Comparison and validation of the recent freely-available ASTER-GDEM ver1, SRTM ver4.1 and GEODATA DEM-9S ver3 digital elevation models over Australia

Running title: Comparison and validation of recent DEMs over Australia

C. HIRT*, M.S. FILMER AND W.E. FEATHERSTONE

Western Australian Centre for Geodesy & The Institute for Geoscience Research,
Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia,
** Tel: (08) 9266 2218; Fax: (08) 9266 2703; Email: C.Hirt@curtin.edu.au*

This study investigates the quality (in terms of elevation accuracy and systematic errors) of three recent publicly available elevation model data sets over Australia: the 9 arc second national GEODATA DEM-9S ver3 from Geoscience Australia and the Australian National University (ANU), the 3 arc second SRTM ver4.1 from CGIAR-CSI, and the 1 arc second ASTER-GDEM ver1 from NASA/METI. The main features of these data sets are reported from a geodetic point of view. Comparison at about 1 billion locations identifies artefacts (e.g., residual cloud patterns and stripe effects) in ASTER. For DEM-9S, the comparisons against the space-collected SRTM and ASTER models demonstrate that signal omission (due to the ~270 m spacing) may cause errors of the order of 100-200 m in some rugged areas of Australia. Based on a set of geodetic ground control points (GCPs) over Western Australia, the vertical accuracy of DEM-9S is ~9 m, SRTM ~6 m and ASTER ~15 m. However, these values vary as a function of the terrain type and shape. Thus, CGIAR-CSI SRTM ver4.1 may represent a viable alternative to DEM-9S for some applications. While ASTER GDEM has an unprecedented horizontal resolution of ~30m, systematic errors present in this research-grade version of the ASTER GDEM ver1 will impede its immediate use for some applications.

KEY WORDS: Heights, DSMs, DEMs, Australia

INTRODUCTION

Digital elevation models (DEM) provide basic information on heights of the Earth's surface and features upon it. The specific terms digital terrain model (DTM) and digital surface model

(DSM) are often used to specify the surface objects described by an elevation model (e.g., Wood 2008). A DTM usually refers the physical surface of the Earth, i.e., it gives elevations of the bare ground (terrain). On the other hand, a DSM describes the upper surface of the landscape. It includes the heights of vegetation, buildings and other surface features, and only gives elevations of the terrain in areas where there is little or no ground cover.

DEMs have become an important data source for a range of applications in Earth and environmental sciences. Examples of applications for elevation data are numerous, such as gravity field modelling, hydrological studies, topographic cartography, orthorectification of aerial imagery, flood simulation and many more. Generally, DEM data sets can be obtained from a range of techniques, such as ground survey (e.g., Kahmen & Faig 1988), airborne photogrammetric imagery (e.g., ASPRS 1996), airborne laser scanning (LIDAR) (e.g., Lohr 1998), radar altimetry (e.g., Hilton et al. 2003) and interferometric synthetic aperture radar (InSAR) (e.g., Hanssen 2001). Quite often, DEMs are constructed from data sourced from several of these methods and are thus of variable quality (e.g., Hilton et al. 2003).

In the past decade, significant advances in global elevation modelling have been made with the release of the space-borne SRTM (Shuttle Radar Topography Mission, cf. Werner 2001, Farr et al. 2007) and ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer; METI/NASA 2009) elevation data sets. The DEM data from these two space missions cover most of the populated regions of the world and are publicly available (at no cost) at spatial resolutions of 3 arc seconds for SRTM (though 1 arc second data are available to the military) and 1 arc second for ASTER.

These new high-resolution data sets considerably improve the knowledge of the Earth's surface in developing regions with poor geospatial infrastructure. However, benefit can also be gained in large countries with low-population regions containing sparse survey infrastructure, such as Australia. SRTM and ASTER thus represent useful supplementary or alternative elevation data sets to the free-of-charge Australian GEODATA DEM-9S elevation model (Hutchinson et al. 2008; www.geoscience.gov.au/gadds) that gives a DEM at a coarser spatial resolution of 9 arc seconds (~270 m in Australia).

Since a number of applications may rely solely on SRTM and/or ASTER DEMs, it is important to assess the quality of these data, i.e., how well does the DEM approximate the shape of the Earth's surface? Quality of elevation data is commonly expressed in terms of vertical accuracy. It can be determined using comparison data that should be based on accurate and independent methods, such as (terrestrial) topographic surveys, airborne laser scanning or photogrammetric techniques, allowing truly external and independent validation. Another issue affecting the quality of space-based DEMs is the presence of systematic error patterns.

For example, this can include artificial structures that are systematically too high or low and therefore not representative of the terrain's surface. Heights of forest regions or buildings, which are often included in space-collected DEM data (i.e., a DSM), represent an error source for applications exclusively interested in elevations of the terrain (i.e., a DTM). Knowledge of these effects is important for several application fields such as hydrology, where the shape and drainage accuracy is of particular importance (Hutchinson and Dowling 1991).

The aim of this paper is to investigate the quality (in terms of elevation accuracy and systematic errors) of the latest releases of SRTM ver4.1 (published in 2009 by CGIAR-CSI, Italy) and ASTER Global Digital Elevation Model (GDEM) ver1 (made available 2009 by NASA, USA and METI, Japan) over Australia in comparison to GEODATA DEM-9S ver3 (published in 2008 by Geoscience Australia and the Australian National University). We begin by describing the main characteristics (e.g., resolution, construction methods, vertical and horizontal datums) of these three data sets. The quality of the models is then assessed in two ways. A comprehensive model-to-model comparison is carried out over Australia, providing insight into random and systematic effects among the elevation data. External validation is carried out based on two sets of geodetic ground control points (GCPs). The present paper represents a follow-up study to Hilton et al. (2003), because we believe that the significant advances – in terms of resolution and coverage – made by SRTM and ASTER justify a new evaluation of elevation data over Australia. Importantly for many users, the three models investigated are publicly available and completely free of charge. We acknowledge that other elevation data sets exist over Australia, such as Global Land One-kilometre Base Elevation (GLOBE) data set (Hastings & Dunbar 1999) or the 30 arc second GTOPO30 data set (US Geological Survey 1997), but they were already found to be deficient in Australia (Hilton et al. 2003).

Importantly, the space-based ASTER and SRTM data sets used here are formally DSMs, i.e. they provide heights of surface features. Opposed to this, the national GEODATA DEM-9S gives the heights of the terrain surface, so is strictly a DTM.

Finally, a number of studies on the quality of SRTM and ASTER elevation data have already been published (e.g. Fujita et al. 2008, Hayakawa et al. 2008, Kervyn et al. 2008, Nikolakopoulos et al. 2006, Jacobsen 2004). However, these studies used preliminary or different releases of SRTM and ASTER, cover regional instead of continental test areas and, importantly, refer exclusively to test areas outside of Australia.

RECENT DEMS OVER AUSTRALIA

The 1" ASTER ver1, the 3" SRTM ver4.1 and the national 9" GEODATA DEM-9S ver3, all of which completely cover Australia, provide elevation data in regularly spaced grids of geographical coordinates. Generally, they contain physically meaningful height data on the Earth's topographic form. To a rough approximation, the model heights refer to local mean sea level (cf. Featherstone & Kuhn 2006, Torge 2001). The individual surfaces used as vertical references for ASTER, SRTM and GEODATA will be explained later.

