

Planimetric Accuracy in Satellite Mapping

Bruce Sharpe, Senior Research Analyst
Kelly Wiebe, Research Analyst
MacDonald Dettwiler,
3751 Shell Road, Richmond B.C.
Canada V6X 2Z9
Commission No. IV

Abstract

The requirements for planimetric accuracy of satellite image correction are discussed in the context of base mapping accuracy standards. Measurements are presented for SPOT and the Landsat Thematic Mapper. Two-dimensional accuracies as low as 7.2 m (TM) and 6.4 m (SPOT) were measured, these being close to the level of noise in the measurement techniques. NATO class A 1:50 000 scale standards were achieved after only 1 (SPOT) or 2 (TM) control points were used in the models.

1 Introduction

Requirements for the accuracy of geometric correction of satellite imagery vary depending on the application. In the case of satellite mapping, the requirements can be obtained from base mapping accuracy standards. In this paper we shall discuss the planimetric accuracy of Landsat Thematic Mapper (TM) and SPOT image correction in the light of those standards.

The cost effectiveness of mapping from space over conventional photogrammetric techniques derives in part from the reduction in the number of ground control points (GCPs) required [2]. Thus we shall pay particular attention to the amount of ground control needed to achieve a given level of accuracy.

In Section 2 we discuss the planimetric accuracy standards used for conventional mapping. This is followed in Section 3 by a discussion of the factors to be considered in geometrically correcting satellite imagery and the way that the MacDonald Dettwiler's Geocoded Image Correction System (GICS) performs these corrections. Sections 4-6 present the methodology and results of measurements of GICS accuracy for TM and SPOT data.

2 Accuracy and Ground Control Requirements

The accuracy standards for NATO class A maps require that 90% of all well-defined points tested be accurate to within 0.5 mm. A similar level of accuracy has been adopted in the U.S. for the National Map Accuracy Standards and in Canada for class A maps. The accuracy standard as stated is a circular map accuracy standard (CMAS).

It is somewhat more common to measure a mean squared error (MSE). If we assume that the errors have a Gaussian distribution with mean zero then the CMAS and MSE accuracies are related by

$$\text{CMAS} = 1.5174 \text{ MSE}$$

The accuracy levels for various scales are given in Table 1.

Scale	CMAS (m)	MSE (m)
1:50 000	25.0	16.5
1:25 000	12.5	8.2
1:20 000	10.0	6.6
1:15 000	7.5	4.9

Table 1: Circular map accuracy standard (CMAS) and mean square error (MSE) values for NATO Class A maps.

Note that these accuracies are two-dimensional. It is common in the literature to report accuracies separately for x and y map directions. These accuracies are typically the same and so are related to a MSE by a factor of $\sqrt{2}$:

$$x \text{ or } y \text{ error} = \text{MSE}/\sqrt{2}$$

In this paper all accuracy figures are two-dimensional unless explicitly stated otherwise.

3 Geometric Correction of Satellite Imagery

Geometric correction is a two stage process. In the first stage, the correspondence between any given pixel in the input imagery and a point on the earth's surface must be established. In the second stage, the input imagery is resampled to a regular grid, for example, a map grid.

There are a large number of factors which must be taken into account when determining this correspondence. The sensor geometry can be quite complicated, especially for the TM, where the following sensor characteristics affect the placement of image pixels on the ground: offsets between bands, offsets between detectors within bands, nonlinear motion of the scanning mirror, motion of the Scan Line Corrector mirror, and others.

Also important are the motion of the satellite along its orbit and changes in the satellite attitude (orientation) during image acquisition. The final ingredients are the shape and location of the earth's surface and rotation of the earth.

GICS is the image processing software originally developed for the Canada Center for Remote Sensing for the ground station in Prince Albert, Saskatchewan [1], with subsequent versions installed in or in preparation for the European Space Agency, Australia Center for Remote Sensing, India National Remote Sensing Agency, Thailand National

Research Council, General Electric's Space Division and TRW. Similar algorithms form part of MacDonald Dettwiler's Meridian image analysis system.

