**CSIR Satellite Applications Centre**

PO Box 395 Pretoria 0001 South Africa

Tel: +27 12 334 5142

Fax: +27 12 334 5001

E-mail: wluck@csir.co.za

# SPOT 5 NATIONAL MOSAIC
# Production and Accuracy Report
# Eastern South Africa

**PREPARED FOR:**

Spot 5 Open Access Stake Holders

**Compiled by:**

Wolfgang Lück

CSIR SAC: EOSC Technology Group

Technology Manager

Tel: (012) 334 5142

Fax: (012) 334 5001

**Date: 09/08/2007**

# CONTENTS

# 1 PROJECT OVERVIEW

The CSIR Satellite Applications Centre (SAC) has negotiated an Open Access agreement for Spot 2, 4 and 5 with Spot Image with the generous support of South African Government stakeholders. A core deliverable to these stakeholders is an annual generation of a national Spot 5 2.5m true colour mosaic.

For the western part of the country including the provinces of the Northern Cape and Western Cape this task has been outsourced to Infoterra France. The CSIR Satellite Applications Centre has purchased an automated processing chain for the radio- and geometric correction of optical satellite imagery called SARMES to process a mosaic for the remaining areas. This system heavily relies on good image reference data to be fully automated. Such reference data was not completely available for the area concerned, so that the resulting mosaic does not give justice to the quality of the automated processing chain. The CSIR SAC is nonetheless of the opinion that a good product has been delivered of which the unit can be proud of.

This document describes the data characteristics of both the raw images and reference data such as the DEM and urban cadastre. It gives a brief overview of the procedures used during orthorectification, radiometric processing and mosaicking. The format of packaging is described for better transparency. The document finally describes the results from an accuracy assessment and quality assurance point of view. It is aimed at the end user of Spot 5 data. For a detailed description of the procedures followed and quality of the mosaic generated by Infoterra France, the reader is referred to "Meyer B & Lück W 2006. Accuracy Acessment on Spot Image 2.5m Mosaic for Northern Cape and Western Cape Provinces. CSIR / SAC / EOSC unpublished document". This document is supplied with the packaged imagery as the file "SAC_QA_of_Infoterra_N_and_W_Cape_and_DOA_Mosaic.pdf".

# 2 SOURCE DATA

All Spot 5 imagery used for the first national mosaic was programmed and down loaded by Spot Image as part of a virtual reception service to SAC. Imagery was delivered on DVD and portable hard drive in the conventional DIMAP format of Spot Image. This file format consists of a series of metadata files and a XML based header file together with a GeoTiff containing the raw image. All imagery was processed to level 1 A which entails no geometric corrections or resampling and some systematic sensor related radiometric artefact removal.

Imagery was acquired between August 05 and June 06 as illustrated by Figure 1 below. Great seasonal differences between aligning imagery have introduced problems during mosaicking.

Image acquisition was not restricted to low sensor angles so that the viewing angle covers the complete possible range from 0 – 30 degrees. In steep terrain with poor reference imagery and an inaccurate DEM this posed a serious problem. Representation of varying viewing geometry is provided with Figure 2 below. Over 50 % of imagery thus has unfavourable viewing geometry for the production of orthorectified (level 3) Spot 5 imagery to be used for cartographic purposes. New acquisitions as tasked by SAC have restrained acquisition angles to 15 degrees as far as possible.

Only imagery with less than 10 % cloud cover where accepted for the mosaic. Areas along the escarpment and the East Coast may exceed this threshold due to a high frequency of cloud cover in these areas. A MS Word document named “2,5mC_Acquisition Plan2.pdf” provided with the spatial metadata for level 1, was created by Spot Image and gives a brief overview of the image quality used for the mosaic.

