The development of a hydrologically improved Digital Elevation Model and derived products for South Africa based on the SRTM DEM

Report to the

WATER RESEARCH COMMISSION

by

HL WEEPENER[1], HM VAN DEN BERG[2], M METZ[2] & H HAMANDAWANA[1]

[1] ARC-Institute for Soil, Climate and Water, Pretoria
[2] IRIS International, Potchefstroom

WRC Report No. 1908/1/11

ISBN 978-1-4312-0217-1

JANUARY 2012

DISCLAIMER

This report has been reviewed by the Water Research Commission (WRC) and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the WRC, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

© **WATER RESEARCH COMMISSION**

EXECUTIVE SUMMARY

Digital Elevation Models (DEMs) represent the topography in a landscape and are an important data source used in numerous hydrological and geo-hydrological studies/projects. There are several DEMs available for South Africa, but usually they require significant editing before they can be used for hydrological modelling, therefore this project is trying to provide a readily usable set of data products for hydrological modelling.

Aim and objectives

The aim of the project was to prepare a hydrologically improved Digital Elevation Model for South Africa from the Shuttle Radar Topography Mission (SRTM) DEM.

The SRTM DEM provides a readily accessible resource that can be effectively exploited to guide regional planning especially in data scarce areas where prohibitive costs preclude the use of high resolution equivalents. Although the utility of the SRTM-derived DEM is incontrovertible, effective use of this dataset is constrained firstly by areas of missing data which are mostly associated with water bodies, mountainous areas and desert regions and secondly by occasionally elevated elevation values in flow paths caused by bridges, high trees with large crowns, steep cliffs, etc. obscuring the flow path below. These ‘blockages’ in the data cause wrong flow path extraction and the formation of ‘inland lakes’ in the data. In very flat areas elevated elevation values can cause serious misrepresentation of large meandering rivers.

The objectives of the project were to:

1. Research methodologies for producing hydrologically improved SRTM datasets.
2. Produce a hydrologically corrected, gap and sink free SRTM dataset for South Africa (SA).
3. Re-interpolate the hydrologically corrected SA DEM to 30 m and derive a hill-shade layer from this DEM for improved visual display.
4. Conduct a comparative analysis of a small study area in the United States of America (USA) where both 30 m and 90 m SRTM datasets are available and use this comparison as a basis for determining the value of re-interpolating the SA hydrologically corrected DEM to 30 m and deriving secondary datasets from this 30 m DEM.
5. Correct inaccurate major river flow patterns for flat areas.
6. Prepare derived products from the DEM.

Study area

Drainage basins have been considered to be some of the most important areal units in environmental decision making (EU 2000, as cited by Oksane and Sarjakoski, 2005), because they delineate the land area of runoff-driven sediment transport. Therefore the study area was selected to cover South Africa and the two transboundary drainage basins of the Limpopo and Orange River basins. The extent of the study area is between 19ºS and 35ºS and 12ºE and 36ºE and includes the whole of Swaziland and Lesotho and large parts of Namibia, Botswana, Zimbabwe and Mozambique.

Methodology

There are currently two global high-resolution DEMs available, namely the SRTM and the Advanced Spaceborne Thermal Emission and Reflection Radiometer Digital Elevation Model (Aster GDEM). Although the Aster GDEM has a better resolution, its full potential would probably only be available with the second release. The Japanese Ministry of Economy, Trade and Industry (METI) and the United States National Aeronautics and Space Administration (NASA) have stated that version 1 of the Aster GDEM should be viewed as "experimental" or "research grade" (Aster GDEM Validation Team, 2009). Additionally contour lines with a 20 m interval at the 1:50 000 scale are available from the Chief Directorate of National Geospatial Information for the whole of South Africa. The SRTM data has an advantage over contour datasets in that it is sampled from regularly spaced data points whereas the contour derived dataset is prone to interpolation errors, particularly in flat areas (Department of Water Affairs and Forestry, 2006). These differences affect the accuracy of streams and catchments derived from elevation data. The SRTM dataset is therefore currently the most appropriate dataset to use.

A continuous DEM was derived by filling voids in the SRTM dataset as follows:

- Inside South Africa: through interpolations from 20 m interval contour lines extracted from 1:50 000 topographical maps;
- For the rest of the study area outside South Africa: the Aster GDEM data was used to fill voids.

The chosen method to perform an automated hydrological correction was based on the impact reduction approach of Lindsay and Creed (2005). For each sink the area to be filled from the spill point and the length of the channel to be carved were determined first. If channel length is shorter than a user-defined threshold (here 4 cells), a channel was carved. If the area to be filled was smaller than a user-defined threshold (here also 4 cells), the sink was filled completely. Otherwise, the sink was incrementally filled up from the bottom and on each step the incremental impact of filling this sink a little and carving out for the partially filled sink was determined. When a least impact equilibrium of partial filling and partial carving was reached, the sink was removed by using the current parameters for partial filling and partial carving. This method reduces the number of modifications necessary to produce a hydrologically correct DEM and avoids typical problems of pure sink filling (artificially flooded terrain) and pure carving. Natural depressions in the landscape must be maintained in the DEM. Examples are areas where natural and man-made water bodies and mines occur. These include fluctuating water bodies like seasonal pans. Caution was taken to avoid filling these depressions.

Flow paths derived from the resulting DEM after the automated hydrological correction mostly corresponded closely to rivers visible on medium and high resolution satellite images (Landsat and SPOT 5). However, a few problems were experienced for larger rivers in flat areas, e.g. the Vet and Vaal rivers in the Free State, and various other inconsistencies in the Limpopo, North West and KwaZulu–Natal provinces.

A study of existing river lines showed that the most appropriate for use in improving the flow paths would be the RQS 1:500 000 rivers coverage (by the Directorate: Resource Quality Services of the Department of Water Affairs). These river lines are based on the rivers from the 1:50 000 topographical maps inside South Africa and various other data sources for neighbouring countries. This dataset was systematically compared with Landsat and SPOT satellite imagery and where the real world rivers were not followed closely the vector dataset was edited. These rivers were burned into the 'raw' DEM at certain threshold values and the process was followed by the described automated hydrological correction.

Results and discussion

The gap filling from both the contour data and Aster data filled the SRTM gaps effectively and even the dune structure in the Namib Desert in Namibia was reconstructed well.

The burning of important rivers from the RQS data followed by the hydrological correction process ensured that DEM-derived flow paths followed real world rivers closely and in particular for the major meandering rivers. Artificially flooded terrain was limited to the bare minimum.

The re-interpolation of 90 m data to 30 m comparison for an area where both native 30 m and 90 m version2_1 data were available showed that some spatial refinement to 90 m

SRTM data is possible that is more than just superficial refinement. However, more analysis is necessary in this regard.

An accuracy assessment was done to determine the accuracy of derived river lines compared to actual river positions as shown on satellite imagery. Five hundred randomly selected locations on river lines were verified. Of these points 99.2 were within two pixels (180 m) of the actual river and 91.8% within one pixel (90 m).

It was demonstrated that SRTM data is currently the best elevation data available at a national level. The modified and corrected SRTM dataset described in the report currently represents the most hydrologically sound dataset for South Africa and the neighbouring countries area.

The following derived products were prepared from the DEM:

- Catchment boundaries (Primary, secondary, tertiary and quaternary)
- Local drain direction
- Upstream number of cells
- Slope
- Aspect
- Hillshade image.

Conclusions

A continuous DEM was successfully created for southern Africa (between 19ºS and 35ºS and 12ºE to 36ºE). The SRTM DEM was used as baseline dataset and voids were filled with elevations from 20 m contour lines and the Aster GDEM dataset. The gap-filled DEM was hydrologically improved and it was demonstrated that products such as catchment boundaries and river lines can be successfully derived from the hydrologically improved DEM.

Appendix 1 provides a list of all the datasets generated in this project. The datasets can be obtained from the Department of Water Affairs, Directorate: Spatial and Land Information Management. Mistakes reported will not be corrected on these datasets but will be noted to be incorporated when higher resolution data becomes available.

Significantly higher resolution RADAR DEMs will become available in the near future for large areas of the world. It is envisaged that the TerraSAR-X / TanDEM-X satellite (operational 2010) will image the entire Earth in unprecedented detail (http://www.infoterra.de/tandem-x_dem). A global DEM at 12 m horizontal resolution with a vertical accuracy of better than 2 m will become available in 2014.

Recommendations

The following potential research themes are recommended:

- The shortcomings of currently available river datasets are outlined in paragraph 2.3. The quality of the DEM produced during the project was limited by the accuracy and completeness of the datasets available. A first step could be to produce a similar dataset as the RQS 1:500 000 rivers on 1:250 000 scale.
- The hydrologically improved DEM will be, amongst others, suitable for the following derived products:
  - Generate geomorphological features, i.e. terrain morphological units (mapping of crests, midslopes, footslopes and valley bottoms) that can be used for soil mapping, erosion potential mapping and flood analysis;
  - Mapping of wetlands.
- Quality checking / evaluating DEMs for hydrological processing. To develop quality criteria for processing DEMs for hydrological applications.

ACKNOWLEDGEMENTS

Sincere appreciation for assistance during the project is extended to the following organizations and individuals:

The Water Research Commission for funding.

The steering committee for their significant inputs to the conceptual development of the project, for their assistance and the constructive discussions during the duration of the project and the finalization of the content of this report. The Steering Committee responsible for this project consisted of the following persons:

Name	**Organization**
Mr. W Nomquphu	WRC
Dr. A Thomas	Council for Geoscience
Dr. M Silberbauer	Department of Water Affairs
Ms. C Rajah	Department of Water Affairs

The following persons from the steering committee and project team of project K5/2020 provided valuable inputs:

Name	**Organization**
Ms. H van Deventer	CSIR
Mr. A Maherry	CSIR
Ms. D Hardwick	University of the Witwatersrand
Mr. M Horan	University of KwaZulu-Natal
Ms. N Fourie	Department of Water Affairs
Prof. R Schulze	University of KwaZulu-Natal
Dr. PM Illgner	Private

Mr. Philip Beukes, Mr. Richard Tswai (ARC-ISCW) and Ms. Leaza van Wyk (IRIS International) for technical assistance.

The support from Mr. Danie van der Spuy and Mr. Pieter Rademeyer of the Department of Water Affairs with the preparation of catchment boundaries is highly appreciated.

Mr. Johan Ferguson, Mr. Philip Beukes and Mr. Phila Sibandze (ARC-ISCW) for packaging the data.

Dr. Thomas Fyfield (ARC-ISCW) for editing this report.

TABLE OF CONTENTS

List of Tables

List of Figures

LIST OF ABBREVIATIONS

ARC – Agricultural Research Council

ASTER – Advanced Spaceborne Thermal Emission and Reflection Radiometer

DEM – Digital elevation model

DWA – Department of Water Affairs

GDEM – Global Digital Elevation Model

GIS – Geographical Information System

IRIS – Integrated Resource Information Systems

KZN – KwaZulu-Natal

METI – Ministry of Economy, Trade, and Industry

MFD – Multiple Flow Direction

NASA – United States National Aeronautics and Space Administration

NED – United States National Elevation Dataset

NGA – National Geospatial-Intelligence Agency

NGI – Chief Directorate of National Geospatial Information

NIMA – National Imagery and Mapping Agency

QA file – Quality Assessment file

RADAR – Radio detection and ranging

RAK – River Awareness Kit

RQS – Resource Quality Services (Department of Water Affairs)

SA – South Africa

SFD – Single Flow Direction

SRTM – Shuttle Radar Topography Mission

STS-99 – A Space Shuttle Endeavour mission, which launched on 11 February 2000

SWBD – STRM Water Body Data

TIFF – Tagged Image File Format

USA – United States of America

WGS84 – World Geodetic System – 1984

WRC – Water Research Commission

GLOSSARY

Aspect	Compass direction of steepest downhill slope
Digital elevation model	A digital elevation model (DEM) represents the spatial distribution of elevations above some arbitrary datum in a landscape
Elevation	Height above mean sea level or local reference
Flow accumulation	The accumulated flow as the accumulated weight of all cells flowing into each downslope cell in the output grid
Hillshade	A shaded relief grid derived from a DEM by considering the illumination angle and shadows
Hydrology	The study of water as a complex but unified system on the earth. This includes the study of both subsurface and surface pathways of water flow. Flow across a surface will always be in the steepest down-slope direction
Pour point	An outlet or pour point is the point at which water flows out of an area. This is the lowest point along the boundary of the watershed
Ridge	Cells with no upstream contributing area
Sink	A sink is defined as a cell or a set of spatially connected cells whose flow direction cannot be assigned one of the eight valid values in a flow direction raster. In other words, a depression is a grid cell that does not have a lower neighbour and therefore represents a barrier to flow
Slope	Rate of change of elevation
Watershed	A watershed is an area that drains water and other substances to a common outlet as concentrated drainage. Other common terms for a watershed are: basin, catchment, drainage region or contributing area. This area is normally defined as the total area flowing to a given outlet, or pour point. The boundary between two watersheds is referred to as a watershed boundary, catchment boundary or drainage divide

1. INTRODUCTION

1.1. Background

Digital elevation models (DEMs) represent the topography in a landscape and are an important data source used in numerous hydrological and geo-hydrological models. There are several DEMs available for South Africa, but usually they require significant editing before they can be used for hydrological modelling. The aim of the project is to prepare a hydrologically improved Digital Elevation Model for South Africa based on the Shuttle Radar Topography Mission (SRTM) dataset.

The SRTM obtained high-resolution elevation data on a near-global scale in February 2000. SRTM was a cooperative project between the United States National Aeronautics and Space Administration (NASA) and the United States National Imagery and Mapping Agency (NIMA) (Farr and Cobrick, 2000). A single-pass radar interferometer onboard the Space Shuttle Endeavour was used during an 11-day mission to cover the earth's surface between latitudes ~60° north and 56° south. A global assessment of the SRTM performance has shown that the absolute height error exceeds the mission goal of 16 m by almost a factor of two (Rodrígues *et al.*, 2006). The SRTM products are being distributed at a resolution of 1 arc sec (~30 m) for the USA and 3 arc sec (~90 m) for the rest of the globe. These products contain areas of missing data which are mostly associated with water bodies, mountainous regions and desert regions.

Another global high-resolution DEM, with an estimated accuracy of 20 m and 30 m in the vertical and horizontal dimensions respectively was released on 29 June 2009 (Aster GDEM Validation Team, 2009). The Aster Global Digital Elevation Model (GDEM) was jointly released by the Japanese Ministry of Economy, Trade, and Industry (METI) and the United States National Aeronautics and Space Administration (NASA). Automated stereo correlation of the entire Aster archive (1.5 million-scenes) was done to produce a DEM between 83° north and 83° south latitudes. The Aster GDEM dataset is available at a resolution of 1 arc sec (~30 m). METI and NASA state that version 1 of the Aster GDEM should be viewed as "experimental" or "research grade". However, they have decided to release it because they feel its benefits outweigh its flaws. The data is accompanied with quality assessment files that show the number of scene-based DEMs contributing to the final DEM value for each pixel. The Aster GDEM includes void pixels for which no cloud-free scenes are available. These voids were patched using SRTM, NED, and other DEM data from different sources. Figure 1 shows the Aster GDEM global stacking number (numbers of

Aster DEMs contributing to the GDEM) by location. From this image it is clear that the stacking number over South Africa is particularly low.

