The EuroSDR project "Automated Checking and Improving of Digial Terrain Models"

Höhle, Joachim

Published in:
ASPRS 2007 Annual Conference

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Höhle, J. (2007). The EuroSDR project "Automated Checking and Improving of Digial Terrain Models" In *ASPRS 2007 Annual Conference: Identifying Geospatial Solutions* American Society of Photogrammetry and Remote Sensing.

THE EUROSDR PROJECT ”AUTOMATED CHECKING AND IMPROVING OF DIGITAL TERRAIN MODELS”

Dr. Joachim Höhle
Aalborg University
9220 Aalborg, Denmark
jh@land.aau.dk

ABSTRACT

The results of the research project “Checking and Improving of Digital Terrain Models” of the European Organization for Spatial Data Research (EuroSDR) are presented. The various solutions to the problem of determining the geometric quality of Digital Terrain Models are explained. Test material consisting of three different Digital Terrain Models (derived by automated photogrammetry, by digitizing of contours and by airborne laserscanning) and stereopairs have been used by several research groups in order to check and improve the provided Digital Terrain Models. Accurate reference data (derived by means of low altitude imagery and by ground surveying) have been used by the pilot centre in order to evaluate the results of the participants in the test.

INTRODUCTION

Digital Terrain Models (DTMs) are required for the production of orthoimages. Orthoimages are produced for large areas and at short intervals of time. The geometric accuracy of the DTM influences the quality of the orthoimage very much. The production process requires quality control, and the DTM has therefore to be checked and updated. Changes in the landscape and construction work make an update of the DTM necessary. Efficient methods for checking are requested by the producers and the users of DTMs. DTMs have nowadays higher demands in accuracy. Improvements of existing DTMs, which were derived by older methods of acquisition, are demanded. Checking **and** improvement of existing DTMs is required for large areas and with a high degree of completeness and speed. Many different types of DTMs exist. Accurate and dense DTMs are today often derived by laser scanning. Many new applications, for example city models, simulation of flooding, mapping tasks and visualizing by means of oblique images (‘pictometry’), are then possible.

In 2004 the European Research Organization for Spatial Data Research (EuroSDR) initiated research on checking of Digital Terrain Models. Participants in the project should develop methods, apply them to the same test material, and the results should be investigated regarding accuracy, completeness, degree of automation, and economy. The analysis of the results should give an indication which method would achieve usable results and which problems will remain unsolved.

METHODS FOR THE GENERATION OF DTMs

The production of orthoimages for large areas and of medium geometric resolution (0.25-0.40 m) uses heights which represent the terrain. Digital Terrain Models are used for this application. The so-called “true orthoimages” require Digital Surface Models (DSMs). They are produced for cities and with a higher geometric resolution, for example 0.05 m - 0.10 m. But DTMs are also used for this task; the displacements of houses are then tolerated. Both types of height models are sometimes called Digital Elevation Models (DEMs). The use of Digital Terrain Models is not restricted to the production of orthoimages. Other tasks such as planning of construction work or flood studies require higher accuracies than the production of orthoimages. Therefore, a demand to check and to improve the accuracy of existing DTMs exists at National Mapping Agencies and other mapping organizations.

Digital Terrain Models can be produced by different methods and each method has its characteristics. In the EuroSDR test the results of three different acquisition methods were investigated: Scanning of contour lines, automated photogrammetry and laser scanning. By means of these three methods, height spots arranged in a grid of equal spacing are produced and can in this form best be used for the production of orthoimages. The heights derived from the contour lines represent the terrain, the heights from automated photogrammetry and laser scanning are on the surface (on top of vegetation, buildings, vehicles, etc.). These raw data can be filtered and thereby be reduced to the terrain (bare earth). The accuracy of the methods for DTM generation are different; contour lines of existing topographic maps, such as 1:25 000, may have standard deviations of σ_h = 1.5–3 m; laser scanning may produce accuracies of σ_h = 0.15-0.3 m. The accuracy of automated photogrammetry is about σ_h = 0.02–0.03% of the flying height (h), for example σ_h=0.8 – 1.2 m at h=4000 m. The mentioned accuracies refer to the accuracies of the grid posts. The spacing of the height points is 1-2 m at laser scanning, but 10-25 m for the DTMs derived from the contour lines. Automated photogrammetry can also produce dense DTMs. All of the three methods of DTM generation may have blunders. They have to be detected and eliminated because they are easily visible in the orthoimages.

PRINCIPLES IN THE DTM CHECKING

The EuroSDR project include three parts in the checking: Accuracy, completeness and the degree of automation. The amount of work involved when checking large areas will require automated methods. The methods should also include the possibility of improvement of the DTMs.

Accuracy of DTMs

The accuracy of grid-DTMs includes the vertical and the horizontal accuracy of the grid posts. If the terrain is hilly or mountainous, a positional error will also result in a height error. Reference values of the check points must have a superior accuracy, which means their accuracy has to be better at least by a factor of 3 to 5. The number of reference values should be as high as possible; the distribution of the check points must be over the whole area. The principle in the checking of the DTMs in the EuroSDR project is depicted in Figure 1. The DTM accuracy is characterized by the Root Mean Square Error (RMSE value), the maximum error, the mean error, the standard deviation, and the number of blunders. The threshold for blunders is defined as 3 * RMSE, and all blunders should be removed before the calculation of the mean and the standard deviation. This threshold is special for each type of landscape and for each checking method. A constant threshold should be used when comparing the results of the different methods. It will be set according to the checking method (for example 0.00015* flying height of the image data available to the participants). Table 1 shows the accuracy measures as they are used in the EuroSDR project.

