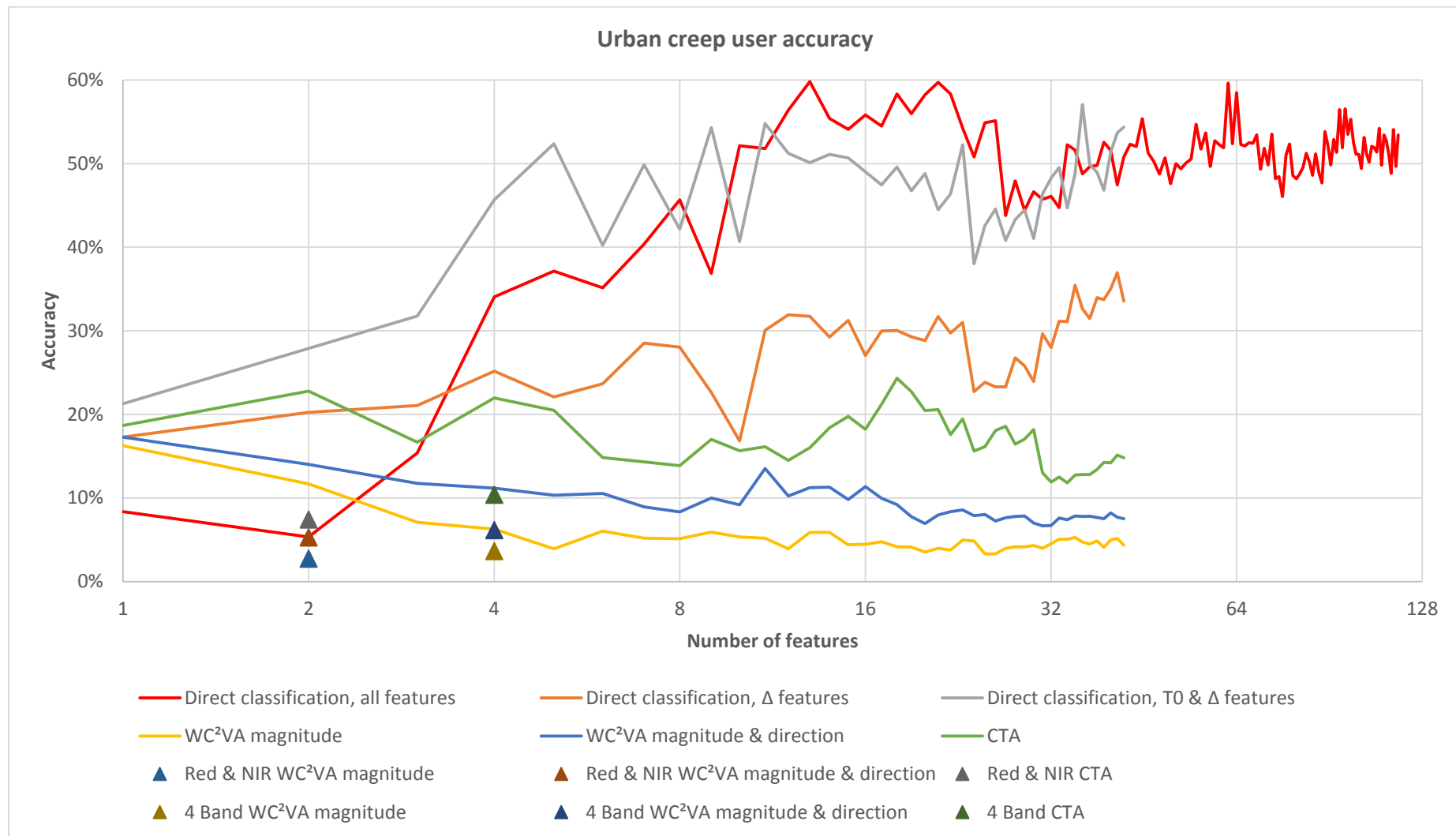


Figure 49: Urban creep user accuracy comparison.

8.3.3 Other change and no change comparison

Overall map accuracy and *other change* identification performance far exceeds that observed with urban creep. The accuracies of *other change* and *no change* are considerably higher than the low to moderate urban creep classification accuracies observed. Figure 50 illustrates, that with a direct classification and 99 input features, the general pattern of *other change* and *no change* is well represented. Figure 51 plots *other change* and *no change* average accuracy for all classification scenarios, while Table 23 summarises the maximum accuracies for each scenario. The maximum average accuracies obtained for *no change*, and *other change*, were 92.89%, and 73.17% respectively, both achieved with a direct classification. Clearly, these are far higher than the most successful urban creep classification average accuracy of 52.15% (Table 22). Comparing urban creep average accuracy response under varying input features (Figure 47), with the response from the other map classes (Figure 51), we see that *other change* accuracy is approximately 25% higher than urban creep. We can also see that the Δ feature direct classification performance is much improved when compared to the urban creep performance. In other words, *other change* is better represented by feature differences than urban creep. CTA performance is also much improved. CTA *other change* average accuracy is typically ~30% higher than equivalent urban creep accuracies. Furthermore, *other change* average accuracies obtained with CTA are typically within 5% to 10% of the comparable direct classifications.

Table 23: Other change and no change summary.

Classification group	Classification scenario	Maximum no change average accuracy		Maximum other change average accuracy	
		Accuracy	Number of features	Accuracy	Number of features
Direct classification	Direct classification, all features	92.89%	99	73.17%	90
	Direct classification, T0 & Δ features	92.64%	33	73.45%	33
	Direct classification, Δ features	91.96%	36	70.07%	32
CVA classification	CTA	89.50%	36	67.57%	36
	WC ² VA magnitude & direction	88.96%	35	56.36%	35
	4 Band CTA	84.73%	4	48.90%	4
	WC ² VA magnitude	84.48%	32	48.89%	33
	Red & NIR CTA	82.68%	2	46.29%	2
	4 Band WC ² VA magnitude & direction	80.54%	4	39.26%	4
	Red & NIR WC ² VA magnitude & direction	79.29%	2	34.99%	2
	4 Band WC ² VA magnitude	74.51%	4	30.51%	4
	Red & NIR WC ² VA magnitude	71.99%	2	28.13%	2

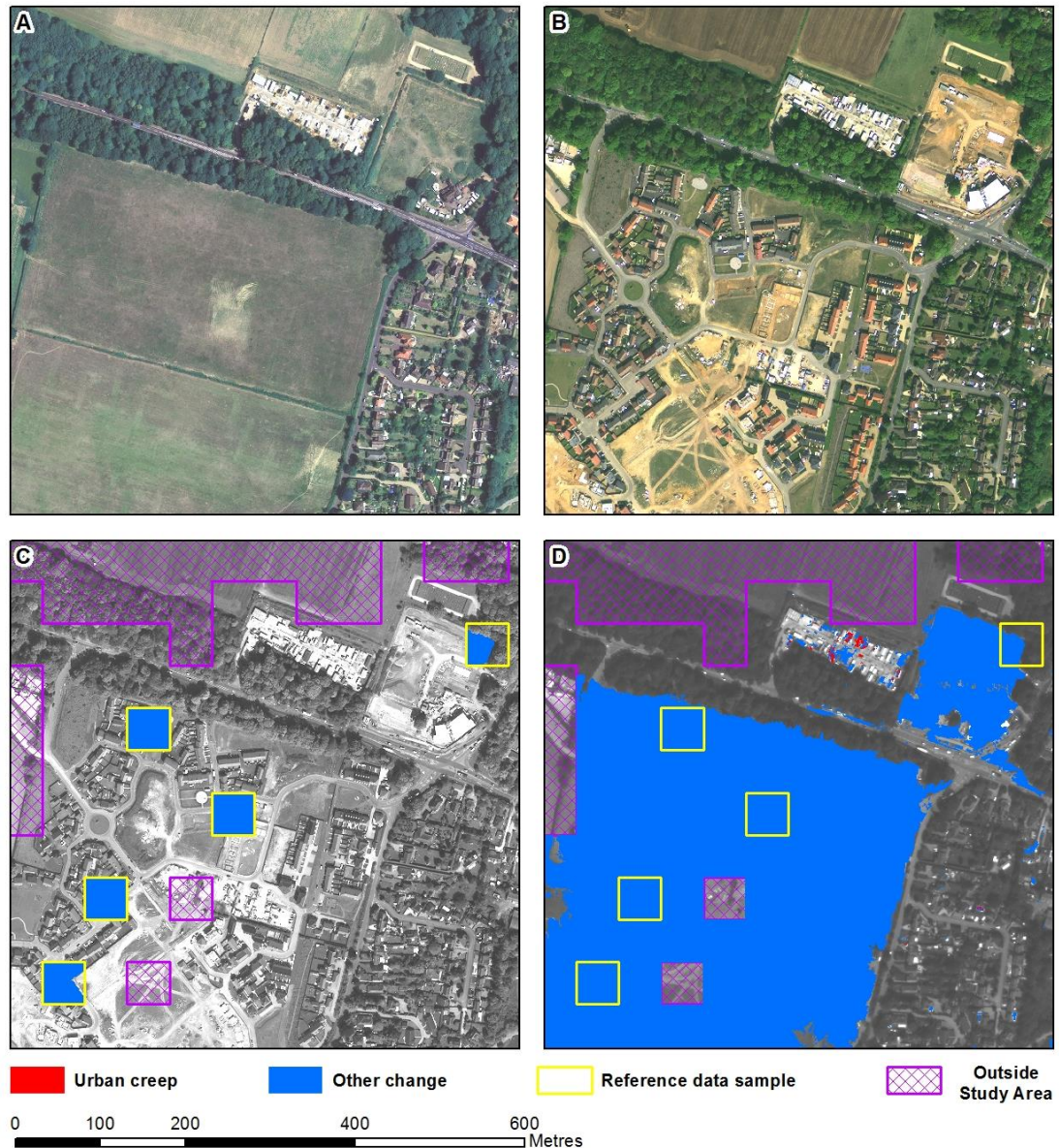


Figure 50: Change classification example. Based on the direct classification of 99 features. Panel A shows the 2006 imagery. Panel B shows the 2010 imagery. Panel C shows the reference data overlaid over the 2006 imagery. Panel D shows the remote sensing classification overlaid over the 2010 imagery.

The results show that the classes *other change* and *no change* show a stronger CVA response to additional classification features than urban creep. For instance, urban creep direct classification user accuracy tends to increase as the number of input features increase; while CVA classification accuracy marginally decreases (Figure 47). Conversely, Figure 51 shows a trend where *other change*, classification accuracy increases with more classification features in all cases. CTA, in particular, shows a marked improvement in performance when compared to urban creep.

A marked difference in the producer accuracy response between the two change classes was observed. Recalling the results from section 8.3.2, urban creep producer accuracy remains relatively stable (Figure 48). Conversely, Figure 52 shows that in general, *other change* producer accuracy increases when more features are added. WC²VA magnitude, and magnitude and direction are the exception to this trend. This broad trend in producer accuracy increase indicates that *other change* omission can be reduced with increasingly complex feature sets.

The novel CVA methods defined in this thesis considerably improve *other change* classification performance compared to simple spectral CVA methods. For instance, considering the four simple spectral CVA classifications in Figure 51, namely the Red & NIR, and 4 band, WC²VA classifications, located in the lower left of the plot. We can see that they are the poorest performing CVA classifications. In other words, the simple spectral CVA variants are all outperformed by CTA, and object-based feature seeded WC²VA and CTA.

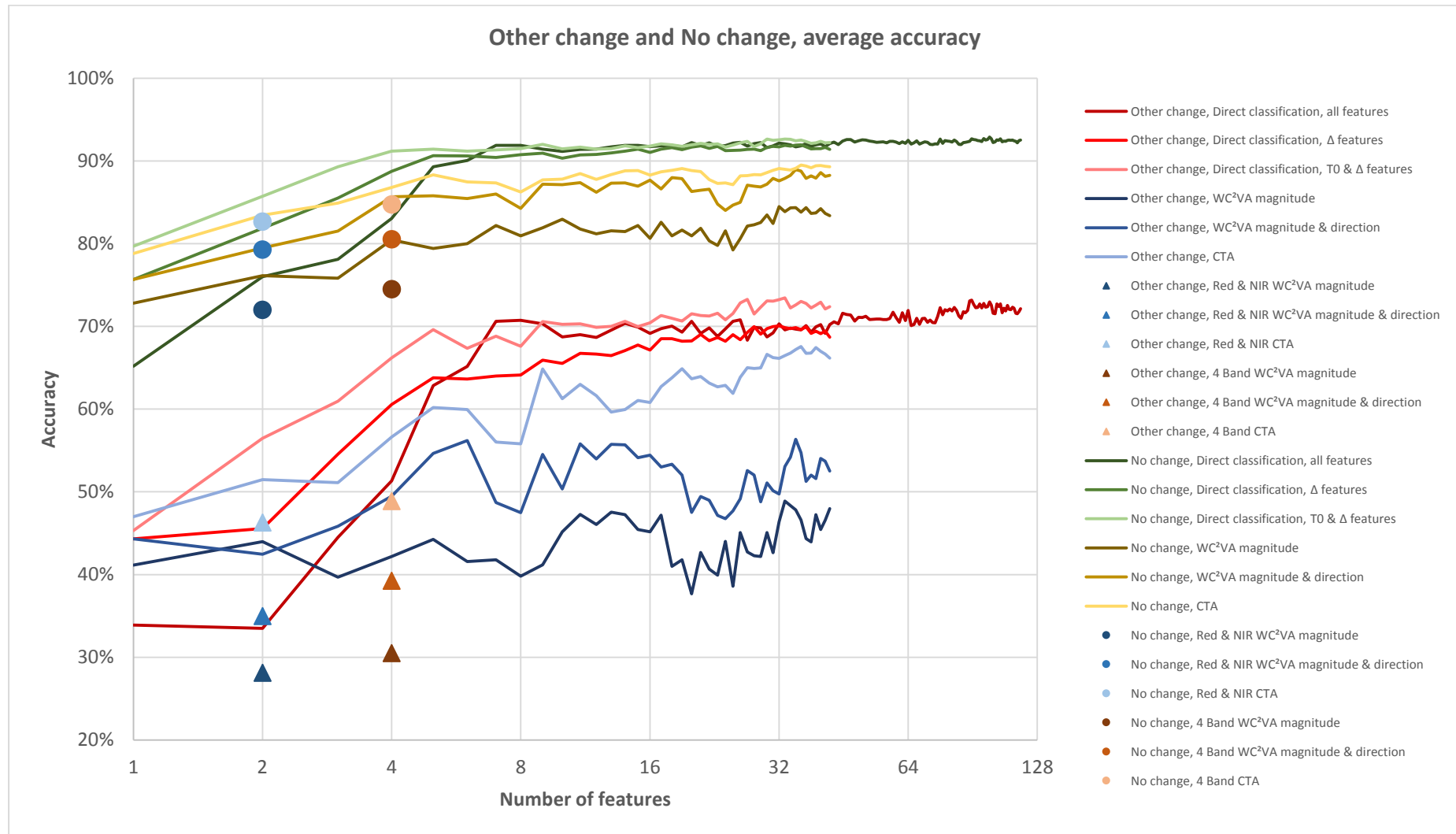


Figure 51: Other change and no change average accuracy comparison.

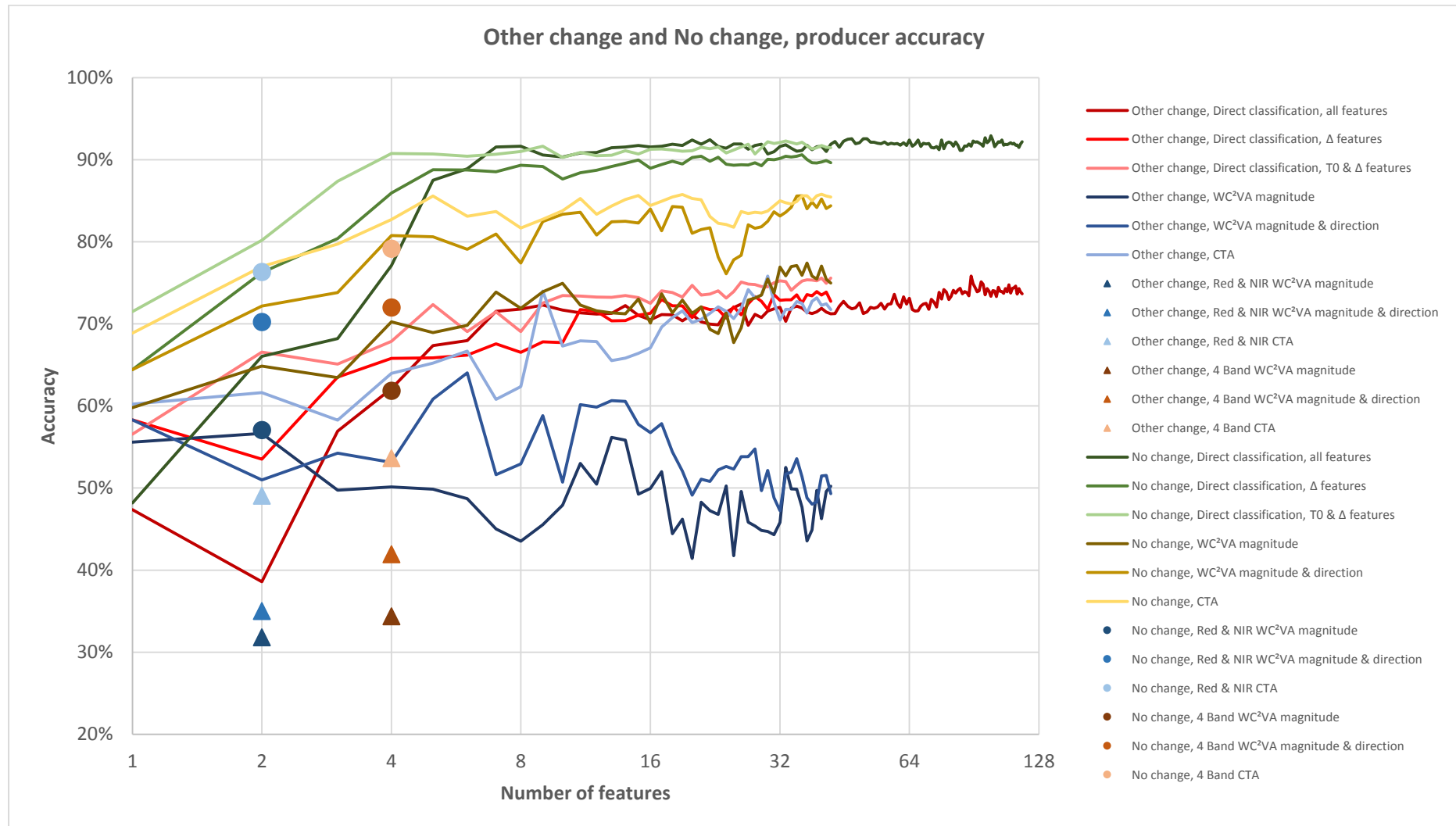


Figure 52: Other change and no change producer accuracy comparison.

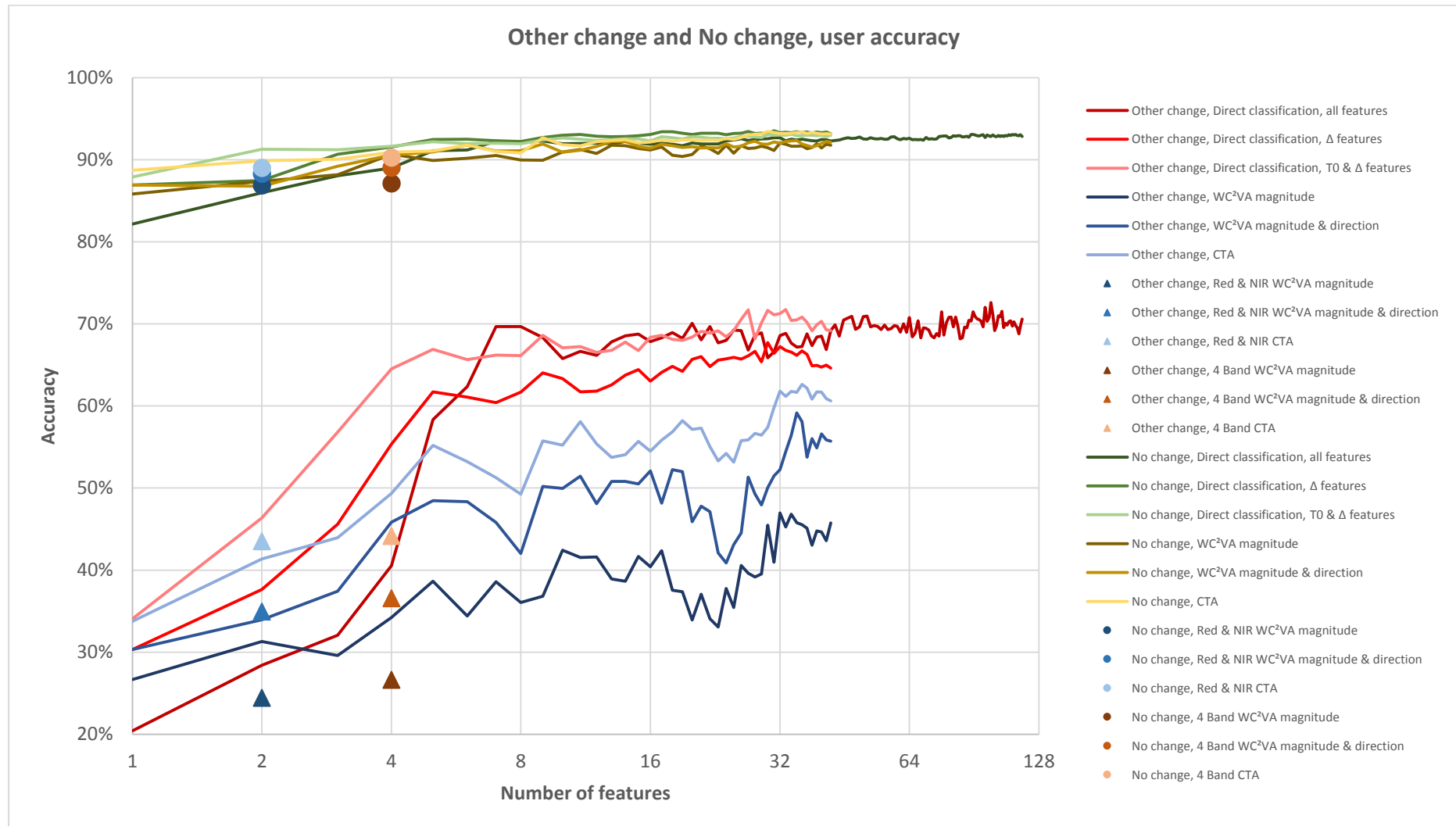


Figure 53: Other change and no change user accuracy comparison.

8.4 Best case classification

This section explores the results of the most successful urban creep classification and identifies the main sources of confusion, and the most important aspects of the classification. The most successful classification was the direct classification of 62 features, from all available features, achieving an urban creep average accuracy of 52.15%. Table 24 contains the error matrix for this classification. Overall, we can see a hierarchy of accuracy. *No change* obtains a high accuracy of 92.33%, *other change* is moderate at 71.66%, while urban creep accuracy is low at 52.15%. There is also a marked difference in urban creep producer and user accuracies, signifying more errors of omission, than errors of commission. There is widespread confusion between urban creep and *no change*. Approximately 40% of the urban creep reference data is incorrectly classified as *no change*, while around 15% is incorrectly classified as *other change*. On the other hand, approximately 34% of the classified urban creep, is, in fact, *no change*, while approximately 6% is *other change*. There is also significant confusion between *other change* and *no change*. Approximately 26% of the *other change* reference data is incorrectly classified as *no change*, while around 0.4% is incorrectly classified as urban creep. On the other hand, approximately 29% of the classified *other change*, is, in fact, *no change*, while approximately 1% is urban creep. Figure 54 shows an example where the urban creep classifications has performed well. Figure 55 shows an example where the instance of urban creep is well classified, but there are *other change* false positives. Figure 56 illustrates an example where urban creep and *other change* are confused with *no change*.

Table 24: The error matrix of the most successful classification. The most successful classification is a direct classification of 62 features, from all available features. Matrix quantities are recorded as study area proportions.

		Reference			
		No Change	Other change	Urban creep	Total
Classification	No Change	0.7260	0.0507	0.0075	0.7841
	Other change	0.0577	0.1414	0.0028	0.2020
	Urban creep	0.0048	0.0008	0.0083	0.0139
	Total	0.7885	0.1929	0.0186	1.0000

Producers accuracy	92.07%	73.31%	44.66%
Users accuracy	92.59%	70.01%	59.64%
Average accuracy	92.33%	71.66%	52.15%

Overall accuracy	87.57%
------------------	--------

Figure 58 shows the random forest classification importance for all features individually and aggregated semantically into image interpretation principle, analysis scale, and source state. The random forest classification importance describes the proportion of decisions in the tree that are made using a particular feature. In other words, it shows how often a feature is used as a discriminator in the decision tree, and therefore how important it is to the overall classification. The moderate scale, green image difference, $\Delta L2\ Green$, is deemed the most important feature. Interestingly, object-based measures of spatial pattern are the second and fifth most important features. Small and moderate-scale texture measures are third and fourth most important respectively. Interestingly, of the texture measures selected, the simple, and computationally lightweight standard deviation measures, are generally more important than the complex, and computationally intensive GLCM measures.

