# DIGITAL ELEVATION MODEL DATABASE W42 - A SCALABLE SYSTEM FOR SPATIAL DATA

M. Habermeyer, U. Marschalk, A. Roth

German Aerospace Center (DLR)，German Remote Sensing Data Center (DFD)
82234 Wessling, Germany，(martin.habermeyer,ursula.marschalk,achim.roth)@dlr.de

**Commission I, SS-1**

**KEY WORDS:** DEM/DTM, Mapping, Mosaic, Software, System, Spatial Databases

**ABSTRACT:**

Digital Elevation Models (DEMs) provide elementary information for various application fields. From a global point of view, there are many different sources to furnish height information of the Earth's surface. These sources can be heterogeneous in coverage, as well as in resolution, information content and quality. A database aiming at the provision of world-wide DEMs has to consider these differences in the design of the system with respect to the structure and the algorithms.The DEM database W42 is a scalable spatial database system capable of holding, extracting, mosaicing, and fusing spatial data. Design aspects for this task can be specified as holding spatial data in a unique data structure and providing unique access functions to the data. These are subject of this paper.The database's structural design is presented in a Four-Layer-Concept. Each layer provides a clearly specified interface to subsequent/preceding layers, as well as each layer is scalable in nature to meet future requirements. The scalability of the system enables the integration of the database into operational processors: at the moment it is serving as part of the geocoding-processor of the TerraSAR-X mission and as part of the generic processing chain for airborne hyperspectral data. It further will be utilized in satellite missions as an element of the Payload Ground Segment in the future. The extensions made for TanDEM-X are explained in this context.

## 1. INTRODUCTION

A design of a scalable system for holding spatial data in general and Digital Elevation Models (DEMs) in specific, has to account for some basic conditions resulting from the different sources DEMs are produced from. The sources can be heterogeneous in coverage, as well as in resolution, information content and quality. A database aiming at the provision of world-wide DEMs has to consider these differences in the design with respect to the structure and the algorithms.

The W42-DEM-Database is a stand-alone database system capable of holding spatial data in a unique system architecture. It provides functions allowing import and export of spatial data in raster and vector-like structures. The purpose of this document is the clarification of the database design and the description of the provided functionality.

This work extends the papers and reports previously published about the database see e.g. (Roth et al. 2002), (DLR, 2006). The database was part of the geocoding system for radar missions: SRTM/X-SAR (Roth et al., 1999) and TerraSAR-X (Huber et al., 2006). It is also used in an operational processing chain for optical airborne data (Habermeyer et al., 2003). The focus in this paper is on design aspects with regard to the TanDEM-X mission, for which the database has been adopted to be a part of the Mosaicking and Calibration Processor (Wessel et al., 2008) and to support accompanying studies.

The document is structured as follows:
Section 2 explains the System Design in a Four-Layer-Concept, comprising Physical Layer, File System Layer, Database Access Layer and File Access Layer.

Three different kinds of data (Data Tiles, Image Chips, and Tracks) can be stored in the database system at the moment. These are specified in section 3.

The data types are abstracted to data formats (section 4).

For each data type (Data Tile, Image Chip, track) the organisation of the file sytem is explained (section 5).

The interaction of the database access routines with the file system is subject of discussion in section 6: the needed routines for import to, export from, update of, inventory of and maintenance of the database are specified and their impact on database components are discussed.

The utilities for the manipulation of the database content are based on programs providing functionalities to handle raster- and vector data. These are explained in section 7, without a consideration of an organisational structure, only working on a file level without spatial semantics.

The functions of the File Access Layer cover file format conversion, file manipulation, information about the data content, visualization of files, fusion of information provided by different layers and different formats, mosaicking them together, and resampling routines to change their resolution.

