2013/2014

REPORT

South African National Land-cover Dataset

SEPTEMBER 2015

environmental affairs
Department:
Environmental Affairs
REPUBLIC OF SOUTH AFRICA

energy
Department:
Energy
REPUBLIC OF SOUTH AFRICA

sanedi
South African National Energy Development Institute

UKaid
from the British people

Strategic Climate Policy Fund

Sustainability of Decentralised Renewable Energy Systems

SCPF007
Prepared for
Department of Environmental Affairs

Prepared by
GeoTerraImage

A project managed by
Cardno Emerging Markets

26 August 2015

The project was funded by the Department for International Development (DFID), UK Government

UKaid from the British people

SCPF

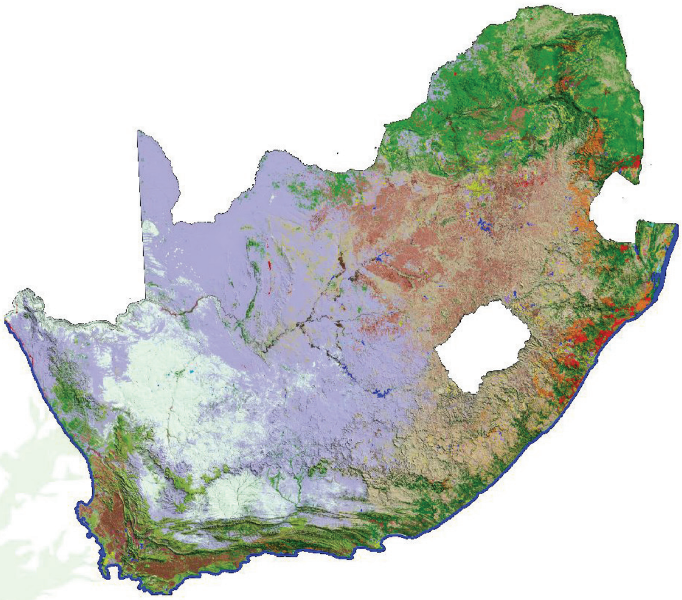
SOUTH SOUTH NORTH

REPORT

2013/2014 South African National Land-cover Dataset

DEA/CARDNO SCPF002: Implementation of Land-use Maps for South Africa

Project-specific Data User Report and Metadata



DATA PRODUCT CREATED BY

GEOTERRAIMAGE (South Africa)

www.geoterraimage.com

July 2015, version 05b *(pivot code-corrected data)*

OPEN ACCESS LICENCE

LICENSE AND COPYRIGHT

The data supplied is licensed by GEOTERRAIMAGE (GTI) to the client, but remains the property of GTI as is protected by copyright laws. The data can be used by the client as per the conditions set forth in the offer. The client does not obtain any ownership or intellectual property rights in the data. "Client" refers to the South African Department of Environmental Affairs (DEA).

The following shall always be clearly mentioned: ©GEOTERRAIMAGE – 2014.

OPEN DATA LICENSING (PERMITTED AND INTENDED DATA USE)

The DEA is granted an Open Data User license, which allows it unrestricted access to and use of the supplied data. The data is defined as the 72 Class GTI South African National Land-cover Dataset (2013/14). The Open Data license allows the DEA to distribute the data unrestricted to third parties with the full understanding that the DEA and third parties may not use the data to develop new products that will compete directly with GTI's existing or "in-progress" commercial data products; and that the DEA and third parties accept and understand the stated data copyright, licensing and liability conditions.

LIABILITY

The client has to verify, prior to his commitment, the adequacy of GTI datasets and products to the specific data and analyses needs as set out by the client. GTI will in no event be liable for any direct, indirect, incidental or consequential damages, including lost savings or profit, legal fees, lost data, lost sales, loss of commitments, business interruption, including loss of or other damage, resulting from its datasets and products, whether or not GTI has been advised of the possibility of such damage. This includes damages incurred by the client or any third party.

GTI's datasets and products are processed from sources like topographic maps and satellite images. The quality of these sources is not the responsibility of GTI. The data and data products are provided "as is", without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose.

Whilst all possible care and attention has been taken in the production of the data, neither GTI nor any of its respective suborganisations or employees accept any liability whatsoever for any perceived or actual inaccuracies or misrepresentations in the supplied data or accompanying report(s), as a result of the nature of mapping from coarse and medium-scale satellite imagery. Any mapped boundaries, either implied or inferred within the data, do not have any legal status whatsoever.

All rights not specifically granted or specified in this agreement, pertaining to this data, are reserved to GEOTERRAIMAGE (Pty) Ltd.

CONTENTS

PLEASE NOTE:
This report provides information on the specific version of the 2013/14 GTI South African National Land-cover Dataset supplied to the South African Department of Environmental Affairs under the DEA/CARDNO Project: SCPF002: *Implementation of Land-use Maps for South Africa*.
The supplied dataset referred to in this report represents a *subset* of the full 2013/14 GTI South African National Land-cover Dataset, in terms of the supplied land-cover/land-use categories and associated class legend.
Use of this DEA/CARDNO data version is governed by the LICENCE AGREEMENT shown previously.

1. INTRODUCTION

Land cover is a key information requirement for a wide range of landscape planning, inventory and management activities, ranging from environmental resource management to telecommunication planning. The recent global availability of Landsat 8 satellite imagery offered the opportunity to create a new, national land-cover dataset[1] for South Africa, circa 2013/14, replacing and updating the previous 1994 and 2000 South African national land-cover datasets. The 2013/14 National Land-cover Dataset is based on 30 x 30 m raster cells, and is ideally suited for approximately 1:75 000- to 1:250 000-scale geographic information system (GIS)-based mapping and modelling applications.

The dataset has been derived from multi-seasonal Landsat 8 imagery, using operationally proven, semi-automated modelling procedures developed specifically for the generation of this dataset, based on repeatable and standardised modelling routines. The dataset has been created by GEOTERRAIMAGE (GTI) and is available as a commercial data product. The production of a comparable 1990 land-cover product is also in progress. This will provide, for the very first time, a standardised and directly comparable set of national land-cover datasets that can be used to determine landscape change in South Africa over a period exceeding 20 years[2].

2. BACKGROUND

The recent global availability of Landsat 8 imagery (April 2013) was the primary catalyst behind the development of new, innovative, automated land-cover modelling procedures, which allowed the creation of the 2013/14 South African National Land-cover Dataset; primarily because Landsat 8 offered a free and regular supply of radiometrically and geometrically standardised, medium-resolution, multi-spectral imagery, suitable for medium to large area mapping. Collectively, this offered an ideal opportunity for GTI to re-examine using automated land-cover mapping techniques as a more efficient alternative to conventional, analyst-assisted per-pixel classifiers, allowing the rapid production of standardised, yet informative land-cover information.

3. OBJECTIVE

The primary objective was to generate a new, national land-cover dataset for the whole of South Africa, and to be able to release this as a commercial data product within about a year of image data acquisition to ensure that the dataset is up to date and current. In support of this primary objective was the requirement to develop operational procedures based on semi-automated (spectral) modelling techniques that would facilitate the rapid production of consistent, standardised land-cover information from multi-seasonal Landsat 8 imagery.

4. PRODUCT DESCRIPTION

The 2013/14 South African National Land-cover Dataset produced by GTI as a commercial data product has been generated from digital, multi-seasonal Landsat 8 multi-spectral imagery, acquired between April 2013 and March 2014. In excess of 600 Landsat images were used to generate the land-cover information, based on an average of eight different seasonal image-acquisition dates, within each of the 76 image frames required to cover South Africa. The land-cover dataset, which covers the whole of South Africa, is presented in a map-corrected, raster format, based on 30 x 30 m cells equivalent to the image resolution of the source Landsat 8 multi-spectral imagery. The full dataset contains 72 land-cover/ land-use information classes, covering a wide range of natural and man-made landscape characteristics. The DEA/CARDNO dataset contains a summarised 35-class legend. Each data cell contains a single code, representing the dominant land-cover class (by area) within that 30 x 30 m unit, as determined from an analysis of the multi-date imagery acquired over that image frame. The original land-cover dataset was processed in UTM35 (north)/WGS84 map-projection format based on the Landsat 8 standard map-projection

1 Note that the term "land cover" is used loosely to incorporate both land-cover and land-use information in the context of the GTI 2013/14 South African National Land-cover Dataset.

2 The existing 1994 and 2000 South African national land-cover datasets were generated independently, using very different source image formats (i.e. paper versus digital), and end-product formats (i.e. 1:250 000 scale digital vector versus 1:50 000 digital raster), despite being derived from equivalent Landsat imagery. See Thompson, MW, 1991. South African National Land-cover Database Project. Data users' manual. Final report (Phases 1, 2 and 3). CSIR Project Report ENV/P/C 98136, February 1999; and Thompson, MW et al, 2001. Guideline procedures for national land-cover mapping and change monitoring. CSIR Client Project Report ENV/P/C 2001-006, March 2001.

format as provided by the United States Geological Survey (USGS). The final product is available in UTM35 (north and south), WGS84 map projections and geographic coordinates.

4.1 Land-cover Legend

The 2013/14 South African National Land-cover Legend is aligned with the SANS 1877[3] South African National Land-cover Classification Standard in terms of class definitions and in hierarchical format. The complete list of information classes that have been mapped and supplied as part of the DEA/CARDNO land-cover dataset and the associated class definitions are supplied in Appendix A.

Note that in recent Chief Directorate: National Geospatial Information (CD: NGI) national land-cover mapping projects, the Food and Agriculture Organisation: Land Cover Classification vs 2 (FAO LCCS2) land-cover classification system has been adopted as an alternative to SANS 1877, although this is not a national standard. However, in September 2014, the South African Group Earth Observation (SA GEO) Group's Land-cover Community of Practice (CoP) task team, which includes CD: NGI representation, proposed that the class detail and content of SANS 1877 be retained/included in a proposed new South African national land-cover standard, which will be based on the latest, internationally accepted Land-Cover Metadata Language (LCML) classification system (ISO1944-2).

5. OVERVIEW OF THE AUTOMATED MODELLING APPROACH

Automated modelling procedures offer significant advantages in terms of ensuring data standards, minimising processing time, allowing easy repeatability and facilitating accurate change detection, when compared to more conventional image mapping approaches where there is a greater reliance on individual image analyst knowledge and inputs. To this end, a series of automated image-based modelling steps were developed that utilise the seasonal dynamics associated with the broad landscape characteristics across South Africa. These were then used to rapidly produce a set of foundation cover classes that could be easily converted into more meaningful land-cover information categories, using predefined geographical masks in the GTI data libraries.

The foundation cover classes represent the basic building blocks associated with all landscape characteristics, namely water, bare ground, grass and tree-bush-shrub cover types, with each being defined in terms of seasonal occurrence or permanence. These basic foundation cover classes represented the initial output from the automated modelling approach. The foundation cover classes are then converted into more conventional land-cover information classes, i.e. urban, forest plantation, etc., as part of the post-automated modelling data-processing steps.

The foundation cover classes are essentially "spectrally dependent" classes, since they are generated from automated modelling procedures that are based directly on the spectral characteristics associated with each image pixel over time (i.e. seasonal) and space (i.e. within an image frame). The final land-cover information classes are referred to as "spectrally independent" classes since different cover classes can share similar foundation class spectral characteristics in a one-to-many-type relationship. For example, the "bare ground" spectrally dependent cover classes could represent non-vegetated built-up urban areas, natural rock exposures, beach sand or a mine pit and tailings dump. Similarly, the tree-bush classes could represent natural vegetation cover, a timber plantation or a fruit orchard. The advantage of this approach is that the conversion of the initial, spectrally dependent foundation cover classes into the final, spectrally independent land-cover information classes can be tailored to suit a variety of end-user information requirements; simply by using a different set of predetermined masks and foundation class subdivisions and amalgamations.

5.1 Model portability to other geographical areas and sensors

Although model development was focused on using Landsat 8 imagery within South Africa, the same models have also been proven to work with equivalent success

3 SANS, 1877. South African Bureau of Standards (SABS)-designated national land-cover classification standard for South Africa.

on Landsat 8 imagery over sites in Botswana, Mozambique, Namibia, Sudan and Zimbabwe, as well as using comparable Landsat 5 archive imagery over South Africa; all of which indicate a high level of model portability. This should allow and support the production of directly comparable historical land-cover datasets for change detection, assuming of course that the required level of seasonal image coverage is available in the data archives.

5.2 Use of object-based modelling

No attempt was made to include or use object-based modelling in the automated mapping process, primarily because the medium-resolution format of multi-spectral Landsat 8 imagery does not lend itself to this type of modelling, since pixel resolution typically precludes the identification of true landscape "objects" in comparison to high- and ultra-high-resolution image formats. The 30 m Landsat resolution pixels typically represent a mix of land-cover characteristics rather than a pure cover surface, i.e. an urban pixel is typically a composite of building roofs, garden vegetation and/or road surfaces. However, object-based modelling may be a useful approach for helping to generate second-level information classes from the primary-level foundation class dataset, such as separating water in rivers (i.e. natural) from water in dams (i.e. artificial), based on size, shape and context.