Tone, texture, shape and pattern are by far the most important image interpretation principles. Tone and texture are deemed the most important, with almost identical

values of 0.257 and 0.255 respectively (Figure 19). Multi-temporal segmentation shape was shown to be the third most important, while interestingly, the object-based measures of spatial pattern are the fourth most discriminate features. The remainder of the interpretation principles, height, size, view angle, association, and shadow all have relatively low importance of 0.019 to 0.041. Disappointingly, the 3D modelled measures of view angle and shadow are not shown to be hugely important, but do contribute to the overall classification result.

Moderate-scale features such as $\Delta L2\ Green$ and $\Delta L2\ GLCM\ Entropy$, are overwhelmingly the most important to the classification. Moderate-scale features sum to an importance of 0.605, compared to 0.247 and 0.148 for small and large-scale features respectively. For the source state, feature differences are the most important, at 0.423. However, T0 and the shared features, T0 & T1, still maintain relatively high importance at 0.312 and 0.265 respectively. This last result indicates that, in this instance, feature difference alone does not comprehensively discriminate change.



Figure 54: Urban creep classification example. Based on the direct classification of 62 features. Panel A shows the 2006 imagery. Panel B shows the 2010 imagery. Panel C shows the reference data overlaid over the 2006 imagery. Panel D shows the remote sensing classification overlaid over the 2010 imagery.



Figure 55: Urban creep classification example. Based on the direct classification of 62 features. Panel A shows the 2006 imagery. Panel B shows the 2010 imagery. Panel C shows the reference data overlaid over the 2006 imagery. Panel D shows the remote sensing classification overlaid over the 2010 imagery.

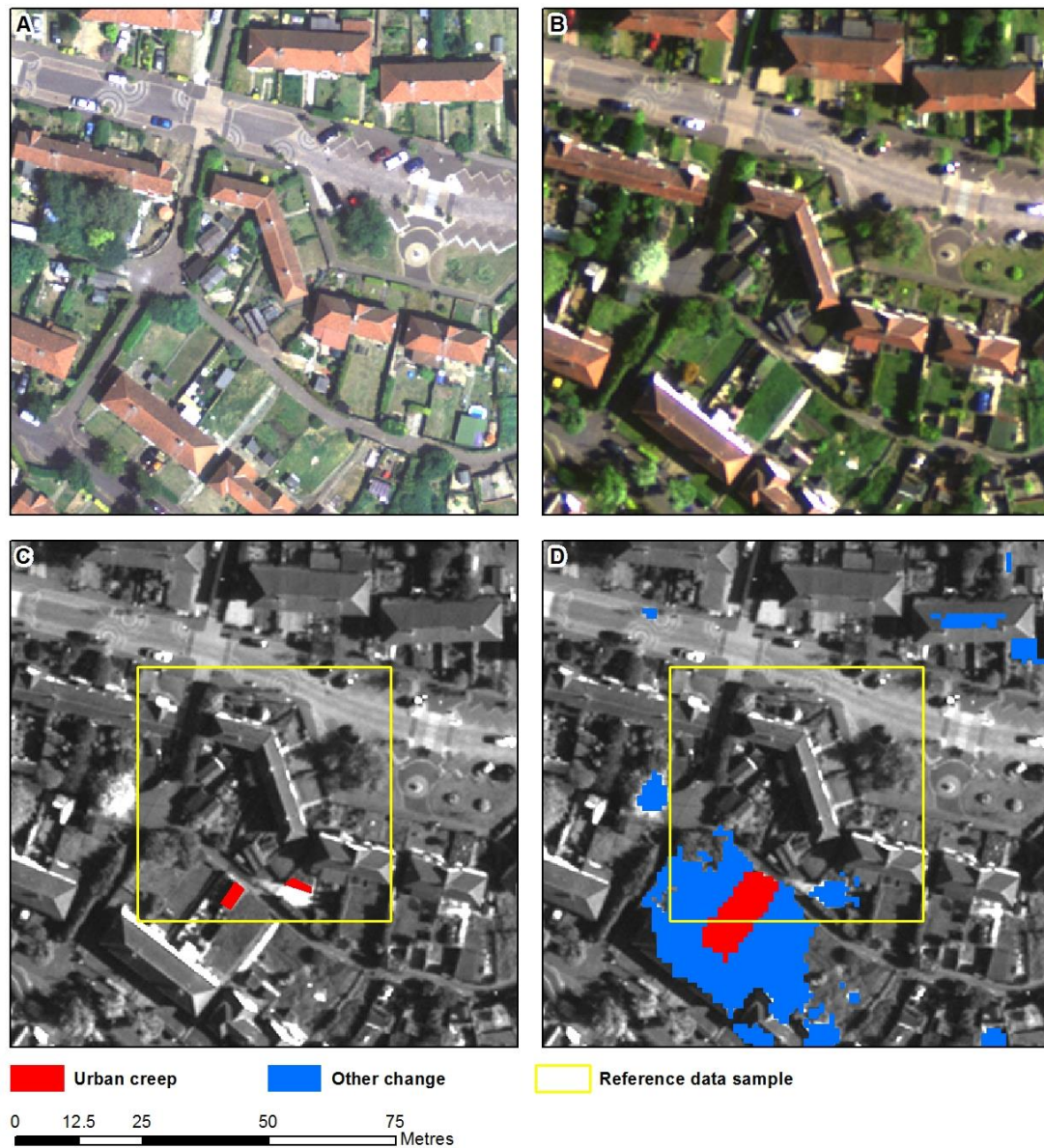


Figure 56: Urban creep classification example. Based on the direct classification of 62 features. Panel A shows the 2006 imagery. Panel B shows the 2010 imagery. Panel C shows the reference data overlaid over the 2006 imagery. Panel D shows the remote sensing classification overlaid over the 2010 imagery.

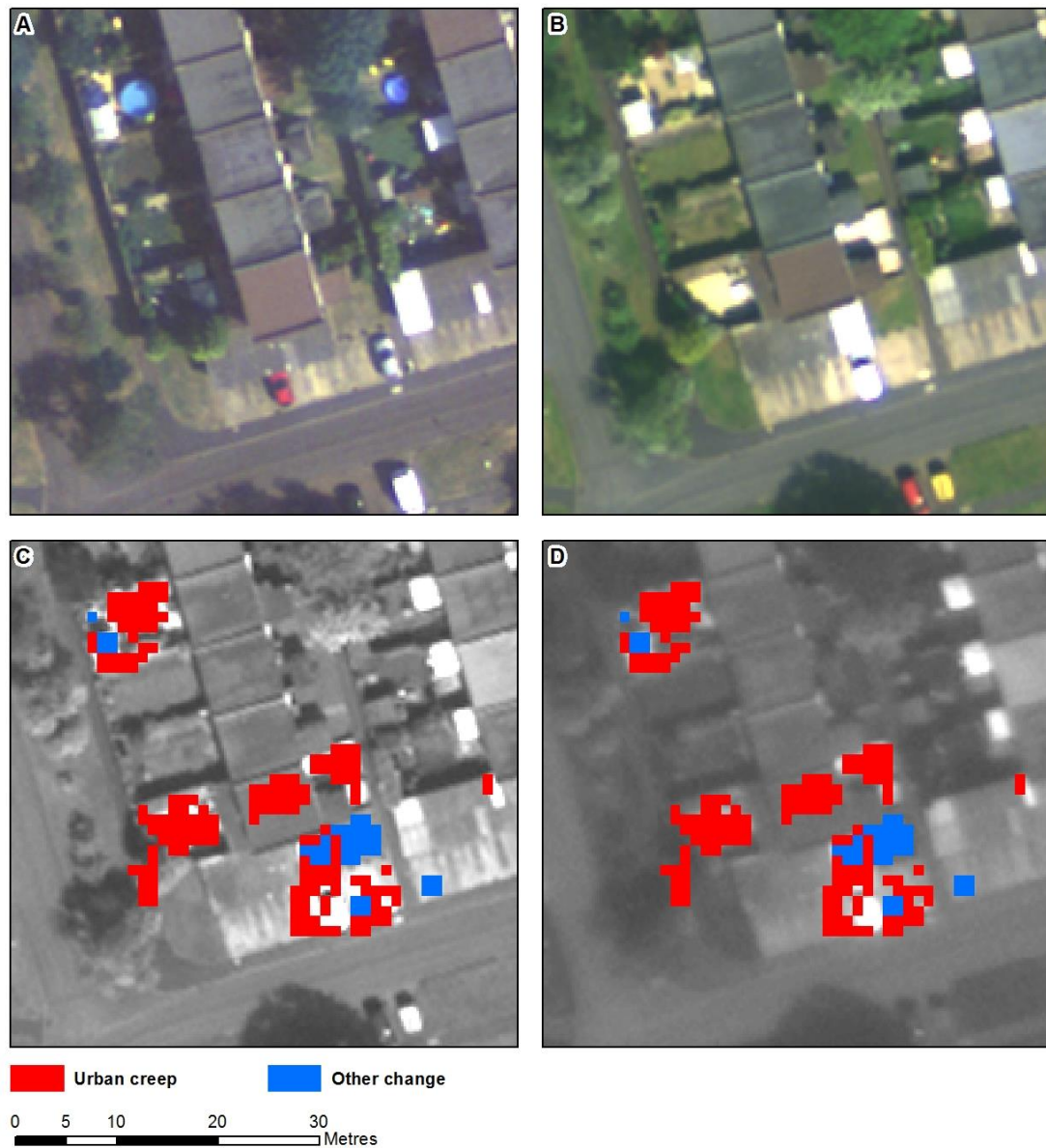


Figure 57: Urban creep classification example discovered from the classification. Based on the direct classification of 62 features. Panel A shows the 2006 imagery. Panel B shows the 2010 imagery. Panel C and D show the remote sensing classification overlaid over the 2006 and 2010 imagery respectively.

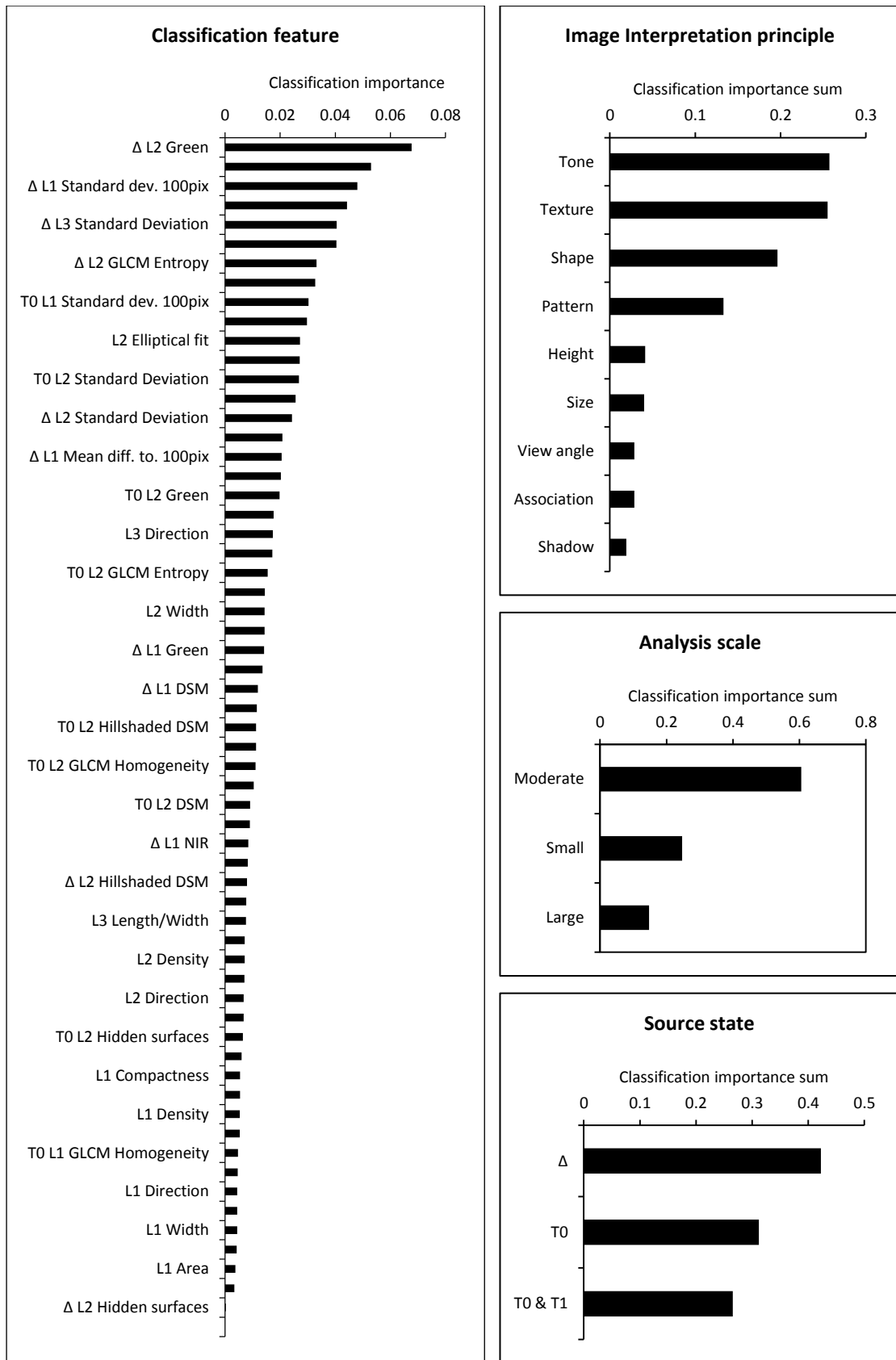


Figure 58: Classification feature importance and semantic summary. The left-hand panel displays the ranked random forest classification importance, by feature. The panels on the right-hand side show summaries of the features semantic content to include, image interpretation principle, scale and source data state. Semantic summaries plots display the feature importance sum of all the features within a particular semantic group.

8.5 Urban creep and CVA

This section presents results illustrating the relationship between urban creep and image differences, or change vectors. Firstly, the magnitude response of a range of different CVA scenarios is compared against the reference data to establish if high change magnitude is a defining characteristic. Secondly, change vector direction and magnitude are assessed qualitatively to evaluate urban creep discrimination. Six CVA scenarios were chosen for this evaluation. Namely: the two most successful CVA feature selections concerning urban creep accuracy (1 and 2 features), the two most successful CVA feature selections concerning overall accuracy map accuracy (39 and 35 features), and the two spectral CVA inputs (Red & NIR, and 4 band).

8.5.1 CVA magnitude

The results indicate that the CVA magnitude response of urban creep has more in common with unchanged areas than different types of change. Figure 59 plots CVA magnitude histograms for six different input combinations. The histograms are prepared by class, and supplemented by a correlation matrix, reporting the Pearson's r score between each histogram pair. Therefore, values of 1 indicate identical histograms and 0, uncorrelated histograms. Visually, we can see that the *other change* histograms appear distinct from urban creep and *no change*. The distinction is especially present in the 35 and 39 feature histograms. For the Red & NIR, and 4 band magnitudes, urban creep and *other change* are very similar to one another, with Pearson's r scores of 0.93 and 0.92 respectively. However, urban creep remains very similar to *no change*, with a Pearson's r score of 0.92 in both cases. Crucially, the remainder of the correlation matrixes show urban creep has a higher correlation to *no change* than *other change*. Urban creep to *no change* correlation is 0.92 or above in all cases, whereas *other change* to *no change* correlation ranges from 0.48 to 0.79. Interestingly, the 35 & 39 feature cases, and to a less extent, 1 & 2 feature magnitudes display bimodal distributions, with the second mode appearing at approximately 4 to 5 standard deviations.

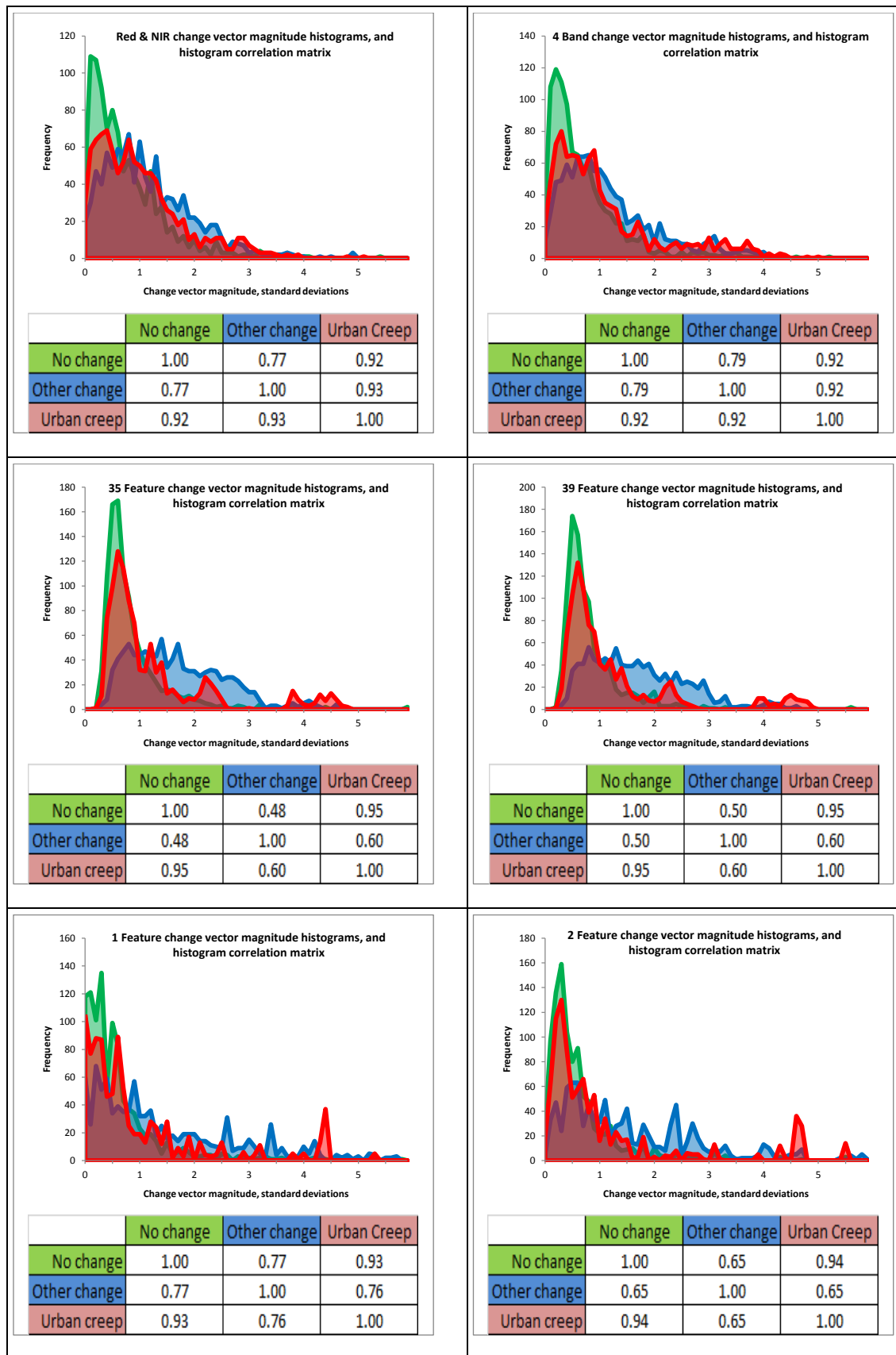


Figure 59: WC²VA magnitude histogram comparison by class. *No change*, *other change* and *urban creep* histograms are plotted in green, blue and red respectively. A correlation matrix accompanies each set of histograms, showing Pearson's *r* correlation scores for all histogram pairs. The upper left panel displays the Red & NIR CVA magnitude histograms; the upper right, 4 band; middle left, 35 feature; middle right, 39 feature; lower left, 1 feature; and lower right, 2 features.

8.5.2 CTA magnitude and direction

This sub-section briefly explores vector magnitude and direction as a unique indicator of urban creep. A full CTA analysis considers baseline and change vector magnitude and direction. Plotting these vectors on polar plots reveals any underlying trends and anomalies. Specifically, we are looking for anomalies unique to urban creep which, may not be confused with *other change* or *no change*. Here, three representative CTA configurations are described, namely Red & NIR, Four bands, and 39 features.

A qualitative assessment of the Red & NIR and Four band configurations yields limited unique anomalies describing urban creep. Figure 60 and Figure 61 show the T0 and CV polar plots for the Red & NIR, and Four band configurations. All of the CV plots show the background, stable image transitions as heavy clustering at low magnitudes, with image changes occurring at greater magnitudes, further out from the centre of the polar plot. For both Red & NIR and Four band configurations, there is a broad cluster in the urban creep change vector plots around 30° , extending from a magnitude of 1.5 to 4.5. This cluster intensifies in the four band case. Figure 62 shows an example of urban creep from this cluster. However, Figure 60 and Figure 61 show that this cluster is not unique to urban creep. Specifically, the CV plots for *no change* and *other change* exhibit similar clusters. Therefore, in this case, there are limited unique anomalies.

A qualitative assessment of the 39 feature configuration yields a unique anomaly linked to a specific type of urban creep. Figure 63 displays the polar plots for the 39 feature CTA. We can see that there is a strong cluster in the urban creep change vector polar plot, centred on approximately 55° and a magnitude of 4.5. Note, that this cluster is the second mode of the histogram described in section 8.5.1 (Figure 59, middle right). There is partial overlap with a cluster in the *other change* plot. Figure 64 shows an example of this cluster. In this case, a large building has been extended using very bright roof materials, which in turn results in a very large magnitude. There are several similar building extensions present in the reference data, which contribute to form this cluster. This particular example of urban creep is a large, strong image expression. Therefore, it can exhibit a considerable change magnitude in many of the 39 features, which in turn compound together effectively. However, this example is not typical of the majority of urban creep features, which do not exhibit strong

compounded magnitudes. For example, even the very distinct example illustrated in Figure 65 does not exhibit extraordinary magnitude and does not fall in unique CTA feature space, when compared to the other change plots in Figure 63. In light of this, it can be concluded that the apparently unique CVA clustering observed is a result of a very specific change exhibiting extraordinary magnitude.

The high dimensional CTA analysis more effectively represents multi-scale change. Observing the change magnitude and direction patterns in Figure 62, we see a very detailed, but noisy result. The level of detail is because the four band CTA utilises large-scale spectral features only. Figure 65 shows 39 feature CTA over the same location. Consequently, the magnitude and direction are now a combination of features from three different scales. In this example, we achieve a robust multi-scale visualisation. Detailed change features, such as cars, are clearly expressed, while noise and artefacts are significantly reduced. From a user perspective, these CVA results appear far clearer than the established spectral techniques.

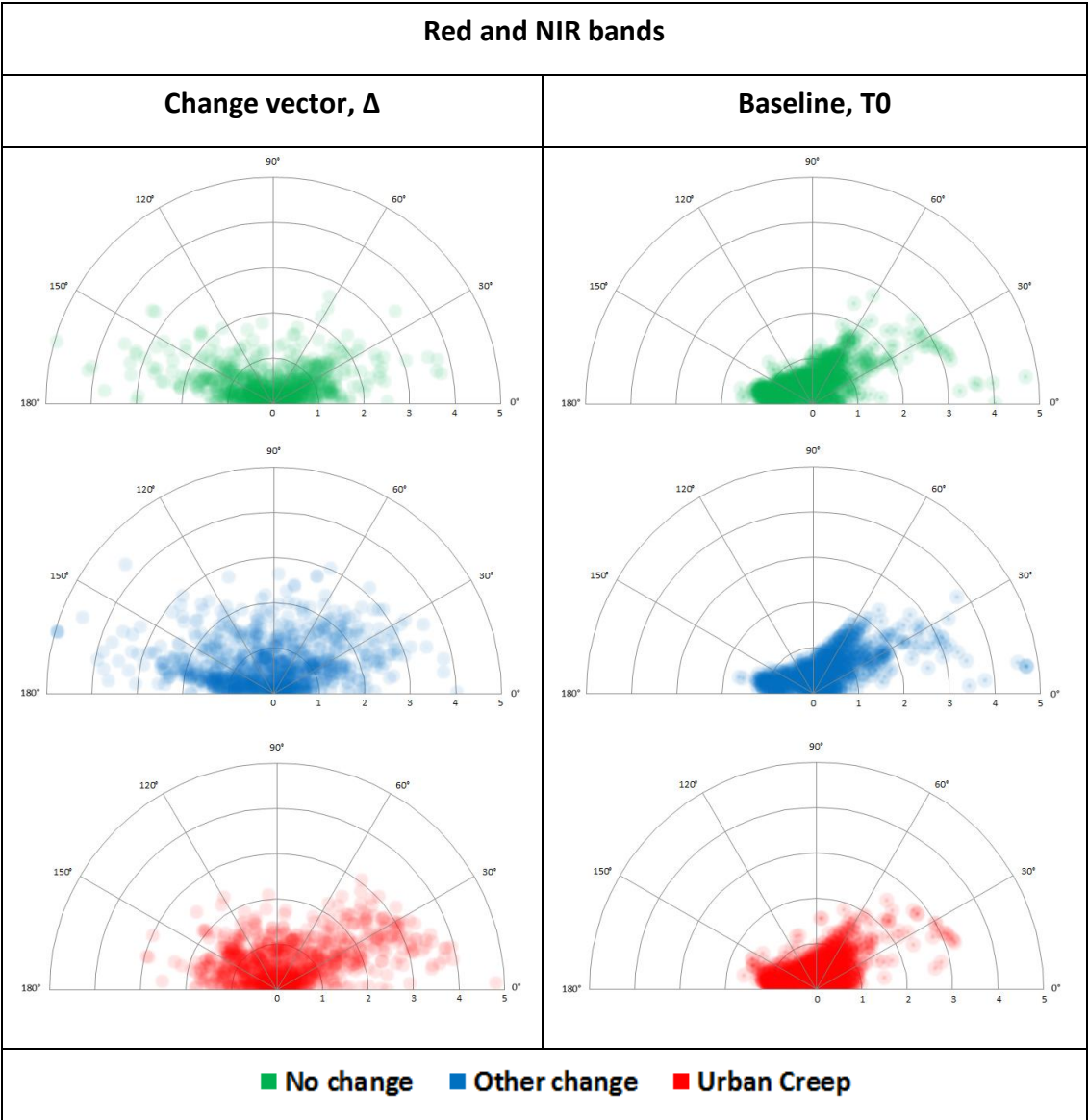


Figure 60: CTA polar visualisation of the Red and NIR bands. The left hand panel shows polar plots of the change vector, separated by reference data class. The right hand panel shows polar plots of the baseline, T_0 vector, separated by reference data class. The polar plots depict vector angle (degrees) rotated around the Y axis and vector magnitude (standard deviations) along the X axis.

