European Commission

JRC SCIENCE AND POLICY REPORTS

Towards a country-wide mapping & monitoring of formal and informal settlements in South Africa

Pilot-study in cooperation with the South African National Space Agency (SANSA)

Thomas Kemper (JRC)
Nale Mudau (SANSA)
Paida Mangara (SANSA)
Martino Pesaresi (JRC)
2015

Report EUR 27024 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information
Thomas Kemper
Address: Joint Research Centre, Via Enrico Fermi 2749, TP 267 21027 Ispra (VA), Italy
E-mail: thomas.kemper@ec.europa.eu
Tel.: +39 0332 785576

https://ec.europa.eu/jrc

JRC92657

EUR 27024 EN

ISBN 978-92-79-44709-9

ISSN 1831-9424

doi:10.2788/970905

Luxembourg: Publications Office of the European Union, 2015

Abstract

This report describes a pilot study that is carried out jointly by the European Commission's Joint Research Centre (JRC) and the South African National Space Agency (SANSA). The pilot study aims at develop a robust methodology for the automated mapping of settlements, formal and informal, in the entire territory of South Africa. It relies on the methodology for automated settlement mapping, also known as Global Human Settlement Layer (GHSL), developed by the JRC and the high resolution satellite data holdings acquired by SANSA. Amongst other uses it aims at supporting the National Department of Human Settlement (NDHS) in its implementation of the Upgrading Informal Settlements Programme (UISP) with the objective of eventually upgrading all informal settlements in the country. The presentation of the results of this pilot study focusses on the Gauteng area, the cities of Durban and Rustenburg as well as a rural area in the Limpopo province. It is planned to release the wall-to wall data set of the settlements of South Africa of 2012 in the second half of 2015.

Table of Contents

Introduction

This report describes an experiment that was carried out jointly between the South African Space Agency (SANSA) and the European Commission's Joint Research Centre (JRC) in the framework of a Collaboration Arrangement between both partners. The experiment aims at preparing the ground for a regular monitoring of the human settlements of South Africa and possibly also beyond.

The JRC has developed in the past years a concept and the corresponding tools for a Global Human Settlement Layer (GHSL), which aims at mapping and characterising automatically human settlements from remotely sensed imagery globally at high spatial resolution. For SANSA the mapping of human settlements is a flagship project due to the strong need for up-to-date information of formal and informal settlements. With the systematic acquisition of SPOT satellites images since 2006, SANSA provides high quality satellite imagery for this pilot-study, which will lead in 2015 to the production of the first wall-to-wall coverage of the settlements of South Africa.

Urban growth in sub-Saharan Africa and South Africa

Understanding the dynamics of human settlements is a pre-requisite for sustainable development and environmental management. Urbanisation is a major change taking place in both developed and developing countries. This growth is attributed to population increase and migration of people from rural areas to urban areas in search of employment and better living conditions. In 2014, 54 % of total population live in urban areas up from 36 % in 1960 and this figure continues to grow. Projections indicate that more than 60% of the world's population will be living in urban areas by 2030. Recent studies reveal that most of the global population growth will occur in smaller towns and cities in the developing countries in the next 30 (Cohen, 2006; UNFPA, 2007; Seto et al., 2011). In the last decade, urban growth in developing world grew by 1.2 million a week. This figure is slightly less than number of people that moving to urban areas in Europe in one full year. In Africa, about 1.23 million people move to urban areas in a week. While African is the least urbanized continent, it is experiencing the most urban growth in the world. About 40% of the population live in urban areas in Africa. This translates to an estimated, 297 million people living in urban areas. This number is expected to exceed to 50% by 2030.

In 2005, Africa had 43 cities with more than one million inhabitants compared to 28 cities a decade earlier. Most of these cities are unable to unable to support the demands of urbanisation due to the fact that old colonial plans and practises did not include the marginalised poor populations during the services and infrastructure investments. With low economic development, most cities cannot cope with high demands of services and environmental management due to urbanisation.

In South Africa, the proportion of people living in urban areas in South Africa increased from 52% in 1990 to 62% in 2011. Both cities and smaller towns are experiencing high growth rates in South Africa. In addition to natural population growth and migration of people from rural areas to cities, urbanisation in South African is also influenced by migration of people from neighbouring and other parts of Africa (STATSSA, Census 2011). Proliferation of informal settlements around the South African cities and towns is evident as poor people settle in informal settlements in search of employment.

Recent studies reveals that on 120 cities in the world reveals that urban areas grew, on average more than double the growth of urban population growth. This shows the need to map and monitor the spatial location of human settlements in order to support sustainable development in the cities.

Informal settlements

Informal settlements remain a challenge in developing countries. Informal settlements, also known as slums, shacks or squatter camps are characterised by insecure tenure, lack of basic services, inadequate building structures, overcrowding, unhealthy or hazardous environmental conditions, poverty and exclusion. According to Millennium Development Goal (MDP) 2012 report, more than 800 million people globally live in informal settlements. Informal settlements developments are evident mostly in the cities. This is due to the increase the influx of immigrants from rural areas to the cities. About 90% of new urban settlements are taking the form of informal settlements in Sub-Saharan Africa. In South Africa, about 1 249 777 households live in informal settlements excluding backyard shacks (StSTATSSA Census 2011). This translates to 4.2 million people of South Africa's 51.7 million living in informal settlements.

Governments of the developing countries with the support of the donor community have implemented numerous projects and programmes aimed at upgrading informal settlements since the 1970s. These programmes range from evictions, demolition to upgrading of the informal settlements. These programmes are funded by international donors, governments, local NGOs and private industry. Countries which have implemented such programmes include Senegal, Zambia, Kenya, Botswana, Nigeria and South Africa.

Since the establishment of Millennium Development Goals (MDG), developing countries have made significant progress in improving the lives of people living in informal settlements. Even though the slum target that aimed at improving the lives of 100 million people living in urban informal settlements by 2020 is reached, the number of people living in urban informal settlements in developing countries continues to grow. The population in urban informal settlements was estimated at 863 million in 2012, compared to 650 million in 1990 and 760 million in 2000. This shows that more efforts needs are required to improve the lives of informal settlements dwellers.

The South African government is party to the United Nations MDG 7 Target 11 which provide for the improvement of people living in informal settlements and has established Upgrading of Informal Settlement Programme (UISP) which is aimed at upgrading all informal settlements in the country using a phased approach. To fast track the service delivery, in 2010, the South African government has established Outcome based approach of which Outcome 8 Output 11 focuses on the upgrade of 400 000 units in informal settlements by 2014. The informal settlement upgrade activities undertaken range from provision of water, sanitation, electricity, refuses removal and adequate housing. The study done by the South African National Space Agency and National Departments of Human Settlement on informal settlement with 45 municipalities shows that the informal settlements within 45 priority municipalities have increased by 191 between 2006 and 2011 despite the provision of adequate housing by the government. There is a need to continuously track the developments of informal settlements to support sustainable development and resources allocation for informal settlement upgrade programmes. One of challenge faced by local authorities responsible for upgrading of the informal settlements is access to timely and consistent spatial information on the informal settlements development.

Possibility of remote sensing

The city managers and planners require update, timely and actionable spatial data to support proactive decision making that encourages sustainable development. The spatial data on human settlement can assist in operation such as allocation of property rights, housing needs, land use management and support management of key problems such as disaster management and environmental management. Remote sensing provides capabilities to map and monitor human settlements developments over larger and inaccessible geographic areas. International I initiatives such as Group on Earth observation (GEO) data democracy provided open access to low to medium resolution satellite imagery including Landsat. These images are suitable for various land applications including land cover and land use monitoring. High resolution is required to support local authorities in planning and monitoring of developments around their area of interest. Some of the challenges faced by city planners and managers in Africa are high costs associated with acquisition of high resolution imagery and expertise to generate solutions to support urban planning and management. Lack of spatial data policies, data access and sharing mechanisms, and metadata discovery tools prohibit the exploitation of spatial technologies in urban planning and management.

In South Africa, there are a number of initiatives that are aimed and capturing human settlements data for different levels of planning and management. These initiatives include Eskom's SPOT Building Count, STATSSA dwelling Frame and land cover and land use by various levels of government departments. The methodologies used for these initiatives are time consuming and resource. To respond to the challenge of access to consistent and up-to-date human settlement data, this project aims at developing the methodology to automatically extract human settlement from high resolution imagery.

Objectives of the pilot-study

The objective of this study is to map the human settlement structures in the South Africa using available nation SPOT 5 imagery using the Global Human Settlement Layer (GHSL) human settlement detection system. With open access to other medium to high resolution imagery, SANSA plans to partner with other African agencies to produce data human settlements data for other African countries that do not have access to high resolution human settlement data. This data will support decision making by governments on sustainable service delivery, informal settlements upgrade, disaster and environmental management.

In South Africa, a national human settlement layer will provide a base for studying urbanisation and its impact on the environment, informal settlements development, low cost housing verification, service delivery planning and monitoring, urban and environmental policy formulation and planning. The use of automated procedures to detect human settlement data will results in saving of resources currently spent in traditional methods of mapping the human settlement data.

Data set

Remote sensing imagery

Due to its availability in South Africa, SPOT 5 acquired in 2012 was used for this project. Since 2006, SANSA has been acquiring the national SPOT 5 imagery annually to support various aspects of government planning and monitoring ranging from infrastructure to environmental management. SPOT 5, launched in May 2002, acquires imagery in multispectral and panchromatic modes. The panchromatic and multispectral modes have has 2.5-5m and 10m spatial resolutions respectively. One scene of SPOT 6 covers 60 X 60 km2 swath. 970 panchromatic and multispectral scenes were used for this study. The imagery used was acquired between January and December 2012. The images were georeferenced using 25cm aerial photography and 20m Digital Elevation Model (DEM) and projected to Universal Transverse Mercator (UTM) system. The aerial photography used was acquired by Department of Rural Development and Land Reform. The imagery covered only the eastern part of the country.
Both the panchromatic and multispectral images were georeferenced using SARMES system that automatically georeference raw imagery. The SARMES system runs on PCI software. The images that yielded high RSME errors were georeferenced manually. The accuracy assessment was done by manually assessing the geographic positions of same points in the satellite imagery and on the aerial photography. An accuracy of 12m at 2-sigma was achieved.
To improve the spatial resolution of the satellite imagery, the 2.5m spatial resolution panchromatic imagery and four multispectral bands with 10m spatial resolution imagery were pansharpened using PCI software to produce 2.5m resolution imagery.