6. LANDSAT IMAGERY

The primary imagery source used in the generation of the 2013/14 South African National Land-cover Dataset was 30 m-resolution Landsat 8 imagery, acquired between April 2013 and March 2014. Within each image frame, a range of seasonal image-acquisition dates were used (within the automated modelling procedures) to characterise the seasonal dynamics across the landscape in terms of the basic tree, bush, grass, water and bare foundation cover classes. Nine image-acquisition dates per image frame per complete seasonal cycle were taken as the optimal number of dates, based on Landsat's 16-day overpass schedule[4]. Unfortunately, due to localised, prolonged cloud-cover problems in some regions during this April 2013 to March 2014 period, it was not possible to achieve this. In such cases, archival Landsat 5 imagery was used as a substitute, but only if it was from a suitable seasonal period to complement the Landsat 8 data, and was almost 100% cloud-free. Figure 1 illustrates the total number of image-acquisition dates per image frame, while Figure 2 illustrates the location and number of Landsat 5 images that were also used.

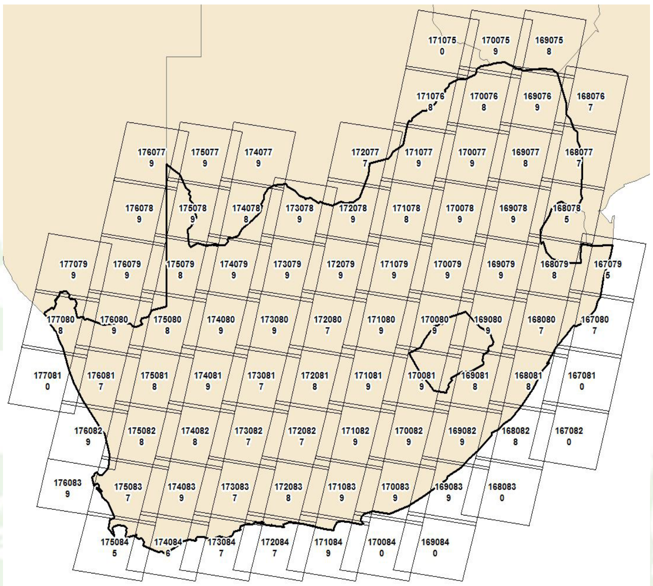


Figure 1. Total number of image-acquisition dates used in the land-cover modelling per image frame.

4 This seasonal window requirement, while improving the overall accuracy of interpretation and land-cover modelling, means that – in practice – between 12 and 16 months are actually required to acquire sufficient cloud-free seasonal imagery for input into the modelling processes. This means that a new land-cover dataset could only be generated approximately every two to three years (minimum), since a shorter period may not provide sufficient separation between the seasonal input image dates.

In summary, 76 Landsat image frames were required to provide complete coverage of Lesotho, South Africa and Swaziland. A total of 616 individual Landsat images were used to model and produce the land-cover data, representing an average of eight acquisitions per image frame per year. Of the 616 images, 592 images were from Landsat 8 (96%) and 24 images (4%) were from Landsat 5. The Landsat 5 images were, however, only used in 10 of the 76 image frames defining the geographic extent of the land-cover dataset.

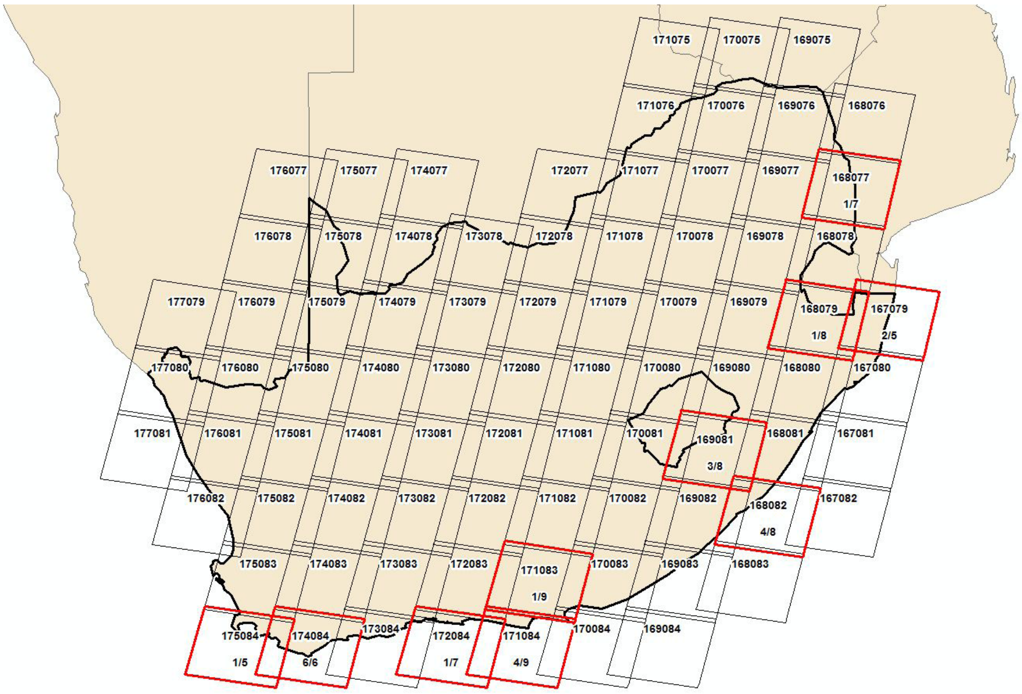


Figure 2. Location of the 10 image frames that required the use of archival Landsat 5 imagery due to the limited availability of Landsat 8 data during the period April 2013 to March 2014.

6.1 Landsat image sourcing

All Landsat 8 and Landsat 5 imagery used in the 2013/14 land-cover data modelling was sourced from the online data archives of the USGS[5]. The data is provided in a precise geo-corrected UTM35 (north)/ WGS84 map-projection format, and was used "as-is" without any further geo-correction. A complete list of the image-acquisition dates used in the modelling of the 2013/14 South African National Land-cover Dataset (for both Landsat 8 and Landsat 5) accompanies this document (see Appendix B).

5 See http://glovis.usgs.gov/.

6.2 Landsat data preparation

All Landsat 8 and Landsat 5 imagery was standardised to 16-bit, top-of-atmosphere (ToA) reflectance values prior to land-cover modelling, using a generic modelling approach. The original USGS-generated UTM35 (north), WGS84 map-projection format was retained without change or modification throughout all image frame-based modelling procedures. As far as possible, only cloud-free or image dates with limited cloud cover were used in the land-cover modelling (i.e. maximum of about 20% terrestrial cloud cover in any one date). Any cloud-affected regions were corrected by merging cloud-affected ToA-corrected imagery with cloud-free ToA-corrected data from preferably the preceding or following overpass date, so that, as far as possible, the final cloud-free merged imagery composite only represented a maximum difference of about 16 days.

Cloud masks were created using either conventional pixel classification procedures (within analyst-defined sub-image areas), or spectral-based modelling using generic thermal and blue light reflectance thresholds[6], depending on cloud characteristics. This was deemed acceptable in terms of minimising any changes in local vegetation cover growth changes. Approximately 35% of the 616 images used in the land-cover modelling were cloud-masked composites. No external atmospheric correction was applied to the image data.

7. LAND-COVER MODELLING

7.1 Spectral modelling

Derived spectral indices, generated from the ToA-corrected imagery, were used as the only inputs into the land-cover models. No original spectral (ToA reflectance) data was used as an input, although collectively, five out of the eight available non-thermal spectral bands were used in the various spectral indices[7]. A standardised set of spectral indices were identified from which the required foundation cover classes – (1) tree dominated, (2) bush dominated, (3) grass dominated, (4) water and (5) bare ground – could be modelled, using predetermined, generic spectral threshold values. The generic threshold values associated with each index and cover type were tested over several landscapes and seasons before being confirmed and accepted as such.

The spectral indices included both existing algorithms, such as the Normalised Difference Vegetation Index (NDVI) and Normalised Difference Water Index (NDWI), as well as algorithms developed in-house specifically for the GTI land-cover modelling requirements. All models were developed in ERDAS Imagine© image-processing software using the Model Maker function.

The modelling and generation of each foundation cover class were undertaken as a separate modelling exercise, i.e. water was modelled separately from bare ground, which was modelled separately from tree and bush cover, etc. This approach simplified the modelling steps and facilitated easier desktop quality control of outputs, compared to attempting to model all foundation cover types simultaneously within a single model workflow. In most cases, the final geographic extent of a particular foundation cover class was generated from the combined use of several different spectral indices, since no single index was found to work well in all landscapes and all seasons.

All modelling was undertaken on an individual image frame-by-frame basis, where the original USGS-supplied UTM35 (north) map-projection format was retained "as-is". This allowed the models to be adapted according to the number of acquisition dates and associated seasonal ranges available for each particular image frame. The model modifications reflected the need to standardise the outputs according to predefined definitions of seasonal permanence, for example, whether a foundation cover class occurred, i.e. in only one image date; in several, but not all image dates; or in all image dates. Obviously, the more image-acquisition dates available per frame, and the wider the seasonal range, the more accurate the modelled interpretation of a particular cover class's seasonal characteristics.

Examples of the separate model outputs for bare ground and tree/bush foundation cover classes are shown in Figure 3 below. The examples show eastern Pretoria, from within Landsat 8 frame 170-078.

6 The spectral modelling approach could only be applied to Landsat 8 imagery (if applicable) due to its enhanced spectral band range, compared to Landsat 5.

7 (2) Green, (3) Red, (4) Red, (5) NIR, (6) SWIR-1 and (7) SWIR-2 Landsat 8 spectral bands.

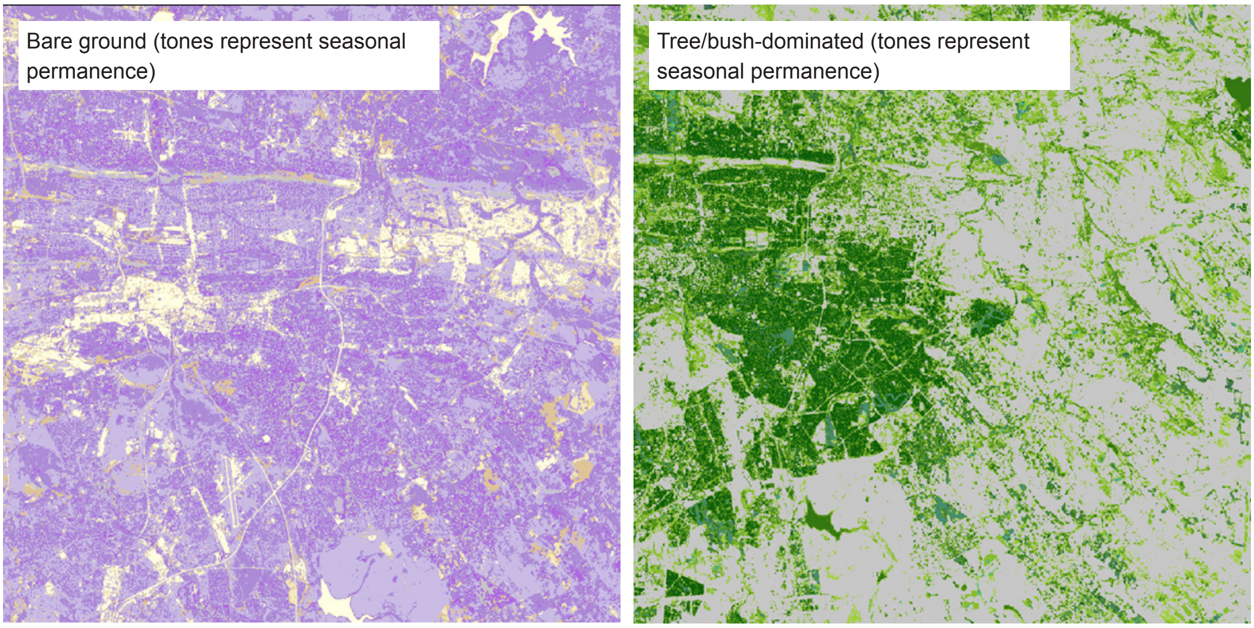


Figure 3. Spectral index-derived model outputs for bare ground and tree/bush-dominated foundation cover classes over eastern Pretoria (frame 170-078).

In some circumstances, modifications or additional modelling steps were used to account for regional landscape characteristics that could not be accurately modelled using only generic models. For example, in some of the far western arid areas, additional tree/bush modelling steps were taken using a slightly modified modelling approach in order to better detect and distinguish between the sparse, low bush and shrub covers in these landscapes.