GICS takes the following approach to geometric correction. The input-ground correspondence is embodied in four models

- sensor
- orbit
- attitude
- earth

These models are developed from manufacturer's data, prelaunch measurements, an understanding of the physical processes involved and *a priori* knowledge. Those model parameters that are not known *a priori* are determined from telemetry data and GCP measurements.

Models that are determined without using ground truth are called *systematic*. They typically have errors on the order of 1 km due to uncertainties in absolute orbit position. However they have a fairly high internal accuracy. Models that are determined with the use of ground control are called *precision*. Errors in the location of GCPs are used as input to a Kalman filter to refine the models.

GICS produces full-scene images in the satellite orientation using systematic models (bulk products) and geocoded images (that is, with map orientation and with map boundaries) with either systematic or precision models.

4 Methodology

The approach used to measure accuracy starts with the marking of well-defined features (check points) in the imagery to determine their input coordinates (line and pixel in the input image). The input coordinates are transformed to ground location using the models, either systematic or precision. The ground locations thus determined are then compared with ground truth, that is, the locations as determined independently, for example, from existing maps or from field surveys.

Some of the marked points are used to determine the model (model points), while the remainder (check points) are used as independent points at which to measure accuracy.

In all our measurements, the height of the check point is used to compensate terrain effects when calculating the accuracy at that point. This is equivalent to incorporating a high-resolution digital elevation model (DEM) into the modelling process. This is a procedure that promises to become available on a routine basis now that DEMs can be automatically derived from SPOT imagery [3].

In the absence of any height information (DEM or even average height), image points are placed as though they lay on a nominal earth surface, usually taken to be an ellipsoid.

This can lead to substantial errors in areas of high terrain relief. For TM data, the error is up to 0.1 times the height; for SPOT, it can be up to 0.5 times the height. Removing these errors allows us to examine the accuracy of the remainder of the modelling.

In this study we investigated the accuracy of the models by using transformations based on them rather than measuring corrected scenes. There is a slight difference between the two methods, because in order to obtain high throughput in a production system, approximations to the exact transformations are used. These approximations are controlled to limit the error they introduce and can be made more accurate by trading off memory usage and performance. For example, in the MOSAICS implementation of GICS, the error is limited to 0.1 pixels RMS. This error combines with the transformation error in an RMS fashion. Thus, for example, an error of 0.5 pixels as determined by the transformations would be less than $\sqrt{0.5^2 + 0.1^2} = 0.51$ pixels in the corrected scene.

4.1 Types of Accuracy

Several different kinds of accuracy were measured for each of the TM and SPOT scenes studied. We now summarize the definitions of these accuracies. Errors are often resolved into one-dimensional components (X and Y), where the dimensions are, for example, map (x, y) or along-track and across-track directions. Let X_i and Y_i denote the measured error at the i^{th} check point and N the number of check points.

The *absolute accuracy* (AA) is the mean square error

$$\text{AA} = \text{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i^2 + Y_i^2)}$$

Relative accuracy (RA) is a measure of internal accuracy. It is based on the difference between true and measured distances between all pairs of check points:

$$\text{RA} = \sqrt{\frac{1}{N(N-1)} \sum_{i \neq j} ((X_i - X_j)^2 + (Y_i - Y_j)^2)}$$

Note that the sum of squared errors is divided by one-half the number of pairs to arrive at a value reflecting the error at one point rather than at both in the pair. It can be shown that relative accuracy is equal to the standard deviation of the errors, and so is a measure of what the absolute accuracy would be if any average errors were removed.

Scale accuracy (SA) is another measure of internal accuracy. It is similar to relative accuracy except that the error in the distance between points is normalized by the true distance (in a given map projection):

$$\text{SA} = \sqrt{\frac{2}{N(N-1)} \sum_{i \neq j} \frac{(X_i - X_j)^2 + (Y_i - Y_j)^2}{D_{ij}^2}}$$

where D_{ij} is the true distance between points i and j .