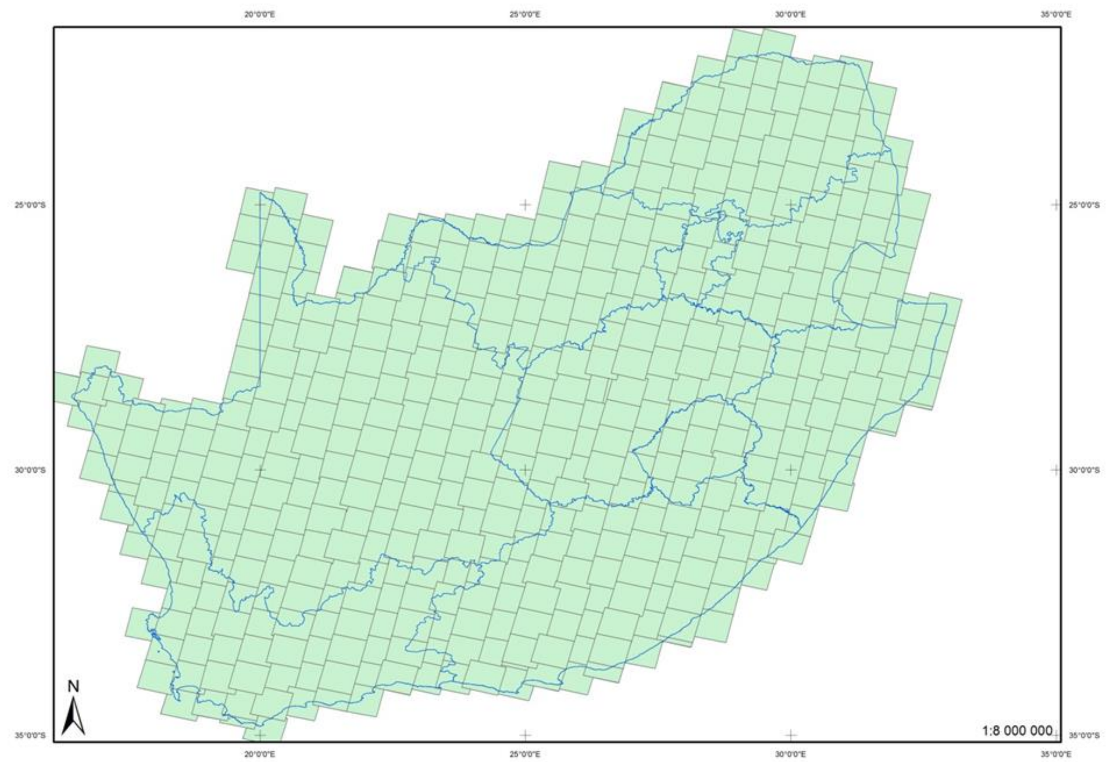

Figure 1. SPOT 5 scenes covering South Africa

Ancillary data

Ancillary data sets used in this project are South African National Land Cover (NLC) 2000 and 2012 SPOT Building Count (SBC). The NLC 2000 data set was derived using multi temporal Landsat 7 ETM imagery acquired in 2000-2003. This data set contains 49 land cover and land use classes of which 16 are urban land use classes. The vegetation and natural environmental land cover classes were mapped using pixel based classification where as human settlements and other spectral heterogeneous land use classes were mapped through manual digitisation. NLC 2000 is the latest detailed land cover and land use data set available over South Africa and was suitable for use as reference for human settlement classification.

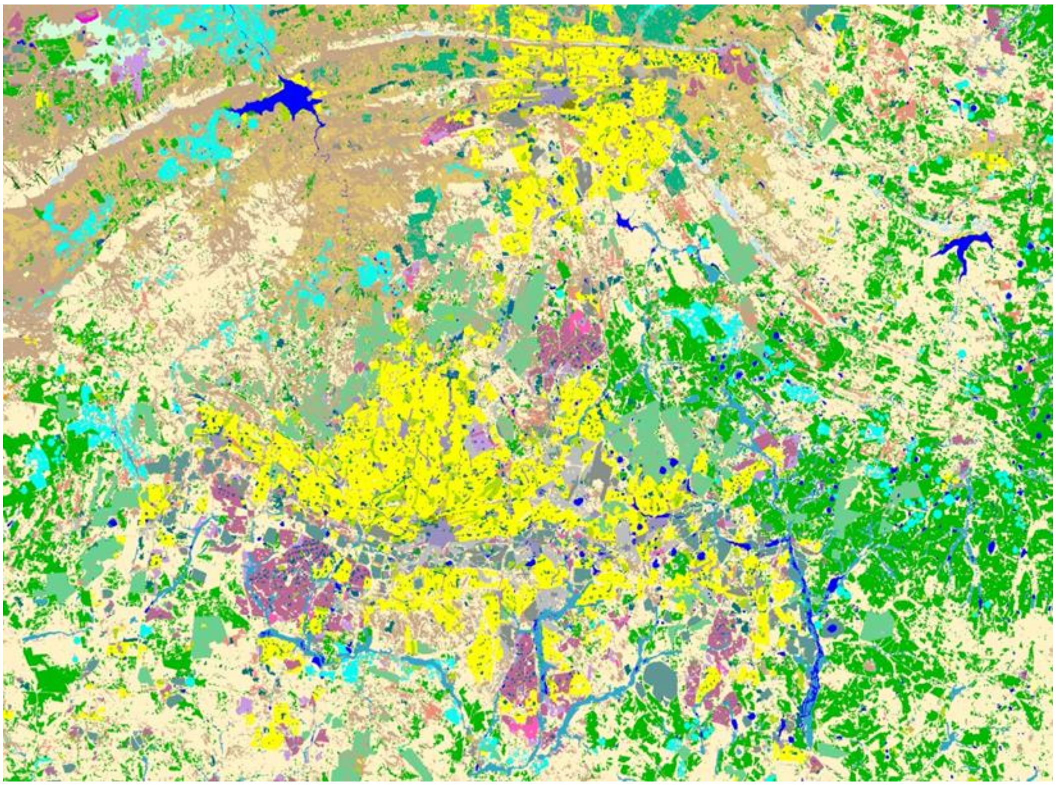

Figure 2. Land cover map over Gauteng province, South Africa

The SBC is a point and polygon data set derived from SPOT 5 satellite imagery and is funded by Eskom, the power utility company. The first layer of SBC was generated using SPOT 5 imagery acquired in 2006. This dataset is updated annually after each national SPOT 5 release. The SBC layer is developed through visual interpretation and manual digitisation of the building structures. The location of individual building structures identifiable from SPOT 5 2.5 natural colour image is mapped by points whereas the extent of informal settlements is mapped by polygon as individual dwellings within informal settlements cannot be identified using 2.5m spatial resolution imagery. The layer used in this study was updated using 2012 SPOT 5 imagery. The SBC layer was used mainly for the validation of the settlement mask. In addition it is used in scarcely populated areas to mark potentially settled areas. Figure 3 gives an example of SBC data over SPOT 5 natural colour imagery.

Figure 3. SBC points over SPOT 5 2.5 natural colour

Methodology

The workflow for the processing of the country-wide settlement information is based on a methodology that was developed by the JRC for the first operational test of an automated image information retrieval system for the production of the first GHSL during June-July 2012. Details of this methodology were published by Pesaresi et al. (2013).[1]

Before describing the methodology it is crucial to understand the conceptual design of the Global Human Settlement Layer, since they are the foundation for the development of the methodology:

- The GHSL is developed with a **multi-scale vision**. The results generated at local level must be comparable to results at regional and global level and – to the extent possible – vice versa.
- The GHSL is based on a **multi-sensor approach**. The methodology is able to produce results from optical air- and satellite-borne sensors with spatial resolutions ranging from 0.5 to 10 m and with a modified methodology up to 30 m.
- Having a global product in mind the GHSL relies on truly **automatic image information retrieval**. GHSL generates mostly quantitative information such as built-up area, building size and density. However, in this report for the first time

[1] An extended version of this paper is available as a report for download here: http://dx.doi.org/10.2788/74059

the generation of land use and land cover information is tested. The automatic information retrieval is key for realistic data volume and quality scenarios.

- Finally the GHSL concept follows an **inclusive concept of human settlements** from small hamlets to megacities including explicitly refugee/IDP camps, slums, rural hamlets, which are often neglected in global/regional mapping activities.

Workflow

For the human settlement layer of South Africa, a new workflow was developed that is taking stock of the experiences since the development of the first GHSL product and the availability of a homogeneous input image data set. The workflow is depicted in Figure 4 and will be described in the following.

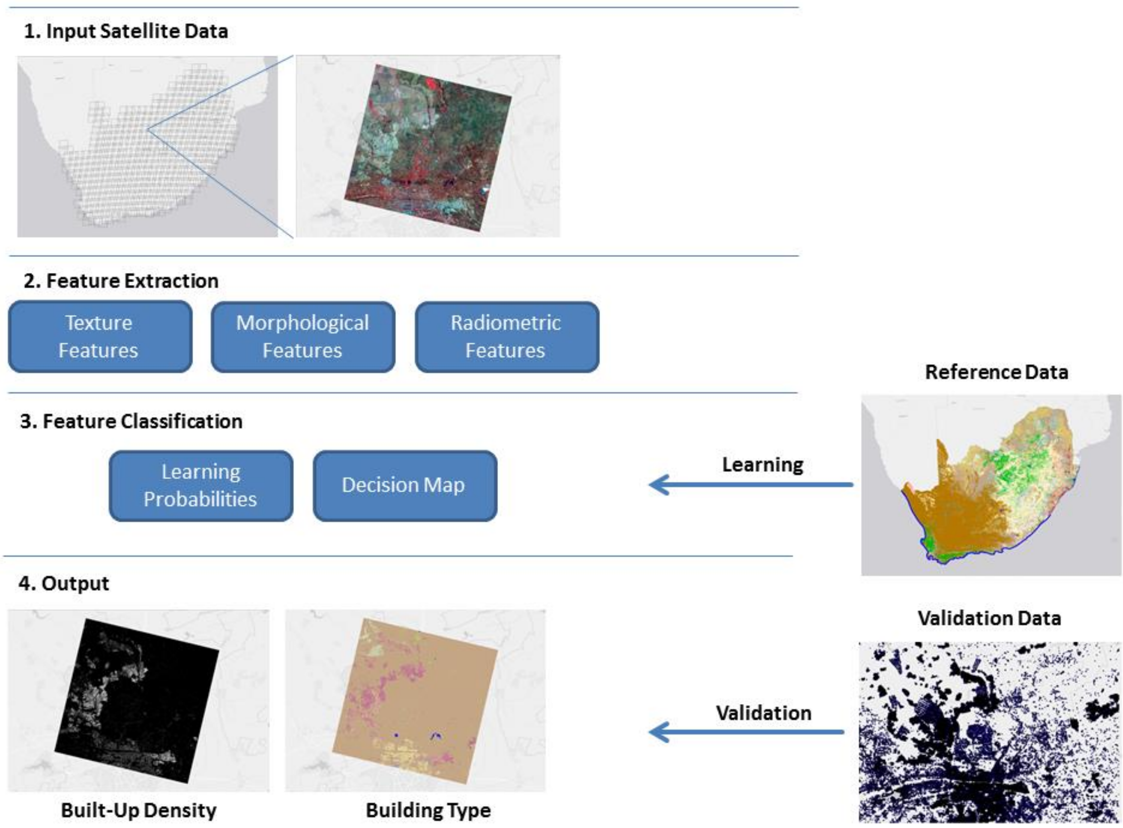

Figure 4. GHSL South Africa workflow

The workflow for the wall-to-wall mapping of human settlements map in South Africa is scene based. All images are processed independently from each other. The system assumes the data are already orthorectified or georeferenced. For the pilot-study the system assumes also cloud-free imagery. For future developments the integration of an automated cloud masking as described in Pesaresi et al. (2013) is possible.
After reading a single scene, the luminance (the maximum reflectance in the visible bands) of the image is calculated. This maximises the settlement contrast with the background and reduces the image size to a single band image. The geographic extend (bounding box) of the image is used to cut the equivalent area from the reference data set (Figure 5).