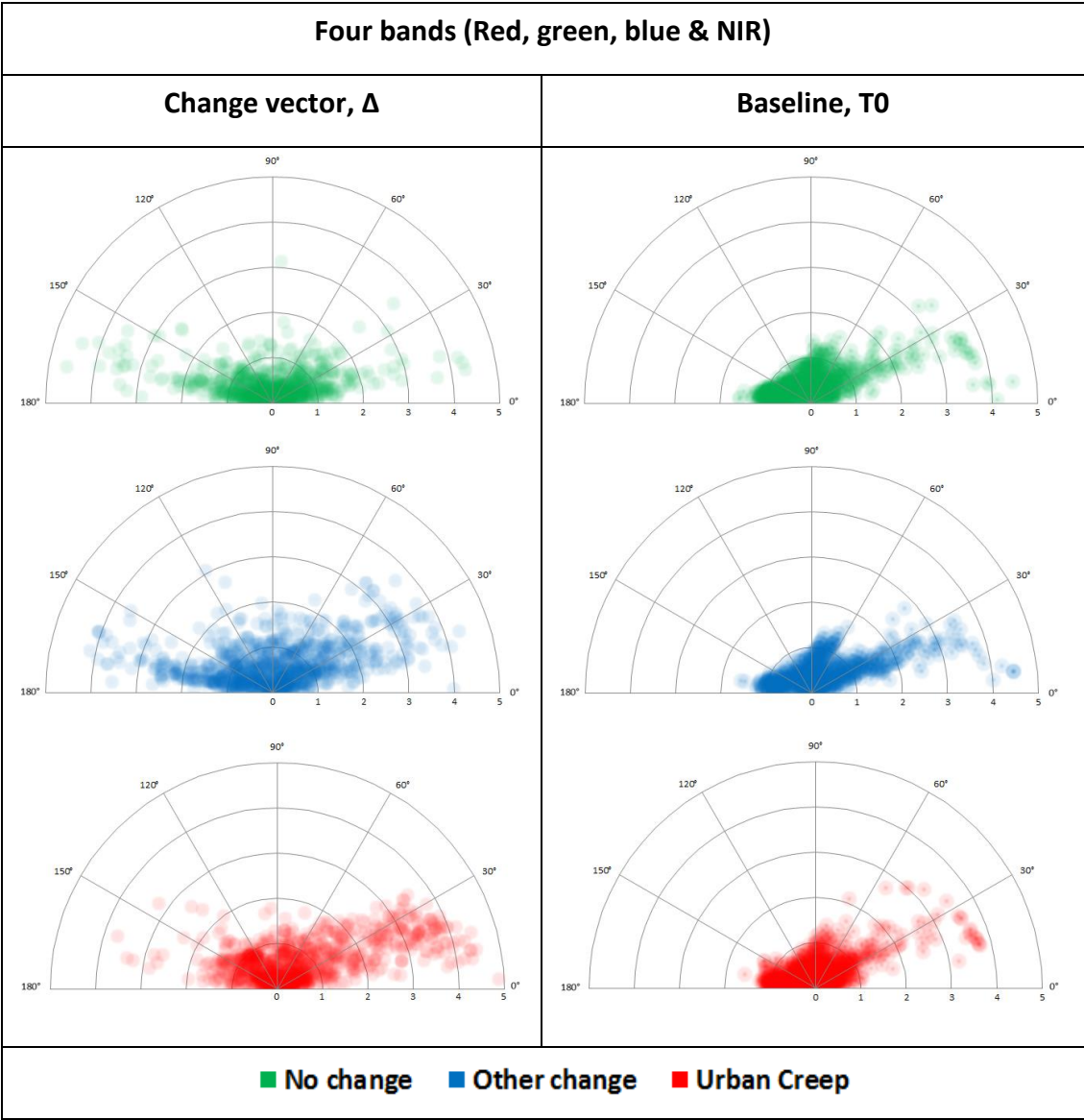


Figure 61: CTA polar visualisation of the four spectral bands. The left hand panel shows polar plots of the change vector, separated by reference data class. The right hand panel shows polar plots of the baseline, T_0 vector, separated by reference data class. The polar plots depict vector angle (degrees) rotated around the Y axis and vector magnitude (standard deviations) along the X axis.

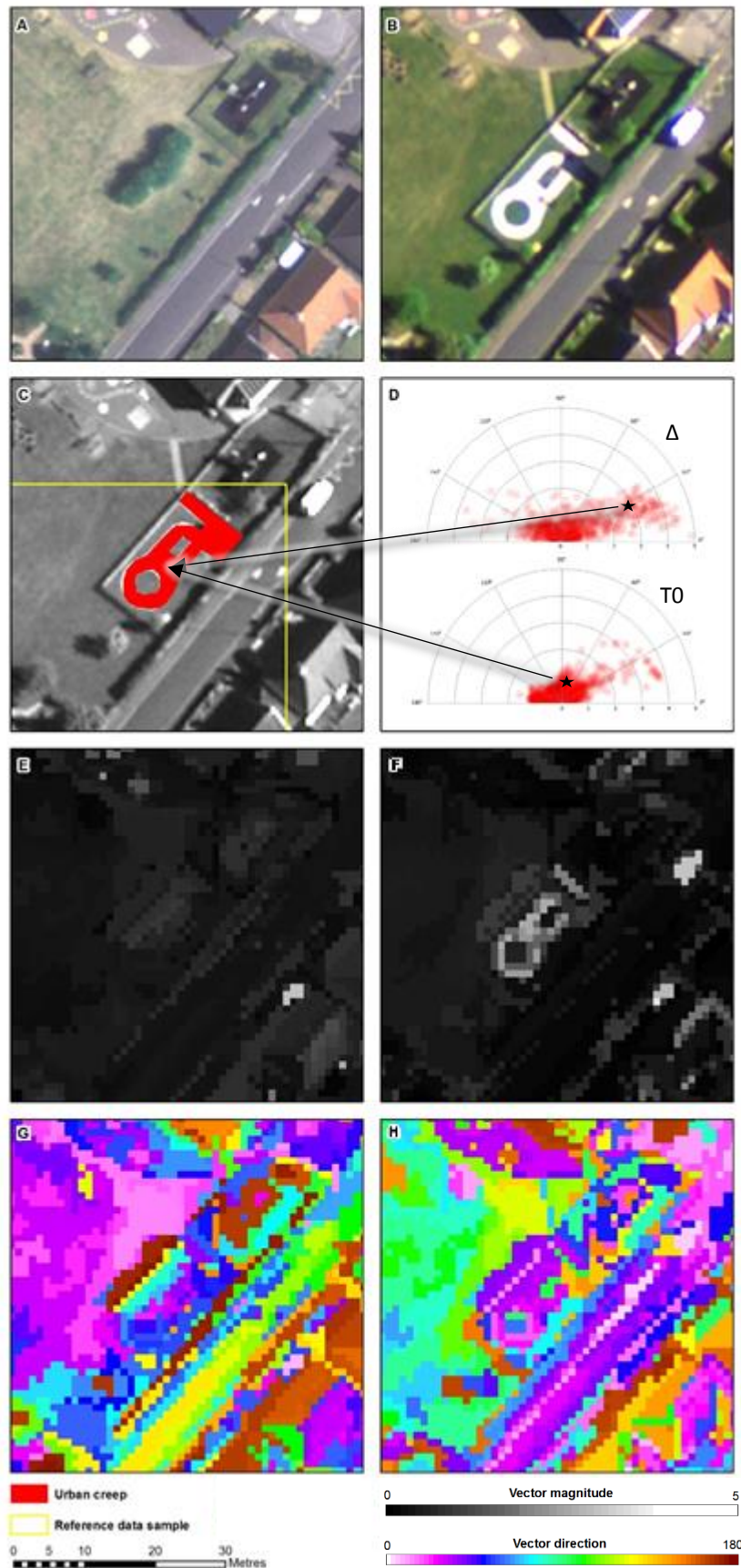


Figure 62: CTA visualisation. The figure shows an example of the CVA anomaly for four bands. Panel A & B show the 2006 and 2010 imagery respectively. Panel C shows the reference data mapping. Panel D shows the Δ and T0 urban creep polar plots. Panel E & F show the T0 and Δ vector magnitude. Panel G & H show the T0 and Δ vector direction

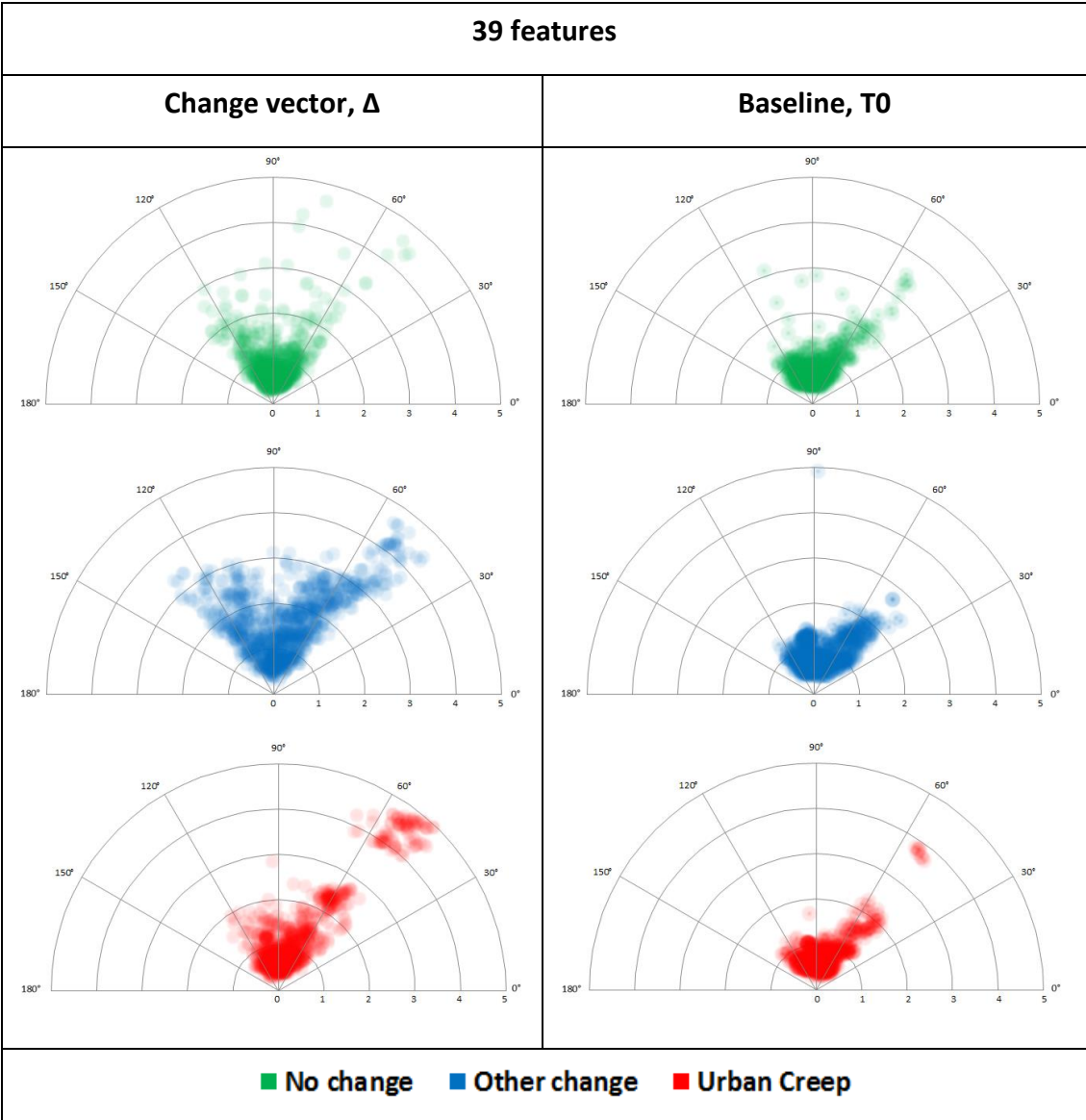


Figure 63: CTA polar visualisation of 39 features. The left hand panel shows polar plots of the change vector, separated by reference data class. The right hand panel shows polar plots of the baseline, T_0 vector, separated by reference data class. The polar plots depict vector angle (degrees) rotated around the Y axis and vector magnitude (standard deviations) along the X axis.

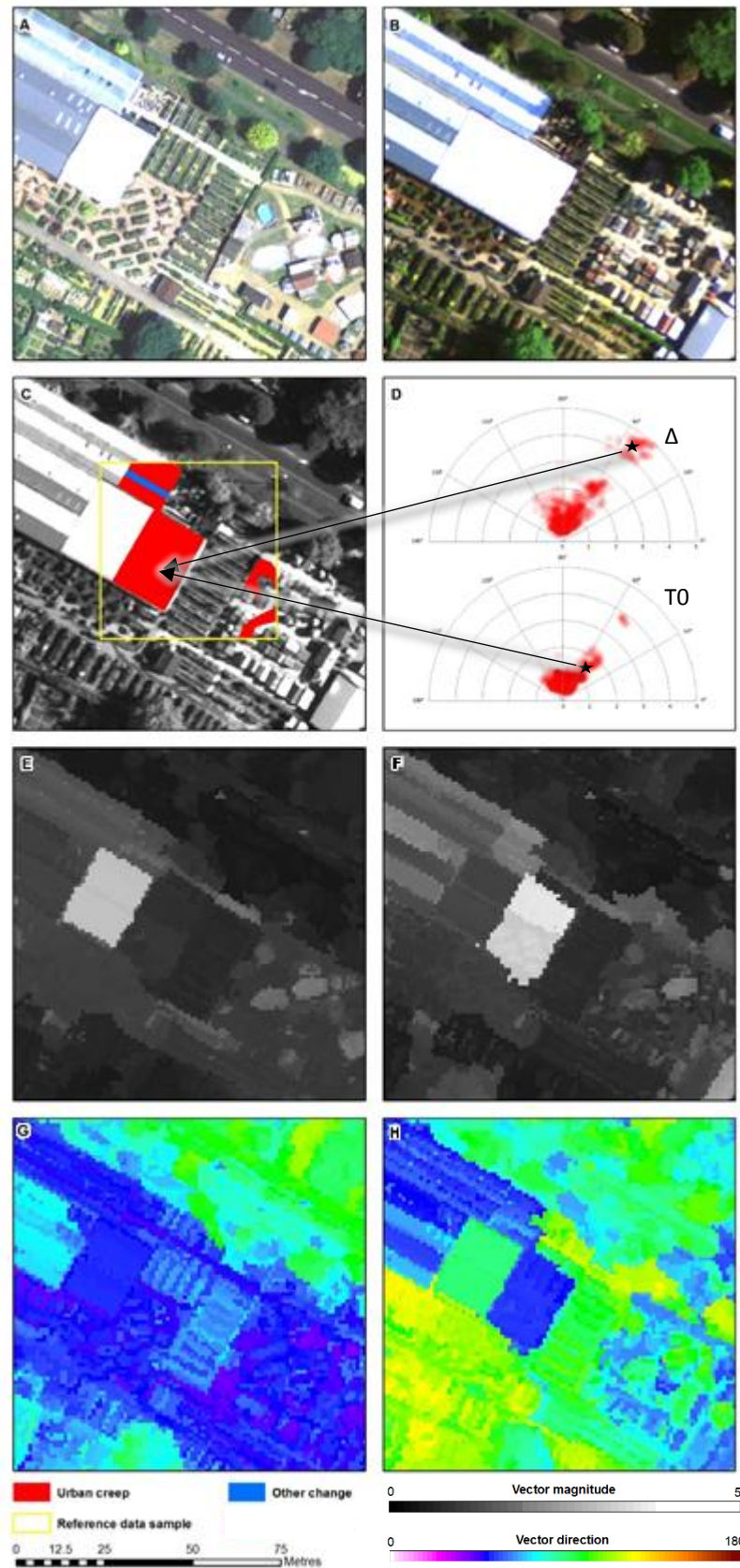


Figure 64: CTA visualisation. The figure shows an example of the CVA anomaly observed with 39 features. Panel A & B show the 2006 and 2010 imagery respectively. Panel C shows the reference data mapping. Panel D shows the Δ and T0 urban creep polar plots. Panel E & F show the T0 and Δ vector magnitude. Panel G & H show the T0 and Δ vector direction.

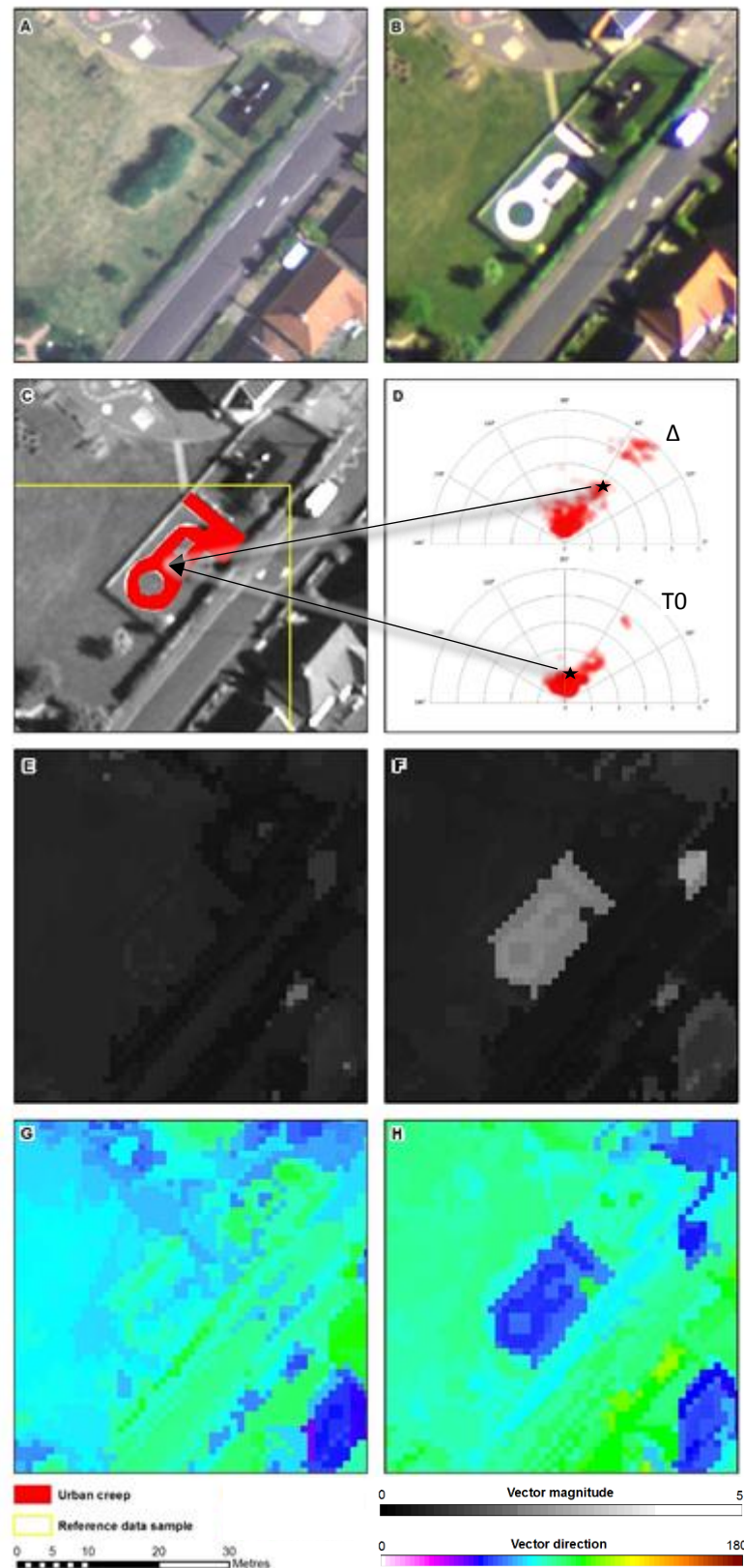


Figure 65: CTA visualisation. The figure shows an example of urban creep observed with 39 features. Panel A & B show the 2006 and 2010 imagery respectively. Panel C shows the reference data mapping. Panel D shows the Δ and T0 urban creep polar plots. Panel E & F show the T0 and Δ vector magnitude. Panel G & H show the T0 and Δ vector direction

8.6 Summary

This chapter has presented the results obtained from the remotely sensed change detection. The change detection work tested the capability of many different classification approaches applied to urban creep identification.

The change detection results are summarised in Table 25. In all, fifteen specific results were obtained although they may be aggregated into broader themes. In particular, Results 10 and 11, are important as they state the difficulty involved in urban creep classification and the considerable performance difference observed between the direct and CVA classifications. Also, Results 22 and 23 provide interesting evidence, highlighting that urban creep is not well described by feature differences, showing considerable confusion with unchanged areas. In light of this; we can make an assessment as to the overall success of remotely sensed urban creep classifications and the relative utility of CVA in this application. These results will be used to support a discussion presented in Chapter 9. The discussion explores the differences between the two main approaches, the types of errors identified and the underlying reasons for those errors.

Table 25: The change detection results summary table.

ID	Description
Result 9	In the majority of the unsupervised selections, moderate-scale features are over-represented at the expense of small-scale features.
Result 10	The most successful direct classification, achieved an average urban creep accuracy of 52.15%, utilising 62 features from all available.
Result 11	Direct classifications of urban creep considerably outperform equivalent CVA classifications.
Result 12	The addition of T0 to an urban creep classification significantly boosts classification accuracy.
Result 13	The CVA methods novel to this thesis, indicate an improvement in classification accuracy when compared to spectral CVA techniques.

ID	Description
Result 14	In general, more than half of urban creep is not identified, and increasing classification complexity does not improve identification rates.
Result 15	In general, when increasing the number of input features, direct classifications respond by reducing urban creep false positives, while CVA classifications increase false positives.
Result 16	Observed urban creep accuracies are considerably lower than those of other change features.
Result 17	In general, other change omission rates do improve as more features are considered.
Result 18	The majority of the classification errors relate to confusion between the change classes and no change, rather than inter-change confusion.
Result 19	Tone, texture, shape and pattern are by far the most important interpretation principles.
Result 20	Height, size, view angle, association, and shadow all have relatively low importance, but do contribute to the overall classification result.
Result 21	Moderate-scale features are by far the most important to the classification followed by small, and large-scale features.
Result 22	The CV magnitude response of urban creep is, in most cases, more correlated with no change than other change.
Result 23	In general, urban creep does not form unique CTA clusters. A separable cluster was only achieved for large, distinct features with extraordinary magnitude.

9 Discussion

9.1 Introduction

This chapter discusses the main results obtained in chapters 6 and 8 to assess their significance in a wider context. The results are discussed in two broad themes: Urban creep estimation and remotely sensed change detection. The first theme discusses the level and extent of urban creep identified during the reference data collection. This theme continues with an explanations of the results, their implications and the success and utility of remote sensing change detection applied to urban creep identification. The second theme discusses the result in a remote sensing context, generalising the very specific application presented to the wider discipline. This chapter contributes to the thesis by comparing the results with other published research, offering explanations for their occurrence, and highlighting their importance. By doing so, the contribution to knowledge is discussed to include implications and recommendations for further research.