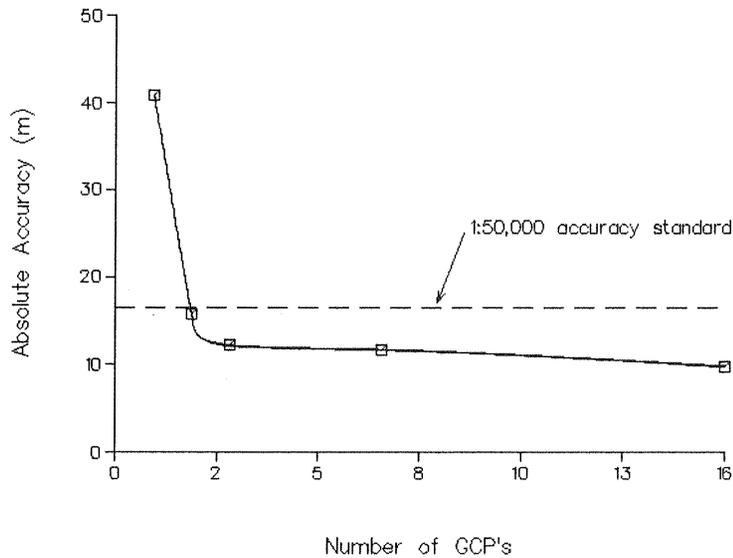


Figure 1: Measured accuracy of precision corrected TM imagery (Iowa 1).

Multitemporal accuracy is a measure of the difference between two scenes acquired at different times. If $X_{i,1}$ and $X_{i,2}$ are the X -coordinates of check point i as determined in scenes 1 and 2 respectively, (and similarly for Y) then the multitemporal accuracy is

$$MA = \sqrt{\frac{1}{N} \sum_{i=1}^N ((X_{i,1} - X_{i,2})^2 + (Y_{i,1} - Y_{i,2})^2)}$$

5 Measurements: TM

Three data sets were used to study TM precision modelling accuracy. They are described in Table 2.

Location	WRS	Date	Number of GCPs
Iowa 1	P26R31	July 23/84	143
Iowa 2	P26R31	November 26/84	143
Minnesota	P28R29	April 16/8	22

Table 2: Data sets for TM accuracy measurements.

The ground truth locations of the GCPs were obtained by digitizing USGS 1:24 000 7.5 minute quadrangles. The results for Iowa 1 are shown in Figure 1. There were 80 check points used to measure the accuracy. In all cases the model points were chosen to have a good distribution throughout the scene. The errors (AA) ranged from 40.9 m with one GCP in the model down to 9.7 m with 16 model GCPs.

The Minnesota scene was studied after Iowa 1, and its higher accuracy (shown in Table 3) reflects our increased expertise in selecting and marking GCPs accurately.

Number of model points	Number of check points	Accuracy (m) (AA)
10	12	7.2

Table 3: Absolute accuracy of precision corrected Minnesota imagery (TM).

The Iowa 2 scene was used to measure multitemporal accuracy. The models for both Iowa scenes were determined using the same GCPs. Rather than comparing each scene to the reference ground truth and then comparing the errors to each other, image chips from the two scenes were correlated to determine the displacement between them. Although correlation was attempted on several chips, the severe seasonal variation between the scenes caused the correlation to fail in all but 17 cases. Band 4 was used for the correlations. The accuracy is shown in Table 4.

Number of model points	Number of check points	Accuracy (m) (MA)
10	17	7.1

Table 4: Multitemporal accuracy of precision corrected Iowa 1 and Iowa 2 imagery.

6 Measurements: SPOT

The SPOT data studied consisted of a panchromatic stereo pair of images of an area near Nottingham, England. They are described in Table 5. The imagery was processed from CNES level 1A tapes. The ground truth was obtained by digitizing British Ordnance Survey 1:10 000 maps. This work was performed as part of a study conducted by the Mapping and Charting Establishment of Great Britain.