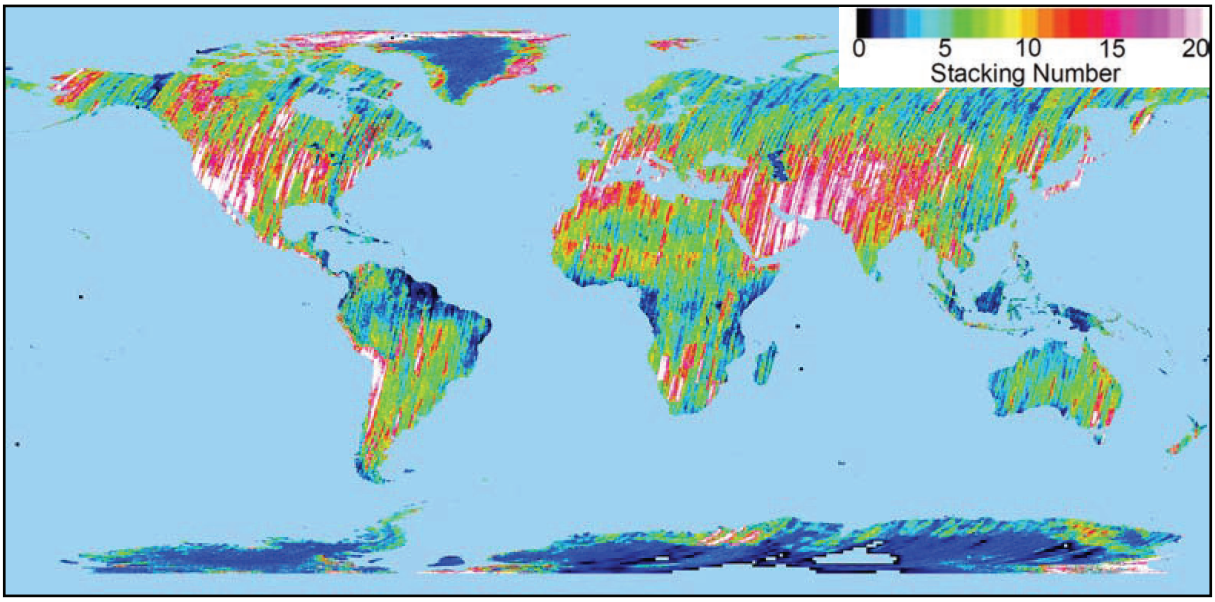


Figure 1: Aster GDEM global stacking number map showing numbers of Aster DEMs contributing to the GDEM by location (ASTER GDEM Validation Team, 2009)

The Chief Directorate of National Geospatial Information provides two types of elevation data for the whole of South Africa, namely:

- Contour lines with a 20 m interval and spot heights from the 1:50 000 topographical maps;
- A regularly spaced grid of data points (400 m x 400 m) derived from stereo aerial photographs.

Hydrology is the study of water as a complex but unified system on the earth (Strahler and Strahler, 1997). This includes the study of both subsurface and surface pathways of water flow. Flow across a surface will always be in the steepest down-slope direction.

Strahler and Strahler (1997: p. 373) describe drainage systems as follows:

> "As runoff moves to lower and lower levels and eventually to the sea, it becomes organised into a drainage system. The system consists of a branched network of stream channels, as well as sloping ground surfaces that contribute overland flow to those channels. Between the channels on the ridges are drainage divides, which mark the boundary between slopes that contribute water to different streams or drainage systems. The entire system is bounded by a drainage divide that outlines a more-or-less pear shaped drainage basin. "

Flow direction of cells in a DEM can be calculated with the raster analyses components of Geographical Information System (GIS) packages such as ArcGIS. Once the direction of flow out of each cell is known, it is possible to determine which and how many cells flow into any given cell. This information can be used to define watershed boundaries and stream networks. Figure 2 provides a schematic illustration of the sequence of steps to be executed in deriving surface characteristics from the DEM (ESRI, 1996).

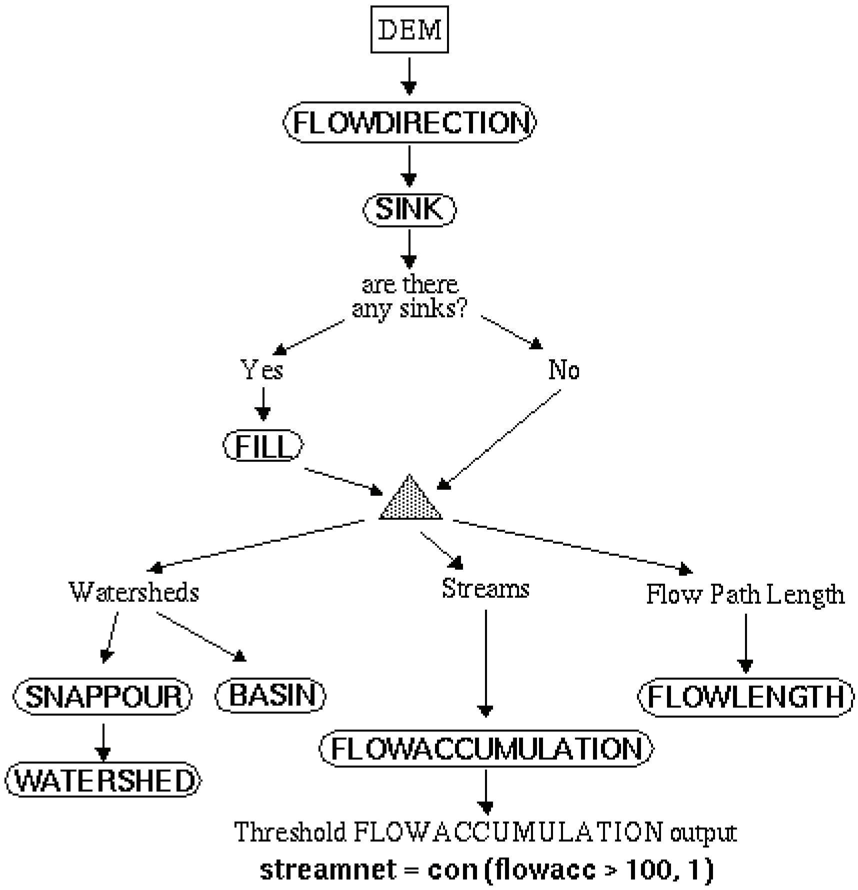


Figure 2: Outline of the steps to derive surface characteristics from a DEM (ESRI, 1996)

Regular-grid DEMs contain numerous depressions or sinks. A depression is defined as a cell or a set of spatially connected cells whose flow direction cannot be assigned one of the eight valid values in a flow direction raster. In other words, a depression is a grid cell that does not have a lower neighbour and therefore represents a barrier to flow (Gallant and Wilson, 2000). In most cases the depressions are artificial and have to be edited to allow flow to continue across the surface. Gallant and Wilson (2000) suggest use of the following methods to handle depressions:

- Filling depressions by increasing internal elevations to their lowest outflow point;

- Reducing by a fixed amount the elevations of cells along mapped stream lines so that the streams are at a lower elevation than the surrounding landscape (also referred to as “burning in” of streams); and
- Incorporating streamlines during the construction of the DEM.

Flow directions are computed using either the MFD (Multiple Flow Direction) model or the SFD (Single Flow Direction, or D8) model. Both methods compute downslope flow directions by inspecting the 3-by-3 window around the current cell. The SFD method assigns a unique flow direction towards the steepest downslope neighbour. The MFD method assigns multiple flow directions towards all down slope neighbours.

1.2. Aim and objectives

The aim of the project is to prepare a hydrologically improved DEM for South Africa based on the SRTM dataset.

The objectives of the project were to:

1. Research methodologies for producing hydrologically correct SRTM datasets.
2. Produce a hydrologically corrected, gap and sink filled SRTM dataset for South Africa.
3. Re-interpolate the hydrologically corrected SA DEM to 30 m and derive a hill-shading layer from this DEM for improved visual display purposes.
4. Do a comparative analysis for a small study area in the USA where both 30 m and 90 m SRTM datasets are available and use this as a basis for making recommendations on the value of re-interpolating the SA hydrologically corrected DEM to 30 m and deriving secondary datasets from the 30 m DEM.
5. Correct inaccurate major river flow patterns for flat areas.
6. Prepare derived products from the DEM.

2. METHODOLOGY

2.1. Study area

The study area includes the whole of South Africa as well as the complete Limpopo and Orange River basins due to the importance of catchment areas in hydrological studies. Drainage basins have been considered to be some of the most important areal units in environmental decision making (EU, 2000, as cited by Oksane and Sarjakoski, 2005), because they delineate the land area of runoff-driven sediment transport. Examples of transboundary drainage basin projects include the Orange-Senqu River Awareness Kit project (ORASECOM, 2009) and the Limpopo Basin Focal Project (Sullivan and Sibanda, 2010).

The Limpopo Basin Focal Project is aimed at reducing poverty and enhancing food security, access to health and environmental security in the Limpopo River Basin (Sullivan and Sibanda, 2010). This will be achieved by analyzing the status of agricultural water use, access and productivity, and identifying opportunities for poverty reduction through improved agricultural water management. The basin has an average rainfall of 530 mm/yr, but extreme variability makes agriculture very risky, and underlies high levels of poverty and malnutrition in all four basin countries, namely Botswana, Mozambique, South Africa and Zimbabwe.

The Orange-Senqu River Awareness Kit (RAK) project created an Awareness Kit for the Orange-Senqu River Basin Commission (ORASECOM, 2009). The Orange-Senqu RAK is an information and knowledge management tool for the Orange-Senqu River Basin, designed to support capacity development and sustainable management of the environment and resources within the basin. The Orange-Senqu River Basin encompasses all of Lesotho, large portions of South Africa, southern Namibia and south-western Botswana.

Figure 3 shows the conterminous extent of the new study area's river basins (bounded by the red box) between 19ºS and 35ºS and 12ºE and 36ºE.

Figure 3: Study area of the project (between 19ºS and 35ºS and 12ºE to 36ºE)

2.2. Acquisition of baseline datasets

As described in the introduction there are currently two global high-resolution DEMs available, namely the SRTM and Aster GDEM. Although the Aster GDEM has a better spatial resolution, the data is still considered as experimental, or research grade. Additionally contour lines with a 20 m interval at a scale of 1:50 000 are available from the Chief Directorate of National Geospatial Information for the whole of South Africa. The SRTM data has an advantage over contour datasets in that it is sampled from regularly spaced data points whereas the contour-derived dataset is prone to interpolation errors, particularly in flat areas (Department of Water Affairs and Forestry, 2006). These differences affect the accuracy of streams and catchments derived from elevation data. The SRTM dataset is therefore the most appropriate dataset to use.

Figure 4 illustrates the differences between 20 m interval contour lines, SRTM DEM and Aster GDEM. A SPOT 5 satellite image from the year 2008 is included as reference. The forest plantations in this example differ between the three images because of the different periods represented. The SRTM data were gathered in February 2000 and the Aster GDEM represents an average of DEMs over several years. It is, however, important to note that for both SRTM and Aster GDEM the plantations are visible, illustrating that tree canopy elevation is sometimes measured instead of ground elevation. The 20 m interval contour lines are drawn as black lines over the three images.

From this example it is clear that the SRTM DEM has an additional level of detail in flat areas; of specific interest for this study are the additional surface pathways of water flow. The typical artifacts of the Aster GDEM can also be seen in this image. According to the ASTER GDEM Validation Team (2009) these pits, or small negative anomalies, occur with regularity and often high frequency in virtually all ASTER GDEM tiles examined during these validation exercises. The magnitude of the negative elevation anomaly associated with the pits varies dramatically from just a few metres to 100 m or more. “Bumps” appear to be the positive elevation-anomaly-artifact-equivalent to pits. They occur with similar frequency as pits, and they also were found to be present in virtually every ASTER GDEM tile examined during the validation studies.

In this example the Aster GDEM totally misrepresented Lake Mzingazi with elevation values ranging from 4 m to 211 m inside the lake.

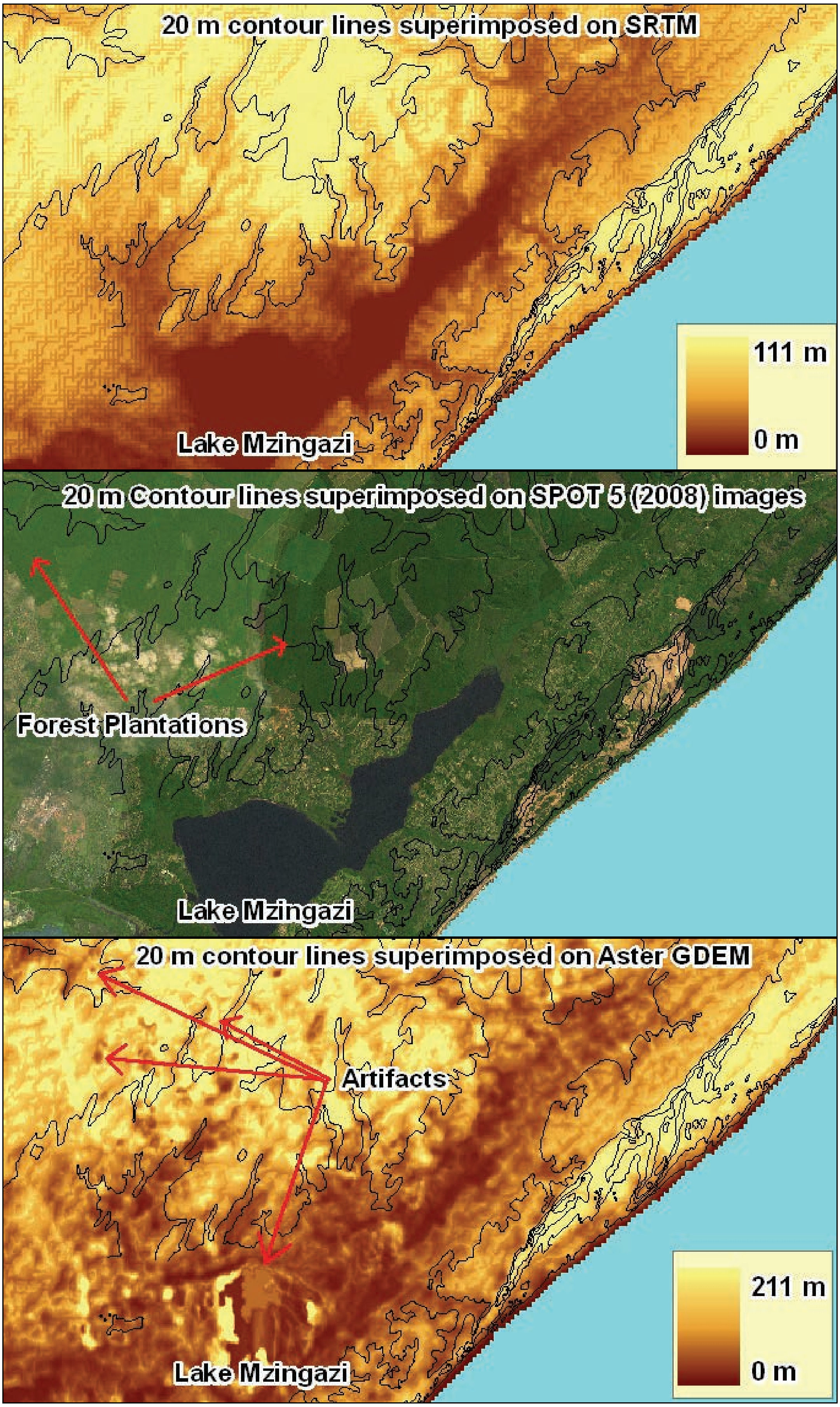


Figure 4: Comparison of 20 m interval contour lines, SRTM DEM and Aster GDEM

2.2.1. SRTM dataset

The following SRTM versions are available:

- Version 1 – consists of the original Digital Elevation Models produced by the SRTM project with data from the STS-99 mission of February 2000 as delivered to the National Geospatial-Intelligence Agency (NGA). These data are unedited and contain spurious data points in areas of low radar backscatter such as water bodies.
- Version 2 – is the result of substantial editing efforts by the NGA and exhibits well-defined water bodies and coastlines and absence of spikes and wells (single pixel errors), although some areas of missing data (voids) are still present. Its directory contains the vectorized coastline mask in shapefile format; the SRTM Water Body Data (SWBD) used by NGA in the editing process.
- Version 2.1 – is a recalculation of the degraded 3 arc sec dataset by 3x3 averaging of the full resolution-edited data (Figure 5). Version 2 was generated by masking edited samples from its lower-resolution counterpart which was released to the public by the NGA, and contained occasional artifacts notably slight vertical “banding” in data beyond 50° latitude. These artifacts have been eliminated in Version 2.1.

