Land use and land cover mapping in the Swan and Canning River Catchments: Technical Report

Ricky van Dongen

April 2020

GOVERNMENT OF
WESTERN AUSTRALIA

Department of **Biodiversity,
Conservation and Attractions**

Department of Biodiversity, Conservation and Attractions
Locked Bag 104
Bentley Delivery Centre WA 6983
Phone: (08) 9219 9000
Fax: (08) 9334 0498

www.dbca.wa.gov.au

This report/document/publication was prepared by Ricky van Dongen

Questions regarding the use of this material should be directed to:

Ricky van Dongen

Research Officer (Remote Sensing)
Remote Sensing and Spatial Analysis
Department of Parks and Wildlife
Locked Bag 104
Bentley Delivery Centre WA 6983
Phone: 9219 9571
Email: Ricky.vanDongen@dbca.wa.gov.au

The recommended reference for this publication is:
Department Biodiversity, Conservation and Attractions, 2019, *Land use and land cover mapping in the Swan and Canning River Catchments: 2019*, Department of Biodiversity, Conservation and Attractions, Perth.

This document is available in alternative formats on request.

Please note: urls in this document which conclude a sentence are followed by a full point. If copying the url please do not include the full point.

Contents

Appendices

Figures

Tables

Summary

Land use and land cover datasets were created for the Swan and Canning river catchments. These datasets will be inputs for nutrient modelling being carried out by the University of Western Australia.

To create the land use and land cover datasets digital aerial photography (captured in 2016) was used along with supplementary datasets including agricultural land use mapping and cadastral layers. The datasets achieved a high degree of accuracy (land use = 91.81% and land cover = 80.7%).

1 Introduction

A land use dataset for the Swan and Canning River catchments was produced by the then Department of Water in 2005. This dataset is an input for nutrient models. Updated nutrient modelling is currently being carried out by UWA and as such updated land use and land cover datasets are required. The extent of the catchments is shown in Figure 1.

Assessing land use in the catchments is a complex task with over 400 000 land parcels to assess. In order to cover this volume, parcels were prioritised according to whether the parcel Id or shape had changed (indicating a likely change in land use). Agricultural land use mapping from the Department of Primary Industries and Regional Development (DPIRD) was utilised and digital aerial photography was used to predict land use. The digital aerial photography was also used to map land cover.