## 2. SYSTEM DESIGN

The database's structure is organized in four layers (fig. 1). The system is based on the "Physical Layer" represented by media mounted on workstations or PCs. The second layer, which is named "File System Layer" spans the directory tree over the

"Physical Layer". It allows unique access to the data and is organized in projects. Inside of each project the geographic coordinate system allows the separation into smaller categories. The access to the "File System Layer" is carried out by the third layer, the "Database Access Layer". This layer provides unique functions to import, export, update, backup and remove data from and to the database. It further allows the display of the information content of the database. These functions are realized by the operating system and by modules of the fourth layer, the "File Access Layer".
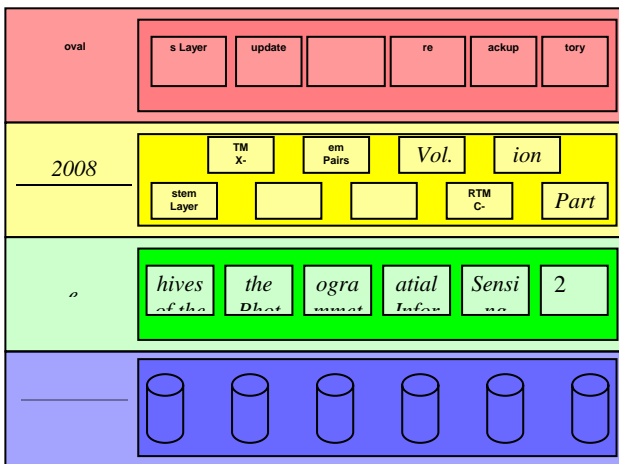


Figure 1. System Design

The system has to be scalable to allow an easy expansion of all four layers. This can be assured as we show for each layer. In case of the "Physical Layer" the system can be expanded by the integration of further media to adjust the available memory capacity to more projects and/or finer resolutions. The "File System Layer" can be adopted to hold further projects and higher resolutions by the introduction of new directories serving as containers to hold the expanded data sources. The "Database Access Layer" can account for the increase of resolution and the availability by the provision of more sophisticated functions. The "File Access Layer" is completely open for an extension of the data available and the enhancement of the available resolutions.

The scalability of the system enables the integration of the database into operational processors: at the moment it is serving as part of the geocoding-processor of the TerraSAR-X mission and as part of the generic processing chain for airborne hyperspectral data. It further will be utilized in satellite missions as an element of the Payload Ground Segment in the future. Currently the database is adopted for the use in the ground segment of the TanDEM-X mission.

As the "File Access Layer" provides the functionality for the manipulation of the files in the database, it is discussed here. There are seven classes in the "File Access Layer" providing the following basic functionalities:

- file format conversion routines for the support of different in- and output formats to keep the system open for future duties, which can go beyond a database system for DEMs,
- modules to display the information content of the data to have a quick tool allowing rudimentary quality control,

- tools to allow the manipulation of file content for the adoption of the data to other tasks,
- visualization utilities to support a quick overview of data capabilities,
- methods to resample the data to different resolutions or coordinate systems to open the system's data to a wider range of applications,
- utilities to allow the mosaicking of data: these comprise the combination of data from different sources as well as from different areas,
- utilities to support the fusion of data, where more than one elevation model for an area is available. Resolution and quality of the candidates have to be taken into account and carefully combined to arrange a product that combines the information content of different layers.

These utilities of the "File Access Layer" form a base of modules that can be combined to new modules. The functionality of the "Database Access Layer" is provided by the combination of functions from the "File Access Layer" and the operating system. These elements are composed to new modules.

## 3. DATA TYPES

Currently three data types are supported: Data Tiles, Image Chips, and Tracks.
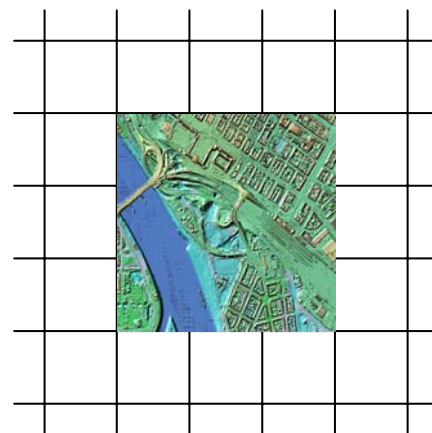
### 3.1 Data Tiles



Figure 2. Data Tiles

Data Tiles contain spatial data representing areal phenomena. Their borders are congruent with the grid spanned by the geographic coordinate. Their resolution of the pixels is one of 30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01 arc-seconds. They are stored in Sun Raster Format with DLR extension (see section 4) and organized as specified in section 5.1.