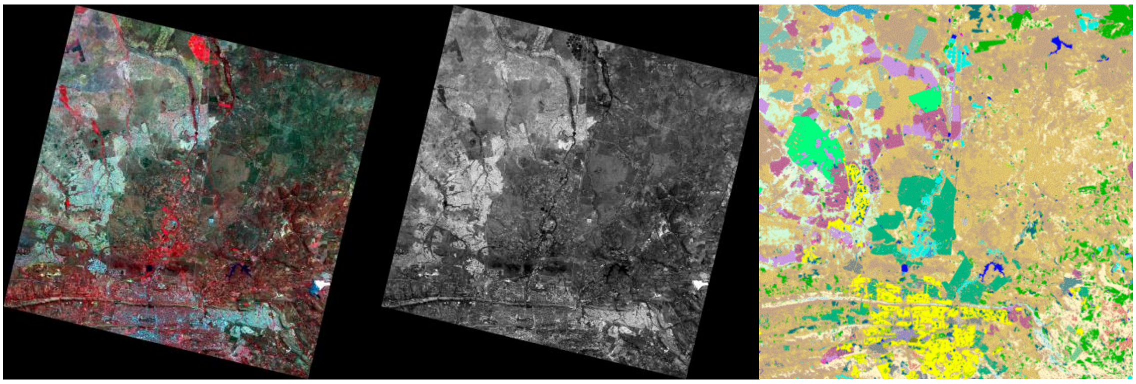

Figure 5. Multispectral input image for radiometric analysis (left), luminance image (middle) for the textural and morphological feature extraction and the subset of the NLC2000 (right) as input for the learning of the corresponding SPOT-5 satellite image.

Feature Extraction

The second step is the feature extraction. The workflow calculates three different types of features: texture, morphological and radiometric features. The following section describes each of the three features.

The textural image features used in this study are derived from grey-level co-occurrence matrix (GLCM) contrast textural measurements (Haralick et al. 1973). The contrast textural measures calculated using anisotropic displacement vectors are combined in a rotation-invariant image feature called PANTEX (Pesaresi et al. 2008) by using extrema operators. PANTEX was demonstrated strongly correlated with the presence of buildings (Pesaresi et al. 2011a) and their spatial density (Pesaresi et al. 2011b). The capacity of PANTEX to discriminate built-up from non-built-up areas is mainly linked to the fact that it is a corner detector (Gueguen et al. 2012) and that the BU areas generate high local image spectral heterogeneity because of heterogeneity of materials used and because of buildings are generally casting shadows. The result of the PANTEX feature extraction is shown in Figure 6.

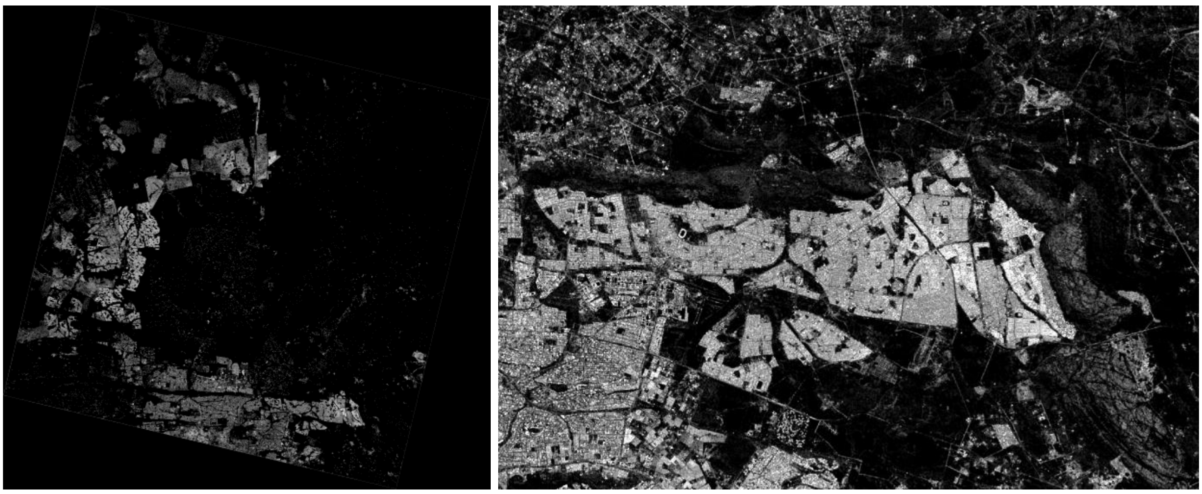

Figure 6. PANTEX features for the test scene full extent (left) and zoom to the Mamoledi township east of Pretoria (right)

The morphological feature analysis extracts single objects of different sizes, which can be used to identify building footprint candidates. The morphological features are summarized in an information layer that is a product of a multi-scale morphological analysis protocol referred to as the "mtDAP" (Ouzounis et al. 2012). The mtDAP protocol computes the Differential Attribute Profile (DAP) vector field (Mura et al. 2010) from the input imagery. DAPs are nonlinear spatial signatures that are used extensively in remote sensing optical image analysis in ways analogous to spectral signatures. The DAP of a pixel is the concatenation of two response vectors. The first registers intensity differences, i.e., contrast, within a top-hat scale-scape of an anti-extensive attribute filter , and the second intensity differences on the bottom-hat scale-space of an extensive attribute filter. The pair defines an adjunction with the first typically being a connected attribute opening and the second being a connected attribute closing. The mtDAP can be configured with any morphological attribute filter but in this case, simple area openings and closings prove to be sufficient. The area attribute is used to order objects based on size and it is computed incrementally. More detailed information is provided in Pesaresi et al. (2013).

The resolution of the DAPs (vector length and the between-scale spacing) is a critical parameter in their utilization as feature descriptors. On the one hand, higher spatial input resolution offers a more detailed profile for each pixel. On the other hand by increasing the vector length the number of DAP vector field planes increases proportionally. This can become an issue with large data sets at higher resolution. Hence, Pesaresi et al. (2012) proposed a compression model was devised to radically reduce the dimensionality of the DAP descriptors. It is called the Characteristic-Saliency-Level or CSL Model and is a medium abstraction semantic layer that can be projected on the HSV colour space for the visual exploration of built-up extracted from VHR satellite imagery. The CSL model is a nonlinear mixture model consisting of three characteristic parameters extracted from the DAP of each pixel. That is the minimum scale at which the maximal contrast response is observed in the DAP (the characteristic), the contrast value (the saliency) and the highest peak component level from which this value is computed. The model is computed directly from the polychotomy of the two tree-based data structures and with no need for exporting the DAP vector fields. It reduces radically the dimensionality of the DAP vector field to a three-band representation in a statistical-model free approach, i.e., it avoids clustering based on the statistical distribution of the DAP features of a given image. It does not require manual tuning and its computation is independent of the length of the DAP.

The output of the morphological feature analysis (Figure 7) identifies building footprints candidates of different sizes. At this point the objects include many non-building features such as roads and open spaces, but in combination with additional layers (e.g. vegetation or road network) it is possible to improve the precision of the feature detection.

The radiometric features are grouping pixels based on their radiometric behaviour. This is obtained by quantization and subsequent sampling of the feature space. Figure 8 shows an example of the feature extraction. Large industrial buildings are clearly discernible, while smaller features, such as residential houses are less apparent.

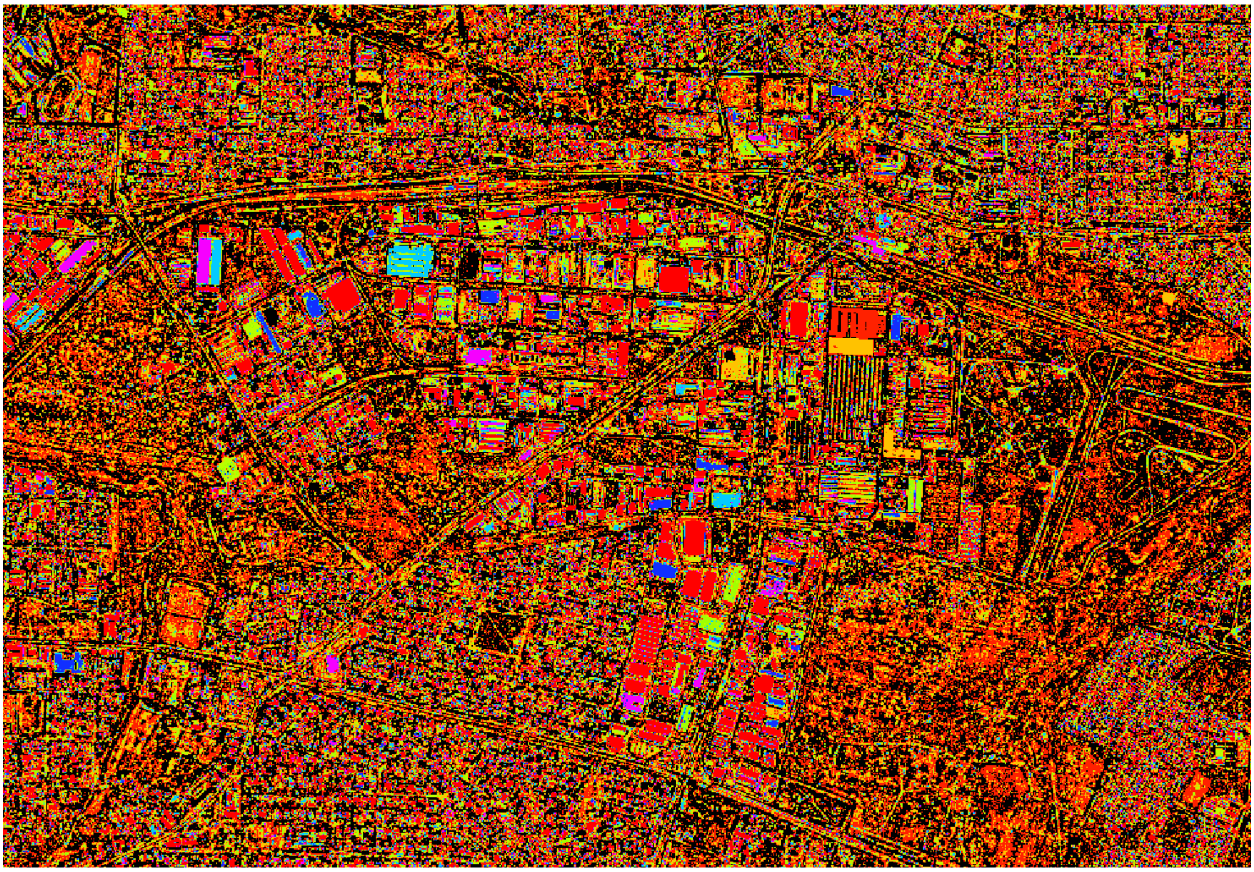

Figure 7. Exemplary output of the morphological feature extraction. The colours mark objects of similar size.

