| | | |
|---|---|---|
| magellium | **L1 GTC MSS Product Geometric VALIDATION** | |
| Ref:<br>IDEAS-TN-05-L5_SupportToDSI-TM_DataValidation | Landsat MSS DSI products<br>SLAP SW V3.03 | Date: 07/07/2014<br>Issue/Rev: 1/1 |

# 1 Subject

The report covers the geometric validation of products outputs from DSI bulk processing project. The objective of this activity has been to validate the multi mission (Landsat1-5) accuracy of MSS data from the MSS Stress test.

# 2 Executive summary

This document, version 1.0, reports the last validation results obtained over sample from MSS stress test of SLAP V3.03 products. Note that it has been difficult to find data from different sensors observed over the same region of interest.

| Validation Item | Comment |
|---|---|
| Multi Temporal geolocation Accuracy (L1T PCD/DEF Product) | The computed accuracy is 1.5 pixels (RMS). No geometric problem has been found.<br>The bias in easting direction is under 1 pixel (from -47 m to 2 m).<br>The bias in northing direction is under half pixel (from -25 m to 9 m)<br>After filtering*, the multi temporal CE90 is 111m (under 2 pixels). |
| Internal Geometry | According to the specification, the internal geometry is correct but is varying depending on the sensor / mission. |
| Miscellaneous | Geolocation anomaly discovered on L1G product and is to be investigated. |

- Executive summary.

*For explanation on the method, in particular filtering process, refer to section below.

# 3 Input data

Different TDS are considered and those are coming from KIRUNA, and FUCINO Archive. We have considered the MSS Stress test as a dataset. The input data covers Landsat 1-5 (all Landsat mission with the MSS sensor). The Worldwide Reference system (WRS) 1 is used for Landsat 1-3 while WRS2 is assigned to Landsat 4-5.

**MSS Stress Test:** Data selected are over the south west of Norway which is an interesting site for Landsat1-5 due to the number of isles and coasts (small temporal changes). The Landsat 1-3 data observed over the WRS1 SCENE 201/18 and WRS2 SCENE 201/18 for Landsat 4-5 have been selected. The reference data is the landsat 5 product. It leads to an assessment of the relative geolocation accuracy along the missions.