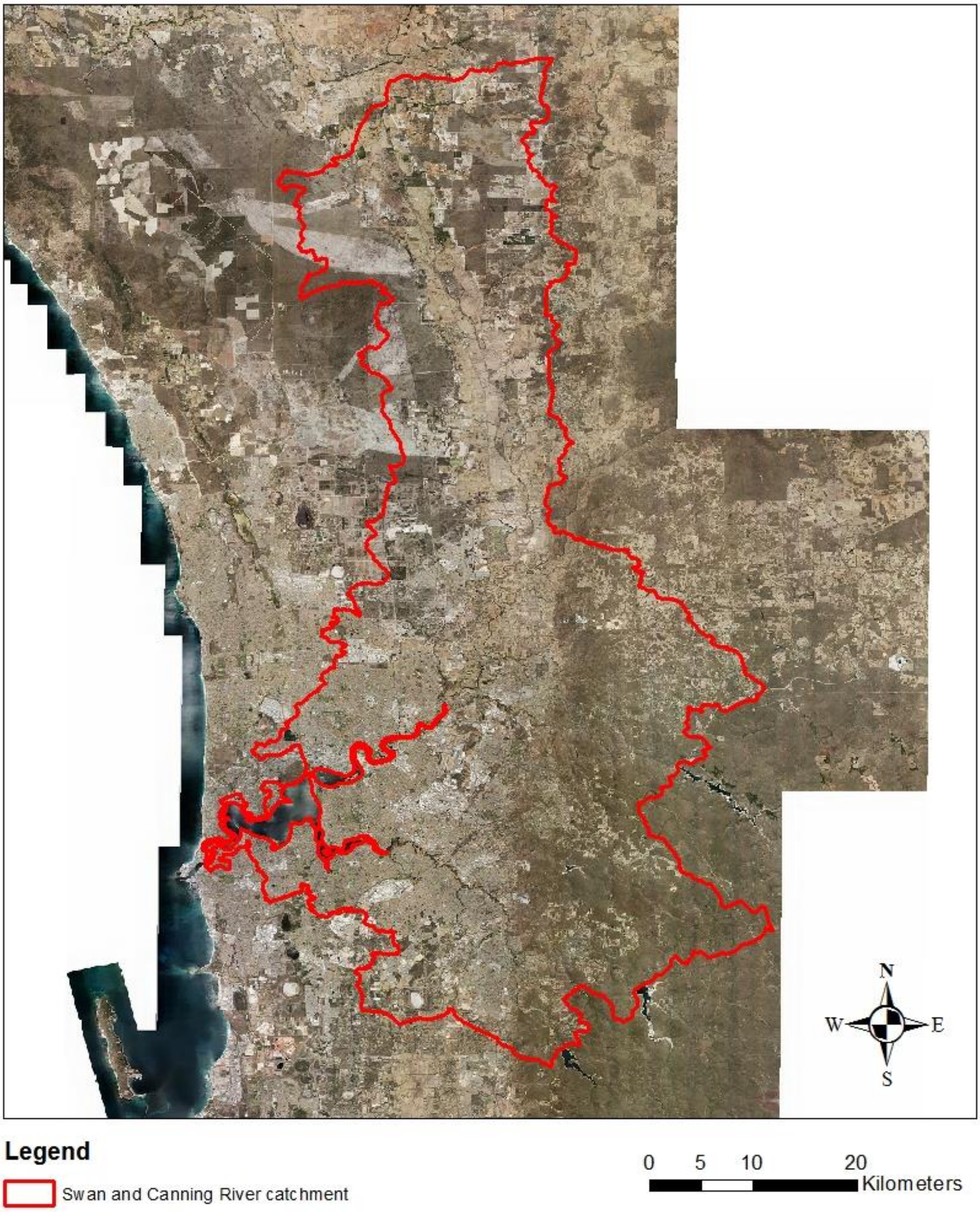

Figure 1: Swan and Canning River catchment boundary.

2 Methodology

2.1 Land use mapping

To handle the volume of land parcels the 2016 cadastre was subset by Local Government Area (LGA). All processing, unless stated otherwise, was carried out in the R statistical environment (R Core Team 2017).

2.1.1 DPIRD mapping

The Department of Primary Industries and Regional Development produced spatial datasets of agricultural land used and remnant vegetation across the south west. See Appendix 1 for DPIRD metadata statement. The remnant vegetation dataset was created through aerial photography interpretations.

Processing steps:

1. For each LGA all parcels (area with a cadastral boundary) which overlapped with a DPIRD land use polygon were selected.
2. The DPIRD mapping was added to the land parcel.
3. Any remnant vegetation within the parcel was also added.
4. The updated parcel attribution was then added to the final 2016 land use dataset.

2.1.2 Roads

The 2016 cadastre has many roads attributed.

Processing steps:

1. Parcels in the 2016 cadastre attributed as roads were selected by LGA.
2. These “road” parcels were added to the final 2016 land use dataset.

2.1.3 Unchanged parcels

Land parcels which did not overlap with the DPIRD mapping were assessed for change. This was done by looking for changes in the physical dimensions of the parcel. Ideally the 2005 cadastre would have been used for this purpose, but it could not be located. A copy of the 2007 cadastre, archived by DBCA, was acquired. The differing time between 2005 and 2007 means that parcels changed (subdivided) between this time were not detected. Many of these errors were corrected in the first round of QA.

Processing steps:

1. Parcels from the 2007 cadastre which the PIN or Parcel ID were different in the 2016 cadastre were selected (parcels which had changed).

2. Of these “changed” parcels, the location of the centroid, area and perimeter (dimensions) of each of these parcels was assessed against all intersecting parcels in the 2007 cadastre.
3. Those with differing dimensions were selected as “changed” and the remainder “unchanged”.
4. Unchanged 2016 parcels were then attributed with the same land use as 2005 and added to the final 2016 land use dataset.

2.1.4 Digital aerial photography predictions

The land use of parcels which have changed between 2007 and 2016 and do not overlap with the DPIRD mapping were predicted using digital aerial photography and sentinel imagery.

Digital aerial photography (captured in 2016) across the Swan coastal plain (including the Swan and Canning river catchments) is available in the form of Urban Monitor (UM) data. Urban Monitor (Caccetta et al. 2015) is a program run by CSIRO which produces calibrated digital imagery across four spectral bands, along with a normalised surface model (NSM). The data is capture across the summer months. The NSM is a raster dataset with each pixel representing the height of objects above ground level.