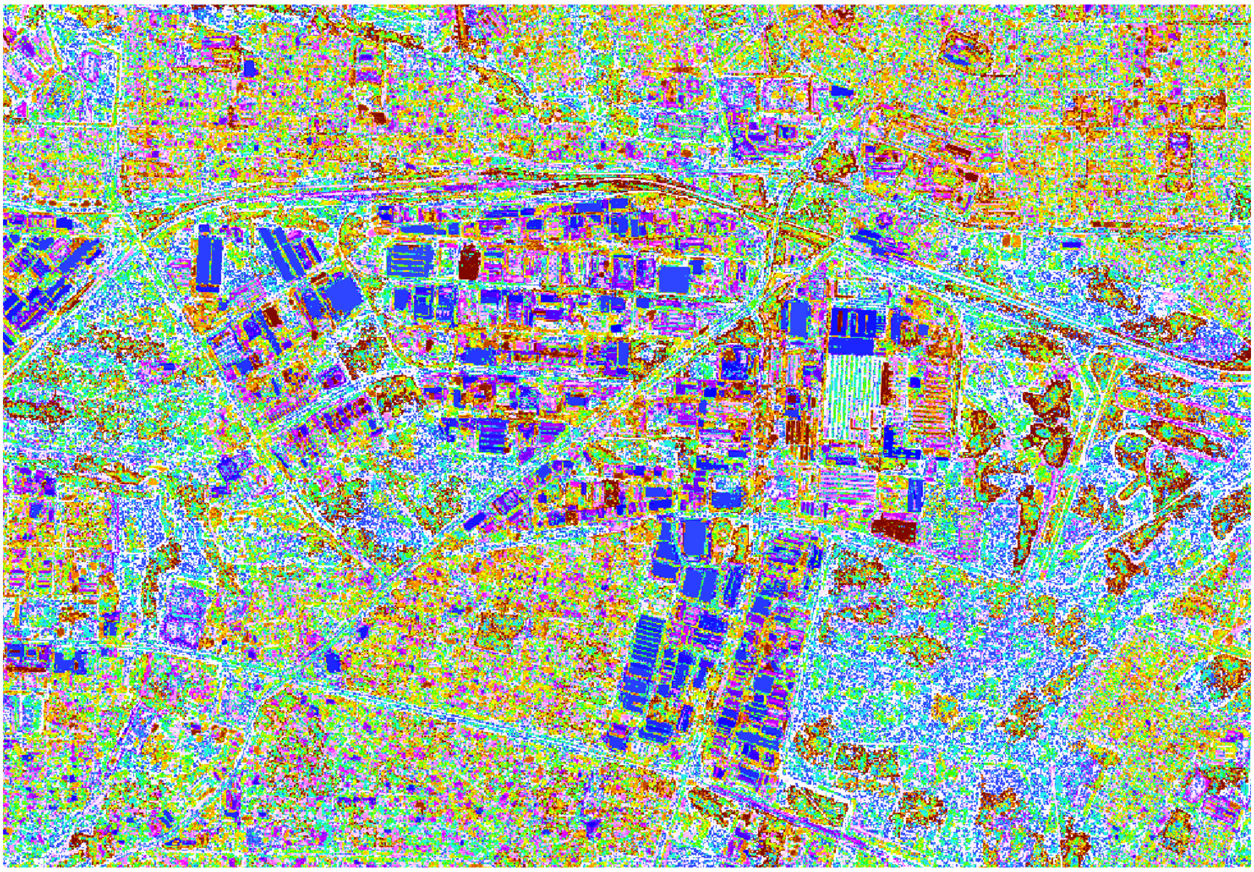

Figure 8. Exemplary output of the radiometric feature extraction. The colours mark similar features.

Feature Classification

In the final learning step the system compares the coarse resolution land use and land cover information with the radiometric, texture and morphological features in order to understand the relation primarily between built-up and non-built-up pattern. During this learning stage the best-threshold between built-up and non-built-up is derived as described in more detail in Pesaresi et al. (2013). For the LULC classes for each pixel a probability to belong to a specific LULC class is calculated and the class with the highest probability is finally selected. Figure 9 illustrates the final results. The first image (top) shows the built-up map. It is interesting to note the differences in building density between dense townships in the East, the residential areas (West) and the areas of smallholdings in the North/NorthEast. On top of this presence/density Information the LULC map (bottom) classifies each settlement into a specific settlement type.

Validation

For a pilot-study, as described in this report, it is indispensable to provide a thorough validation of the results in order to allow a true evaluation of the results in view of a potential application of the method to the entire country. Assessing the accuracy of land cover data is not trivial. Often there is no adequate, spatially consistent, up-to-date and area-wide reference data available. On top there is still a discussion in the scientific community regarding the best agreement measures (McPherson et al, 2004; Allouche et al., 2006; Taubenböck et al., 2011). Nevertheless, it is now widely recognized in today's scientific community that no classification is valid or complete until a certain degree of confidence of the mapping accuracy has been gained. Error or confusion matrices are an often used approach, which mostly use randomly distributed test sites leading to a descriptive evaluation by standard measures of agreement between the validation data and the classification output (Taubenböck et al., 2011). These standard measures often include the producer's accuracy to determine the error of commission of allocated pixels as well as the user's accuracy as a measure of the omission. However, there is broad consensus in the scientific literature that metrics based on the entire error matrix add significant value to the accuracy assessment beyond these basic measures. A thorough review of the standard accuracy measures, problems encountered when using them and remarks on meaningful interpretation is given by Foody (2008).

As a first step to assess the capabilities of mapping of urban areas, both maps of global urban extent are visually and quantitatively compared to the reference dataset on object level to determine the share of buildings captured by these layers. From this, the error of omission regarding missed buildings is calculated. These basic descriptive measures establish a general degree of completeness of the classified settlement pattern as defined by building inventory, however, do not comprise information on its overall correctness and quality.

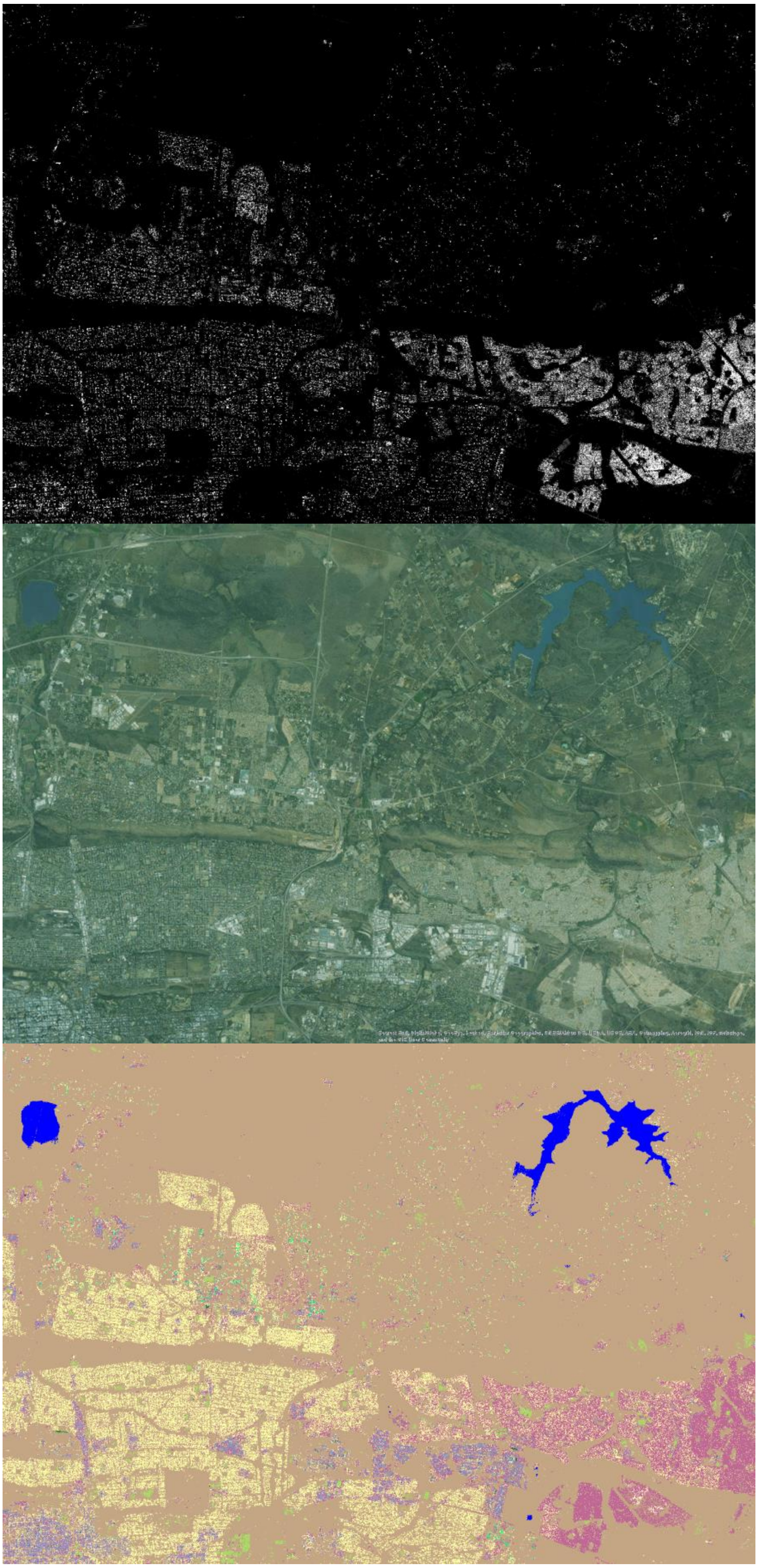

Figure 9. Example of the final output of the workflow: built-up mask (top), reference image (middle) and LULC (bottom)

Table 1. Contingency table for a two-class map comparison

		Validation data	
		Presence	Absence
Settlement layer	Presence	*a*	*b*
	Absence	*c*	*d*

For the validation of the GHSL layer a wider set of accuracy measures is calculated on a per-pixel basis as recommended by Foody (2008). These are based on two-class contingency tables resulting from the overlay of the settlement layers and the high resolution validation data and record the number of true positives (a), false positives (b), true negatives (c) and false negatives (d) (Table 1). These allow for the calculation of several quantitative measures of agreement (Table 2):

- **Overall accuracy** measures the classification accuracy as the share of all correctly classified urban and non-urban pixels in the error matrix and thus, gives general information regarding the overall map accuracy. However, this measure does not take into account unequal class distributions and thus, does not paint a detailed picture of the accuracy across individual land cover classes.