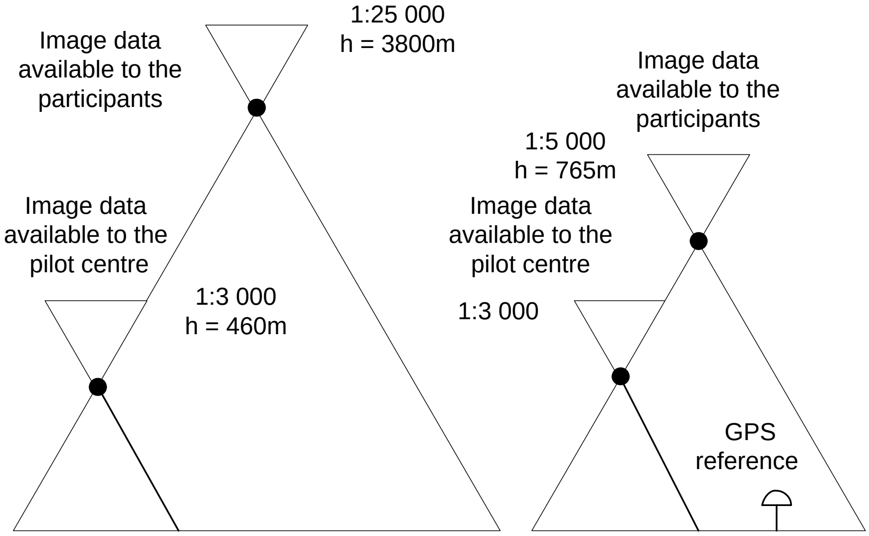

Figure 1. Principle of checking DTMs in the EuroSDR project. The left configuration is used for DTMs derived by digital photogrammetry and by scanning of contour lines, the right one for DTMs derived by laser scanning.

Special attention has to be given to the mean error because a systematic shift of the heights may lead to problems in the calculation of volumes or the directions of water flow.

Completeness of DTMs

The DTM cannot be determined in all areas. Excluded areas can be areas with objects above the terrain, for example houses and trees. Photogrammetric methods require contrast and structure in the imagery, and if these conditions do not exist, such areas must be excluded. Laser scanning requires roughness at the terrain surface and blunders may occur due to multi-path in the neighborhood of buildings. This may also lead to "excluded areas" and grid posts without heights. The amount of the missing areas in relation to the whole area in percent will give a value for the absolute completeness. Furthermore, the applied checking method should give an overview where problematic areas are situated. By means of a thematic map, the zones of intolerable errors can be visualized and the efforts for updating of the DTM can then be estimated.

Difference from reference data	Δh
Number of tested points	n
Root Mean Square Error	$RMSE = \sqrt{\frac{\sum (\Delta h)^2}{n}}$
Maximum difference	$\mid \Delta h_{\max} \mid$
Definition of a blunder (threshold)	S > 3* RMSE
Number of blunders	N
Number of points without blunders	$n' = n - N$
Mean	$\mu = \frac{\sum \Delta h}{n'}$
Standard deviation	$\sigma = \sqrt{\frac{\sum (\Delta h - \mu)^2}{(n'-1)}}$

Table 1. Accuracy measures for DTMs as they are used in the EuroSDR project.

Economy of the checking methods

The economy of the checking methods is influenced by the degree of automation and the amount of manual work. Other factors are investments for tools (e.g. software packages) and training of the personnel. The easiness in use of the method is also an economic factor. Most important for the economy is the availability or generation of reference data with the required accuracy and distribution. If DTMs are used for orthoimage production, the same imagery may be used for the checking and improving of the existing DTMs. This will, of course, save a lot of costs. Orthoimages for the open land are produced with a pixel size between 25-40 cm using images in the scale of 1:15000 - 1:33000; orthoimages for cities are produced with a pixel size between 5 - 10 cm using images in the scale of 1:3000 - 1: 8000. Methods and procedures which avoid reference data and test for blunders and inconsistencies in the DTMs only, are then more economic, but they are not complete checking methods.

OUTLINE OF THE EUROSDR PROJECT

The objective of the EuroSDR project "DTM checking" has been to gain experience with various automated methods of checking and improving of DTMs. The pilot centre of the project, the research group Geoinformatics of Aalborg University, collected three different data sets (consisting of DTMs, stereopairs and ground control points) and made them available on the Internet. The participants in the project applied their methods with all or some of the test data. They delivered results (accuracy measures for the delivered DTM, a corrected DTM and/or a list of blunders) which were then analyzed by the pilot centre by means of accurate reference data. A report was compiled and this article is based on that publication (Höhle, Potuckova, 2006). The **test material** included three DTMs, which were derived by the three different methods of data collection: Automated photogrammetry, digitizing of contour lines, and laser scanning. Beside the DTMs, two stereopairs and some control points for the orientation of the imagery were available for the participants. Also some profiles were marked in orthoimages. The participants of the test with the laser scanning data had to check the height values in these locations. If possible, the method should also be used for improving the delivered DTM. The investigation was separated into **three tasks** (A, B and C). In each of the tasks the data acquisition technology, the spacing of the grid points and the magnitude of errors are different.

