Beata Hejmanowska

Data inaccuracy in Geographical Information System - propagation of DTM and ortophotomap errors in the spatial analysis .

First, there were numerical maps, usually digitized on the basis of paper maps. Digital maps were often treated as reliable and accurate due to lack of the knowledge of their uncertainty. Such maps are of course burdened with a number of errors made: starting from errors during digitizing process, deformation of the digitized map, through mapping errors to the field measurement errors. It is often believed that a digitized geodesic map has been made according to the technical specifications and, therefore, meets the requirements on accuracy. However, this is often far from the true, especially when the data from different sources are overlaid each other. Until now as yet this problem is inconvenient causing the doubt of a professionalism of land surveyors.

The problem of data base reliability, GIS data bases including, is an important issue in situations when such data are to be used for decision making support (DMS), when a decision with the awareness of its consequences has to be made. The degree to which we trust data is one of the elements which may increase both our consciousness and security when making decisions. It is not a good idea to ignore this problem on behalf of making popular the employing GIS as DMS.

Two issues are discussed in this paper. The first one concerns the Digital Terrain Model (DTM), and the other one – the inaccuracy of an orthophotomap.

The inaccuracy of the DTM was analyzed in the context of determining potential river flood zones, and transposition of errors onto maps from DTM, i.e. slopes and aspects.

In the discussion of the inaccuracy of an orthophotomap, the problem of a parcel area determination with a specific accuracy is undertaken.

Both questions are related with two projects currently realized in Poland, i.e. creation of Co-ordination-Information Centers for flood prevention, and a controlling system for direct subsidies for agriculture (Integrated Administration Control System - IACS). Both they are GIS systems, employing orthophotomaps and DTM as source data.

Inaccuracy of the Digital Terrain Model

Terrain morphology data are frequently applied in analyzing the spatial data bases. This is due to the fact that photogrammetry methods very often supply the ortophotomaps to the data base. A side-effect of an orthophotomap production is a Digital Terrain Model. The presence of DTM in GIS gives a variety of possibilities for making analyses. One of them is to generate derivative maps, i.e. maps of slopes and azimuths. More advanced analyses lie in, e.g. selecting a locality of an object, accounting for boundary slope values and definite aspects. Another example is the selection of an optimum connection between objects, minimizing the cost of making such a connection (e.g., choose such a way between points A and B as to make it pass through slopes of less than 1.5% of inclination; in such a case earthworks can be much reduced).

DTM is also used for flood prevention systems, e.g. for finding flood zones.

Generally, the accuracy of DTM is not homogeneous spatially. It depends on, e.g. technology of its production. DTM can be generated by photogrammetry methods, supplemented with data from direct surveys in the most vital places (e.g., levees), and digitization of the existing cartographic materials (e.g., in forests). Hence, places of high (+/- 0.2 m) and low (> +/- 1.5 m) DTM accuracy are obtained.

It is crucial for GIS analyses to take into account errors in DTM. Two examples where a DTM error had an influence on determination of flood ones, slopes and aspects are presented at length in the below sub-chapters.

Modelling of flood zones on the basis of a Digital Terrain Model, taking into account for its inaccuracies

DTM in flood protection systems is mainly used for determining flood zones. Water level in points on a river is determined on the basis of hydrometeorological data, as well as hydrological and hydraulic models. After generating the water surface in GIS, the crossing of the surface with DTM can be made. As a result, a line determining the boundary of the flood area is obtained. This "hard" method neither accounts for erroneous data (DTM, meteorological data) nor errors in hydro-hydraulic models.

Accounting for the inaccuracy of data bases, the "soft" method determines a probability that a given terrain will be flooded. Assuming the value of DTM error and/or hydro-hydraulic model, the probability that the value of a parameter will be exceeded is calculated.

As a result of crossing (separation) of the surface of water table and DTM, a "difference" map is obtained. There are areas, for which values of difference are either negative (flooded areas) or positive (not flooded areas). The depth of water is determined by the value of the difference. After transforming the difference map and selecting only the zero-difference points, the flood zone boundary is obtained.

Let's assume such a spatially homogeneous model error (having a normal resolution), that the expected value is 0 m, and standard deviation 1 m.