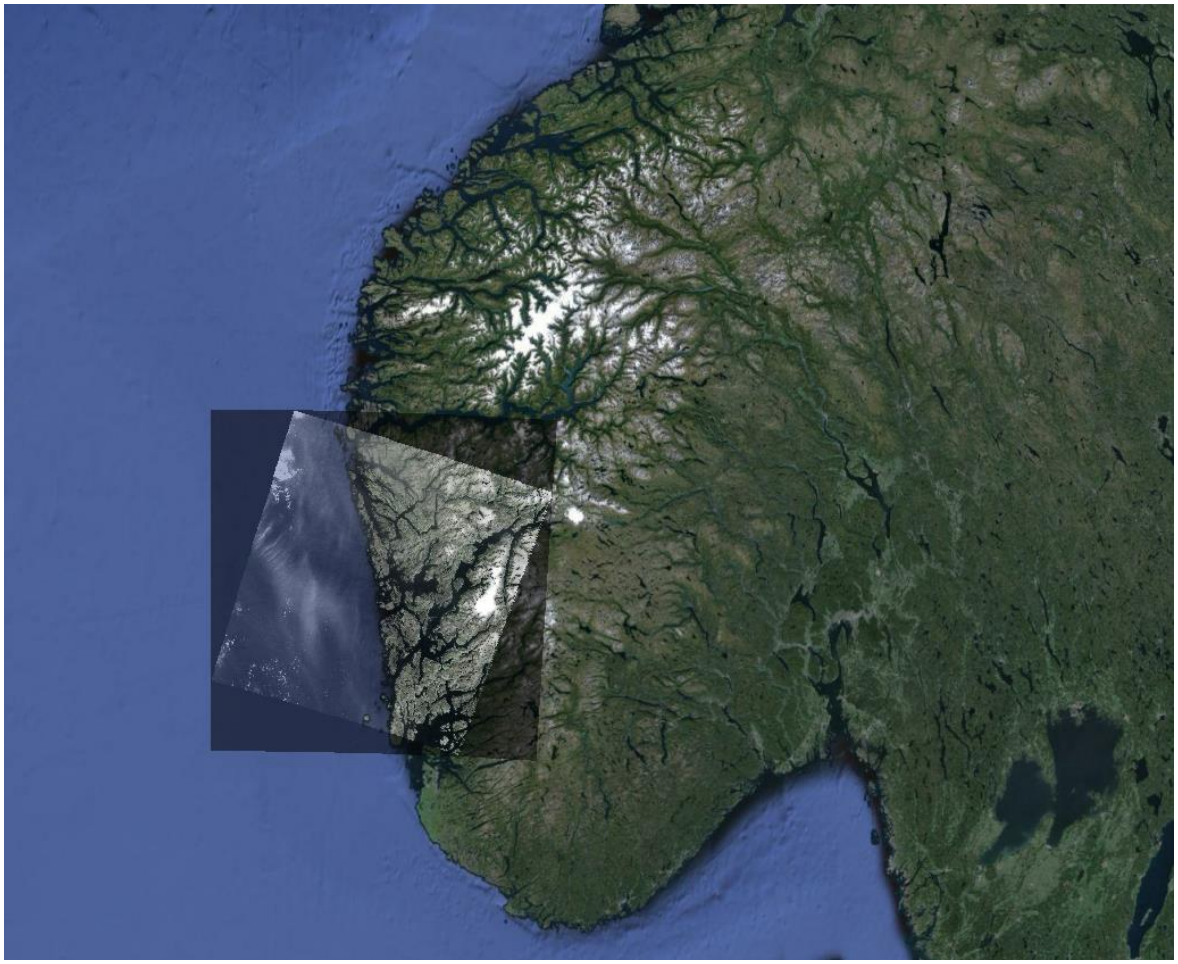


*fig 1. A Landsat scene over Norway Area, Geometric Test site.*

# 4 Analysis Methodology

The multi temporal geo-positional accuracy is associated to data observed over the same region of interest with Landsat1-5 data involved. The accuracy of each product is assessed by considering as reference either one product from the dataset or an external raster reference.

As outputs of the study there are different categories of information,

- Overall Statistics associated to the dataset
- Statistics associated to each products
- Internal geometry analysis.

The standard statistics, applicable at point/product/dataset level used to report the geopositionnal accuracy are defined as follow;

- The standard deviation
- The Root Mean Square Error (RMSE) in one direction
- Two dimensional RMSE
- The Circular Error probable error at 90 percentile (CE90) or 95 percentile (CE95)

The methodology to validate the multi temporal registration accuracy is not straightforward; in particular because several products, observed at different locations over the Earth, according to different configuration, are needed in order to obtain a robust estimate of this quality parameter. For cost effective approach, it is therefore more convenient to automatize process by using image matching technics. The success of image matching depends on matching factors, which can be summarized as follow;

- The quality of the images (noise, blurring);
- The seasonal variation and meteorological / atmospheric properties,
- The properties of the terrain, relief, surface reflectance, and information content,
- The similarities of spectral bands (in case of different cross comparison),
- The scene content (frequency content).

In addition, the results of image matching should be carefully checked and filtered because a significant number of GCPs can be retrieved from image matching, but kept depending on the matching confidence and on the spatial distribution. Actually, the accuracy should be associated to the full extent of the product footprint. It should not include an over contribution of errors located in a certain part of the image. Aside,

when addressing the multi-temporal case, it sounds important to consider, for a same area of interest, the same GCP, whilst the spatial content of an image is changing with time.

Therefore; the proposed methodology is basically broken down into three mains stages; the dense matching process between an input images and a reference images, the filtering of the correlation grid and the accuracy analysis.

The following main outputs are given and the document is organized accordingly.