### 3.2 Image Chips

Image Chips contain spatial data representing areal phenomena. Their borders can be incongruent with the grid spanned by the geographic coordinate system (see fig. 3). The resolution of the chips can range from 30 to 0.01 arc-seconds. Image chips are stored in Sun Raster Format with DLR extension (see section 4) and organized as specified in section 5.2.
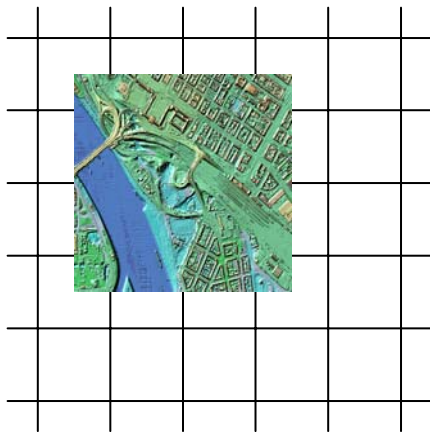
Figure 3. Image Chips

### 3.3 Tracks

Tracks contain spatial data representing line phenomena (see fig. 4).
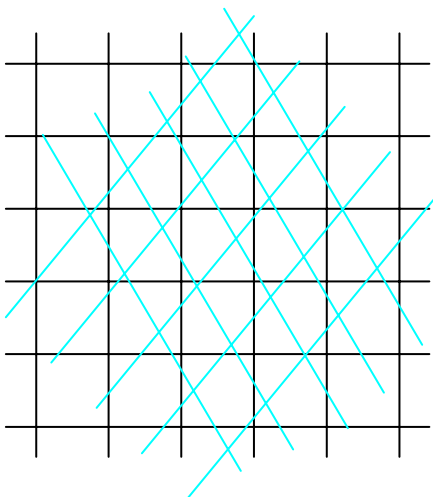


Figure 4. Tracks

## 4. DATA FORMAT SPECIFICATION

For the representation of Image Chips and Data Tiles the Sun Raster File with DLR extension (SRF) (DLR, 2002) is foreseen. The files are stored in the database in big-endian byteorder format strictly to avoid confusion. For the purpose of enabling fast processing the files are kept in machine-specific byteorder after the export and prior to the import to the database. The DLR footer specifies statistics and the annotation, allowing a fast display of statistical information and quick addressing for Data Tiles and Image Chips.

The vector and point data is represented in ESRI's (ESRI, 1998) shapefile format.

## 5. ORGANISATION OF THE FILE SYSTEM

The database structure is based on a UNIX file system. The different sources of data stored are accounted for by separate branches (called projects in the following) in the directory tree. Below the project level the grid spanned by the geographic coordinate system (fig. 5) is projected onto the directory tree of the file system in the order North (two digits ranging from 00-89) to South (two digits ranging from 01-90) and East (three digits ranging from 000-179) to West (three digits ranging from 001-180) in 1° steps.



Figure 5. Geographic coordinate system

For resolutions finer than 1 arc-second the directory tree is further subdivided in 0.01 arc-second steps in latitudes and longitudes as well (in both directions represented by two digits ranging from 00-99). At each level a logfile is kept to keep track of changes made to each subdirectory (fig. 6).
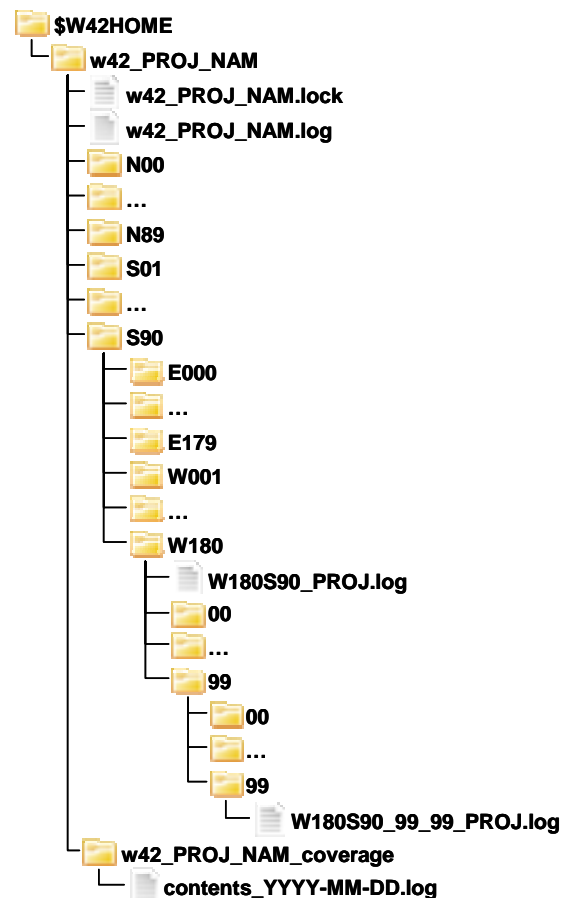


Figure 6. Directory Structure of the file system

The different data formats (Data Tiles, Image chips, and Tracks) sort into this file system in specific manners.