A weak inter-dependency exists between SRTM ver4.1 and DEM-9S ver3, in that 'holes' (i.e., no-data areas, mainly in mountainous regions) in SRTM have been filled with auxiliary data supplied by Geoscience Australia (cf. CGIAR-CSI 2009). Apart from this, they provide elevations independent of each other. Table 1 gives the model resolutions, basic storage requirements, and lists the URLs of the data distributors. A first impression of the spatial information delivered by the three models is given by Figure 1, showing Uluru (Ayers Rock), Northern Territory. Due to their higher spatial resolution, SRTM and, particularly, ASTER provide considerably more information on topographic details than DEM-9S.

Table 1 URLs of the data distributors, spatial resolution and storage requirements (model size). The metric resolution (e.g., 270 m for GEODATA DEM-9S) is valid in North-South direction and varies in East-West direction as a function of latitude. The storage requirements are rough estimates based on 2 byte storage per elevation include only the land areas of Australia.

Elevation Model	Resolution	Storage requirements	URL
GEODATA DEM-9S	9" (270 m)	0.2 GB	http://www.geoscience.gov.au/
CGIAR-CSI SRTM ver4.1	3" (90 m)	2 GB	http://srtm.csi.cgiar.org/
ASTER GDEM ver1	1" (30 m)	18 GB	http://www.gdem.aster.ersdac.or.jp/ https://wist.echo.nasa.gov/api/

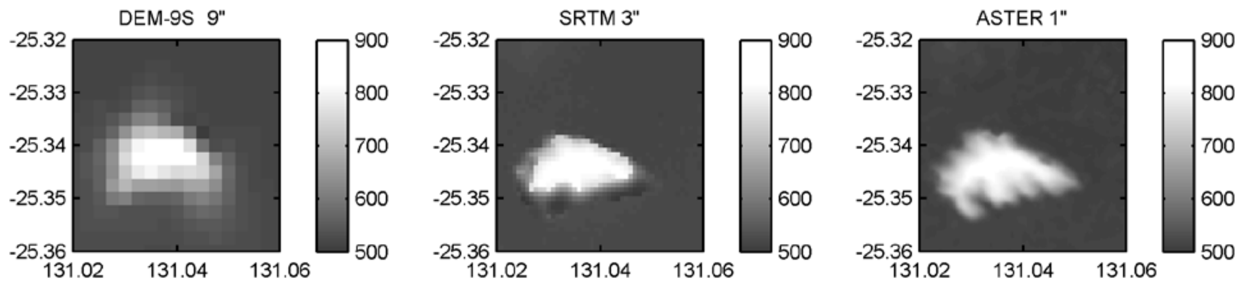

Figure 1 Uluru (Ayers Rock) as represented by 9" GEODATA DEM-9S, 3" SRTM and 1" ASTER. Units in metres.

GEODATA DEM-9S ver3

The GEODATA 9" Digital Elevation Model (DEM-9S) version 3 model (Hutchinson et al. 2008) represents the current national elevation data set of Australia and is publicly available via www.geoscience.gov.au/gadds. This model resulted from a joint effort between the Fenner School of Environment and Society, Australian National University (ANU) and Geoscience Australia (GA). The grid of elevations is based on a variety of input data sets, most of which originate from terrestrial surveying and photogrammetry. This comprises ~5.2 million spot heights, ~2 million water course lines and cliff lines, water bodies and, additionally, altimetry-derived elevations (Geoscience Australia 2008). The approach used to construct DEM-9S is geomorphology-based because of the explicit consideration of Australian drainage patterns (Hutchinson 2007; Hutchinson et al. 2008). Most of the existing terrain structures with scales of 9" and larger are represented.

According to Hutchinson et al. (2008, p.16), DEM-9S provides approximate elevations at the centre of each 9" by 9" cell. Another description of the elevation type is found in Hutchinson et al. (2008, p.17), suggesting that DEM-9S provides average (mean) elevations for a 9" by 9" cell. As such, the definition of elevations provided by the DEM-9S model is ambiguous, although the differences between both definitions may only be significant in complex terrain. Hutchinson (2009 pers. comm.) clarified this by saying "…Formally the DEM values are estimates of the average height across the cell, but mostly there was no more than one source elevation data point per grid cell. So in grid cells with a data point, it tends to be close to the data value in the cell, wherever it was located. In grid cells without data points (the majority), the continuous surface represented by the grid is fairly smooth, so that as far the model is concerned there is little distinction between centre and average, and in reality it's probably somewhere in between".

The vertical accuracy of DEM-9S (standard deviation, 1 sigma) is specified to be 10 m and better in low-elevation terrain, which holds for about 50% of Australia. In rugged or complex terrain, however, the accuracy may deteriorate to about 60 m, which holds for approximately 1% of the data. (Hutchinson et al. 2008). This is due to the rapid variation of elevation across a 9" cell in complex terrain. In other words, the fine structure of the topography is not sufficiently sampled by a 9" grid, which is termed omission error.

DEM-9S is horizontally georeferenced to the Geocentric Datum of Australia (GDA94), but the methods used to realise this and hence the horizontal accuracy are unknown. While GDA94 is claimed to be compatible with WGS84, the latest realisation of WGS84-G873 (NIMA 2004) will differ by about a metre due to the northeast-ward tectonic drift of the Australian continent. Given the uncertainty of the horizontal georeferencing and the grid resolution of 9", this effect is negligible. A sea mask has been applied to DEM-9S, which distinguishes between land and sea points since some heights on the Australian Height Datum (AHD; Roelse et al. 1971) can be below mean sea level (e.g., Lake Eyre). DEM-9S is technically a DTM. For the precise interpolation of DEM-9S, particularly in complex terrain, it is recommended to use higher order methods such as bicubic or biquadratic interpolation (Hutchinson et al. 2008; Hutchinson 2009, pers. comm.).

CGIAR-CSI SRTM ver4.1

The SRTM elevation data cover most land regions between 60 degrees North and 56 degrees South in February 2000 (Werner 2001). It was the first space-borne mapping mission to produce a consistent near-global high-resolution elevation data set. The sensor used for the acquisition was a C-band InSAR, which gives heights of the surface including topographic objects (cf. Farr et al. 2008), i.e., a DSM.

Following the first release of a research-grade SRTM data set in 2004, a finished-grade release became available in 2006. Several research groups subsequently worked on improving the original releases (see the review by Gamache 2004). The improvements concern both the introduction of precise coastline and water-body information, as well as the filling of no-data areas (also called data voids or ‘holes’) in the official releases (e.g., Reuter et al. 2007), an issue that previously impeded the straight-forward use of SRTM elevation grids in certain applications such as gravity field modelling (e.g., Denker 2004).

From a variety of post-processed releases, the freely available CGIAR-CSI SRTM ver4.1 elevation data base (Jarvis et al. 2008) was selected for this study, purely because of its currency. This is the latest post-processed SRTM release by the Consortium for Spatial Information (CSI) of the Consultative Group of International Agricultural Research (CGIAR), Italy. The CGIAR-CSI SRTM ver4.1 data set is based on the official 2006 finished-grade release of SRTM from NASA. An important feature of CGIAR-CSI SRTM ver4.1 is the availability of high-resolution information on shorelines, thus allowing the user to distinguish between land and ocean areas. The shoreline information used is from the SRTM Water Body Dataset, produced by the US Geological Survey (2003).