Spot scenes are identified by the standard scene ID of Spot Image with the following naming convention:

The Scene ID "511140805120809100081T" Consists of 21 characters with the following meaning:

1 Satellite (Spot 5)

2 - 4 K (Path ID 111)

5 - 7 J (Row ID 408)

8 - 13 Aquisition Date (YYMMDD, 051208)

14 - 19 Aquisition Time (HHMMSS, 091008)

20 Camera (either 1 or 2)

21 Aquisition mode (J multispectral, A or B 5m pan, T 2.5m supermode pan, JT pansharpened multispectral 2.5m)

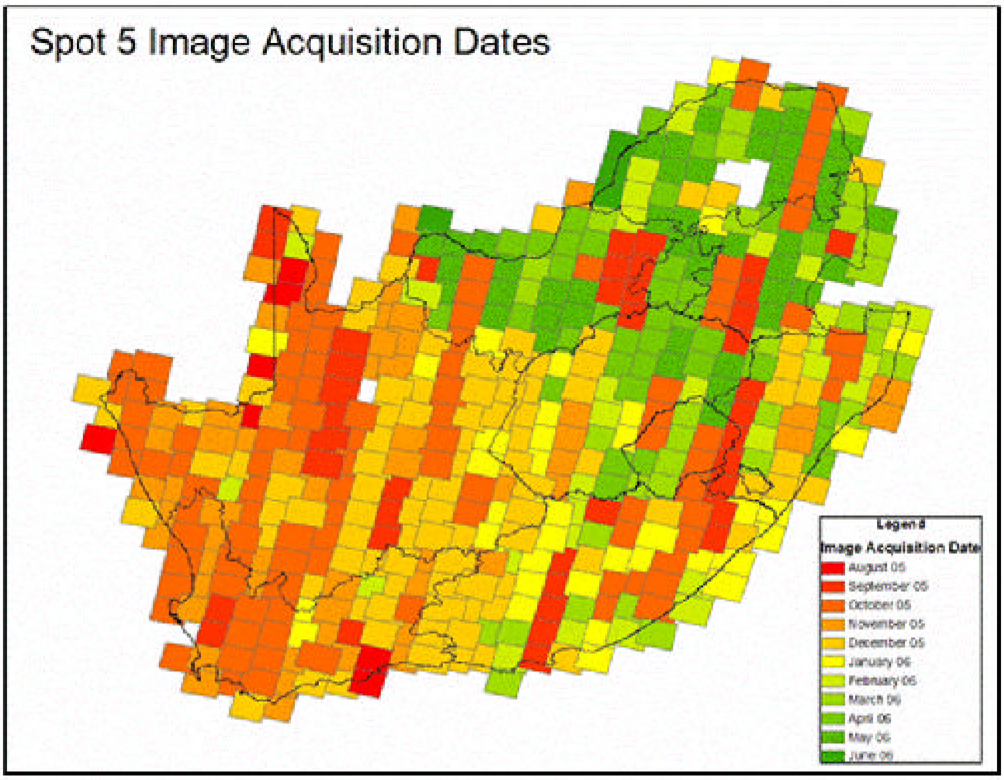


*Figure 1: Variable acquisition dates for Spot 5 imagery acquired between August 2005 and June 2006*