9.2 Purpose

The purpose of this thesis was to explore urban creep detection using remote sensing techniques. This both tests the state-of-the-art and revisits the relevance of traditional differencing methods. It was hypothesised that urban creep could not be adequately modelled with classical remote sensing techniques such as CVA. By conducting the research, it is hoped that the capability of remote sensing to detect and classify large-scale urban change features will be clarified. Specifically, the thesis sought to find out if effective urban creep classification is possible using state-of-the-art techniques, represented by a direct classification. As a counterpoint it also seeks to assess the relevance of traditional differencing methods, here represented by CVA. A case study of detecting urban creep in Norwich between 2006 and 2010 from VHR imagery was chosen. To assess the success of the case study, reference data was collected to describe the occurrence and rate of urban creep throughout Norwich. The rate of urban creep calculated from the reference data is in itself a valuable finding. Therefore the secondary purpose of the thesis is to assign social importance to the observed rate of urban creep.

9.3 Results summary

To focus the discussion, a subset of the most important results identified in Chapters 6 and 8 will be carried forward for discussion. Table 26 lists the results selected for further discussion.

Table 26: The subset of the results for discussion.

ID	Description
Result 2	Urban creep clustering is visible in New Costessey, close to Dereham Road and to a lesser extent, Sprowston.
Result 5	An urban creep rate of 24.11 m ² /Ha/year was observed, equivalent to 0.24 % per year.
Result 6	The addition of new permeable gravel surfaces was widespread and easily confused in VHR imagery as impermeable surface creep. A more rigorous, multi-source interpretation yielded urban creep commission error rates of 21%.
Result 7	The contextual assignment of urban creep was complicated and subjective to a certain degree.
Result 10	The most successful direct classification, achieved an average urban creep accuracy of 52.15%, utilising 62 features from all available.
Result 13	The CVA methods novel to this thesis, indicate an improvement in classification accuracy when compared to spectral CVA techniques.
Result 14	In general, more than half of urban creep is not identified, and increasing classification complexity does not improve identification rates.
Result 15	In general, when increasing the number of input features, direct classifications respond by reducing urban creep false positives, while CVA classifications increase false positives.
Result 16	Observed urban creep accuracies are considerably lower than those of other change features.
Result 18	The majority of the classification errors relate to confusion between the change classes and no change, rather than inter-change confusion.

ID	Description
Result 19	Tone, texture, shape and pattern are by far the most important interpretation principles.
Result 20	Height, size, view angle, association, and shadow all have relatively low importance, but do contribute to the overall classification result.
Result 21	Moderate scale features are by far the most important to the classification followed by small and large scale features.
Result 22	The CVA magnitude response of Urban creep is, in most cases, more correlated with No change than Other change.
Result 23	In general, urban creep does not form unique CTA clusters. A separable cluster was only achieved for large, distinct features with extraordinary magnitude.














9.4 Urban creep estimation

9.4.1 Relation to existing research

The urban creep rate estimated in this thesis by sampling is broadly consistent with those observed in other studies. Table 27 catalogues all known published urban creep rates alongside the thesis findings. The published results range from 0.12% per year for Newcastle 2002 to 2007 (UKWIR 2010), through to 1.31% per year for suburban Newcastle 1996 to 2005 (Newcastle City Council 2008). Mean and median rates, inclusive of the thesis result, are 0.46% and 0.29% respectively. Firstly, it is clear the thesis result is in the same order of magnitude as the other estimates. Moreover, it is the most recent. The new result from Norwich is considerably lower than the mean but very close to the median rate. The two highest rates (from Newcastle and London) are outstanding and could be interpreted as outliers. Therefore, the median value may be a more appropriate point of comparison. On this basis, we could conclude that the thesis rate of urban creep is very similar to the median of the other published results and does not show any marked change. In fact, the thesis result is identical to the UKWIR average. However, when compared to the previous estimate for Norwich,

there is a marked increase from 0.14% to 0.24%. In light of this we can conclude that it appears that for Norwich at least, the rate of urban creep has increased considerably.

Table 27: Urban creep rate comparison. The thesis result is highlighted in red, while, the previous estimate for Norwich is highlighted in blue.

Reference	Location	Time period	Urban creep rate %/year	
Newcastle City Council (2008)	Samples from sub-urban Newcastle, UK	1996 to 2005	1.31%	
London Wildlife Trust et al. (2011)	Greater London, UK (Garden areas only)	1998-99 to 2006-08	1.12%	
Cutting (2003)	Oakwood, Derby, UK	Build to 2003	0.67%	
UKWIR (2010)	Leicester, UK	1999 to 2006	0.39%	
	Maidstone, UK	2003 to 2006	0.34%	
Perry & Nawaz (2008)	Sub-urban Leeds, UK	1971 to 2004	0.29%	
This study	Norwich, UK	2006 to 2010	0.24%	
UKWIR (2010)	UKWIR average	circa 1999 to circa 2007	0.24%	
	Chester, UK	2003 to 2007	0.23%	
UKWIR (2010)	Norwich, UK	1999 to 2006	0.14%	
UKWIR (2010)	Newcastle, UK	2002 to 2007	0.12%	
Mean			0.46%	
Median			0.29%	

9.4.2 Explanation of the results and their importance

The apparent increase in the rate of urban creep from 1999 - 2006 to 2006 - 2010 could be broadly explained by overriding economic factors. The concept of linking urban creep to social and economic factors is not new. For example, Duckworth (2005) suggested possible correlations with paving material sales figures, while UKWIR (2010) explored demographic links. Considering the global economic crisis in 2008, overriding economic conditions would not appear conducive to increased rates of urban creep at

this time. In the UK there was a significant drop in GDP growth, from approximately 1.25% in the third quarter of 2007, to -2.2% in the fourth quarter of 2008, with no recovery until the fourth quarter of 2009 (Office for National Statistics 2012). Intuitively, we would expect these conditions to equate to a drop in development, and consequently a reduction in urban creep. This hypothesis could be compounded by the fact that in England, house building fell from approximately 165 000 completions in 2006-7 to approximately 115 000 completions in 2010-11 (Department for Communities and Local Government 2016). However, reduced housebuilding offers a possible explanation for the apparent acceleration of urban creep rate. With house building reduced, house moves are also reduced. In this case, householders may be more likely to undertake home improvement projects, some of which will add impermeable surface. Interestingly, between 2006 and 2010 household expenditure remained static (Office for National Statistics 2016), providing some evidence supporting this hypothesis. In light of these conflicting pieces of evidence, it can be concluded that a possible explanation for the increase in the rate of urban creep was a focus on home improvement rather than house moves, prompted by a depressed economic outlook.

Another explanation for the apparent increase in urban creep rate is that there is, in fact, no increase, and we are observing the cumulative effect of inaccuracy in the estimates. This thesis estimated the rate at $0.24\% \pm 0.065\%$ (95% confidence limits) between 2006 and 2010. Assuming the previous estimate from 1999 to 2006 of 0.14% (UKWIR 2010) is completely accurate, it is highly unlikely that the apparent increase is false. Of course, there will be some element of imprecision and bias in the UKWIR result, but accuracy was not published. Assuming both estimates have the same accuracy as the thesis result, the 95% confidence limits do overlap. Therefore, it is possible that there is, in fact, no difference between the rates. In summary, although it is more probable that an increase in urban creep rate was observed, more research is recommended to corroborate this finding.

The newly derived urban creep rate is important as it provides a much-needed addition to the literature in support of urban drainage planning and policy. In 2010 UKWIR (2010) recommended that urban creep be regularly monitored, with rates

revised at approximately five yearly intervals. Since then, only one further estimate has been published (London Wildlife Trust et al. 2011), although the monitoring period concerned (1998-9 and 2006-8) does not advance the chronology of the UKWIR research (circa 1999 to circa 2007). Therefore a revision is overdue. The thesis result is an important addition to the literature because of existence, timeliness and data continuity. Firstly, there is limited urban creep research published, and this work expands the quantitative evidence base available by approximately 10% (based on the research catalogued in Table 27). Secondly, with a monitoring period of 2006 to 2010, the thesis result is the most recent urban creep research published. Lastly, this thesis provides continuity with previous urban creep research in Norwich. Hence important insight is gained into the changing nature of urban creep.

The findings will be of importance locally to stakeholders of urban drainage in Norwich. Chapter 5 discussed the threat of surface water flooding in Norwich and described six areas of concern identified by Norwich County Council and Anglian Water. The results of the urban creep sampling presented in Chapter 6 identified elevated urban creep activity in one of these areas, Sprowston, to the Northeast of the city. However, increased activity was also identified in New Costessey, to the West of the city which was not highlighted as an area of concern. Therefore, this finding could be used to inform future surface water management planning. The Norfolk strategic flood risk assessment (Wilson 2010, p50, 51) voices the desire to limit urban creep to mitigate any further flood risk. Similarly, the Norwich surface water management plan (Ogilvie et al. 2011, p5) highlights the importance of raising community awareness of surface water flood risk and the importance of Sustainable Urban Drainage Systems (SUDS). Therefore, the thesis findings could be used to illustrate the scale of the problem and harbour public support for more sustainable development and home improvement. While it is beyond the scope of this thesis to speculate on the implications of accelerated urban creep on local flooding, it is clear the results will help to inform decision makers and could raise public awareness towards more sustainable development.

The urban creep sampling methodology developed in this thesis is a valuable contribution to the literature, building upon previous works. Ground-breaking studies

by Cutting (2003), Duckworth (2005), and Richard Allitt Associates Ltd. (2008) focussed on relatively small extents aiming to establish urban creep as a real, measurable phenomenon. These works laid the foundations with which other research has followed. In particular, Cutting (2003) demonstrated that a field survey referencing historical mapping and aerial imagery could identify urban creep. Richard Allitt Associates Ltd. (2008) applied sampling to indicate the rate of urban creep across wider areas, while Duckworth (2005) showed how aerial imagery alone could be used as an effective indicator of urban creep. Subsequently, further work focused on city-wide estimates. London Wildlife Trust et al. (2011) showed how a large random sample of 1292 garden plots was representative of the whole of London. On the other hand, UKWIR (2010) used a semi-automatic object-based remote sensing classification to quantify urban creep across five entire cities. This thesis builds upon these studies by demonstrating an efficient, city-wide approach with quantitative accuracy estimates. Specifically, a novel pre-stratification using Ordnance Survey MasterMap™ and multi-temporal imagery facilitated the application of a random stratified sample, reducing sample sizes while retaining urban creep estimation accuracy. This resulted in an accurate estimate of the rate of urban creep for Norwich with only 154 samples. Lastly, statistical inference was applied to derive quantitative estimates of urban creep accuracy, a first in the literature. In summary, the sampling methodology derived in this thesis could be an important asset for researchers seeking updated, city-wide urban creep estimates.

To the Author's knowledge, this is the first study that rigorously investigates the capability of remote sensing change detection to identify urban creep. In previous research Allitt & Tewkesbury (2009) demonstrated how a rule-based semi-automatic OBIA workflow could be used to assist urban creep identification. This thesis goes beyond existing work in three specific areas:

- Demonstrated a fully automatic, albeit supervised, classification workflow
- Compared the performance of two broad methodological categories: Direct classification and CVA
- Provide a rigorous accuracy assessment with quantitative results

The change detection classification results provided by this thesis will likely prove interesting reading to those concerned with urban drainage planning keen to understand the changing nature of urban creep. While Chapter 4 discusses the importance of urban creep in urban drainage planning and flood prevention, existing estimates of creep extent and rate are limited, with new estimates difficult to obtain. Given that in the UK at least, rates of urban creep are overdue a revision (UKWIR 2010), the findings and methods described in this thesis are a timely contribution to the literature. Firstly, the results inform relevant stakeholders that fully automatic, remotely sensed estimates of urban creep rate and extent with an accuracy exceeding the common ‘standard’ of 85% (Foody 2008) is still beyond the current state-of-the-art. However, with a user accuracy of up to 61%, end users can find valid instances of urban creep far more quickly and derive quantitative information regarding the nature and extent of the change. To derive meaningful quantitative statistics and effectively map urban creep, the automatic classification results would still require considerable manual correction. In light of these findings, urban drainage stakeholders would be able to make an informed decision when considering remote sensing technology for urban creep research.

9.4.3 Limitations

The accuracy of the sampled result reported is a combination of bias and precision (Atkinson & Foody 2002). Imprecision was observed and quantified as a confidence limit by analysing the variance in the sample. The data collection method described in Chapter 5 was strictly implemented to minimise bias. However, without multi-temporal field observations, it is not possible to eliminate bias. Specifically, bias could remain in three ways: permeability confusion, data gaps, or contextual confusion. Firstly, in a few extreme cases, it could be possible that the interpretation of permeability change was incorrect. For instance, permeable gravel could be incorrectly labelled as an impermeable hard surface, leading to an overestimate of urban creep. Secondly, it is possible that small occluded portions of the image, or inaccessible field check sites were misinterpreted. Lastly, since the classification of urban creep is subjective to some degree (Evans & Eadon 2007), there may be a contextual bias between estimates. In other words, marginal cases could be

interpreted as either urban creep or *other change* depending on the interpreter. As with any research, it is all but impossible to eliminate bias, and while it is clear that some bias may exist, there is no direct evidence to suggest that significant bias remains.

9.4.4 Implications

The newly defined urban creep rate for Norwich could be the impetus for further work and a review of planning policy. The ‘first wave’ of research estimating urban creep (Cutting 2003; Duckworth 2005; Richard Allitt Associates Ltd. 2008; UKWIR 2010) shaped current government planning policy (Department for Communities and Local Government 2008b) and guided urban drainage system stakeholders to minimise flood risk (Ofwat 2011). Chapter 4 of this thesis established the significant urban flood risk posed by circa 1999 to circa 2007 rates of urban creep. This thesis provides evidence that the rate of urban creep, for Norwich at least, could be accelerating. This finding poses two questions:

1. Is accelerated urban creep evident anywhere else in the UK?
2. What are the urban flooding implications of accelerated creep?

To answer the first question would, of course, require further revised estimates across the UK. A direct implication of this work is that it provides evidence that a revision of urban creep rates called for in previous work (UKWIR 2010) is both overdue and urgent. In answer to the second question, an assessment of the significance and flood risk implications of an increase are beyond the scope of this thesis. However, existing research (Ofwat 2011, p30) does outline measures that could be deployed to mitigate urban creep. Specifically, these include: sewer system modification and upgrade, stricter enforcement of existing building regulations, and removal of inappropriate sewer connections, thereby reversing urban creep. The findings of this thesis alone are unlikely to directly influence planning policy, but, they serve to highlight the need for a policy review. Specifically, the findings indicate that planning legislation passed in 2008 (Department for Communities and Local Government 2008b) in order to curb urban creep could be ineffectual.

9.4.5 Recommendations for further research

The rate of urban creep estimated in this thesis puts into question the relevance of existing rates used to manage urban sewer systems across the UK. Furthermore, the findings question the effectiveness of legislation and planning policy in this area. In light of these points further work is recommended in three areas:

1. A wider revision of urban creep rates across the UK should be undertaken to corroborate the increase identified in Norwich. The sampling methodology presented in this thesis would be an ideal approach to provide quantitative statistics with defined error bounds.
2. These revised rates should then inform research investigating their impact. Specifically, this would include hydraulic modelling assessing flood risk and an assessment of the level of investment required to mitigate these risks.
3. Pending a revised, national urban creep impact assessment, government planning policy should be revisited. Specifically, the effectiveness of urban creep control legislation introduced in 2008 (Department for Communities and Local Government 2008b) should be reviewed.

9.5 Remotely sensed change detection

9.5.1 Relation to existing research

Overall change map accuracy obtained in this thesis is broadly consistent with other VHR image change detection research published in the literature. The most successful classification in this thesis, a direct classification, obtained an overall map accuracy of 88.26%. Table 28 lists a selection of VHR image change detection overall accuracy results reported in the literature. The list is not an exhaustive catalogue and does not describe a definitive accuracy ranking. The point being that the literature reports a range of accuracies above and below the thesis result, not in contrast to it. For example, Bovolo (2009) presented a hugely successful multi-level CVA technique that detected a variety of changes in a VHR image pair with an overall accuracy of 98.05%. The literature also presents a range of different accuracies as low as approximately 50% (Tang et al. 2013). By no means are these works less important contributions to the literature, in fact, they serve to illustrate the high variability in VHR imagery and

the challenge of different applications. For example, Tang et al. (2013) present novel, useful and rigorous building change detection research with comprehensive accuracy reporting. The lower reported accuracies appear to be the result of false positives induced by the challenging imagery and application. There is no evidence that their proposed method is deficient. In fact, there is contrary evidence, as the method proposed by Tang et al. (2013) was compared against, and outperformed Bovolo's (2009) multi-level CVA method topping the list in Table 28. This particular example highlights the limitation of benchmarking VHR image change detection against one another because of the differing challenges posed by image quality and the nature of the application. While the results obtained in this thesis do not push classification accuracy boundaries, they do make a contribution by application and method. Specifically, presenting a novel and extremely challenging application gives other researchers an insight into the methodological design and the expected performance of very granular urban change detection systems.

Table 28: A brief literature survey of VHR change detection overall accuracy.

Study	Overall accuracy
Bovolo (2009)	98.05%
Marchesi et al. (2010)	96.69%
Hao, Shi, Deng, Zhang, et al. (2016)	94.29%
Bruzzone & Bovolo (2013)	93.91%
Falco et al. (2013)	89.10%
Jian et al. (2016)	88.21% to 93.88%
This study	88.26%
Tang et al. (2015)	80.80%
Wu et al. (2016)	76.43%
Vakalopoulou et al. (2016)	74.40%
Tang et al. (2013)	49.51% to 52.40%

In contrast to some other published research, this thesis showed a larger difference between basic, spectral CVA and a particular state-of-the-art technique under

investigation. Given the complexity of many satellite scenes, -it is an intuitive and expected result that CVA can be outperformed by more sophisticated methods. For example, Tan et al.'s (2016) spectral-spatial method captured urban change with accuracies 95.87% and 97.76% compared to CVA at 91.30% and 91.71%. Similarly, Hao et al.'s (2016) Super Pixel Active Contour Model outperformed CVA at 93.5% compared to 87.2% and Shi et al. (2014) improved forest clearance identification from 93.38% with CVA to 95.5% with a chi-squared transformation. Moreover, Shah-Hosseini et al. (2015) improved VHR image change detection to as high as 97.73% from a baseline of 86.57%. This selection of studies showed differences between CVA and proposed state-of-the-art methods in the range of 2.12% to 11.16%. In contrast, this thesis reports a large gulf between the most successful method and standard formulations of CVA. For example, a direct classification achieved overall map accuracy of up to 88.26% and urban creep average accuracy of up to 52.15%. In stark contrast, Red & NIR CV Magnitude yielded equivalent accuracies of 51.78% and 18.46%, in this case, differences of 36.51% and 33.69%. Considering up to four spectra bands and CV magnitude and direction, the gap closes to 22.51% for overall accuracy and 24.06% for urban creep. Nonetheless, a large difference remains. An initial explanation for this disparity with the literature is that urban creep is not well described by spectral differences when compared to many of the change targets presented in the literature. For example, VHR image changes identified by Ma et al. (2016) were well described by spectral features, and they found no improvement to detection rates by adding additional object-based features. However, the class *other change* also shows a high disparity between the direct classification and simple spectral CVA of between 45.04% and 33.91%. This is unexpected because, within this thesis, *other change* is more typical of changes reported in the majority of the change detection literature such as construction activities and broad land cover changes. This slightly confusing result could be explained by scale. Specifically, because the reference data and training data reside at a very large cartographic scale, only the more sophisticated classifications can reconstruct the change in a similar form as the reference data. The disparity with the literature could then be explained by the challenge of change detection at the very large-scale undertaken within this thesis (1:500, with 25cm imagery). At this scale, these findings indicate that simple spectral

CVA accuracy is considerably less than smaller scale results from the literature. Furthermore, the very large ‘boost’ in accuracy provided by the direct classification indicates the effectiveness of this method in overcoming the challenges of large-scale urban change detection.

9.5.2 Explanation of the results and their importance

Part of the difficulties of urban creep identification can be explained by analysis scale. The results showed a profound difference in accuracy between generally small, urban creep features at up to 52.15%, and generally larger *other change* features at up to 73.45%. Furthermore, Result 23 highlights that only larger, distinct urban creep features exhibit strong, distinguishable change signals. This result corroborates work by Weber & Chen (2010) who found a strong correlation between VHR image target size and classification accuracy, with larger targets achieving higher accuracies. Also, Ma et al. (2016) reported better VHR image urban change detection performance when analysing the images with moderate scale segmentations compared to fine scale segmentations. While work by Padilla et al. (2015) showed that burnt area was very challenging to classify due to its size and rarity. In fact, the respective commission and omission errors observed by Padilla et al. (2015) at >40% and >65% respectively, are broadly in line with the rates for urban creep presented in this thesis. Difficulties identifying small, rare classes were also reported by Dingle Robertson & King (2011). The thesis results add further evidence to the literature highlighting the difficulty of capturing small, rare change classes. More importantly, it adds evidence against any relationship between image resolution and classification accuracy. Atkinson & Foody (2002) tentatively suggest, with appropriate caveats, classification accuracy may improve as image resolution increases. The result of this thesis rejects any such relationship, instead, corroborating with Huiping et al. (2003) that accuracy is class dependent.

Part of the difficulty in urban creep identification will also be down to the difficulty in identifying impermeable surfaces from remotely sensed images. As urban creep is an addition of impermeable surface, correct identification is highly sensitive to surface type. It is commonly reported in the literature that impermeable surface is difficult to detect from remotely sensed images, and is prone to confusion with automated image

analysis (Weng 2012; Lu et al. 2011; Lu & Weng 2006; Ridd 1995), and manual image interpretation (Perry & Nawaz 2008; Duckworth 2005; London Wildlife Trust et al. 2011). The results of the reference data collection, in particular, the field work confirm that this is the case, and in some cases, it was impossible to differentiate surface permeability by image colour and texture alone. It is highly likely that urban creep identification is more likely to suffer from surface type confusion than larger developments because of context. Urban roads and roofs form continuous networks which add contextual confidence to individual elements. On the other hand, urban creep is often isolated and lacks contextual support. Clearly, surface type confusion has contributed to the limited classification accuracy, in accordance with trends observed in the literature. However, the effect of scale compounds the confusion by removing contextual support in the classification.

The limitations of the contextual modelling are also a factor in contributing to poor urban creep classification accuracy. Urban creep is differentiated from other impermeable surface additions such as new housing development by complex contextual relationships. The experimentation in this thesis modelled the contextual relationships by defining a large variety of object-based features describing image colour, texture and structure at three different scales. For direct classifications, in general, urban creep classification accuracy improved with the addition of more features, with maximum accuracy achieved at 62 features. This shows that the most successful classifications were dependent on a large number of criteria. The CVA urban creep classifications, by comparison, did not perform as well. Crucially, in general, the CVA classifications performed worse with more features. The important point is that the simple, compounded feature differences represent urban creep very poorly, typically below 30% in accuracy. In other words, feature differences positively correlated with urban creep will to some extent be cancelled by uncorrelated or negatively correlated features. On the other hand, a direct classification can model some of the complicated multi-criteria relationships that better describe the class, boosting accuracy by 20-30%. However, no matter how many features are added the level of omission errors remain relatively static at approximately 40-45%. In summary, the data shows that correct urban creep identification is a complicated combination of

a large number of object-based features showing a range of different responses. It is not simply a response to unanimous feature difference.

The classification results showed that the majority of the errors were due to confusion between the change classes and *no change*, rather than inter-change confusion. In other words, in most cases, if urban creep was incorrectly classified, it was more likely to be labelled as *no change*, rather than *other change*, and vice-versa for *other change*. This result is important and interesting because it indicates that the classifier, informed by the object-based features, is making meaningful contextual assignments. Essentially, we would expect urban creep and other change to be readily confused, since both exhibit land cover change. However, the results indicate that in most cases the classifier is prohibiting this and so is ‘recognising’ the contextual separation. In a further result, it was found that the CV magnitude response of urban creep has more in common with *no change* than *other change*. In other words, urban creep often exhibits low CV values, more akin to *no change* while *other change* shows stronger CV magnitude. Therefore urban creep often exhibits little or no CV magnitude. Clearly, correct classification, as demonstrated by the improved direct classification results, is a result of finding patterns in a combination of subtle differences, and the baseline state T0. This finding is quite profound because observing changes as discernible radiometric differences is a fundamental principle of remote sensing change detection (Singh 1989). Furthermore, a large amount of change detection research investigates threshold selection (Adar et al. 2014; Wang 2014a; Patra et al. 2011; Sinha & Kumar 2013) with the specific aim of identifying changes with a large image difference, or CV magnitude. This is an important finding as it demonstrates that, in this application, the fundamental concept of identifying strong image change is of little use, and instead a complicated combination of subtle feature differences and baseline position is important.