7.2 Terrain modifications

Terrain-based modifications were included within some of the foundation cover class modelling procedures (i.e. water, bare and tree/bush classes) in order to minimise seasonally induced terrain shadowing effects, using a combination of solar illumination and slope parameters. These parameters were modelled independently using the 90 m-resolution Shuttle Radar Topography Mission (SRTM) dataset[8], with outputs being resampled to a Landsat-comparable 30 m-resolution format, prior to being incorporated into the foundation cover class models. For example, solar illumination and Digital Elevation Model (DEM) slope masks were used to minimise any spectral confusion between dark terrain shadow areas and comparable water body reflectance levels. Note that when creating the DEM-derived parameters, such as shadow extents, modelling was based on the full range of seasonal solar azimuth and zenith angles and not a specific date or season in order to better align with seasonal differences evident across any given range of image-acquisition dates.

8 Although the 30-m SRTM terrain dataset would have been preferable, it did not become publicly accessible until after the completion of the bulk of the 2013/14 land-cover modelling had been completed.

7.3 Individual frame spectral models

The basic modelling parameters used to create each of the foundation cover classes are described below. In each case, the modelled output would typically represent a (qualitative) gradient of spatial densities, as well as temporal, i.e. seasonal occurrence. For example, the bare ground model output consisted of a set of classes that represented a range of conditions from “pure bare” to “mixed bare/sparsely vegetated”, with each described in terms of seasonal, permanent or temporary occurrence. The foundation water class modelling was the only exception to this, with the model output describing water extent in terms of only permanent or seasonal occurrence.

7.3.1 Spectral modelling of water

Water was modelled using a combination of five spectral indices, including a burn index, and DEM-derived solar illumination and slope masks. The burn index was used to minimise any spectral confusion between dark burnt areas and comparable water body reflectance. The solar illumination and slope masks were used to minimise any spectral confusion between dark terrain shadow areas and comparable water body reflectance. Generic spectral water models were developed for both flat (standard) and hilly terrain, and for shallow water (pan)-dominated landscapes.

7.3.2 Spectral modelling of bare ground

Bare ground was modelled using a combination of two spectral indices, and DEM-derived solar illumination and slope masks. The solar illumination and slope masks were used to minimise any spectral confusion between dark terrain shadow areas and comparable bare ground reflectance. The output represented a gradient of bare ground from pure bare to sparsely vegetated cover.

7.3.3 Spectral modelling of tree/bush cover

Tree and bush cover was generated within the same spectral model. Both woody covers were modelled using a combination of four spectral indices, and a DEM-derived slope mask. The slope mask was used to minimise any spectral confusion between dark terrain shadow areas and comparable woody cover in two of the four spectral indices. “Trees” refers to dense, typically tall woody cover, such as natural and planted forests, dense woodland and thickets, whereas “bush” refers to more open, often lower mixed tree/bush communities, such as typical bushveld or open woodland. The spectral modelling procedure also separated woody cover spectral classes that showed similar characteristics to other vegetated covers, such as crops, sportsfields, golf courses (especially if irrigated), from those that did not show similar characteristics. An additional “desert” tree/bush model was developed as an extra model to be run in addition to the standard model in more arid areas to increase the representation of bush cover in these regions where the modelling thresholds used in the standard model were not sensitive enough.

7.3.4 Spectral modelling of grass

Grass cover was modelled using only a single spectral index. No DEM-derived solar illumination or slope mask modifiers were used.

7.3.5 Spectral modelling of burnt areas

Burnt cover was modelled using only a single spectral index. No DEM-derived solar illumination or slope mask modifiers were used.

7.4 Seasonally defined spectral land-cover per image frame

The modelled outputs for each foundation cover class were combined into a single composite dataset for each frame in a predetermined hierarchical order so that the final output cover class combinations could reflect the dominant and subdominant cover types in terms of seasonal occurrence. For example:

- water in all dates (permanent)
- water in many dates (seasonal)
- dominated by trees in all dates
- dominated by trees in many but not all dates, plus bare ground in one date, etc.

Figure 4 shows the output after all the separate foundation land-cover classes have been combined into a composite set of seasonally defined foundation land-cover classes. The illustrated area is the same eastern Pretoria area as shown in Figure 3. This interim land-cover product consists of 51 seasonally defined foundation cover classes, as listed in Appendix C. These seasonally combined foundation cover classes are still essentially “spectrally dependent” classes, since they have been generated using automated modelling procedures on only the spectral characteristics associated with each image pixel over time (i.e. seasonal) and space (i.e. within an image frame).

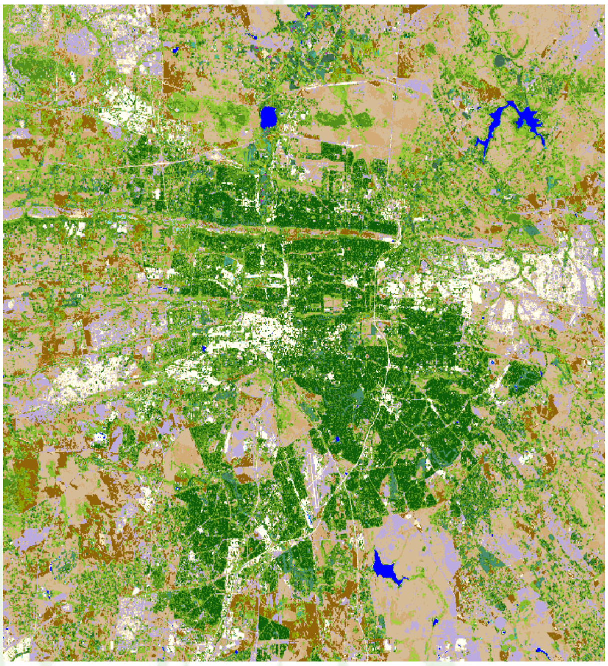


Figure 4. Interim spectral land-cover product after combining all the seasonally defined foundation cover classes (eastern Pretoria, frame 170-078).

Note that, in addition to the basic tree, bush, grass, water and bare ground foundation classes already described, fire scar extent and occurrence were also modelled and incorporated into the final combined foundation cover composite classes in order to provide an indication of whether or not fire was a factor in the modelled seasonal cover profiles, i.e. a single-date, bare ground area could be the result of a wildfire event that then revegetated later in the season.

The spectrally dependent and seasonally defined foundation cover classes in this base dataset represent the "building blocks", from which more information-focused land-cover classes can be derived, especially if combined with ancillary datasets such as vegetation or land-use maps to facilitate further class subdivisions or area-based class modifications. The entire spectral modelling procedure is summarised graphically in Figure 5 below.

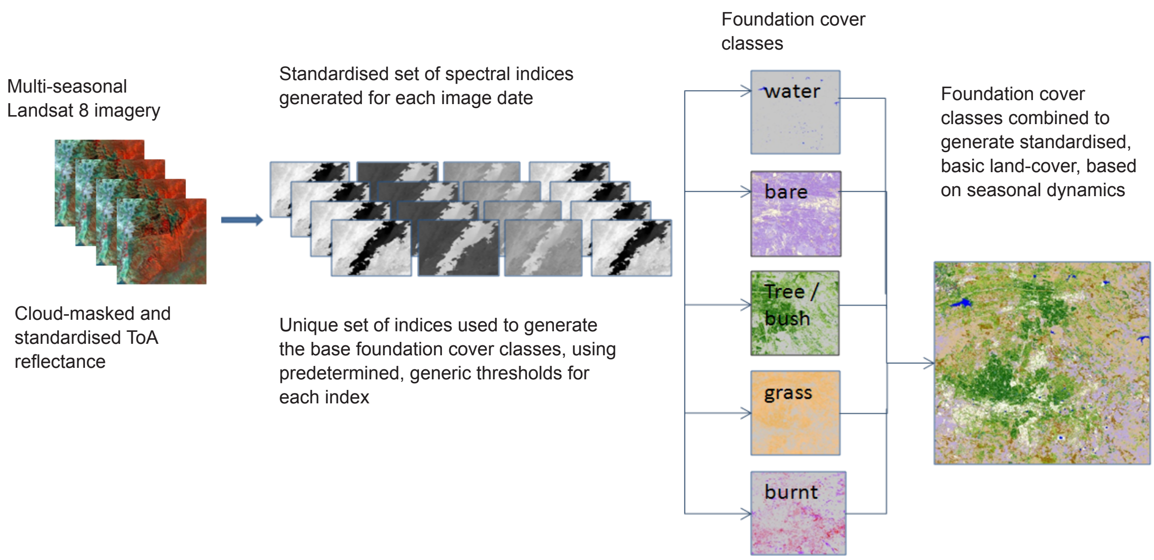


Figure 5. Summary of the full spectral land-cover modelling process.

7.5 Generation of full South African spectral class mosaic

Once the spectral modelling for the individual frames had been completed, the next step was to merge these into a single, seamless coverage for the whole of South Africa. Achieving this involved several processing steps. Individual image-frame datasets were first edge-cleaned to ensure that no frame-edge pixel anomalies would be carried over into the final countrywide mosaic. Cleaned image frames were then re-projected were necessary to a standardised UTM35 (north), WGS84 map projection, since the majority of image datasets covering South Africa were sourced from the USGS data archives, and were in this UTM zone.

Individual image frames were merged into a single, countrywide spectral foundation class composite, based on an analyst-determined sequence of image border overlap rules in order to ensure the best possible seamless integration of the individual frames. This approach was necessary since some image-to-image edge differences occur between image frames where spectral foundation classes have been generated from significantly different input acquisition dates, simply as a result of image data availability.

Although the standardised modelling approach used to generate the spectral classes supported the generation of spatial and content-comparable land-cover data in adjacent image frames (if both frames utilised comparable input image-acquisition dates), if there was a significant difference in the number or seasonal range of input images, then adjacent images often showed edge-matching differences. In such cases, the adjacent images were often spatially comparable, but not necessarily comparable in terms of content. In these cases, the edge-matching process was set up, as far as possible, so that the most prominent and spatially extensive landscape pattern was maintained across the image borders in order to provide a more visually seamless image frame overlap.

Note that for water ("seasonal" and "permanent") and bare ground ("pure bare, all dates") countrywide coverages, the maximum extent of each class, from all images, was transferred across to the final countrywide spectral class composite, regardless of any frame overlap edge-matching applied to other spectral classes.

7.5.1 Grassland corrections: inland

Persistent cloud-cover problems on several image-acquisition dates during the 2013/14 wet season period over the Highveld region resulted in the unavailability of usable, early wet-season imagery for modelling inputs. This resulted in the standard spectral models under-estimating the grassland extent. In order to correct this, an alternative grassland model was developed, based on a single – rather than multiple – image-acquisition date (the optimal single date representing the best dry-season image for visually separating green woody vegetation from senescent grassland vegetation, with minimal burn scars). For image frames covering areas of significant topographical relief, such as over the escarpment, the single-date grass-corrective model included the use of solar illumination and slope parameters to improve modelling outputs. For standardisation purposes, the inland grassland corrective model was applied to all image frames overlapping or contained within the South African National Botanical Institute (SANBI) grassland biome boundary. The resultant grassland extents for each image frame where then merged into a composite geographical mask and integrated back into the original countywide spectral class dataset to improve the overall grassland delineation.

7.5.2 Grassland corrections: coastal

In some high-rainfall coastal regions, the standard spectral models were unable to accurately distinguish between grassland and woody cover, where the grass biomass was exceptionally high as a result of high local rainfall during the 2013/14 wet season, even if suitable image-acquisition dates were available. This was especially the case around Pondoland (Eastern Cape) and Zululand and Maputoland (KwaZulu-Natal). To correct this, another grassland correction model was developed with alternate spectral thresholds to better delineate the correct grassland extent in these areas. The resultant frame-specific outputs were then merged into a composite geographical mask and integrated back into the original countywide spectral class dataset to improve the overall grassland delineation.

8. LAND-COVER INFORMATION GENERATION

Multi-seasonal Landsat imagery has been used to generate a comprehensive set of spectrally based outputs that describe the seasonal characteristics of tree, bush, grass, bare and water cover characteristics within each image pixel, as represented by the 51 spectral foundation classes. These were converted into more meaningful and informative land-cover and land-use information classes, aligned with the South African National Standard (SANS) 1877 South African National Land-cover Classification Standard (see Section 4.1). For example, a pixel may be described as having detectable tree cover on more than one image-acquisition date, but not all dates, and grass cover for several, but not all dates. This would then be interpreted as an open, deciduous tree-covered woody community (i.e. woodland), where the grass cover became evident when the tree canopy showed less than maximum foliage cover.

8.1 Conversion from spectral to information classes

The creation of both land-cover and land-use information classes involves the recoding and re-grouping of selected spectral foundation classes, either in a controlled or uncontrolled geographic area. In a controlled recoding, independently sourced and/or generated geographical masks are used to define the location and extent of the recode process. This is required because, in many situations, different landscape features can be represented by similar spectral characteristics. For example, a spectrally modelled area of permanent bare ground (spectral characteristics are representative of a non-vegetated surface in all image-acquisition dates) could be representative of a beach, a mine or a large building.