Name	Date	Sensor	Off-nadir angle	Number of GCPs
Nott1	April 17, 1987	HRV1 P	L20.9°	84
Nott2	April 24, 1987	HRV1 P	R20.0°	68

Table 5: Data sets for SPOT accuracy measurements.

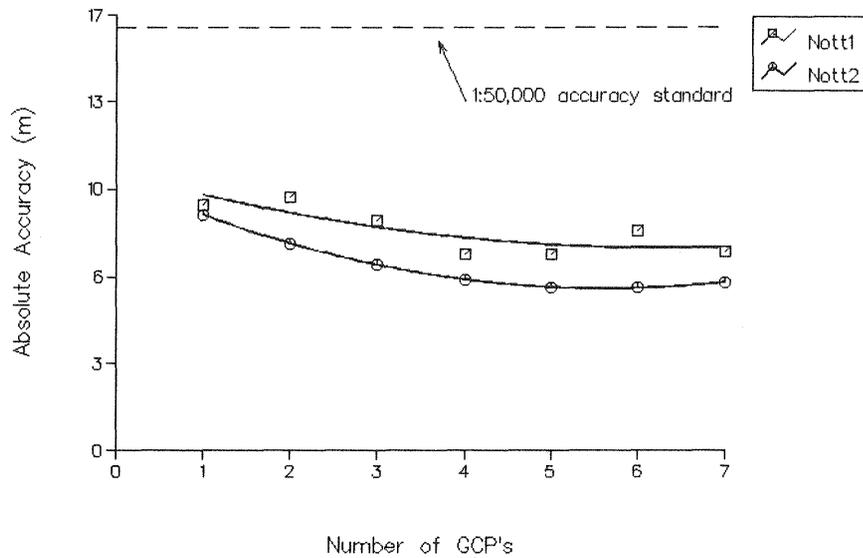


Figure 2: Absolute accuracy of precision corrected SPOT imagery. Points marked with squares are Nott1 accuracies, those marked with circles are Nott2 accuracies.

The absolute accuracy measurements were performed for different numbers of model GCPs and different sets of model GCPs for each number. Each set of model points had a good distribution over the scene. The resulting accuracies were then averaged. They are shown in Figure 2. All points not used in the model were used as check points. The accuracies ranged from 9.9 m to 7.7 m for Nott1 and 9.2 m to 6.4 m for Nott2.

The Nott1 scene was hazy and had low contrast. Consequently, it was more difficult to mark GCPs there and the lower accuracy is a result.

The accuracy of the systematic models was also measured. The results are shown in Table 6.

Scene	Absolute Accuracy (m) AA	Relative Accuracy (m) RA	Scale Accuracy
Nott1	444.1	9.8	1.1×10^{-3}
Nott2	746.0	6.9	0.9×10^{-3}

Table 6: Accuracy of systematic corrections of SPOT imagery.

In a final test, the two scenes were corrected using 9 model GCPs each and the heights of 45 check points were deduced by making use of the stereoscopic view obtained with the two scenes. The RMS error of the heights thus obtained was 6.9 m when compared with the heights obtained from the maps.

7 Inferred Accuracy

The level of measured error is determined by three things: the accuracy of the models, the accuracy of the reference ground truth, and the accuracy of marking features in the imagery. These errors can be expected to be independent and so combine in an RMS way. That is,

$$E_{\text{measured}} = \sqrt{E_{\text{model}}^2 + E_{\text{reference}}^2 + E_{\text{marking}}^2}$$

Of these error sources, the model error is of the most interest and if the other two dominate, as we shall show is the case, the magnitude of the first can only be inferred.

For the TM data, the USGS maps conformed to the NMAS which state that 90% of the features are accurate to within 1/50 inch. This corresponds to an MSE of 8.0 m. We estimated the marking error to be about 0.25 pixels or 7.5 m. The combined map and marking error come to 11.0 m. This is greater than the lowest measured error and so does not tell us much about the best accuracy that can be achieved.