Sentinel 2 satellite imagery, captured 29/9/2016 was also used. The red and near-infrared bands were used to create a normalised difference vegetation index (NDVI). This dataset was used to complement the UM data and provide a measure of spring vegetation cover.

Processing steps:

1. Produce urban monitor and Sentinel 2 imagery tiles for each parcel.
2. Classify and calculate compositional and geometry variables for each parcel using R and eCognition. Variables are listed below.
    - Catchment name
    - area of small buildings
    - area of medium sized buildings
    - area of large buildings
    - area of single trees
    - area of tree patches
    - area of bare ground
    - area of grass
    - number of small buildings
    - number of medium sized buildings
    - number of large buildings
    - building maximum height
    - number of single trees
    - number of tree patches
    - percentage of parcel area covered by buildings
    - percentage of parcel area covered by trees
    - percentage of parcel area covered by bare ground
    - percentage of parcel area covered by grass
    - total area of parcel
    - percentage green in spring (satellite)
    - length/width of parcel
    - rectangular fit of parcel
    - shape index of parcel
    - radius of smallest enclosed ellipse of parcel
    - compactness of parcel

3. Land use classes which were likely to have a similar composition were consolidated as follows:

Table 1: Original and merged land use classes.

Original (n = 27)	**Merged classes (n = 18)**
Animal keeping - non-farming	Animal keeping - non-farming
Office - without parkland	Commercial
Storage / distribution	Commercial
Manufacturing / processing	Commercial
Commercial / service - centre	Commercial
Commercial / service - residential	Commercial
Community facility - education	Community facility - education
Community facility - non-education	Community facility - non-education
Drainage	Drainage
Farm	Farm
Horticulture	Horticulture
Lifestyle block / hobby farm	Lifestyle block / hobby farm
Recreation - turf	Recreation - turf/grass
Recreation - grass	Recreation - turf/grass
Residential - multiple dwelling	Residential - multiple dwelling
Residential - aged persons	Residential - multiple dwelling
Residential - temporary accommodation	Residential - multiple dwelling
Residential - single / duplex dwelling	Residential - single / duplex dwelling
Rural residential / bush block	Rural residential / bush block
Transport / access - non-airport	Transport / access - non-airport
Recreation / conservation - trees / shrubs	Trees / shrubs
Unused - uncleared - trees / shrubs	Trees / shrubs
Unused - cleared - bare soil	Unused - cleared
Unused - cleared - grass	Unused - cleared
Utility	Utility
Viticulture	Viticulture
Water body	Water body

4. Training data was generated by:
    a. Attributing parcels in the 2007 cadastre with land use from the 2005 dataset.
    b. Randomly selecting 400 parcels from within each land use class. Where less than 400 existed, all were selected. This resulted in 8894 parcels.
    c. Variables for these parcels were calculated from the 2007 UM data. This provided a compositional "signature" for each land use.
5. The training data was supplemented with 2060 manually selected parcels from the 2016 cadastre with parameters from the 2016 UM data. This was required to add property types not present in the 2007 sample – primarily parcels in the small block and house estates now seen.
6. This training data was then split into a training and test set and run in the ranger random forest model (Wright and Ziegler 2017). The overall accuracy against test data was 92.4%.
7. The output was applied to the parcels then added to the final 2016 land use dataset.

2.2 Land cover classification

The land cover classification was carried out in eCognition and R using digital aerial photography (Urban Monitor) and supplementary polygon datasets. The supplementary datasets included roads datasets from the 2016 cadastre and waterbodies from the “Hydro Polygons” dataset on SLIP (https://catalogue.data.wa.gov.au/dataset/medium-scale-topo-water-polygon-lgate-016). Permeable and non-permeable ground was also identified using the land use classifications.

Land uses where ground was classified as non-permeable

- Transport / access - non-airport
- Transport / access - airport
- Commercial
- Residential - multiple dwelling
- Manufacturing / processing
- Storage / distribution
- Commercial / service - centre
- Community facility - non-education
- Office - without parkland
- Residential - aged persons
- Residential - temporary accommodation
- Commercial / service - residential
- Office - with parkland
- Yacht facilities

Land cover classes with technical descriptions are shown in Table 2. The NSM is a raster where pixel values represent the height of objects above ground level. However small amounts of noise in the datasets and general difficulties in determining if a feature is ground or above ground mean that a threshold of 400 mm was set to separate ground from above ground.