Task A dealt with a DTM derived by digital photogrammetry. The provided DTM had a relatively low density (grid spacing = 25 m), a relatively low accuracy (σ_h= 0.7 m) and only a few blunders (N=41). The images for the investigation had a scale of 1: 25000. The reference data for the evaluation were derived by low altitude photography (m_b=1:3000) and its accuracy amounted to σ_h = 0.09 m, which means that the accuracy of the reference data is better than the DTMs' accuracy by a factor of 8. The subareas I-IV have 2033 reference points, that is 59% of all grid posts in these sub areas. The landscape type is an open area.

The DTM of **task B** was derived from 5m contours of a topographic map by scanning. The contours themselves are compiled from different sources (plane table surveying and analogue photogrammetry). These older data acquisition methods surveyed only a few points and lines or plotted the contours continuously. A generalization of the contours took place in the cartographic process afterwards. The scanning of the final contour lines required automated labeling of the contours which is a difficult procedure and which may produce blunders. However, improvements thereafter took place by means of data fusion with more accurate height data from new photogrammetric mapping. The DTM delivered to the participants has a modest density (grid spacing = 10 m), a relatively low accuracy (σ_h=1.4 m) and a relatively large number of blunders (N=165). The reference data for the evaluation by the pilot centre were again derived from low altitude photography and its accuracy is σ_h = 0.09 m, which means that it is better by a factor of 15 in regard to the accuracy of the DTM. The sub-areas (I-IV) have 10390 reference points, which is 47% of all the grid posts. The images for the investigation by the participants are the same as in task A (m_b = 1: 25000).

Task C checked the DTM derived by laser scanning. This DTM has a high density (grid spacing = 1 m), a relatively high accuracy (σ_h= 0.10 m) and very few blunders (N=9). The reference data for the evaluation are derived by GPS/RTK and low altitude photography (m_b = 1: 3000). The accuracy of the reference data is relatively high (σ_h= 0.02 m and σ_h = 0.06 m, respectively); that means the accuracy is better by a factor of 5 and 1.7, respectively. The condition of superior accuracy is fulfilled for the GPS derived reference data only. The GPS data are arranged in 9 profiles. The provided images for the investigation by the participants have a scale of 1: 5000. The landscape type of Task C is built-up area.

DESCRIPTION OF THE APPLIED METHODS

The following research groups participated in the EuroSDR project:
D. Skarlatos and A. Georgopoulus (National Technical University of Athens, School of Surveying, Laboratory of Photogrammetry, Athens, Greece)
M. Potuckova (Aalborg University, Department of Development and Planning, Research Group of Geoinformatics, Denmark)
Z. Paszotta and M. Szumilo (University of Warmia and Mazury, Department of Photogrammetry and Remote Sensing, Olsztyn, Poland)
R. Fiala and J. Sima (University of West Bohemia in Pilsen, Department of Geomatics, Czech Republic)
T. Jancso and J.Zavoti (University of West Hungary, Faculty of Geoinformatics, Szekésfehérvár)
J. S. Kim and J. Shan (Purdue University, West Lafayette, Indiana, USA).

Details of the methods can be read in the articles of each research group which are published in (EuroSDR 2006). A short summary of the methods is presented in the following.

Skarlatos/Georgopoulos (S/G)

Errors in the DTM appear as parallaxes between corresponding points in two overlapping orthoimages. Corrections of the DTM can be derived from the measured parallaxes (Norvelle 1996). By applying a rigorous mathematic solution, height corrections **and** shifts in position from the DTM posts are calculated. Corresponding points are found by means of area-based matching. A novel sub-pixel matching technique using elliptical templates was used. Moreover, an adaptive template size and a skipping of homogeneous areas are applied. Height corrections are calculated with a higher density than the original DTM grid. Due to positional shifts, a

TIN of corrections or of corrected heights is derived. In order to improve the reliability and to decrease the number of mismatches, only points with a correction within a defined interval are accepted. All calculations are carried out fully automatically. The areas, which are not suitable for correlation, are excluded. The high density of corrections gives the advantage that the terrain can be modeled accurately.

Potuckova (P)

The applied method is based on the same principle as in (Norvelle, 1996). Corresponding points are found in two overlapping orthoimages and height corrections are calculated from x-parallaxes between the orthoimages.