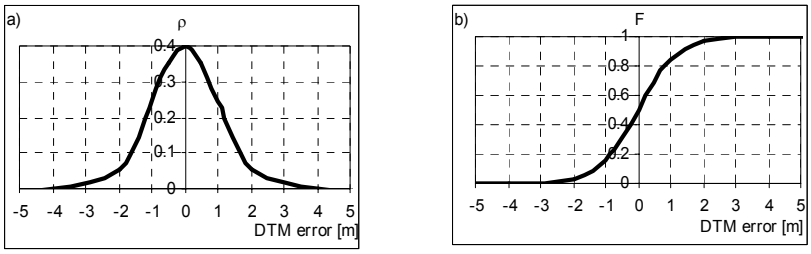

Fig. 1 Distribution of probability (a) and cumulative distribution function (b) of DTM error.

The reliability of the flood zone on the map, obtained by the "hard" method is 50%. Thus, we have a 50% chance that the range of the zone will differ. In the "soft" method, the boundary of the flood zone may be determined at a 10% risk of making an error. Values marked on the difference map are treated as DTM errors, i.e. on the abscissa axis in Fig. 1b, and re-calculate them to probability values (cumulative distribution function – on the ordinate axis in Fig. 1b). It follows from the cumulative distribution function in Fig. 1b, that a 90% probability of flooding can be obtained, provided one considers points in the difference map, for which the difference is less than 1.28 m (not 0 m as in the "hard" method). Analogously, for 95% probability, points below 1.64 m should be selected.

An example of flood zone determination is presented for two cases in Fig. 2; (a) data base errors are not taken into account, and (b) data base errors are taken into account.

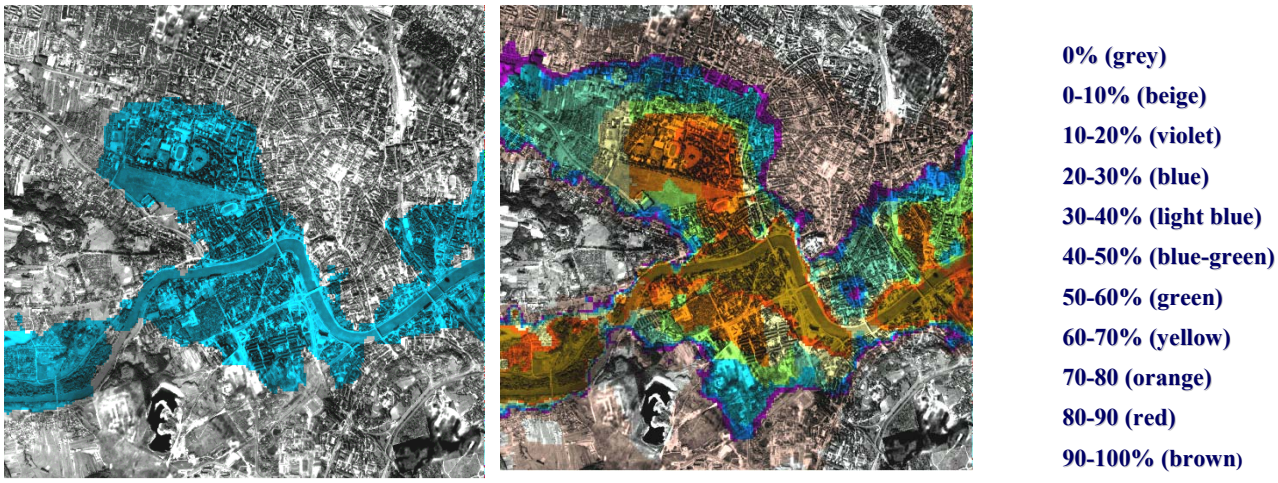

Fig. 2. An example of a flood zone generated on the basis of the "hard" (a) and "soft" (b) methods (Vistula River, center of Cracow).