Table 2: Land cover classes with technical definitions

Class	**Technical definition**
Trees/shrubs	Green vegetation above ground level NSM > 400 mm and NDVI > 0.3 and brightness (sum of all spectral bands) < 3000 and NIR – Red > 0
Grass	Green vegetation at ground level NSM < 400 mm and NDVI > 0.3 and (Sentinel 2) NDVI > 0.35
Buildings	Above ground level (> 400 mm) that are not trees/shrubs
permeable ground	Ground level (< 400 mm) that is not grass and not in a “non-permeable” land use
Non-permeable ground	Ground level (< 400 mm) that is not grass and is in a “non-permeable” land use
Roads	Any classes, except Trees/shrubs, that overlap with the roads dataset
Waterbodies	Buildings and unclassified areas that overlap with the waterbodies dataset

To assess the accuracy of the land cover classes 50 points were systematically placed in each class using aerial photography as a reference. Theses actual classes were then compared with the modelled classes.

3 Results

3.1 Land use classes from classification

An example of the land use dataset is shown in Figure 2.

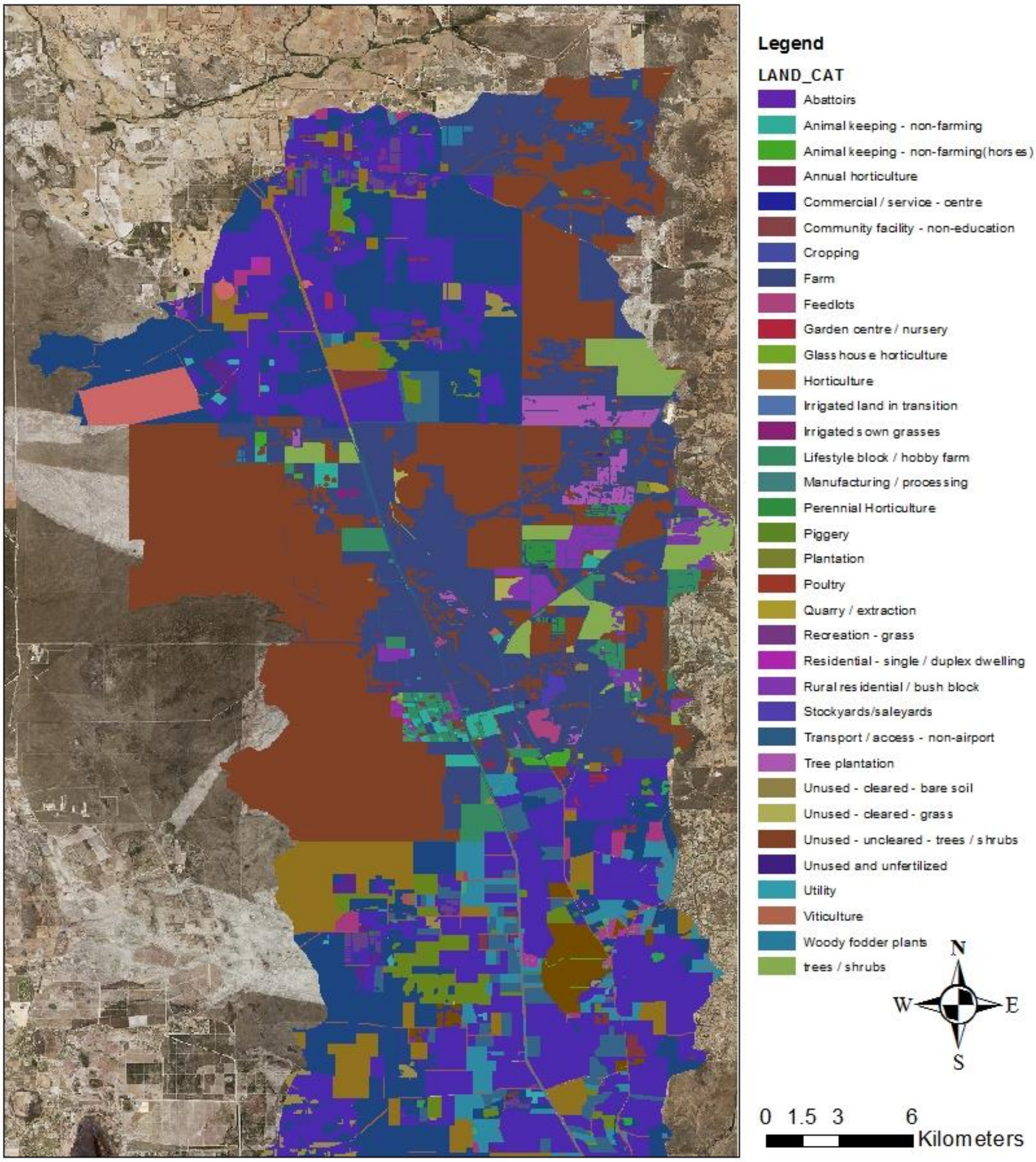

Figure 2: The land use dataset over the Gin Gin and Chittering LGAs.

From the land use classification "water bodies" were mapped to the highest degree of accuracy followed by "residential – single/duplex dwelling" (Table 3). The number of training samples from the "residential – single" class dominate the training dataset with 69% of samples in this class. This is however representative of the proportion of this class in the land use dataset.

Table 3: Classification accuracy per cover class.

Class	Accuracy	Number of training samples
Water body	98.56	355
Residential - single / duplex dwelling	97.23	7608
Unused - cleared - bare soil	89.15	721
Commercial	87.03	763
Rural residential / bush block	84.71	441
Trees / shrubs	76.12	317
Lifestyle block / hobby farm	67.97	116
Unused - cleared - grass	67.17	182
Transport / access - non-airport	66.18	186
Residential - multiple dwelling	66.14	145
Community facility - education	46.67	32
Recreation - turf/grass	35.71	53
Community facility - non-education*		17
Drainage*		6
Farm*		8

*insufficient number of training samples to include in the model

The accuracy of the random forest model as assessed within the model on “out of bag” (OOB) data was 91.81% (Table 4). The OOB data is withheld for testing in each fold (split, n = 10) within the model. This figure can be considered a good representation of the model performance. The model was also tested against a test set of data, which was excluded from the training process. The assessments of accuracy using the test data and OOB against overall accuracy and kappa score are high (> 0.8).

Table 4: Overall model accuracy against the test data and from the out of bag calculation.

Measure	Against test data	out of bag
Accuracy	0.928	0.918
Kappa	0.853	0.829