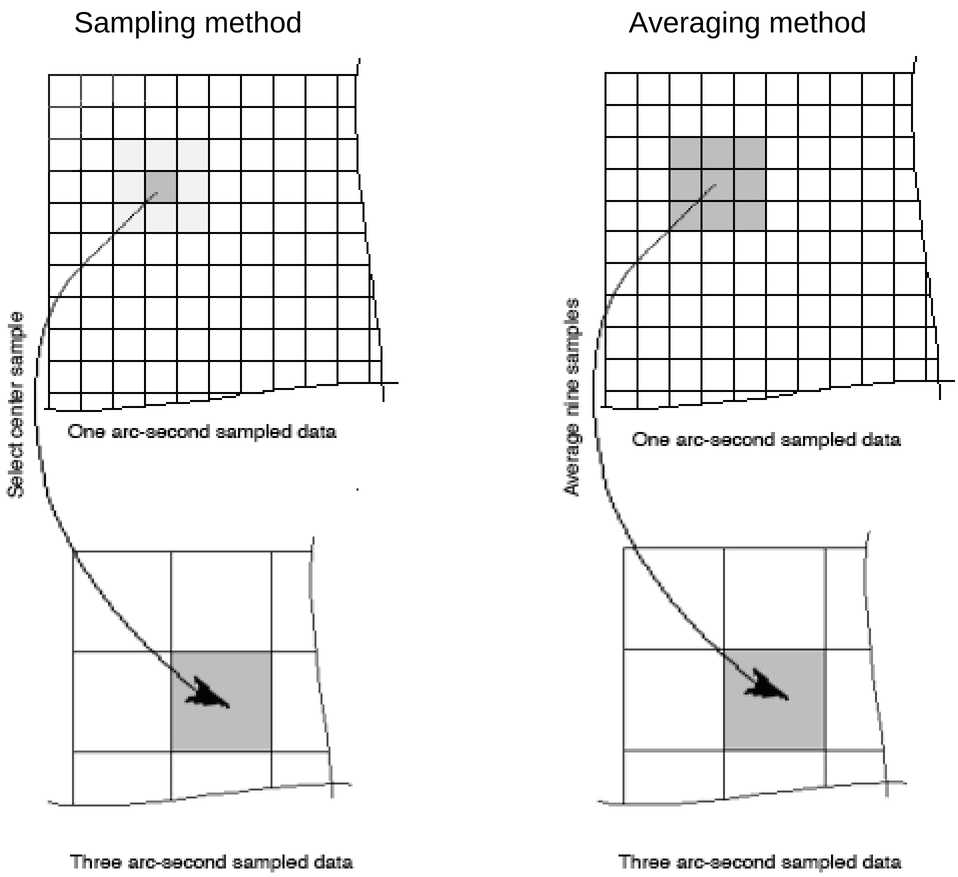


Figure 5: Sampling methods for SRTM version 1 and version 2.1 (USGS, s.a.)

Most analysts agree that the averaging method produces a superior product by decreasing the high frequency ‘noise’ that is characteristic of radar-derived elevation data (USGS, s.a.).

This is similar to the conventional technique of 'taking looks', or averaging pixels in radar imagery to decrease the effects of speckle and to increase radiometric accuracy, although at the expense of horizontal resolution.

SRTM data can be downloaded as tiles from http://dds.cr.usgs.gov/srtm/. Version 2.1 (the official version of the data) was used for this project. Various unofficial versions of the data exist, e.g. version 3 and 4 (http://srtm.csi.cgiar.org), but these are heavily processed and positional shifts have been introduced. Figure 6 shows the 288 tiles that were downloaded for the study area. Individual data tiles are named by reference to the longitude and latitude of the centre of the lower-left (southwest) corner cell of each tile (USGS, s.a.). For example, the centre coordinates of the lower-left corner cell of tile S30E017 are latitude 30 degrees south and longitude 17 degrees east.

The SRTM tiles are downloaded as .hgt files without header information. Tile S30E017.HGT is used as an example to illustrate the procedure followed before the file can be imported into a GIS software program. The file extension has to be renamed from .hgt to .bil (i.e. S30E017.HGT to S30E017.BIL). The world and header files can be created with a text editor as illustrated in Figures 7 and 8.

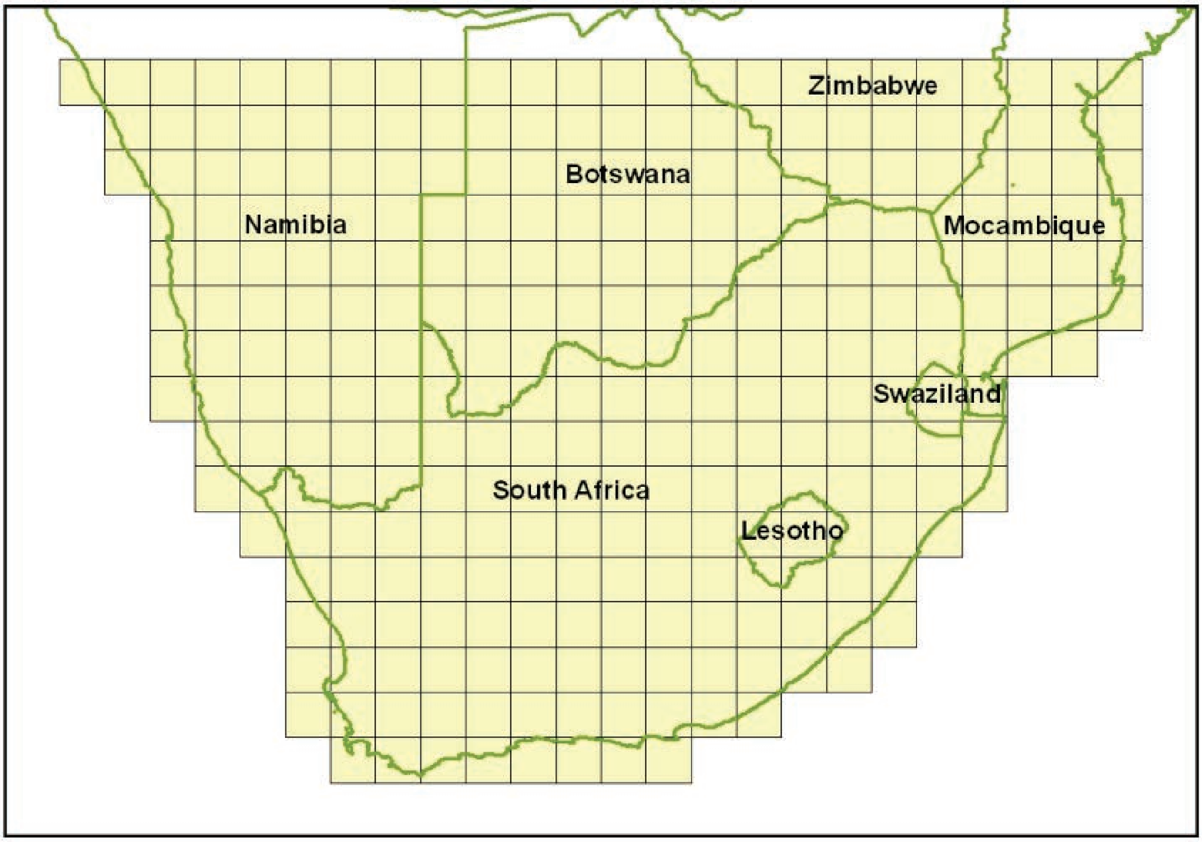


Figure 6: Index grid of SRTM tiles downloaded for the study area

```
0.000833333333333
0
0
-0.000833333333333
17
-29
```

Figure 7: Content of the world file (S30E017.BLW)

```
BYTEORDER      M
LAYOUT       BIL
NROWS        1201
NCOLS        1201
NBANDS       1
NBITS        16
BANDROWBYTES        2402
TOTALROWBYTES       2402
BANDGAPBYTES        0
NODATA       -32768
ULXMAP     17
ULYMAP     -29
XDIM         0.000833333333333
YDIM         0.000833333333333
```

Figure 8: Content of the header file (S30E017.HDR)

A major potential source of error in SRTM data is vegetation (Van Niel *et al.*, 2008). According to Weydahl *et al.* (2007), SRTM elevations may show ground elevations in bare agricultural landscapes, while dense forest areas will introduce an additional elevation height approximately corresponding to tree canopy height.

2.2.2. Aster GDEM dataset

The Aster GDEM dataset can be downloaded from: http://www.gdem.Aster.ersdac.or.jp/index.jsp. The Aster GDEM data is also distributed in 1°- x -1° tiles as illustrated in Figure 5 for SRTM data and distributed as a zip-compressed DEM and quality assessment (QA) file (Figure 9). These files are stored in GeoTIFF format and can be opened directly with GIS software packages. The total size of the 288 uncompressed DEM files for the study area is 6.8GB. The size of QA files is also 6.8GB.

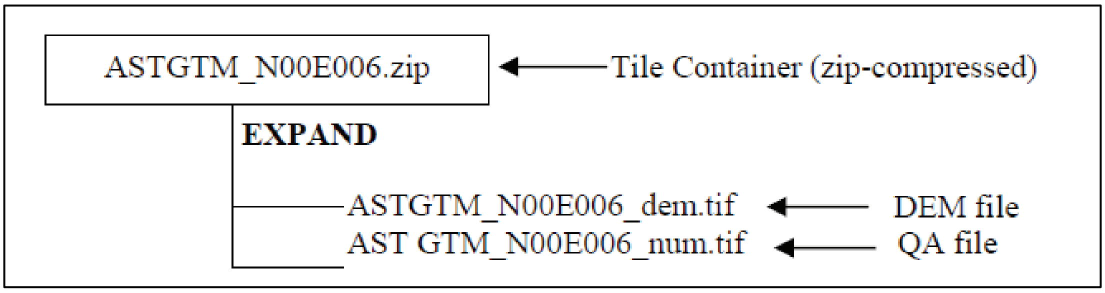


Figure 9: GDEM file structure

The 30 m Aster GDEM was degraded to 90 m and used to patch areas of the study area outside South Africa where SRTM data are not available.

2.2.3. Contour lines

GISCOE created a 20 m regular-grid DEM from the 1:50 000 scale contour lines and spot heights (GISCOE, 2001) by using the 'topogrid' function in ArcInfo 7.x (ESRI, 1996). This function is optimized for creating hydrologically correct DEMs and is based on Hutchinson's (1989) interpolation method. The 20 m spatial resolution DEM created by GISCOE was degraded to 90 m and used to patch areas with SRTM data holes over South Africa.

2.2.4. Comparison of 30 m and 90 m resolution SRTM DEMs

One arc-second (30 m resolution) DEM data were acquired for an area over the USA (where native 30 m data and degraded 90 m data are available) and gap filling was done on the data. SRTM Version 2 data at 90 m resolution were resampled to 30 m and compared with the 90 m data and the original 30 m data. The differences between the 90 m, the interpolated 30 m and the original 30 m data are shown in Figures 10 and 11 respectively. Flow paths were calculated from the interpolated 30 m data and the original 30 m data and compared (Figure 11). As shown in Figure 12, there is little difference between the flow paths of the interpolated 30 m data and native 30 m data.

The hydrologically corrected DEM was resampled to 30 m (1 arc second) and a hillshade was derived for this layer. It is important to note that though this layer improves visual display it does not improve on the detail of the 90 m DEM.