Importantly, CGIAR SRTM ver4.1 represents a significant improvement over previous releases because ‘holes’ are filled using sophisticated interpolation and patching methods. Depending on the type of terrain, a range of hole-filling interpolation algorithms were applied, such as Kriging, inverse distance weighting and spline interpolation (Reuter et al. 2007). Larger holes (e.g., occurring in steep terrain due to limitations in the SRTM observation principle, see Gamache 2004) were patched by means of auxiliary data sets.

Over Australia, CGIAR-CSI used the GEODATA TOPO 100k contour data from GA (CGIAR-CSI 2009) to fill a total of 255,471 no-data pixels in the SRTM data (Reuter 2009, pers. comm.). This corresponds to less than 0.03% of the SRTM elevations over Australia and causes an, albeit weak, correlation between SRTM and DEM-9S.

The quality of SRTM elevations has been analysed by Rodriguez et al. (2005) in terms of 90% linear and absolute and relative errors. More common accuracy estimates are root mean square errors (RMSEs), which correspond to 1 sigma (68.3% confidence) when sufficiently precise ground truth data is available. These measures have been used by several other authors (e.g., Denker 2004, Marti 2004, Jacobsen 2005, Bildirici et al. 2008). The vertical accuracy estimates (1 sigma or 68.3 % of the elevations) – obtained from comparisons with national ground truth data – vary between 4-6 m in low-elevation terrain and deteriorates to 11-14 m in rugged terrain. It is acknowledged that these figures refer to earlier SRTM releases, but with the improvements by CGIAR-CSI, no deterioration in accuracy is expected for SRTM ver4.1.

SRTM 3D positions are referred to the WGS84 ellipsoid with the heights transformed to a gravity-related physical height using the EGM96 geoid model (Lemoine et al. 1998). The 3" CGIAR-CSI SRTM ver4.1 release is distributed in 5 degree x 5 degree tiles containing 6001 x 6001 (mean) elevations. According to the SRTM observation principle (Farr et al. 2008), the SRTM gives average values for each 3"x3" cell rather than point values and is technically a DSM.

ASTER GDEM ver1

ASTER GDEM ver1 is a new global 1" elevation data set that was released in June 2009 by METI (Ministry of Economy, Trade and Industry), Japan and NASA. The ASTER GDEM is based on optical imagery collected in space with the METI ASTER imaging device that was operated on NASA's Terra satellite. The approach used for constructing the GDEM is correlation of stereoscopic image pairs (e.g., Shapiro and Stockman 2001).

The complete ASTER GDEM covers land surfaces between 83 degrees South and 83 degrees North, which is an improvement over the SRTM coverage. During an observation period of more than 7 years (2000-2007), a total of about 1,260,000 scenes of stereoscopic DEM data of 60 km x 60 km ground areas were collected, so the topography of most regions has been sampled several times. For the 2009 public release, all sets of individual scene-based DEM data were merged and portioned to 1 degree x 1 degree tiles (3601 x 3601 mean elevations).

The overall vertical accuracy of ASTER elevations is specified to vary between 10 m and 25 m (ASTER Validation Team 2009). Like SRTM, ASTER refers to WGS84, with the heights transformed via EGM96 to a physical height. Importantly, no accurate information on land or marine areas is contained in ASTER, nor was an inland water mask applied. This may pose problems (e.g., for hydrological applications) unless external information on water bodies is used as a supplement.

ASTER has the highest formal spatial resolution (1" or ~30 m) and best available coverage to date. Some characteristics of this data set over Australia can be seen in Figure 2. In Figure 2A, series of sand ridges (Great Sandy Desert, Western Australia) can be seen, demonstrating the detail captured by ASTER. It is important to note that ASTER GDEM ver1 is considered to be research-grade (ASTER Validation Team 2009) because a number of artefacts (systematic errors) remain in the elevation data.

Probably the most disturbing effect over Australia is unremoved cloud patterns (Figure 2B), which falsify the elevation model by several kilometres. Fortunately, these artefacts are only over small areas (in particular over Tasmania) and may be easily removed with statistical outlier detection algorithms. Another frequently occurring systematic error is the stripe effect (Figure 2C, see van Ede (2004) for details). Such structures with steps of 10-20 m are generally present over the whole of the Australian continent. For further, but probably less significant, systematic effects detected in the ASTER GDEM, we refer to the report by the ASTER Validation Team (2009).

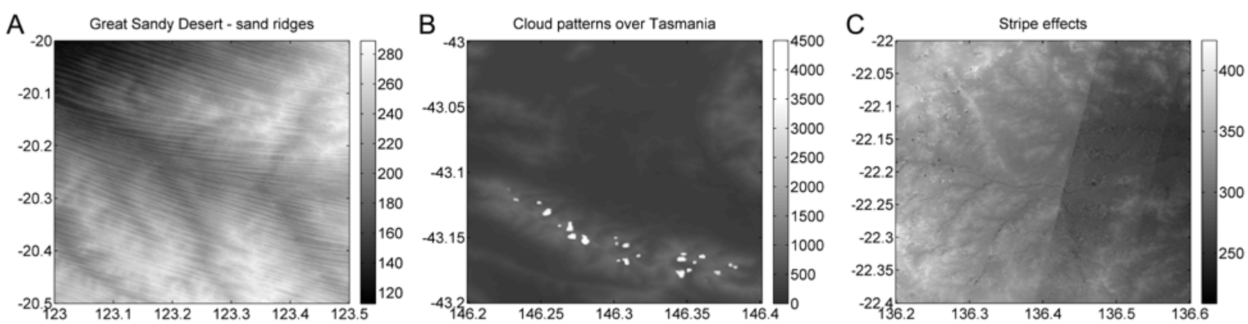

Figure 2 Selected examples of the 1" ASTER elevation data. A: Sand-dune ridges in the Great Sandy Desert. B: Cloud patterns over Tasmania contained in the ASTER data set. C: Stripe-effects contained in the ASTER data set. Units in metres.

DEM EVALUATION

Data preparation and georeferencing

The three DEMs were converted into square tiles of identical binary data format but different spatial coverage (ASTER: 1 degree x 1 degree, DEM-9S and SRTM: 5 degrees x 5 degrees) and stored in a 16 bit integer format, which is a sufficiently precise digital representation of the elevations. For the comparisons among the elevation models, a set of Matlab functions was used that allow for seamless data extraction of arbitrary areas.

When working with elevation data sets, correct georeferencing is an important issue. Previous investigations showed that systematic horizontal shifts can exist among DEMs (e.g., Denker 2004, ASTER Validation Team 2009). Such a shift, sometimes referred to as 'geolocation' errors (Rodriguez et al. 2005), might originate from erroneous georeferencing inherent in the DEM observations. Also, horizontal shifts of 0.5 or 1 cells can be encountered in practice by ambiguous or changing definitions of the position to which elevation refers to (cell corner or centre), as is documented in CGIAR-CSI (2009).