Moderate-scale object-based features were an important component of the thesis results. Moderate-scale features were over-represented in the unsupervised feature selection and highly important in the most successful random forest classification. This means that the moderate-scale features were most representative of the scene and provided the most discriminatory information. These findings align with the results

obtained by Ma et al. (2016), where moderate-scale objects outperformed fine-scale objects. In the scope of this research, this outcome is surprising. Given the focus on large-scale change detection and the highly heterogeneous imagery, it is intuitive to assume that large-scale features would be best placed to geometrically define the change features and adequately describe the scene content. An explanation of these outcomes are two-fold. Firstly, the large-scale, L1 image-objects are over-segmented, and there is redundancy in the objects. This is expected to some degree as over-segmentation is a feature of multi-temporal image objects (Ma et al. 2016). Secondly, context is important to the classification process. This explanation is interesting because the urban creep targets are inherently contextual and would not be identified in a simpler classification scheme.

This thesis contributes to the literature by building upon and extending CVA research. Bovolo et al. (2012) introduced a framework for n-dimensional CVA, termed 'C²VA' designed to extend detection capability by considering multiple spectral bands. This thesis builds upon this work in four areas:

- 1) C²VA is extended to include weights, denoted as WC²VA. Tian et al. (2013) applied weights to CVA to balance out the contributions of image and DSM difference. This thesis extends this concept, applying to n-dimensions (WC²VA) using the weightings as a mechanism to normalise the outputs under varying numbers of features. Although not demonstrated in this thesis, WC²VA can also be used to assign varying levels of importance to specific features.
- 2) WC²VA is tested against a complex urban change detection application. Previous research has hypothesised that CVA (Bovolo & Bruzzone 2007) and later C²VA (Bovolo et al. 2012) are methods to identify specific thematic changes. This research further tests this hypothesis against urban creep identification, concluding that such a specific thematic subdivision was not possible.
- 3) The thesis application of CVA goes beyond image spectral bands to include a range of object-based features. He et al. (2011) demonstrate the use of spectral and textural features within CVA, while Bovolo (2009) showed the significant benefit of a multi-scale, object-based CVA. This thesis has extended

this to combine both of these concepts along with features describing spatial pattern, scene illumination and sensor view angle. Furthermore, unsupervised feature selection was employed to identify the most representative features, while WC²VA was used to normalise CV magnitude under varying numbers of input features.

- 4) The framework is extended further to include the baseline vector, T₀, giving CTA. This thesis defines a definitive framework with which to remove CV ambiguity previously reported by Johnson & Kasischke (1998) and Tewkesbury et al. (2015).

The novel CVA methods developed in this thesis show small improvements in urban creep classification accuracy when compared to simple spectral methods. With one to three input features CTA displayed improvements in urban creep accuracy of between 5.985 and 4.39% when compared to Red & NIR WC²VA, and 4 band WC²VA respectively. The main differences in methodology are that CTA is considering T₀ values and a variety of object-based features, as opposed to spectral differences alone. In theory, the object-based features should improve classification accuracy when compared to the spectral features. This is because the object-based features were feature selected, and offer optimum representations of the scene information content. However, the urban creep classification results do not support this hypothesis. Specifically, the feature selected WC²VA –equivalent to CTA without T₀– shows no significant accuracy loss or gain when compared to the spectral WC²VA. Therefore, these limited results imply that the improvement in accuracy between spectral WC²VA and CTA is accounted for by the inclusion of T₀ describing the ‘baseline’ state of the change detection. However, urban creep classification accuracy was very limited in all cases, achieving between 18.34% and 34.07%. Therefore, although improvements were seen in different formulations of CVA, these results demonstrate that the technique is not suitable to detect small, complex urban changes from VHR imagery.

The novel CVA methods developed in this thesis show strong improvements in *other change* classification accuracy when compared to simple spectral CVA methods. When comparing Red & NIR WC²VA, and 4 band WC²VA to the feature selected WC²VA

respective gains of 7.42% and 10% are observed. This comparison is summarised in Table 29. This result shows that the feature selected object-based features are a marked improvement when compared to the simple spectral features. This result also corroborates Myint et al.'s (2011) view that spectral information has limited use in the classification of VHR images. In interpreting the difference between the two, there is a shift in scale, with the majority of the selected features at a moderate-scale compared to the large-scale spectral features. Furthermore, we can see that DSM and texture proved to be more effective discriminators of change compared to the spectral features. CTA improved accuracies further, with gains of up to 17.39% when compared to spectral WC²VA. This indicates that the addition of T0 produces a similar boost to accuracy than substituting spectral inputs for selected object-based features. The largest gain was seen when considering high dimensional CTA. When classifying *other change* CTA responded well to increases in the number of features, achieving a maximum of 67.57% at 36 features. This result far exceeds any of the other methods, with a gain of 32.58% compared to Red and NIR WC²VA. These results, in particular, CTA, show promise for future change detection research with marked *other change* accuracy gains compared to spectral WC²VA. This is because the class *other change* encompasses elements typical to change detection research such as new buildings, roads and construction sites. Therefore, it is expected that WC²VA and CTA, when combined with object-based features, could be beneficial to the wider discipline and other change detection applications.

Table 29: A comparison between spectral and feature selected C²VA.

Number of features	2	4
Spectral features	Δ L1 Red, Δ L1 NIR	Δ L1 Blue, Δ L1 Green, Δ L1 Red, Δ L1 NIR
Feature selected features	Δ L2 DSM, Δ L2 GLCM Mean	Δ L1 Blue, Δ L2 DSM, Δ L2 GLCM Mean, Δ L3 Standard Deviation
Accuracy shift from Spectral to Feature selected WC²VA	+7.42%	+10%
Accuracy shift from Spectral WC²VA to Spectral CTA	+11.30%	+9.64
Accuracy shift from Spectral WC²VA to Feature selected CTA	+16.50%	+17.39%

9.5.3 Limitations

The accuracy of the results will be limited by the DSM resolution and variable data quality. Chapter 4 discussed the importance of an effective isolation between scene illumination and view angle effects and granular changes such as urban creep.

Therefore, in theory, shadow modelling and hidden surface detection are of high importance. However, the classification feature importance analysis does not support this. Specifically, shadow and view angle features did contribute to the classification but were considerably less important than features such as tone, texture and shape. It could be that these features are genuinely less important. Though the most likely explanation lies in the DSM quality limitations and concerns raised in Chapter 7.

Firstly, the limited resolution of the DSMs results in modelled layers far coarser than the image resolution, limiting the shadow and hidden surface identification to larger features only. Secondly, although both DSMs are within an RMSE tolerance of 1.5m, the 2010 DSM is of poorer quality than the 2006 DSM. This second point will limit the

effectiveness of the modelling. This thesis established innovative methods to mitigate the effects of scene illumination and view angle differences in change detection. The innovative illumination and view angle modelling methods derived in this thesis have contributed towards the most successful classification result. However, the improvement is muted by data quality issues. Therefore, there is an opportunity to revisit these methods in future research with more accurate DSM data.

The reference data presented in Chapter 6 highlighted that urban creep is very diverse. The diversity is expressed as a range of different land cover transitions and contextual settings. For instance, added impermeable surfaces could include concrete, tarmac and a variety of different types of block paving. This diversity is further compounded by a range of initial land covers including vegetation, bare earth and gravel. Furthermore, contexts can range from a commercial unit extension to ‘leafy’ residential garden. The reference and training data sampling methodology were treated as rigorously as was possible, within the practical limitations of the research, and satisfying guidelines presented in the literature (Congalton & Green 2009; Stehman 2001). However, the variability experienced during the data collection mean that the diversity may not be adequately sampled. The implications are that the 100 urban creep training samples may not provide sufficient data for the random forest classifier. Given the rarity and difficulty of identification, there is no simple solution to this problem, but further research could investigate active learning to more efficiently capture rare class diversity (Hospedales et al. 2013; Crawford et al. 2013).

9.5.4 Implications of the research

The main implication of this research is that it provides further clarification concerning the limitations of change detection techniques when applied to urban VHR images. It is a long-standing tenet of remote sensing change detection that changes on the ground relate to radiometric differences in the imagery, and that important changes can be isolated because they exhibit large radiometric differences (Singh 1989; Bruzzone & Prieto 2000; Adar et al. 2014; Wang 2014b; Patra et al. 2011; Sinha & Kumar 2013). With the increasing uptake of VHR images in change detection research, there is what could only be described as a sense of ‘unease’ in relation to this tenet, but no clear position. For example, the challenges of VHR image change detection are

often discussed (Bruzzone & Bovolo 2013; G. Chen et al. 2012; Listner & Niemeyer 2011b), there are calls for new methods (Lu et al. 2014) and there is strong support for the use of OBIA, instead of pixels for VHR change detection (G. Chen et al. 2012; Hussain et al. 2013; Boldt et al. 2012). Bruzzone & Bovolo (2013) stress the importance of separating radiometric change into meaningful and noisy components, while Myint et al. (2011) voice concern that spectral information has limited use in VHR image classification. This research acts to bolster and clarify this theme. The experimental results are in direct conflict with the long-standing radiometric differencing tenet. Specifically, in the case of urban creep at least, spectral differences alone have almost no utility to identify meaningful change. Instead, meaningful change was only identified using the complete radiometric change trajectory and a large set of contextual measures. Of course, urban creep detection is a highly specialised application, and this conclusion is far from universal. In fact, spectral differences do represent many VHR changes well (Ma et al. 2016). However, this thesis provides strong evidence that for large-scale VHR urban change detection, the correlation between real land cover change and spectral differences is very limited.

The results also serve to clarify the discussion surrounding object-based change detection. The object vs. pixel debate centres on the fact that differencing image pixel spectral values is unsuitable for VHR imagery because of within class spectral variability, while OBIA is a framework to solve this with spatial and contextual modelling (G. Chen et al. 2012; Hussain et al. 2013). However, Chapter 3 discussed the shortfall between the theoretical potential of OBIA and the reality. Specifically, that object spectral differences are still the most prevalent change discriminator. The change detection results obtained in this thesis conclude that object spectral differences are not an appropriate discriminator for the application of urban creep detection. Recalling the review by Tewkesbury et al. (2015), and Chapter 2 of this thesis argue, it is argued that the unit of analysis and comparison method should not be conflated. Based on the results of this thesis, we could conclude a profound contribution. Specifically, that the pixel is not the problem, it is the differencing of spectral values under any unit of analysis that is problematic.

There is an opportunity for the methods developed in this thesis to be carried forward in further research to ratify their effectiveness and ultimately improve the state-of-the-art. Chiefly among these are the illumination and view angle modelling described in Chapter 7. The contribution and improvement of these novel methods were muted in this research because of DSM resolution and quality. However, the modelling methods have great potential to improve urban change detection result given Lidar data or other suitably accurate 3D data. The novel additions to CVA, WC²VA and CTA, show much potential and further work is recommended. While the urban creep application proved too challenging for this group of techniques, WC²VA and CTA showed considerable accuracy gains when compared to simple spectral CVA. These new techniques could have good utility in lower-scale land cover change detection applications using sensors such as Landsat. Lastly, the thesis proposed an advancement to Multi-Resolution Segmentation, namely, Compactness Enforcement Multi-Resolution Segmentation (CEMRS) which qualitatively improved the formulation of meaningful moderate and small scale objects in urban VHR imagery. There is scope to qualitatively evaluate CEMRS in terms of feature representation and classification accuracy when compared to established segmentation algorithms.

9.5.5 Recommendations for future work

The findings of this thesis highlight the difficulties posed in large-scale VHR image change detection. Specifically, that with fine targets, generalisation, scale filtering or broad structural similarity measures cannot be applied and that the targets of interest bare little relation to observed image spectral differences. While this thesis has gained new insight here, questions remain, and support or rejection of the somewhat controversial findings is needed. Therefore, the following three research areas are recommended to follow on from this thesis:

- Follow up this research with more large scale VHR change detection. This would help to bridge the gap between the resolution of images currently available and the scale of change information reported from them. Especially brave researchers could even undertake further urban creep research.
- More research investigating the applicability of spectral differencing in large scale urban change detection. As well as clarifying detection capacities this

could also help to move the debate in the literature away from ‘pixel vs. object’ towards ‘spectral vs. contextual’ to focus future research.

- Further explore the novel thesis methods. Follow-up research is recommended ratifying the effectiveness in other applications of the shadow and view angle modelling, WC²VA and CTA and CEMRS.

9.6 Summary

This chapter has discussed the main results of the thesis in two broad themes: Urban creep estimation and remotely sensed change detection. The first theme discussed the level and extent of urban creep identified during the reference data collection, explanations of the results and their implications and finally, the success and utility of remote sensing change detection applied to urban creep identification. The main findings of the urban creep discussion are:

- There is a likely increase in the rate of urban creep in Norwich from 0.14% per year (UKWIR 2010) between 1999 and 2006, to 0.24% per year between 2006 and 2010 as derived in this thesis.
- Evidence of a potential increase in the rate of urban creep is very valuable to urban drainage stakeholders both locally in Norwich and nationally. This thesis could be the impetus for further work and a review of UK planning policy.
- The apparent increase could be explained by broad economic factors at the time of the observation. Specifically, there was a reduction in housebuilding paired with static household expenditure which could lead to elevated home improvement.
- The automatically classified urban creep results were not accurate enough to derive meaningful quantitative statistics. However, the automatically classified results could form the basis of a semi-automatic workflow or the derivation of qualitative results.

Secondly, the results were discussed in a remote sensing context, generalising the very specific application presented to the wider discipline. The main findings of the remote sensing change detection discussion include:

- Overall the change detection was successful with overall accuracies consistent with the literature. However, urban creep accuracy was comparably low and remains an extremely challenging application beyond the state-of-the-art.
- There was a very large difference in accuracy between the direct classification and the CVA techniques. In fact, in this application, simple spectral CVA techniques had almost no utility, returning overall accuracy as low as 51.78% and urban creep accuracy as low as 18.46%.
- This thesis provides strong evidence that for large-scale VHR urban change detection, the relationship between real land cover change and spectral differences are very limited. This helps to clarify the position of current change detection techniques when applied to VHR imagery. It also helps to clarify the debate regarding object-based change detection, moving on from ‘pixel vs. object’ towards ‘spectral vs. contextual’.
- The novel methods proposed in this thesis are promising additions to the literature and would warrant further research. These include CVA formulations WC²VA and CTA, the illumination and view angle modelling and the urban segmentation improvement CEMRS.

10 Conclusion

10.1 Introduction

The thesis has explored urban creep detection using remote sensing to test the state-of-the-art, and review the relevance of CVA for large-scale urban change detection. The research has focused exclusively on bi-temporal optical VHR change detection. Therefore, concepts such as multi-temporal data analysis, sub-pixel mapping or active sensors to include synthetic aperture radar, or Lidar have not been investigated. The literature review was split into three chapters, describing remote sensing change detection techniques, the challenges to change detection and the importance of urban creep in terms of social impacts and as a challenge to change detection techniques. The methods and their results were addressed in two phases. Firstly, the Norwich case study was introduced, and a baseline reference data set was established using statistical sampling techniques to quantify the extent of urban creep. Secondly, the remote sensing data was classified using an exhaustive set of experiments exploring the capability of a direct classification, and a classification of CVA features. The purpose of these classifications was two-fold. Firstly, the direct classification represented the state-of-the-art and sought to identify if meaningful urban creep classification was possible. Secondly, the CVA classification represented differencing methods and sought to investigate the relevance of such techniques in this application, and large-scale urban VHR change detection in general.

The experimentation yielded many interesting results, and of these five main findings emerged from the discussion in Chapter 9:

- While the change detection was successful, and broadly comparable with results published in the literature urban creep accuracy was comparatively low, and the application remains a challenge to the state-of-the-art.
- There was a large drop in accuracy observed between the direct, and CVA classifications. In fact, simple spectral CVA techniques had almost no utility returning urban creep accuracy as low as 18.46%.

- The results do not conform to the principle tenet of remote sensing change detection that changes on the ground relate to strong spectral or feature differences.
- The methods WC²VA and CTA, developed in Chapter 7, appreciably increased classification accuracy when compared to simple spectral CVA.
- The reference data indicates a probable increase in the rate of urban creep in Norwich from 0.14% per year (UKWIR 2010) between 1999 and 2006, to 0.24% per year between 2006 and 2010.

10.2 Research questions

Can a direct classification of multi-temporal VHR imagery adequately identify urban creep in Norwich?

No. After exhaustive experimentation, the resulting urban creep classifications were not deemed adequate. Specifically, the most successful classification achieved an urban creep accuracy of 52.15%, and would not provide meaningful quantitative results. This finding states that automatic urban creep classification is beyond the state-of-the-art. However, despite the application being extremely challenging urban creep was still identified, with a moderate user accuracy. Therefore, the methods demonstrated in this research could support qualitative evaluation and act as the foundation for semi-automated workflows.

Can a Change Vector Analysis of multi-temporal VHR imagery adequately identify urban creep in Norwich?

No. The experimentation exhaustively investigated common CVA approaches in the literature as well as the new methods developed for this thesis, WC²VA and CTA. None of these approaches delivered satisfactory results. However, it was found that, in this application, simple spectral CVA performed extremely poorly, having almost no utility, and the CVA results lagged far behind the direct classification results with a difference greater than other published comparisons. Furthermore, it was found that the new methods WC²VA and CTA significantly outperformed spectral CVA.

Has the rate of urban creep in Norwich changed since the last published estimate?

Yes, a probable increase has been observed. The latest estimate for Norwich was calculated as 0.24% per year between 2006 and 2010 compared to the previous estimate of 0.14% per year between 1999 and 2006. Because the previous estimate was not published with error bounds, we cannot say for certain that an increase has occurred. However, assuming both estimates possess similar accuracies, it is highly probable that the rate has increased.

10.3 Contribution to knowledge

Clarifies the capabilities of change detection techniques in detailed urban applications

It is a long-standing tenet of remote sensing change detection that changes on the ground relate to radiometric differences (Singh 1989), and that important changes can be isolated because they exhibit large radiometric differences (Bruzzone & Prieto 2000; Adar et al. 2014; Wang 2014b; Patra et al. 2011; Sinha & Kumar 2013). The findings are in direct conflict with this tenet. Spectral differences and compounded object-based feature differences describe urban creep very poorly, with improvements a result of mining complicated patterns in a large number of classification features. Of course, this finding originates from a very specific case, but it may be generalised to other large-scale urban VHR change detection research. Specifically, future researchers have direct evidence that small instances of land cover change, on a comparable spatial scale to scene illumination and view angle effects, may not be described by compounding spectral, or object-based feature differences.

This contribution is important because it contributes to clarify the status and limitations of VHR image urban change detection. Critics could argue that the outcome of this research is expected and it is entirely predictable that urban creep would be heavily confused with static land cover and multi-temporal VHR image artefacts. In fact, this hypothesis is discussed in Chapter 4. However, the discrepancy between the established tenet of remote sensing change detection and its application to VHR images is not clear. Chapter 9 has discussed a sense of ‘unease’ in the literature but no clear position. It is the position of this thesis that a strong relationship between image difference and real land cover change is scale and class dependent, and that the

relationship breaks down in large-scale urban VHR applications. This position strengthens and extends the views expressed by Bruzzone & Bovolo (2013) and Myint et al. (2011), in relation to the limitation of radiometric differencing. In light of this future research can be focused towards mining complicated patterns concerning context, texture and spatial pattern. Furthermore, the concept of avoiding UOA conflation introduced by Tewkesbury et al. (2015) and Chapter 2 is extended to help to clarify the ‘pixel vs object’ debate as discussed in the literature (G. Chen et al. 2012; Hussain et al. 2013; Boldt et al. 2012). Specifically, a further position of this thesis is that differencing pixel spectral values are not the problem, differencing spectral values alone under any UOA is the problem. Solving complex urban change detection problems is only achievable by identifying complex patterns in a range of object-based features.

Builds upon and extends existing CVA research

The CVA methods defined in this thesis build upon existing works in four steps:

1. They further test the hypothesis that CVA is an ideal tool to identify specific thematic changes (Bovolo & Bruzzone 2007; Bovolo et al. 2012). While, in most cases, urban creep could not be uniquely identified; some specific examples did form unique WC²VA clusters.
2. This research builds upon the work of He et al. (2011) by not only considering spectral and textural features, but considering a range of object-based features.
3. Defines WC²VA as a mechanism to normalise output magnitude and direction under varying numbers of features. This concept builds upon the work of Bovolo et al. (2012) and Tian et al. (2014), applying weights to n-dimensional CVA.
4. CTA extends WC²VA further to include the baseline vector. This formulation solves CV ambiguity as reported by Johnson & Kasischke (1998) and Tewkesbury et al. (2015).

Contributes to urban creep knowledge

This thesis provides a much-needed update on the status of urban creep in the UK. Specifically, nationwide estimates of urban creep are overdue a revision (UKWIR 2010), and the findings of this research not only contribute to this revision but also provide more evidence of the need for a revision. Specifically, the findings indicate an acceleration in urban creep rates for Norwich, which could be wider reaching. This information would be of great interest to environmental engineers and urban planners tasked with minimising urban flooding.

Furthermore, the remote sensing results demonstrate and quantify the practical limitations of urban creep identification. Specifically, as a result of this work, urban drainage stakeholders would be able to make an informed decision when considering remote sensing technology for future urban creep research.

10.4 Implications

This research has an academic impact, clarifying the status and limitations of VHR urban change detection. This work helps to clarify the position and utility of change detection when applied to complex urban change problems by providing clear evidence that radiometric differences cannot be relied upon to derive large-scale change information. Moreover, neither can the compounded differences of object-based features. Instead, research should focus on mining complex patterns in a range of object-based features, at multiple scales, many of which will not contain appreciable differences. In turn, this position helps to clarify the ‘pixel vs. object’ debate by not conflating the UOA. Specifically, pixels are not the problem. Relying on spectral differences alone under any UOA is the problem.

This research has an academic impact extending the definition of CVA. It was found that the new extensions to CVA, WC²VA and CTA considerably improved classification performance. While it was shown that even these techniques are not suited to large-scale urban VHR image change detection, the techniques would be of interest to researchers investigating other monitoring applications. Specifically, those monitoring broad land cover and land use change with moderate resolution, fixed frame sensors such as Landsat and Sentinel-2. These new techniques would facilitate more exploration with CVA, investigating the effect of mixing spectral inputs with a range of

different textural and contextual features. Furthermore, in CTA, a clear framework is established to bypass CV ambiguity.

Lastly, this research has a societal and economic impact. Specifically, urban drainage system stakeholders would view a potential acceleration of the rate of urban creep with great interest because an increase could elevate urban flood risk. Therefore, this research could be the impetus of a nationwide review of urban creep rates to establish any national trends. If an increase is identified, then this would result in updated flood risk impact assessments and possibly a policy review. Furthermore, there is an opportunity to help reduce the causes of urban creep via public engagement. The publication of statistics reporting increased urban creep, and the potential implications this may have to local flooding and environmental quality could help to engage the public towards more sustainable development.

10.5 Recommendations for future work

The findings of this research have opened up avenues for future work to both corroborate these findings and further advance their principles. In light of this, future work could follow in four areas:

More research investigating large scale VHR image change detection

The broad aim of such research is to bridge the gap between the resolution of images currently available and the scale of change information reported from them. Such research would act to corroborate the thesis findings regarding the limitation of spectral differencing and the importance of a variety of different object-based features.

Further explore the novel methods derived in this research

WC²VA and CTA should be tested against other applications and datasets to see if the accuracy boosts observed in this research can be replicated. The specific effectiveness of the shadow and viewing angle modelling and the segmentation method CEMRS should also be ratified by further research.

Appendix



Figure A-1: Example field map, sample ID 7388, panel 1. Panel 1 of the field map design contains the 2006 reference image, overlaid with MasterMapTM (top) and smaller scale overview map (bottom) to aid with location. Backing imagery ©Airbus Defence and Space (2016). Overview base mapping courtesy of ESRI and its data contributors 2016.