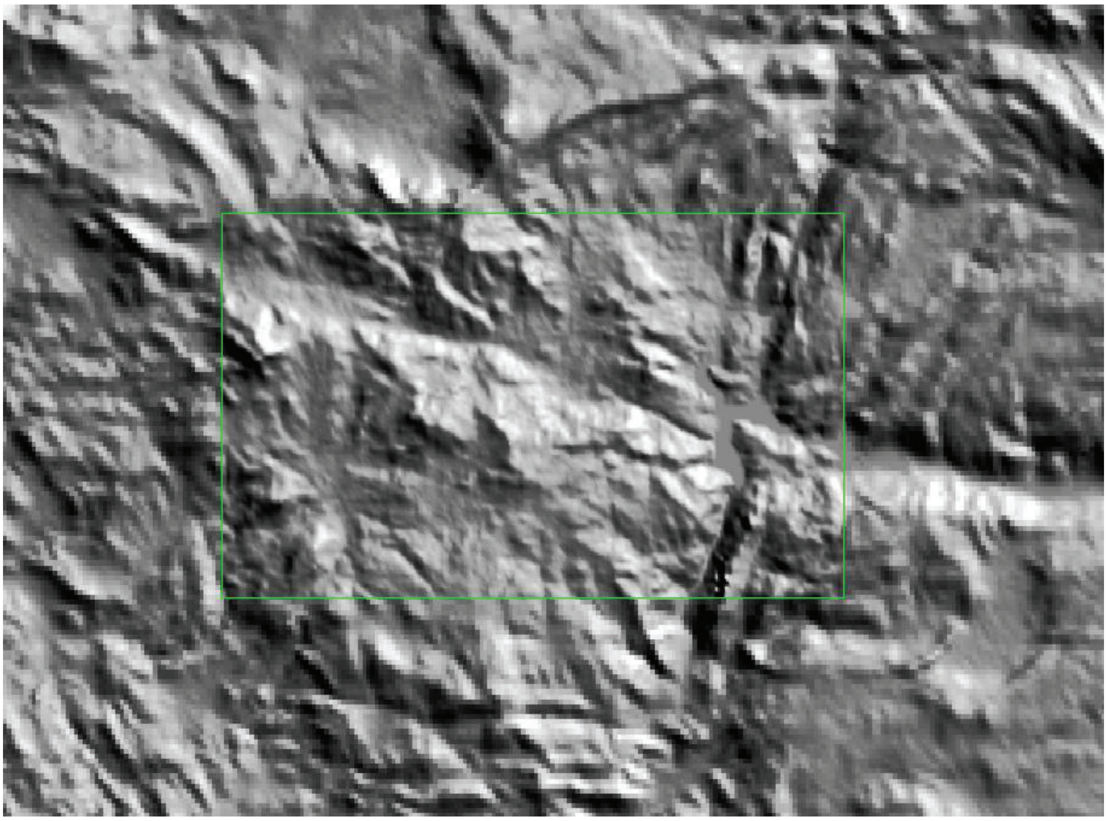


Figure 10: The 90 m version 2 data are in the background and the 90 m data interpolated to 30 m are in the green rectangle

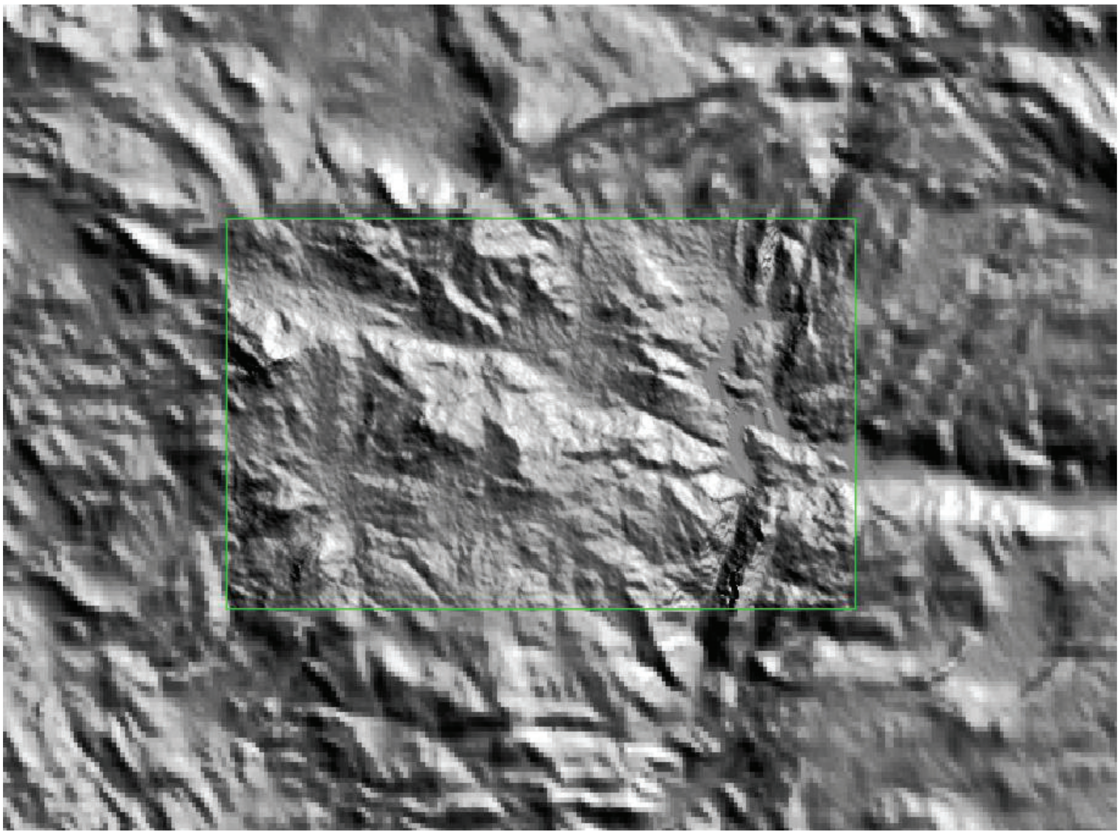


Figure 11: The 90 m version 2 data are in the background and the original 30 m data are in the green rectangle

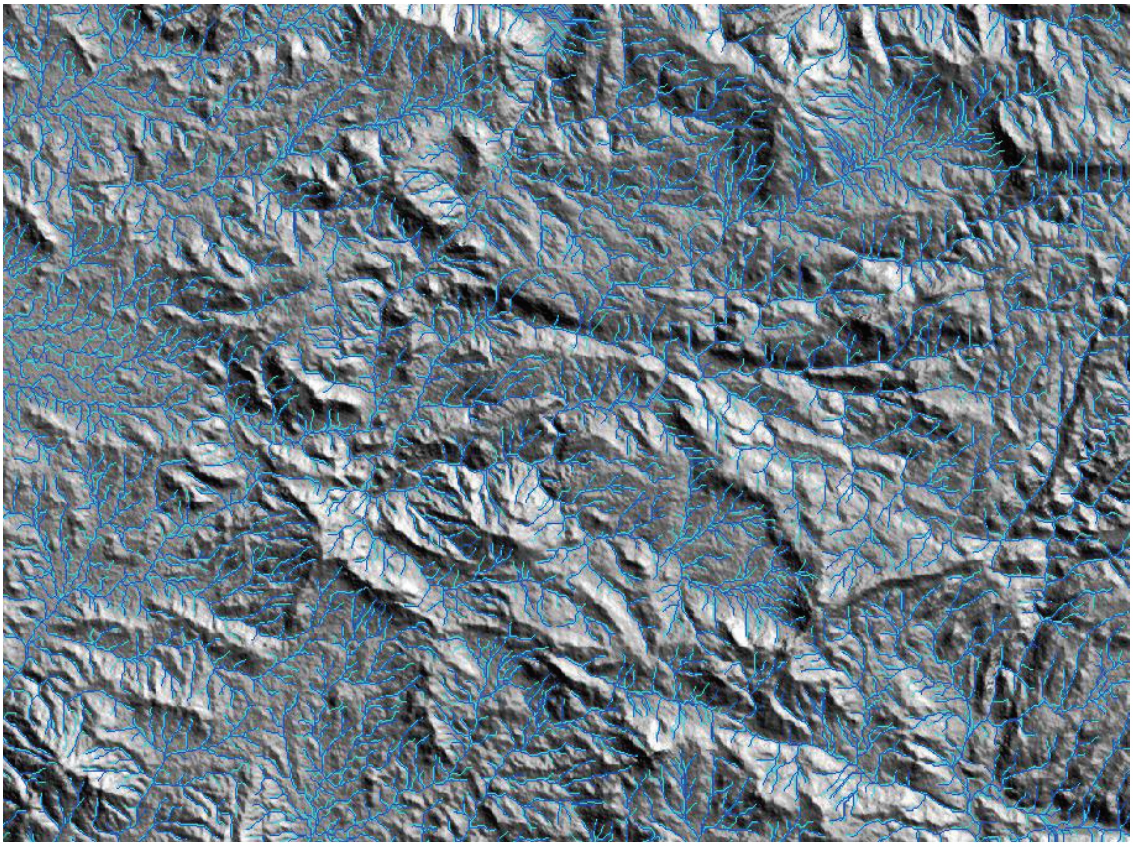


Figure 12: Flow paths from 30 m interpolated data (blue) over 30 m original data (cyan) showing a very close match

2.3. Hydrological correction

GRASS software (http://grass.osgeo.org/) was used for performing an automated hydrological correction based on the impact reduction approach of Lindsay and Creed (2005). In the original approach, the impacts of sink filling and channel carving on each sink are compared and the method causing fewer modifications is chosen. In the approach applied on the DEM, for each sink the area to be filled from the spill point and the length of the channel to be carved are determined first. If channel length is shorter than a user-defined threshold (here 4 cells), a channel is carved. If the area to be filled is smaller than a user-defined threshold (here also 4 cells), the sink is filled completely. Otherwise, the sink is incrementally eliminated from the bottom and on each step the impact of filling this sink a little and carving out for the partially filled sink is determined. When a least impact equilibrium of partial filling and partial carving is reached, the sink is removed by using the current parameters for partial filling and partial carving. This method reduces the number of modifications necessary to produce a hydrologically correct DEM and avoids typical problems of pure sink filling (artificially flooded terrain) and pure carving broader and visible one cell wide streams with Multiple Flow Direction (surface flow accumulation). Natural depressions in the landscape must be maintained in the DEM. Examples are areas where natural and man-made water bodies and mines occur. These include fluctuating water bodies like annual pans. A great deal of care was taken to avoid filling these depressions.

Flow paths derived from the resulting DEM after the automated hydrological correction were mostly corresponding closely to rivers visible on medium and high resolution satellite images (Landsat and SPOT 5). However, a few problems were experienced for larger rivers on flat land, e.g. the Vet and Vaal rivers in the Free State, and various other inconsistencies in the Limpopo, North West and KwaZulu-Natal provinces.

To solve this problem it was decided to ‘burn’ the RQS 1:500 000 rivers (Department of Water Affairs and Forestry, 2006) into the DEM before the stream carving and filling process was applied.

The RQS river dataset was chosen because it represented only significant river lines and was compared to the 1:50 000 topographical river data for South Africa (Figure 13). Figure 13 shows an overwhelming amount of small streams in the 1:50 000 dataset that would significantly alter the SRTM data should it be applied in the correction process. The other problem with the 1:50 000 dataset is that there are numerous interpreter inconsistencies for the minor streams between adjacent 1:50 000 standard SA map sheets. The 1:50 000 rivers are the data source for rivers in the RQS river dataset for the South Africa and areas close to the South African border. In the South African part of the dataset there is a near perfect fit of the RQS rivers on the 1:50 000 rivers (mostly ± 50 m). For neighbouring countries other data sources of varying scales (as small as 1: 1 000 000) have been used in the RQS river dataset. The RQS river dataset was also systematically compared to Landsat and SPOT satellite data for river line accuracy. The pre-river line flow accumulation layer was also compared simultaneously to the RQS rivers. Where there were significant differences the satellite image river was taken as the reference and the RQS river line was either edited or deleted depending on how close the flow accumulation layer was to the satellite image river. Outside the South African area many rivers were completely removed and a limited number were re-digitized. Only minor edits were made over the South African area.

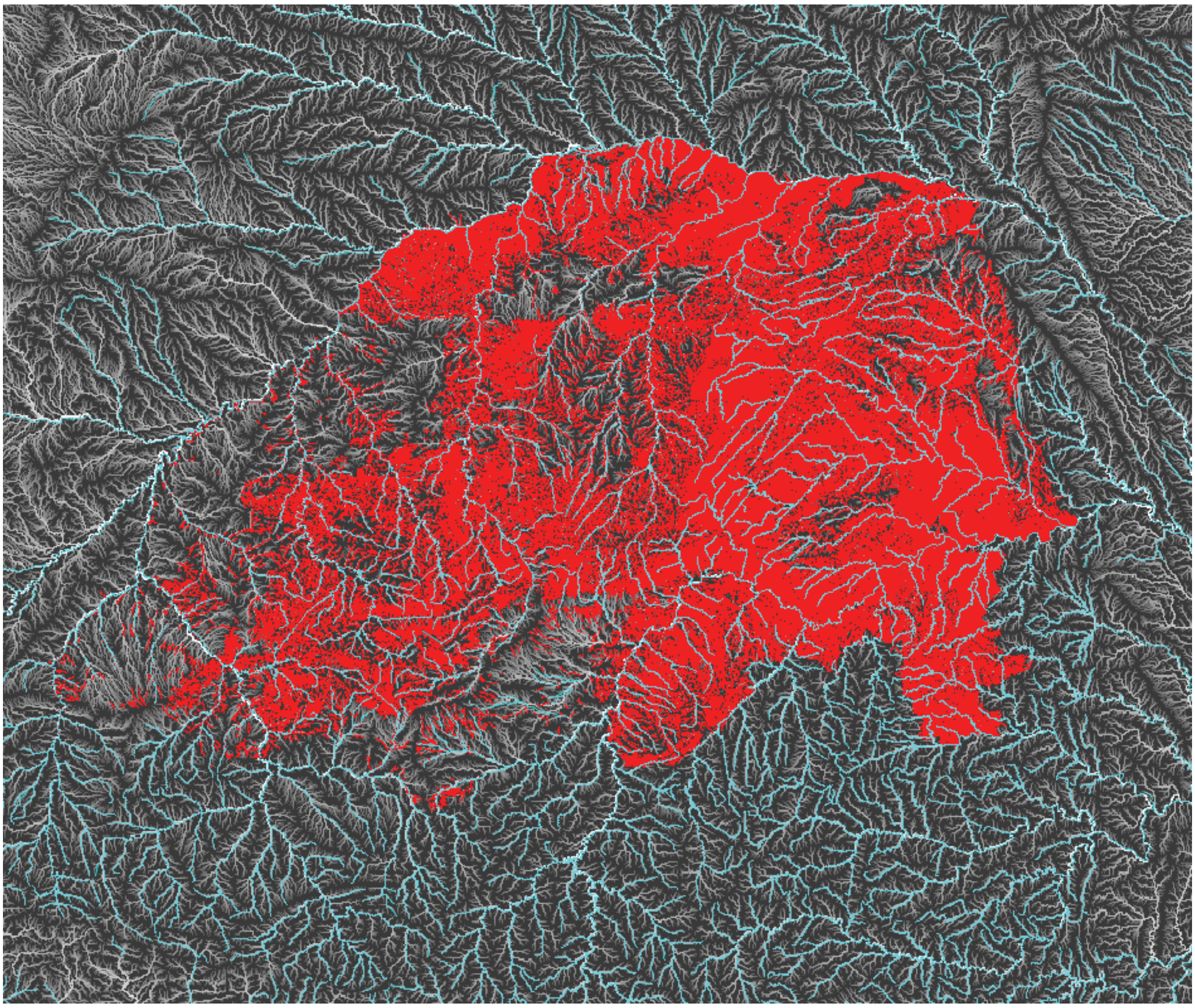


Figure 13: The 1:50 000 and the RQS river datasets for the Limpopo Province are overlaid on the SRTM flow accumulation. The 1:50 000 rivers are shown in red and the RQS rivers in cyan

The original gap-filled SRTM elevation was lowered by 7 m under a river line only if the lowered elevation was larger than the minimum elevation for this stream segment, otherwise the minimum segment elevation was used. This prevented artifacts of too deeply carved rivers and avoided carving a river into a lake. The minimum threshold of 7 m was important in order to avoid missing some rivers in coastal plains. A threshold larger than 7 m would not have improved the river courses and would have led to heavier but unnecessary modifications of the original data. The final stream carving and filling process was applied to this dataset to create the hydrological corrected SRTM dataset.