In total there were 304871 parcels which did not change their pin or shape between 2007 and 2016. These parcels were allocated the 2005 land use as it was assumed that their land use had not changed. The validity of this assumption has not been tested as it would require extensive field checking.

Parcels which changed land use between 2005 and 2007 were manually identified using aerial photography and the appropriate class was allocated. These occurred at several developments on the urban fringe.

In total 60 classes were generated. The composition of these classes is shown in Table 5.

Table 5: Land use composition of the Swan and Canning River catchments.

Land use class	**Area (Ha)**	**%**		**Land use class**	**Area (Ha)**	**%**
Unused - uncleared - trees / shrubs	47259.75	21.85		Irrigated land in transition	265.935	0.123
Farm	36805.97	17.017		Office - without parkland	220.464	0.102
Recreation / conservation - trees / shrubs	22833.34	10.557		Residential - aged persons	213.758	0.099
Residential - single / duplex dwelling	21753.4	10.057		Recreation - turf/grass	200.329	0.093
Transport / access - non-airport	17480.58	8.082		Poultry	185.477	0.086
trees / shrubs	14777.33	6.832		Unused and unfertilized	185.718	0.086
Lifestyle block / hobby farm	11273.02	5.212		Utility	167.651	0.078
Rural residential / bush block	8857.684	4.095		Stockyards/saleyards	126.987	0.059
Unused - cleared - grass	6352.974	2.937		Woody fodder plants	119.932	0.055
Animal keeping - non-farming	2647.772	1.224		Feedlots	114.863	0.053
Transport / access - airport	2531.437	1.17		Garden centre / nursery	85.58	0.04
Recreation - grass	2040.48	0.943		Residential - temporary accommodation	74.263	0.034
Unused - cleared - bare soil	1693.868	0.783		Turf farm	74.15	0.034
Recreation - turf	1665.049	0.77		Commercial / service - residential	62.513	0.029
Commercial	1656.311	0.766		Office - with parkland	55.46	0.026
Residential - multiple dwelling	1627.321	0.752		Landfill	54.081	0.025
Manufacturing / processing	1559.443	0.721		Plantation	54.809	0.025
Community facility - education	1496.296	0.692		Irrigated sown grasses	38.066	0.018
Tree plantation	1440.184	0.666		Glasshouse horticulture	26.412	0.012
Perennial Horticulture	1254.938	0.58		Turf Farm	24.484	0.011
Viticulture	1253.544	0.58		Cropping	21.241	0.01
Storage / distribution	1245.439	0.576		Abattoirs	15.33	0.007
Commercial / service - centre	895.175	0.414		Mixed grazing	15.822	0.007
Community facility - non-education	640.145	0.296		Intensive animal husbandry	13.535	0.006
Water body	525.178	0.243		Sewage - treatment plant	11.641	0.005
Horticulture	512.395	0.237		Yacht facilities	8.681	0.004
Animal keeping - non-farming(horses)	493.077	0.228		Aquaculture	7.261	0.003
Quarry / extraction	440.599	0.204		Land in transition	7.158	0.003
Drainage	373.649	0.173		Sewage - non-treatment plant	2.354	0.001
Annual horticulture	336.385	0.156		Piggery	0.675	0