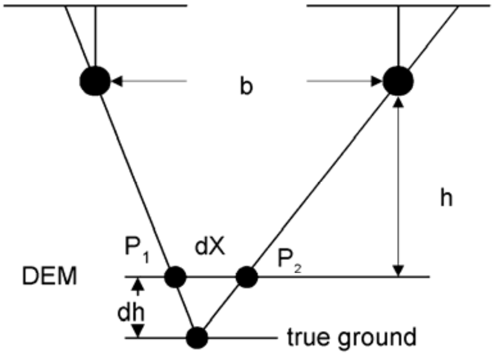

Figure 2. Method of two overlapping ortho images. A parallax (dX) between the two overlapping orthoimages will appear when the DTM/DEM has an error. These x-parallaxes are automatically determined and the height errors (dh) are accurately derived after formula (1):

$$dh \approx \left(dX \cdot \frac{h}{b} \right) + \left(dX \cdot \frac{h}{b} \right)^2 \cdot \frac{1}{h} \qquad (1)$$

Cross-correlation and least squares matching are techniques applied for searching corresponding points. The original color images were converted to gray scale images first. In order to eliminate errors in image matching, the following approaches were used:

- setting thresholds for the correlation coefficient and for the standard deviation of the shift parameters derived in least squares matching
- searching corresponding points along epipolar lines (which are parallel to the line of flight)
- setting thresholds for differences between matching from the left orthoimage (template) to the right one (search area) and vice versa (L/R method)
- calculating of corrections also in the surrounding of the DTM posts and statistical evaluation of these corrections (histogram method)

The histogram method combined with epipolar geometry and thresholds for the correlation coefficient and the accuracy of least squares matching gives the best results. Height corrections are applied directly at DTM posts. Built-up and forest areas can be excluded in advance using map data. In the EuroSDR test, all calculations are restricted to DTM posts which have reference data. After applying the proposed method, each point of the checked DTM is assigned into one of two groups:

- points where the method is applied and the DTM is corrected and
- points where the matching criteria are not fulfilled.

Superimposition of these two groups of points in different colors on the orthoimage gives a quick overview where the problems within the DTM and correlation occur. The principle of the method is depicted in Figure 2.

Paszotta/Szumilo (P/S)

The presented approach uses a stereo-pair of aerial images for deriving a DTM of the same point density and size as the DTM to be checked. Deviations between the original and the derived DTMs are evaluated.
First, control points are measured and orientation parameters of the stereo-pair are determined. The measurement itself as well as the automatic derivation of the DTM is based on area-based matching. An analysis of height differences between the derived and original DTMs follows. A hypothesis about a mean error of the sample is tested at a certain level of significance provided a sufficient sample size is available. At the same time a map of height differences is plotted. It gives a quick overview where the deviations of the compared DTMs are largest. The algorithms are implemented at the Internet.

Fiala/Sima (F/S)

This method caries out a statistical evaluation of height differences between two TINs. The first TIN represents a DTM to be evaluated. The second one is produced as reference data set with higher accuracy. It can be obtained by digitizing contour lines from large-scale maps or by accurate stereoscopic measurements in the stereomodels. In the test, manual photogrammetric measurements of characteristic spot heights and break lines were carried out. Three statistical measures, namely average error, mean error and root mean square error are derived from differences in volume between the TINs. Contour lines of the height differences between TINs give a quick overview on errors in the checked DTM.

Jancso/Zavoti (J/Z)

The proposed DTM checking method is based on back projection of a grid point into the original images. Area-based matching is used for finding corresponding points. Several matching methods are tried by (J/Z). In this paper the optimal solution is used for comparison. Matching is applied in the red, green and blue channels at reduced image resolution. The searching occurs along an epipolar line with a variable size of the search window. In order to ensure good conditions for image matching, a texture coefficient is calculated and evaluated. A calculation of sub pixel values is not applied. The software solution gives a user the possibility of setting several parameters such as the size of the correlation matrix, thresholds for correlation and texture coefficients, maximal height error, etc. Based on values of these parameters each point is evaluated and assigned into one of three categories – accepted, rejected, or skipped point. An output text file comprises, among others, corrected heights. Checking of the exterior orientation of an aerial image is done before the checking of a DTM.

Kim/Shan (K/S)

In the method of K/S a height at a DTM post (Z) is compared with the surrounding heights. A mean value (Z_{mean}) of the heights of eight neighboring DTM posts as well as a standard deviation (σ) are calculated. The value $c = |Z - Z_{mean}| / \sigma$ is compared with two thresholds, c1 and c2. Only DTM points where $c \leq c2$ are considered as correct. DTM posts where $c2 \leq c \leq c1$ are labeled as 'caution area' and values where $c > c1$ indicate blunders. The thresholds c1 and c2 are determined empirically based on diagrams of blunder ratio decrements. They depend on the quality of an original DTM. There is no need for additional data sets (e.g. aerial images) when applying this checking method.

CATEGORIZING OF THE APPLIED METHODS

The described methods of the participants can be categorized in three classes of application: 'Checking for relative accuracy' (C), 'Checking and Improving' (C&I), and 'Checking for Blunders' (CB).
'Checking for relative accuracy' means that the accuracy is checked with small samples using reference data which are derived with the help of the delivered stereopairs. 'Checking and improving' leads to new heights at the majority of the grid posts. When the improved DTM is compared with very accurate reference data an absolute accuracy is obtained. At 'Checking for blunders' the heights of all DTM posts are compared with the heights of their surroundings using an algorithm. The three applications (C&I, C and CB) use three different technologies (photogrammetry, ground surveying and statistics). Statistical methods do not introduce new measurements. Ground surveying can hardly be used for extensive DTMs. Photogrammetry can solve the checking **and** the improving of DTMs. It is a universal method. Table 2 shows the applied technologies in the three applications.