- Overall statistics, this is a summary of the product accuracy to be communicated to the general audience. The statistics are based on accuracy obtained for each product. The data stack, as input, has been filtered in a way that selected results are relevant for the GCP covering the full extent of the scene and a multi temporal GCP set is considered.
- Circular Error(CE) at 90 percentile is computed at point level, considering GCP sample, and deducing multi temporal accuracy from this sample. It is somehow more informative approach, and there is no accuracy specification output from this approach.
- Circular Error (CE) at 90 percentile is computed at Product level, considering sample before and after filtering process. A figure is provided showing the CE90 and the error distribution. The accuracy specification is computed based on the filtered sample data.
- Multi-temporal analysis of geolocation accuracy reports accuracy metric depending on the observation date, report in fractional day of year. The graphic plot includes the accuracy report in the Landsat product metadata file as well.
- Internal geometry section of the images is dedicated to the analysis of image matching results, in particular by computing statistics on line / pixel displacement profiles.
- In depth analysis of some products is dedicated to products that have been, in most cases, discarded from statistics the overall statistics because of quality issues.

# 5 MSS Stress Test dataset – Results

## 5.1 Overall Statistics

The results listed in table below are **mean value over statistics** from products that belongs to the filtered dataset the Circular Error at 90 percentile for the dataset is given below.

The multi temporal geolocation accuracy is appreciated according to several accuracy items. For each item, specific comment has been added. The overall results are very correct. Note that geometric distortions have been observed and they are discussed below.

| Accuracy Item | Value | Comment |
|---|---|---|
| Error Mean (Easting) | -26.07 m | Easting Centring accuracy is correct, no systematic effect detected. The centring is better for images from Landsat 1 and Landsat 3 |
| Error Mean (Northing) | -2.94 m | Northing Centring accuracy is correct, no systematic effect detected. |
| Std (Easting) | 53.64 m | The mean precision is about one pixel, from 46 m to 63 m. |
| Std (Northing) | 65.71 m | The mean precision is about one pixel, from 53 m to 80 m. |
| RMSE (Easting) | 62.71 m | Major contributor to the mean RMSE value is the precision. |
| RMSE (Northing) | 68.32 m | Major contributor to the mean RMSE value is the precision. |
| RMSE 2D | 92.90 m | The quadratic sum of RMSE in both directions does not exceed 90m (1,5 pixel). |
| Circular Error (Empirical)* | 111.85 m | The circular error at 90 percentile is ranging from 78m to 140 m. The mean value is within specification (120 m). |
| % points within Circle* | 94% | If normal distribution is assumed, the percentage of points within CE90 circle deduced from RMSE accuracy values is about |

| Accuracy Item | Value | Comment |
|---|---|---|
| | | 90%. |
| % points within Ellipse* | 78% | In case of normal distribution, 68% of points should be within 1 sigma ellipse. The 78% indicated a major part of points are highly concentrated around the mean values. |

*Table 1. Overall Statistics (4 products).*

## 5.2 Filtered points

The image below represents the good distribution of the filtered points (definition in the chapter 4).