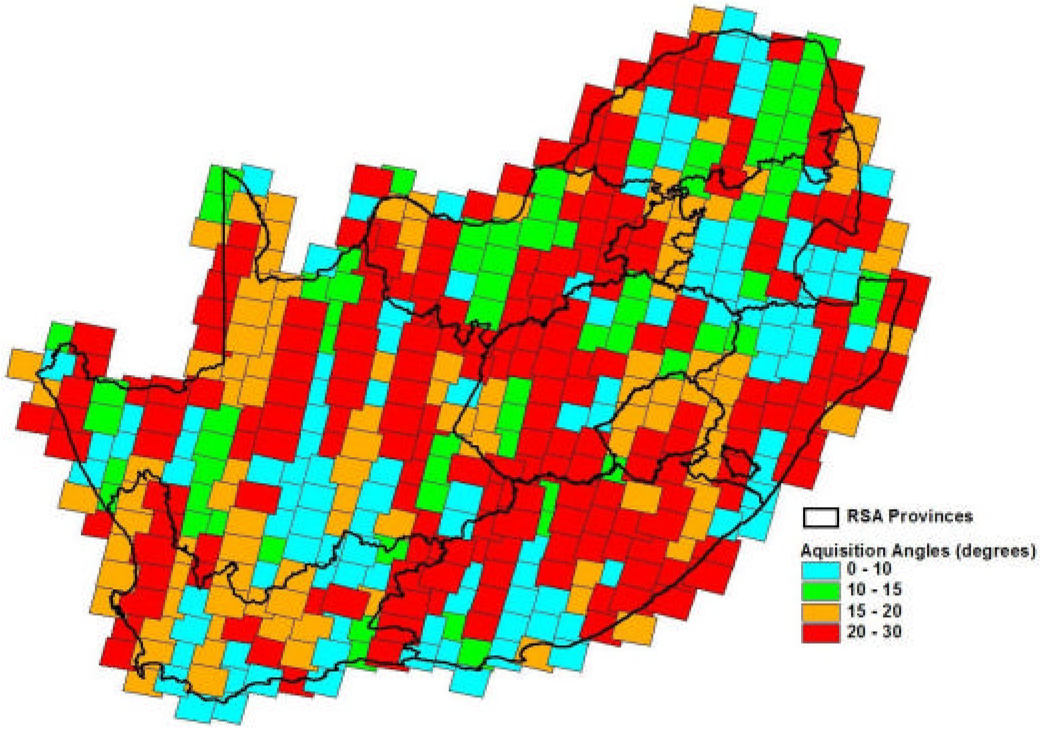


*Figure 2 Variable acquisition angles for imagery used in the national mosaic.*

## 3 REFERENCE DATA

A wide range of reference data is on the marked and could be used for the generation of a national mosaic. Considering that SAC aims at generating a dataset acceptable to all of government, including the strict accuracy standards of CDSM for 1:50 000 mapping, only a limited set of such reference data could be used after a thorough inspection. This consists of:

1. 20m DEM in 16bit supplied by Computamaps interpolated from CDSM 20m vector contours and patched with a combination of SRTM 90m version 1 & 2. This DEM is by no means perfect as it presents elevation in metre intervals (and not floating point) and contains a wide range of artefacts attributed to incorrectly assigned contour heights. SAC is eagerly awaiting a new 25m DEM generated by CDSM which is expected to be of much higher quality and will be used for future mosaics.

2. Orthophotos obtained from the CDSM generated from 1: 30000 and 1:50000 aerial photography. This was the only raster data with a certified accuracy, suitable for completely automatic Ground Control Point collection in the automated processing chain. Imagery where mosaicked by flight campaign, then degraded to 2m, reprojected to UTM and finally added to larger mosaics per UTM zone. The extent of imagery used is illustrated by Figure 3 below.

3. Of the cadastral vector layers available on the market, only the urban "erven" cadastre was found accurate and reliable enough to be used as reference in areas not covered by CDSM orthophotos. Due to the low population densities in many rural areas of South Africa this was not all ways sufficient for this project.

4. True colour reconstruction and colour balancing as a mosaicking step can be rather subjective. SAC chose to rather use physical reflectance measurements as represented by a 16 day image composite obtained from the MODIS sensor to give even the mosaic a scientific backing. This topic will be discussed in greater detail under Section 5 of this report.