2.4. Derived products

2.4.1. Local drain direction

One of the keys to deriving hydrologic characteristics about a surface is the ability to determine the direction of flow from every cell in the grid. The Flowdirection function in

ArcGIS was used to calculate local flow direction (ESRI, 2008). This function takes a surface as input and outputs a grid showing the direction of flow out of each cell. There are eight valid output directions, relating to the eight adjacent cells into which flow could travel. Figure 14 illustrates the values that will be allocated to a cell depending on the direction of flow.

32	64	128
16	**X**	1
8	4	2

Figure 14: Values that are allocated to indicate the eight valid flow directions from cell X

2.4.2. Upstream number of cells

For Single Flow Direction (SFD) surface the Flowaccumulation tool in ArcGIS (ESRI, 2008) was used to calculate the upstream number of cells. This function calculates accumulated flow as the accumulated weight of all cells flowing into each downslope cell in the output grid.

A second surface calculated from Multiple Flow Direction (MFD) was created using GRASS (http://grass.osgeo.org/grass62/manuals/html62_user/r.terraflow.html). Quinn *et al.* (1991) and Tarboton (1997) give an overview of the potential benefits of MFD for certain applications, e.g. erosion modelling. The MFD flow accumulation performed well over pans (Figure 15) and over dunes (Figure 16).

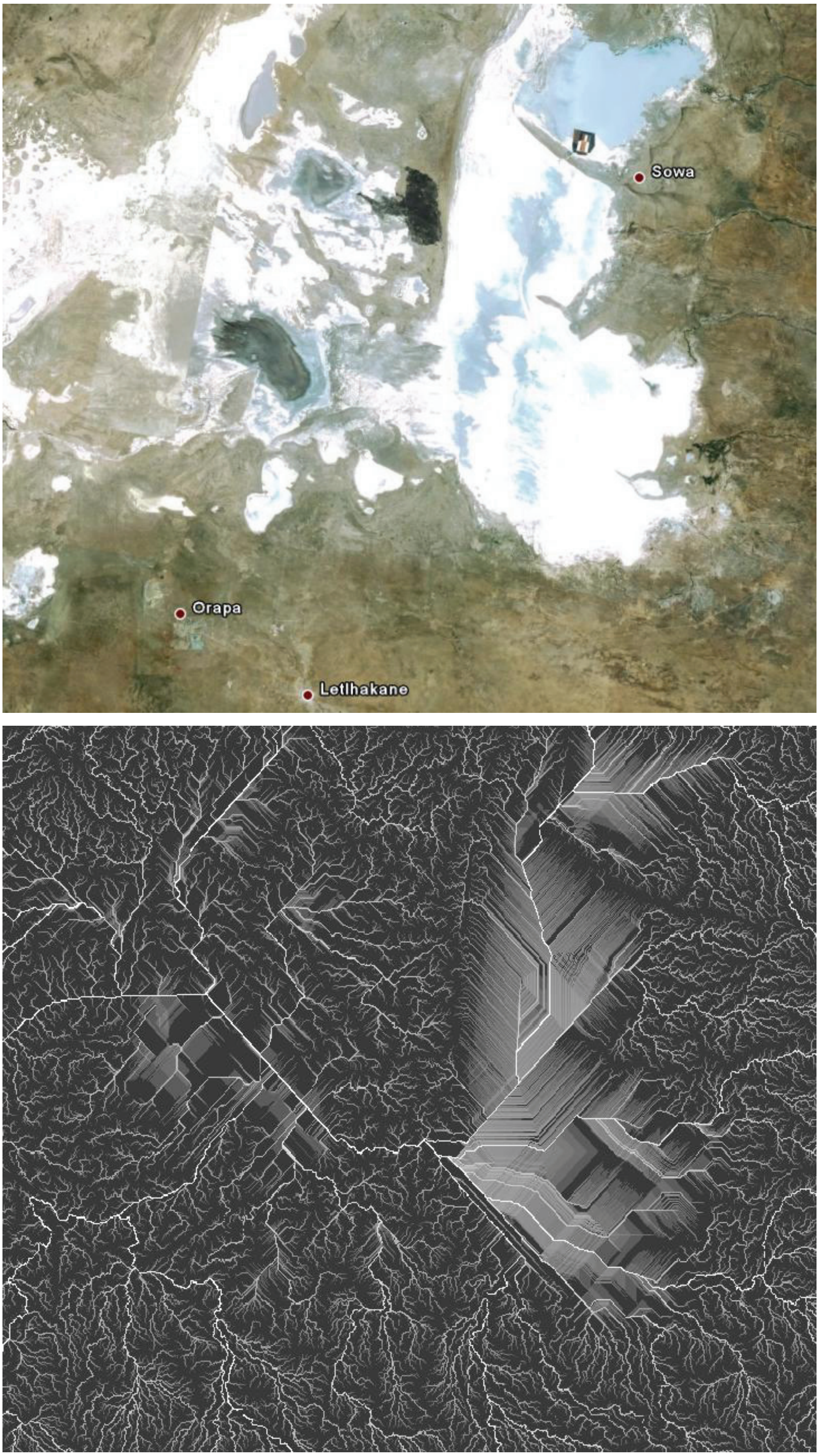


Figure 15: Quickbird image (top) and flow accumulation (bottom) of the Makgadikgadi pans

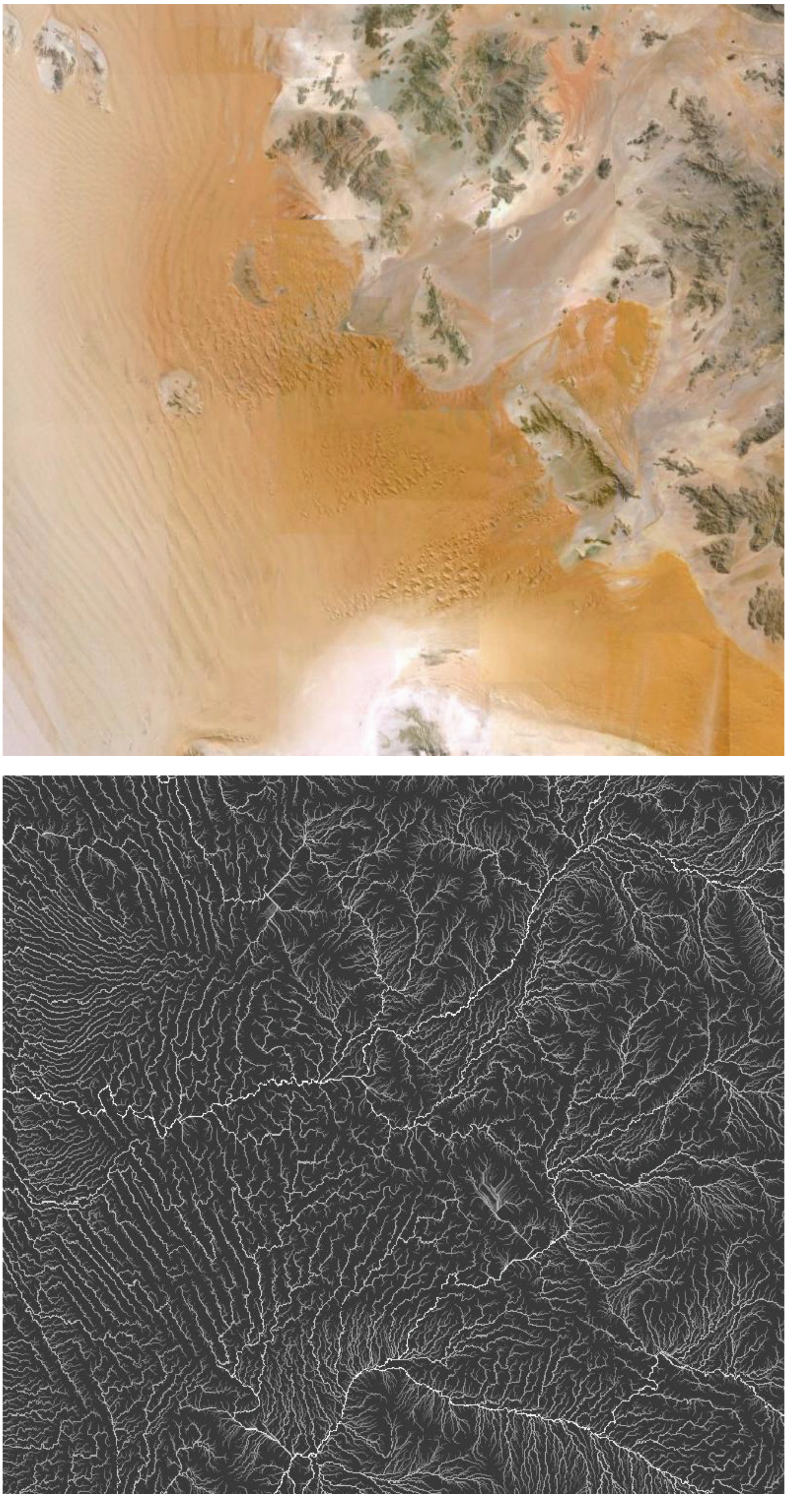


Figure 16: Quickbird image (top) and flow accumulation (bottom) of the Namib dunes

2.4.3. Slope

In order to accurately calculate slope the DEM was projected to the Albers Equal Area projection with standard parallels 18˚ South and 32˚ South. A central meridian of 24° East was used. The final product was projected back to Geographic (Datum: WGS84).

The slope function in ArcGIS (ESRI, 2008) was used. This function identifies the rate of maximum change in z value from each cell. Two slope products were prepared by calculating slope in degrees and as percent rise (also referred to as percent slope).

2.4.4. Aspect

Aspect identifies the down-slope direction of the maximum rate of change in value from each cell to its neighbours. (Aspect can be thought of as the slope direction.)

The aspect function in ArcGIS (ESRI, 2008) was used. The values of the output grid represent the compass direction of the aspect (see Figure 17).

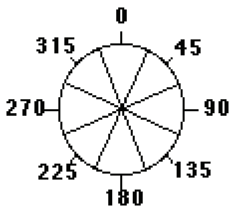


Figure 17: The values of the output grid of the aspect function represent the compass direction

2.4.5. Hillshade

The hillshade function in TNTmips (www.microimages.com) was used to create a shaded relief grid from the DEM by using an elevation angle of the sun of 50 degrees and a direction of 45 degrees.

2.4.6. Catchment boundaries

An improved set of catchment boundaries (primary, secondary, tertiary and quaternary) were prepared using the hydrologically improved DEM. Pour points were defined for each of the quaternary catchments by hydrologists from the Department of Water Affairs (Danie van der Spuy and Pieter Rademeyer). Mostly geographic recognisable points such as the intersection of tributaries with the main stream or dam walls were used as pour points. Furthermore the pour points were selected in such a way that they did not deviate too much from the previous catchments.

At the coast small rivers were often grouped together. Where possible a pour point was selected for at least every second catchment along the coastline, which then defined the left or right catchment boundary of the grouped catchment.

Figure 18 show the primary catchment boundaries.

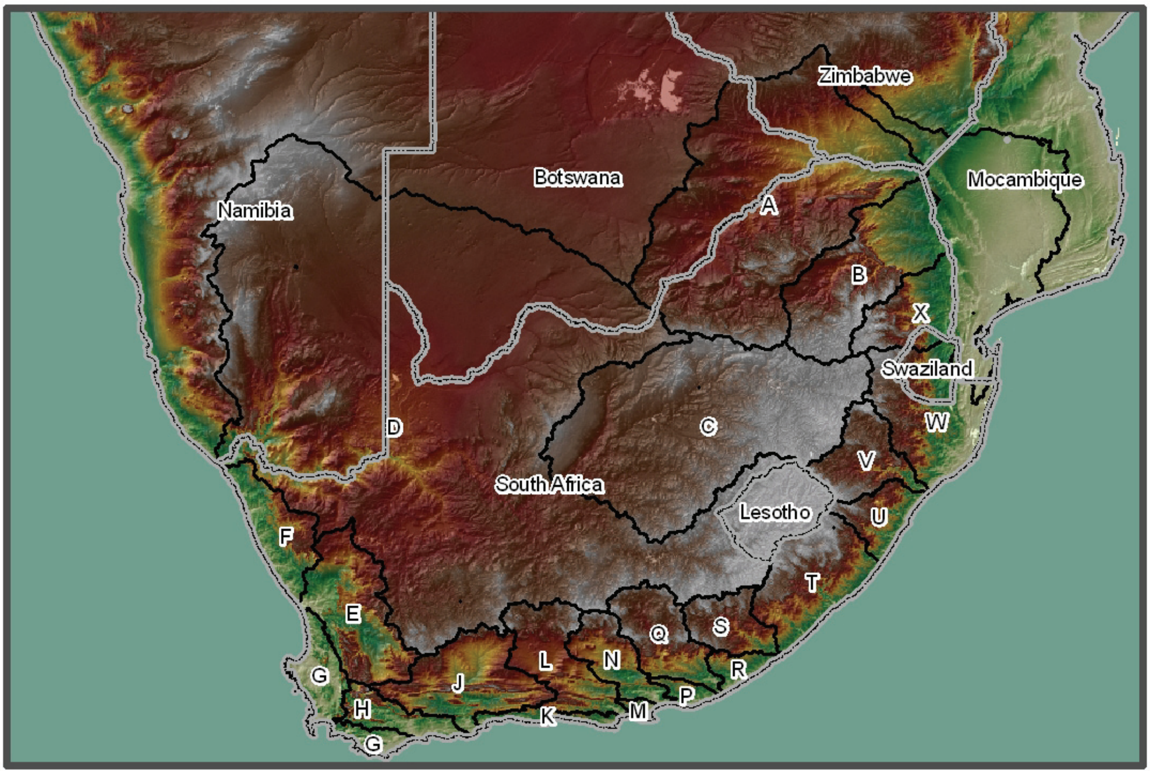


Figure 18: Primary catchment boundaries

2.5. Accuracy assessment of flow paths

The positional accuracy of flow paths derived from the hydrological corrected DEM was compared with flow paths visible on SPOT 5 and Landsat pan sharpened data GeoCover circa 2000 (https://zulu.ssc.nasa.gov/mrsid/). Five hundred random points were extracted over flow paths with a flow accumulation of 2000 or more cells (Figure 19). The 2000 cell flow accumulation was done to ensure that the more significant flow accumulation was also included and not the overwhelming number of very small flow accumulations. The distance between the position of each accuracy assessment point and the middle of the nearest flow path line as visible on the satellite imagery was measured. When the nearest flow path was on Kalahari or Namib Desert sand it was noted with a separate symbol. The same was done when the point fell on or near to a water body or pan.