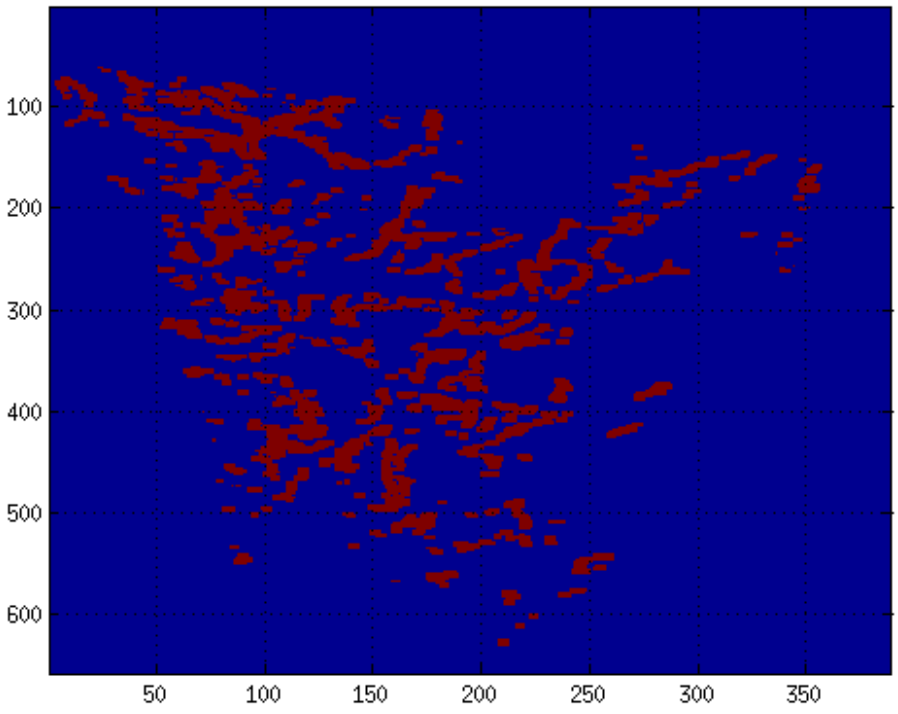



*fig 2. Distribution of the filtered points.*

## 5.3 Circular error at Point level

Figure below shows all the points from multi temporal dataset. Each product has its own contribution to the overall errors; the image below shows directional components that are actually strongly related to the input product.

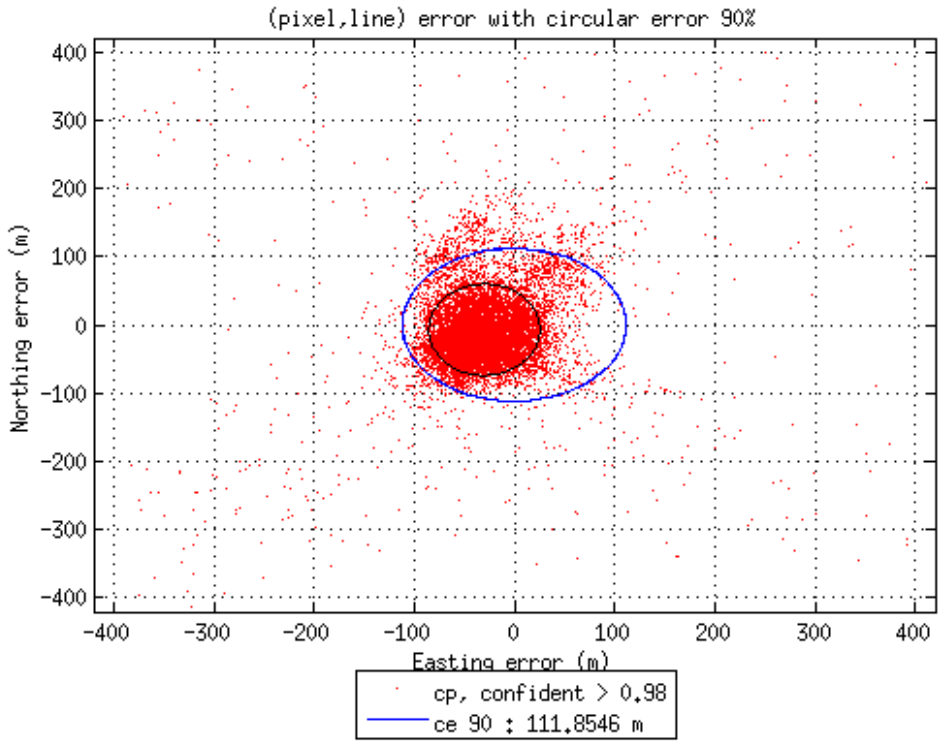



*fig 3. Product Circular Error @90 Plot – (Filtered sample), the black ellipse, is 1d sigma error ellipse (normal distribution hypothesis).*

## 5.4 Circular error at Product level

When applying filtering; with at least 90% of spatial coverage over scene footprint extent and same GCP across multi temporal images, the **CE90 accuracy of the dataset is about 49 m**, which is extremely correct.