By using an appropriate geographical mask (essentially an independently defined or sourced polygon coverage), it is possible to code all bare ground pixels within the mask to, for example, mining, while allowing all other bare areas outside the mask to represent other bare ground surfaces, which can then be further recoded as and where required.

The key to this process is that the spectral modelling defines the exact footprint of the spectral cover class (in a user-independent manner), whereas the geographical mask defines the information content (based on user expertise and knowledge). This approach overcomes problems where spectral characteristics are shared between two or more separate land-cover/land-use classes, which is self-evident in many image-based mapping applications. Figure 6 illustrates the various procedural pathways that can be used to convert the spectral foundation classes into land-cover/land-use information classes.

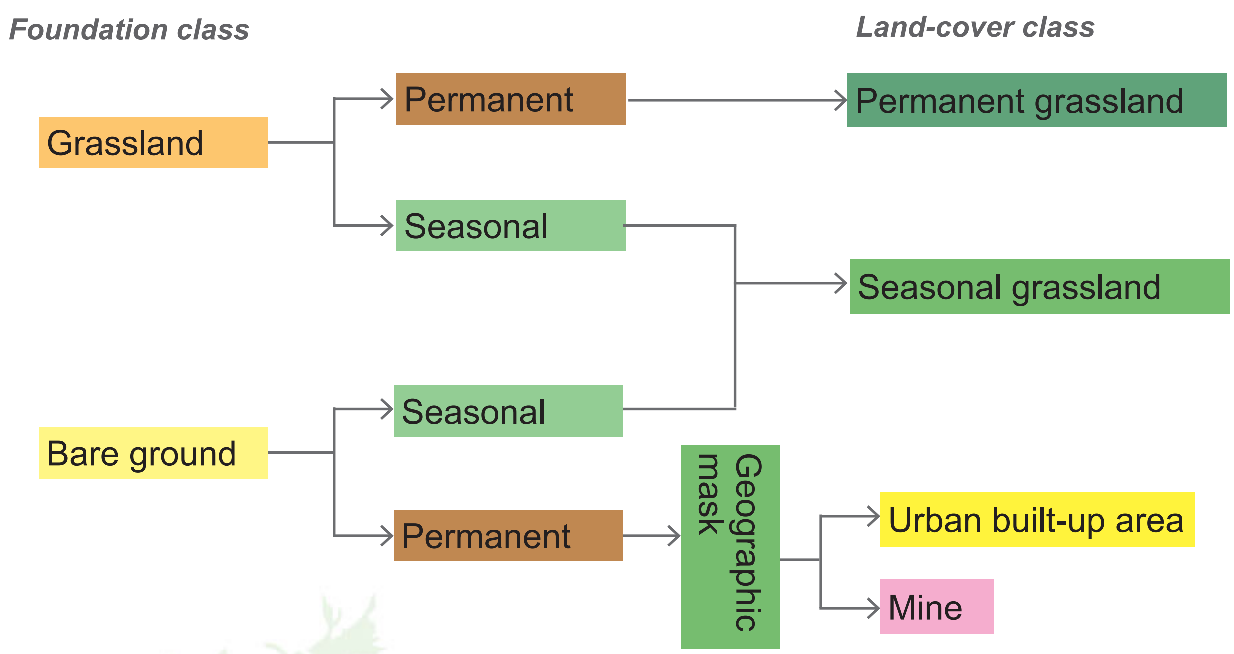


Figure 6. Processing pathways for converting spectral foundation classes to land-cover/land-use information classes, based on seasonal permanence and shared spectral characteristics.

8.2 Creating base vegetation land-cover classes

The natural vegetation base classes, namely thicket/dense bush, woodland/open bush, grassland, low shrubland and bare, non-vegetated areas, were created by merging selected groups of spectral classes. Generally, selected spectral classes were merged into the required vegetation cover classes without geographical control, using the original spectral class extent and distribution “as-is”. However, in some regions, and for some spectral classes, it was necessary for the recoding and merging process to be geographically controlled, and only allow the recode procedure within predetermined geographical regions. Controlled geographical recoding (i.e. masking) used pre-selected SANBI bioregion[9] boundaries to define where and how selected spectral classes were modified. Note that the final vegetative land-cover class boundary is always defined by the original spectral foundation class boundaries contained in the SANBI bioregion boundary, and not by the bioregion boundary itself, other than for the fynbos: low shrubland subclass.

9 Mucina, L and Rutherford, MC, 2006. *The vegetation of South Africa, Lesotho and Swaziland. Strelitzia* 19. SANBI: associated digital GIS database. Note that SANBI biome and bioregion boundaries, while generically applicable for 1:250 000 digital map scales, are often considerably more accurate on a local scale, even down to the nearest 100 m in some cases.

8.2.1 Fynbos

The sensitivity of the automated spectral models was insufficient to determine the transition between structurally similar fynbos: low shrubland communities and non-fynbos: low shrubland communities, except for the boundary with the structurally different Namaqualand (SKn and SKs) and Knersvlakte (Skk) bioregions. In these areas, the boundary between the fynbos and non-fynbos: low shrubland communities has been defined on the basis of the original spectral class boundaries[10].

8.2.2 Indigenous forests

The indigenous forest land-cover class was generated in a similar manner to the previously described base vegetation classes, although the geographical masks used to control the recoding of the selected spectral foundation classes were derived from a combination of independent datasets. These included the SANBI forest biome boundaries, higher detail regional masks created in-house in support of previous provincial-level land-cover mapping, shadow and altitude terrain parameters modelled from the higher 30 m-resolution SRTM DEM terrain data[11] and NDVI spectral datasets. The NDVI dataset represented a nationwide summary of the seasonal maximum NDVI recorded in each pixel, calculated from all image-acquisition dates used in the spectral land-cover modelling. Note that separate models were required to recode the different forest vegetation types (i.e. Northern Afromontane, sand, scarp, etc.), with each one based on different geographical mask combinations and associated data thresholds in order to correctly model the forest patch distributions in these regions. The final modelled forest extents were then integrated into the base vegetation land-cover classes.

10 This is primarily because satellite imagery such as Landsat, with its combination of medium spatial resolution and limited spectral capabilities (compared to hyper-spectral sensors), is primarily responsive to structural vegetation characteristics rather than community floristics. Accurate delineation of biome-related boundaries, especially if the boundary is a transition zone, rather than a distinct border, is – at best – challenging with imagery such as Landsat, especially if floristic differences play a greater role than structural differences in defining the border. This is likely to be the case between the fynbos, succulent Karoo, Nama Karoo and surrounding grassland and desert biomes.

11 The 30-m resolution SRTM terrain data for Africa was available for use at this time within the 2013/14 South African National Land-cover Project, despite not having been available for the initial spectral modelling phase.

8.2.3 Wetlands

The wetland class was created using independent and separate modelling routines that did not include reference to or recoding of any of the original 51 spectral foundation classes. Depending on the available image-acquisition dates per frame, wetland extent was modelled using either a dual or single image-acquisition-date process. The single-date approach used the best "wettest" image date, whereas the dual-date approach was based on the difference between wettest and driest image dates, where the single (wettest)-date modelling was only used as an alternative approach if the available acquisition dates did not provide suitable seasonal differences for the dual-date, wet-dry-date approach. Both modelling approaches only used inputs from multiple spectral indices, and did not use the 51 spectral foundation cover classes as used in the generation of the other vegetation class. A combination of eight indices were used in the dual-date approach, with each being applied to both dates, whereas only four spectral indices were used in the single (wettest)-date approach. Both modelling approaches generated a dataset that represented the likelihood of wetlands occurring in the landscape, within which a threshold was determined on a frame-by-frame basis in order to best represent actual wetland extent.

The single (wettest)-date modelling approach used terrain shadow masking to minimise spectral confusion with non-wetland regions, whereas the dual-date approach used more complex slope and floodplain modelling to limit the topographic areas within which wetlands were likely to occur. Both approaches used the 90 m SRTM terrain data as the source for topographic modelling.

The slope and floodplain model generated potential flooding zones, based on potential flood heights, which were taken to be indicative of areas most likely to be wet, and thus to contain the highest probability of having wetlands. The flood heights were created by estimating the water expansion level in river channels against a modelled stream surface height layer (in comparison to actual terrain surface). The modelling provides an indication of potential lateral water extent at a range of heights above the stream base height in relation to the stream channel profile at that location, irrespective of flood duration or time periods. The modelling does not take account of the surface roughness or flow blockage, which may result in the additional lateral extension of flood waters through backup, etc. The accuracy of the output is directly dependent on the scale, resolution and detail contained in the DEM. The modelled potential flooding zones were then used to restrict the dual-date wetland modelling where wetlands are most likely to occur.

Note that, in some localised areas, the results of both single- and dual-date wetland modelling approaches required post-modelling edits to improve the wetland delineation, especially if wetlands were not showing saturation or significant (vegetation) flushes. In such cases, analyst-assisted conventional unsupervised classifications were used to locally improve the modelled wetland output.

The modelled wetland extents were then integrated into the base vegetation land-cover classes.

8.2.4 Nama and succulent Karoo

The Nama and succulent Karoo categories and associated subclasses are included as a result of a specific DEA information requirement in the DEA/CARDNO land-cover data product. Since the sensitivity of the automated spectral models was insufficient to determine these biome boundaries, they have been defined solely on the SANBI biome boundaries, as per the previous fynbos explanation.

However, within these SANBI boundaries, the original level of modelled land-cover detail has been retained in terms of the forest, dense bush, open bush, low shrubland, grassland and bare ground subclasses. This is because each biome can contain a wide range of structural vegetation types, over and above the dominant structural form. For example, localised patches of thicket, tall bush and grasslike structural communities all occur within the general low shrub matrix that comprises the Nama Karoo biome, according to local species composition, terrain location and fire history.

8.3 Creating land-use classes

Land-use classes, such as cultivated lands, forest plantations, mines and settlements, were all derived as separate modelling procedures, typically using independently sourced and generated geographical masks to control where and how the original 51 seasonally defined spectral foundation classes were recoded and modified into the final land-use classes. The masking process also allowed independent, class-specific recoding to be applied, irrespective of spectral class data content, if required.

8.3.1 Cultivated lands

The Department of Agriculture, Forestry and Fisheries (DAFF) public domain 2013 national field boundary dataset (captured from 2.5 m-resolution, pan-merge SPOT5 imagery) was the initial source of the cultivated land-use classes. This existing dataset was updated nationally to represent the latest cultivated land patterns present on the latest 2014 Landsat 8 imagery used in each image frame. This typically involved capturing the distribution of new, commercially cultivated fields, especially centre pivot irrigation units, although all field types were included in the 2014 updating process, including subsistence cultivation areas, if clearly apparent.

In addition, the location, extent and distribution of all commercial pineapple and sugar cane crops were mapped from the same 2014 Landsat 8 imagery to increase the level of crop-specific subclass detail, in line with several of the previous provincial land-cover datasets generated by GTI.

All new cultivated lands (and crop types) mapped from the 2014 Landsat 8 imagery were captured manually using on-screen photo-interpretation techniques, since this facilitated and maintained similar interpretation accuracies to the source SPOT5-derived field boundary data. Furthermore, it also eliminated post-classification editing necessities typically associated with field boundary (as opposed to crop type) classification attempts.

The final cultivated land boundaries were used "as-is" to define the final geographical extent of all the cultivated land-use classes incorporated into the land-cover dataset. No use of or reference to the original 51 seasonally defined spectral foundation classes was made.

8.3.2 Plantations

Forestry plantations were derived from selected classes within the 51-class spectral foundation cover dataset, which were recoded and regrouped within controlled geographical areas, using independently sourced and generated geographical masks. The plantation area masks were sourced from several in-house provincial mapping projects, updated nationally to the plantation extent visible on the 2014 Landsat imagery.

Note that the final plantation class boundaries were always defined by the original spectral foundation class boundaries contained within the geographical masks, and not by the geographical mask boundaries themselves.

The geographical masks used to define the areas to be recoded for the clear-felled plantation subclass were only mapped off the latest 2014 Landsat imagery in order to ensure that the representation and interpretation was current.

8.3.3 Mines

Mines were derived from selected classes within the 51-class spectral foundation cover dataset, which were recoded and regrouped within controlled geographical areas, using independently sourced and generated geographical masks. The mine area masks were sourced from previous in-house, provincial mapping projects and national 1:50 000-scale map datasets[12], all of which were then updated nationally to the mine activity extent visible on the 2014 Landsat imagery. Note that mine-water subclasses were generated by identifying water classes that were located within the final mine geographical masks.

Note that major road and rail features were excluded from the mine area footprint if they intersected the mine mask.

Note that the final mine class boundaries were always defined by the original spectral foundation class boundaries contained within the geographical masks, and not by the geographical mask boundaries themselves.

