# Representation of rivers and Lakes within the forthcoming ACE2 Global Digital Elevation Model

**Richard G. Smith[(1)], Philippa A.M. Berry[(1)], Jerome Benveniste[(2)]**
[(1)]*E.A.P.R.S Laboratory, De Montfort University, The Gateway, Leicester, LE1 9BH, UK*
*Email: rgs@dmu.ac.uk, pamb@dmu.ac.uk*
[(2)] *European Space Agency, Earth Observation Applications Department,ESRIN, Via Galileo Galilei, I-00044, Frascati (RM), Italy*
*Email: Jerome.Benveniste@esa.int*

## ABSTRACT

The Shuttle Radar Topographic Mission has generated a unique and valuable dataset, allowing the generation of a Digital Elevation Model of unprecedented accuracy and spatial resolution over much of the earth's land surface. However, one of the limitations of the SRTM dataset is the representation of River and Lake heights. Whilst several researchers have attempted to derive lake and river height information from the SRTM dataset, the accuracy of these heights over large bodies of water is uncertain and over smaller water bodies no values are given. One of the goals of the ACE2 development is to provide the best possible mean height estimates over lakes and to verify or replace the SRTM river values using heights derived from multi-mission Satellite Radar Altimetry. This paper presents case studies outlining the approach taken and shows the results of fusing the Altimeter based data with the SRTM heights for inclusion in ACE2.

## INTRODUCTION

The new ACE2 Global Digital Elevation Model (GDEM) currently under development aims to provide the most accurate GDEM of the Earth. The GDEM will be a fusion of Orthometric height data derived from the Shuttle Radar Topographical Mission (SRTM), the ERS1-Geodetic Mission (ERS1-GM), ERS-2, EnviSat Ku-band and Jason-1. However in this initial study only the SRTM and ERS-1 GM datasets were included.
One of the major elements of this development is obtaining the best possible surface heights for the Earth's inland water bodies. There are a number of important factors requiring examination before the decision of which dataset to use for this particular purpose can be made. Over many rivers there is little choice as the SRTM was unable to return any values and so the dataset contains only voids for those areas. Over the lakes the SRTM solution was obtained by combining the SRTM dataset for the surrounding terrain with other data sources.
This paper presents the findings of the initial intercomparison of the SRTM and ERS1-GM derived heights over Lakes.

## DATASETS

### SRTM

For this study the 'Finished' 3 arc second SRTM dataset derived from C-band radar was used (Hensley et al., 2000; USGS, 2005). The coverage of this dataset is near global extending from 60º North to 56º South. The data were obtained as individual 1º x 1º tiles and combined in order to facilitate direct comparison with the altimeter data on a continental scale. The SRTM documentation states the accuracy of the dataset for absolute heights as ±16m.

### Altimetry

In 1990 the ERS-1 satellite was put into two 168 day interleaved orbits, to provide comprehensive coverage of the world. The data produced along track was at a resolution of 350m (approximately). Whilst the primary focus of most Altimeters is on acquiring Oceanographic heights, an additional mode of the ERS-1 altimeter meant that it was able to retrieve data over varying terrain using a hard wired mask set just inside continental boundaries. In consequence, data are available for much of the Earth's land mass [1,2].
The data from ERS-1 were then processed using an expert system method [3]. Using this method each echo is assigned to one of eleven retracking algorithms chosen according to echo shape. These algorithms identify the point of mean surface reflectivity within the echo and set the time to this point, correcting the range to surface accordingly. These data are then fused with precise orbit data, atmospheric and instrumental corrections to produce ellipsoidal heights. Once combined with a geoid model, in this case EGM96 [4], orthometric heights are returned.

### Lakes Mask

In order to separate the points derived from lakes and those from the surrounding terrain a mask driven approach was utilised. The mask was derived from the Global Land Cover Characterisation database (GLCC) version 2.0 which provides a 30 arc seconds resolution classification of land coverage derived primarily from AVHRR [5]

## METHODOLOGY

### Lakes

In order to make a comparison between the two datasets a test region was designated. This test region was made large enough that a number of different lakes of varying sizes would be present. This paper presents results from this initial investigation and the comparison is to be extended globally.

To ensure that only data over lakes were analysed, a mask of the lakes within each region was created. The ERS1-GM and SRTM data were then checked against this mask and only those values that corresponded to lakes and were present in both datasets were carried through. The lakes were divided up into 3 different categories to see if there was any correlation between the size of the lake and the quality of data being returned. Large lakes were categorised as having an area of greater than 10000 $Km^2$, medium lakes had an area of between 1000 and 10000 $Km^2$ and finally small lakes had an area of less than 1000 $Km^2$.