Since 'geolocation' errors deteriorate the vertical accuracy of the elevation data, the three models were initially trialled for correct georeferencing using a simple but effective approach. For selected, sufficiently rugged test areas, such as the Australian Alps or the Stirling Range (Western Australia), 0.25 degree x 0.25 degree DEM grids, were extracted. In order to test relative horizontal offsets among the models, one grid was systematically shifted by small increments of a half cell size (e.g. 1.5 arc seconds with SRTM) in North-South and East-West directions in all combinations and compared against another, unshifted grid. The best fit, i.e. the lowest RMS (root mean square) computed from the differences among the shifted and the unshifted grid indicates the shifts needed for the correct georeferencing among the models. Our testing did not reveal any horizontal offsets with respect to the officially stated location of the grid points (i.e., for DEM-9S, the centres of 9" cells with the edges aligned to whole degrees; for SRTM, the centres of 3" cells with the centres aligned to the whole degrees). The detailed analysis or modelling of regional variations of geolocation errors (cf. ASTER Validation Team 2009, p.9) is beyond the scope of the present study.

Model heights over Australia

Table 2 gives the descriptive statistics of the heights of the Australian continent as implied by the DEMs. In all three cases, the SRTM land mask was applied to extract the land points only, thus making the statistics comparable. The elevation of Australian's highest mountain (Mt. Kosciuszko, 2228 m) is well approximated by DEM-9S and SRTM, while the smallest elevation of DEM-9S represents Australia's lowest region well (Lake Eyre, -16 m, the location of the extreme values were checked).

Furthermore, the mean values of the SRTM, ASTER and GEODATA DEM-9S statistics show - in good agreement - an average height of the Australian continent of about 270-277 m and the RMS values of about 335 m, demonstrating the relative smoothness of most of the Australian topography. Maximum values of about 5 km reveal gross errors from unremoved clouds in the ASTER data set (cf. Figure 2B).

Table 2 Statistics of heights across Australia implied by GEODATA DEM-9S, SRTM and ASTER. The ASTER statistics contain gross errors due to unremoved cloud reflections. Units in metres.

Model	Data points	min	max	mean	RMS
DEM-9S	111,582,167	-16.0	2228.0	272.5	333.9
SRTM	1,001,033,318	-188.0	2220.0	277.5	338.4
ASTER	9,000,069,182	-314.0	5268.0	269.7	331.6

It should be noted that descriptive statistics of the GEODATA (ver1) Australian heights in Hilton et al. (2003) refer to land and ocean points and not to the land surfaces only, as stated in that publication. As such, their mean value is an underestimate.

Comparison among the models

The aim of the comparisons among the three DEMs is to show how they fit to each other, to locate areas of larger discrepancies, and to detect large-scale systematic effects (cf. Hilton et al. 2003). Due to the different spatial resolutions, the comparison requires interpolation. As a compromise, the SRTM resolution of 3" was chosen as resolution for the comparisons. DEM-9S was bicubically interpolated to a denser grid, while the 1" ASTER model was generalised by arithmetically averaging nine adjoining cells. The SRTM land mask was applied consistently to the elevation data of the three models, thus preventing the ocean points from giving unrepresentative statistics.

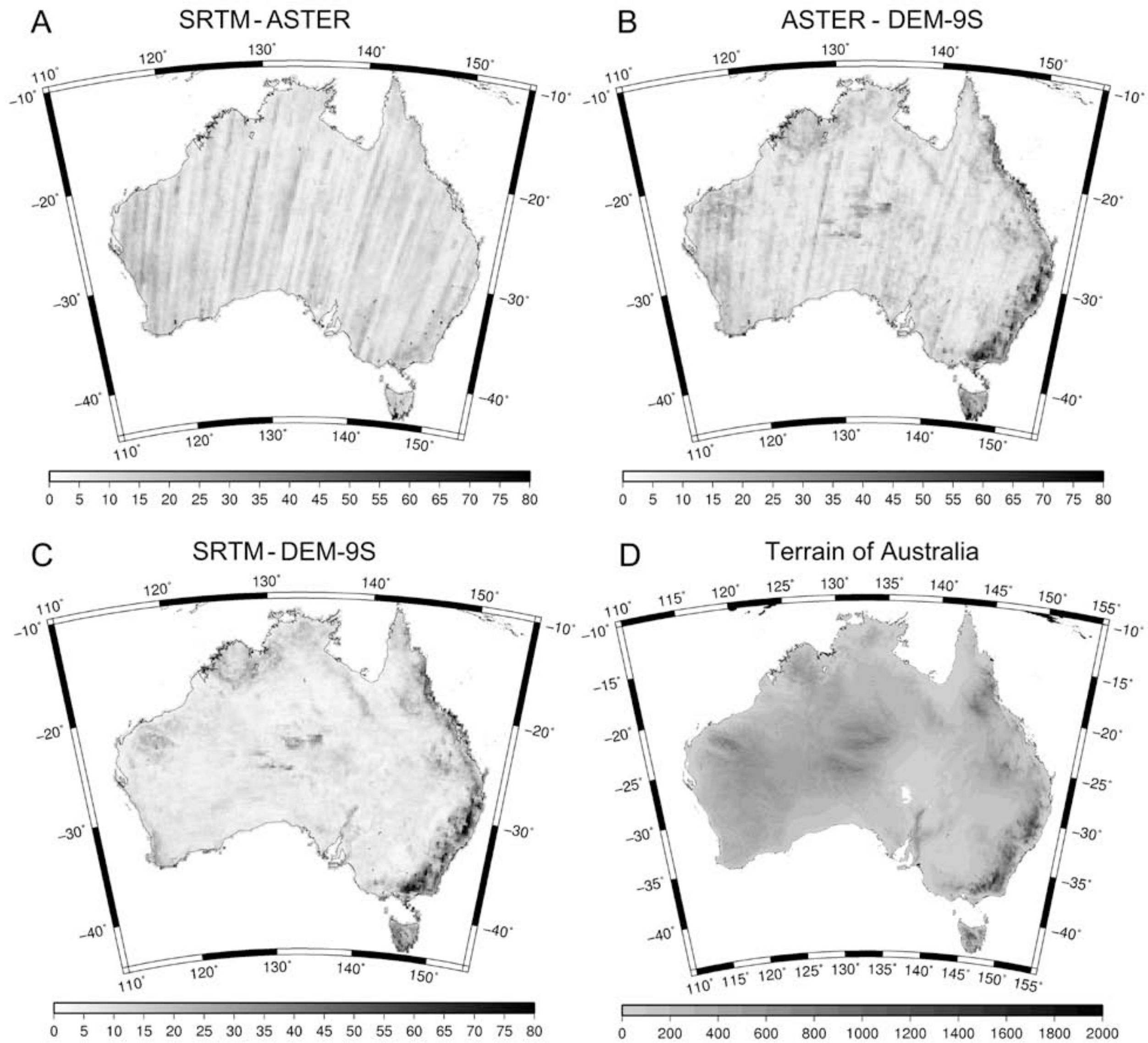

Figure 3 Results of the model-to-model comparisons over Australia. A: RMS differences between SRTM and ASTER, B: RMS differences between ASTER and GEODATA DEM-9S, C: RMS differences between SRTM and GEODATA DEM-9S, D: Terrain of Australia (from SRTM). Units in metres, Lambert projection.

Table 3 Statistics of the model to model comparison at 1,008,271,495 data points (at 3" resolution). Units in metres.