				NA	116.501	0.054

3.2 Land Cover

An example of the land cover dataset is shown in Figure 3.

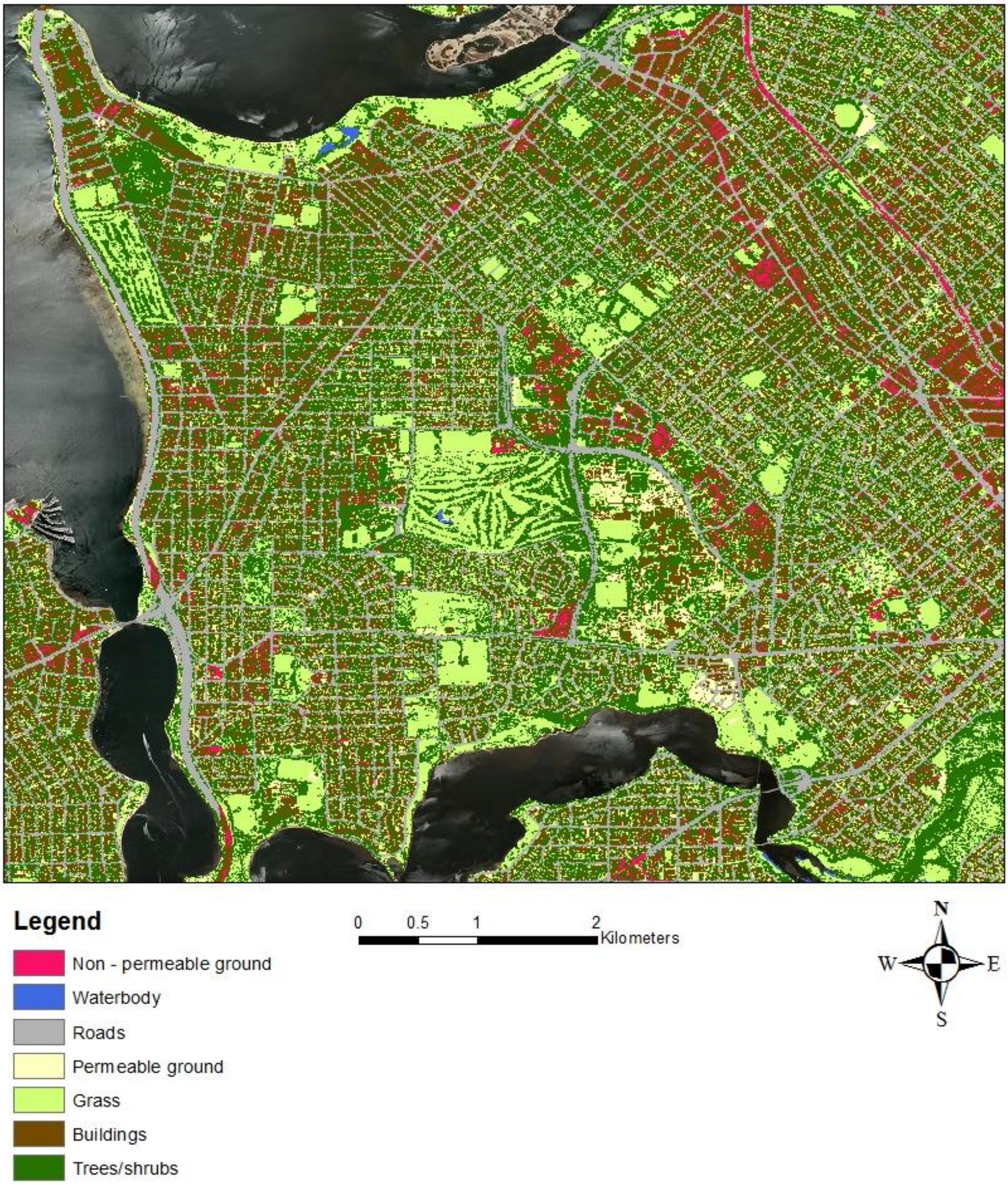

Figure 3: An example of the land cover dataset.

The land cover classification achieved an overall accuracy of 80.7%. Classes with the highest accuracy were trees/shrubs, buildings and roads (Table 6). Non - permeable ground and waterbodies recorded the lowest accuracies. These two classes also occupy the smallest proportions of the study area. The trees/shrubs (33.14%) and grass (37.52 %) classes occupied the largest proportions.

Table 6: Land cover accuracy by class.

Class	Accuracy	% of study area
Trees/shrubs	0.936	33.14
Buildings	0.94	9.16
Grass	0.886	37.52
Permeable ground	0.844	12.26
Roads	0.957	5.77
Waterbodies	0.835	0.25
Non - Permeable ground	0.813	1.9

A confusion matrix of land cover classes is shown in were also classified as buildings. This appears mainly due to variable topography, where areas of raised land are classified as above ground (Table 7).

The largest source of error appears to be waterbodies being mapped as trees/shrub. This may be due to emergent vegetation within waterbodies, or due to anomalous reflectance and height values that occur over water in Urban Monitor imagery. In terms of the overall dataset this error is quite small as waterbodies only account for 0.25 % of the study area.

Errors in the roads class include misclassification as trees/shrubs and buildings. The misclassification of roads as buildings is due to a combination of slight errors in the NSM and roads visible in the imagery which are not in the cadastre. A number of the non-permeable ground points (n = 9) were also classified as buildings. This appears mainly due to variable topography, where areas of raised land are classified as above ground.