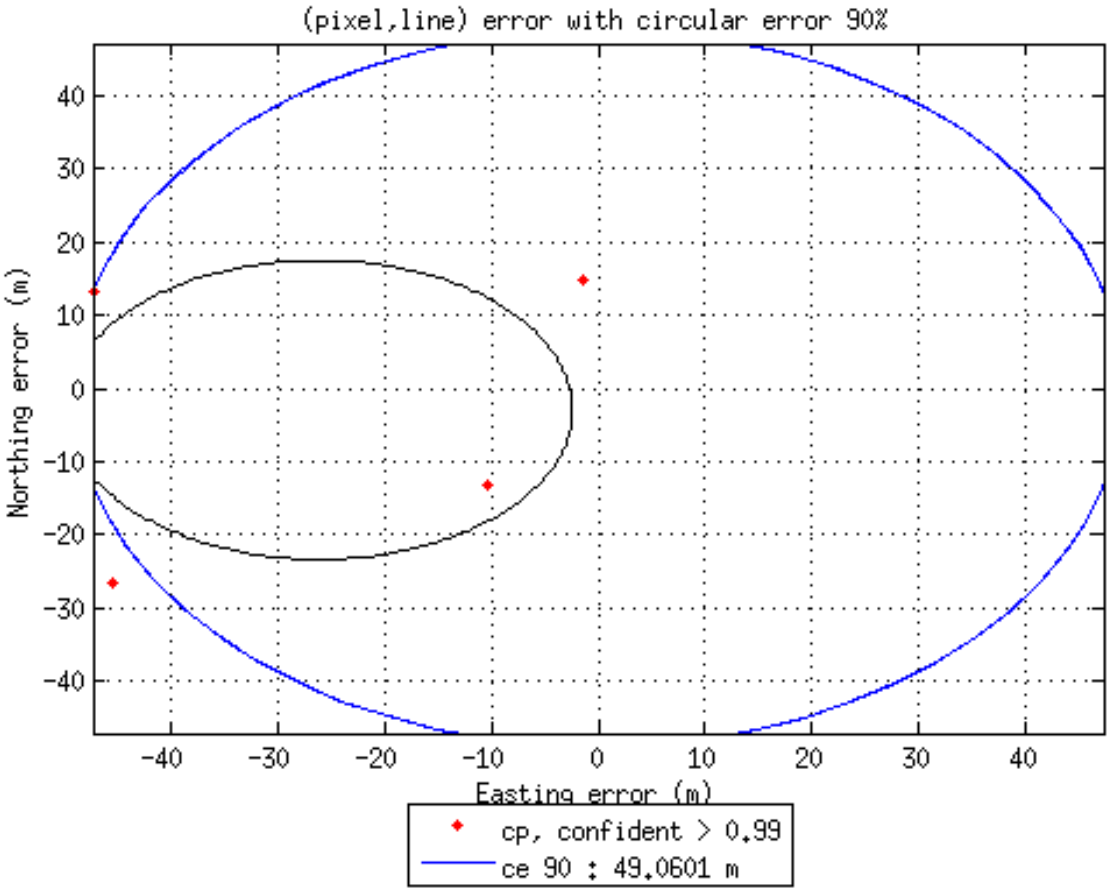



***fig 4. Product Circular Error @90 Plot – (Filtered sample), the black ellipse, is 1d sigma error ellipse (normal distribution hypothesis).***

## 5.5 Multi-temporal analysis

Figure below shows the temporal evolution of geolocation accuracy together with the accuracy extracted from product metadata file (*rmx_mtl, rmsy_mtl*). Herein, **the input sample includes data from different missions, L1-4** (L5 is the reference data). The rms values in the easting direction are between 48 meters and 70 meters. There is a bias observed about 1 pixel, it is mainly due to the contribution of the precision.

The accuracy values extracted from metadata file are mainly below the accuracy obtained from this validation exercise and do not follow the same behavior.