- **Sensitivity (Completeness)** relates to the ability/probability to classify urban pixels as defined by the building reference correctly. It is the percentage of the building reference data which corresponds to the classification output of the respective urban extent layer and is closely related to the error of omission (1-sensitivity). The ideal value for the completeness is 100 percent. In turn, the ability of classifying the absence of urban areas correctly is called specificity.

- **Specificity** relates to the test's ability to identify negative results.

- **Precision (Correctness)** relates to the classifier's ability to exclude non-urban areas correctly from the urban extent classification as defined by the building reference. This measure is closely related to the error of commission (1-precision) and reaches and ideal value of 100 percent.

- **Balanced Accuracy** avoids inflated performance estimates on imbalanced datasets. This is usually true for settlement maps, which often occupy a small percentage of area.

- **True Skill Statistic (TSS) or Informedness** is designed to measure the agreement between the classification and the building reference layer. It is calculated as the specificity (fraction of correctly classified urban pixels) plus the sensitivity (fraction of correctly classified non-urban pixels) minus one. Compared to Kappa it has the advantage of being independent from unequal class distributions, i.e. prevalence ($(a+c)/n$ which is the proportion of pixels assigned to buildings in the reference dataset), and thus provides a more robust measure of classification accuracy (Allouche et al., 2006). Its range spans from negative values (systematic disagreement) to 1 (perfect agreement), with a value of 0 indicating a random classification result.

Table 2 Map agreement measures used in this work; in all formulae n = a + b + c + d (Allouche et al., 2006)

Measure	Formula
Overall accuracy	$\frac{a+d}{n}$
Sensitivity (Completeness)	$\frac{a}{a+c}$
Specificity	$\frac{c}{b+c}$
Precision (Correctness)	$\frac{a}{a+b}$
Balances Accuracy	$\frac{sensitivity + specificity}{2}$
True Skill statistic (TSS)	$sensitivity + specificity - 1$

In this pilot-study the Spot Building Count (SBC) is used as reference data set. It is used for the validation of the built-up mask (Figure 9, top). At this point the LULC map (Figure 9, bottom) is not validated due to the lack of appropriate high resolution data.

Results

The results of the pilot study will be presented for the area covered by five SPOT-5 satellite images (Figure 10). In this way it is possible to discuss a range of settings (e.g. rural, urban or coastal environments), while at the same time keeping a high level of detail, which might be difficult to maintain in a country wide assessment.
The selected areas are:

- **Thohoyandou (Limpopo Province):** this is a rural setting with smaller villages and hamlets including a considerable amount of thatched roof dwellings (rondavels).
- **Rustenburg (North West Province):** this is one of the fastest growing cities in South Africa with a lot of mining areas, informal settlements and low cost housing projects.
- **Pretoria (Gauteng):** the capital of South Africa with is characterised by very distinct settlement in the west, east and north, relative to the central business district.
- **Johannesburg (Gauteng):** the provincial capital of Gauteng, the wealthiest province in South Africa, having the largest economy of any metropolitan region in Sub-Saharan Africa. Johannesburg is one of the 50 largest urban agglomerations in the world, and is also the world's largest city not situated on a river, lake, or coastline.
- **Durban (Kwazulu Natal):** it is the second most important manufacturing hub in South Africa after Johannesburg and has the busiest port in South Africa and Africa.

Qualitative analysis of the results

The qualitative analysis of the results focusses mostly on the visual assessment of the information produced. It tries to answer questions such as the following: Is the settlement information reproducing the settlement pattern according to the expectations? Is it providing added value to the available information? What are the observed shortcomings?

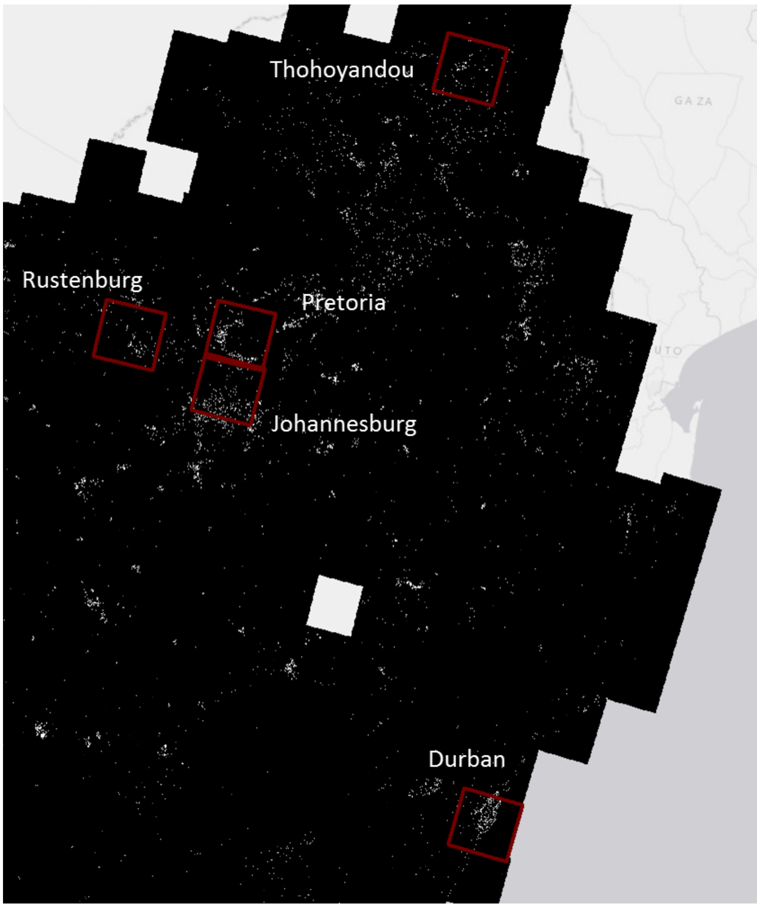

Figure 10. Location of the five areas selected for the presentation of the results (red boxes) superimposed on the built-up mask.

The settlement pattern for Pretoria (Figure 11) reveals clearly the structure of the city, where we see low density residential areas around the CBD, which is spreading in East-West direction. The map also reveals very well the different building densities. The Mamoledi township in the East shows the highest densities with a strong contrast to the smallholding farmland North of it. Also the Central Business District and the residential suburbs are recognisable.

Using the advanced workflow developed for this pilot-study it possible to refine the coarse NLC classification and to derive a much more detailed LULC. This experimental LULC map uses aggregated classes from the NLC 2000 (see Annex 2 for details). With this new methodology it is in principle possible to map any LULC class given an appropriate coarse resolution learning data set is available. However, more research is needed prior to a release of such a data set.

Figure 11. Built-up area of Pretoria.

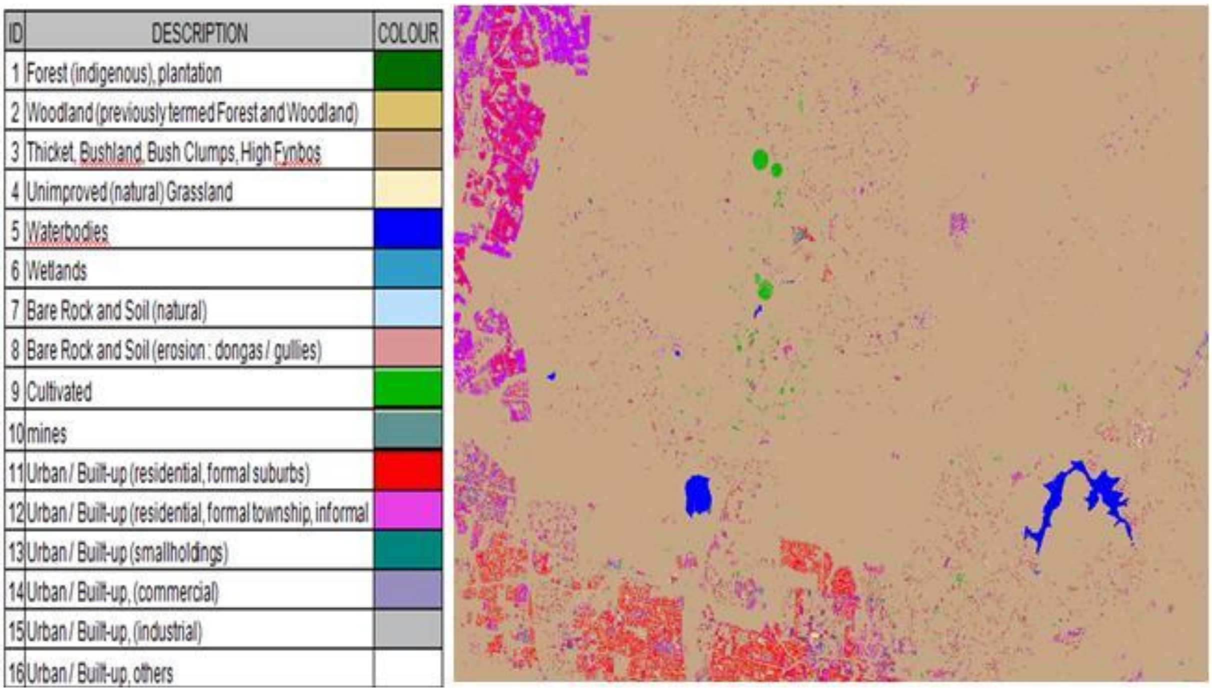

Figure 12. New experimental LULC map for Pretoria.

Quantitative analysis of the results

The quantitative assessment aims for providing statistical measurements that describe the accuracy and reliability of the produced information, which is crucial information for any application.

For this report the SPOT Building count data set was used as reference data set. In order to prepare the contingency table a reference built-up raster layer was created that has the same dimensions as the corresponding SPOT scene. Each pixel that is touched by a SBC point is marked as built-up. This reference layer is then used in a ROC analysis to determine the threshold in the built-up layer that minimises the error (Table 3).

Table 3. Results of the ROC Analysis to determine the best threshold for the built-up layer

ROC Analysis	Durban	Johannes-burg	Pretoria	Thoho-yandou	Rustenburg
Minimal Error Rate (MER)	0.0884	0.0513	0.0943	0.0544	0.0195
MER Threshold	84	79	87	49	60

After the application of the threshold the contingency table is created and the accuracy measures are calculated (Table 4). The overall accuracies range from 80% to 94% with Durban, Pretoria and Johannesburg in a group around 80% and Rustenburg and Thohoyandou above 90%. The other error measures follow the same trend. This range of values is in accordance with prior experiments (Pesaresi et al. 2013). An apparent source of error is the omission (1-sensitivity), which reaches values of more than 20% for the more urbanised scenes.

Table 4. Accuracy measures for the built-up in comparison to the SBC reference data

Accuracy Measures	Durban	Johannes-burg	Pretoria	Thoho-yandou	Rustenburg
Accuracy	0.7998	0.8098	0.8106	0.9031	0.9371
Sensitivity	0.7572	0.7859	0.7762	0.8973	0.9369
Specificity	0.9901	0.9835	0.9646	0.9494	0.9452
Precision	0.9971	0.9971	0.9899	0.9930	0.9981
Balanced_Accuracy	0.8736	0.8847	0.8704	0.9234	0.9410
TSS/Informedness	0.7473	0.7694	0.7408	0.8467	0.8821

When analysing in depth the error distribution it can be observed that the current workflow is ignoring large industrial buildings. In fact the system uses a maximum size limit in order to speed-up the processing and avoiding erroneous classification of, for example, agricultural areas. In a future workflow this threshold needs to be adapted to map also larger structures entirely.