Table 7: Land cover confusion matrix.

	Trees/shrubs	Buildings	Grass	Permeable ground	Roads	Waterbodies	Non - permeable ground
Trees/shrubs	46	3	1	0	1	15	0
Buildings	0	46	0	0	3	0	9
Grass	3	1	41	8	0	1	1
Permeable ground	0	0	6	36	0	0	8
Roads	0	0	2	0	46	0	0
Waterbodies	0	0	0	1	0	33	0

Non-permeable ground	0	0	0	4	0	0	32

4 Discussion

Identifying unchanged parcels is a crucial part of the analysis. Processing speeds and similarities between many of the classes mean that classifying all parcels using the Urban Monitor imagery is not feasible. The accuracy achieved by assigning unchanged parcels an unchanged land use has not been assessed but it is assumed to be greater than what could be achieved if the UM modelling method could be applied.

The overall model accuracy (92.4%) is dominated by the “Residential - single / duplex dwelling” class (69% of training data). This is representative of the changed land parcels – the major change is to residential dwellings.

Sources of error include:

- Land parcels which were assumed to have not changed in land use but have.
- Error in the DPIRD land use mapping.
- Parcels attributed as road in the cadastre which are not yet road.
- Classification error in the prediction using UM data.

The classification accuracy of the land cover map was high (80%). This accuracy figure is quite conservative given that the classes with the lowest accuracy, waterbodies and non-permeable ground, occupy the smallest proportions of the study area. The dominant sources of error appear to be due to the variable reflectance of waterbodies, which can be influenced by algae, emergent vegetation and sun glint. Variability in the NSM where raised features are considered above ground is also a source of error in the roads and non-permeable ground classes.