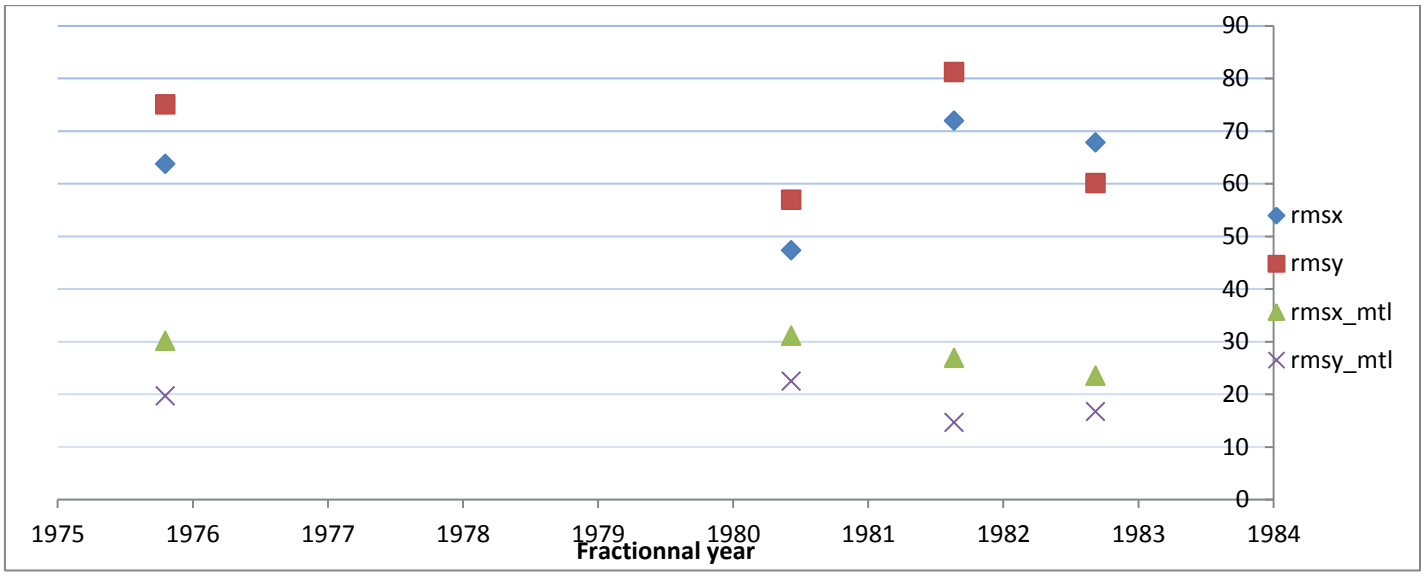
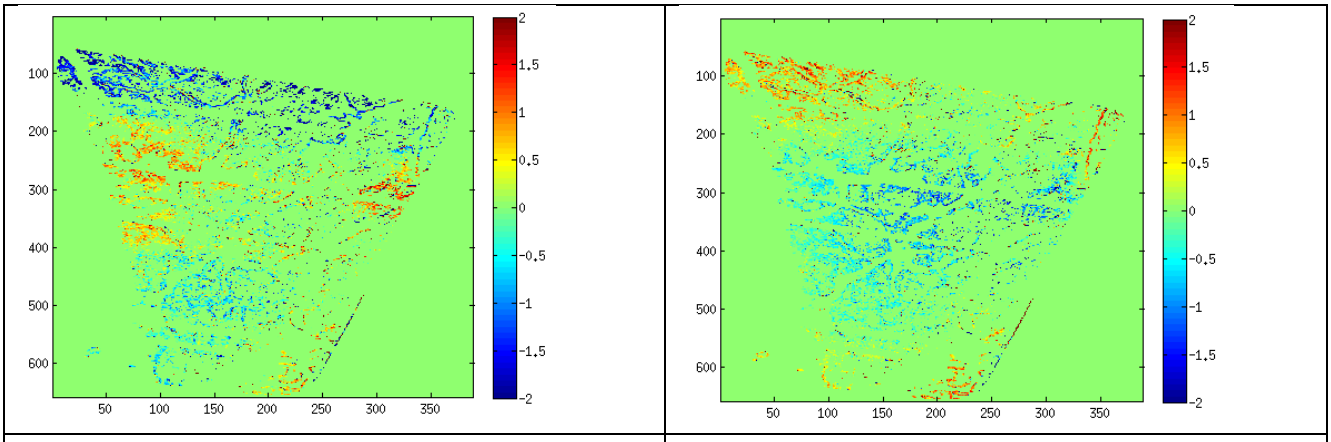