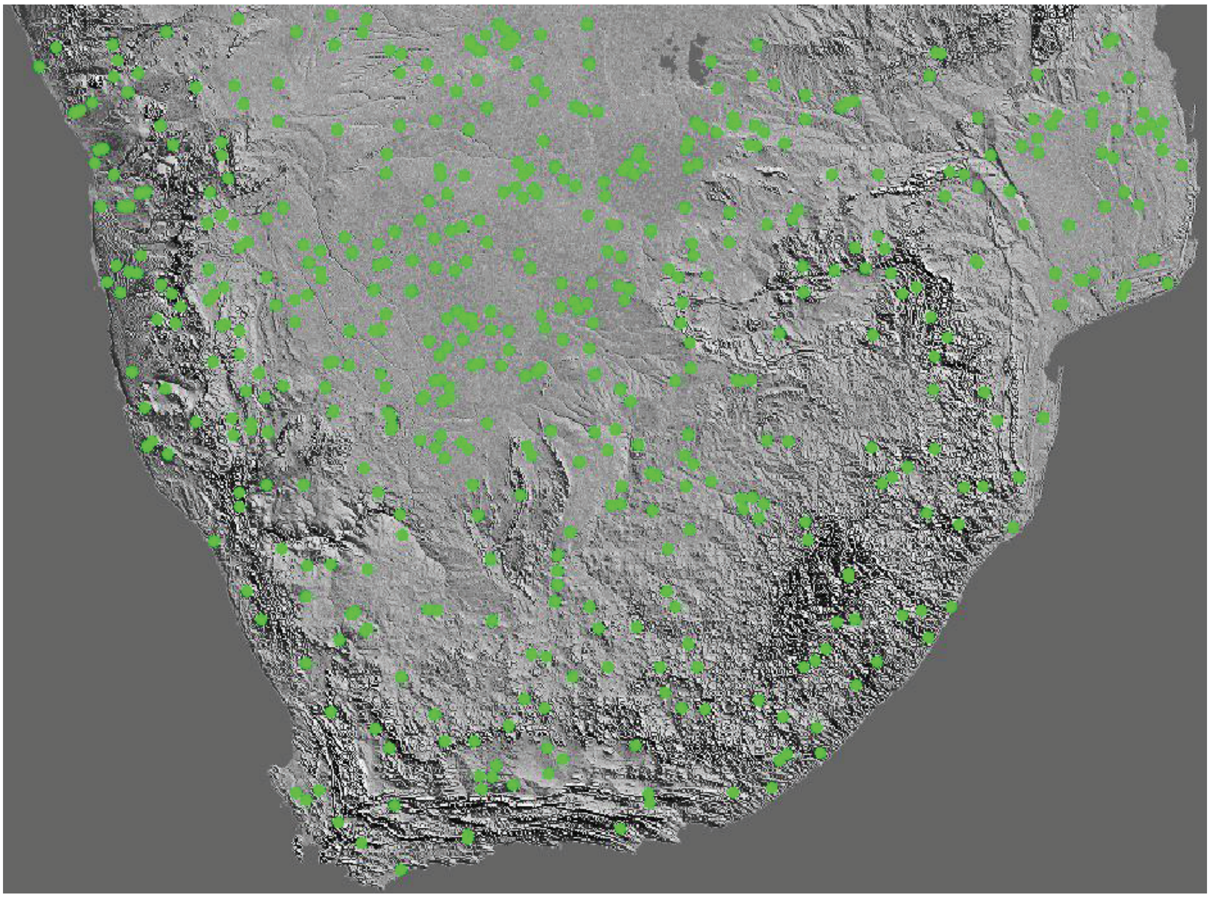


Figure 19: Random accuracy assessment points displayed over the hill-shading of the DEM

3. RESULTS AND DISCUSSION

3.1. Filling of voids (gaps)

A continuous DEM was derived by filling the SRTM dataset's voids over South Africa with interpolations from the contour lines. The Aster GDEM data was used to fill voids over the rest of the study area outside South Africa. Datasets indicating the source of each cell in the DEM are provided with the DEM. Table 1 provides a summary of the number of cells per data source.

Table 1: Summary of the number of cells per data source

Value	Source	Nr of cells	Percentage
1	SRTM	412486731	99.44%
2	Aster GDEM	2262303	0.55%
3	Contour Lines	66953	0.02%
4	Edited by hand	17	0.00%

3.2. Correction of flow paths

The burning of important rivers from the RQS data followed by the sink filling and channel carving process ensured that DEM derived flow paths followed real world rivers closely, particularly for the major meandering rivers. The gap filling from both the contour data and Aster data filled the SRTM gaps effectively and even the dune structure in the Namib Desert in Namibia was reconstructed closely to the real world situation.

The hydrological correction ensured that artificially filling of basins was avoided; however, natural depressions (e.g. pans) were sometimes filled. Some of the most significant natural depressions (mostly large pans) were restored afterwards to their original values. Depending on the type of software these will be filled again when flow accumulation is done on the hydrologically improved DEM. Some software programs allow for flow accumulation into a depression where flow will stop completely. The flow accumulations created as derived products (SFD and MFD) allowed drainage for all depressions.

3.3. Comparison of 30 m and 90 m resolution SRTM DEMs

As shown in Figure 12, there is little difference between the flow paths of the interpolated 30 m data and native 30 m data. It is important to note that although this layer improves visual display it does not necessarily improve on the detail of the 90 m DEM. However, the interpolation process did soften the large step like effect of elevation over 90 m x 90 m cells by using the elevation of the eight surrounding cells. For this reason the block-like appearance of SRTM data on e.g. a 1:50 000 scale was significantly reduced. It is possible that flow path position might also be positionally more accurate but this needs more testing over different landscapes. Over the test area this seemed to be the case based on a visual comparison between the native 30 m SRTM flow paths, the 30 m SRTM interpolated flow paths and the 90 m SRTM flow paths.

3.4. Accuracy assessment

The flow path positional accuracy assessment results (Tables 2 to 4) show as expected a very high correspondence between the reference satellite imagery and flow paths derived from the hydrologically corrected DEM. A very high accuracy was obtained regardless of when the uncertainty of flow paths in sandy areas, e.g. the Kalahari and Namib Desert, and pans and water-bodies were included or excluded from the accuracy analysis. In all the scenarios shown by Tables 2 to 4 there was an accuracy of approximately 90% or less than one SRTM cell deviation from the reference data. The average distance deviation was 26 m with a standard deviation of 42 m.

Table 2. Accuracy analysis with all 500 random points.

Distance class	Number of points per class	Total number of points	%
< 1 Cell (90 m)	459	500	91.8
1-2 Cells (90-180 m)	37	500	7.4
> 2 Cells	4	500	0.8
Total	**500**		**100**

Table 3. Accuracy when points over sandy areas have been removed.

Distance class	Number of points per class	Total number of points	%
< 1 Cell (90 m)	390	425	91.8
1-2 Cells (90-180 m)	31	425	7.3
> 2 Cells	4	425	0.9
Total	**425**		**100**

Table 4. Accuracy when points over both sandy areas and pans/water-bodies have been removed.

Distance class	Number of points per class	Total number of points	%
< 1 Cell (90 m)	274	306	89.5
1-2 Cells (90-180 m)	28	306	9.2
> 2 Cells	4	306	1.3
Total	**306**		**100**

3.5. Derived products

The following products were derived from the DEM:

- Local drain direction
- Upstream number of cells
- Slope
- Aspect
- Hillshade image
- Catchment boundaries (Primary, secondary, tertiary and quaternary).

The derived products are illustrated in the maps provided in Appendix 2.

4. CONCLUSIONS

A continuous DEM was successfully created for southern Africa (between 19ºS and 35ºS and 12ºE to 36ºE). The SRTM DEM was used as baseline dataset and voids were filled with elevations from 20 m contour lines and the Aster GDEM dataset. The gap-filled DEM was hydrologically improved and it was demonstrated that products such as catchment boundaries and river lines can be successfully derived from the hydrologically improved DEM.

Appendix 1 provides a list of all the datasets generated in this project. The datasets can be obtained from the Department of Water Affairs, Directorate: Spatial and Land Information Management. Mistakes reported will not be corrected on these datasets but will be noted to be incorporated when higher resolution data becomes available.

Significantly higher resolution RADAR DEMs will become available in the near future for large areas of the world. It is envisaged that the TerraSAR-X / TanDEM-X satellite (operational 2010) will image the entire Earth in unprecedented detail (http://www.infoterra.de/tandem-x_dem). A global DEM at 12 m horizontal resolution with a vertical accuracy of better than 2 m will become available in 2014.

5. RECOMMENDATIONS

The following potential research themes are recommended:

- The shortcomings of currently available river datasets are outlined in paragraph 2.3. The quality of the DEM produced during the project was limited by the accuracy and completeness of the datasets available. A first step could be to produce a similar dataset as the RQS 1:500 000 rivers on 1:250 000 scale.
- The hydrologically improved DEM will be, amongst others, suitable for the following derived products:
    - Generate geomorphological features, i.e. terrain morphological units (mapping of crests, midslopes, footslopes and valley bottoms), that can be used for soil mapping, erosion potential mapping and flood analysis;
    - Mapping of wetlands.
- Quality checking / evaluating DEMs for hydrological processing. To develop quality criteria for processing DEMs for hydrological applications.

REFERENCES

ASTER GDEM Validation Team (2009). Aster Global DEM Validation Summary Report. http://www.ersdac.or.jp/GDEM/E/image/ASTERGDEM_ValidationSummaryReport_Ver1.pdf

Department of Water Affairs and Forestry (2006). The construction of a hydrologically-correct, annotated 1:500 000 spatial dataset of the rivers of South Africa and contiguous basins, Report Number N/0000/00/REH/0701. DWAF: Resource Quality Services. Pretoria, South Africa.

ESRI (1996). ArcDoc *Version 7.0*, Environmental Systems Research Institute, Inc., Redlands, CA.

ESRI (2008). ArcMap 9.3. Environmental Systems Research Institute, Inc., Redlands, CA.

Farr TG and Kobrick M (2000). *Shuttle Radar Topography Mission produces a wealth of data,* EOSTrans. AGU, 81, 583-585.

Gallant JC and Wilson JP (2000). Primary topographic attributes. In: JP Wilson and JC Gallant (Eds.), Terrain Analyses – principles and applications: 51-85. Wiley, New York.

GISCOE (2001). GISCOE Digital Terrain Models. GISCOE, South Africa.

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

Lindsay JB and Creed IF (2005). Removal of artifact depressions from digital elevation models: towards a minimum impact approach. *Hydrological Processes*, 19: 3113–3126

Oksane J and Sarjakoski T (2005). Error propagation analysis of DEM-based drainage basin delineation. *International Journal of Remote Sensing*, 26 (14): 3085-3102.

ORASECOM (2009). The Orange-Senqu River Awareness Kit. Orange-Senqu River Commission, Pretoria.

Quinn P, Chevallier P and Planchon O (1991). The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, Vol.5, 59-79.

Rodrígues E, Morris CS and Belz JE (2006). A global assessment of the SRTM performance. *Photogrammatic Engineering & Remote Sensing* 72 (3): 249-260

Strahler A and Strahler A (1997) Physical Geography: Science and systems of the human environment. John Wiley and Sons, New York.

Sullivan A and Sibanda LM (2010). Project Synthesis Report. Limpopo Basin Focal Project. Project Number 62. Challenge Programme for Water and Food. Sri Lanka.

Tarboten G (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water research*, Vol. 33, No. 2. 309-319.

USGS (s.a.) SRTM Topography. Available online at: http://dds.cr.usgs.gov/srtm/version2_1/Documentation/SRTM_Topo.pdf. Last date accessed: 2 June 2010.

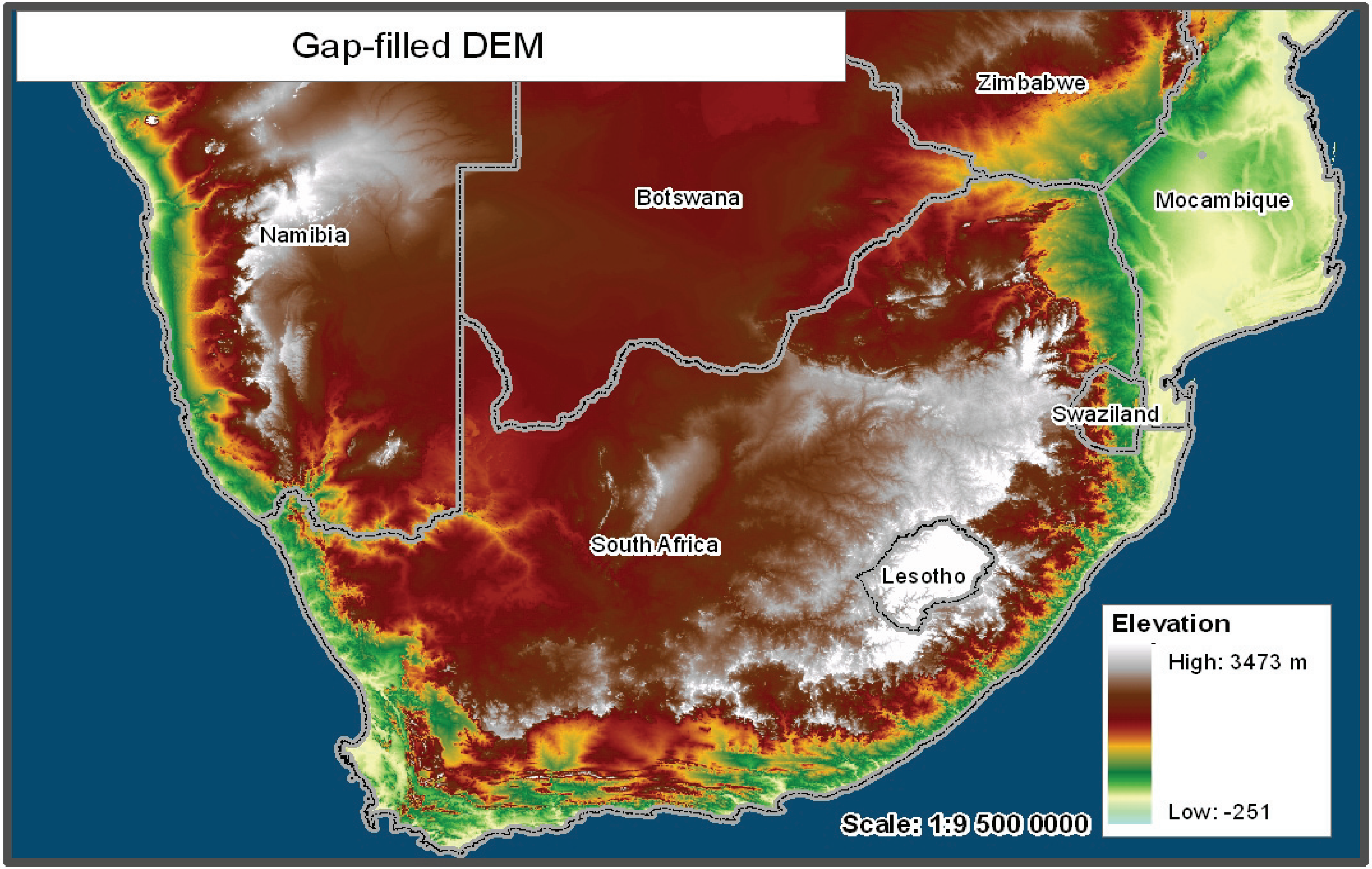
Van Niel TG, McVicar TR, Li L, Gallant JC and Yang Q. (2008). The impact of misregistration on SRTM and DEM image differences. *Remote Sensing of Environment*, 112 (5): 2430-2442.

Weydahl DJ, Sagstuen J, Dick ØB and Rønning H (2007). SRTM DEM accuracy assessment over vegetated areas in Norway. *International Journal of Remote Sensing*, 28 (16): 3513-3527.