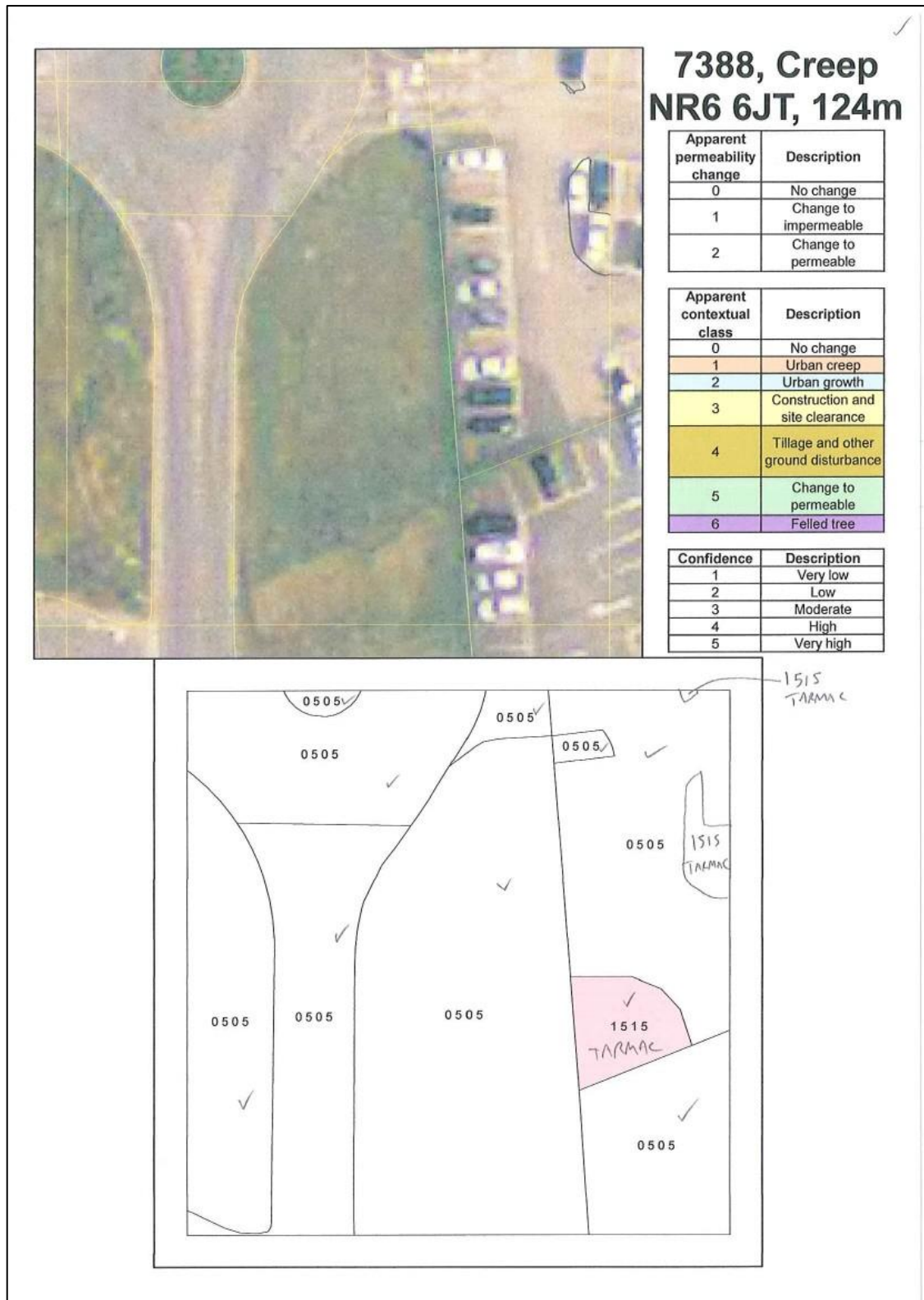


Figure A-2: Example field map, sample ID 7388, panel 2. Panel 2 of the field map design contains the 2010 reference image, overlaid with MasterMap™ (top) and the reference mapping (bottom) marked up in the field. This panel also contains sample information such as its stratification class, nearest postcode and distance and a nomenclature key. Backing imagery © Airbus Defence and Space (2016).

Table A-1: Change proportions by stratum.

Property	Notation	Stratum								
		No change candidate strata			Urban creep candidate strata			Other change candidate strata		
Number of PSU units	N_h	17377			6977			902		
Area	m^2	43442500			17442500			2255000		
Stratum weight	$W_h = \frac{N_h}{N}$	0.6880			0.2763			0.0357		
PSU samples	n_h	50			50			54		
Sampling fraction	$f_h = \frac{n_h}{N_h}$	0.0029			0.0072			0.0599		
Class	n/a	No change	Urban creep	Other change	No change	Urban creep	Other change	No change	Urban creep	Other change
PSU class proportion mean	$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$	0.9569	0.0023	0.0408	0.9625	0.0250	0.0125	0.4708	0.0278	0.5014
PSU class proportion variance	$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$	0.0229	0.0000	0.0229	0.0008	0.0009	0.0003	0.0939	0.0086	0.1105

Table A-2: Change proportions by population.

Property	Notation	Population		
Number of PSU units	N	25256		
Area	m^2	63140000		
PSU samples	n	154		
Sampling fraction	$f = \frac{n}{N}$	0.0061		
Class	n/a	No change	Urban creep	Other change
Class proportion mean	$\bar{y}_{st} = \sum_{h=1}^L W_h \bar{y}_h$	0.9411	0.0095	0.0494
Class area	m^2	59420739	597082	3122225
Variance of the class proportion mean	$s^2(\bar{y}_{st}) = \sum_{h=1}^L \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^L \frac{W_h s_h^2}{N}$	0.0002190	0.0000017	0.0002194
Standard error of the class proportion mean	$s(\bar{y}_{st})$	0.01480	0.00132	0.01481
95% Confidence limit	$\mp 1.96 \times s(\bar{y}_{st})$	0.029001	0.00259	0.029003

```

pif_count_threshold = 50 #The target number of PIFs per radiometric bin
tile_dim = 5000 #The target tile row and column dimension
radiometric_bins = 20 #The number of bins with which to split the radiometric range by
SAM_threshold_start = 0.01 #The initial SAM threshold
SCM_threshold_start = 0.99 #The initial SCM threshold
threshold_increment = 0.01 #The amount to increment the threshold by
pif_size = 2 #The minimum number of pixels for a PIF
tiles = tile_images(T0,T1, tile_dim) #Create spatial tiles from the input images
pif_array = [ ] #Create empty array to store PIFs

for tile in tiles: #Iterate through each of the spatial tiles
    SAM = calc_SAM(T0tile, T1tile) #Calculate the SAM image
    SCM = calc_SCM(T0tile, T1tile) #Calculate the SCM image
    bins = bin_radiometry(T0tile, T1tile, radiometric_bins) #Split the radiometric range into bins

    for bin in bins: #Iterate through each of the radiometric bins
        pif_count = 0 #Initiate with no PIFs
        SAM_threshold = SAM_threshold_start #Initiate the threshold
        SCM_threshold = SCM_threshold_start #Initiate the threshold
        pif_array_temp = [ ] #Create temp empty array to store PIFs for bin

        while pif_count <= pif_count_threshold: #Iterate until enough PIFs are identified
            SAM_mask = threshold_by_percentile(SAM, SAM_threshold) #SAM PIFs by percentile
            SCM_mask = threshold_by_percentile(SCM, SCM_threshold) #SCM PIFs by percentile
            if SAM_mask == 1 and SCM_mask == 1: #Where the two intersect
                pif_array_temp.append() #Create a PIF
            pif_count = count(pif_array_temp) #Update the number of PIFs per bin
            SAM_threshold += threshold_increment #Increase the spectral angle
            SCM_threshold -= threshold_increment #Decrease the correlation
        pif_array.append(pif_array_temp) #Retain the PIFs

```

Figure A-3: PIF identification algorithm pseudo code.

```

flight_path #The approximate 3D flight path of the aircraft
FOV_extent_par = 2 metres #The sensor field of view parallel to the flight path. This is set to mirror the DSM resolution
FOV_extent_lat = 5000 metres #The lateral extent of the field of view either side of the sensor.
observations = get_obs(flight_path, FOV_extent_par) #Break the continuous flight line into observation coordinates
                                                    every 2m
FOV_hidden_surface_list = [ ] #Create an empty list for the individual FOV visibility raster images

for observation in observations: #Iterate along the flight path, one observation at a time
    flight_az = get_az(flight_path, observation) #Get the flight line azimuth for the observation
    FOV_line = get_FOV_line(observation, flight_az, FOV_extent_lat) #Draw a line either side of the observation
                                                                    perpendicular to the flight line
    FOV = buffer(FOV_line, FOV_extent_par/2) #Buffer FOV line by one DSM pixel to simulate the sensor FOV
    DSM_FOV = clip(DSM, FOV) #Clip the DSM to the FOV
    FOV_hidden_surface = viewshed(DSM_FOV, observation) #Create a binary raster depicting the surfaces
                                                         hidden from the observation FOV
    FOV_hidden_surface_list.append(FOV_hidden_surface) #Gather the hidden surface raster images together

hidden_surfaces = mosaic(FOV_hidden_surface_list) #Mosaic the observations into a single raster image

```

Figure A-4: Pseudo code outlining the hidden surface detection algorithm.

```

Scale #The MRS scale parameter for the new parent IOL
Shape #The MRS shape parameter for the new parent IOL
Compactness #The MRS compactness parameter for the new parent IOL
CompactnessThr #The compactness threshold for the segmentation. Objects with lower compactness will be re-
shaped.
Ln-1 #The child IOL to be aggregated to larger, compact objects

Ln = MRS(Ln-1, Scale, Shape, Compactness) #Establish the parent IOL using the MRS algorithm

continue_loop = 1 #Initiate the loop control variable
while continue_loop == 1:
    complex_objects = [ ] #Create an empty list to store complex objects
    compact_objects = [ ] #Create an empty list to store the compact objects
    for object in Ln: #Iterate through each object
        if compactness > CompactnessThr: #Identify complex objects
            complex_objects.append(object) #Store complex objects
        else:
            compact_objects.append(object) #Store compact objects
    if count(complex_objects) > 0: #Test the number of complex objects
        Convert_to_sub-objects(complex_objects, Ln-1) #Split the complex objects into the smaller, child
                                                         objects
        for object in complex_objects: #Iterate through each complex object
            eliminate_by_border(object, compact_objects) #Merge object with the compact object
                                                         with the greatest shared border
        for object in Ln: #Iterate through each object
            if object shared_border > 0.5: #Identify objects largely surrounded by another object
                eliminate_by_border(object, Ln) #Merge object with its surroundings
    else:
        continue_loop = 0 #End the loop

```

Figure A-5: Pseudo code describing the CEMRS algorithm.

Table A-3: Full feature listing.

Feature	Description	Image interpretation principle	Scale	Source data state	Analysis unit comparability
T0 L1 DSM	Level 1 object mean, 2006 DSM	Height	Large	T0	Comparable
T0 L1 Mean diff. to. 0pix	Level 1 object spatial auto-correlation, 2006 mean albedo difference to adjacent neighbours	Pattern			
T0 L1 Standard dev. 0pix	Level 1 object spatial auto-correlation, 2006 standard deviation of albedo differences to adjacent neighbours				
T0 L1 Hill shaded DSM	Level 1 object mean, 2006 hill shaded relief and shadows	Shadow			
T0 L1 Standard Deviation	Level 1 object texture, 2006 standard deviation	Texture			
T0 L1 GLCM Entropy	Level 1 object texture, 2006 GLCM entropy				
T0 L1 GLCM Homogeneity	Level 1 object texture, 2006 GLCM homogeneity				
T0 L1 GLCM Mean	Level 1 object texture, 2006 GLCM mean				
T0 L1 GLCM Std Dev	Level 1 object texture, 2006 GLCM standard deviation				
T0 L1 Blue	Level 1 object mean, 2006 blue	Tone			
T0 L1 Green	Level 1 object mean, 2006 green				
T0 L1 Red	Level 1 object mean, 2006 red				
T0 L1 NIR	Level 1 object mean, 2006 near infra-red				
T0 L1 Hidden surfaces	Level 1 object mean, 2006 visibility and feature parallax	View angle			
T0 L2 DSM	Level 2 object mean, 2006 DSM	Height	Moderate		
T0 L1 Mean diff. to. 100pix	Level 1 object spatial auto-correlation, 2006 mean albedo difference to adjacent neighbours within 100 pixels	Pattern			
T0 L1 Standard dev. 100pix	Level 1 object spatial auto-correlation, 2006 standard deviation of albedo differences to adjacent neighbours within 100 pixels				
T0 L2 Hill shaded DSM	Level 2 object mean, 2006 hill shaded relief and shadows	Shadow			
T0 L2 Standard Deviation	Level 2 object texture, 2006 standard deviation	Texture			
T0 L2 GLCM Entropy	Level 2 object texture, 2006 GLCM entropy				
T0 L2 GLCM Homogeneity	Level 2 object texture, 2006 GLCM homogeneity				
T0 L2 GLCM Mean	Level 2 object texture, 2006 GLCM mean				
T0 L2 GLCM Std Dev	Level 2 object texture, 2006 GLCM standard deviation				
T0 L2 Blue	Level 2 object mean, 2006 blue	Tone			

Feature	Description	Image interpretation principle	Scale	Source data state	Analysis unit comparability
T0 L2 Green	Level 2 object mean, 2006 green				
T0 L2 Red	Level 2 object mean, 2006 red				
T0 L2 NIR	Level 2 object mean, 2006 near infra-red				
T0 L2 Hidden surfaces	Level 2 object mean, 2006 visibility and feature parallax	View angle			
T0 L3 DSM	Level 3 object mean, 2006 DSM	Height			
T0 L1 Mean diff. to. 200pix	Level 1 object spatial auto-correlation, 2006 mean albedo difference to adjacent neighbours within 200 pixels	Pattern			
T0 L1 Standard dev. 200pix	Level 1 object spatial auto-correlation, 2006 standard deviation of albedo differences to adjacent neighbours within 200 pixels				
T0 L3 Hill shaded DSM	Level 3 object mean, 2006 hill shaded relief and shadows	Shadow			
T0 L3 Standard Deviation	Level 3 object texture, 2006 standard deviation	Texture			
T0 L3 GLCM Entropy	Level 3 object texture, 2006 GLCM entropy				
T0 L3 GLCM Homogeneity	Level 3 object texture, 2006 GLCM homogeneity				
T0 L3 GLCM Mean	Level 3 object texture, 2006 GLCM mean				
T0 L3 GLCM Std Dev	Level 3 object texture, 2006 GLCM standard deviation				
T0 L3 Blue	Level 3 object mean, 2006 blue	Tone			
T0 L3 Green	Level 3 object mean, 2006 green				
T0 L3 Red	Level 3 object mean, 2006 red				
T0 L3 NIR	Level 3 object mean, 2006 near infra-red				
T0 L3 Hidden surfaces	Level 3 object mean, 2006 visibility and feature parallax	View angle			
Δ L1 DSM	Level 1 object mean, DSM, 2006 to 2010 difference	Height			
Δ L1 Mean diff. to. 0pix	Level 1 object spatial auto-correlation, mean albedo difference to adjacent neighbours, 2006 to 2010 difference	Pattern			
Δ L1 Standard dev. 0pix	Level 1 object spatial auto-correlation, standard deviation of albedo differences to adjacent neighbours, 2006 to 2010 difference				
Δ L1 Hill shaded DSM	Level 1 object mean, hill shaded relief and shadows, 2006 to 2010 difference	Shadow			
Δ L1 Standard Deviation	Level 1 object texture, standard deviation, 2006 to 2010 difference	Texture			

Feature	Description	Image interpretation principle	Scale	Source data state	Analysis unit comparability
Δ L1 GLCM Entropy	Level 1 object texture, GLCM entropy, 2006 to 2010 difference				
Δ L1 GLCM Homogeneity	Level 1 object texture, GLCM homogeneity, 2006 to 2010 difference				
Δ L1 GLCM Mean	Level 1 object texture, GLCM mean, 2006 to 2010 difference				
Δ L1 GLCM Std Dev	Level 1 object texture, GLCM standard deviation, 2006 to 2010 difference				
Δ L1 Blue	Level 1 object mean, blue, 2006 to 2010 difference	Tone			
Δ L1 Green	Level 1 object mean, green, 2006 to 2010 difference				
Δ L1 Red	Level 1 object mean, red, 2006 to 2010 difference				
Δ L1 NIR	Level 1 object mean, near infra-red, 2006 to 2010 difference				
Δ L1 Hidden surfaces	Level 1 object mean, visibility and feature parallax, 2006 to 2010 difference	View angle			
Δ L2 DSM	Level 2 object mean, DSM, 2006 to 2010 difference	Height			
Δ L1 Mean diff. to. 100pix	Level 1 object spatial auto-correlation, mean albedo difference to adjacent neighbours within 100 pixels, 2006 to 2010 difference	Pattern			
Δ L1 Standard dev. 100pix	Level 1 object spatial auto-correlation, standard deviation of albedo differences to adjacent neighbours within 100 pixels, 2006 to 2010 difference				
Δ L2 Hill shaded DSM	Level 2 object mean, hill shaded relief and shadows, 2006 to 2010 difference	Shadow			
Δ L2 Standard Deviation	Level 2 object texture, standard deviation, 2006 to 2010 difference	Texture			
Δ L2 GLCM Entropy	Level 2 object texture, GLCM entropy, 2006 to 2010 difference				
Δ L2 GLCM Homogeneity	Level 2 object texture, GLCM homogeneity, 2006 to 2010 difference				
Δ L2 GLCM Mean	Level 2 object texture, GLCM mean, 2006 to 2010 difference				
Δ L2 GLCM Std Dev	Level 2 object texture, GLCM standard deviation, 2006 to 2010 difference				
Δ L2 Blue	Level 2 object mean, blue, 2006 to 2010 difference	Tone			

Feature	Description	Image interpretation principle	Scale	Source data state	Analysis unit comparability
Δ L2 Green	Level 2 object mean, green, 2006 to 2010 difference				
Δ L2 Red	Level 2 object mean, red, 2006 to 2010 difference				
Δ L2 NIR	Level 2 object mean, near infra-red, 2006 to 2010 difference				
Δ L2 Hidden surfaces	Level 2 object mean, visibility and feature parallax, 2006 to 2010 difference	View angle			
Δ L3 DSM	Level 3 object mean, DSM, 2006 to 2010 difference	Height	Small		
Δ L1 Mean diff. to. 200pix	Level 1 object spatial auto-correlation, mean albedo difference to adjacent neighbours within 200 pixels, 2006 to 2010 difference	Pattern			
Δ L1 Standard dev. 200pix	Level 1 object spatial auto-correlation, standard deviation of albedo differences to adjacent neighbours within 200 pixels, 2006 to 2010 difference				
Δ L3 Hill shaded DSM	Level 3 object mean, hill shaded relief and shadows, 2006 to 2010 difference	Shadow			
Δ L3 Standard Deviation	Level 3 object texture, standard deviation, 2006 to 2010 difference	Texture			
Δ L3 GLCM Entropy	Level 3 object texture, GLCM entropy, 2006 to 2010 difference				
Δ L3 GLCM Homogeneity	Level 3 object texture, GLCM homogeneity, 2006 to 2010 difference				
Δ L3 GLCM Mean	Level 3 object texture, GLCM mean, 2006 to 2010 difference				
Δ L3 GLCM Std Dev	Level 3 object texture, GLCM standard deviation, 2006 to 2010 difference				
Δ L3 Blue	Level 3 object mean, blue, 2006 to 2010 difference	Tone			
Δ L3 Green	Level 3 object mean, green, 2006 to 2010 difference				
Δ L3 Red	Level 3 object mean, red, 2006 to 2010 difference				
Δ L3 NIR	Level 3 object mean, near infra-red, 2006 to 2010 difference				
Δ L3 Hidden surfaces	Level 3 object mean, visibility and feature parallax, 2006 to 2010 difference	View angle			
L1 Direction	Level 1 object orientation, direction/azimuth	Association	Large	T0 & T1	Non-comparable
L1 Asymmetry	Level 1 object shape, asymmetry	Shape			

Feature	Description	Image interpretation principle	Scale	Source data state	Analysis unit comparability
L1 Compactness	Level 1 object shape, compactness				
L1 Density	Level 1 object shape, density				
L1 Elliptical fit	Level 1 object shape, elliptical fit				
L1 Length/Width	Level 1 object shape, length/Width				
L1 Rectangular fit	Level 1 object shape, rectangular fit				
L1 Roundness	Level 1 object shape, roundness				
L1 Area	Level 1 object size, area	Size			
L1 Length	Level 1 object size, length				
L1 Width	Level 1 object size, width				
L2 Direction	Level 2 object orientation, direction/azimuth	Association	Moderate		
L2 Asymmetry	Level 2 object shape, asymmetry	Shape			
L2 Compactness	Level 2 object shape, compactness				
L2 Density	Level 2 object shape, density				
L2 Elliptical fit	Level 2 object shape, elliptical fit				
L2 Length/Width	Level 2 object shape, length/Width				
L2 Rectangular fit	Level 2 object shape, rectangular fit				
L2 Roundness	Level 2 object shape, roundness				
L2 Area	Level 2 object size, area	Size			
L2 Length	Level 2 object size, length				
L2 Width	Level 2 object size, width				
L3 Direction	Level 3 object orientation, direction/azimuth	Association	Small		
L3 Asymmetry	Level 3 object shape, asymmetry	Shape			
L3 Compactness	Level 3 object shape, compactness				
L3 Density	Level 3 object shape, density				
L3 Elliptical fit	Level 3 object shape, elliptical fit				
L3 Length/Width	Level 3 object shape, length/Width				
L3 Rectangular fit	Level 3 object shape, rectangular fit				
L3 Roundness	Level 3 object shape, roundness				
L3 Area	Level 3 object size, area	Size			
L3 Length	Level 3 object size, length				
L3 Width	Level 3 object size, width				

Table A-4: The top 15 representative features from all available features.

Number of selected features	1	TO L2 Blue															
	2	Δ L1 GLCM Mean	TO L1 Blue														
	3	Δ L2 GLCM Mean	L1 Length	TO L1 Blue													
	4	Δ L2 GLCM Mean	L3 Length	TO L1 Blue	TO L3 NIR												
	5	Δ L3 Green	Δ L1 Mean diff. to. 100pix	L3 Length	TO L1 Blue	TO L3 NIR											
	6	Δ L3 Green	Δ L1 Mean diff. to. 100pix	L3 Length	TO L2 Green	TO L3 NIR	TO L1 Mean diff. to. 100pix										
	7	Δ L1 Blue	Δ L3 Green	Δ L3 Standard Deviation	L3 Length	TO L2 Green	TO L3 NIR	TO L1 Mean diff. to. 100pix									
	8	Δ L1 Blue	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation	L3 Length	TO L2 Green	TO L3 NIR	TO L1 Mean diff. to. 100pix								
	9	Δ L1 Blue	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation	L3 Length	TO L1 Blue	TO L2 Green	TO L2 GLCM Std Dev	TO L3 NIR							
	10	Δ L1 Blue	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation	L3 Length	L1 Length/Width	TO L1 Blue	TO L2 Green	TO L2 GLCM Std Dev	TO L3 Hillshaded DSM						
	11	Δ L1 Blue	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation	L3 Length	L1 Length/Width	TO L1 Blue	TO L2 Green	TO L2 DSM	TO L2 GLCM Std Dev	TO L3 Hillshaded DSM					
	12	Δ L1 Blue	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation	L3 Elliptical fit	L3 Length	L1 Length/Width	TO L1 Blue	TO L2 Green	TO L2 DSM	TO L2 GLCM Std Dev	TO L3 Hillshaded DSM				
	13	Δ L1 Blue	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation	L3 Elliptical fit	L3 Length	L1 Length/Width	L1 Rectangular fit	TO L1 Blue	TO L2 Green	TO L2 DSM	TO L2 GLCM Std Dev	TO L3 Hillshaded DSM			
	14	Δ L1 Blue	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation	L3 Elliptical fit	L3 Length	L1 Length/Width	L1 Rectangular fit	TO L1 Blue	TO L2 Green	TO L2 NIR	TO L2 DSM	TO L2 GLCM Std Dev	TO L3 Hillshaded DSM		
	15	Δ L1 Blue	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation	L3 Elliptical fit	L3 Length	L1 Length/Width	L1 Rectangular fit	TO L1 Blue	TO L2 Green	TO L2 NIR	TO L2 DSM	TO L2 GLCM Homogeneity	TO L2 GLCM Std Dev	TO L3 Hillshaded DSM	

Table A-5: The top 15 representative features from all difference features.

Number of selected features	1	Δ L2 GLCM Mean															
	2	Δ L2 DSM	Δ L2 GLCM Mean														
	3	Δ L2 DSM	Δ L2 GLCM Mean	Δ L1 Mean diff. to. 100pix													
	4	Δ L1 Blue	Δ L2 DSM	Δ L2 GLCM Mean	Δ L3 Standard Deviation												
	5	Δ L1 Blue	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L3 Green	Δ L3 Standard Deviation											
	6	Δ L1 Blue	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation										
	7	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation									
	8	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 Hidden surfaces	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation								
	9	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 NIR	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 Hidden surfaces	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation							
	10	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 NIR	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 Hidden surfaces	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation	Δ L1 Standard dev. 100pix						
	11	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 Green	Δ L2 NIR	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 Hidden surfaces	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation	Δ L1 Standard dev. 100pix					
	12	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 Green	Δ L2 NIR	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 Hidden surfaces	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation	Δ L1 Standard dev. 100pix	Δ L1 GLCM Homogeneity				
	13	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 Green	Δ L2 NIR	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 Hidden surfaces	Δ L2 Standard Deviation	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation	Δ L1 Standard dev. 100pix	Δ L1 GLCM Homogeneity			
	14	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 Green	Δ L2 NIR	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 Hidden surfaces	Δ L2 Standard Deviation	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L3 Standard Deviation	Δ L1 Standard dev. 100pix	Δ L1 Hidden surfaces	Δ L1 GLCM Homogeneity		
	15	Δ L1 Green	Δ L1 GLCM Std Dev	Δ L2 Green	Δ L2 NIR	Δ L2 DSM	Δ L2 Hillshaded DSM	Δ L2 Hidden surfaces	Δ L2 Standard Deviation	Δ L2 GLCM Homogeneity	Δ L3 Green	Δ L1 NIR	Δ L3 Standard Deviation	Δ L1 Standard dev. 100pix	Δ L1 Hidden surfaces	Δ L1 GLCM Homogeneity	

References

- Abd El-Kawy, O.R., Rød, J.K., Ismail, H.A. & Suliman, A.S., 2011. Land use and land cover change detection in the western Nile delta of Egypt using remote sensing data. *Applied Geography*, 31(2), pp.483–494.
- Adar, S., Shkolnisky, Y. & Ben Dor, E., 2014. A new approach for thresholding spectral change detection using multispectral and hyperspectral image data, a case study over Sokolov, Czech republic. *International Journal of Remote Sensing*, 35(4), pp.1563–1584.
- Adelabu, S., Mutanga, O. & Adam, E., 2015. Testing the reliability and stability of the internal accuracy assessment of random forest for classifying tree defoliation levels using different validation methods. *Geocarto International*, 30(7), pp.810–821.
- Aguirre-Gutiérrez, J., Seijmonsbergen, A.C. & Duivenvoorden, J.F., 2012. Optimizing land cover classification accuracy for change detection, a combined pixel-based and object-based approach in a mountainous area in Mexico. *Applied Geography*, 34(5), pp.29–37.
- Airbus Defence & Space, 2016. Airbus Defence & Space DSM. Available at: <http://www.intelligence-airbusds.com/en/6064-elevation1-highly-precise-altimetric-information>.
- Allitt, M. & Tewkesbury, A.P., 2009. Investigations into Urban Creep at 5 Cities. In *WaPUG Autumn Conference*. pp. 1–10.
- American Congress on Surveying and Mapping, 1994. *Glossary of the mapping sciences*,
- Anglian Water, 2010. *Anglian Water Final Business Plan Part B5 Maintaining the Wastewater Supply / Demand Balance*,
- Atkinson, P.M., 2006. Resolution Manipulation and sub-Pixel Mapping. In S. M. De Jong & F. van der Meer, eds. *Remote Sensing Image Analysis: Including the Spatial Domain*. Springer, pp. 51–70.