The region selected was defined to ensure a large enough area so that all of the Rift Valley lakes in Eastern Africa could be incorporated. This area was chosen for a number of reasons; it has a diverse range of both large and small lakes, there are a number of various types of lakes including finger lakes, long thin lakes and more rounded lakes. Coupled to this the terrain surrounding these lakes is incredibly variable which will provide an insight into whether or not the surrounding terrain had any effect on the performance of either instrument.

Data from both missions were returned from 23 lakes within the region. There are 3 "large" lakes in the region, which are commonly referred to as the "African Great Lakes" consisting of Lave Victoria, Lake Tanganyika and Lake Malawi. There are 12 medium sized lakes including Lake Turkana, Lake Albert and Lake Mweru. Finally there were 8 small lakes including Lake Eyasi, Lake Manyaya and Lake Natron.

The statistics for this region are shown in Table 1. It is evident there is actually very strong correlation between the two datasets as 94.7% of points agree to within the stated accuracy of the SRTM dataset. The vast majority of points in the SRTM dataset read as higher than the corresponding ERS1-GM, however only 6% of these points show a difference of greater than 16 metres. These figures show that there is a good general agreement between the two datasets. To further investigate the difference between the two datasets 3D plots depicting both sets of data, were generated for each of the lakes to see if any problems were apparent.

Looking at the overall results from the region (Fig.1) the individual waterbodies are clearly defined; the SRTM is showing rapidly rising terrain around the edges of the lakes, however these values are not supported by the ERS1-GM dataset.

The results over Lake Victoria (1 South 33 East)(Fig.2) are typical of the results seen within the test area. For a large percentage of the lake near perfect agreement between the two datasets is obtained, however the North bank of the lake clearly shows the SRTM values rising rapidly. Since the lake surface could not be exhibiting these rapid height variations the most likely reason is that the SRTM data introducess some horizontal misplacement. The SRTM places the actual start of the lake's surface slightly to the South, so is returning values for the banks rather than the actual water surface.

Using a profile over Lake Victoria (Fig. 3), the scale of the SRTM differences can clearly be seen with some points over 250m from the main lake surface and there is some off ranging visible in the Altimeter data at the edges. The Altimeter has been able to form a much more stable, coherent surface than the SRTM, which has a much larger dispersion of points around the mean lake level. Some pixellation in the SRTM dataset would occur due to the resolution, with values being binned to the nearest meter but even factoring this in the Altimeter results still form a more coherent surface. The variations in the Altimeter lake level heights occur due to the timeframe over which the data were gathered, with the changing level of the lake's surface over the year of measurement introducing this height variation.

Over nearly every lake in this initial sample the Altimeter values' surface is consistently in the lower reaches of the wider SRTM (Fig. 3). The majority of the results though still lie within the SRTM's stated ±16m limits.

Over some lakes such as Lake Turkana (Fig. 4) there is very little difference between the two datasets even at the edges of the lake. Once more the remarkably good cohesion between the SRTM and ERS-1 GM is clearly visible. Two large peaks however occur in the SRTM dataset over the South East of the lake. The farthest east peak is again a case of horizontal misplacement. The second peak (approx 2.7N 33.6 East) however shows a new characteristic, which is a problem within the SRTM dataset. The heights rise in a sharp spike of 400 metres above the level of the lake. While this was at first taken as simple misplacement, the track then returns to a comparable level with the rest of the lake heights in the selected area.

As the data will be present in the new ACE2 GDEM it was necessary to remove the outliers that occurred in both datasets. In order to facilitate this a test lake was chosen with surrounding terrain that was less extreme than some of the cases in the African Rift Valley. Lake Kariba was chosen as the most appropriate lake to fulfil these criteria. The initial unfiltered results (Fig.5) are consistent with the majority of lakes in this test, with some outliers and the Altimeter forming a coherent surface at the lower edge of the SRTM surface. In order to remove the outliers 2 sigma filtering was performed (Fig. 6). With the outliers removed the Altimeter points cover a range of 5m as opposed to the 15m range of the SRTM. The SRTM values are dispersed in such a way that there is little chance of obtaining a clear mean value, the Altimeter data meanwhile are fairly evenly distributed within the 5m range.