8.3.4 Built-up areas/settlements

Built-up areas were generated independently of the spectral foundation cover dataset. The primary source for the cover class were several internally developed GTI urban map products and associated databases. These provided detailed information on the extent, distribution and land-use for all settlements nationally. This was verified in rural areas with the National Dwelling Frame Dataset, available from Statistics South Africa. This information was then spatially remodelled to represent built-up area outlines, which were then further updated and corrected nationally, using the 2014 Landsat imagery for visual reference. This was especially so in terms of rural village outlines.

The independently generated built-up area class boundaries were used “as-is” to define the final geographical extent of the settlement patterns incorporated into the land-cover dataset. No use of or reference to the original 51 seasonally defined spectral foundation classes was made.

8.3.5 Erosion

The erosion (donga) class was derived from selected classes (representative of bare ground) from within the 51-class spectral foundation cover dataset, which were recoded and regrouped within controlled geographical areas, using independently sourced and generated geographical masks. The erosion area masks were sourced from previous in-house, provincial mapping projects and other national erosion datasets, all of which were then updated nationally to represent the current extent of major dongas visible on the 2014 Landsat imagery. Note that, as a result of spectral modelling sensitivities and the need to be able to separate the bare ground within donga features from the surrounding non-eroded areas, the final modelled extent of erosion features is significantly better represented both spatially and numerically in the wetter, more vegetatively lush regions of the country, where the non-vegetated erosion surface is significantly different from the surrounding vegetation cover (i.e. bushveld and grassland regions). Donga feature detection in the drier, more arid region is not as accurate.

As with previous land-use classes, the final donga/erosion class boundaries were always defined by the original spectral foundation class boundaries contained within the geographical masks, and not by the geographical mask boundaries themselves.

12 Geographical masks from the national 1:50 000-scale topographic digital map data were created by extracting the relevant vector features and buffering them by one Landsat pixel extent. These were integrated with the in-house provincial mine masks, before updating them to the 2014 geographic extents off the 2014 Landsat 8 imagery.

8.3.5.1 Degraded areas

As part of the DEA-commissioned version of the 2014 GTI South African National Land-cover Dataset, an additional subclass defining “degraded” areas was also generated. This subclass is, however, not part of the standard land-cover data product. Degraded areas are defined in terms of this dataset as areas of significantly reduced vegetation cover compared to immediately adjacent pristine or semi-pristine natural areas. Degraded areas were modelled from selected classes from within the 51-class spectral foundation cover dataset, which were recoded and regrouped within controlled geographical areas, using independently generated geographical masks. The basic degraded area mask was created from image-derived, seasonally summarised NDVI maximum and standard deviation datasets. The extent of this mask allocation was limited to areas outside formally protected (conservation) areas[13], within non-arid biome regions[14], on terrain slopes less than 21 degrees[15], and not overlapping major roads, dry river beds, beaches and dune fields[16].

This modelling approach, while relevant in terms of practical, available input data, excluded the identification of cover-rich, species-poor degraded areas, as well as degraded low vegetation areas within naturally occurring low-vegetation regions. In the latter case, alternate modelling options were considered, based on buffered threshold distances around settlements and mines, but the generic outputs were not considered to be universally reliable or accurate.

As with previous land-use classes, the final degraded area boundaries were always defined by the original spectral foundation class boundaries contained within the geographical masks, and not by the geographical mask boundaries themselves.

13 As defined in the DEA Protected Areas 2014 Protected and Conservation Areas (PACA) Database.

14 To ensure that degraded areas were clearly identifiable as having a lower vegetation cover than the surrounding non-degraded areas, unless they were located within the boundaries of the former TBVC states (Transkei, Bophuthatswana, Venda and Ciskei) and closely associated with areas of dense rural settlements.

15 To exclude natural rocky slopes and cliff faces with similar non-vegetated spectral characteristics.

16 To exclude other land-cover and land-use features with similar non-vegetated spectral characteristics

9. ACCURACY ASSESSMENT

The 2013/14 South African National Land-cover Dataset has been verified in terms of mapping accuracy to provide a measure of end-user confidence in data use. The satellite image-generated land-cover/land-use information was verified visually, as part of a desktop-only procedure, against equivalent-date, high-resolution imagery and photography in Google Earth©. Accuracies are reported using industry-standard error (confusion) matrices, and include producer, user and kappa values.

9.1 Design and approach

Sample points for verification were selected using two separate approaches that considered the validity of statistical representation, the spatial resolution of Landsat imagery in relation to landscape features, the national reporting frame extent and the structure of the mapped land-cover information. Note that 33 land-cover/land-use classes were verified, in some cases representing primary rather than subclass land-cover or land-use characteristics. Where necessary, these were then amalgamated into the broader DEA/CARDNO land-cover/land-use classes for statistical reporting. For example, water, mines and plantations were verified at the primary level, since subclass detail was linked to short-term temporal effects that could not be verified on the single Google Earth© imagery.

The selection of samples representing potential commission errors was achieved by selecting about 150 samples for each of the 33 classes to be verified, given a total of approximately 3 500 points. Points were selected randomly across the full national land-cover dataset. Typically, about 100 samples were selected from thematic class units smaller than 100 ha and about 50 samples were selected from thematic class units greater than 100 ha in order to ensure a balanced representation of thematic unit areas. In all cases, the actual sample point represented the centre of the selected thematic unit, with placement set to exclude areas within 50 m of the thematic unit boundary to minimise class-edge effects.

The selection of samples representing potential omission errors was achieved by selecting approximately 2 500 samples without reference to class type initially with a stratified approach using the image frame extents, and then randomly, within the selected frames, across the original spectral foundation class dataset (as opposed to the final land-cover dataset). The boundaries of the spectral foundation classes were just used as a sampling frame. Points represented the centre of the selected thematic spectral unit, with placement set to exclude areas within 50 m of the thematic unit boundary to minimise class-edge effects. This approach was deemed appropriate since it ensured that the sample points at least reflected the spatial variability of the multi-date input imagery, which in turn reflected the seasonal landscape characteristics. Approximately 1 000 samples were selected from thematic class units between 100 and 200 ha in size, 1 000 from thematic class units between 200 and 500 ha, 400 from thematic class units between 500 and 1 000 ha, and 100 from thematic class units greater than 1 000 ha in order to ensure a balanced representation of thematic unit areas.

Note that the results presented below represent the DEA/CARDNO land-cover legend content, and not the full land-cover legend from which it has been derived.

Note that no map accuracy statistics are provided at the biome level, other than for the fynbos: low shrubland class, since the biome-defined boundaries have been derived from fixed, independent data and simply overlaid on the spectrally modelled structural vegetation categories. The fynbos: low shrubland class is retained within the accuracy assessment since its delineation has, in part, been defined on modelled spectral characteristics.

Note that no map accuracy is provided for the degraded class due to the difficulty of reliably determining – visually – that the vegetation at a specific sample point is significantly lower than that in surrounding, undisturbed natural vegetation areas, without in-field observation over wider areas. It is suggested that the calculated accuracy for bare ground and erosion is taken as indicative of the degraded mapping accuracy, since the degraded class is a modelled derivative of these classes.

Figure 7 illustrates the number and distribution of sample points across the full country extent.

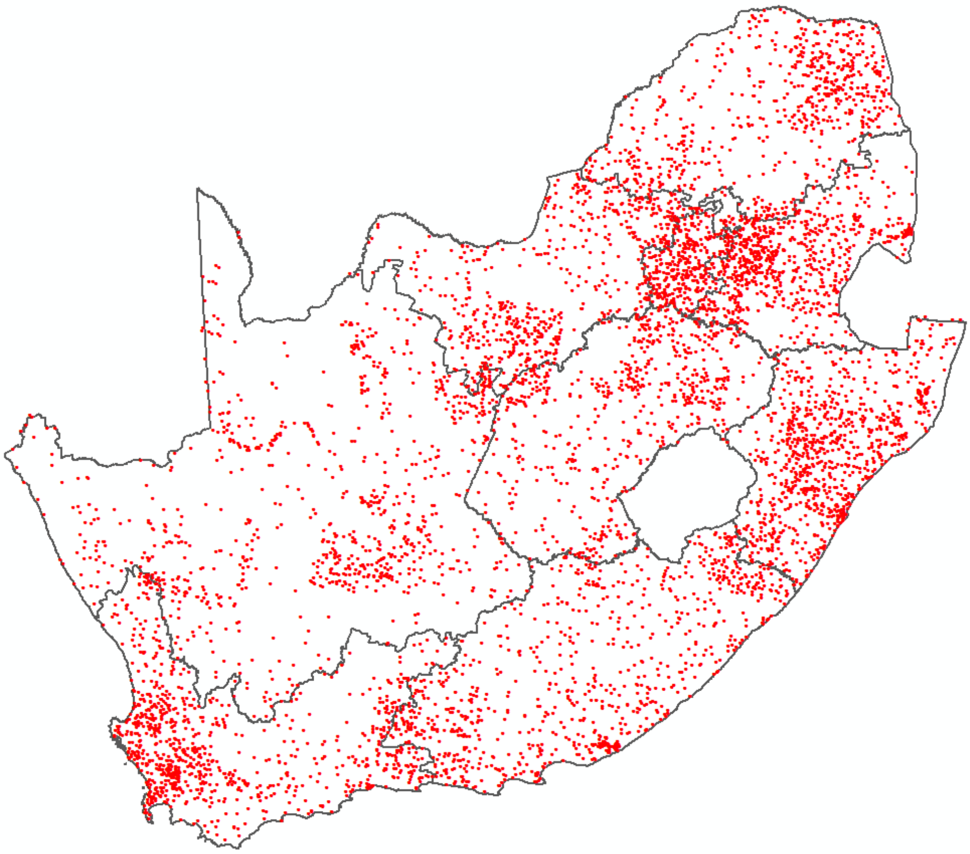


Figure 7.Number and distribution of 6 415 sample points across the full country extent.

9.2 Results

The map accuracy results for the 2013/14 South African National Land-cover Dataset, modelled from multi-seasonal Landsat 8 imagery, based on the modified-legend format representing the DEA/CARDNO information requirements, are as follows:

OVERALL SUMMARY		
Overall Map Accuracy	82.53	%
Mean Class Accuracy	88.36	%
90 % confidence limits	81.91	83.14
	low	high
Kappa Index	80.87	
Number of classes present (enter)		17
Number of sample sites		6415

The overall map accuracy for the 2013/14 South African National Land-cover Dataset in the DEA/CARDNO format, modelled from multi-seasonal Landsat 8 imagery, is 82.53%, with a mean land-cover/land-use class accuracy of 88.36%. This has been determined from 6 415 sample points representing 17 DEA/CARDNO land-cover/land-use classes. The Kappa Index value of 80.87 indicates that these results are very unlikely to be the result of chance occurrence.

A breakdown of individual class accuracies is provided below:

CLASS SUMARY STATISTICS			User Acc %	Prod Acc %	90% C.L low	high	Omm Error	Comm Error
LAND-COVER CLASS	Indignenous Forest	1	72.60	94.64	91.78	97.50	0.05	0.27
	Dense Bush, Thicket	2	53.74	83.64	80.72	86.55	0.16	0.46
	Woodland, Open Bush	3	60.84	54.13	51.27	57.00	0.46	0.39
	Low Shrub: Other	4	70.59	61.82	59.37	64.27	0.38	0.29
	Forest Plantation	5	89.30	94.35	92.31	96.39	0.06	0.11
	Cultivated Commercial Crop	6	91.91	99.54	98.69	100.00	0.00	0.08
	Cultivated Pivot Crop	7	95.38	92.42	91.18	93.66	0.08	0.05
	Cultivated Orchards	8	92.18	95.29	93.62	96.95	0.05	0.08
	Cultivated Vineyard	9	91.61	97.26	95.37	99.15	0.03	0.08
	Cultivated Subsistence	10	89.00	94.42	92.22	96.61	0.06	0.11
	Settlements	11	93.90	98.68	97.93	99.43	0.01	0.06
	Wetlands	12	88.07	91.18	88.30	94.05	0.09	0.12
	Grasslands	13	84.56	69.82	68.11	71.53	0.30	0.15
	Fynbos: Low Shrubland	17	79.64	93.31	91.15	95.46	0.07	0.20
	Water	33	98.77	97.58	95.89	99.27	0.02	0.01
	Bare Ground	34	78.69	85.92	83.86	87.97	0.14	0.21
		Total		82.53	81.91	83.14		

Note: Class numbers refer to the class numbers in the digital land-cover data, which includes additional subclass information relating to the fynbos, Nama and succulent Karoo biome groupings.

These results exceed the requirements as defined in the terms of reference, that "a minimum map accuracy of 70% is to be achieved at the 90% confidence limits, with a minimum 70% kappa index value ... with a minimum of 30 samples per class ... (and) a target of 1 000 samples in total".

