

The use of statistical analysis for estimation of positional accuracy of large-scale digital maps

Adam Dorskocz

University of Warmia and Mazury in Olsztyn
Chair of Surveying
12 Heweliusza St., 10-724 Olsztyn, Poland
e-mail: adam.dorskocz@uwm.edu.pl

Received: 21 June 2005/Accepted: 28 September 2005

Abstract: The paper presents estimation of positional accuracy of digital maps using statistical analysis. Investigations have been performed for four large-scale digital maps made using different methods of producing digital map data: new total station survey (object A), re-calculation of previous direct measurements (orthogonal and polar surveys) (object B), manual vectorisation of a raster orthophotomap image (object C) and graphical-and-digital processing of analogue maps (object D).

Analysis has been performed for large statistical samples of sets of vectors of shift ε_L of control points and their components, i.e. true errors ε_X , ε_Y of increments of co-ordinates. In the case of a map produced by means of new survey with an electronic tacheometer, the true errors were represented by differences between co-ordinates of control points obtained from two separate set outs. In the case of other methods of data collection for digital map production true errors were represented by differences of co-ordinates acquired from an investigated map and co-ordinates calculated from new direct surveys.

Keywords: Digital map, accuracy estimation, statistical analysis

1. Introduction and background

Implementing of computerised information systems for administration and organisation of surveying documentation centres should be accompanied by activities aiming at assurance of an appropriate accuracy of data, which are acquired and distributed in the digital form (the digital map database).

The digital data quality may be specified as a set of features connected with the data ability to meet the current and future demands of users. The basic features that describe the digital data quality include: genealogy, accuracy, completeness, logical consistency and temporal information. It is commonly understood that data should be complete, logically consistent and updated to the maximum possible level.

The feature of positional accuracy of the digital map database is not understood explicitly. Discrepancies may be visible at the level of general, national legal regulations,

related to the field of surveying. The Technical Instruction G-4, being the obligatory technical standard in Poland, states, that location of particular details from the 1st accuracy group (including location of boundary beacons) should be set out with the accuracy of at least 0.10 m with respect to the closest sites of the horizontal geodetic control network (Instruction, 1983). On the other hand, the Technical Guidelines K-1.2, recommended for surveying practice, state that the content of the base map should be developed with the accuracy not worse than 0.3 mm (for the details of the 1st accuracy group), at the map scale, with respect to the closest point of the horizontal geodetic control network (Guidelines, 1981). At the same time, the decree on register of lands and buildings within urban areas allows for the use of co-ordinates of boundary beacons, which location has been determined with the accuracy not worse than 0.60 m with respect to sites of the detailed horizontal geodetic control network (Decree, 2001).

The discussed issue does not reflect a local problem that occurs in Poland only. Accuracy of digital data has already been investigated by numerous researchers (e.g. Husár, 1996). Those investigations are also performed with the use of statistical analysis (Bolstad et al., 1990), as well as other, modern research methods (Podobnikar, 1999). They concern, however, maps produced at medium and small scales available in particular countries. In Poland, the large-scale maps exist for the entire country in the form of the base map; the investigations were thus focused on that map.

The aspect of direct relations between results of analysis and correctness (accuracy) of data (the GIGO principle, i.e. "Garbage In, Garbage Out") (Urbański, 1997) has been often discussed in literature. Correlations between data quality and costs of data acquisition are also stressed. Costs of survey depend, in general, on the accuracy of survey (Gaździcki, 1995).

The author of this paper suggests that combination of accuracy requirements with the planned map scale would result in appropriate specification of the required accuracy of digital map data. The Technical Instruction K-1 from 1998 specifies the, so-called, base scales with respect to the base map (specifying principles of map production in a conventional or digital form) (Instruction, 1998). The map in the analogue form at a specified scale is, at present, the base form of map; it will remain one of forms of digital map data presentation in the future.

The present work is the continuation of author's investigations of accuracy of large-scale digital maps using statistical analysis (Doskocz, 2002).

2. The subject matter and methodology of research

2.1. Characteristics of investigated objects

Four digital maps for the cities of Olsztyn and Zielona Góra were the subject of research.

The object A is the digital map in the scale 1:500, produced on the basis of direct survey performed at the university campus in Olsztyn in 1995–96 by student teams with use of an electronic tacheometer (total station survey). The survey was based on the restorable 3rd order geodetic control network. 481 control points covering the area of approximately 200 ha were used for estimation of accuracy.