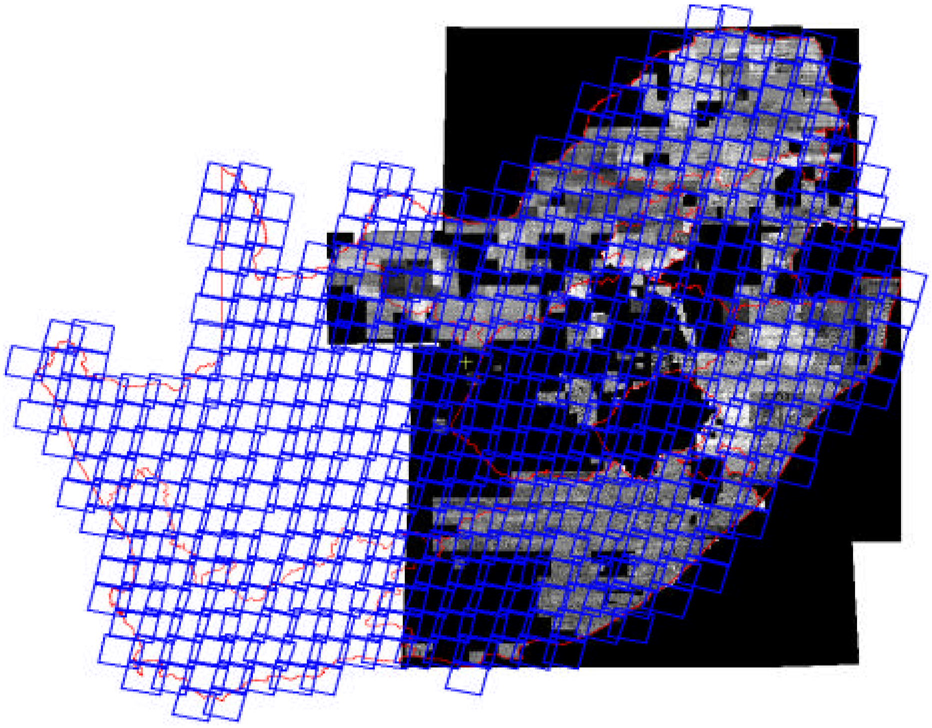

*Figure 3 Orthophotos used as reference for the orthorectification of imagery*

## 4 ORTHORECTIFICATION PROCEDURES

The CSIR / SAC has purchased the SARMES system for future automatic processing of a wide range of satellite imagery. SARMES is based on PCI Geomatica 10 libraries. Where manual user intervention is required, the user interface of PCI Geomatica 10 can be used as part of SARMES.

Imagery was processed in 6 phases. SARMES imported L1A Spot5 imagery in batch mode, automatically naming files according to scene ID as extracted from the header file. During the image import a report file was generated consisting of vector polygons covering scene extents with an attribute table attached containing all the relevant information extracted from the scene header files, required for future radio and geometric processing as well as fields providing information on various processing steps undertaken.

In the case where adjoining imagery was acquired during the same satellite pass, imagery was stitched forming a long image segment consisting of several scenes. In such a case the scene ID from the most northerly scene was used as ID. Long image segments drastically reduce the amount of ground control points required for triangulation, and the amount of files required to be handled.

Ground control points (GCP's) were collected for the panchromatic imagery automatically where orthophotos covered the required area with SARMES, and manually where only the cadastre was available using the PCI Geomatica 10 interface.

Only in areas with poor reference data, tie points were collected between imagery for a stable triangulation. As absolute accuracy is rated higher than relative accuracy, tie points were not collected between imagery with a near complete orthophoto coverage. This implies that although edges of adjoining level 3 products do not necessarily match, the spatial deviation between orthorectified Spot 5 images (level 3) and the corresponding orthophotos is low.

After the orthorectification of the panchromatic imagery, GCP's where automatically collected from the orthorectified panchromatic image for the corresponding multispectral image and this in turn orthorectified using the SARMES interface.

Multispectral and panchromatic imagery was fused producing a multispectral product of 2.5m resolution. As statistical fusion, simulating the point spread function was used, the resulting pansharpened product has very similar spectral properties as the original multispectral image. This is the same algorithm used by Digital Globe for Quick Bird imagery and retains the products suitability for quantitative image analysis.

As the Spot HRV sensor has not got separate blue and a green bands, but one bluegreen band, and most users prefer a true colour composite, the blue and the green band had to be reconstructed. The blue and green band where split using spectral characteristics from the remaining red and infrared bands into consideration.

## 5 MOSAICKING PROCEDURES

Creating a seamless mosaic from imagery acquired over a time period of almost one year is a complex task. The main difficulty being phonological changes of land-cover types over time. Most mosaics are purely cosmetic products, as colour schemes of imagery have been set to that of neighbouring imagery based on areas of overlap. As mosaic's get larger, images tend to loose definition.

SAC tried a rather unconventional approach by matching images to a MODIS MOD09 16 day composite for the period of March 2006. This was a particular good rain season which is the reason for the image being very green. As the MOD09 product is provided in 16bit and the Spot imagery is in 8bit, the Modis product had to be rescaled. This was done by dividing the MODIS image by 10 and recoding values out of the 8bit 0 - 255 range to ether 8bit extreme. Figure 4 below illustrates the MODIS 16day composite, rescaled as mentioned above. From user response SAC has realised that a scale factor of 10 produces too dark an image, which is why a scale factor of 7 will be used in future as illustrated by Figure 5 below.

As a global matching was performed, large clouds present in a mosaic member, lead to the compression of values of the relevant image. This reduces definition and introduces a haze in the mosaic segment covered by the relevant image. A proper cloud masking will eliminate this effect in future.