	Statistics	**Photogrammetry**	**Ground surveying**
C&I		x	
C		x	x
CB	x	x	

Table 2. Applicable technologies for the checking of DTMs. It means:
C&I…checking and improving,
C…checking for relative accuracy,
CB…checking for blunders.

RESULTS WITH THE APPLIED METHODS

Relative vertical accuracy

The participants used 'own' reference data, which they derived from the delivered images. The results are here named "relative accuracy". A comparison of the relative accuracy achieved by different participants is given in Figure 3. Three methods (P, S/G, and J/Z) are nearly the same for the RMSE value at test area AI. When comparing all the values, it is obvious that the results for area A differ between 0.7 m and 1.5 m in the

RMSE values. This is a considerable difference. There exists a good agreement between the results of P and F/S at the BII and BIV areas. A large difference of 0.5 m in the RMSE value exists between the results of S/G and P at the area C.

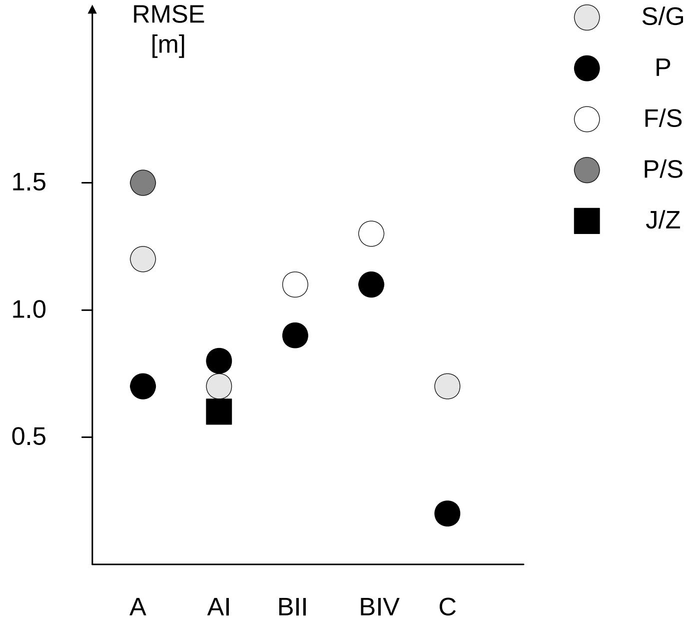

Figure 3. Relative accuracy as determined by the different methods.

Absolute vertical accuracy

The absolute accuracy of the DTMs can be determined by comparing the heights at the grid posts with the accurate reference heights of the pilot centre. Absolute accuracy of the delivered DTMs (DTM_orig) are presented in Table 3.

area	**n**	**RMSE [m]**	**$|\Delta h_{max}|$ [m]**	**N**	**μ [m]**	**σ [m]**
AI	907	0.7	3.9	30	+0.2	0.7
A	2033	0.7	4.9	59	+0.2	0.7
B	10390	1.4	10.9	1636	+0.3	1.4
C	1412	0.13	1.18	10	-0.07	0.10

Table 3. Absolute accuracy of the delivered DTMs (DTM_orig) derived by comparison with the accurate reference DTMs (DTM_ref). The threshold for blunders has been set to S =1.7 m (area A and B) or S = 0.34 m (area C). (The results for area C are given in cm due to a higher accuracy.)

The results of Table 3 confirm the fact that the DTM derived by laser scanning is the most accurate one (RMSE=0.13 m). The checking with reference data from photogrammetry gives nearly the same RMSE value as with GPS/RTK data as reference data ($RMSE_{GPS}$=0.08 m). Automated photogrammetry with images 1:25 000 produced an accuracy of RMSE = 0.7 m or 0.018% of the flying height. The DTM derived from 5 m contours has only a modest accuracy in this test area (RMSE=1.4 m).

Improved DTMs (DTM_cor) were delivered by S/G, P and J/Z only (compare Table 4). The position of the height values should be the same at the delivered (original) DTM and at the DTM used as the absolute reference. A small difference in the position was tolerated (2 m in the DTM of task A and B and 0.5 meter in task C). If the DTM has been delivered as a TIN model, a grid model has first been derived by linear interpolation.

At Table 4 can be observed that the accuracies of the DTMs in tasks A and B at method ‘P’ are the same. The same imagery is used for the correction of the two different DTMs and the same accuracy is therefore achieved.

Participant	S/G			P				J/Z
Test Area	AI	A	C	AI	A	B	C	A1
n	841	1789	1398	726	1729	8973	1099	414
RMSE [m]	1.2	1.6	0.72	0.5	0.5	0.4	0.26	0.7
$\mid \Delta h_{\max} \mid$ **[m]**	5.4	9.7	2.54	2.6	4.6	6.0	2.69	1.9
N	97	363	840	5	8	27	57	4
μ [m]	0.2	0.1	0.00	-0.1	0.1	0.1	0.15	-0.2
σ [m]	0.8	0.8	0.20	0.5	0.4	0.4	0.08	0.6

Table 4. Results of the participants regarding the absolute vertical error in the test areas of task AI, A, B and C. It means: N …number of blunders, μ…systematic shift, σ**…**standard deviation.