A significant majority of individual land-cover/land-use classes achieved user and producer accuracies of 70% or more (see yellow highlights). Fourteen out of 17 (82%) classes achieved a user accuracy greater than or equal to 70%. Thirteen out of 17 (76%) classes achieved a producer accuracy greater than or equal to 70%.

Producer accuracy (omission error) represents how well reference pixels of the true land cover are classified. User accuracy (commission error) represents the probability that a pixel classified in a given category actually represents that category on the ground. Producer accuracy therefore represents the accuracy of the sample points, whereas user accuracy represents what the data user will experience when using the land-cover data. The Kappa Index is a measure of statistical chance, with higher values (greater than 0.7) typically representing repeatable and reliable results.

The contingency matrix showing the per-cover class breakdown of these map accuracy statistics is shown in Appendix E.

9.3 Analysis and conclusions

Overall, the map accuracies for the 2013/14 South African National Land-cover Dataset are very good and indicate that the dataset is accurate and that data users can have confidence in the information contained in it. There are, however, some individual classes with less than 80% user and/or producer accuracy that need further analysis and comment. All of these, bar one, are associated with spectrally modelled base vegetation classes, rather than the post-modelling-derived land-use classes.

9.3.1 Dense bush/thicket

The dense bush/thicket class achieved a 83.64% producer accuracy, but only a 53.86% user accuracy. This appears to be primarily the result of confusion with woodland/open bush (76 out of 428 samples), grassland (51 out of 428 samples) and low shrubland[17] (45 out of 428 samples). There is no clear regional pattern associated with this misclassification.

Within the woodland/open bush confused samples, 52 of the 76 samples had a woody cover greater than or equal to 55%, and 26 of the 76 samples had a woody cover greater than or equal to 65%, which may explain this inter-class confusion, since the (Google Earth©) observed canopy cover amounts could be considered close to the transition to dense bush cover. The confusion with observed grassland areas is difficult to explain since very little woody cover is observed in these samples. A possible explanation may be that localised areas of extremely high grassland biomass are (spectrally) confused in the modelling process with woody cover. The dense bush/thicket confusion with low shrubland may be associated, in some cases, with the tall woody cover density of low shrubland[18], since 24 out of 45 of the incorrect sample sites have an observed woody cover greater than or equal to 15%, 15 out of 45 have an observed woody cover greater than or equal to 25%, nine out of 45 have an observed woody cover greater than or equal to 35% and five out of 45 have an observed woody cover greater than or equal to 45%.

If all the dense bush/thicket samples are grouped into tall woody versus non-woody (or low woody shrub) cover[19], then 349 out of 428 (81%) samples have a tall woody canopy cover greater than or equal to 15%, regardless of observed land-cover type, 331 out of 428 (77%) samples have a tall woody canopy cover greater than or equal to 25%, 323 out of 428 (75%) samples have a tall woody canopy cover greater than or equal to 35%, 310 out of 428 (72%) samples have a tall woody canopy cover greater than or equal to 45%, and 197 out of 428 (46%) samples have a tall woody canopy cover greater than or equal to 65%.

9.3.2 Woodland/open bushland

The woodland/open bushland class achieved a 60.84% user accuracy and a 54.13% producer accuracy. This appears to be primarily the result of confusion with grassland (59 out of 452 samples) and low shrubland[20] (80 out of 452 samples). There is no clear regional pattern associated with this misclassification.

17 Low shrubland fynbos and other subclasses combined.

18 The observed woody cover percentage for low shrubland (fynbos and other) is indicated in the accuracy spreadsheet, and represents the tall woody cover and not the general low matrix of woody shrubs.

19 Note that forest plantations do not have a woody canopy cover value recorded for either observed or mapped plantation sample sites, so it is likely that the accuracy of the modelled extent of dense woody cover is actually higher, since there are eight dense bush/thicket sample points that were observed to be plantation forests.

20 Low shrubland fynbos and other subclasses combined.

The confusion with grassland is difficult to explain, other than that it is the result of modelling and interpretation errors associated with lower woody canopy cover densities and associated spectral characteristics, with 14 out of 59 (23%) of the observed grassland samples having a woody cover of greater than or equal to 15%. Similar confusion was found with low shrubland, with 17 out of 80 (21%) of the observed low shrubland samples having a tall woody cover of greater than or equal to 15%, irrespective of the background low woody shrub matrix.

If all the woodland/open bushland samples are grouped into tall woody versus non-woody (or low woody shrub) cover, then 320 out of 452 (70%) samples have a tall woody canopy cover greater than or equal to 15%, regardless of observed land-cover type, 289 out of 452 (63%) samples have a tall woody canopy cover greater than or equal to 25%, 284 out of 452 (62%) samples have a tall woody canopy cover greater than or equal to 35%, 221 out of 452 (48%) samples have a tall woody canopy cover greater than or equal to 45%, and 80 out of 452 (17%) samples have a tall woody canopy cover greater than or equal to 65%.

9.3.3 Grassland

The grassland class achieved a 84.56% user accuracy, but only a 69.82% producer accuracy. This appears to be primarily the result of confusion with woodland/open bush (82 out of 1 004 samples) and low shrubland[21] (55 out of 1 004 samples). There is no clear regional pattern associated with this misclassification.

The confusion with woodland/open bush is difficult to explain, other than that it is the result of modelling and interpretation errors, since the majority of observed woodland/open bush sample sites all have tall woody cover densities greater than or equal to 55%. However, on the majority of observed low shrubland sites, the woody cover density was less than 5%, irrespective of the background low woody shrub matrix, which is probably representative of a transitional zone.

If all the grassland samples are grouped into tall woody versus non-woody cover, then 813 out of 1 004 (80%) samples have a tall woody canopy cover less than or equal to 15%, regardless of observed land-cover type, and 711 out of 1 004 (70%) samples have a tall woody canopy cover less than or equal to 5%, regardless of observed land-cover type.

9.3.4 Low shrubland (fynbos and other subclasses combined)

The low shrubland classes for fynbos and other respectively achieved user accuracies of 79.64% and 70.59%, and producer accuracies of 93.31% and 61.82%. These two classes have been assessed in combination, since they are only differentiated on the basis of the overlaid SANBI biome boundary. In both cases, confusion was primarily with grassland (142 out of 858 samples). There is no clear regional pattern associated with this misclassification.

If all the low shrubland samples are grouped into (tall) woody versus non-woody cover, then 791 out of 858 (92%) samples have a tall woody canopy cover less than 15%, regardless of observed land-cover type, and 691 out of 858 (80%) samples have a tall woody canopy cover less than 5%, regardless of observed land-cover type.

9.3.5 Bare ground

The bare ground class achieved a user accuracy of 73.54% and a producer accuracy of 77.54%. This appears to be primarily the result of confusion with grassland (26 out of 378 samples) and low shrubland[22] (65 out of 378 samples). There is no clear regional pattern associated with this misclassification. The confusion with low shrubland is difficult to explain, other than that it is the result of modelling and interpretation errors, since 57 of the 65 observed low shrubland sample sites all have tall woody cover densities between 5 and 55%, with the majority being between 15 and 25%. However, these percentages, as indicated, only refer to the tall woody cover component, and not to the low woody shrub component, which may have been very sparse. The confusion with grassland may be due to a comparable, overall low vegetation base cover, although unlike with the observed low shrubland samples, the observed grassland samples all had, bar one, 0% tall woody cover densities.

9.3.6 Conclusion

In conclusion, it would seem that the majority of mapping errors associated with the base natural

21 Low shrubland fynbos and other subclasses combined.

22 Low shrubland fynbos and other subclasses combined.

vegetation cover classes can be broadly associated with transitional vegetation conditions, represented by structural vegetation gradients, within which the spectral-based modelling has defined a fixed boundary.

10. DATA USE

The digital raster data is supplied without any post-modelling or classification spatial filtering. Data users may wish to consider applying spatial cleaning techniques, such as applying a moving 3 x 3 pixel window filter to remove isolated single-class pixels from the dataset, especially if the raster data is to be converted to vector format.

11. METADATA

2013/14 GTI South African National Land-cover Dataset DEA/CARDNO 35-class Legend: Core metadata elements (SANS 1878)

1(M) Dataset title: 2013/14 GTI SA National Land-cover (DEA/CARDNO version) (*dea_cardno_2014_sa_lcov_utm35n_vs2b_pivot-corr.img*)

2(M) Dataset reference date: April 2013 to March 2014

3(O) Dataset responsible party: Produced by GeoTerra Image (GTI) (Pty) Ltd, South Africa

4(C) Geographic location of the dataset. MBR
WestBoundLongitude: -717294.00 (Upper Left X)
EastBoundLongitude: 1301256.00 (Lower Right X)
NorthBoundLongitude: -2239230.00 (Upper Left Y)
SouthBoundLongitude: -4046670.00 (Lower Right Y)

Projection coordinates based on Universal Transverse Mercator (UTM) 35 north, WGS84 (datum), metres.

5(M) Dataset language: “English” (eng)

6(C) Dataset character set: UTF8 (8-bit data)

7(M)Dataset topic category: 010 = Base map earth coverage

8(O) Scale of the dataset: Land cover mapped from 30 m-resolution Landsat satellite imagery, therefore recommended for approximately 1:75 000 to 1:90 000 scale or coarse mapping and modelling applications.

9(M) Abstract describing the dataset: The 2013/14 South African National Land-cover Dataset produced by GTI as a commercial data product has been generated from digital, multi-seasonal Landsat 8 multispectral imagery, acquired between April 2013 and March 2014. In excess of 600 Landsat images were used to generate the land-cover information, based on an average of eight different seasonal image-acquisition dates, within each of the 76 image frames required to cover South Africa. The land-cover dataset, which covers the whole of South Africa, is presented in a map-corrected, raster format, based on 30 x 30 m cells equivalent to the image resolution of the source Landsat 8 multi-spectral imagery. This specific version of the 2013/14 GTI South African National Land-cover Dataset is supplied to the South African Department of Environmental Affairs under the DEA/CARDNO Project: SCPF002: *Implementation of Land-use Maps for South Africa*. The supplied dataset referred to in this report represents a subset of the full 2013/14 GTI South African National Land-cover Dataset, in terms of supplied land-cover/land-use categories. Use of this DEA/CARDNO data version is governed by the LICENCE AGREEMENT contained in this report. The DEA/CARDNO dataset contains 35 land-cover/land-use information classes, covering a wide range of natural and man-made landscape characteristics. The original land-cover dataset was processed in UTM35 (north)/ WGS84 map-projection format based on the Landsat 8 standard map-projection format as provided by the USGS. The data remains the property of GTI, and is protected by copyright laws. All intellectual property rights pertaining to the data remain with GTI at all times.

This updated 2013/14 dataset corrects the coding error for DEA/CARDNO classes 6 and 7 (cultivated fields and pivot cultivated fields), which were inadvertently swapped in the previous version. Apart from the recoding of the two field-type classes, all other details and class codes remain exactly the same as the previously delivered datasets.

10(O) Dataset format name: ERDAS Imagine© *img raster formats

11(O) Dataset format version: version 01 (file#22)

12(O) Additional extent information for the dataset: (vertical and temporal)

Vertical extent:

Minimum value: N/A
Maximum value: N/A
Unit of measure: N/A
Vertical datum: N/A

Temporal extent: Land-cover datasets generated in January 2015, based on April 2013 to March 2014 multi-seasonal Landsat 8 and Landsat 5 satellite imagery.

14(O) Reference system: Universal Transverse Mercator (UTM) 35 north

CRS:

Projection used: Universal Transverse Mercator (UTM) 35 north
Spheroid used: WGS84
Datum used: WGS 84

Ellipsoid parameters:

Ellipsoid semi-major axis
Axis units
Denominator of flattening ratio

Projection parameters:

UTM Zone: 35 (north)
Standard parallel
Longitude of central meridian: 27:00:00.00 east
Latitude of projection origin: 00:00:00.00 east
False easting: 500 000.00 metres
False northing: 0.00 metres
Scale factor at equator: 0.999600
Projection units: metres

15(O) Lineage statement: Land-cover dataset generated in-house by GTI (Pretoria) in January 2015, based primarily on multi-date Landsat 8 imagery acquired between April 2013 and March 2014.

16(O) Online resource: N/A

17(O) Metadata file identifier: N/A

18(O) Metadata standard name: SANS I878

19(O) Metadata standard version: version 01

20(C) Metadata language: English (eng)

21(C) Metadata character set: 021 (UsAscii)

22(M) Metadata point of contact:

Name:	Mark Thompson
Position name:	Director Remote Sensing
Organisation name:	GeoTerraImage (Pty) Ltd

Physical address:

Building	Grain Building (1st Floor)
Street	Witherite
Street Suffix	Street
Street No	477
Suburb	Die Wilgers
City	Pretoria
Zip code	0041
State	Gauteng
Country	South Africa

Postal address:

Box	295
Suburb	Persequor Park
City	Pretoria
Zip code	0020
State	Gauteng
Country	South Africa

23(M) Metadata time stamp: 31 July 2015

APPENDIX A: The 2013/14 South African National Land-cover Legend (DEA/CARDNO legend)

The table below describes the 35 land-cover/land-use classes contained in the DEA/CARDNO-supplied subset of the 2013/14 South African National Land-cover Dataset, generated from multi-seasonal Landsat 8 imagery.
See Appendix D for visual examples of percentage cover densities per unit area.