After the radiometric colour balancing of all imagery in a mosaic, cutlines demarcating the transition from one image to the next were generated automatically using a minimum relative spectral difference algorithm and the mosaic finally assembled.

*Figure 4 Modis MOD09 16 Day composite, rescaled from 16bit to 8bit by division by 10, used for colour matching 2.5m Spot 5 true colour images*

*Figure 5 Modis MOD09 16 Day composite, rescaled from 16 bit to 8 bit by division by 7. ( To be used for future matching.)*

## 6 DATA PACKAGING

Data has been packaged with the convenience of the customer in mind. The GeoTIF format has been chosen for its generic nature and widest support by a wide range of GIS software. This format however has the limitation that images my not exceed 4G in size. In instances where this may occur, they have been clipped. Imagery should be provided with the GeoTIF file only, not being accompanied by any auxillary files generated by third party software (this is particularly relevant to .rrd and .aux files which may cause problems with image import into catalogues and databases). There are two main data products:

1. The level 3 orthorectified images consisting of the 10m multispectral (J), the 2.5 m panchromatic (T) and the 2.5m pansharpened (JT) product. This imagery is projected to UTM and resides in folders named according to UTM zones. In cases where an image lies between UTM zones, it might be found in two folders, of which one version might be cut off at the edge of the DEM used to orthorectify it. Multispectral imagery is ordered according to wavelength from shortest to longest (Bluegreen, Red, NIR, SWIR).

2. The true colour mosaic consisting of 30" tiles named according to equivalent map-sheets. Tiles are of 2.5m resolution with three bands only in a Red, Green and Blue order for greatest display convenience. Although it is seen as very poor practice to provide raster data in "Geographic" projection, mosaic tiles have been projected to "Geographic" due to overwhelming demand by clients. This highlights the need for further fundamental GIS education in our country. Mosaic tiles reside in two major folders split into a Western part of the country, covering the Northern and Western Cape and a second folder covering the remaining provinces. Folders are named by province abbreviations.

A range of ancillary and metadata accompany the imagery. The **"Documentation"** folder contains a range of license agreements that need to be adhered to when using these products. The Metadata folder contains a **"readme"** file with a detailed description of all its contents. In brief a polygon shape file delineating image extents of each panchromatic and multi-spectral file exists with attribute tables containing the essential metadata for each image. A polygon shapefile delineating cutlines used for mosaicking, a point coverage for SPOT scene indexing and a province shape file are provided. Where obtainable, thumbnails of products are provided as georeferenced jpeg files. These can only be displayed in advanced GIS / RS software.

## 7 QUALITY ASSURANCE

Many remote sensing data providers use the RMS error calculation during the triangulation from GCP's as a measurement of spatial accuracy for a final orthorectified product. This method is very convenient but incorrect as it mearly describes how well the GCP's fit a model irrespective if that model is correct or not. In order to get a true reflection of spatial accuracy, a new set of GCP's need to be collected between the orthorectified product and the reference data. The difference between input and reference coordinates then gives an indication of spatial inaccuracies of the product. A dedicated SAC technician collected up to 5 GCP's per image evenly distributed over the image where reference data was available. This data has been statistically analysed and linked spatially to the scene ID shape file. Figure 6 below illustrates the mean error calculated per scene, while Figure 7 illustrates the error at 2 sigma which is the mean plus twice the standard deviation of the error. The error at 2 sigma gives the certainty that at a 95% confidance interval all points within the image are closer than the given value to their true position. From the two figures we can assume that most images have met the goal requirements of 12m at 2 sigma, but that further improvements need to be made in certain instances to generate an accurate dataset for future reference in automatic processing chains. Main factors for inaccuracies have been mentioned previously, and a new aquisiton of Spot5 imagery with an acquisition angle restrained to 15 degrees as well as better reference data is being accumulated. That Chief Directorate of Surveys and Mapping have offered to assist with reference points for the next mosaic in areas where no reference could be obtained in the previous round.

Accuracy results per scene are appended in the polygon shape-file and should give the user an indication on the spatial trustworthiness of each scene or parts of the mosaic. Values as provided are conservative estimates as interpreter errors during QA GCP collection where not subtracted from the published values.