## 5.1 Data Tiles

DEMs are provided in 1° x 1° degree tiles in geographic coordinates in Sun Raster Format. The various DEM sources plus additional data are organized as directory-trees from North to South and East to West in 1 degree steps. For resolutions below 1 arc-second the directory tree is further split up into 0.01 degree tiles. For performance reasons each tile is held in lower resolutions as well. This enables an incorporation of data with fine resolution (e.g. height models from laser scanners) as well as quick access to data only available in relatively coarse resolution (e.g. GLOBE data).

Data Tiles are provided in exact 1° x 1°-tiles (geographic coordinates) in Sun Raster Format with DLR extension. The data is zipped prior to storage. As mentioned before, the various DEM sources plus additional data are organized in directory-trees from North to South and East to West in 1° steps, and for resolutions below 1° in 0.01° steps (fig. 6). The finest resolution supported is 0.01 arc-seconds. For performance reasons each tile of a fine resolution is held in coarser resolutions as well. For an overview of the resolution of the stored data tiles check table 1.

Each tile contains an overlap of 30 arc-seconds to the South and the West to the adjacent tiles at the 1° x 1°-directory level and an 0.3-arc-second-overlap at the 0.01° x 0.01°-directory level.
For the upper left and the lower right coordinate the pixel centers are annotated. As lower right coordinate the first coordinate outside the image is specified. For changing resolutions the tile border is kept, which means that the coordinates change in the different layers as well.

| Resolution (per pixel, in arc-seconds) | File Size (in pixels) | Directory Level |
|---|---|---|
| 30 | 121 x 121 | 1° x 1° |
| 10 | 363 x 363 | 1° x 1° |
| 3 | 1210 x 1210 | 1° x 1° |
| 1 | 3630 x 3630 | 1° x 1° |
| 0.3 | 121 x 121 | 0.01° x 0.01° |
| 0.1 | 363 x 363 | 0.01° x 0.01° |
| 0.03 | 1210 x 1210 | 0.01° x 0.01° |
| 0.01 | 3630 x 3630 | 0.01° x 0.01° |

Table 1. Resolutions of data tiles provided in the database

At the moment the following different elevation data with a global coverage are stored as Data Tiles in the database:
- the GLOBE dataset (Hastings and Dunbar, 1998) with a global coverage and a resolution of 1km,
- data from the Shuttle-Radar-Topography-Mission (SRTM) (Pessagno, 2000). SRTM mapped 58 million km² with DLR's X-band system and 119 km² of the land surface with NASA's C-band system within the 56° Southern and 60° Northern latitudes,
- data resulting from repeat-pass-interferometric DEM production with the ERS-1 and ERS-2 satellites,
- a reference dataset used in the ground segment of the TerraSAR-X mission, composed of SRTM and GLOBE data.

As the nature of the data has an impact on the requirements for the algorithms and consequently the design of an elevation model database the characteristics of the DEMs provided by the different sensors/missions is summarized in table 2. The accuracies to be provided by TanDEM-X are also listed.

| Source/Format | Accuracy | | Resolution (in arc-seconds) |
|---|---|---|---|
| | relative | absolute | |
| DTED-1 | 20 m | 30 m | 1 |
| DTED-2 | 20 m | 30 m | 1 |
| ERS | 20 m | 30 m | 1 |
| SRTM/C-Band | 8 m | 16 m | 3 |
| SRTM/C-Band | 8 m | 16 m | 1 |
| SRTM/X-Band | 6 m | 16 m | 1 |
| TanDEM-X | 2 m | 10 m | 0.4 |

Table 2. Resolutions and accuracies of different sources held in the database

## 5.2 Image Chips

Image chips are stored in the given directory structure (fig. 6). They are referenced by their center coordinate. Different from the data tiles they are only stored in their original resolution and not in the resolutions specified by table 1. Their resolution needn't fulfil any specification except for being specified in degrees or seconds.
The directory structure has to be expanded by two components:
- an identifier file ensuring a unique addressing of the image chips,
- and a file containing a list where all image chips are tracked that are completely or in parts situated in the 1°x1°- tile to guarantee a quick search for tiles.