**Table 1 Statistics of comparison between SRTM and ERS1-GM points for the Rift Valley Region**

| Type | Number of results | Percentage |
|---|---|---|
| Altimeter higher | 21495 | 22.8% |
| SRTM higher | 72692 | 77.2% |
| Within +/-16m | 89192 | 94.7% |
| Altimeter higher difference >16m | 665 | 3.1% of alt points |
| SRTM higher difference >16m | 4330 | 6.0% of SRTM points |
| Total | 94187 | |

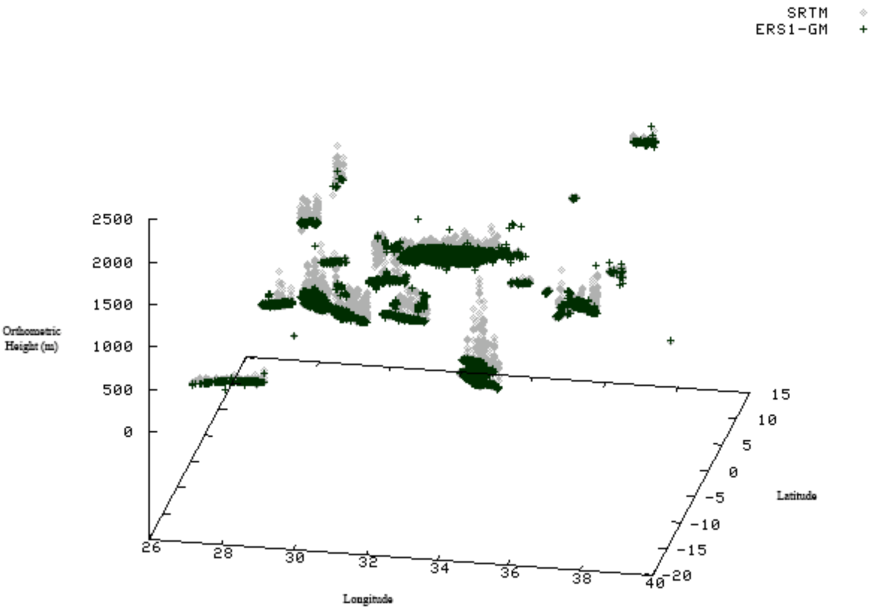


**Figure 1 3-D representation of all SRTM/ERS1-GM heights within the Rift Valley region**

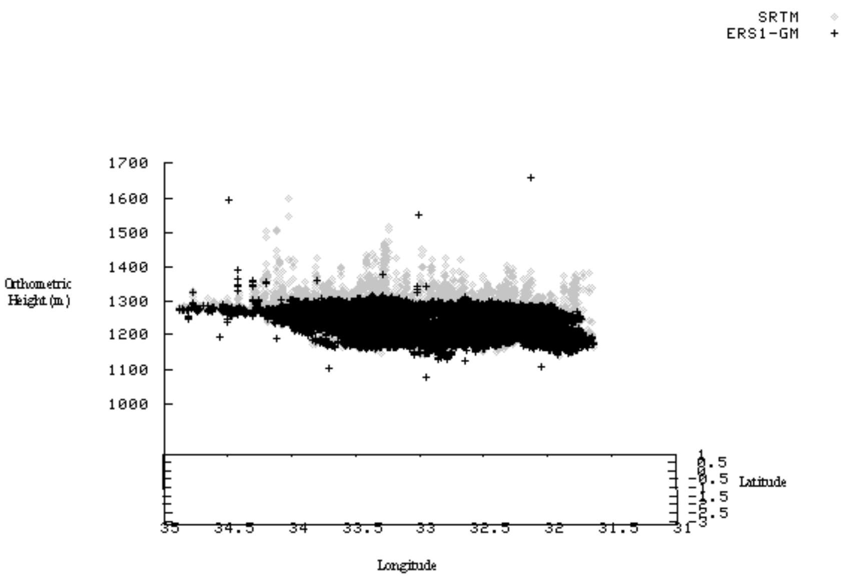


**Figure 2 3-D representation of SRTM/ERS-1GM heights over Lake Victoria**

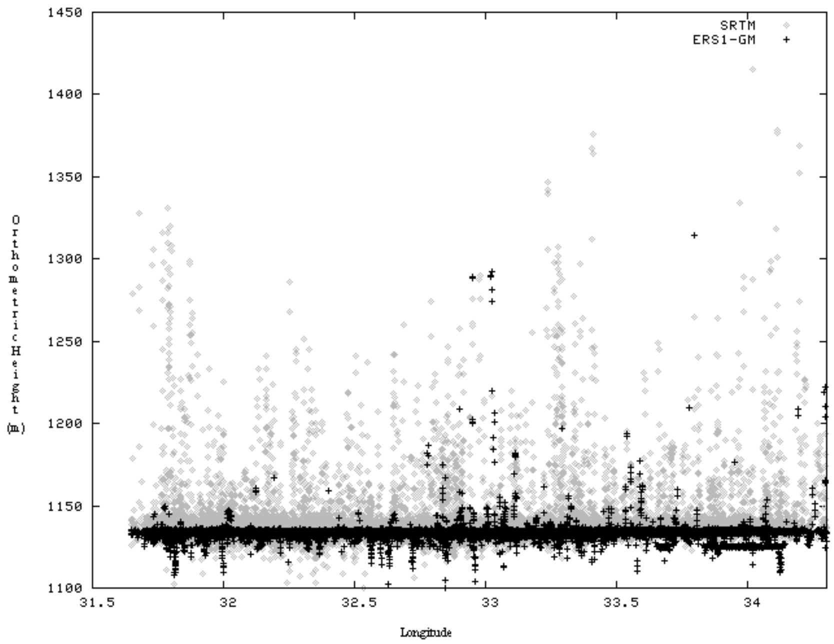


**Figure 3 Profile plot of SRTM/ERS1-GM heights over Lake Victoria**

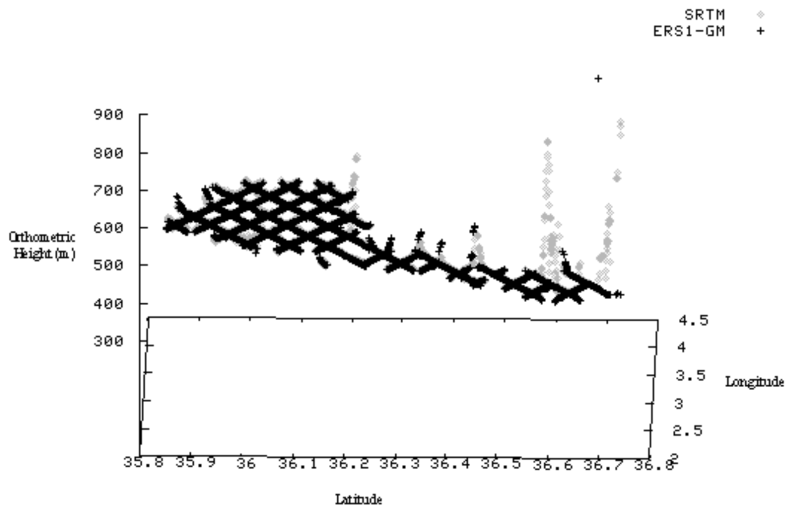


**Figure 4 3-D representation of SRTM/ERS-1GM heights over Lake Turkana**

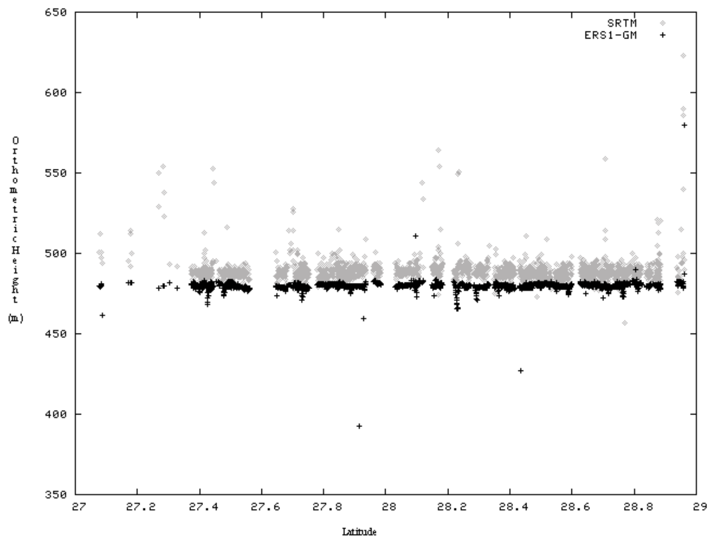


**Figure 5 Unfiltered SRTM/ERS1-GM heights over Lake Kariba**

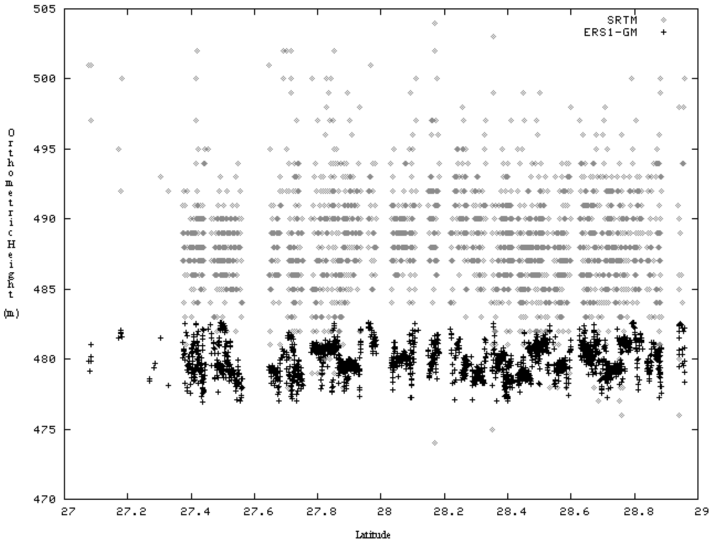


**Figure 6 SRTM/ERS1-GM heights over Lake Kariba after 2Sigma filtering**

**Rivers**

The initial test area for the comparison over rivers is a 15° region with 15° South 75° West as the Southwestern most co-ordinate. Due to the nature of the ERS1-GM track spacing interpolation was required to remove the areas where no direct comparison could take place. The results (Fig. 7) show that on the main rivers there is relatively good agreement, with the difference being sometimes only a couple of meters but rarely above 10 meters. This is because there are numerous tide gauges on the main stream of the river and most of the major tributaries which provided numerous measurements to include within the SRTM derived values.