- Atkinson, P.M. & Foody, G.M., 2002. Uncertainty in Remote Sensing and GIS: Fundamentals. In G. M. Foody & P. M. Atkinson, eds. *Uncertainty in Remote Sensing and GIS*. Wiley, pp. 1–18.
- Baatz, M., Hoffmann, C. & Willhauck, G., 2008. Progressing from object-based to object-orientated image analysis. In T. Blaschke, S. Lang, & G. J. Hay, eds. *Object Based Image Analysis*. Springer-Verlag, pp. 29–42.
- Bang, K.-I., Habib, A., Kim, K. & Shin, S., 2007. Comprehensive analysis of alternative methodologies for true ortho-photo generation from high resolution satellite and aerial imagery. In *ASPRS 2007 Annual Conference*.
- Baraldi, A. & Boschetti, L., 2012. Operational Automatic Remote Sensing Image Understanding Systems: Beyond Geographic Object-Based and Object-Oriented Image Analysis (GEOBIA/GEOOIA). Part 1: Introduction. *Remote Sensing*, 4(12), pp.2694–2735.
- Barazzetti, L., 2016. Sliver Removal in Object-Based Change Detection from VHR Satellite Images. *Photogrammetric engineering and remote sensing*, 82(2), pp.161–168.
- Barazzetti, L., Brumana, R., Cuca, B. & Previtali, M., 2015. Change detection from very high resolution satellite time series with variable off-nadir angle. In *Third International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2015)*.
- Bazi, Y., Melgani, F. & Al-Sharari, H.D., 2010. Unsupervised Change Detection in Multispectral Remotely Sensed Imagery With Level Set Methods. *IEEE Transactions on Geoscience and Remote Sensing*, 48(8), pp.3178–3187.
- Bennett, S., Woods, T., liyanagec, W.M. & Smith, D.L., 1991. A Simplified General Method for Cluster-Sample Surveys of Health in Developing Countries. *World health statistics quarterly*, 44(3), pp.98–106.
- Bianchetti, R.A. & Maceachren, A.M., 2015. Cognitive Themes Emerging from Air Photo Interpretation Texts Published to 1960. *ISPRS International Journal of Geo-Information*, 4, pp.551–571.

- Biging, G.S., Colby, D.R. & Congalton, R.G., 1998. Sampling systems for Change Detection Accuracy Assessment. In R. S. Lunetta & C. D. Elvidge, eds. *Remote Sensing Change Detection*. ANN ARBOR PRESS, pp. 281–308.
- Blair, R.B., 1996. Land use and avian species diversity along an urban gradient. *Ecological applications*, 6(2), pp.506–519.
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), pp.2–16.
- Bluesky, 2016. Bluesky DSM. Available at: <http://www.bluesky-world.com/#!standard-height-data/csdq>.
- Boldt, M., Thiele, A. & Schulz, K., 2012. Object-based urban change detection analyzing high resolution optical satellite images. In *PROCEEDINGS- SPIE THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING: Earth Resources and Environmental Remote Sensing/GIS Applications III Vol 8538*. p. 85380E–1 to 85380E–9.
- Bontemps, S., Langer, A. & Defourny, P., 2012. Monitoring forest changes in Borneo on a yearly basis by an object-based change detection algorithm using SPOT-VEGETATION time series. *International Journal of Remote Sensing*, 33(15), pp.4673–4699.
- Bouziani, M., Goïta, K. & He, D.-C., 2010. Automatic change detection of buildings in urban environment from very high spatial resolution images using existing geodatabase and prior knowledge. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), pp.143–153.
- Bovolo, F., 2009. A Multilevel Parcel-Based Approach to Change Detection in Very High Resolution Multitemporal Images. *IEEE Geoscience and Remote Sensing Letters*, 6(1), pp.33–37.
- Bovolo, F. & Bruzzone, L., 2007. A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain. *IEEE Transactions on Geoscience and Remote Sensing*, 45(1), pp.218–236.

- Bovolo, F. & Bruzzone, L., 2015. The Time Variable in Data Fusion: A Change Detection Perspective. *IEEE Geoscience and Remote Sensing Magazine*, 3(3), pp.8–26.
- Bovolo, F., Marchesi, S. & Bruzzone, L., 2012. A framework for automatic and unsupervised detection of multiple changes in multitemporal images. *IEEE Transactions on Geoscience and Remote Sensing*, 50(6), pp.2196–2212.
- Bradford Metropolitan District Council, Leeds City Council, Yorkshire Water Services, The Environment Agency & Pennine Water Group, 2008. *TR344 – River Aire Strategic Studies Report*,
- Bruzzone, L. & Bovolo, F., 2013. A Novel Framework for the Design of Change-Detection Systems for Very-High-Resolution Remote Sensing Images. *Proceedings of the IEEE*, 101(3), pp.609–630.
- Bruzzone, L. & Prieto, D.F., 2000. Automatic analysis of the difference image for unsupervised change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 38(3), pp.1171–1182.
- Butler, D. & Davies, J.W., 2011. *Urban Drainage* Third Edit., Spon Press.
- Canty, M.J., Nielsen, A.A. & Schmidt, M., 2004. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sensing of Environment*, 91(3–4), pp.441–451.
- Cao, G., Li, Y., Liu, Y. & Shang, Y., 2014. Automatic change detection in high-resolution remote-sensing images by means of level set evolution and support vector machine classification. *International Journal of Remote Sensing*, 35(16), pp.6255–6270.
- Carbó-Ramírez, P. & Zuria, I., 2011. The value of small urban greenspaces for birds in a Mexican city. *Landscape and Urban Planning*, 100(3), pp.213–222.
- Carleer, A.P. & Wolff, E., 2006. Urban land cover multi-level region-based classification of VHR data by selecting relevant features. *International Journal of Remote Sensing*, 27(6), pp.1035–1051.
- Carvalho Júnior, O.A., Guimarães, R.F., Gillespie, A., Silva, N.C. & Gomes, R., 2011. A

- New Approach to Change Vector Analysis Using Distance and Similarity Measures. *Remote Sensing*, 3(12), pp.2473–2493.
- Carvalho Júnior, O.A., Guimarães, R.F., Silva, N.C., Gillespie, A., Gomes, R., Silva, C. & de Carvalho, A., 2013. Radiometric Normalization of Temporal Images Combining Automatic Detection of Pseudo-Invariant Features from the Distance and Similarity Spectral Measures, Density Scatterplot Analysis, and Robust Regression. *Remote Sensing*, 5(6), pp.2763–2794.
- Castilla, G. & Hay, G.J., 2008. Image objects and geographic objects. In T. Blaschke, S. Lang, & G. J. Hay, eds. *Object-Based Image Analysis*. Springer, pp. 91–110.
- Chehata, N., Orny, C., Boukir, S. & Guyon, D., 2011. Object-based forest change detection using high resolution satellite images. *Remote Sensing and Spatial Information Sciences*, 38.
- Chen, C.J., 2011. *Physics of Solar Energy*, Hoboken : Wiley.
- Chen, G., Hay, G.J., Carvalho, L.M.T. & Wulder, M.A., 2012. Object-based Change Detection. *International Journal of Remote Sensing*, 33(14), pp.4434–4457.
- Chen, X., Chen, J., Shi, Y. & Yamaguchi, Y., 2012. An automated approach for updating land cover maps based on integrated change detection and classification methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, 71, pp.86–95.
- Chou, T.Y., Lei, T.C., Wan, S. & Yang, L.S., 2005. Spatial knowledge databases as applied to the detection of changes in urban land use. *International Journal of Remote Sensing*, 26(14), pp.3047–3068.
- Cichosz, P., 2014. *Data Mining Algorithms : Explained Using R*, Wiley.
- Cochran, W.G., 1977. *Sampling Techniques* Third edit., John Wiley & Sons, Inc.
- Cohen, W.B. & Fiorella, M., 1998. Comparison of Methods for Detecting Conifer Forest Change with Thematic Mapper Imagery. In R. S. Lunetta & C. D. Elvidge, eds. *Remote Sensing Change Detection*. ANN ARBOR PRESS, pp. 89–99.
- Comber, A., Brunsdon, C. & Farmer, C.J.Q., 2012. Community detection in spatial networks: Inferring land use from a planar graph of land cover objects.

- International Journal of Applied Earth Observation and Geoinformation*, 18, pp.274–282.
- Comber, A., Fisher, P.F. & Wadsworth, R., 2004. Assessment of a Semantic Statistical Approach to Detecting Land Cover Change Using Inconsistent Data Sets. *Photogrammetric Engineering & Remote Sensing*, 70(8), pp.931–938.
- Comber, A., Law, A.N.R. & Lishman, J.R., 2004. Application of knowledge for automated land cover change monitoring. *International Journal of Remote Sensing*, 25(16), pp.3177–3192.
- Comber, A., Umezaki, M., Zhou, R., Ding, Y., Li, Y., Fu, H., Jiang, H. & Tewkesbury, A.P., 2012. Using shadows in high-resolution imagery to determine building height. *Remote Sensing Letters*, 3(7), pp.551–556.
- Congalton, R.G. & Green, K., 2009. *Assessing the Accuracy of Remotely Sensed Data* Second Edi., CRC Press, Taylor & Francis Group.
- Coops, N.C., Gillanders, S.N., Wulder, M. a., Gergel, S.E., Nelson, T. & Goodwin, N.R., 2010. Assessing changes in forest fragmentation following infestation using time series Landsat imagery. *Forest Ecology and Management*, 259(12), pp.2355–2365.
- Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B. & Lambin, E., 2004. Review Article Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25(9), pp.1565–1596.
- Coulter, L., Hope, A.S. & Stow, D.A., 2011. Time-space radiometric normalization of TM/ETM+ images for land cover change detection. *International Journal of Remote Sensing*, 32(20), pp.7539–3556.
- Crawford, M.M., Tuia, D. & Yang, H.L., 2013. Active learning: Any value for classification of remotely sensed data? *Proceedings of the IEEE*, 101(3), pp.593–608.
- Cutting, J., 2003. Property Creep: A Case Study. In *WaPUG Autumn Conference*.
- Czaplewski, R., 2000. Accuracy assessments and areal estimates using two-phase

- stratified random sampling, cluster plots, and the multivariate composite estimator. In H. T. Mowrer & R. G. Congalton, eds. *Quantifying spatial uncertainty in natural resources: Theory and Applications for Remote Sensing and GIS*. pp. 79–100.
- Dai, X. & Khorram, S., 1998. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5), pp.1566–1577.
- Dams, J., Dujardin, J., Reggers, R., Bashir, I., Canters, F. & Batelaan, O., 2013. Mapping impervious surface change from remote sensing for hydrological modeling. *Journal of Hydrology*, 485, pp.84–95.
- Davis, C.H., 2011. Pixel-Based Invariant Feature Extraction and its Application to Radiometric Co-Registration for Multi-Temporal High-Resolution Satellite Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(2), pp.348–360.
- Dazhao, F., Song, J., Rong, L. & Yongsheng, Z., 2007. Automatic DSM Generation form Aerial Three Line Array ADS40 Digital Images. In *2007 8th International Conference on Electronic Measurement and Instruments*. Ieee, pp. 2-870-2–875.
- DEFRA, Birmingham City Council & Hydra Consulting, 2008. *Making Space for Water Urban Flood Risk and Integrated Drainage Pilots: Upper Rea Catchment including Longbridge, Northfield and Rubery Districts of Birmingham. Volume One - Pilot Report*,
- Demir, B., Bovolo, F. & Bruzzone, L., 2013. Updating land-cover maps by classification of image time series: A novel change-detection-driven transfer learning approach. *IEEE Transactions on Geoscience and Remote Sensing*, 51(1), pp.300–312.
- Deng, J.S., Wang, K., Deng, Y.H. & Qi, G.J., 2008. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, 29(16), pp.4823–4838.
- Department for Communities and Local Government, 2016. *House Building: March*

2016, England,

Department for Communities and Local Government, 2008a. *Impact Assessment – Permeable Surfaces*,

Department for Communities and Local Government, 2008b. *The Town and Country Planning (General Permitted Development) (Amendment) (No . 2) (England) Order 2008*, London: Department for Communities and Local Government.

Desclée, B., Bogaert, P. & Defourny, P., 2006. Forest change detection by statistical object-based method. *Remote Sensing of Environment*, 102(1–2), pp.1–11.

Dickinson, G.C. & Shaw, M.G., 1977. What is ' land use '? *The Royal Geographical Society*, 9(1), pp.38–42.

Diesing, M., Mitchell, P. & Stephens, D., 2016. Image-based seabed classification: what can we learn from terrestrial remote sensing? *ICES Journal of Marine Science*, 73(7).

DigitalGlobe, 2016. DigitalGlobe DSM. Available at:
<https://www.digitalglobe.com/products/advanced-elevation-series>.

Dingle Robertson, L. & King, D., 2011. Comparison of pixel- and object-based classification in land cover change mapping. *International Journal of Remote Sensing*, 32(6), pp.1505–1529.

Doxani, G., Karantzalos, K. & Tsakiri-Strati, M., 2011. Monitoring urban changes based on scale-space filtering and object-oriented classification. *International Journal of Applied Earth Observation and Geoinformation*, 15, pp.38–48.

Duckworth, C., 2005. *Assessment of urban creep rates for house types in Keighley and the capacity for future urban creep*. University of Manchester.

Duro, D.C., Franklin, S.E. & Dubé, M.G., 2013. Hybrid Object-based Change Detection and Hierarchical Image Segmentation for Thematic Map Updating. *Photogrammetric Engineering and Remote Sensing*, 79(3), pp.259–268.

Dy, J.G. & Brodley, C.E., 2004. Feature Selection for Unsupervised Learning . *Journal of Machine Learning Research*, 5, pp.845–889.

- Ehlers, M., Sofina, N., Filippovska, Y. & Kada, M., 2014. Automated Techniques for Change Detection Using Combined Edge Segment Texture Analysis, GIS, and 3D Information. In Q. Weng, ed. *Global Urban Monitoring and Assessment through Earth Observation*. CRC Press, Taylor & Francis Group, pp. 325–351.
- European Space Agency, 2015. *Copernicus Space Component Data Access system (CSCDA)*,
- Evans, T.D. & Eadon, A.R., 2007. *Urban Runoff Modelling: why not do it properly?*,
- Falco, N., Mura, M. & Bovolo, F., 2013. Change Detection in VHR Images Based on Morphological Attribute Profiles. *IEEE Geoscience and Remote Sensing Letters*, 10(3), pp.636–640.
- Fisher, P.F., 1997. The pixel: a snare and a delusion. *International Journal of Remote Sensing*, 18(3), pp.679–685.
- Foody, G.M., 2008. Harshness in image classification accuracy assessment. *International Journal of Remote Sensing*, 29(11), pp.3137–3158.
- Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Remote sensing of environment*, 80(1), pp.185–201.
- Frohn, R.C., 2006. The use of landscape pattern metrics in remote sensing image classification. *International Journal of Remote Sensing*, 27(10), pp.2025–2032.
- Gamanya, R., De Maeyer, P. & De Dapper, M., 2009. Object-oriented change detection for the city of Harare, Zimbabwe. *Expert Systems with Applications*, 36(1), pp.571–588.
- Gangkofner, U., Weichelsbaum, J., Kuntz, S., Brodksy, L., Larsson, K. & Pasquale, V.D.E., 2010. Update of the European High-resolution Layer of Built-up Areas and Soil Sealing 2006 with Image2009 Data. In *30th EARSeL Symposium: Remote Sensing for Science, Education and Culture*. EARSeL.
- Gao, F., Colstoun, E. de & Ma, R., 2012. Mapping impervious surface expansion using medium-resolution satellite image time series: a case study in the Yangtze River Delta, China. *International Journal of Remote Sensing*, 33(24), pp.7609–7628.

- Gehrke, S., Uebbing, R., Downey, M. & Morin, K., 2011. Creating and using Very High Density Point Clouds Derived from ADS Imagery. In *ASPRS 2011 Annual Conference*.
- Gerard, F., Petit, S., Smith, G.M. & Thomson, A., 2010. Land cover change in Europe between 1950 and 2000 determined employing aerial photography. *Progress in Physical*, 34(2), pp.183–205.
- Ghosh, S., Roy, M. & Ghosh, A., 2014. Semi-supervised change detection using modified self-organizing feature map neural network. *Applied Soft Computing*, 15, pp.1–20.
- Gill, E., 2008. *Making Space for Water Urban flood risk & integrated drainage(HA2). IUD pilot summary report*,
- Greater London Authority, 2005. *Crazy-paving: The environmental importance of London's front gardens*, London.
- Green, K., Kempka, D. & Lackey, L., 1994. Using remote sensing to detect and monitor land-cover and land-use change. *Photogrammetric Engineering & Remote Sensing*, 60(3), pp.331–337.
- Gueguen, L. & Hamid, R., 2016. Toward a Generalizable Image Representation for Large-Scale Change Detection: Application to Generic Damage Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6), pp.3378–3387.
- Guo, J., Liang, L. & Gong, P., 2010. Removing shadows from Google Earth images. *International Journal of Remote Sensing*, 31(6), pp.1379–1389.
- Hamada, Y., Stow, D.A., Roberts, D.A., Franklin, J. & Kyriakidis, P.C., 2013. Assessing and monitoring semi-arid shrublands using object-based image analysis and multiple endmember spectral mixture analysis. *Environmental monitoring and assessment*, 185(4), pp.3173–90.
- Hame, T., Heiler, I. & San Miguel-Ayanz, J., 1998. An unsupervised change detection and recognition system for forestry. *International Journal of Remote Sensing*, 19(6), pp.1079–1099.

- Hamedianfar, A. & Shafri, H.Z.M., 2013. Development of fuzzy rule-based parameters for urban object-oriented classification using very high resolution imagery. *Geocarto International*, (March 2015), pp.1–25.
- Han, D., 2011. *Flood Risk Assessment and Management*, Bentham Science Publishers.
- Hao, M., Shi, W., Deng, K. & Feng, Q., 2016. Superpixel-based active contour model for unsupervised change detection from satellite images. *International Journal of Remote Sensing*, 37(18), pp.4276–4295.
- Hao, M., Shi, W., Deng, K., Zhang, H. & He, P., 2016. An Object-Based Change Detection Approach Using Uncertainty Analysis for VHR Images. *Journal of Sensors*, 2016, p.17.
- Hayes, D. & Sader, S., 2001. Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series. *Photogrammetric engineering and remote sensing*, 67(9), pp.1067–1075.
- He, C., Wei, A., Shi, P., Zhang, Q. & Zhao, Y., 2011. Detecting land-use/land-cover change in rural–urban fringe areas using extended change-vector analysis. *International Journal of Applied Earth Observation and Geoinformation*, 13(4), pp.572–585.
- Hester, D.B., Nelson, S.A.C., Cakir, H.I., Khorram, S. & Cheshire, H., 2010. High-resolution land cover change detection based on fuzzy uncertainty analysis and change reasoning. *International Journal of Remote Sensing*, 31(2), pp.455–475.
- Hölbling, D., Friedl, B. & Eisank, C., 2015. An object-based approach for semi-automated landslide change detection and attribution of changes to landslide classes in northern Taiwan. *Earth Science Informatics*, pp.327–335.
- Homer, C. & Xian, G., 2011. Change of Impervious Surface Area Between 2001 and 2006 in the Conterminous United States. *Photogrammetric Engineering & Remote Sensing*, 77(8), pp.758–762.
- Hospedales, T.M., Gong, S. & Xiang, T., 2013. Finding rare classes: Active learning with generative and discriminative models. *IEEE Transactions on Knowledge and Data*

- Engineering*, 25(2), pp.374–386.
- Huang, G., Zhou, W. & Cadenasso, M.L., 2011. Is everyone hot in the city? Spatial pattern of land surface temperatures, land cover and neighborhood socioeconomic characteristics in Baltimore, MD. *Journal of environmental management*, 92(7), pp.1753–9.
- Huang, W. & Bu, M., 2015. Detecting shadows in high-resolution remote-sensing images of urban areas using spectral and spatial features. *International Journal of Remote Sensing*, 36(24), pp.6224–6244.
- Huiping, H., Bingfang, W. & Jinlong, F., 2003. Analysis to the relationship of classification accuracy, segmentation scale, image resolution. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*. pp. 3671–3673.
- Hussain, M., Chen, D., Cheng, A., Wei, H. & Stanley, D., 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80, pp.91–106.
- İlsever, M. & Ünsalan, C., 2012. *Two-Dimensional Change Detection Methods*, Springer.
- Im, J. & Jensen, J.R., 2005. A change detection model based on neighborhood correlation image analysis and decision tree classification. *Remote Sensing of Environment*, 99(3), pp.326–340.
- Im, J., Jensen, J.R. & Tullis, J.A., 2008. Object-based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing*, 29(2), pp.399–423.
- James, P., Tzoulas, K., Adams, M.D., Barber, a., Box, J., Breuste, J., Elmqvist, T., Frith, M., Gordon, C. & Greening, K.L., 2009. Towards an integrated understanding of green space in the European built environment. *Urban Forestry & Urban Greening*, 8(2), pp.65–75.
- Jensen, J.R. & Toll, D., 1982. Detecting residential land-use development at the urban

- fringe. *Photogrammetric Engineering and Remote Sensing*, 48(4), pp.629–643.
- Jian, P., Chen, K. & Zhang, C., 2016. A hypergraph-based context-sensitive representation technique for VHR remote-sensing image change detection. *International Journal of Remote Sensing*, 37(8), pp.1814–1825.
- Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J. & Xian, G., 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment*, 132, pp.159–175.
- Johnson, R.D. & Kasischke, E.S., 1998. Change vector analysis: A technique for the multispectral monitoring of land cover and condition. *International Journal of Remote Sensing*, 19(3), pp.411–426.
- Jones, C.B., Ware, J. & Miller, D., 1999. A probabilistic approach to environmental change detection with area-class map data. *Lecture Notes in Computer Science*, 1737, pp.122–136.
- Kasetkasem, T. & Varshney, P.K., 2011. An Optimum Land Cover Mapping Algorithm in the Presence of Shadows. *IEEE Journal of Selected Topics in Signal Processing*, 5(3), pp.592–605.
- Khorram, S., Biging, G.S., Chrisman, N., Colby, D.R., Congalton, R.G., Dobson, J., Ferguson, R., Goodchild, M.F., Jensen, J.R. & Mace, T., 1999. *Accuracy Assessment of Remote Sensing-Derived Change Detection*, ASPRS Monograph Series.
- Klaric, M.N., Claywell, B.C., Scott, G.J., Hudson, N.J., Sjahputera, O., Li, Y., Barratt, S.T., Keller, J.M. & Davis, C.H., 2013. GeoCDX: An Automated Change Detection and Exploitation System for High-Resolution Satellite Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4), pp.2067–2086.
- Kuncheva, L.I., 2014. *Combining Pattern Classifiers : Methods and Algorithms*, Wiley.
- Kuntz, S., Braun, M., Ernst, C., Gallaun, H., Santos, C., Hame, T., Jacob, P., Herrera-Cruz, V., Katagis, T., Keil, M., Lamb, A., Larsson, K., Rosengren, M., Markus, T., Smith, G.M., Weichelsbaum, J. & Kotarba, A., 2011. *Geoland 2 - Technical Note - Change Detection Status Report*,

- Li, X., Zhang, Y., Liu, X. & Chen, Y., 2012. Assimilating process context information of cellular automata into change detection for monitoring land use changes. *International Journal of Geographical Information Science*, 26(9), pp.1667–1687.
- Linke, J. & McDermid, G.J., 2012. Monitoring landscape change in multi-use west-central Alberta, Canada using the disturbance-inventory framework. *Remote Sensing of Environment*, 125, pp.112–124.
- Linke, J., McDermid, G.J., Laskin, D.N., McLane, A.J., Pape, A.D., Cranston, J., Hall-Beyer, M. & Franklin, S.E., 2009. A Disturbance-Inventory Framework for Flexible and Reliable Landscape Monitoring. *Photogrammetric Engineering & Remote Sensing*, 75(8), pp.981–995.
- Linke, J., McDermid, G.J., Pape, A.D., McLane, A.J., Laskin, D.N., Hall-Beyer, M. & Franklin, S.E., 2009. The influence of patch-delineation mismatches on multi-temporal landscape pattern analysis. *Landscape Ecology*, 24(2), pp.157–170.
- Listner, C. & Niemeyer, I., 2011a. Object-based Change Detection. *Photogrammetrie - Fernerkundung - Geoinformation*, 2011(4), pp.233–245.
- Listner, C. & Niemeyer, I., 2011b. Recent advances in object-based change detection. In *Geoscience and Remote Sensing Symposium (IGARSS), IEEE International*. pp. 110–113.
- Liu, C., Frazier, P. & Kumar, L., 2007. Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, 107(4), pp.606–616.
- Liu, W. & Yamazaki, F., 2012. Object-based shadow extraction and correction of high-resolution optical satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4), pp.1296–1302.
- Liu, Z., Gong, P., Shi, P., Chen, H., Zhu, L. & Sasagawa, T., 2010. Automated building change detection using UltraCamD images and existing CAD data. *International Journal of Remote Sensing*, 31(6), pp.1505–1517.
- London Cabinet Office, 2008. *The Pitt Review: Learning lessons from the 2007 floods*,

London.

