SUMMARY OF INDEPENDENT EVALUATIONS OF STAR-3i DEMs

1. Foreword

This document presents a summary of the published results of tests performed by several independent organizations to establish the vertical accuracy of DEMs created by Intermap's STAR-3i system. The tests were conducted over different areas and under various conditions. The accompanying table summarizes the test characteristics and results. Additionally there is a brief verbal description of each of the test situations.

Several factors affect the results of any test. Some of the more important factors are discussed in the following section.

1.1 Relevant Test Factors:

- Flight Altititude Above Ground Level (AGL): Although the image resolution and the DEM sample spacing do not change as a result of the STAR-3i flight altitude, the vertical accuracyof the DEM does depend upon the radar height above ground. The errors observed in data from lower altitude flights (eg 20,000' AGL) are expected to be somewhat smaller than those from flights at higher altitudes (eg 40,000' AGL). This is in fact observed in the results tabulated below.

- Comparative 'Truth' Data: Elevation data from STAR-3i are compared with ground data representing an approximation to 'truth'. In order to approximate 'truth' the ground data must be at a level of accuracy significantly better than that expected from STAR-3i. Ideally, the data would be in the form of a raster DEM of the same sample spacing as STAR-3i. Alternatively, the ground data may consist of a well-distributed collection of individual points obtained with GPS or other means. These are good for statistical evaluation, but small systematic effects may remain undetected.

- Terrain Cover: The effect of terrain cover such as vegetation, buildings, etc, is reflected in the STAR-3i DEM. For purposes of accuracy assessment the tests should be performed on 'bald earth' terrain – that is, the terrain being comparatively tested should be free of vegetation or other objects that could distort the analysis.

- Terrain Slope: Vertical accuracy is somewhat dependant on slope, so it is useful to group results with respect to terrain slope. Errors increase slowly with increasing slope.

- Size of Test Area: Some sources of error may be slowly varying in time; therefore larger test areas are desireable in order to include such effects. Ideally, test areas should cover strips several km long.

- DEM Spacing: At sample spacings greater than 5 meters the elevation noise will be reduced somewhat as a result of averaging.

- Accuracy Reporting: Accuracy reporting is generally done with statistical measures; typically the mean offset and standard deviation (σ), or alternatively the Root Mean Square Error (RMSE). The mean offset generally represents residual systematic error which can be largely removed with ground control if available. The standard deviation (σ) represents mostly random noise, approximately normally distributed about the mean offset. A 90% confidence interval corresponds to 1.65 σ, while 95% confidence interval corresponds to 2 σ. The RMSE represents an overall absolute error and can be calculated from the mean offset and σ. Confidence intervals cannot be calculated directly from RMSE. In general σ corresponds to the relative or in-scene accuracy, while the RMSE corresponds to the absolute accuracy over the test area.

1.2 Organizations that have conducted evaluations of the STAR-3i data:

- NASA Stennis (Stennis Space Center), Commercial Remote Sensing Program
  - Author: Gregory Terrie
  - Report : STAR-3 DEM Accuracy Evaluation (Part of a Proprietary Document)
  - Date: June, 1998

- Institute of Navigation, University of Stuttgart, Germany
  - Authors: Professor A. Kleusberg and Dr. H.G. Klaedtke
  - Report: Accuracy Assessment for the STAR-3i Derived DHM in Baden-Wurttemburg
  - Date: May, 1998

- US Army Topographical Engineering Center (TEC), Data BaseTechnology Division, Alexandria, Virginia
  - Author: F. Raye Norvelle
  - Report: Evaluation of ERIM's IFSARE (1) Digital Elevation Models of Cholame Hills and Camp Roberts in Supprt of GeoSAR
  - Date: December, 1996

- Decision-Science Applications / DARPA, Arlington, Virginia
  - Authors: Robert Carlisle and Mark Davis
  - Report: GeoSAR Program: IFSAR Validation and Terrain Classification from Polarimetry
  - Date: 1995

Notes:
(1) STAR-3i was developed by ERIM under the name IFSARE

2. Summaries of Evaluations

The following summaries provide an overview of the tests that have been conducted and the results of the tests. The details and results are compiled together for convenience of cross-referencing in section 3. below.

2.1 NASA - Stennis Space Center

The STAR-3i data were acquired in December 1997 in the Walnut Gulch area of Arizona. An evaluation was performed by NASA Stennis of the Upper San Pedro Watershed Sub-Basin area and reported in July, 1998. The test area was 2.9 km x 1.3 km in size (this was all area for which an existing government DEM overlapped the STAR-3i coverage). The area is characterized by mountainous terrain of about 600 meters maximum relief. The data were collected from the aircraft platform at an altitude of <u>24,000 feet AGL</u> and processed to <u>10 meter postings</u>. Ground truth were in the form of a DEM supplied by the USDA–ARS Watershed Research Center. This DEM was posted at 15 meters and was specified at 0.4 meters vertical accuracy. In the evaluation process, the USDA DEM was re-sampled to 10 meter posts to correspond to the STAR-3i DEM. The surfaces were differenced and the results are summarized as follows:

Relative Accuracy σ = 0.8 meters (Mean Offset = 0.2 meters)

Absolute Accuracy RMSE = 0.8 meters

2.2 Institute for Navigation (INS), University of Stuttgart, Stuttgart, Germany

STAR-3i data were collected near Freiburg, Germany in September, 1997 and an evaluation was conducted by INS and reported in May, 1998. The test area was about 9 km x 18 km in size. The data were collected at an altitude of <u>20,000 feet AGL</u> and processed to <u>2.5 meter postings</u>. Ground control was acquired by INS using several different methods of varying accuracy. The most accurate and unambiguous was a collection of 44 points on bald terrain using Differential GPS (DGPS). About half of these points were collected on sloped terrain, the others being well distributed over flat terrain. A separate set of 78 points were taken from the existing, trigonometrically-derived, geodetic control network.