More revealing is the multitemporal accuracy. If we make the reasonable assumptions that errors in the two scenes are independent, have mean zero and the same magnitude, then we deduce that

$$\begin{aligned} \text{AA} &= \text{RA}/\sqrt{2} \\ &= 7.1 \text{ m}/\sqrt{2} \\ &= 5.0 \text{ m} \quad (\text{inferred}) \end{aligned}$$

For SPOT, the map errors were estimated in the following way. Fifteen GCPs with a location accuracy of better than 1 m were available in a small region of the scene. Their locations were compared with the positions on the maps leading to an estimate of 4.9 m RMS for the map error.

The marking accuracy was estimated by having different operators mark the same points and measuring the difference in the marked coordinates. This led to an estimate of 3.0 m RMS.

The combined map and marking error is thus 5.7 m. For the best accuracies measured (6.4 m) this suggests a model error of

$$\text{AA} = \sqrt{6.4^2 - 5.7^2} = 2.9 \text{ m} \quad (\text{inferred})$$

Alternatively, we can use the accuracy of the height determinations from the stereo pair to infer a planimetric accuracy. It can be shown that an across track error of E leads to a height error of

$$\frac{E}{\tan \theta_1 + \tan \theta_2}$$

where θ_1 and θ_2 are the incidence angles of the two scenes. If we assume that the along-track error has the same magnitude as the across-track error and that the errors in the two models are independent, then we can infer that

$$\text{AA} = E_{\text{height}}(\tan \theta_1 + \tan \theta_2)$$

where E_{height} is the error of the deduced height. If we compensate the measured height error (6.9 m) for the error of the reference heights, which we estimated to be 3–4 m, we can estimate

$$E_{\text{height}} = \sqrt{6.9^2 - 3.5^2} = 6.0 \text{ m}$$

so that

$$\begin{aligned} \text{AA} &= 6.0 \times 0.75 \\ &= 4.5 \text{ m} \quad (\text{inferred}) \end{aligned}$$

8 Conclusions

We have measured two-dimensional accuracies as low as 7.2 m for TM and 6.4 m for SPOT. The corresponding one-dimensional accuracies are 5.1 m for TM and 4.5 m for SPOT. In all cases, the best accuracies are close to the noise level of our measurement techniques.

The actual error is less than the measured error because the latter includes map and marking inaccuracies. This was illustrated by the fact that the inferred accuracies were as low as 5.0 m for TM and 2.9–4.5 m for SPOT. The lower inferred errors also shows how the Kalman filter can reduce errors in the ground truth, resulting in products that are more accurate than the ground truth.

The best measured accuracies were better than the map accuracy standards for 1:25 000 scale (TM) and 1:20 000 scale (SPOT). The accuracy standard for 1:50 000 scale was measured for both TM and SPOT imagery after only 1 (SPOT) or 2 (TM) GCPs were used in the model. This compares favorably with conventional photogrammetric techniques, where about 20 GCPs are required to create a typical 1:50 000 scale mapsheet. A SPOT scene covers a ground area of about four 1:50 000 mapsheets; a TM scene covers a ground area of about 36 mapsheets.

These results show that satellite imagery has the potential to substantially reduce the cost of accurate base mapping by reducing the requirements for ground control.

References

- [1] J.P. Friedel and T.A. Fisher. MOSAICS—a system to produce state-of-the-art satellite imagery for resource managers. *Geocarto International*, 2(3):5, September 1987.
- [2] R. Swann, D. Hawkins, A. Westwell-Roper, and Wm. Johnstone. The potential for automated mapping from geocoded digital image data. *Photogrammetric Engineering and Remote Sensing*, LIV(2):187, February 1988.
- [3] R. Swann and D. Kauffman. Results of automated digital elevation model generation from SPOT satellite data. In *these proceedings*.