An influence of DTM inaccuracy on slope and aspect errors

DTM is often used not directly, but through the derivative maps of slopes and aspects. The procedure of determining slope inclination on the basis of a contour map is well known, though tedious to perform. Contrary to the height map, which requires interpolation, grid maps with DTM describe the height of the terrain in every point, therefore can be used for automatic generation of slopes and aspects in every point. More or less complex algorithms for slopes and aspects calculation are implemented in the GIS systems. Such algorithms are always related with an analysis of an area around a point, for which slope and aspects are to be determined (Fig. 3), i.e. point having z_5 height.

z_1	z_2	z_3
z_4	$\mathbf{z_5}$	z_6
z_7	z_8	z_9

---Δh---

Fig. 3. A fragment of the DTM grid.

The algorithm may employ an area approximated on the basis of 9 points and determine the direction and angle of slope on the basis of the greatest slope line. To simplify things, the angle of slope along x and y axes may be determined and the slope and aspect calculated as a result:

$$S = \arctan\left(\sqrt{\frac{(z_3 - z_1)^2 + (z_2 - z_4)^2}{2\Delta h}}\right) \tag{1}$$

$$A = \arctan\left(\frac{z_3 - z_1}{z_2 - z_4}\right) \tag{2}$$

where:

S – slope

A – aspect

dh – grid size

z_i – DTM height, as in Fig. 3

The problem of the influence of DTM error on slope (S) and aspect (A) values is frequently ignored in GIS software packages (MGE, ARCINFO, ARCVIEW, MAPINFO, MFWORKS). The IDRISI software, where new advanced techniques of GIS analysis are implemented, is equipped with additional information saying that after introducing a DTM error and calculating slope and aspect maps, there appears one parameter that can be interpreted as slope error. This parameter is common for the whole map (not commented in the instruction to the program).

Papers discussing this problem can be found in foreign literature. The transposition of DTM error onto S and A values is interpreted in a number of ways. Due to the lack of analytical solutions even for the simplest cases, the author analyzed equations (1, 2), obtaining a relation between error of slope (S) and aspect (A) on DTM error:

$$M_S = \frac{\sqrt{2}}{2\Delta h(1 + tg^2 S)} M_Z \qquad M_A = \frac{1}{\sqrt{2}\Delta h tg S} M_Z \tag{3}$$

where:

M_Z – CRT error

M_s – slope error

M_A – aspect error

Knowing the CMT error (M_z) and using equations (3), slope and aspect errors (M_s and M_A) can be determined. Areas, for which error exceeds S and A values may be excluded from the analysis.

Land parcel selection in the context of possibilities of parcel area's calculation using orthophotomap in IACS with the accuracy recommended by EU.

The direct subsidies system for agriculture (IACS) offers help in breeding and crops production. The Land Parcel Identification System (LPIS) is an important part of the IACS. A number of land parcel identification systems exist across Europe. Agricultural parcels are spatially determined by one type of crop. The LPIS is used for managing the agriculture parcels. In Poland, cadastre (alphanumerical and graphical part of ground register) is a bases of LPIS. Agricultural parcels is define as a part of cadastre parcel but cultivated by one crop. Sometimes the agricultural parcel is equal the cadastre parcel, but sometimes only part of cadastre parcel is cultivated and another is covered by bushes, forest or wasteland. The agricultural parcel may also consist of a few cadastre parcels. The actual area of a agricultural parcel can be determined by removing uncultivated areas basing on the orthophotomap.

The cadastre data (alphanumerical and graphical) will be supplied to LPS from Surveying Resource Centers using Data Exchange Standards (SWDE).

The structure of both agricultural and cadastre parcels vary in different parts of Poland. The cadastre maps in Poland in agricultural regions are generally in two scales (1:5 000 and 1: 2 000 not taking to account ex. the old Austrian maps). Therefore LPIS will use orthophotomaps of different spatial resolution. In the South of Poland, the orthophotomaps will be produced using pixel 0.25 m, whereas in the remaining area – 0.5 m.

To function properly, the LPIS must have an informatic system and data.

The existence of the IACS and partly LPIS is a starting point for EU direct subsidies to agriculture.