The DGPS results for all terrain classes showed:

Relative Accuracy σ = 0.9 meters (Mean Offset = - 1.3 meters)

Absolute Accuracy RMSE = 1.6 meters

2.3 US Army Topographical Engineering Center (TEC)

This was the first, and also the most extensive and rigorous of all the external tests performed. The data were acquired in September, 1995 and the report, authored by Ray Norvelle of TEC, was dated December, 1996. Two study areas in California were tested (Camp Roberts and Cholame Hills) each about 10 km x 16 km in area and incorporating mosaics of three and two swaths of STAR-3i data , respectively. The aircraft altitude was 39,000' AGL, and the data were processed to 5 meter postings in these tests.

Ground truth data were obtained in the form of a photogrametrically-derived DEM of 0.5 m (RMSE) vertical accuracy specification and 5 meter posts. One of the areas was relatively flat, the other largely hilly. Trees were relatively sparse and masked out of the 'bald-earth' analysis. The photo-DEM was rigorously checked and anomalous data also masked out as were zero-fills in the radar DEMs. An analysis of error as a function of grazing angle (which incorporates both slope and look-angle) showed that σ, the standard deviation, tends to increase with large positive or negative slopes. In this summary, all data are included, which comprises some six million points for each study area.

Camp Roberts Test Area:

Relative Accuracy σ = 1.5 meters (Mean Offset = - 1.4 meters)

Absolute Accuracy RMSE = 2.1 meters

Cholame Hills Test Area:

Relative Accuracy σ = 1.7 meters (Mean Offset = - 1.4 meters)

Absolute Accuracy RMSE = 2.2 meters

2.4 Decision-Science Applications / DARPA

This was a study performed on the same data set (Cholame Hills) as described in section 3. above, but by different authors and with a different set of objectives. The basic findings were identical with those reported in the TEC document.

3. Evaluation Summary Table:

Acq'n Date	Report Date	Test Region	Terrain Type	Flight Alt.(AGL)	DEM Spacing	Size of Test Area (kmxKm)	Author	Truth	No. of Pts.	s (m)	Mean Offset (m)	RMSE (m)
Sept,'95	Dec,'96	Cholame Hills	Mixed[1]	39,000'	5m	10 x 16	TEC[2,5]	Photo DEM	6M	**1.7**	**-1.4**	**2.2**
Sept,'95	Dec,'96	Camp Roberts	Mixed	39,000'	5m	16 x 10	TEC	Photo DEM	6M	**1.5**	**-1.4**	**2.1**
Sept,'97	May '98	Freiburg	Mixed	20,000'	2.5m	8 x 18	INS[3]	GPS	44	**0.9**	**-1.3**	**1.6**
			Flat	*20,000'*	*2.5m*	*9 x 18*	*INS*	*GPS*	*20*	*0.7*	*-1.1*	*1.3*
			Sloped	*20,000'*	*2.5m*	*7 x 20*	*INS*	*GPS*	*24*	*1.0*	*-1.4*	*1.7*
			Flat	*20,000'*	*2.5m*	*8 x 20*	*INS*	*Trig Pts*	*78*	*1.3*	*-0.9*	*1.5*
Dec'97	June '98	Walnut Gulch	Mountains	24,000'	10m	2.9 x 1.3	NASA[4]	Photo DEM	40K	**0.8**	**0.2**	**0.8**

Notes:

1	Mixed terrain refers to a mixture of flat and moderately sloped terrain, with slopes up to +/- 35 degrees
2	TEC is the US Army Topographic Engineering Center - Ray Norvelle was report author.
	A separate report on this data set was written by Carling and Davis, with similar results
3	INS is the Institute of Navigation at the University of Stuttgart, Germany - A. Kleusberg and H.G. Klaedtke were authors
4	NASA Stennis Space Center did the evaluation. Ground truth DEM supplied by USDA-ARS for the San Pedro Watershed Sub-Basin 11
5	Norvelles' report was very detailed and in it he noted a number of anomalies in certain situations (ravines, slopes, trees)

4. Conclusions:

The results of the studies are mutually consistent and provide a performance envelope over a range of conditions. The RMSE results represent the absolute accuracies to be expected in the absence of ground control. If ground control is available, the standard deviation (σ) represents the ultimate limiting accuracy. At lower acquisition altitudes, the accuracies are superior to those at higher altitudes as expected.

These results demonstate that for acquisition at the lower flight altitudes (near 20,000 ft AGL), and <u>with</u> adequate ground control available, absolute accuracies better than 1 meter (RMSE) are obtainable for bald-earth DEMs with 5 meter sample spacing in moderate terrain. At higher altitudes (40,000' AGL) and <u>without</u> ground control, absolute accuracy levels better than 2.5 meters (RMSE) have been demonstrated in moderate terrain.