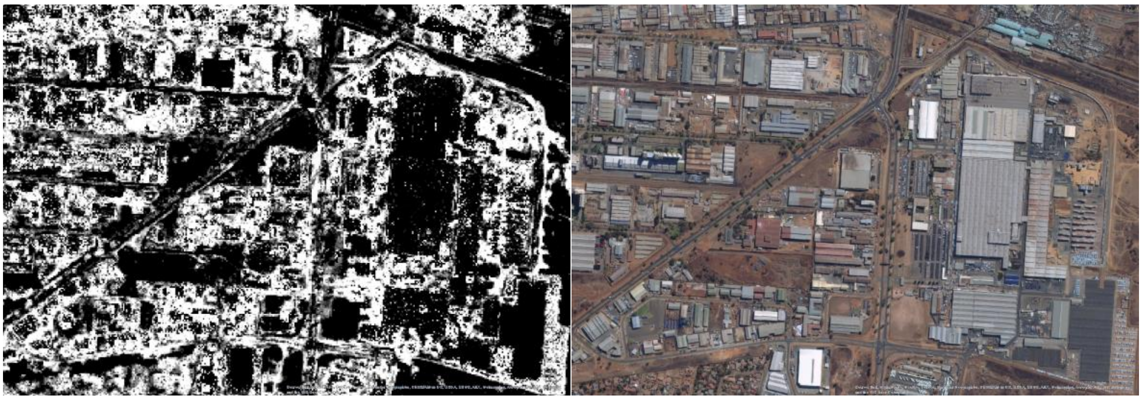

Figure 13. Example of an industrial complex. The built-up layer (left) is systematically ignoring larger industrial structures.

The comparison of the SBC data with the built-up mask highlights nicely the added value of the automated assessment. Figure 14 shows an area North of the Mamoledi township, Pretoria (Gauteng). There is a big potentially new informal settlement that was not mapped by the SBC (apart from few dwellings that might have been digitized earlier). Being a fully automatic procedure the built-up mask will help monitoring the settlement area to identify newly settled areas to guide the visual analysis.

Application domains

The National Development Plan (NDP), South Africa's strategic blueprint recognises the value of geospatial information in national spatial development and calls for the establishment of a national observatory for spatial data and analysis. Remote sensing is an integral part of any national spatial observatory since satellite imagery and aerial photography are an effective and reliable means of monitoring spatial infrastructure developments over time. Globally, satellite imagery is being used to generate geo-spatial information for spatial planning and monitoring. The NDP aims to transform human settlements through strong and efficient spatial planning systems and to upgrade informal settlements on suitable land by 2030. The densification of cities and development of public transport infrastructure has also been prioritized. The 2013 Gauteng Growth Management Strategy also clearly highlights the significance of human settlements in the long-term development plan for the province. The role of human settlements in the development South Africa's development agenda is well articulated in many national, provincial and municipal strategic planning documents. The mapping of human settlements and built up areas is therefore important to provide urban and rural planners working in the different spheres of government with spatio-temporal information critical to monitor urban and rural development. Remote sensing can be used effectively to monitor urban sprawl and to assess the impact of development policies on the environment.

The Housing Act 107 of 1997 states that 'housing development' means the establishment and maintenance of habitable, stable and sustainable public and private residential environments to ensure viable households and communities in areas allowing convenient access to economic opportunities, and to health, educational and social amenities. The vision of the South African government is to have a nation housed in sustainable Human Settlements and aims to facilitate the creation of sustainable human settlements and improved quality of life. The department's main functions are guided by the Comprehensive Plan for Sustainable Human Settlement 2005 which places emphasis on quality of housing and housing environments by integrating communities and settlements. Government Outcome 8 of 2010 aims to provide South African citizens with sustainable human settlements and improved quality of household life. Priority outcomes in the department include accelerated delivery of housing opportunities. The department is doing this through the provision of low cost housing, upgrading of informal settlement and improving access to basic services (water, sanitation, refuse removal, electricity).

The human settlement layer derived from SPOT 5 is substantially contributing to enabling the Department of Human Settlements in meeting its mandate. Human settlements maps are being used to assess and monitor informal settlements and in estimating household numbers based on the dwelling count. The value of SPOT 5 has also been accredited in the mapping of housing development, especially low cost housing. In the North-West province SPOT 5 data was used to inventory the informal settlements, backyard shacks and low cost housing. SPOT 5 and aerial photography are being to validate the number of low cost houses in the North-West province and has been accepted as valuable data source to independently verify government progress and expenditure. The information extracted from SPOT 5 imagery is now being used for performance monitoring and audit reporting. Further than supplying SPOT 5 imagery SANSA is assisting the National Department of Human Settlements and the North-West Department of Human Settlement in mapping the low cost houses and informal settlements in 45 cabinet prioritized municipalities and in developing an informal settlements atlas for 2009/10. It is foreseen that satellite imagery supplied by SANSA will continue to contribute significantly to the Integrated Residential Development Programme, Informal Settlements Upgrading Programme and Rural Housing Programme. SANSA's contribution in the attainment of government outcome 8 of 2010 is resolutely reaching new frontiers since the role of satellite imagery has permeated spatial planning and decision making in human settlements than ever before.

The human settlements information that is being developed through the Global Human Settlement Layer through the collaboration between JRC and SANSA has far reaching applications and will support the a plethora of legislative mandates assigned to the different government departments and public entities in South Africa. Some of the most prominent legislative acts that will be supported by the human settlement information include: Electoral Act through the demarcation of voting districts and verification of voting stations, the Statistics Act through supporting the dwelling frame and census planning, National Human Settlements Land Inventory Act through the quantifying of areas occupied by human settlements, Conservation of Agricultural Resources Act by monitoring encroachment of human settlements in fertile agricultural land, the Spatial Planning and Land Use Management Act through the provision of information relating to the spatial extends of human settlement and the Disaster Management Act since information on human settlements is critical for post disaster verification, disaster risking profiling and assessment , and for monitoring and evaluating the impacts of passive and active disasters.

In addition to the strong demand for settlement information in South Africa the information generated will also support the generation of the GHSL. The GHSL concept relies on a collaborative approach based on exchange of methods and data for the mutual exploitation of the derived settlement operation.

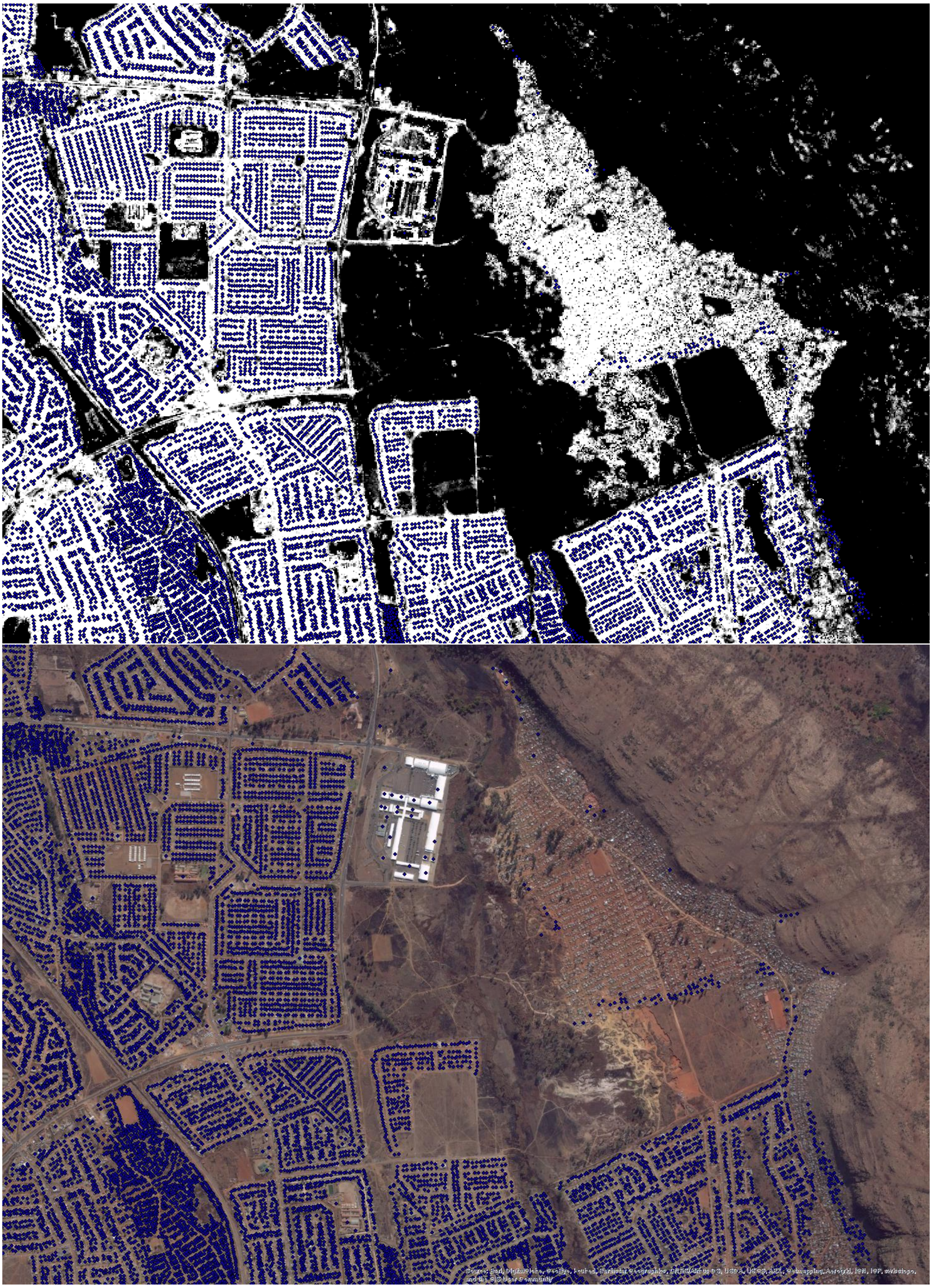

Figure 14. Comparison of SBC points (blue dots) with built-up mask (top) and reference image (bottom).

Conclusions

There is a strong demand for up-to-date information on formal and informal settlements in South Africa. Hence, this pilot-study comes very timely. This pilot-study demonstrated the potential of remote sensing to map and, very important, monitor regularly the settlements of South Africa in a sustainable way. The tools developed by the JRC for the Global Human Settlement Layer proofed to be reliable also in the South African context. With the new workflow developed for South Africa it is even possible to go beyond the state of the art. Instead of providing information on the presence of settlements, the new workflow allows also the classification of major LULC classes.