Parent	DEA class name (and digital code)	Definition
Water	Water (33)	Areas of open, surface water that are detectable on all image dates used in the Landsat 8-based water modelling processes. The mapped extent represents the maximum detectable water extent from all available imagery acquired within the 2013/14 assessment period. Includes both natural and man-made water features.
Wetland	Wetland (12)	Wetland areas that are primarily vegetated on a seasonal or permanent basis. Defined on the basis of seasonal image-identifiable surface vegetation patterns (not subsurface soil characteristics). The vegetation can be either rooted or floating. Wetlands may be either daily (i.e. coastal), temporarily seasonal or permanently wet and/or saturated. Vegetation is predominately herbaceous. Includes, but is not limited to wetlands associated with seeps/springs, marshes, floodplains, lakes/pans, swamps, estuaries and some riparian areas. Wetlands associated with riparian zones represent image-identified vegetation along the edges of watercourses that show similar spectral characteristics to nearby wetland vegetation. Excludes mangrove swamps. Permanent or seasonal open-water areas within the wetlands are classified separately. Seasonal wetland occurrences within commercially cultivated field boundaries are not shown, although they have been retained within subsistence-level cultivation fields.
Forest	Indigenous forest (1)	Natural/semi-natural indigenous forest, dominated by tall trees, where tree canopy heights are typically greater than 5 m and tree canopy densities are typically more than 75%, often with multiple understory vegetation canopies. Note this class refers to grassland areas not within the fynbos, Nama Karoo and succulent Karoo biome boundaries (as defined in the SANBI 2006 *Vegetation of South Africa, Swaziland and Lesotho* map data).
Thicket and dense bush	Dense bush, thicket and tall dense shrubs (2)	Natural/semi-natural tree and/or bush-dominated areas, where typically canopy heights are between 2 and 5 m, and canopy density is typically more than 75%, but may include localised, sparser areas down to about 60%. Includes dense bush, thicket, closed woodland, tall, dense shrubs, scrub forest and mangrove swamps. Can include self-seeded bush encroachment areas if there is sufficient canopy density. Note: This class refers to grassland areas not within the fynbos, Nama Karoo and succulent Karoo biome boundaries (as defined in the SANBI 2006 *Vegetation of South Africa, Swaziland and Lesotho* map data).
Woodland/ open bush	Woodland and open bushland (3)	Natural/semi-natural tree and/or bush-dominated areas, where typically canopy heights are between 2 and 5 m, and canopy densities are typically between 40 and 75%, but may include localised, sparser areas down to 15 to 20%. (*Note: Normally, it is preferred that land-cover class definitions are mutually exclusive in terms of content and associated landscape characteristics. Due to the nature of the multi-seasonal spectral modelling approach used in compiling the land-cover dataset, it seems that there is some transitional overlap between "bushy/woody grassland" and "grassy woodland/ bushland" in terms of woody cover densities. For this reason, the class definitions for both "woodland/open bush" and "grassland" include both a class-specific core woody cover definition and a shared transitional woody cover definition. The sample points used to verify the map accuracy record the observed woody cover percentage for these vegetation types should the user wish to undertake further analysis.*)

Parent	DEA class name (and digital code)	Definition
Woodland/ open bush	Woodland and open bushland (3)	Includes sparse open bushland and woodland, including transitional wooded grassland areas. Can include self-seeded bush encroachment areas if canopy density is within the indicated range. In the arid western regions (i.e. Northern Cape), this cover class may be associated with a transitional bush/shrub cover that is lower than typical open bush/woodland cover, but higher and/or more dense than typical low shrub cover. Note: This class refers to grassland areas not within the fynbos, Nama Karoo and succulent Karoo biome boundaries (as defined in the SANBI 2006 *Vegetation of South Africa, Swaziland and Lesotho* map data).
Grassland	Grassland (13)	Natural/semi-natural grass-dominated areas, where the tree and/or bush canopy densities are typically less than 20%, but may include localised denser areas up to 40% (regardless of canopy heights).[7] Includes open grassland and sparse bushland and woodland areas, including transitional wooded grasslands. May include planted pasture (i.e. grazing) if not irrigated. Irrigated pastures will typically be classified as cultivated, and urban parks and golf courses under urban. Note: This class refers to grassland areas not within the fynbos, Nama Karoo and succulent Karoo biome boundaries (as defined in the SANBI 2006 *Vegetation of South Africa, Swaziland and Lesotho* map data).
Low shrubland	Low shrubland (4)	Natural/semi-natural low shrub-dominated areas, typically with a canopy height up to 2 m. Includes a range of canopy densities encompassing sparse to dense canopy covers. Very sparse covers may be associated with the bare ground class. Typically associated with low, woody shrub, Karoo-type vegetation communities, although can also represent locally degraded vegetation areas where there is a significantly reduced vegetation cover in comparison to surrounding, less impacted vegetation cover, including long-term wildfire scars in some mountainous areas in the Western Cape. Note that taller tree/bush/shrub communities within this vegetation type are typically classified separately as one of the other tree- or bush-dominated cover classes. Note: This class refers to low shrubland areas not within the fynbos, Nama Karoo and succulent Karoo biome boundaries (as defined in the SANBI 2006 *Vegetation of South Africa, Swaziland and Lesotho* map data).
Cultivated	Commercial annuals (rain-fed) (6)	Cultivated lands used primarily for the production of rain-fed, annual crops for commercial markets. Typically represented by large field units, often in dense local or regional clusters. In most cases, the defined cultivated extent represents the actual cultivated or potential extent. Includes sugarcane crops.
	Commercial pivot (7)	Cultivated lands used primarily for the production of centre pivot-irrigated, annual crops for commercial markets. In most cases, the defined cultivated extent represents the actual cultivated or potential extent. Includes sugarcane crops.
	Commercial permanent (orchards) (8)	Cultivated lands used primarily for the production of both rain-fed and irrigated permanent orchard crops for commercial markets. Includes both tree, shrub and non-woody crops, such as citrus, tea, coffee, grapes, lavender and pineapples. In most cases, the defined cultivated extent represents the actual cultivated or potential extent.
	Commercial permanent (vines) (9)	Cultivated lands used primarily for the production of both rain-fed and irrigated permanent vine (grape) crops for commercial markets. In most cases the defined cultivated extent represents the actual cultivated or potential extent.
	Subsistence (10)	Cultivated lands used primarily for the production of rain-fed, annual crops for local markets and/or home use. Typically represented by small field units, often in dense local or regional clusters. The defined area may include intra-field areas of non-cultivated land, which may be degraded or use-impacted if the individual field units are too small to be defined as separate features.

Parent	DEA class name (and digital code)	Definition
Forest plantation	Forest plantations (5)	Planted forestry plantations used for growing commercial timber tree species. The single class represents a combination of mature, young and temporary clear-felled stands. The class includes spatially smaller woodlots and windbreaks with the same cover characteristics. Note that young saplings are very difficult to identify on 30 m-resolution Landsat imagery if the actual tree canopy cover density is below 30 to 40%, because the background cover, for example, grassland, then dominates the spectral characteristics in that pixel area.
Mine	Mine (32)	Mining activity footprint, which includes extraction pits, tailings, waste dumps, flooded pits and associated surface infrastructure, such as roads and buildings (unless otherwise indicated), for both active and abandoned mining activities. Class may include open-cast pits, sand mines, quarries and borrow pits.
Bare	Bare (non-vegetated) (34)	Bare, non-vegetated ground with little or very sparse vegetation cover (i.e. typically up to 5 to 10% vegetation cover), occurring as a result of either natural or man-induced processes. Includes, but is not limited to natural rock exposures, dry river beds, dry pans, coastal dunes and beaches, sand and rocky desert areas, very sparse low shrublands and grasslands, erosion areas and major road networks. May also include long-term wildfire scars in some mountainous areas in the Western Cape.
	Degraded (35)	Sparsely vegetated areas, occurring as a result of man-induced processes, which show significantly lower overall vegetation cover compared to surrounding, natural undisturbed areas.
Built-up	Settlements (11)	All built-up areas, represented as a single class, including, but not limited to commercial, industrial, heath, education, religion, transport and residential land uses, including both formal and informal structures, across a range of structural densities from high to low. Includes high- and low-density areas containing high-density buildings and other built-up structures mainly associated with non-residential, commercial, administrative, sport, health, religious or transport (i.e. train station) activities. Includes agricultural smallholdings on the urban periphery.
Fynbos	Fynbos: forest (14)	Forest areas as per class (1) within the SANBI fynbos biome boundary.
	Fynbos: thicket (15)	Dense bush and thicket areas as per class (2) within the SANBI fynbos biome boundary.
	Fynbos: open bush (16)	Woodland and open bush areas as per class (3) within the SANBI fynbos biome boundary.
	Fynbos: low shrub (17)	Low shrubland areas as per class (4) within the SANBI fynbos biome boundary.
	Fynbos: grassland (18)	Grassland areas as per class (13) within the SANBI fynbos biome boundary.
	Fynbos: bare ground (19)	Bare ground areas as per class (14) within the SANBI fynbos biome boundary.
Nama Karoo	Nama Karoo: forest (20)	Forest areas as per class (1) within the SANBI Nama Karoo biome boundary.
	Nama Karoo: thicket (21)	Dense bush and thicket areas as per class (2) within the SANBI Nama Karoo biome boundary.
	Nama Karoo: open bush (22)	Woodland and open bush areas as per class (3) within the SANBI Nama Karoo biome boundary.
	Nama Karoo: low shrub (23)	Low shrubland areas as per class (4) within the SANBI Nama Karoo biome boundary.
	Nama Karoo: grassland (24)	Grassland areas as per class (13) within the SANBI Nama Karoo biome boundary.
	Nama Karoo: bare ground (25)	Bare ground areas as per class (14) within the SANBI Nama Karoo biome boundary.

Parent	DEA class name (and digital code)	Definition
Succulent Karoo	Succulent Karoo: forest (26)	Forest areas as per class (1) within the SANBI succulent Karoo biome boundary.
	Succulent Karoo: thicket (27)	Dense bush and thicket areas as per class (2) within the SANBI succulent Karoo biome boundary.
	Succulent Karoo: open bush (28)	Woodland and open bush areas as per class (3) within the SANBI succulent Karoo biome boundary.
	Succulent Karoo: low shrub (29)	Low shrubland areas as per class (4) within the SANBI succulent Karoo biome boundary.
	Succulent Karoo: grassland (30)	Grassland areas as per class (13) within the SANBI succulent Karoo biome boundary.
	Succulent Karoo: bare ground (31)	Bare ground areas as per class (14) within the SANBI succulent Karoo biome boundary.

Note: the fynbos, Nama Karoo and succulent Karoo vegetation subclasses are provided with additional class codes representing the parent primary vegetation class in order for different levels of class detail to be used and accessed by data users. This is illustrated on the next page.