Comparison	Min	Max	Mean	RMS
SRTM – ASTER	-5552.7	437.2	7.7	11.7
ASTER – DEM-9S	-592.8	5675.9	-3.7	15.4
SRTM – DEM-9S	-502.4	553.3	4.0	13.6

The models were compared elevation by elevation: SRTM–ASTER, ASTER–DEM-9S and SRTM–DEM-9S. Accounting for large numbers of elevation points over Australia (about 1 billion at a 3" resolution), the comparisons were performed by means of small tiles of 0.25 degree x 0.25 degree in size, giving 810,000 differences per tile. The RMS (root mean square) of the differences indicating the (dis)agreement among the models is shown for each tile in Figure 3 A-C. The descriptive statistics (of the complete comparison at about 1 billion points) is given in Table 3.

A visual interpretation of Figure 3 shows that the space-based SRTM and ASTER elevation models (Figure 3A) agree well with the RMS values mostly between 5 m and 20 m and an overall RMS of 11.6 m (Table 3). However, large-scale stripe effects are visible all over Australia (Figs. 3A and 3B).

The plot of the RMS differences between ASTER and DEM-9S (Figure 3B) also shows stripe effects, indicating that the source of the stripes is in ASTER. Additionally, significant discrepancies with RMS values as large as 60-80 m are found throughout most of Australia's rugged areas: The Great Dividing Range along the Eastern seaboard (New South Wales and Queensland), the Australian Alps between Victoria and New South Wales (centred at 148W, 37S), the mountains of Tasmania and the MacDonnell Ranges (centred at 132W, 23S), Northern Territory, cf. Figure 3D which illustrates Australia's topography.

Figure 3C shows the RMS differences between SRTM and DEM-9S with similarly large error patterns in all mountainous regions of Australia, but without the stripe artefacts.

Based on the three RMS difference plots, the stripe patterns are unambiguously associated with the ASTER model, and the large discrepancies seen in rugged terrain are attributable to DEM-9S. Interestingly, the ASTER stripe effects are not localised phenomena, but occur on scales of several thousand kilometres.

The cause for the considerable differences in the DEM-9S elevation data present in rugged terrain is signal omission. In these areas, the fine structure of the terrain significantly varies over scales shorter than the model resolution of 9". Errors of the order of 200 m and more may be introduced, which is acknowledged by Hutchinson et al. (2008). The effect of omitted high-frequency terrain signals in DEM-9S also manifests in the larger RMS errors in Table 3.

Table 4 Statistics of large differences (based on analysis of 1,008,271,495 data points at 3" resolution).

Comparison	Number of differences ΔH			Number of differences ΔH		
	100 m < ΔH ≤ 500 m	500 m < ΔH ≤ 1000 m	ΔH >1,000 m	-100 m > ΔH ≥ -500 m	-500 m > ΔH ≥ -1000 m	ΔH < -1,000 m
SRTM – ASTER	68,342	0	0	11,347	321	1,052
ASTER – DEM-9S	1,330,300	314	1037	693,725	2	0
SRTM – DEM-9S	1,729,889	21	0	430,690	1	0

Further interesting insight into the errors of the DEMs is given in Table 4, showing the complete statistics of large discrepancies over Australia, i.e. differences which exceed 100 m, 500 m, 1000 m or fall below -100 m, -500 m and -1000 m, respectively. From Table 4, it can be concluded that about 1400 outliers (discrepancies of 500 m or larger) are contained in the ASTER data set (at a reduced resolution of 3"). Furthermore, it can be seen from SRTM–DEM-9S and ASTER–DEM-9S that the differences of roughly about 1.3-1.7 million points fall into the range 100 m to 500 m, while a smaller number (-0.4 to -0.7 million) range between 100 m and 500 m. This provides some evidence that interpolating DEM-9S elevations in Australia's mountain regions gives differences that are often systematically too small. It should be noted that the results in Table 4 are subject to interpolation (DEM-9S) and generalisation (ASTER).

Model validation with ground truth data

As opposed to comparisons among the DEMs, model validation using ground truth data can deliver reasonable accuracy estimates, provided that the height data are independent and sufficiently precise (say, 1 m or better). Two such data sets, available at the Western Australian Centre for Geodesy, were selected to serve as ground control points (GCPs) because of their higher-order accuracy of the height component, and because of a sufficiently precise horizontal position.

An accurate height is required for comparison to DEMs, but a large uncertainty in the horizontal coordinates will lead to the serious degradation of the height. For example, the horizontal positions of benchmarks on the AHD were originally scaled from 1:250,000 map sheets and recorded to the nearest arc minute of latitude and longitude (Roelse et al. 1971).

Thus, the maximum error in horizontal position could be 30" (~ 900 m in latitude). In hilly or mountainous terrain, the height difference between the benchmark and the topography at the actual position of the benchmark coordinates could be hundreds of metres; in relatively flat country it could still amount to a few metres. Ideally, the horizontal positional uncertainty of the benchmarks used as ground truth should be no more than several metres.

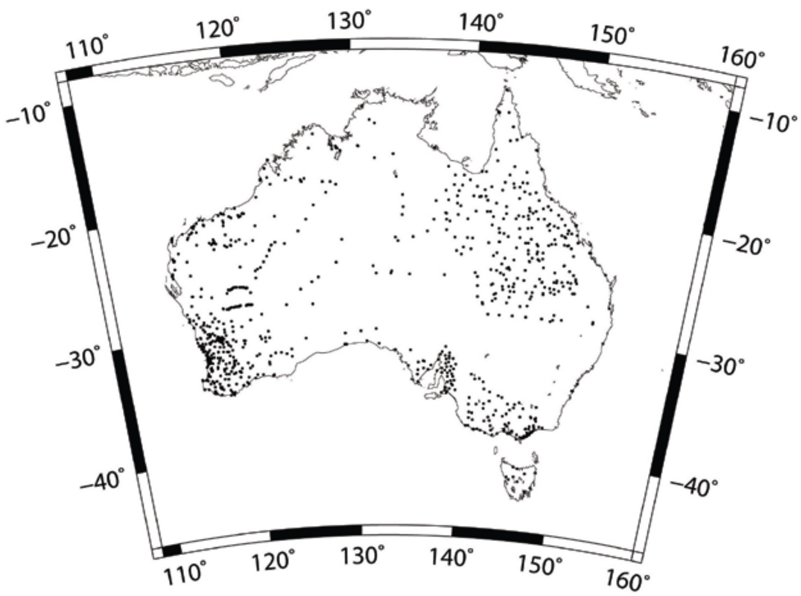

Figure 4 911 GCPs (GPS/levelling; provided by GA) over Australia. Lambert projection.