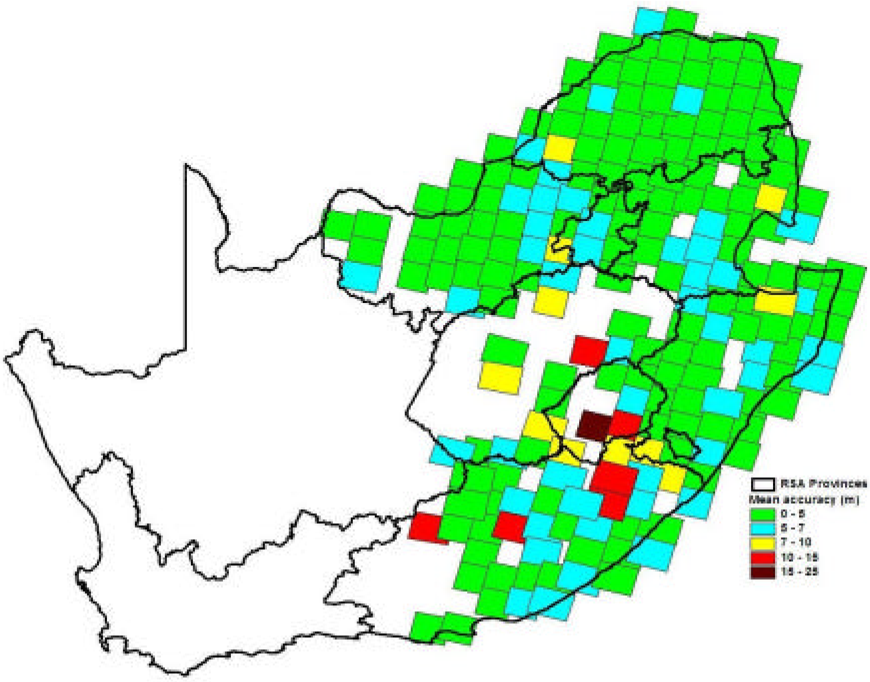


*Figure 6: Accuracy checked L3 imagery. Mean accuracy values in metres, compared to reference data.*

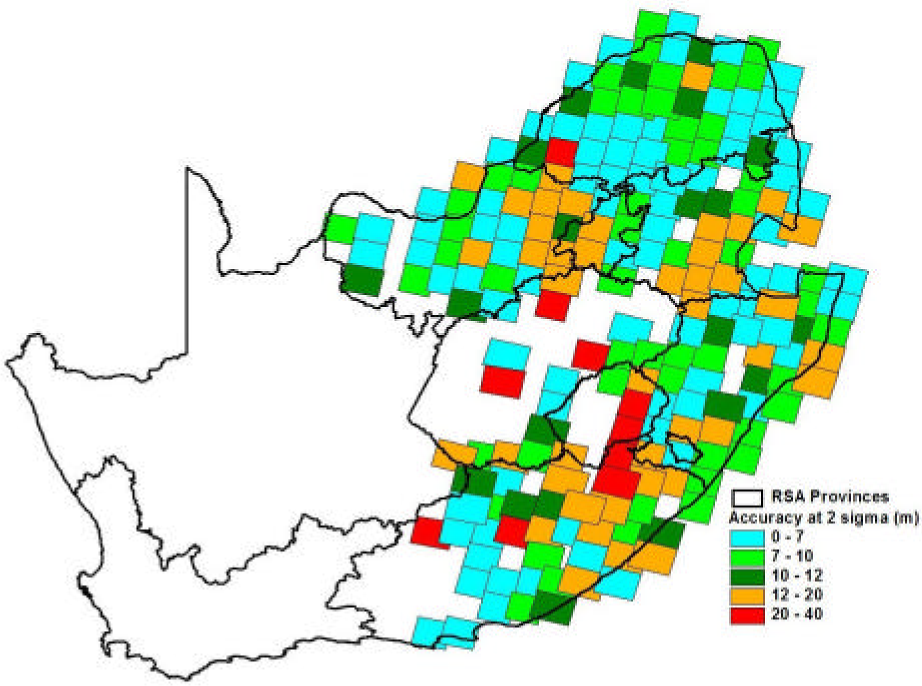


*Figure 7: Accuracy checked L3 imagery. Accuracy of imagery at 2 Sigma (95% confidence interval).*