With the LPIS it will be possible to manage the data bases of agricultural parcels. In the direct subsidies system, the farmer will have to apply for a subsidies, giving information on the type of crop and area covered by it on the base of reference parcel. In Poland, the reference parcel will be a cadastre parcel, eventually spatial corrected basing on ortopfotomap to remove no cultivated area. To minimize the risk of abuse and make sure that the applications will be submitted only once per parcel, documents will be verified formally. At this stage, the total area within the cadastre parcel as well as the total declared area will be verified. Every year a group of applications will be selected for the controlling of area and

type of culture. The verification is usually made with the use of remote sensing data: aerial orthophotomaps (to determine the area of the agriculture parcel) and satellite images (to determine the land use of the parcel). Multispectral satellite data will be obtained on the basis of a phenological calendar of crops.

The EU Regulation (Article 9 par. 2 of Regulation No. 3887/97) determines the tolerance to the difference between declared and actual area of agricultural parcel. A 3% error is acceptable without any legal consequences. In the case of 3 to 20% error, the farmer has to pay a fine equal to a doubled obtained subsidies. More than 20% errors are fined otherwise. It should be stressed that the controls are made in stages. After a pre-selection of agricultural parcels on the basis of remote sensing data, parcels with exceeded discrepancy between declared and actual values are directly controlled. Farmers are fined only in the case of a discrepancy confirmed by a direct survey of the parcel.

The accuracy of measurement is significant not only to the farmers but also to the whole country and its credibility abroad. In the case of mass violation of the tolerance parameter, the LPIS looses its reliability, which has a negative effect on the magnitude of subsidies granted by the EU. Consequently, a majority of direct subsidies **may have to be paid from the country's budget.**

The IACS control is a broad question, but emphasis will be put on the accuracy of agricultural/cadastral parcel determination.

Resolution of an orthophotomap and a discussion on possibilities of parcel area determination with definite tolerance

EU issued specific regulations on the LPIS and remote sensing control [Discussion Paper "Land Parcel Identification System in the framework of the Regulation EC 1593/2000, Common technical specifications for the 2002 campaign of remote-sensing control of arable and forage land area-based subsidies"].

The minimum admissible accuracy of an orthophotomap corresponds to a scale of 1:10 000, i.e. pixel of orthophotomap = 1m, and RMS of orthophotomap = +/-2.5 m.

It has also been stated that these are the only boundary conditions, and the orthophotomap parameters should be so adjusted to local conditions as to meet the generally defined tolerance.

Depending on the shape and size of the parcels, remote sensing data of various resolution can be applied. Orthophotomaps can be made on the basis of aerial images, high resolution satellite images (<= 1 m) and medium-resolution panchromatic images (SPOT,

IRS). It is also possible to use co-processed high resolution aerial or satellite images (<= 1 m) and medium resolution satellite images. Applying merging techniques and using higher spatial resolution images, the resolution of satellite SPOT and IRS images can be improved. To automatize the process of parcel surveying and verification of declared areas, a simple buffering around the parcel boundary was introduced. The width of the buffer around the parcel depends on the type of source data, on the basis of which the orthophotomap has been made.

If the difference between the declared and surveyed area is smaller or equal to the buffer area, the declared area is accepted as equal to the measured, within the tolerance limit. Otherwise, the parcel is selected for further verification. This is a simple and quick method for parcel verification with the use of standard buffering operations available in GIS.

Table. 1. Width of tolerance buffer proposed for parcel area measurements [3].

Source of data	Width of buffer [m]
Aerial images, 1 m resolution	+/- 1.5
Aerial images merged with panchromatic satellite images	+/- 3
Panchromatic images SPOT/IRS	+/- 5

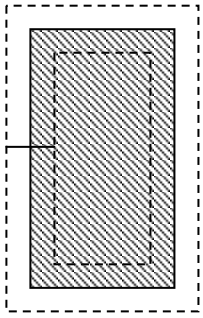

Table 2. Technical tolerance applied to the perimeter
([5] Part 3: technical tolerances and categorisation rules", Point 5.1.4)