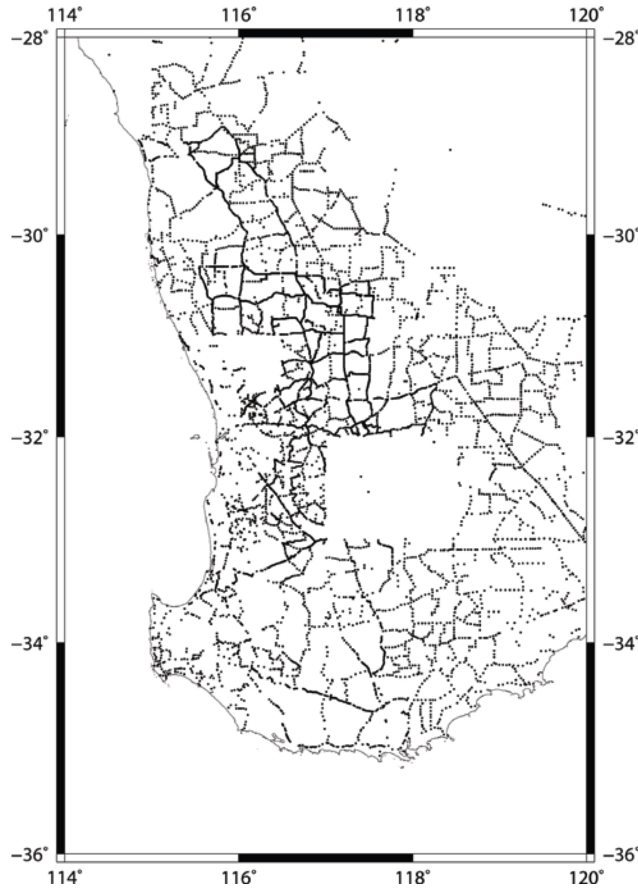

Figure 5 Distribution of 6392 AHD levelling benchmarks (provided by Landgate) with horizontal coordinates accurate to 3 m or less. Mercator projection.

The first dataset comprises 911 GPS/levelling points provided by GA (N. Brown pers. comm. 2009), which has good coverage over Australia (Figure 4). These data have recently been reprocessed in ITRF2005 at epoch 2000.0 and are expected to have horizontal and vertical accuracy of a few centimetres with respect to the reference frame ITRF2005 (Hu 2009). For the comparison with the DEM data, the GPS ellipsoidal heights were transformed to physical heights using EGM96 (Lemoine et al. 1998). This has the advantage of being consistent with the vertical georeferencing of SRTM and ASTER.

The second dataset comprises 6392 AHD levelling benchmarks (Figure 5) provided by the Western Australian Land Information Authority Landgate (G. Holloway pers. comm. 2009) which cover the south-western part of Western Australia. While AHD benchmark coordinates generally have a horizontal accuracy to the nearest arc minute in the Australian Geodetic Datum 1966 (AGD66), Landgate, where possible, have been gradually updating the accuracy of horizontal benchmark coordinates, often with differential GPS to an accuracy of 3 m or less (G. Holloway pers. comm. 2009).

However, the AHD is known to suffer from a north-south slope of ~1 m (e.g., Featherstone 2004) and distortions of up to ~±0.5 m in the levelling network due to gross and systematic levelling errors (e.g., Filmer and Featherstone 2009). We consider a reasonable vertical accuracy estimate of absolute AHD heights to be ~ 1 m, plus an unknown bias with respect to global geoid models such as EGM96. Because of the connection to the AHD, the 6392 GPCs are more consistent with the vertical georeferencing of DEM-9S than with the space-based ASTER and SRTM models.

For this aspect of the DEM evaluation, the model elevations were interpolated bicubically from the surrounding grid points of the original spatial resolution of each model to each GCP. The descriptive statistics of the differences against the 911 GPS GCPs (ellipsoidal heights referred to EGM96) is reported in Table 5, the histograms are found in Figure 6. In open terrain (mostly without forest or buildings), SRTM gives good results with RMS differences as small as 5.0 m. The other models show larger residuals with RMS differences of 10.5 m (DEM-9S) and 13.1 m (ASTER).

Table 5 Statistics of the model comparison with 911 GPS-EGM96 GCPs. Units in metres.

Comparison	**Resolution ["]**	**Min**	**Max**	**Mean**	**RMS**	**Std.dev**
DEM-9S – GPS-EGM96	9	-78.3	35.6	-3.7	10.5	9.8
SRTM – GPS-EGM96	3	-36.6	15.4	1.3	5.0	4.9
ASTER – GPS-EGM96	1	-60.0	75.4	-8.2	13.1	10.2

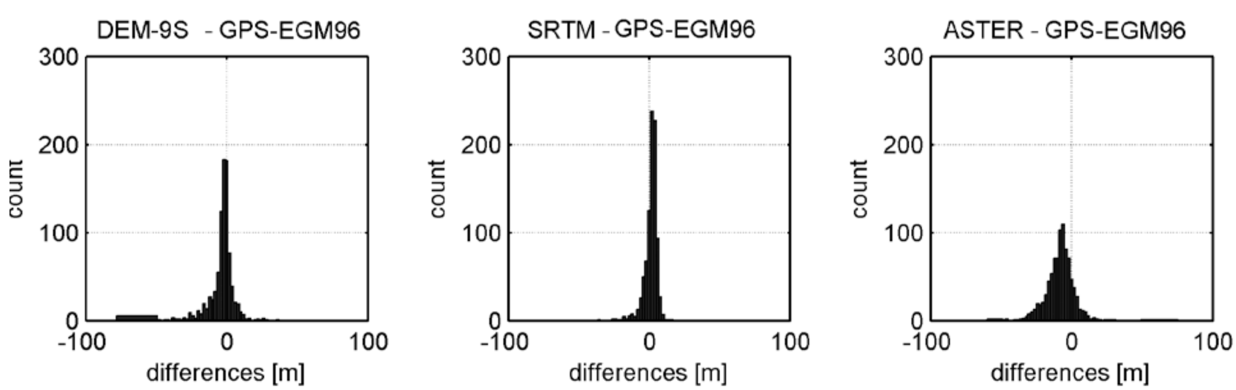

Figure 6 Distribution of the differences among DEM-9S, SRTM and ASTER and 911 Australian GPS-EGM96 GCPs.

The results from the comparisons at the 6392 levelling GCPs are given in Table 6 and Figure 7. Again, SRTM elevations produce the lowest residual errors with an RMS (1 sigma) of 6.1 m and a low standard deviation of 3.2 m. These values provide some evidence of the reasonably good quality of the SRTM elevation data set by CGIAR-CSI over Australia.

The accuracy of DEM-9S, as determined using our benchmarks is about 9 m (RMS and STD) and the ASTER accuracy is lower with about 16 m RMS and 13 m standard deviation. The analysis of mean values of differences shows a very good fit among the GCPs and the DEM-9S elevations. Recalling that the vertical datum of both the 6392 GCPs and the DEM-9S is the AHD, the good agreement is an endorsement of the modelling and interpolation methods used for computing DEM-9S (Hutchinson 1989, 2007).

The mean values of SRTM and ASTER differences reflect a number of effects: (1) the incompatibility of the AHD and WGS84-EGM96 heights, (2) satellite-collected elevation data tend to be too high (DSM vs. DTM), and (3) ASTER elevations are subject to large-scale stripe-like error patterns (shown earlier). At our 6392 GCPs, SRTM elevations are around 5 m too high, while the heights from the ASTER model are about 9 m too low. Further analysis will be required (i.e. larger areas with dense sets of GCPs) in order to corroborate these results.