Dual legend system for the DEA CARDNO 2013/14 South African land-cover classes (statistics for pivot code corrected dataset)

- Class_Names (Color) represents the full 35-class land-cover/land-use legend (with biome-level information)
- Class_Names2 (Color2) represents the amalgamated 17-class land-cover/land-use legend (no biome-level information)
- Area refers to hectares (ha) and histogram to the number of 30 x 30 m image pixels.

dea_cardno_2014_sa_lcov_utm35n_vs2b_pivot-corr.img

Row	Color	Class_Names	Color2	Class_Names2	Area	Histogram
0					0	0
1		Indigenous Forest		Indigenous Forest	395720	4396888
2		Thicket /Dense bush		Thicket /Dense bush	7.09698e+006	78855335
3		Woodlan/Open bush		Woodlan/Open bush	1.09087e+007	121207366
4		Low shrubland		Low shrubland	1.80007e+007	200008268
5		Plantations / Woodlots		Plantations / Woodlots	1.8737e+006	20818902
6		Cultivated commercial annual crops non-pivot		Cultivated commercial annual crops non-pivot	1.06108e+007	117898196
7		Cultivated commercial annual crops pivot		Cultivated commercial annual crops pivot	782049	8689435
8		Cultivated commercial permanent orchards		Cultivated commercial permanent orchards	346950	3855005
9		Cultivated commercial permanent vines		Cultivated commercial permanent vines	188711	2096791
10		Cultivated subsistence crops		Cultivated subsistence crops	2.04053e+006	22672521
11		Settlements		Settlements	2.90828e+006	32314218
12		Wetlands		Wetlands	1.0259e+006	11398887
13		Grasslands		Grasslands	2.37573e+007	263970479
14		Fynbos: forest		Indigenous Forest	32724	363600
15		Fynbos: thicket		Thicket /Dense bush	691164	7679603
16		Fynbos: open bush		Woodlan/Open bush	307674	3418605
17		Fynbos: low shrub		Low shrubland	5.32864e+006	59207100
18		Fynbos: grassland		Grasslands	381413	4237918
19		Fynbos: bare ground		Bare Ground	210113	2334586
20		Nama Karoo: forest		Indigenous Forest	0	0
21		Nama Karoo: thicket		Thicket /Dense bush	328724	3652485
22		Nama Karoo: open bush		Woodlan/Open bush	540805	6008941
23		Nama Karoo: low shrub		Low shrubland	1.40578e+007	156197442
24		Nama Karoo: grassland		Grasslands	1.31982e+006	14664675
25		Nama Karoo: bare ground		Bare Ground	9.38285e+006	104253894
26		Succuelent Karoo: forest		Indigenous Forest	0	0
27		Succulent Karoo: thicket		Thicket /Dense bush	174801	1942236
28		Succulent Karoo: open bush		Woodlan/Open bush	677790	7530997
29		Succulent Karoo: low shrub		Low shrubland	4.44011e+006	49334527
30		Succulent Karoo: grassland		Grasslands	335397	3726633
31		Succulent Karoo: bare ground		Bare Ground	1.97062e+006	21895741
32		Mines		Mines	328973	3655254
33		Waterbodies		Waterbodies	2.04562e+006	22729092
34		Bare Ground		Bare Ground	1.49435e+006	16603927
35		Degraded		Degraded	944061	10489566

APPENDIX B: List of accompanying documents and files

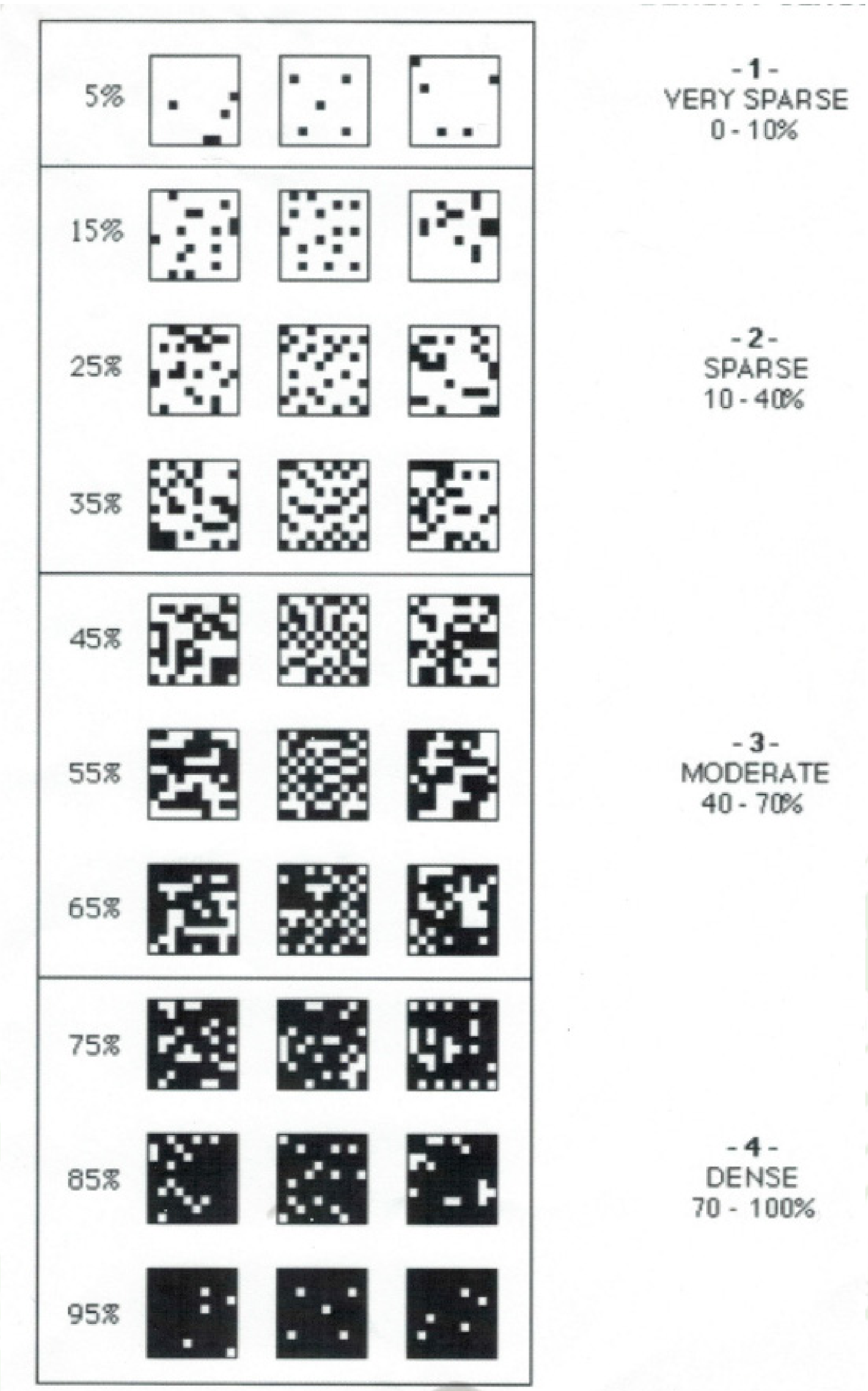
- An Excel spreadsheet containing the full list of Landsat 8 and Landsat 5 acquisition dates, per image frame, as used in the generation of the 2013/14 South African National Land-cover Dataset. Note that the listed acquisition dates represent the primary dates used in the land-cover modelling, and not any additional image dates that were only used for localised cloud masking on the primary image date. Additional image dates are, however, listed if they themselves were also considered primary acquisition dates and used as modelling inputs.

- ESRI ArcGIS point coverage (UTM35 north) representing the sample points used to verify the 2013/14 South African National Land-cover Dataset (version 4), and to calculate the DEA-CARDNO legend format mapping accuracies.

APPENDIX C: The 51-class Seasonally Defined Spectral Foundation Class Legend

Row	Color	Class_Names
0		
1		water seasonal
2		water permanent
3		bare mix 1 date only, no overlap with any pure bare class
4		bare mix >1 but < all dates, no overlap with any pure bare class
5		bare mix all dates, no overlap witrh any pure bare class
6		bare pure 1 date only
7		bare pure >1 but < all dates
8		bare pure all dates
9		Tree/Bush mix 1 date only
10		Tree 1 date only
11		Tree/Bush mix >1 but < all dates
12		Tree >1 but <all dates
13		Tree/Bush mix all dates
14		Tree all dates
15		golf-irrig combo Tree/Bush mix 1 date only
16		golf-irrig combo Tree 1 date only
17		golf-irrig combo Tree/Bush mix >1 but < all dates
18		golf-irrig combo Tree >1 but < all dates
19		golf-irrig combo Tree/Bush mix all dates
20		golf-irrig combo Tree all dates
21		Tree/Bush 1 date only + Bare pure 1 date only
22		Tree 1 date only + Bare pure 1 date only
23		Tree/Bush multi dates + Bare pure 1 date only
24		Tree multi dates + Bare pure 1 date only
25		Tree/Bush multi dates + Bare pure multi dates
26		Tree multi dates + Bare pure multi dates
27		Tree/Bush 1 date only + Bare pure multi dates
28		Tree 1 date only + Bare pure multi dates
29		golf-irrig combo / Tree/Bush mix 1 date only + Bare pure 1 date
30		golf-irrig combo / Tree 1 date only + Bare pure 1 date only
31		golf-irrig combo / Tree/Bush mix multi dates + Bare pure 1 date
32		golf-irrig combo / Tree multi dates + Bare pure 1 date only
33		golf-irrig combo / Tree/Bush mix multi dates + Bare pure multi dates
34		golf-irrig combo / Tree multi dates + Bare pure multi dates
35		golf-irrig combo / Tree/Bush mix 1 date only + Bare multi dates
36		golf-irrig combo / Tree 1 date only + Bare pure multi dates
37		Grass (all dates)
38		Grass (any date) / Bare pure 1 date
39		Grass (any date) / Bare mix 1 date
40		Grass (any date) / Bare pure multi dates
41		Grass (any date) / Bare mix multi dates (model overlap)
42		Grass (any date) / Bare mix all dates (model overlap)
43		Tree/Bush mix 1 date only + Bare 1 date only + burn (any date)
44		Tree 1 date only + Bare pure 1 date only + burn (any date)
45		Tree/Bush mix multi dates + Bare pure 1 date only + burn (any date)
46		Tree multi dates + Bare pure 1 date only + burn (any date)
47		golf-irrig combo / Tree/Bush mix 1 date only + Bare pure 1 date only + burn (any date)
48		golf-irrig combo / Tree 1 date only + Bare pure 1 date only + burn (any date)
49		golf-irrig combo / Tree/Bush mix multi dates + Bare pure 1 date only + burn (any date)
50		golf-irrig combo / Tree multi dates + Bare pure 1 date only + burn (any date)
51		grass (all dates) + burn (any date)

APPENDIX D: Visual representation of canopy cover density percentages per unit area



Adapted from Paine, D.P. 1981. *Aerial photography and image interpretation for resource management*. New York: John Wiley & Sons, 422 p.

APPENDIX E: Contingency matrix representing the map accuracy results for all 33 land-cover/land-use classes

			CLASSIFIED LAND-COVER (IMAGE)																			
			Indigenous Forest	Dense Bush, Thicket	Woodland, Open Bush	Low Shrub: Other	Forest Plantation	Cultivated Commercial Crop	Cultivated Pivot Crop	Cultivated Orchards	Cultivated Vineyard	Cultivated Subsistence	Settlements	Wetlands	Grasslands	Fynbos: Low Shrubland	Mines	Water	Bare Ground			
			1	2	3	4	5	6	7	8	9	10	11	12	13	17	32	33	34	S	Pr	
	Indigenous Forest	1	106	2	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	112	0.017	6
	Dense Bush, Thicket	2	22	230	7	3	5	0	0	1	0	0	2	1	1	3	0	0	0	275	0.043	45
	Woodland, Open Bush	3	5	76	275	10	7	4	3	0	0	7	4	11	82	12	5	0	7	508	0.079	233
	Low Shrub: Other	4	0	33	80	408	0	0	1	0	0	1	0	0	54	9	0	0	74	660	0.103	252
	Forest Plantation	5	1	8	1	0	217	0	0	0	0	0	1	1	0	0	0	1	0	230	0.036	13
ACTUAL LAND COVER	Cultivated Commercial Crop	6	0	0	0	0	0	216	1	0	0	0	0	0	0	0	0	0	0	217	0.034	1
	Cultivated Pivot Crop	7	0	3	2	1	2	14	744	18	7	2	4	1	5	2	0	0	0	805	0.125	61
	Cultivated Orchards	8	0	0	0	1	1	1	5	283	5	1	0	0	0	0	0	0	0	297	0.046	14
	Cultivated Vineyard	9	0	1	0	0	0	0	0	3	142	0	0	0	0	0	0	0	0	146	0.023	4
(FIELD)	Cultivated Subsistence	10	0	1	5	0	0	0	0	0	0	186	0	0	3	0	0	0	2	197	0.031	11
	Settlements	11	0	0	0	1	0	0	2	0	0	0	523	0	2	0	1	0	1	530	0.083	7
	Wetlands	12	0	5	7	3	0	0	0	0	0	0	0	155	0	0	0	0	0	170	0.027	15
	Grasslands	13	12	51	59	112	9	0	24	2	0	11	20	6	849	30	3	0	28	1216	0.190	367
	Fynbos: Low Shrubland	17	0	12	0	1	0	0	0	0	0	0	0	0	1	223	0	0	2	239	0.037	16
	Mines	32	0	0	0	1	0	0	0	0	0	0	1	0	0	0	155	1	0	158	0.025	3
	Water	33	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	161	0	165	0.026	4
	Bare Ground	34	0	6	14	37	0	0	0	0	0	1	1	0	7	1	2	0	421	490	0.076	69
		S	146	428	452	578	243	235	780	307	155	209	557	176	1004	280	167	163	535	6415		459
		Pc	0.023	0.067	0.070	0.090	0.038	0.037	0.122	0.048	0.024	0.033	0.087	0.027	0.157	0.044	0.026	0.025	0.083		0.08659	
			40	198	177	170	26	19	36	24	13	23	34	21	155	57	12	2	114	1121		

GEOTERRA IMAGE

The project was funded by the Department for International Development (DFID), UK Government

UKaid from the British people

Department of Environmental Affairs

Environment House
473 Steve Biko
Arcadia
Pretoria, 0083
South Africa

DEA call centre:
+27 86 111 2468