Ratio between the relative and the absolute accuracy of the tested DTMs

In order to find out if there is a relation between the relative and the absolute accuracy, Table 5 is compiled**.** It can be seen from Table 5 that the relative accuracy and the absolute accuracy in the DTMs of task A and B are approximately the same. This means that the DTMs of task A and B can be checked and improved by imagery at 1:25 000. This is possible by all of the three methods (S/G, P, and J/Z). In area C, however, the ratio between the relative and the absolute accuracy is much higher than 1.0. This means that the DTM derived by laser scanning data in this test cannot be improved reliably by means of images at 1:5000.

		Relative accuracy		**Absolute accuracy**		**Ratio**	
		cor		**DTM_orig**		**cor/DTM_orig**	
method	**area**	**RMSE [m]**	**σ [m]**	**RMSE [m]**	**σ [m]**	**RMSE**	**σ**
S/G	AI	0.7	0.7	0.7	0.7	1.0	1.0
P	AI	0.8	0.8	0.7	0.7	1.1	1.1
J/Z	AI	0.6	0.6	0.5	0.5	1.2	1.2
S/G	A	1.2	1.2	0.7	0.7	1.7	1.7
P	A	0.7	0.7	0.7	0.7	1.0	1.0
P	B	1.5	1.4	1.4	1.4	1.1	1.0
S/G	C	0.73	0.68	0.13	0.10	5.6	6.8
P	C	0.30	0.16	0.13	0.10	2.3	1.6

Table 5. Ratio between the relative and the absolute accuracy of the tested DTM by different methods. It means: cor/DTM_orig…ratio corrections/original DTM.

Ratio between the delivered DTMs and the improved DTMs

In order to find out whether the delivered DTM could be improved, Figure 4 is compiled. From this figure it is obvious that improvement in the DTM was achieved by method 'P' at task A (DTM derived by digital photogrammetry) by 29% and by 71% at task B (DTM derived by contour lines). The final accuracies of the DTMs are then σ_h=0.013 % (task A) and σ_h=0.011% (task B). The other methods (S/G and J/Z) could not improve the delivered DTMs of task A and B. In task C (DTM derived by laser scanning) no improvement could be achieved by any of the methods with the provided images in the scale of 1:5000. The final accuracy of the 'improved' DTM amounted to σ_h=0.19 m or 0.025 % of the flying height.

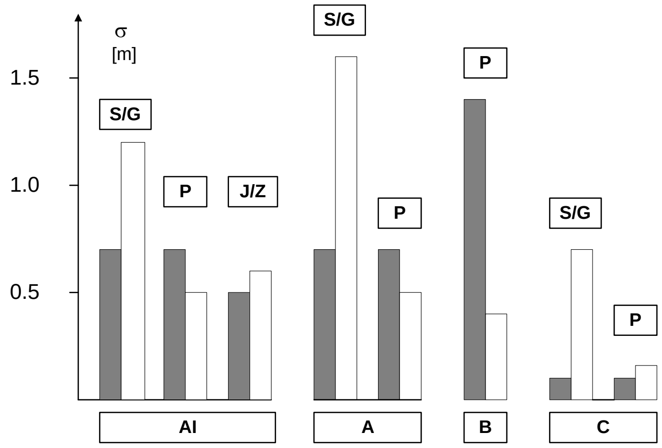

Figure 4. Comparison between the delivered DTM (DTM_orig) and the corrected DTM (DTM_cor).

Additional investigations by the pilot centre with imagery 1:3000 demonstrated that better results in the correction of a DTM can be achieved (n =116, RMSE = 0.06 m, σ_h = 0.04 m = 0.009 % of the flying height). Reference data were the heights of three profiles measured by GPS/RTK. The delivered DTM of task C is already very accurate (RMSE=0.08 m) so that improvements in the standard deviation (σ_h) are difficult to achieve and very likely not required in most of the applications.

Number of blunders

The checking method should be able to detect blunders in the DTM. In order to compare the results of the participants the same threshold was applied. Furthermore, the number of blunders detected and corrected by the method was of interest. A blunder detection ratio and a blunder ratio was derived which also characterizes the methods. The sizes of the RMSE* values for the detected blunders help to interpret the results. Table 6 shows results with areas AI, A and B (threshold S=1.7 m) and the results for area C (threshold S=0.34 m). With the selected threshold (1.7 m) up to 6% of the blunders were present in the DTM derived by digital photogrammetry and up to 16% in the DTM derived from contours of a topographic map. In the DTM derived from laser scanning data less than 1% of the heights were over the selected threshold (0.34 m). For testing the DTMs the chosen thresholds correspond to 3*0.015% of the flying height from which images are taken. The thresholds are then equal for A and B, but different for C.

method	area	N	N [%]	N*	N'	N'/N	N*/N	RMSE* [m]
J/Z	AI	1	0.2	4	1	1.0	4.0	1.9
P	AI	27	3.7	5	25	0.9	0.2	2.7
S/G	A	50	2.9	8	47	0.9	0.2	2.6
P	A	50	2.9	8	47	0.9	0.2	2.6
P	B	1389	15.5	27	1385	1.0	0.0	3.0
S/G	C	10	0.7	840	1	0.1	84.0	0.49
P	C	6	0.5	57	2	0.3	9.5	0.45

Table 6. Number of blunders for test areas in task A, B and C.