Table 6 Statistics of the model comparison with 6392 levelled benchmarks (levelling GCPs). Units in metres

Comparison	**Resolution ["]**	**Min**	**Max**	**Mean**	**RMS**	**Std.dev**
DEM-9S – Lev	9	-79.8	63.8	0.5	8.9	8.9
SRTM – Lev	3	-23.0	36.9	5.2	6.1	3.2
ASTER – Lev	1	-167.1	123.4	-9.1	15.7	12.8

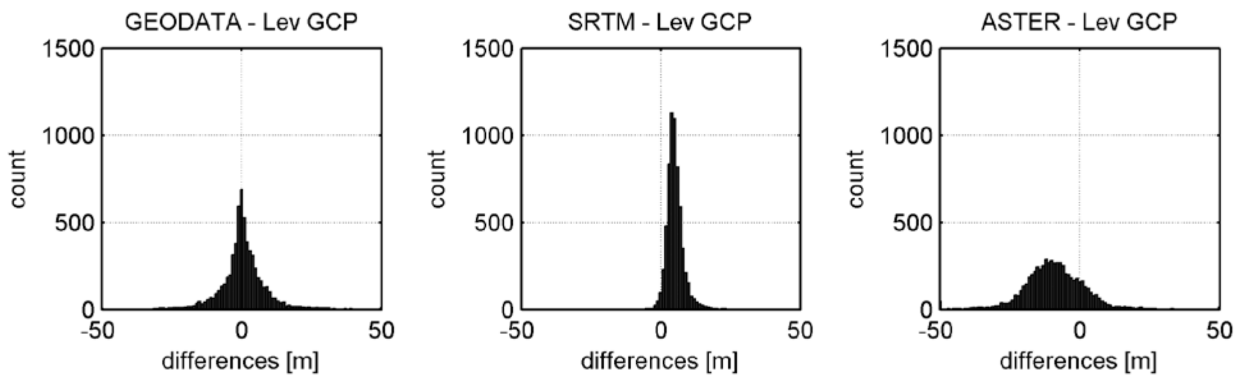

Figure 7 Distribution of the differences among DEM-9S, SRTM and ASTER and 6392 Australian levelling benchmarks over WA.

CONCLUSIONS

This study has investigated the quality of three new digital elevation models GEODATA DEM-9S ver3, CGIAR-CSI SRTM ver4.1 and NASA/METI ASTER GDEM ver1 over Australia, all of which are available free of charge. The basic characteristics of the models were described, comparisons among the three models drawn, and accuracy estimates by means of comparisons against GCPs derived. All models have strengths and weaknesses, which can be summarised as follows.

The national GEODATA DEM-9S ver3 elevation model that mainly relies on terrestrial survey data represents the Australian topography with particular focus on the proper inclusion of drainage patterns. The DEM-9S elevations are provided on the AHD. The vertical accuracy of DEM-9S elevations is found to be around 9 m from the comparison with levelling GCPs, which corroborates the official accuracy estimate by Hutchinson et al. (2008) valid for less-elevated terrain. Because of the relatively coarse resolution of 9" (as compared to the space collected models), DEM-9S shows large errors of up to a few 100 m in rugged terrain. These errors reflect signal omission and may limit its suitability for certain applications.

The CGIAR-CSI SRTM ver4.1 elevation data set from InSAR observations comes at a 3" resolution. It performs best in both the model-to-model comparisons and in the comparisons with GCPs (RMS values of about 6 m). However, this good result is possibly related to the fact that our GCPs are located in rather less-vegetated areas. In areas with dense vegetation, systematically too high SRTM heights are generally to be expected based on experiences in other countries (e.g., Denker 2004, Marti 2004). According to CGIAR-CSI (2009), holes in mountainous areas – the most crucial part in earlier SRTM releases – were filled using auxiliary data from GA. In summary, we consider the SRTM ver4.1 data to be a serious alternative to GEODATA for a range of DEM applications in Australia. For hydrological applications, however, the drainage accuracy remains to be assessed.

The ASTER GDEM ver1 elevation data set constructed from optical stereo imagery is provided at a very high grid resolution of 1". The model contains artificial error patterns (stripes and cloud anomalies), which is why METI/NASA consider it to be research-grade only. Moreover, the ASTER elevations showed the lowest accuracy in the GCP comparison with RMS values of about 15 m. However, this agrees with the formally stated accuracy range of ASTER elevations (10-25 m, cf. ASTER Validation Team 2009).

The currently available DEM-9S or SRTM releases are preferred over ASTER for most applications, unless the ASTER model can be improved (e.g. outliers and stripes removed) by the user. It is hoped that efforts towards data cleaning (previously seen with the SRTM data) will lead to better, post-processed ASTER versions. In particular, it is the unprecedented detail that will be beneficial for a number of applications.

Acknowledgements

CH and WEF thank the Australian Research Council for financial support through grant DP0663020. MSF receives financial support from an Australian Postgraduate Award (APA), The Institute for Geoscience Research and the Cooperative Research Centre for Spatial Information. Special thanks go to Andreas Lindau for providing useful Matlab scripts and to Mike Hutchinson for emailed clarifications on DEM-9S, his thorough review and helpful comments. We would also like to thank John Dawson for his useful remarks on this manuscript. We would like to thank Geoscience Australia and Landgate (the Western Australian geodetic agency) for supplying GPS and AHD benchmark data respectively, particularly Graeme Holloway (Landgate) for helpful clarification regarding WA benchmark information. Figures 3, 4 and 5 were produced using Generic Mapping Tools (GMT; Wessel and Smith, 2007). This is The Institute for Geoscience Research (TIGeR) publication number XX.

References

ASPRS 1996. Digital Photogrammetry – An Addendum to the Manual of Photogrammetry (ed. C. Greve). Publication of the American Society for Photogrammetry and Remote Sensing, Maryland.

ASTER Validation Team 2009. ASTER global DEM validation summary report. Published by the ASTER GDEM Validation Team: METI, NASA and USGS in cooperation with NGA and other colloborators. June 2009, 28 pages. Available online: https://lpdaac.usgs.gov (accessed on 28.07.2009).

Bildirici O., Ustun A., Ulugtekin N., Zahit Selvi H., Abbak A., Bugdayci I. & Ozgur Dogru A. 2008. Comparison of SRTM and 25K topographic maps in Turkeye. Proceedings Second International Conference on Cartography and GIS, Borovets (Bulgaria), 219-227. Available online: http://www.mmf.selcuk.edu.tr/~aabbak/pubs/8.pdf (accessed on 28.07.2009).

CGIAR-CSI 2009. SRTM Data Processing Methodology. *Web document.* Available at: http://srtm.csi.cgiar.org/SRTMdataProcessingMethodology.asp (accessed on 28.07.2009).

Denker H. 2004. Evaluation of SRTM3 and GTOPO30 Terrain Data in Germany. *GGSM 2004 IAG International Symposium Porto, Portugal* (ed. C. Jekeli et al.), Springer, Heidelberg: 218-223

Farr, T.G., Rosen P.A., Caro E., Crippen R., Duren R., Hensley S., Kobrick M., Paller M., Rodriguez E., Roth L., Seal D., Shaffer S., Shimada J., Umland J., Werner M., Oskin M., Burbank D. & Alsdorf D. 2007. The Shuttle Radar Topography Mission. *Rev. Geophys.* **45**, RG2004, doi:10.1029/2005RG000183.

Featherstone, W.E. 2004. Evidence of a north-south trend between AUSGeoid98 and the AHD in southwest Australia, *Survey Review* **37**(291): 334-343.

Featherstone W.E. & Kuhn M. 2006. Height systems and vertical datums: a review in the Australian context. *Journal of Spatial Science* **51**(1), 21-42

Filmer, M.S. & Featherstone W.E. 2009. Detecting spirit-levelling errors in the AHD: recent findings and some issues for any new Australian height datum, *Australian Journal of Earth Sciences* **56**(4), 559-569

Fujita K., Suzuki R., Nuimura T., Sakai A. 2008. Performance of ASTER and SRTM DEMs and their potential for assessing glacial lakes in the Lunana region, Bhutan Himalaya. *Journal of Glaciology* **54**(185), 220-228.