## 5.3 Tracks

Tracks are also stored in the directory structure shown in figure 6 and discussed in section 5.2. This structure is expanded by:
- a file holding a unique number for assignment of new tracks during the import process,
- a list of tracks situated in the specific tile, containing the unique identifier, the file name plus attributes that can be searched for,
- a zipped file for each track.

Tracks are not further referenced and are identified by their unique identifier, which is part of the filename.

## 6. DATABASE ACCESS ROUTINES

For each datatype (tiles, chips, and tracks) unique routines have to be foreseen allowing the import, export, inventory and maintenance of the data.

## 6.1 Import

For all data types exclusive write access is guaranteed by lock files. Log files are kept for a later reproduction of changes to the

database. The inventory is created or updated for a quick overview of the database's content.

Data Tiles have to be resampled to lower resolutions to provide a fast access to the needed accuracy later. In case of the Image Chips the chip list is updated, no resampling of the data is done. Tracks are partitioned to fit in the tiles where they are stored, and the track list is updated.

For each data type the components are zipped to save memory space.

### 6.2  Export

#### 6.2.1  Data Tiles

The export routine for data tiles comprises the following steps:
- get DEM tiles from the addressed project(s),
- mosaick DEM tiles to one file per project,
- cut out a specified part for each project,
- fuse information from different layers to one file,
- apply coordinate transformation to target projection/ coordinate system.

#### 6.2.2  Image Chips

The export routine for image chips accomplishes the following function: "for a given region of the earth's surface (specified by the corner coordinates) provide me with all image chips intersecting this region".

#### 6.2.3  Tracks

The export routine for tracks accomplishes the following two functions:
- "for a given region of the earth's surface (specified by the corner coordinates) provide me with all points stored in tracks within this region" , or
- "for a given track provide me with all points situated within this region".

Both functions are assured by the help of the track list stored in each tile.

### 6.3  Update

Update routines for all data types have the following steps in common: creation and deleting of lock files providing exclusive file access, and the update of log files and inventory.

Data tiles can be updated at any level. The coarser resolutions in this case have to be updated as well. The update of an image chip has to be interpreted as a replacement of an existing chip by a new chip. The identification of the chip to be updated is done by its unique identifier. In this case the tile specified with its unique identifier is replaced.

For updating tracks insertion and deletion of points into or from a track have to be foreseen.

### 6.4  Inventory

For each file type an inventory function is foreseen that is capable of providing an overview of the data stored within (parts of) the project.

### 6.5  Maintenance

The maintenance module for each data type should be capable of providing the following functionality: undoing the last creation/update step, removing files from archives and inventory lists, erasing complete directory trees, backing up

(parts of) the directory tree, copying (parts of) the directory tree. All of the above functions can only be simulated. Only listings are prepared in this case.

## 7.  FILE ACCESS ROUTINES

The functionalities provided for the handling of raster and vector data are displayed in table 3. This comprises seven classes of basic functions:
- file format conversion routines for the support of different in- and output formats,
- routines to allow the manipulation of the file content ,
- modules to display the information content of the data,
- visualization utilities, e.g. for the generation of color shaded presentations,
- algorithms to fuse overlaying tiles in different resolution and qualities,
- utilities to allow the mosaicking of data,
- methods to resample the data to different resolutions or coordinate systems.

|  | Raster Data | Raster and Vector Data | Vector Data |
|---|---|---|---|
| Data Conversion | x |  | x |
| Data Manipulation | x |  | x |
| Data Information | x |  | x |
| Data Visualisation | x |  | x |
| Information Fusion | x | x | x |
| Mosaicking | x |  | x |
| Resampling | x | x | x |

Table 3. File Access Functions

### 7.1  Data Conversion

Conversion routines between common generic image processing formats with a geographic representation have to be foreseen. At the moment the exchange between files in SRF, GeoTiff, and DTED format is supported.

Conversion routines from sensor specific formats (e.g. IceSat data) to shapefiles has to be implemented for Tracks.