It means: N…number of blunders after checking the original DTM with reference values applying a threshold S = 1.7 m for areas of task A and B. The threshold S=0.34m has been used for the areas of task C.

N [%]…number of blunders in % of the total number of points in the original DTM
N*…number of blunders in the corrected DTM
N'…number of blunders detected and corrected by the method
N'/N … blunder-detection-ratio (where 1.0 = 100% detection rate)
N*/N …blunder-ratio
RMSE*…root mean square errors for blunders detected and corrected by the method

The methods for checking and improving of DTMs were different with respect to finding of blunders. Method ‘P’ could detect and correct 93% of the blunders in the DTMs of task A, 100% in the areas of task B, but only 30% in the areas of task C. The other methods (S/G, J/Z) were also successful in the areas of task A. As it can be seen from the RMSE* values of the detected and corrected blunders, the values are different in the three tasks (2.6 m in A, 3.0 m in B and 0.47 m in C).

The evaluation for blunders derived by the **method ’K/S’** has to be dealt with separately. No reference values were derived from the available imagery, but the heights of the surroundings were used for the detection of blunders. The positions (co-ordinates) of detected blunders were delivered to the pilot centre. True reference values in the same position were derived for the evaluation. The results revealed that the success rate for finding blunders has been close to 0% for this method.

Completeness of the checking

The amount of elevations which is tested in comparison with the total number of available reference points is a measure for the completeness of checking. In this investigation the delivered DTMs consisted of all grid posts in the area, but reference values were not available for all of the grid posts. The pilot centre measured only at such grid posts where the bare ground could be measured and where good conditions for correlation existed. The measure for completeness is therefore a relative number. In this case it will be calculated as a percent value which is related to the amount of available reference points. The automated procedures for checking will lead to a large number of check points. The sample size will therefore be large and the statistical values (RMSE, μ, σ) will be reliable.

If the DTM should also be improved, it will then be important to produce an overview about how many points are accepted (regarding a DTM specification) and how many are not tested and improved. Such a graphical display will help in the decision making for further actions. The density of the data is another item for the completeness of the checking. Automated methods can achieve a high density in the checking. If the checking is carried out manually, only a few points of the DTM can be checked.

The checking of completeness has been handled differently by the participating research groups. All test areas (A, B and C) have been checked by method ‘P’ which therefore will be presented here with some further details. Values for the relative completeness are given in Table 7 for all three test areas

area	total number of reference points	accepted points	completeness [%]
A	2033	1729	85
B	10390	8973	86
C	1412	1099	78

Table 7. Evaluation of the (relative) completeness of the checking by method ‘P’.

In addition, all grid posts are plotted on top of orthoimages. They visualize how many points of the DTM are corrected and how many could not be checked and therefore not be improved (compare the example in Figure 5).

In order to give a summarized evaluation of the completeness of the checking a table with three categories (good, medium, poor) is compiled (see Table 8).

method	degree of completeness		
	good	medium	poor
S/G		x	
P		x	
P/S	x		
F/S			x
J/Z		x	
K/S	x		

Table 8. Comparison of the methods with respect to completeness of the checking.