As the tributaries get smaller the differences start to increase. Those easily visible to the eye have differences in the region of 10 – 16 meters. The smallest tributaries have a difference ranging between 16 – 30+ meters. This is because the SRTM was able to get values from the banks of the tributaries as they were close enough together. However the SRTM cannot significantly penetrate the Rainforest canopy, whereas the Altimeter sees straight through to the ground (Berry et all, 2006).


**Figure 7 Differences between ERS1-GM/SRTM datasets over Amazon River in 15 degree region 15S 75W**

## CONCLUSIONS

This initial study area has revealed a number of interesting characteristics. The SRTM dataset appears to show some horizontal misplacement of the edge of many lakes. Even with this horizontal misplacement 94.7% of all points compared lay within the stated ±16m accuracy of absolute heights within the SRTM dataset. Consistently the Altimeter heights are slightly lower than the SRTM mean height values, this is consistent with other Altimeter results over land. The Altimeter data have a closer vertical distribution than the SRTM data, both before and after applying a 2sigma filter, therefore the Altimeter data are more suitable for the formation of a mean surface of a lake. Altimeter derived values will therefore be used within ACE2 for the mean heights of the African lakes considered in this work.

This comparison is to be extended globally and a number of other test regions will be selected for closer investigation. The areas will be chosen to obtain as wide a variety of terrain and waterbody types as possible. These areas will therefore incorporate mountainous terrain and flat terrain, deep lakes and shallow lakes to investigate whether the results obtained from this comparison apply globally or if they are affected by the surrounding terrain.

## REFERENCES

[1] Berry, P.A.M. ,Global Digital Elevation Models - Fact or Fiction? *Astronomy and Geophysics* Vol. 40, Issue 3, 3.10-3.13, 1999.

[2] Hilton, R.D., Featherstone, W.E., Berry, P.A.M., Johnson, C.P.D. and Kirby, J.F.,Comparison of Digital Elevation Model (DEMs) over Australia and External Validation using ERS Satellite Altimetry, *Australian Journal of Earth Science*, 50, 157-168, 1999.

[3] Berry, P.A.M., Jasper, A., Bracke, H., Retracking ERS-1 altimeter waveforms over land for topographic height determination: an expert system approach. *ESA Pub.* SP-414, Vol. 1, 403-408, 1997.

[4] Lemoine, F.G., Smith, D.E., Kunz, L., Smith, R., Pavlis, E.C., Pavlis, N.K., Klosko, S.M., Chinn, D.S., Torrence, M.H., Williamson, R.G., Cox, C.M., Rachlin, K.E., Wang, Y.M., Kenyon, S.C., Salmon, R., Trimmer,R., Rapp , R.H. & Nerem , R.S., The development of the NASA GSFC and NIMA Joint Geopotential Model , in International Association of Geodest Symposia, Gravity, Geoid and Marine Geodesy , vol 117, ed J. Segawa et al., pp 461 – 469, Springer Verlag, Berlin, 1997.

[5] Loveland T.R., Reed B.C., Brown J.F., Ohlen D.O., Zhu, J., Yang L., and Merchant J.W., Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data: *International Journal of Remote Sensing* , v. 21, no. 6/7, p. 1,303-1,330, 2000