Gamache M. 2004. Free and low cost data sets for international mountain cartography. Paper presented at the Workshop of the Commission on Mountain Cartography if the International Cartographic Association, Vall de Nuria, Spain 2004, 39 pages. Online available at: http://www.mountaincartography.org/publications/papers/papers_nuria_04/gamache.pdf (accessed 28.July 2009).

Geoscience Australia 2008. GEODATA 9 Second Digital Elevation Model (DEM-9S) Version 3. Web document. http://www.ga.gov.au/bin/htsqr?file=/oracle/geomet/geomet2.htsqr&datasetno=11541 (accessed 28.July 2009).

Hayakawa Y.S., Oguchi T., Lin Z. 2008. Comparison of new and existing global digital elevation models: ASTER G-DEM and SRTM-3. *Geophysical Research Letters* **35**(17), L17404.

Hanssen R.F. 2001. Radar interferometry: Data Interpretation and error analysis. Springer, New York.

Hastings D.A. & Dunbar, P.K. 1999. Global Land One-kilometre Base Elevation (GLOBE) Digital Elevation Model, Documentation, 1.0 Key to Geophysical Records Documentation (KGRD). 34. National Oceanic and Atmospheric Administration, National Physical Data Center, Boulder.

Hilton R.D., Featherstone W.E., Berry P.A.M., Johnston C.P.D & Kirby J.F. 2003. Comparison of digital elevation models over Australia and external validation using ERS-1 satellite radar altimetry, *Australian Journal of Earth Sciences* **50**(2): 157-168. doi: 10.1046/j.1440-0952.2003.00982.x

Hu, G. 2009. Analysis of Regional GPS Campaigns and their Alignment to the International Terrestrial Reference Frame (ITRF) , *Journal of Spatial Sciences*. **54**: 15-22

Hutchinson, M.F. 1989. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. Journal of Hydrology 106: 211-232

Hutchinson M.F. & Dowling T.I. 1991. A continental hydrological assessment of a new grid-based digital elevation model of Australia, *Hydrological Processes* **5**: 45-58

Hutchinson M.F. 2007. ANUDEM Version 5.2.2 Fenner School of Environment and Society, Australian National University, Canberra. http://fennerschool.anu.edu.au/publications/software/anudem.php (accessed 29.October 2009)

Hutchinson M.F., Stein J.A., Stein J.L., Anderson, H. & Tickle, P. 2008. Geodata 9 Second DEM and D8 – Digital Elevation Model Version 3 and Flow Direction Grid. User Guide. Fenner School of environment and society, ANU and Geoscience Australia, 43 pages.

Jacobsen K. 2004. Analysis of digital elevation models based on space information. *EARSeL conference proceedings*, Dubrovnik. 8 pages. Available at: http://www.ipi.uni-hannover.de/uploads/tx_tkpublikationen/JAC_dubrov04.pdf (accessed 28.July 2009)

Jacobsen K. 2005. Analysis of SRTM Elevation Models. *EARSeL conference proceedings*, Porto, 7 pages. Available at: http://www.ipi.uni-hannover.de/uploads/tx_tkpublikationen/ASEjac.pdf (accessed 28.July 2009).

Jarvis A., Reuter H.I., Neson, A. & Guevara, E. 2008. Hole-filled SRTM for the globe Version 4. Available from the CGIAR-SXI SRTM 90m database: http://srtm.csi.cgiar.org

Kahmen H. & Faig W. 1988. Surveying. W. de Gruyter, Berlin, New York.

Kervyn M., Ernst G.G.J., Goosens R., Jacobs P. 2008. Mapping volcano topography with remote sensing: ASTER vs. GDEM. *International Journal of Remote Sensing* **29**(22), 6515-6538.

Lemoine F.G., Kenyon S.C., Factor, J.K., Trimmer, R.G., Pavlis, N.K., Chinn D.S., Cox C.M., Klosko S.M., Luthcke S.B., Torrence M.H., Wang Y.M., Williamson, R.G., Pavlis E.C., Rapp R.H. & Olson T.R. 1998. The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) geopotential model EGM96, *NASA/TP-1998–206861*. National Aeronautics and Space Administration, Greenbelt

Lohr U. 1998. Digital Elevation Models by Laser Scanning. *Photogrammetric Record* 16(91), 105-109.

Marti U. 2004. Comparison of SRTM data with the national DTMs of Switzerland. Paper presented at *GGSM 2004 IAG International Symposium Porto, Portugal.* Also published by Swisstopo, Wabern, Switzerland.

METI/NASA 2009. ASTER Global Digital Elevation Model by Ministry of Economy, Trade and Industry of Japan (METI) and the National Aeronautics and Space Administration (NASA) Available at: http://asterweb.jpl.nasa.gov/gdem.asp and http://www.gdem.aster.ersdac.or.jp/

Nikolakopoulos K.G., Kamaratakis E.K., Chrysoulakis N. 2006. SRTM vs ASTER elevation products. Comparison for two regions in Crete, Greece. *International Journal of Remote Sensing* **27**(21), 4819–4838

NIMA 2004. Department of Defense – World Geodetic System 1984. *National Imagery and Mapping Agency Technical Report* 8350.2. Published by the US National Imagery and Mapping Agency NIMA (now National Geospatial Intelligence Agency NGA), 175 pages

Reuter H.I., Nelson A., & Jarvis A. 2007. An evaluation of void filling interpolation methods for SRTM data. *International Journal of Geographical Information Science* **21(9)**: 983-1008

Rodríguez E., Morris C.S., Belz J.E., Chapin E.C., Martin J.M., Daffer W. & S. Hensley 2005. An Assessment of the SRTM Topographic Products, *Technical Report JPL* D-31639, Jet Propulsion Laboratory, Pasadena, California, 143 pp.

Roelse, A., Granger, H.W. and Graham, J.W. 1971. The adjustment of the Australian levelling survey 1970-1971, *Technical Report 12*, Division of National Mapping, Canberra, Australia, 81 pp.

Shapiro L.G. & Stockman G.C. 2000. Computer Vision. Prentice Hall, New Jersey.

Torge W. 2001. Geodesy. Third edition. W. de Gruyter, Berlin, New York.

U.S. Geological Survey 1997. GTOPO30 Digital Elevation Model. Web document. http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html (accessed 28.July 2009).

Us Geological Survey 2003. SRTM Water Body Data Set. Web document. http://edc.usgs.gov/products/elevation/swbd.html (accessed 28.July 2009).

van Ede R. 2004. Destriping and Geometric Correction of an ASTER Level 1A Image. Thesis, Faculty of GeoSciences, Utrecht University, Netherlands. Available at: http://www.scribd.com/doc/3752774/Aster-Correction-Thesis (accessed 18. August 2009).

Werner M. 2001. Shuttle Radar Topography Mission. Mission overview. *Journal of Telecommunication* **55**, 75-79.

Wood J. 2008. Digital Elevation Model (DEM). In: Encyclopedia of Geographic Information Science (ed. K.K. Kemp). SAGE Publications, London.

Wessel, P. & W. H. F. Smith 1998. New, improved version of the Generic Mapping Tools released, *EOS Trans. AGU*, 79, 579.