	Technical tolerances applied to the perimeter							
	1.5 m		2.0 m		3.0 m		5.0 m	
Parcel area	Tolerance in area		Tolerance in area		Tolerance in area		Tolerance in area	
(rectangle b=4a)	ha	%	ha	%	ha	%	ha	%
0.5 ha	0.06	***11.0%***	0.07	***14.0%***	0.11	***22.0%***	0.18	***35.3%***
1 ha	0.08	***7.5 %***	0.10	***10.0%***	0.15	***15.0 %***	0.25	***25.0%***
2 ha	0.11	***5.8 %***	0.14	***7.0 %***	0.21	***10.5%***	0.35	***17.7%***
5 ha	0.17	***3.4 %***	0.22	***4.5 %***	0.34	***6.7 %***	0.56	***11.1%***
10 ha	0.24	***2.4 %***	0.32	***3.2 %***	0.47	***4.7 %***	0.79	***7.9 %***
20 ha	0.34	***1.7 %***	0.45	***2.3 %***	0.67	***3.4 %***	1.12	***5.6 %***
50 ha	0.53	***1.0 %***	0.70	***1.4 %***	1.06	***2.1 %***	1.75	***3.5 %***
100 ha	0.75	***0.7 %***	1.00	***1.0 %***	1.50	***1.5 %***	2.50	***2.5 %***

It is known that the measurement error depends not only on its area but also on shape, or to be more precise, on its elongation. Error of a square parcel area will be lower than in an

elongated parcel of the same area. It is important to pay attention to the analysis of accuracy for parcels having various shapes. For the analysis sake, a rectangular parcel was assumed (longer to shorter side ratio equal to 4). It can be seen from Table 2 for the orthophotomap (1 m resolution, RMS +/- 2.5 and width of buffer +/- 1.5 m) that the tolerance to a 0.5 ha parcel is 0.06 ha, i.e 11%. For parcels under 5 ha, the tolerance exceeds the admissible 3%, above which farmers are fined. In Poland, the parcels are usually much smaller, especially in the South of Poland (often < 0.3 ha.) It should be emphasized in the subsidy model that depending on the type of culture, the smallest declared parcel area is 0.3 m. It is possible that, e.g. one agricultural parcel with the same crop comprises three cadastre parcels).

A simulation of tolerance analysis was made for Polish conditions. The following assumptions were made:

- a 0.5 ha parcel, rectangular, longer to shorter side ratio equal to 4,
- tolerance pattern as in Table 1,
- simulation of buffer width for an orthophotomap made on the basis of:
  - high resolution satellite images of IKONOS type, 1 m pixel, assumed RMS of orthophotomap +/- 1 m
  - aerial images 1:26 000, resolution 0.5 m, RMS +/- 0.5.

It was knowingly to assume equal values for the root mean square error and the orthophotomap pixel. The root mean square error of the orthophotomap depends on a number of factors and can change from sub-pixel to 3 pixel values. This, however, has not been taken into account in the analysis. The effect of above mentioned simulation is shown in Table 3.

Table 3. Simulated on the basis of Table 1 buffer width with the calculated relative error of parcel area.

Source of data	**Accuracy of orthophotomap**	**Width of buffer**	**Relative error of area**
Aerial images, 1 m resolution (EU)	+/- 2.5 m	1.5 m	11%
IKONOS	+/-1.0 m	0.6 m	4.2%
Aerial photos 1: 26 000	+/-0.5 m	0.3 m	2.1%

It follows from the above Table 3 that in the case a 0.5 ha parcel only an orthophotomap made from aerial images in scale 1:26 000, 0.5 m pixel meets the requirements of < 3% tolerance limit. To obtain an accuracy of 1 pixel on an orthophotomap may be difficult, especially when we treat it as accuracy of identification of a point on the parcel's border.

The above exemplary values are given to illustrate the problem of accuracy of measurement of a land parcel and minimum area that can be measured from an orthophotomap with a definite tolerance limit.

Determination of a minimum orthophotomap accuracy for a parcel area calculation (the relative error of the area is to be lower than the tolerance limit)

Although the above method of accurate parcel area selection by means of a buffering method is very efficient, it does not give information whether or not a given parcel is measurable on a given orthophotomap. This section of the paper will be devoted to a method for determining minimum orthophotomap accuracy for a land parcel calculation. Here the relative parcel area error should be lower than the admissible limit.