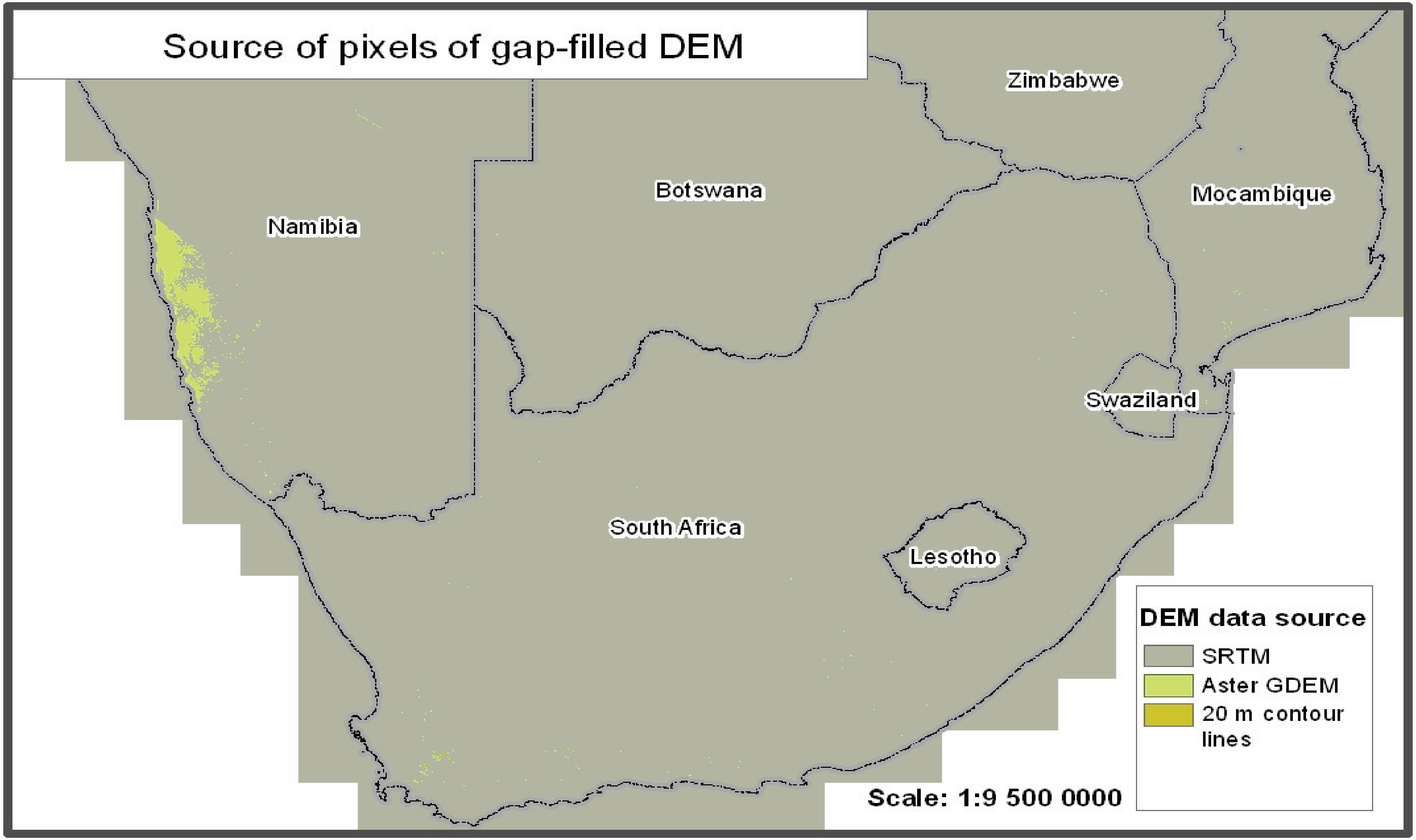
APPENDIX 1: List of products

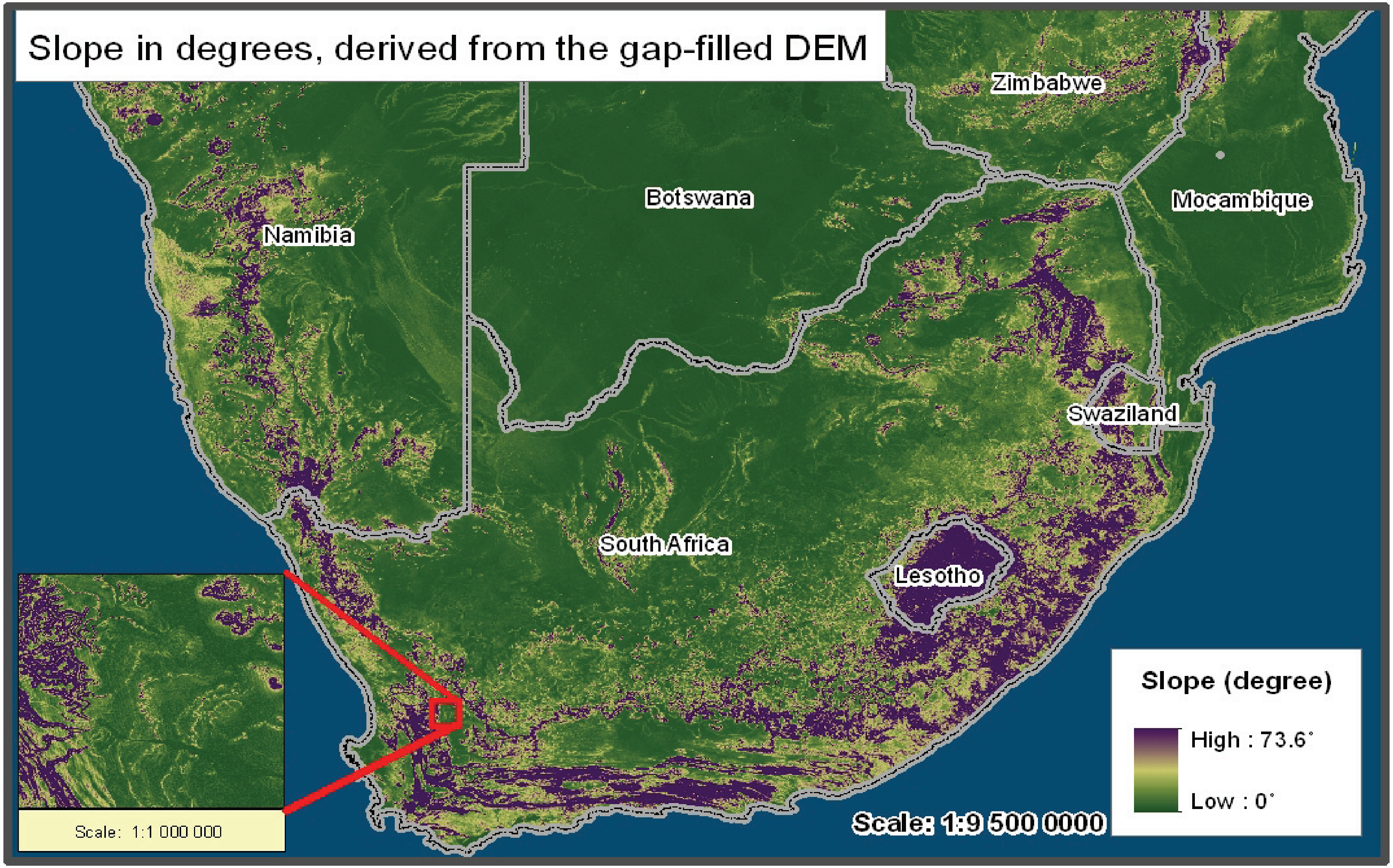
Filename	Description
SRTM90 m_Gapfilled.tif	Gap-filled DEM based on SRTM (Voids filled with contour information and Aster GDEM)
SRTM90 m_Gapfilled_Source.tif	Source of pixels of gap-filled DEM
SRTM90 m_Flowpath_improved.tif	Flow path improved DEM
SRTM90 m_Hillshade.tif	Hillshade, derived from the flow path improved DEM
SRTM90 m_Aspect.tif	Aspect, derived from the gap-filled DEM
SRTM90 m_Slope_Degree.tif	Slope in degrees, derived from the gap-filled DEM
SRTM90 m_Slope_Percent_rise.tif	Slope as percentage rise, derived from the gap-filled DEM
SRTM90 m_Flowdir_sfd.tif	Single flow direction
SRTM90 m_Flowdir_mfd.tif	Multiple flow direction
SRTM90 m_Flowacc_sfd.tif	Upstream number of cells (single flow direction)
SRTM90 m_Flowacc_mfd.tif	Upstream number of cells (multiple flow direction)
SRTM1km.tif	1km DEM derived from SRTM dataset
SRTM1km_Hillshade.tif	Hillshade of 1km DEM
SRTM90 m_flow_paths.shp	Flow paths, derived from the flow path improved DEM
SRTM90 m_Quaternary.shp	Quaternary catchment boundaries
SRTM90 m_Tertiary.shp	Tertiary catchment boundaries
SRTM90 m_Secondary.shp	Secondary catchment boundaries
SRTM90 m_Primary.shp	Primary catchment boundaries

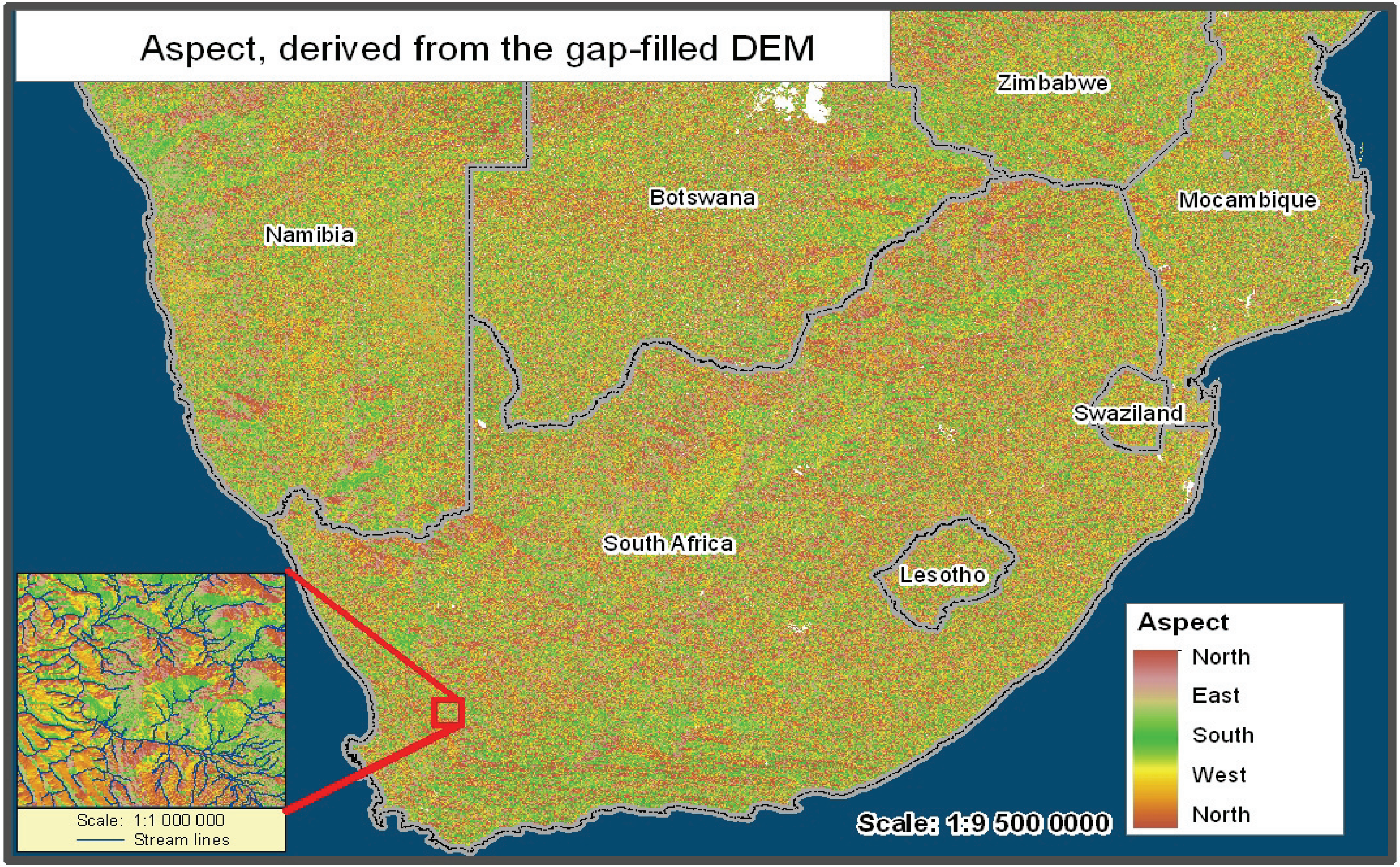
APPENDIX 2: Maps

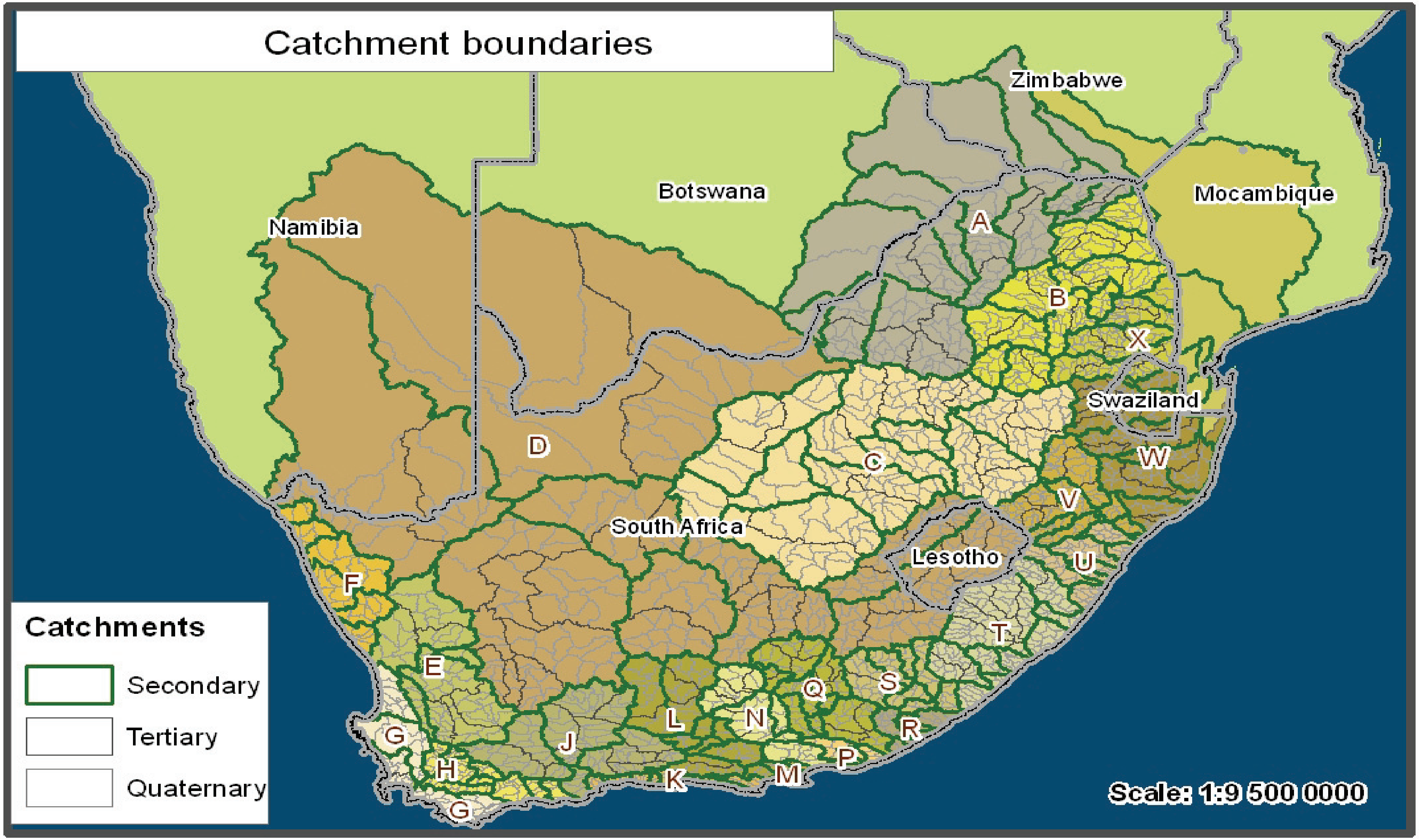


Gap-filled DEM
Zimbabwe
Botswana
Mocambique
Namibia
Swaziland
South Africa
Lesotho
Elevation
High: 3473 m
Low: -251
Scale: 1:9 500 0000