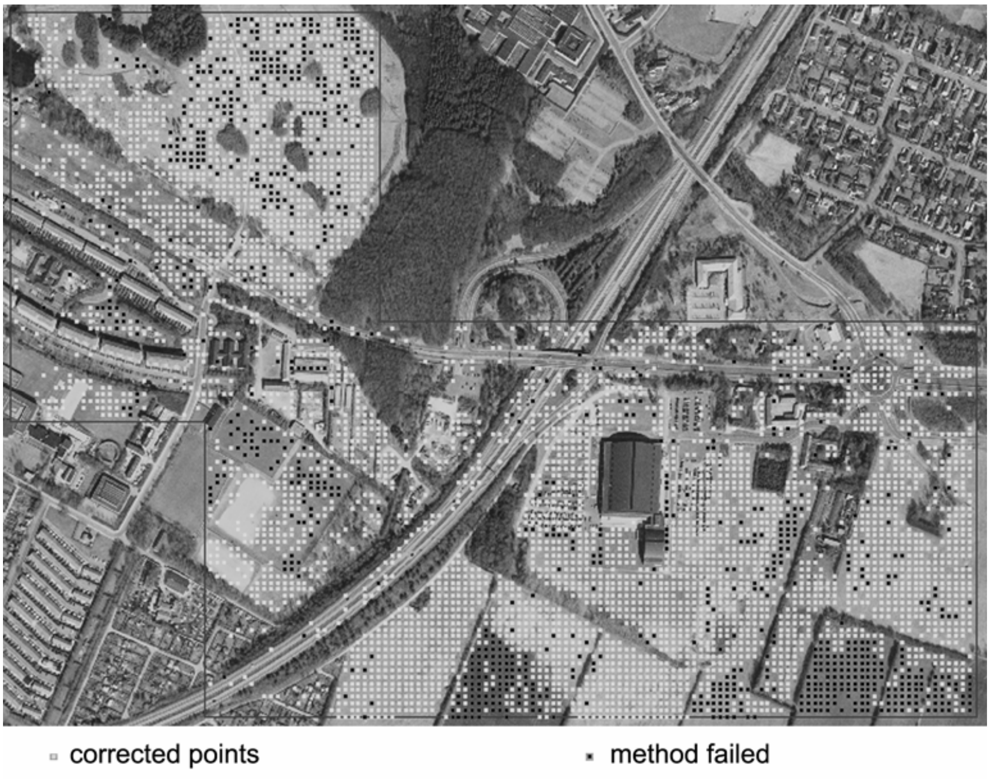

Figure 5. Division of the tested DTM into two categories, example of test B, sub area A_I. The white points fulfilled all set criteria and are corrected; dark points must be checked by other methods. The distance between the grid posts is 10 m. (Source: Potuckova, M., 2006)

A good degree of completeness is possible in the methods ‘P/S’ and ‘K/S’. The manual measurement has a poor degree of completeness. The other methods (S/G, P, J/Z) are characterized as ‘medium”. They test only in positions where good conditions for correlation exist.

DISCUSSIONS

The objectives in this project were to obtain experience with different methods of checking and improving of DTMs. Test material consisting of three DTMs and auxiliary material (overlapping aerial images and control points) has been made available to interested research groups. The provided imagery is in a scale which is adequate both for the production of orthoimages **and** for the checking and improving of DTMs. The three DTMs have had different acquisition methods, accuracy, and density. Besides automated photogrammetry, also manual photogrammetry and a statistical method without reference values have been used by the participants. The participating research groups applied their methods to some of the delivered test areas. A comparison of the methods could therefore be carried out with a few test areas only.

Most of the proposed methods require good conditions for correlation. Reliable results cannot be achieved at all positions of the DTM, but for checking of DTMs such areas with good conditions for automated measurements can be found. The sample size will still be large enough.

An overall **comparison of the investigated methods** can be done by means of Table 9. It contains the items of the investigation and the judgment of their performance.

The method ‘P’ has a positive evaluation in most items of the investigation. It uses two overlapping orthoimages which have to be derived from aerial images first. Original images are used by J/Z. It can be assumed that the original images give better results in the matching and thereby in the accuracy of the corrections, but this could not be confirmed by the results of the EuroSDR investigation.

method	detection & correction of blunders	detection & correction of systematic errors	improvement of standard deviation	completeness of checking & improvement	potential for full automation
S/G	+	+	-	-	+
P	+	+	+	-	+
P/S	-	+	-	+	+
F/S	+	+	-	+	-
J/Z	+	+	-	-	+
K/S	-	-	-	+	+

Table 9. Advantages (+) and disadvantages (-) of the participating methods.

Nearly all methods have the potential for full automation. It needs more programming before the methods can be used in production. Only then a comparison regarding the time consumption can be carried out.

CONCLUSIONS

At National Mapping Agencies and other mapping organizations the quality control of DTMs has become an important task today. The EuroSDR project concentrated on the automatic and semi-automatic checking of the DTM grid posts and the derivation of adequate quality measures. Several new methods for checking and for improving of DTMs have been developed by six different research groups. Their methods have been applied to three different DTMs, and the results of the participants were checked by means of accurate reference data produced by the pilot centre. The majority of the participating research groups used the photogrammetric approach and the provided auxiliary material consisting of a stereomodel and control points. The result of the investigation proved that automatic checking **and** improving of DTMs is possible by aerial images which can also be used for the orthoimage production. When using the best method the automatically corrected DTMs had then an accuracy of σ_h=0.013 % (task A), σ_h=0.011% (task B) and σ_h=0.025% (task C) of the flying height. 85% of the available reference points could then be checked and improved. The grid posts which could not be checked were visualized graphically in this method. At the tasks A and B the used image scale (1:25000) could be used for the checking and improving of the DTMs as well as for the orthoimage production. The DTMs derived by laser scanning needed a larger image scale (1:3000 instead of 1:5000) in order to improve the delivered DTM.

Regarding the number of blunders which could be detected and corrected, the methods of S/G, P and J/Z are equal. At least 93% of the blunders could be detected and corrected. This supplement gives a quick overview what measures have to be taken for completing the improvements. Five of the six proposed methods have the potential for full automation.

ACKNOWLEDGEMENTS

The author of this paper thanks the researchers who participated in the EuroSDR project for their contributions.

REFERENCES

EuroSDR 2006. Official *Publication No 51*, ISSN 0257-0505, December 2006.

Höhle, J. and M. Potuckova, 2006. The EuroSDR test "Checking and Improving of Digital Terrain Models, 136 p., In: (EuroSDR, 2006).

Norvelle, F.R., 1996, Using Iterative Orthophoto Refinements to Generate and Correct Digital Elevation Models. *Digital Photogrammetry*, *an Addendum to the Manual of Photogrammetry*, ASPRS, ISBN 1-57083-037-1, pp. 151-155.

Potuckova, M., 2006, Checking and Improvements of DTMs in the EuroSDR test. In: (EuroSDR, 2006), pp. 63-71.