London Wildlife Trust, Greenspace Information for Greater London & Greater London Authority, 2011. *London: Garden City?*, London.

Lu, D., Hetrick, S. & Moran, E., 2011. Impervious surface mapping with Quickbird imagery. *International Journal of Remote Sensing*, 32(9), pp.2519–2533.

Lu, D., Li, G. & Moran, E., 2014. Current situation and needs of change detection techniques. *International Journal of Image and Data Fusion*, 5(March 2015), pp.13–38.

Lu, D., Mausel, P., Brondizio, E.S. & Moran, E., 2004. Change detection techniques. *International Journal of Remote Sensing*, 25(12), pp.2365–2401.

Lu, D., Moran, E. & Hetrick, S., 2010. Detection of impervious surface change with multitemporal Landsat images in an urban–rural frontier. *ISPRS Journal of Photogrammetry and Remote Sensing*.

Lu, D. & Weng, Q., 2006. Use of impervious surface in urban land-use classification. *Remote Sensing of Environment*, 102(1–2), pp.146–160.

Ma, L., Li, M., Blaschke, T., Ma, X., Tiede, D., Cheng, L., Chen, Z. & Chen, D., 2016. Object-Based Change Detection in Urban Areas : The Effects of Segmentation Strategy, Scale, and Feature Space on Unsupervised Methods. *Remote Sensing*, (September), pp.1–18.

Marchesi, S., Bovolo, F. & Bruzzone, L., 2010. A context-sensitive technique robust to registration noise for change detection in VHR multispectral images. *IEEE transactions on image processing*, 19(7), pp.1877–89.

Martha, T.R., Kamala, P., Jose, J., Vinod Kumar, K. & Jai Sankar, G., 2016. Identification of new Landslides from High Resolution Satellite Data Covering a Large Area Using Object-Based Change Detection Methods. *Journal of the Indian Society of Remote Sensing*, 44(August), pp.1–10.

McDermid, G.J., Linke, J., Pape, A.D., Laskin, D.N., McLane, A.J. & Franklin, S.E., 2008. Object-based approaches to change analysis and thematic map update:

- challenges and limitations. *Canadian Journal of Remote Sensing*, 34(5), pp.462–466.
- Mishra, N.B. & Crews, K. a, 2014. Mapping vegetation morphology types in a dry savanna ecosystem : integrating hierarchical object-based image analysis with Random Forest. *International Journal of Remote Sensing*, 35(3), pp.1175–1198.
- Moughal, T.A. & Yu, F., 2014. An Automatic Unsupervised Method Based on Context-Sensitive Spectral Angle Mapper for Change Detection of Remote Sensing Images. In X. Luo, J. X. Yu, & Z. Li, eds. *Advanced Data Mining and Applications*. Springer International Publishing, pp. 151–162.
- Murphy, K.P., 2012. *Adaptive Computation and Machine Learning : Machine Learning : A Probabilistic Perspective*, The MIT Press.
- Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S. & Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, 115(5), pp.1145–1161.
- Nakajima, T., G, T. & Y, Y., 2002. Simulated recovery of information in shadow areas on IKONOS image by combing ALS data. In *Proceeding of Asian Conference on Remote Sensing*.
- National Oceanic & Atmospheric Administration, 2016. NOAA solar calculator. Available at: <http://www.esrl.noaa.gov/gmd/grad/solcalc/index.html>.
- Natural England, 2010a. *NCA 78 : Central North Norfolk Key Facts & Data Map of Central North Norfolk*,
- Natural England, 2010b. *NCA 80 : The Broads Key Facts & Data*,
- Natural England, 2010c. *NCA 84 : Mid Norfolk Key Facts & Data*,
- Newcastle City Council, 2008. *Urban flood risk and integrated drainage - Ouseburn and North Gosforth Pilot Project*,
- Norfolk County Council, 2011. *Preliminary Flood Risk Assessment Report*,
- Ochoa-Gaona, S. & Gonzalez-Espinosa, M., 2000. Land use and deforestation in the

- highlands of Chiapas, Mexico. *Applied Geography*, 20(1), pp.17–42.
- Office for National Statistics, 2013. 2011 Census for England and Wales. Available at: <http://www.neighbourhood.statistics.gov.uk>.
- Office for National Statistics, 2012. *Economic Review, April 2012*,
- Office for National Statistics, 2016. *Economic Review: April 2016*,
- Ofwat, 2011. *Future Impacts on Sewer Systems in England and Wales - Summary of a Hydraulic Modelling Exercise Reviewing the Impact of Climate Change, Population and Growth in Impermeable Areas up to Around 2040*,
- Ogilvie, J., Graham, M., Cox, S., Longworth, A., Skilton, D. & Ogden, M., 2011. *Norwich Surface Water Management Plan Stage 2: Final Report*,
- Olofsson, P., Foody, G.M., Herold, M., Stehman, S. V., Woodcock, C.E. & Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148, pp.42–57.
- Olver, P.J. & Shakiban, C., 2006. *Applied Linear Algebra*, Pearson Prentice Hall.
- Oshiro, T., Perez, P. & Baranauskas, J., 2012. How many trees in a random forest? In *Machine Learning and Data* pp. 154–168.
- Pacifici, F. & Frate, F. Del, 2010. Automatic change detection in very high resolution images with pulse-coupled neural networks. *IEEE Geoscience and Remote Sensing Letters*, 7(1), pp.58–62.
- Padilla, M., Stehman, S. V., Ramo, R., Corti, D., Hantson, S., Oliva, P., Alonso-Canas, I., Bradley, A. V., Tansey, K., Mota, B., Pereira, J.M. & Chuvieco, E., 2015. Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation. *Remote Sensing of Environment*, 160, pp.114–121.
- Pan, J., Shi, P. & Zheng, F., 2011. Analysis on Spatial Pattern of Urban Heat Island and Impervious Surface Using Linear Spectral Mixture Analysis. *Advance Materials Research*, 216, pp.600–604.

- Park, C., 2013. A Feature Selection Method Using Hierarchical Clustering. In *Mining Intelligence and Knowledge Exploration*. pp. 1–6.
- Parmentier, B. & Eastman, J.R., 2014. Land transitions from multivariate time series: using seasonal trend analysis and segmentation to detect land-cover changes. *International Journal of Remote Sensing*, 35(2), pp.671–692.
- Patra, S., Ghosh, S. & Ghosh, A., 2011. Histogram thresholding for unsupervised change detection of remote sensing images. *International Journal of Remote Sensing*, 32(21), pp.6071–6089.
- Peiman, R., 2011. Pre-classification and post-classification change-detection techniques to monitor land-cover and land-use change using multi-temporal Landsat imagery: a case study. *International Journal of Remote Sensing*, 32(15), pp.4365–4381.
- Pelling, C. & Kelly, S., 2010. *Greater Norwich Development Partnership Stage 2b Water Cycle Study Non Technical Planning Report*,
- Perry, T. & Nawaz, R., 2008. An investigation into the extent and impacts of hard surfacing of domestic gardens in an area of Leeds, United Kingdom. *Landscape and Urban Planning*, 86(1), pp.1–13.
- Poli, D. & Toutin, T., 2012. Review of developments in geometric modelling for high resolution satellite pushbroom sensors. *Photogrammetric Record*, 27(137), pp.58–73.
- Qin, R., Huang, X., Gruen, A. & Schmitt, G., 2015. Object-Based 3-D Building Change Detection on Multitemporal Stereo Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(5), pp.2125–2137.
- Radke, R.J., Andra, S., Al-Kofahi, O. & Roysam, B., 2005. Image change detection algorithms: a systematic survey. *IEEE transactions on image processing*, 14(3), pp.294–307.
- Rahman, A., Aggarwal, S.P., Netzband, M. & Fazal, S., 2011. Monitoring Urban Sprawl Using Remote Sensing and GIS Techniques of a Fast Growing Urban Centre, India.

- IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(1), pp.56–64.
- Reese, H. & Olsson, H., 2011. C-correction of optical satellite data over alpine vegetation areas: A comparison of sampling strategies for determining the empirical c-parameter. *Remote Sensing of Environment*, 115(6), pp.1387–1400.
- Remote Aerial Surveys, 2016. Remote Aerial Surveys DSM. Available at: <http://www.remoteaerialsurveys.co.uk/photogrammetry>.
- Reulke, R., Becker, S., Haala, N. & Tempelmann, U., 2006. Determination and improvement of spatial resolution of the CCD-line-scanner system ADS40. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(2), pp.81–90.
- Richard Allitt Associates Ltd., 2008. *Wessex Water final business plan 2010 - 2015: Appendix C6.ii Urban Creep Study*,
- Richter, R., 1998. Correction of Satellite Imagery Over Mountainous Terrain. *Applied Optics*, 37(18), p.4004.
- Ridd, M.K., 1995. Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities†. *International Journal of Remote Sensing*, 16(12), pp.2165–2185.
- Rokach, L., 2009. *Series in Machine Perception and Artificial Intelligence : Pattern Classification Using Ensemble Methods*, WSPC.
- Ruiz, J.J., Ariza, F.J., Ureña, M. a. & Blázquez, E.B., 2011. Digital map conflation: a review of the process and a proposal for classification. *International Journal of Geographical Information Science*, 25(9), pp.1439–1466.
- Salehi, B., Zhang, Y., Zhong, M. & Dey, V., 2012. Object-based classification of urban areas using VHR imagery and height points ancillary data. *Remote Sensing*, 4(8), pp.2256–2276.
- Scalenghe, R. & Marsan, F.A., 2009. The anthropogenic sealing of soils in urban areas. *Landscape and Urban Planning*, 90(1–2), pp.1–10.
- Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S. & Martonchik, J. V.,

2006. Reflectance quantities in optical remote sensing-definitions and case studies. *Remote Sensing of Environment*, 103(1), pp.27–42.
- Schipperijn, J., Stigsdotter, U.K., Randrup, T.B. & Troelsen, J., 2010. Influences on the use of urban green space – A case study in Odense, Denmark. *Urban Forestry & Urban Greening*, 9(1), pp.25–32.
- Schneider, A., 2012. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. *Remote Sensing of Environment*, 124, pp.689–704.
- Schroeder, T.A., Cohen, W.B., Song, C., Canty, M.J. & Yang, Z., 2006. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. *Remote Sensing of Environment*, 103(1), pp.16–26.
- Scikit Learn, 2016a. RandomForestClassifier. Available at: <http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>.
- Scikit Learn, 2016b. Scikit-Learn Webpage. Available at: <http://scikit-learn.org/stable/>.
- Serra, P., Pons, X. & Sauri, D., 2003. Post-classification change detection with data from different sensors: some accuracy considerations. *International Journal of Remote Sensing*, 24(16), pp.3311–3340.
- Seto, K.C., Woodcock, C.E., Song, C., Huang, X., Lu, J. & Kaufmann, R.K., 2002. Monitoring land-use change in the Pearl River Delta using Landsat TM. *International Journal of Remote Sensing*, 23(10), pp.1985–2004.
- Shah-Hosseini, R., Homayouni, S. & Safari, A., 2015. A Hybrid Kernel-Based Change Detection Method for Remotely Sensed Data in a Similarity Space. *Remote Sensing*, 7(10), pp.12829–12858.
- Shahtahmassebi, A., Yang, N., Wang, K., Moore, N. & Shen, Z., 2013. Review of shadow detection and de-shadowing methods in remote sensing. *Chinese Geographical Science*, 23(4), pp.403–420.

- Shalaby, A. & Tateishi, R., 2007. Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt. *Applied Geography*, 27(1), pp.28–41.
- Shao, Y., Taff, G. & Walsh, S., 2011. Shadow detection and building-height estimation using IKONOS data. *International Journal of Remote Sensing*, 32(22), pp.6929–6944.
- Sheng, Y., 2007. Minimising algorithm-induced artefacts in true ortho-image generation: a direct method implemented in the vector domain. *The Photogrammetric Record*, 22(June), pp.151–163.
- Shi, A., Huynh, D.Q., Huang, F.C., Shen, S.H., Lu, W.P. & Ma, Z.L., 2014. Unsupervised change detection based on robust chi-squared transform for bitemporal remotely sensed images. *International Journal of Remote Sensing*, 35(21), pp.7555–7566.
- Singh, A., 1989. Review Article Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, 10(6), pp.989–1003.
- Sinha, P. & Kumar, L., 2013. Independent two-step thresholding of binary images in inter-annual land cover change/no-change identification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 81, pp.31–43.
- Slonecker, E.T., Jennings, D.B. & Garofalo, D., 2001. Remote sensing of impervious surfaces: A review. *Remote Sensing Reviews*, 20(3), pp.227–255.
- Sofina, N., Ehlers, M. & Michel, U., 2012. Integrated data processing of remotely sensed and vector data for building change detection. In *Proceedings of SPIE The International Society for Optical Engineering: Earth Resources and Environmental Remote Sensing/GIS Applications III Vol 8538*. p. 85380F–1 to 85380F–8.
- Stehman, S. V., 2009. Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30(20), pp.5243–5272.
- Stehman, S. V., 2001. Statistical rigor and practical utility in thematic map accuracy assessment. *Photogrammetric Engineering and Remote Sensing*, 67(6), pp.727–734.

- Stehman, S. V., Sohl, T.L. & Loveland, T.R., 2005. An evaluation of sampling strategies to improve precision of estimates of gross change in land use and land cover. *International Journal of Remote Sensing*, 26(22), pp.4941–4957.
- Stone, B., 2004. Paving over paradise: how land use regulations promote residential imperviousness. *Landscape and Urban Planning*, 69(1), pp.101–113.
- Stow, D., Hamada, Y., Coulter, L. & Anguelova, Z., 2008. Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery. *Remote Sensing of Environment*, 112(3), pp.1051–1061.
- Tabib Mahmoudi, F., Samadzadegan, F. & Reinartz, P., 2013. Object oriented image analysis based on multi-agent recognition system. *Computers & Geosciences*, 54, pp.219–230.
- Tan, B., Masek, J.G., Wolfe, R., Gao, F., Huang, C., Vermote, E.F., Sexton, J.O. & Ederer, G., 2013. Improved forest change detection with terrain illumination corrected Landsat images. *Remote Sensing of Environment*, 136, pp.469–483.
- Tan, K., Jin, X., Plaza, A., Wang, X. & Xiao, L., 2016. Automatic Change Detection in High-Resolution Remote Sensing Images by Using a Multiple Classifier System and Spectral – Spatial Features. , 9(8), pp.1–13.
- Tang, Y., Huang, X. & Zhang, L., 2013. Fault-Tolerant Building Change Detection From Urban High-Resolution Remote Sensing Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(5), pp.1060–1064.
- Tang, Z., Tang, H., He, S. & Mao, T., 2015. OBJECT-BASED CHANGE DETECTION MODEL USING CORRELATION ANALYSIS AND CLASSIFICATION FOR VHR IMAGE. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, pp. 4840–4843.
- Teo, T.-A. & Shih, T.-Y., 2013. Lidar-based change detection and change-type determination in urban areas. *International Journal of Remote Sensing*, 34(3), pp.968–981.
- Tewkesbury, A.P., 2011. Mapping the extent of urban creep in Exeter using OBIA. In

- Proceedings of RSPSoc Annual Conference*. p. 163.
- Tewkesbury, A.P. & Allitt, M., 2010. Urban creep mapping from Remote Sensing data. In *Proceedings of RSPSoc Annual Conference and Irish Earth Observation Symposium*. p. 34.
- Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A. & Fisher, P.F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, 160, pp.1–14.
- The Government Office for Science, 2010. *Foresight Land Use Futures Project. Final Project Report.*, London.
- Tian, J., Cui, S. & Reinartz, P., 2014. Building Change Detection Based on Satellite Stereo Imagery and Digital Surface Models. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1), pp.406–417.
- Tian, J., Reinartz, P., D'Angelo, P. & Ehlers, M., 2013. Region-based automatic building and forest change detection on Cartosat-1 stereo imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79, pp.226–239.
- Torres-Vera, M. a., Prol-Ledesma, R.M. & Garcia-Lopez, D., 2009. Three decades of land use variations in Mexico City. *International Journal of Remote Sensing*, 30(1), pp.117–138.
- Toure, S., Stow, D., Shih, H., Coulter, L., Weeks, J., Engstrom, R. & Sandborn, A., 2016. An object-based temporal inversion approach to urban land use change analysis. *Remote Sensing Letters*, 7(5), pp.503–512.
- Trimble, 2015. *eCognition v9.1 Reference Book*,
- Trioulet, L., 2012. Impact of Urban Creep on the hydrology of a catchment. In *1st Civil and Environmental Engineering Student Conference*. Imperial College London.
- UKWIR, 2010. *Impact of Urban Creep on Sewerage Systems*, London.
- Vakalopoulou, M., Karantzalos, K., Member, S., Komodakis, N. & Paragios, N., 2016. Graph-Based Registration , Change Detection , and Classification in Very High Resolution Multitemporal Remote Sensing Data. *IEEE Journal of Selected Topics in*

- Applied Earth Observations and Remote Sensing*, 9(7), pp.2940–2951.
- Verbeeck, K. & Van Orshoven, J., 2012. External geo-information in the segmentation of VHR imagery improves the detection of imperviousness in urban neighborhoods. *International Journal of Applied Earth Observation and Geoinformation*, 18(1), pp.428–435.
- Verbeeck, K., Van Orshoven, J. & Hermy, M., 2011. Measuring extent, location and change of imperviousness in urban domestic gardens in collective housing projects. *Landscape and Urban Planning*, 100(1–2), pp.57–66.
- Volpi, M., Tuia, D., Bovolo, F., Kanevski, M. & Bruzzone, L., 2013. Supervised change detection in VHR images using contextual information and support vector machines. *International Journal of Applied Earth Observation and Geoinformation*, 20, pp.77–85.
- Van de Voorde, T., De Genst, W., Canters, F., Stephenne, N., Wolff, E. & Binard, M., 2004. Extraction of land use/land cover related information from very high resolution data in urban and suburban areas. In *Remote Sensing in Transition. Proceedings of the 23rd Symposium of the European Association of Remote Sensing Laboratories*, edited by R. Goossens (Millpress, Rotterdam). pp. 237–244.
- Walter, V., 2004. Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3–4), pp.225–238.
- Wang, B., Choi, S., Byun, Y., Lee, S. & Choi, J., 2015. Object-Based Change Detection of Very High Resolution Satellite Imagery Using the Cross-Sharpening of Multitemporal Data. *Ieee Geoscience and Remote Sensing Letters*, 12(5), pp.1151–1155.
- Wang, G., 2014a. Study of change detection based on Poisson's mixture model and component-maximum threshold algorithm. *International Journal of Remote Sensing*, 35(1), pp.5–15.
- Wang, G., 2014b. Study of change detection based on Poisson's mixture model and component-maximum threshold algorithm. *International Journal of Remote*

- Sensing*, 35(1), pp.5–15.
- Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), pp.236–244.
- Warner, T., 2011. Kernel-Based Texture in Remote Sensing Image Classification. *Geography Compass*, 5(10), pp.781–798.
- Warner, T., Almutairi, A. & Lee, J.Y., 2009. Remote Sensing of Land Cover Change. In T. A. Warner, D. M. Nellis, & G. M. Foody, eds. *The SAGE Handbook of Remote Sensing*. SAGE Publications, pp. 459–472.
- Weber, K.T. & Chen, F., 2010. Detection Thresholds for Rare, Spectrally Unique Targets within Semiarid Rangelands. *Photogrammetric Engineering and Remote Sensing*, 76(11), pp.1253–1259.
- Wen, D., Huang, X., Zhang, L. & Benediktsson, J.A., 2016. A Novel Automatic Change Detection Method for Urban High-Resolution Remotely Sensed Imagery Based on Multiindex Scene Representation. *IEEE Transactions on Geoscience and Remote Sensing*, 54(1), pp.609–625.
- Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment*, 117, pp.34–49.
- Weng, Q., Rajasekar, U. & Hu, X., 2011. Modeling Urban Heat Islands and Their Relationship With Impervious Surface and Vegetation Abundance by Using ASTER Images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(10), pp.4080–4089.
- Wessex Water, 2009. *Wessex Water final business plan 2010 - 2015: C6. Sewer Flooding*,
- Whitehand, J.W.R. & Carr, C., 1999. The Changing Fabrics of Ordinary Residential Areas. *Urban Studies*, 36(10), pp.1661–1677.
- Willis, K.G. & Whitby, M.C., 1985. The Value of Green Belt Land. *Journal of Rural Studies*, 1(2), pp.147–162.

- Wilson, C., 2010. *Norwich City Council Strategic Flood Risk Assessment Level 2*,
- Van de Wouw, D.W.J.M., Dubbelman, G. & DeWith, P.H.N., 2016. Hierarchical 2.5D scene alignment for change detection with large viewpoint differences. *IEEE Robotics and Automation Letters*, 1(1), p.1.
- Wright, G.B., Arthur, S., Bastien, N., Bowles, G. & Unwin, D., 2010. *Extent and cost of designing and constructing small areas of hardstanding around new and existing , domestic and non-domestic buildings*,
- Wright, G.B., Arthur, S., Bowles, G., Bastien, N. & Unwin, D., 2011. Urban creep in Scotland: stakeholder perceptions, quantification and cost implications of permeable solutions. *Water and Environment Journal*.
- Wu, C., Zhang, L. & Zhang, L., 2016. A scene change detection framework for multi-temporal very high resolution remote sensing images. *Signal Processing*, 124, pp.184–197.
- Wu, S., Silvánhyphen, J. & Wang, L., 2007. Per-field urban land use classification based on tax parcel boundaries. *International Journal of Remote Sensing*, 28(12), pp.2777–2801.
- Wulder, M. a., Butson, C.R. & White, J.C., 2008. Cross-sensor change detection over a forested landscape: Options to enable continuity of medium spatial resolution measures. *Remote Sensing of Environment*, 112(3), pp.796–809.
- Xian, G. & Homer, C., 2010. Updating the 2001 National Land Cover Database Impervious Surface Products to 2006 using Landsat Imagery Change Detection Methods. *Remote Sensing of Environment*, 114(8), pp.1676–1686.
- Yu, W., Zhou, W., Qian, Y. & Yan, J., 2016. A new approach for land cover classification and change analysis: Integrating backdating and an object-based method. *Remote Sensing of Environment*, 177, pp.37–47.
- Yuan, D., Elvidge, C.D. & Lunetta, R.S., 1998. Survey of Multispectral Methods for Land Cover Change Analysis. In R. S. Lunetta & C. D. Elvidge, eds. *Remote Sensing Change Detection*. ANN ARBOR PRESS, pp. 21–35.

- Zhan, Q., Shi, W. & Xiao, Y., 2005. Quantitative analysis of shadow effects in high-resolution images of urban areas. In *3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas (URBAN) and 5th International Symposium Remote Sensing of Urban Areas (URS)*. pp. 1–6.
- Zhou, W., Huang, G., Troy, A. & Cadenasso, M.L., 2009. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. *Remote Sensing of Environment*, 113(8), pp.1769–1777.
- Zhuang, H., Deng, K., Fan, H. & Yu, M., 2016. Strategies Combining Spectral Angle Mapper and Change Vector Analysis to Unsupervised Change Detection in Multispectral Images. *IEEE Geoscience and Remote Sensing Letters*, 13(5), pp.681–685.