However, at this point only the built-up area product was validated with satisfactory results. Also here some shortcomings were identified. In particular the identification of larger buildings is currently not optimal and needs improvement. In addition, there is more work needed for the LULC classes: To what level can different settlement types (e.g. formal vs. informal township) be separated? How do we assure that LULC classes are consistent in neighbouring scenes?

The above mentioned points will be taken into consideration in the improvements of the workflow for the creation of a wall-to-wall settlement map for South Africa of the year 2012, which is planned for 2015. Subsequently it is foreseen to process on the one hand older archived imagery and on the other hand to process newer information layers, possibly including the new SPOT 6 and 7 data acquired by SANSA.

References

O. Allouche, A. Tsoar, R. Kadmon, Assessing the accuracy of species distribution models: prevalence, Kappa and true skill statistic (TSS). Journal of Applied Ecology, 43, 1223-1232, 2012

Ellen M. Bassett, Sumila Gulyani, Catherine Farvarque-Vitkovik, and Sylvie Debomy, Informal Settlement Upgrading in Sub-Saharan Africa: Retrospective and Lessons Learned, 2013

Cohen B, 2006, 'Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability', Elsevier, vol. 28, pp. 63-80, 2006

G.M. Foody, Harshness in image classification accuracy assessment. International Journal of Remote Sensing, 29, 3137–3158, 2008

Gauteng Provincial Government, Gauteng Management Strategy, 2014

L. Gueguen, P. Soille, and M. Pesaresi, "A new built-up presence index based on density of corners," in IEEE IGARSS, Munich, Germany, 2012

R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," IEEE Trans. Syst., Man, Cybern., vol. 3, no. 6, pp. 610–621, 1973

International Federation of Surveyors, Rapid Urbanization and Mega Cities: The need for Spatial Information Management, Publication no. 48, 2010

J.M. McPherson, W. Jetz, D.J. Roegers, The effects of species' range sizes on the accuracy of distribution models: Ecological phenomenon or statistical fact? Journal of Applied Ecology, 41, 811-823, 2004

Mudau, N., Mhangara, P., & Gebreselasie, M, Monitoring urban growth around Rustenburg, South Africa, using SPOT 5. South African Journal of Geomatics, 185-196., 2014

M. D. Mura, J. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 10, pp. 3747–3762, 2010

Odindi J, Mhangara P, KakembonV, 'Remote sensing land-cover change in Port Elizabeth during South Africa's democratic transitions', South African Journal of Science, vol. 108, no 5, pp. 1-7, 2012

G. K. Ouzounis, M. Pesaresi, and P. Soille, "Differential area profiles: Decomposition properties and efficient computation," IEEE Trans. Pattern Analy. Machine Intell., vol. 34, no. 8, pp. 1533–1548, 2012

M. Pesaresi, A. Gerhardinger, and F. Kayitakire, "A robust built-up area presence index by anisotropic rotation-invariant textural measure," J. Earth Observ. Applicat., vol. 1, no. 3, pp. 180–192, 2008

M. Pesaresi, D. Ehrlich, I. Caravaggi, M. Kauffmann, and C. Louvrier, "Towards global automatic built-up area recognition using optical VHR imagery," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 4, no. 4, pp. 923–934, 2011a

M. Pesaresi, M. Halkia, and G. Ouzounis, "Quantitative estimation of settlement density and limits based on textural measurements," in Proc. 2011 Joint Urban Remote Sensing Event, JURSE 2011, pp. 89–92, 2011b

M. Pesaresi, G. K. Ouzounis, and L. Gueguen, S. S. Shen and P. E. Lewis, Eds., "A new compact representation of morphological profiles: Report on first massive VHR image processing at the JRC," in Proc. SPIE 8390, Baltimore, MD, USA, Apr. 2012, vol. 8390, p. 8390025, 2012

M. Pesaresi, Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kauffmann, M., Kemper, T., Lu, L., Marin-Herrera, M.A., Ouzounis, G.K., Scavazzon, M.,Soille, P., Syrris, V., Zanchetta, L., 2013. A global human settlement layer from

optical HR/VHR RS data: concept and first results. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 6, 2102–2131, 2013
Statistics South Africa, Census 2011, http://beta2.statssa.gov.za/, 2013
Seto K, Fragkias M, Güneralp B, Reilly M, 'A Meta-Analysis of Global Urban Land Expansion', PlosOne, vol. 6, no 8, pp. 1-9, 2011
Solo, T. , The Bank's Experience in Urban Projects in Africa: An Overview. Washington, DC, The World Bank, 1991
H. Taubenböck, T. Esch, A. Felbier, A. Roth, S. Dech, Pattern-based accuracy assessment of an urban footprint classification using TerraSAR-X data. IEEE Geoscience and Remote Sensing Letters, 8, 278-282, 2011
Tshikotshi V,' The Challenges of Eradicating Informal Settlements in South Africa by 2014-The Case of Seraleng Sustainable Human Settlement, Rustenburg Local Municipality, North West Province', Master's thesis, Faculty of Engineering and Built Environment, University of Witwatersrand, 2009
United Nations Population Fund (UNFPA), 'Linking population, poverty and development; Urbanisation: A majority in Cities', 2007
United Nations, World Urbanisation Prospects: The 2014 revision, 2014
United Nations, Economic Commission for Africa: Contribution to the 2014 United Nations Economic and Social Council (ECOSOC) Integration Segment, 2014
United Nations, The Millennium Development Goals Report 2014, 2014

Annex 1 – Examples of built-up detection

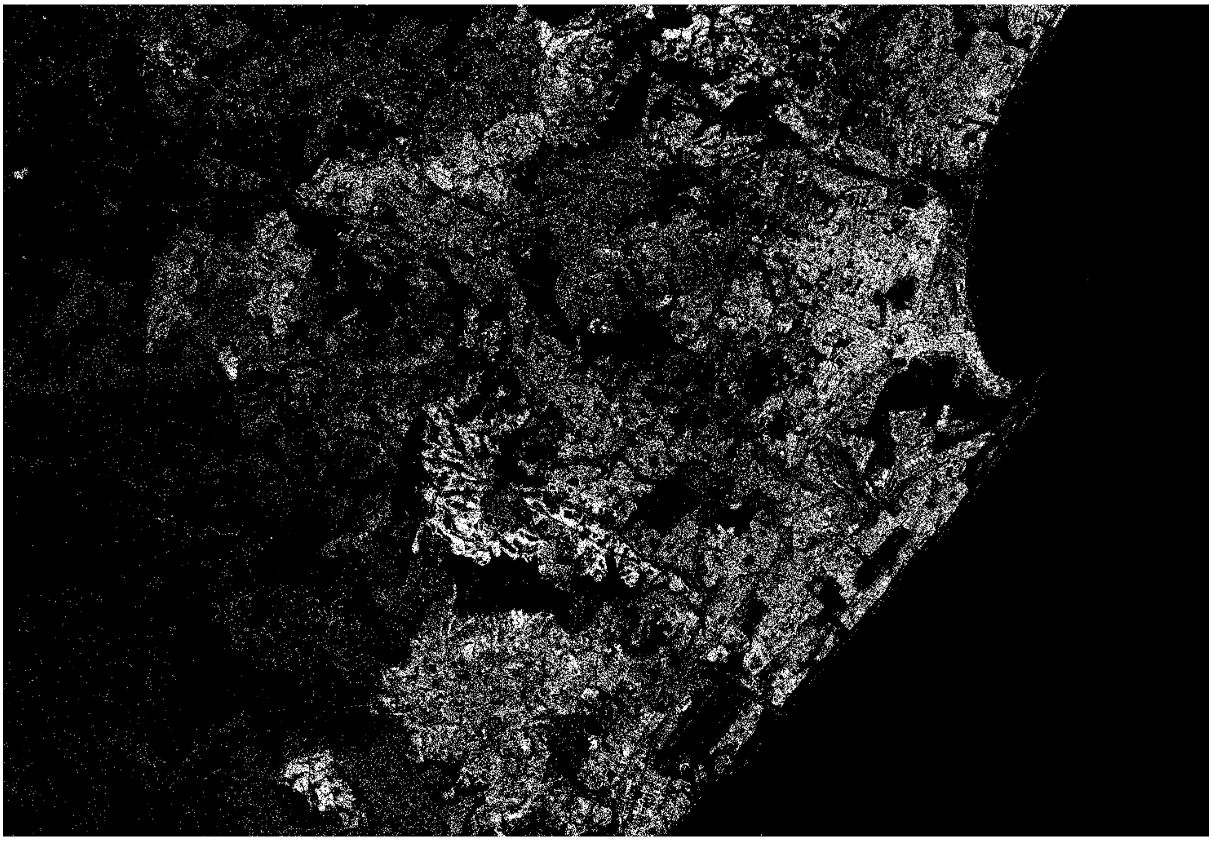

Figure 15. Built-up area map of Durban (Kwazulu Natal)

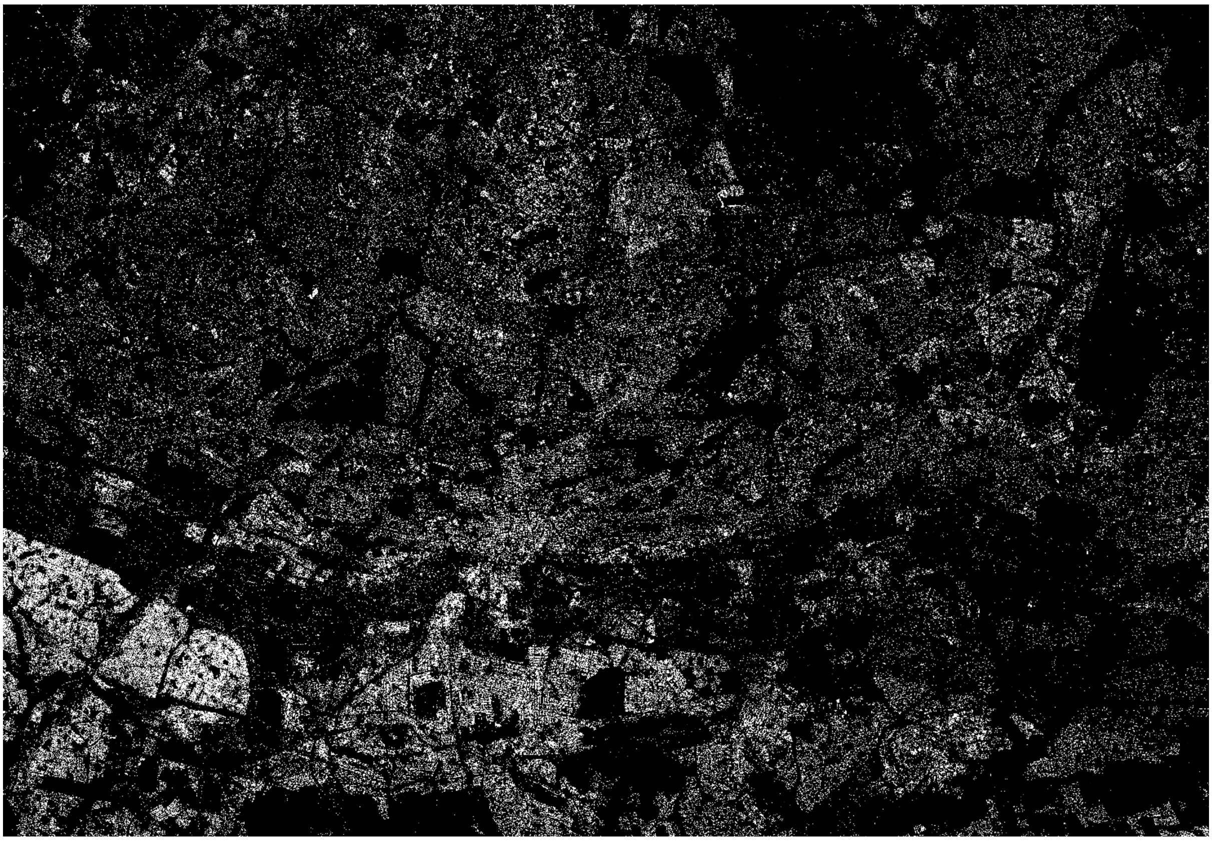

Figure 16. Built-up area map of Johannesburg (Gauteng)