The parcel area can be calculated from the Gauss formula:

$$P = \frac{1}{2}\sum_{i=1}^{n} x_i (y_{i+1} - y_{i-1}) \tag{4}$$

where:

P – parcel area

x_i, y_i – co-ordinates of the i-th corner of the parcel

n – number of corners in the parcel

The error of corners location has an effect on the parcel area error:

$$m_p = m_{pkt}\sqrt{\sum_{ii=1}^{n} \frac{(y_{i+1} - y_{i-1})^2 + (x_{i-1} - x_{i+1})^2}{8}} \tag{5}$$

where:

m_{pkt} – error of point location on an orthophotomap

m_p – parcel area error.

Analyzing the parcel area error (5), the following assumption was made:

$$\frac{m_p}{P} <= 0{,}03 \tag{6}$$

where:

m_p – parcel area error

To meet the above requirement, the error of location of point on an orthophotomap should fulfill the condition:

$$m_{pkt} =< \sqrt{\frac{0{,}0072P^2}{\sum_{ii=1}^{n}(y_{i+1}-y_{i-1})^2+(x_{i-1}-x_{i+1})^2}} \qquad (7)$$

Example.

Let's assume that a 0.3 m parcel, 20 x 150 m, i.e. sides ratio equal to 7.5. Having assumed the orthophotomap RMS to +/- 1 m, the parcel area error can be calculated from eq. (5): m_p = +/- 107, i.e. 3.6% of the parcel's area.

To provide the relative area error < 3%, the root mean square error (error in point location on the orthophotomap) should be calculated from eq. (7): m_{pkt} ,+ 0.87 m.

In the case of a parcel of the same area and size 10 x 300 m, this error should be m_{pkt} <= 0.42, respectively.

Knowing the root mean square error of the orthophotomap and having the graphical part of cadastre, it is possible to classify parcels (eq. (7)), the area of which can be established from an orthophotomap with a definite accuracy (e.g., 3%).

Resume

In the times of spatial data bases (GIS), it is important to learn more about the accuracy of data making up these bases. Depending on the type of gathered data, various accuracy levels can be encountered, e.g.

- spatial location of objects
- qualitative and quantitative reliability of attributes

Data gathered in a spatial data base are often modelled, e.g. slopes and aspects are determined on the basis of DTM, water level in rivers from hydraulic models. Thus, data in the GID data base can be burdened with data acquisition errors. These errors are continued in the course of various procedures applied to the source data. Thus, to consciously use data bases, it is crucial to learn about these errors. In the paper the following suggestions of the author were discussed:

- to use the DTM error for flood zones generation
- to assess error of height map derivatives, i.e. maps of slopes and aspects on the basis of a defined DTM error
- to classify land parcels in the IACS (LPIS) to select parcels, for which area can be measured from orthophotomaps with a tolerance limit established by EU.

Literature.

1. Jan J. Konieczny – „Adaptation of Integerated System for management and control system of agricultulral area (IACS) to Polish conditions", Institut of Geodesy and Cartography Publication, Monograph series Nr 2, Warszaw 2001 r (in Polish)

2. http://mars.jrc.it/presentations/control/demonstrateur/run_demonstrator.htm

3. „Discussion Paper: Land Parcel Identification Systems in the frame of Regulation (EC)1593/2000", http://mars.jrc.it/documents/lpis/2580_LPIS_discussion_EN_v_1_4.pdf

4. "Common technical specifications for the 2002 campaign of remote - sensing control of arable and forage land area – based subsidies" http://mars.jrc.it/documents/control/specifications/2903_final_specs_2002_en.pdf

5. Technical recommendations: "Control with remote sensing of arable and forage land"
6. Hejmanowska B. - „Influence of Digital Terrain Uncertainty of slope and aspect errors", Arch. of Photogrammetry, Cartography and Remote Sensing Nr 10, 2000, Kalwaria, Poland, (in Polish),
7. Hejmanowska B., Drzewiecki W., Pirowski T. - " Spatial analysis on the base of Digital Atlas of Cracow Province (Poland)" – Arch. of Photogrammetry, Cartography and Remote Sensing Nr 9, 1994, Olsztyn-Mierki, Poland, (in Polish).
8. Heuvelink G. B.M. – "Error Propagation in Environmental Modelling with GIS), Taylor&Francis, 1998,
9. Weihl R.C., Smith J.L – "The influence of cell slope computation algorithms on a common forest management" – Digital Terrain Modelling II.