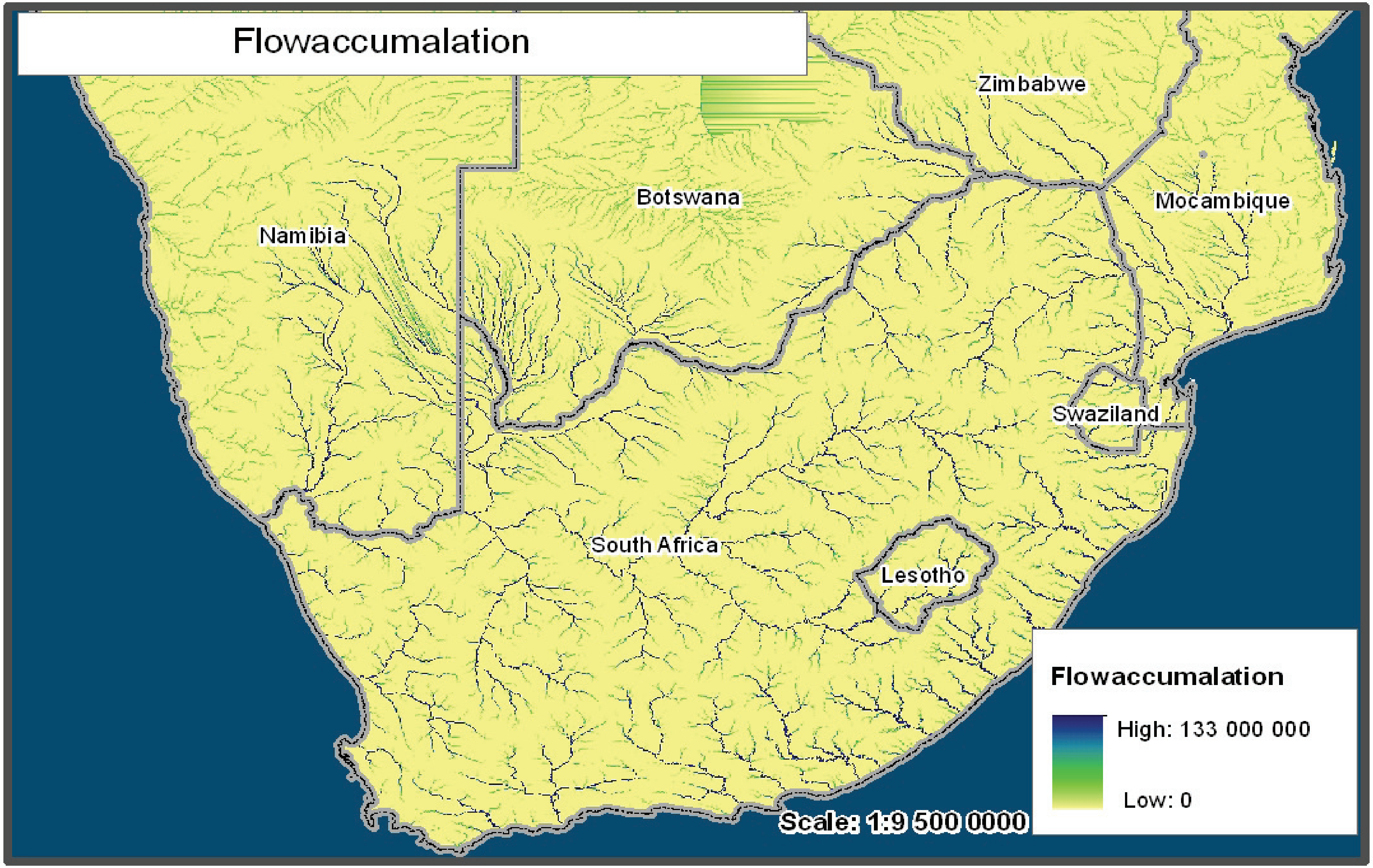

Source of pixels of gap-filled DEM
Zimbabwe
Botswana
Mocambique
Namibia
Swaziland
South Africa
Lesotho
DEM data source
SRTM
Aster GDEM
20 m contour lines
Scale: 1:9 500 0000

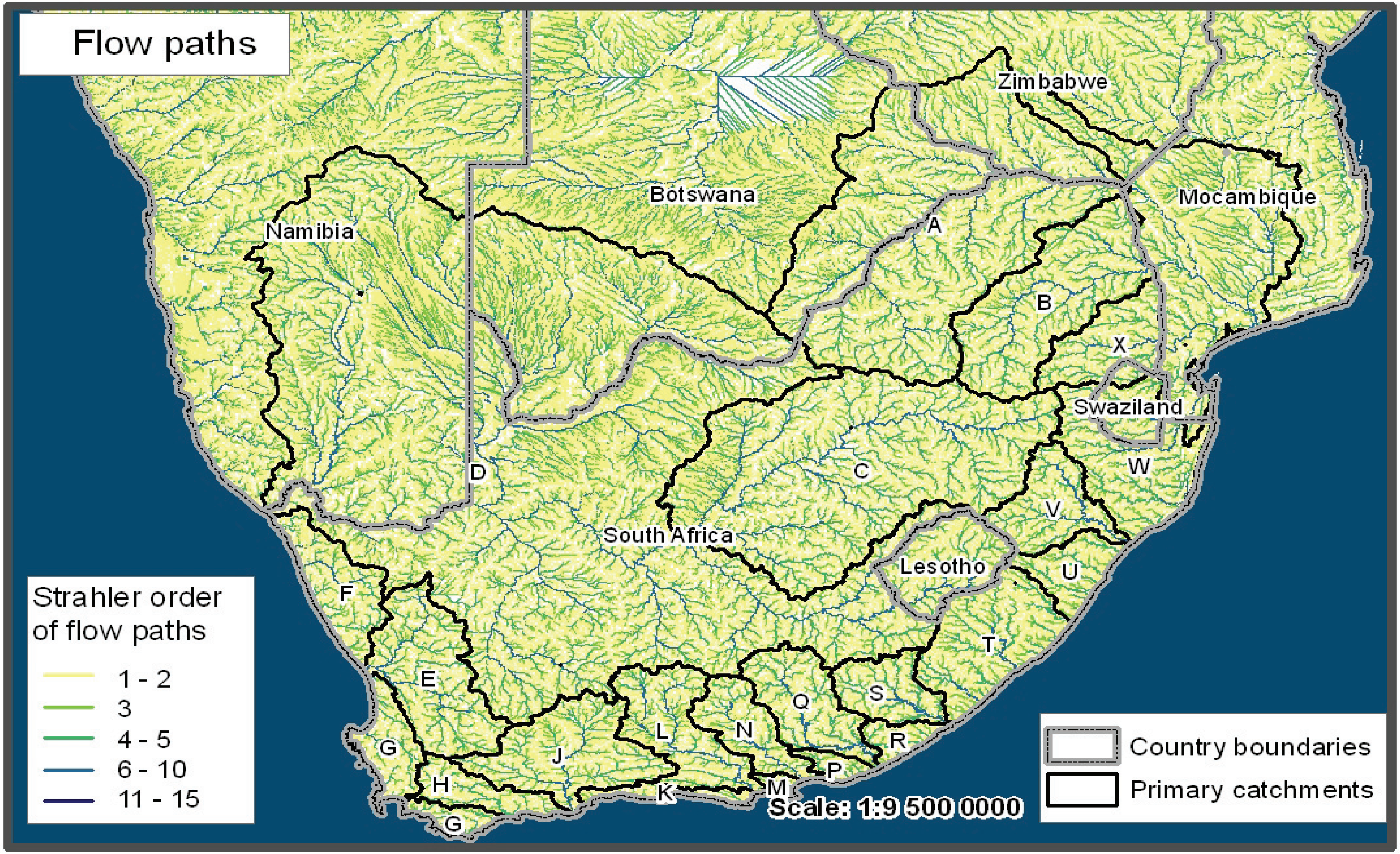

Slope in degrees, derived from the gap-filled DEM
Zimbabwe
Botswana
Mocambique
Namibia
Swaziland
South Africa
Lesotho
Slope (degree)
High : 73.6˚
Low : 0˚
Scale: 1:9 500 0000
Scale: 1:1 000 000

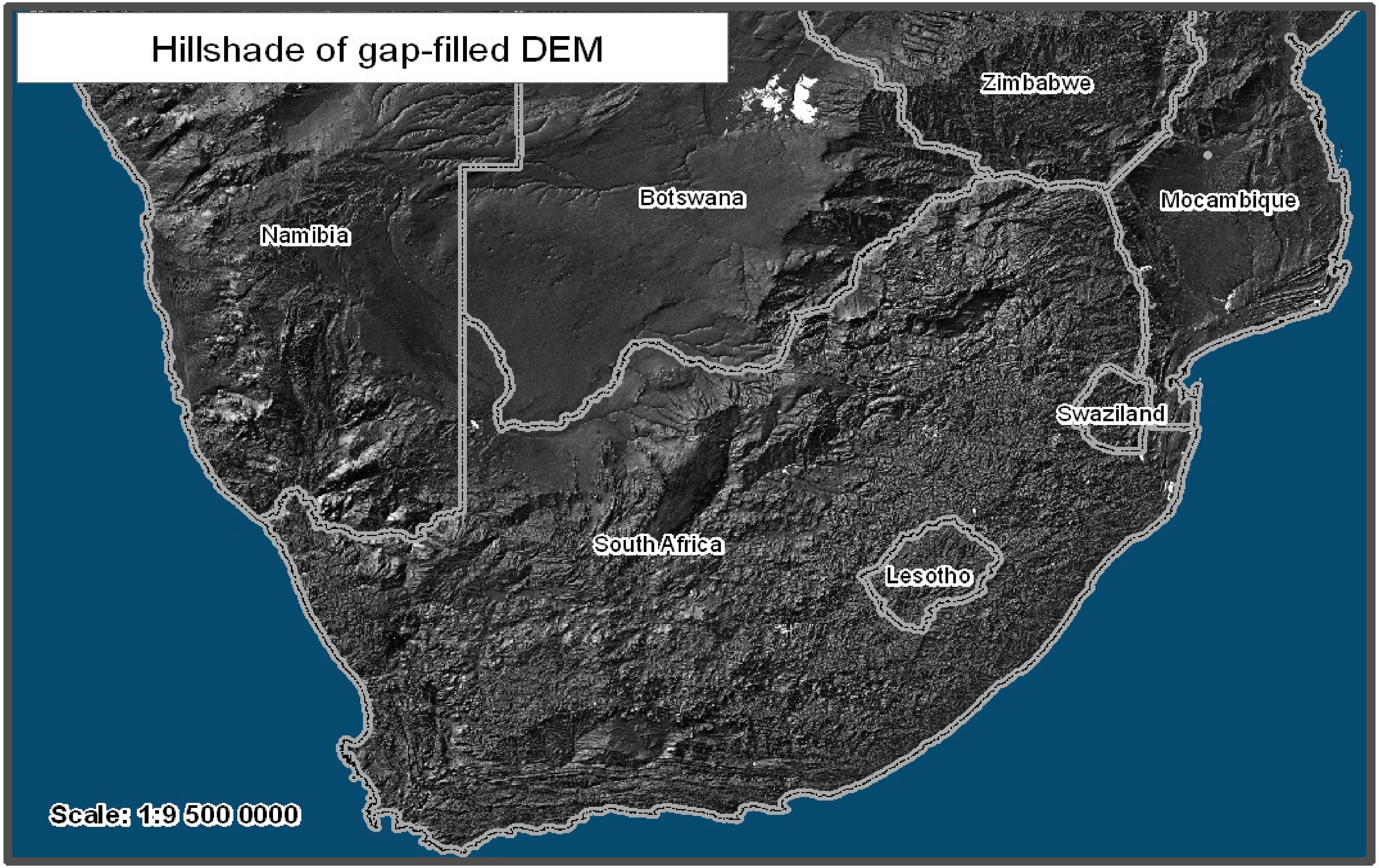

Aspect, derived from the gap-filled DEM
Zimbabwe
Botswana
Namibia
Mocambique
Swaziland
South Africa
Lesotho
Aspect
North
East
South
West
North
Scale: 1:9 500 0000
Scale: 1:1 000 000
Stream lines



Catchment boundaries
Namibia
Botswana
Zimbabwe
Mocambique
Swaziland
Lesotho
South Africa
A
B
C
D
E
F
G
G
H
J
K
L
M
N
P
Q
R
S
T
U
V
W
X
Catchments
Secondary
Tertiary
Quaternary
Scale: 1:9 500 0000


Flowaccumalation
Zimbabwe
Botswana
Mocambique
Namibia
Swaziland
South Africa
Lesotho
Flowaccumalation
High: 133 000 000
Low: 0
Scale: 1:9 500 0000


Flow paths
Zimbabwe
Botswana
Mocambique
Namibia
A
B
X
Swaziland
C
D
W
V
South Africa
Lesotho
U
F
T
E
S
Q
N
L
R
G
J
P
H
M
K
G
Strahler order
of flow paths
1 - 2
3
4 - 5
6 - 10
11 - 15
Scale: 1:9 500 0000
Country boundaries
Primary catchments


Hillshade of gap-filled DEM
Zimbabwe
Botswana
Mocambique
Namibia
Swaziland
South Africa
Lesotho
Scale: 1:9 500 0000