*fig 5. Multi Temporal Plot (raw sample, without filtering applied).*

## 5.6 Internal geometry

Density maps from matching process are inspected and analyzed for detecting geometric anomaly and validating ortho corrections applied. This process has been done for all products from the dataset. The image below is an example of density maps (Landsat 1 images).




| Line displacement | Pixel displacement |
|---|---|

*fig 6. Results of the correlation: Line and pixel displacement (Landsat1 and Landsat5 as reference)*

Image deformation is observed. The "mean row & mean column" analysis are done to evaluate the magnitude of the deformations. Images bellows shows profiles, the variations does not exceed 2 pixels. The variations are nonetheless very different depending on the product.

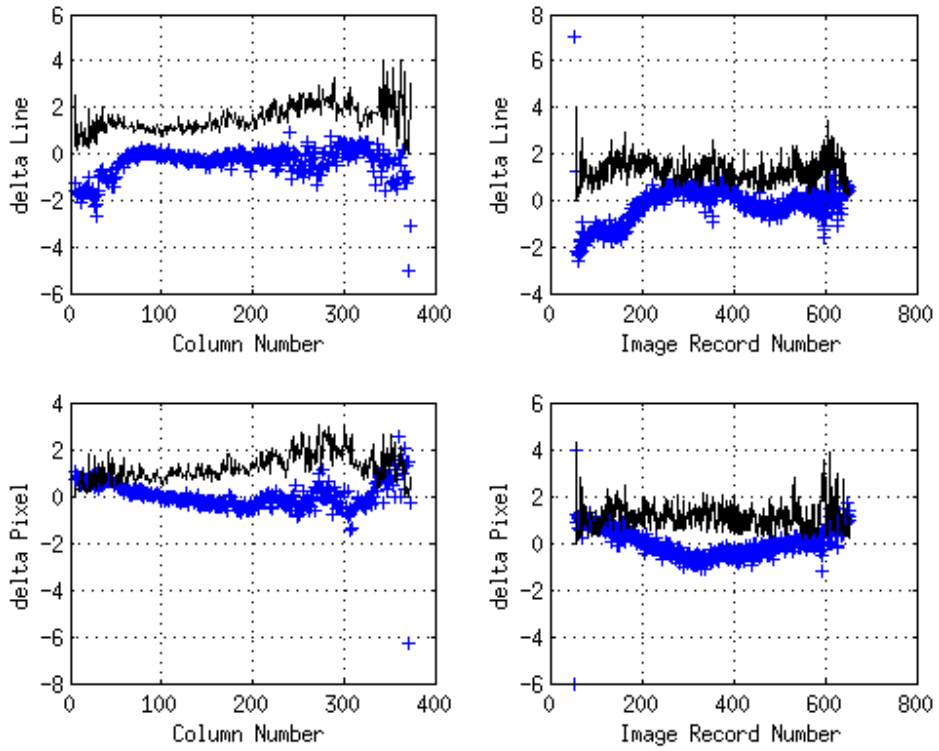



*fig 7. Delta line / pixel against column / image record number (Landsat1 and Landsat5 as reference)*

## 5.7 In depth analysis

The MSS Stress data analysis has showed some Landsat 3 products are misplaced (at least 2). It should be carefully monitor. For example the following product is affected by a corruption of geolocation information.
LS03_RKSE_MSS_GEO_1P_19800620T080618_19800620T080648_011684_0196_0032_BFF9