### 7.2  Data Manipulation

Modules are implemented to
- allow various file manipulation actions for SRFs,
- add or subtract the geoid undulation to or from a given SRF,
- manipulate SRF header and/or footer content,
- create simple synthetic patterns in SRFs,
- rotate a given SRF clockwise by a given angle,
- mirror given SRF at the horizontal and/or vertical axes,
- allow the separation of shapefiles according to spatial or attributive criteria.

### 7.3  Data Information

At the moment the following functionalities are implemented to show the information contained in the elevation data:
- creation of difference images of given SRFs,
- calculation and display of statistical values,

- display of SRF header and footer information, and of data values around a specified point,
- creation of histograms and statistics for SRFs as well as for shapefiles.

### 7.4 Data Visualization

The data can also be visualized. The following functionalities lie at hand:
- creation of color shaded DEM models from given SRFs,
- creation of DEM models from given SRFs,
- visualization of information contained in shapefiles.

### 7.5 Information Fusion and Mosaicking

Information fusion in our case means the combination of evidences contained in different layers for the same area, whereas mosaicking is understood as the process of producing a dataset for a wider area from sources not intersecting each other. Information fusion has to be differentiated into fusion between raster and vector data and fusion of information contained in vector data, and fusion between raster file layers. The latter is implemened as documented in (Knöpfle et al., 1998).

### 7.6 Resampling

There are functions for the calculation of cartographic transformations for Sun Raster Files. They have to be adopted to work on shapefiles as well. These functions not only invoke a resampling process but also lead to a re-projection of the values onto a different coordinate system.

## 8. CONCLUSIONS

The design of a scalable database for Digital Elevation Models has been shown. The concept for efficiently holding and accessing the data has been discussed. The data types supported at the moment have been presented. An overview of the datasets currently held in the database has been given. With this background algorithms that already have been implemented or which have to be implemented have been discussed.

The performance of the algorithms in preparation and support of the TanDEM-X Mosaicking and Calibration Processor has to be shown.

## REFERENCES

DLR, 2002. File Formats used by DFD-Geocoding Software EGEO and GEMOS, Oberpfaffenhofen, Germany.

DLR, 2006. TerraSAR-X TMSP W42 DEM Database Interface Document, Oberpfaffenhofen, Germany.

ESRI, 1998. ESRI Shapefile Technical Description, Redlands, California, USA.

Habermeyer, M., Müller, A., Holzwarth, S., Richter, R., Müller, R., Seitz, K.-H., Seifert, P. and Strobl, P., 2003. Developing a fully automatic processing chain fort he upcoming hyperspectral scanner ARES. In: M. Habermeyer, A. Müller, S. Holzwarth (eds), Proceedings of the 3rd EARSeL Workshop on Imaging Spectroscopy, Herrsching, Germany, pp. 82-87.

Hastings, D. and Dunbar, P., 1998. Development and assessment of the global one-km base evelevation digital elevation model (GLOBE). International Archives of Photogrammetry and Remote Sensing 32(4), pp. 218-221

Huber, M., Wessel, B. and Roth, A., 2006. The TerraSAR-X orthorectification service and its benefits for land use applications. In: Proceedings of IGARSS '06.

Knöpfle, W., Strunz, G. and Roth, A., 1998. Mosaicking of digital elevation models derived from SAR interferometry. International Archives of Photogrammetry and Remote Sensing 32(4), pp. 306-313.

Pessagno, C., 2000. SRTM status report. In: I. Dowman (ed.), Validation of DEMs from satellite data. Reports and proceedings of 7th Meeting of Mapping Sub Group, CEOS WG Calibration and Validation, Gaithersburg, USA.

Roth, A., Knöpfle, W. Rabus, B., Gebhardt, S. and Scales, D., 1999. GEMOS – a system for the geocoding and mosaicking of interferometric digital elevation models. In: Proceedings of IGARSS '99. Hamburg, Germany.

Roth, A., Knöpfle, W., Strunz, G., Lehner, M. and Reinartz, P., 2002. Towards a global elevation product: combination of multi-source digital elevation models. In: Symposium on Geospatial Theory, Processing and Applications, Ottawa, Canada.

Wessel, B., Marschalk, U., Gruber, A., Huber, M., Hahmann, T., Roth, A. and Habermeyer, M., 2008. Design of the DEM mosaicking and calibration processor for TanDEM-X. In: Proceedings of the 7th European Conference on Synthetic Aperture Radar, 2-5 June 2008, Friedrichshafen, Germany.