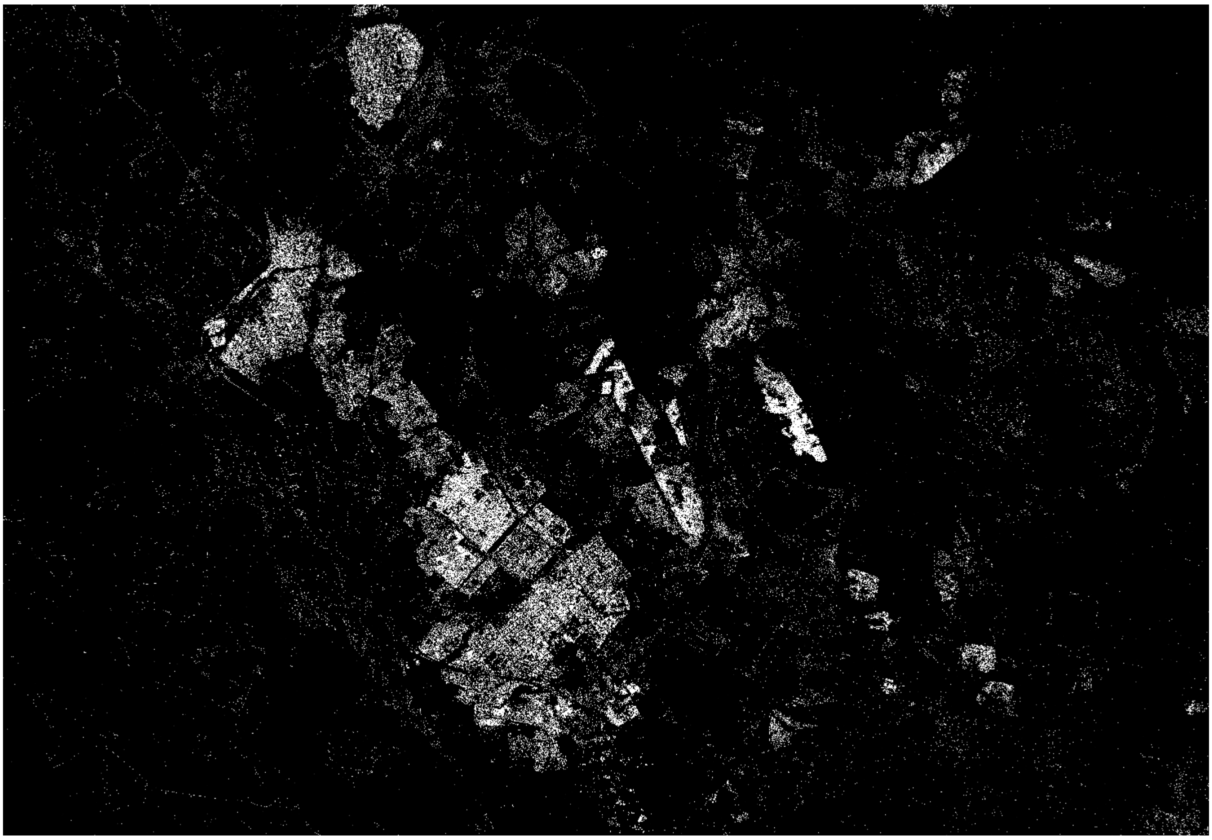

Figure 17. Built-up area map of Rustenburg (North-West Province)

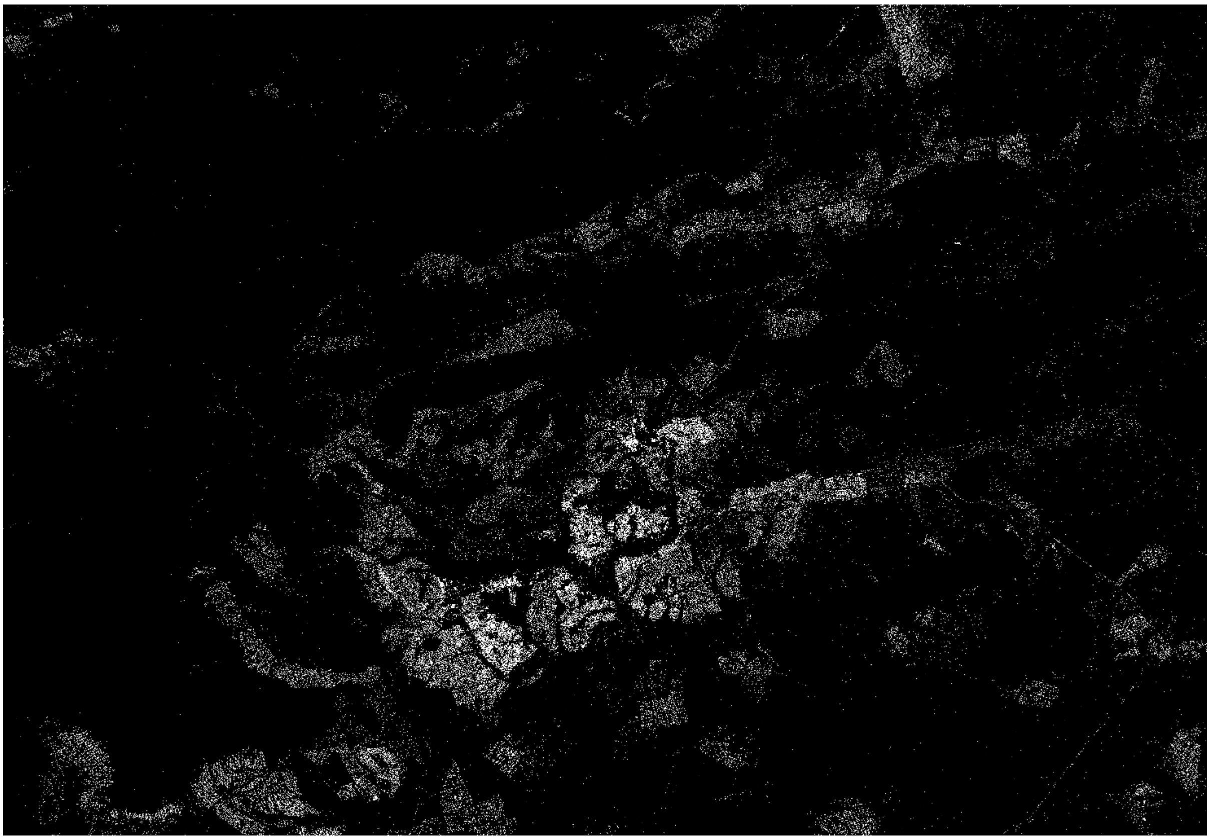

Figure 18. Built-up area map of Thohoyandou (Limpopo)

Annex 2 – National Land Cover legend and adapted GHSL legend

Class Number	Class Name	Recoding
1	Forest (indigenous)	1
2	Woodland (previously termed Forest and Woodland)	2
3	Thicket, Bushland, Bush Clumps, High Fynbos	2
4	Shrubland and Low Fynbos	3
5	Herbland	4
6	Unimproved (natural) Grassland	4
7	Improved Grassland	5
8	Forest Plantations (Eucalyptus spp)	1
9	Forest Plantations (Pine spp)	1
10	Forest Plantations (Acacia spp)	1
11	Forest Plantations (Other / mixed spp)	1
12	Forest Plantations (clearfelled)	6
13	Waterbodies	7
14	Wetlands	9
15	Bare Rock and Soil (natural)	9
16	Bare Rock and Soil (erosion : dongas / gullies)	10
17	Bare Rock and Soil (erosion : sheet)	10
18	Degraded Forest & Woodland	11
19	Degraded Thicket, Bushland, etc	11
20	Degraded Shrubland and Low Fynbos	11
21	Degraded Herbland	11
22	Degraded Unimproved (natural) Grassland	11
23	Cultivated, permanent, commercial, irrigated	5
24	Cultivated, permanent, commercial, dryland	5
25	Cultivated, permanent, commercial, sugarcane	5
26	Cultivated, temporary, commercial, irrigated	5
27	Cultivated, temporary, commercial, dryland	5
28	Cultivated, temporary, subsistence, dryland	5

Class Number	Class Name	Recoding
29	Cultivated, temporary, subsistence, irrigated	5
30	Urban / Built-up (residential)	13
31	Urban / Built-up (rural cluster)	13
32	Urban / Built-up (residential, formal suburbs)	13
33	Urban / Built-up (residential, flatland)	13
34	Urban / Built-up (residential, mixed)	13
35	Urban / Built-up (residential, hostels)	13
36	Urban / Built-up (residential, formal township)	14
37	Urban / Built-up (residential, informal township)	14
38	Urban / Built-up (residential, informal squatter camp)	15
39	Urban / Built-up (smallholdings, woodland)	16
40	Urban / Built-up (smallholdings, thicket, bushland)	16
41	Urban / Built-up (smallholdings, shrubland)	16
42	Urban / Built-up (smallholdings, grassland)	16
43	Urban / Built-up, (commercial, mercantile)	17
44	Urban / Built-up, (commercial, education, health, IT)	17
45	Urban / Built-up, (industrial / transport : heavy)	18
46	Urban / Built-up, (industrial / transport : light)	18
47	Mines & Quarries (underground / subsurface mining)	12
48	Mines & Quarries (surface-based mining)	12
49	Mines & Quarries (mine tailings, waste dumps)	12

European Commission
EUR 27024 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: Towards a country-wide mapping & monitoring of formal and informal settlements in South Africa

Authors: Thomas Kemper, Nale Mudau, Paida Mangara, Martino Pesaresi

Luxembourg: Publications Office of the European Union

2015 – 27 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424

ISBN 978-92-79-44709-9

doi:10.2788/970905

LB-NA-27024-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

doi:10.2788/970905

ISBN 978-92-79-44709-9

Publications Office