Appendices

Appendix 1 DPIRD Land Use Metadata Statement

Department of Agriculture and Food, Western Australia

Title: Intensive Agricultural Land Use in Western Australia ALUM v7
Job Number: 2015003, 2015151
Custodian: Department of Agriculture and Food, Western Australia (DAFWA)
Jurisdiction: Western Australia

Abstract: This dataset provides a major revision to the DAFWA Land Use in Western Australia Australian Land Use and Management (ALUM) v6 dataset for intensive horticultural classes under perennial and seasonal horticulture classifications 3.4, 3.5, 4.4 and 4.5 and under classification 5.2 – Intensive Animal Husbandry - and class 5.3.5 - Abattoirs. This is the first attempt using ALUM methodology to map specific land use classes rather than mapping a complete region or catchment at one time.

Recommendation: It is recommended that any land use area figures derived from the data be used only as a preliminary guide due to land use capture being undertaken primarily by aerial photo interpretation and at various scales of capture.

Geographic Extent: Western Australia, as is made available.

Beginning Date: January 2015
Ending Date: Current
Photography Date: Various – from 2008 to 2015 depending on the region.

Progress: The capture process has been completed, however potential future projects may update / refine the data further or expand the classification captured in this dataset.

Maintenance and Update Frequency: At this stage this is a static dataset but this may change depending on demand for updates. Updates would be most likely to be undertaken as new imagery becomes available so timeframes and frequency will be region dependant.
Potential future projects may update / refine the data further or expand the classification captured in this dataset.

Stored Data Formats: Oracle Spatial Database UTM Zone 50 GDA94
Available Data Formats: On request
Access Constraint: License on request

Lineage: A preliminary base dataset was derived by extracting data from the DAFWA Client, Property and Event (CPE) database and existing ALUM v6 data overlaid on aerial photography.

An area of interest (AOI) was digitised for specified land uses and an ALUM v7 code was assigned by the land use mapping project team.

Land not involved in Intensive Agriculture land uses was not captured.

Aerial photography dates vary across geographic extent and are incorporated into individual polygons in the final dataset.

Positional Accuracy: ~ + or – 50 m.

Attribute Accuracy:
Land Use interpretation is subjective and based on experience and knowledge of the GIS operator capturing land use.

Land Use coding interpretation is open to discrepancies based on the GIS operators' interpretation of the coding system used (ALUM).

The ALUM code is a generalised coding system and does not allow for 'variety' level information to be incorporated into defining horticultural activity. Where more information was available and was deemed relevant it was added as a comment in the Description field or as a DAFWA Activity Code.

Spatial and Attribute verification of the dataset was undertaken via:
- aerial photography interpretation
- industry expert confirmation
- some comparison of custodial/corporate data

Completeness:
The dataset covers all intensive agriculture and abattoir sites in the state of WA of which DAFWA staff are aware.

Supplemental Information:
Database Fields:
1. AOI ID - Area of Interest Identifier
2. ALUM Code - Australian Land Use and Management (ALUM) v7 classification code
3. Description - Additional Descriptive Information
4. Reliability - Reliability of attribute
5. Audit Comments – Additional comments made during audit process
6. Orthophoto Date - Year of run - if not in file name
7. Orthophoto Filename - Orthophoto name
8. Area (ha) - size of AOI in hectares (autopopulated)
9. Created by - Autopopulated field through Oracle
10. Creation Date - Autopopulated field through Oracle
11. Last Update by - Autopopulated field through Oracle
12. Last Update date - Autopopulated field through Oracle
13. Activity Code – DAFWA activity code

Reference Tables:
1. ALUM Classification
2. Reliability
3. DAFWA Property Activity Codes

Contact Organisation: Geographic Information Services Group
Department of Agriculture and Food, Western Australia
Contact Project Person: Avril Russell-Brown / Jeff Watson
Mail Address: 3 Baron-Hay Court
Locality: South Perth
State: Western Australia
Country: AUSTRALIA
Postcode: 6151
Telephone: (08) 9368 3728
Email Address: gis@agric.wa.gov.au

Metadata Date: 26th April 2016

References

Caccetta, P, S Collings, A Deveraux, K Hingee, D McFarlane, A Traylen, X Wu, and Zheng-Shu Zhou. 2015. “Monitoring Land Surface and Cover in Urban and Peri-Urban Environments Using Digital Aerial Photography.” *International Journal of Digital Earth*. https://doi.org/10.1080/17538947.2015.1046510.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wright, Marvin N., and Andreas Ziegler. 2017. “Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R.” *Journal of Statistical Software* 77 (1): 1–17. https://doi.org/10.18637/jss.v077.i01.