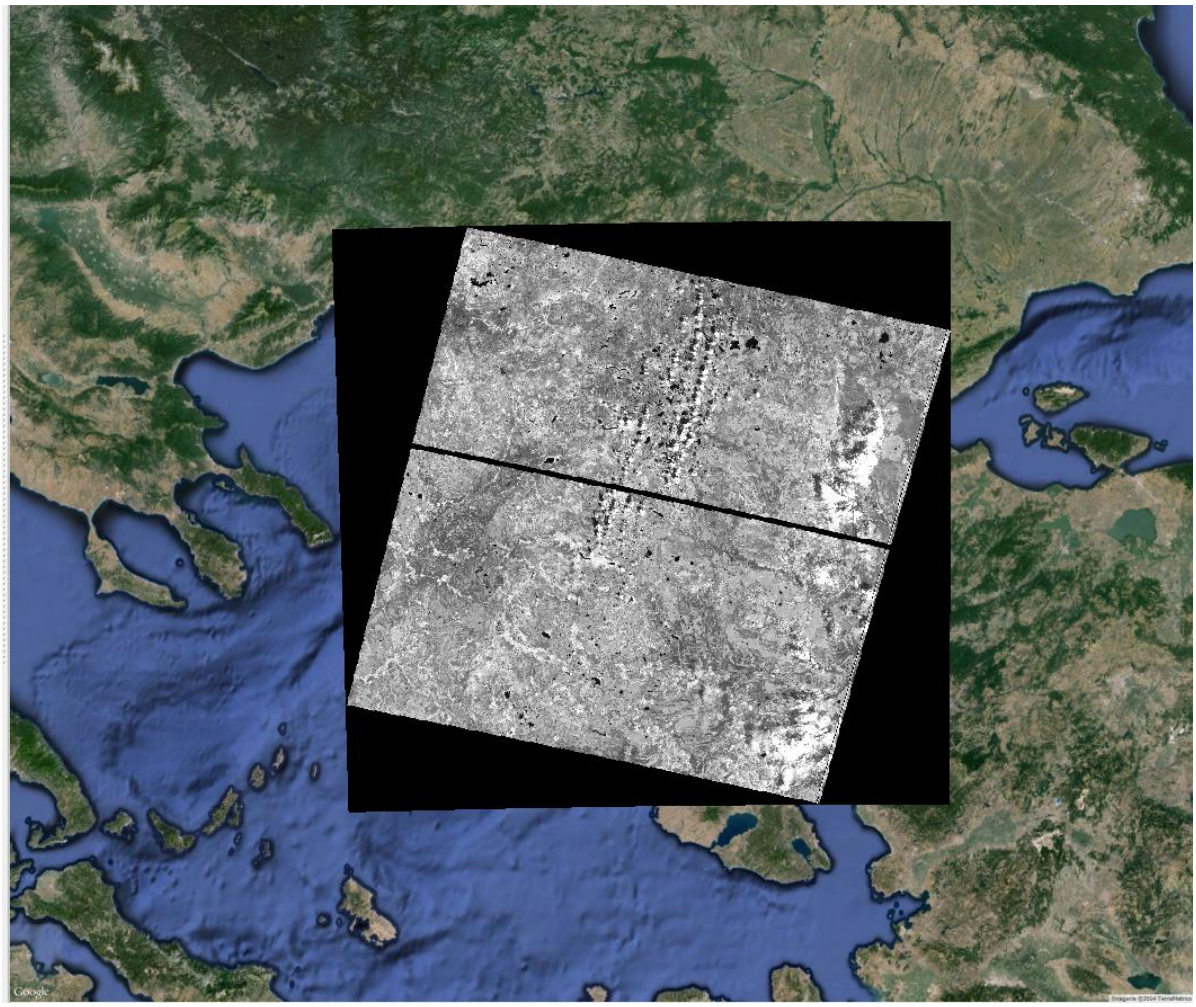


*fig 8. The Lansdat 3 image (band 6) overlays part of google raster.*

*fig 9. By applying image transparency, the mis registration is evident.*

Root causes of this anomaly are still within investigations. It might be due to the processing, refer to DSI report regarding the analysis of Stress Test Working directory.