The object B is the digital base map of the City of Zielona Góra. Materials for testing the accuracy were submitted by the Municipal Surveying Documentation Centre of Zielona Góra. The digital map has been produced basing on existing results of surveys, performed in the period of 1974–1999 (basing on technical traversing of the 2nd order from 1973–74, developed in accordance with the B-III Instruction (Instruction, 1965)), by means of the method of orthogonal measurements, and, in the recent period, by means of the polar method, using an electronic tacheometer (in relation to the restorable 3rd order network). The map accuracy was evaluated for the area of approximately 330 ha (out of 5800 ha of the total area of the city), using 1619 control points.

The object C is the digital orthophotomap of the City of Olsztyn, submitted for testing by the Municipal Department of Surveying in Olsztyn. The digital orthophotomap was produced basing on 1:5000 aerial photographs taken within the Phare Programme in 1995. Aerial photographs were processed to the digital form using a matrix scanner with the resolution of 1000 dpi; then the orthophotomap was developed at the scale of 1:2000. The orthophotomap accuracy was evaluated basing on 311 control points from the area of approximately 115 ha.

The object D is the digital base map of the City of Olsztyn. Selected layers of the vector digital map were submitted by the Municipal Surveying Documentation Centre. The digital map was produced using the method of graphical-and-digital processing of the analogue base map at the scale of 1:500, with the layers of utilities at the scales of 1:500 and 1:1000. The base map was produced basing on technical traversing of the 2nd order from 1974. Survey of details was performed by means of a photogrammetric method, for initial conditions of October 1977. Locations and the terrain relief were delineated by means of Wild A8 stereo plotter and a protractor of details. The map was updated by means of direct surveys, connected to the control network from 1974, and, after 1986 – to the newly established restorable control network of the 3rd order. Accuracy of the digital map was estimated for the area of approximately 355 ha (out of 8800 ha, being the total area of the city), using 2282 control points.

2.2. Research methodology

Accuracy of digital maps was estimated basing on true errors ε_X , ε_Y of increments of plain coordinates of geometric details of the 1st accuracy group (of selected details of the 1st group, being the so-called, control points). In the case of a map produced on the basis of measurements with the use of a total station (object A), the differences of coordinates of control points, surveyed twice (following the theory of measurements in pairs) were used as true errors. With respect to other analysed methods of acquisition of location details used for creation of large-scale digital maps (objects B, C and D), true errors of control points were calculated basing on differences of coordinates obtained from the investigated map and coordinates determined from a new field survey.

Coordinates of analysed details of locations of the objects A, B and D were acquired in the form of text listings or database reports; in the case of the object C coordinates were acquired by means of manual vectorisation of the raster image of the orthophotomap (Doskocz, 2003).

Reference coordinates of control points (if possible, equally distributed within the areas of investigated objects) were determined by means of new field surveys, performed with an electronic tacheometer, in connection to the restorable control network of the 3rd order of the accuracy of $m_p < 0.03$ m, containing redundant observations. Basing on rigorous adjustment of measurements (using the SIEC 95 software tool developed by Gajderowicz) and estimation of their accuracy it has been stated that control points were surveyed with the accuracy of $m_p < 0.02$ m in Olsztyn and with the accuracy of $m_p < 0.03$ m in the case of Zielona Góra.

In the first stage of research, coordinate systems were settled and unified. To do that the coordinates of control points, acquired from sets of evaluated digital maps were transformed into the system of the national horizontal control network (which was the basis for control measurements). Affine transformation was performed basing on points of the horizontal control network of the 3rd order, which existed in the control network of the seventies (technical traversing of the 2nd class: of the 1st order $m_p \leq 0.10$ m or the 2nd order $m_p \leq 0.15$ m (Instruction, 1965)), and then they were included in the newly established restorable network. The mean square error of transformation did not exceed 0.10 m in any of objects investigated.

Basing on differences of coordinates of control points, expressed in the unified coordinate system, true errors of increments of coordinates of analysed geometric details, were determined. The following types of details were considered in research: corners of buildings, boundary beacons and characteristic points of technical utilities.

3. Statistical analysis of sets of true errors

In the course of statistical analysis of sets of true errors, estimation of parameters of distribution of the random variable and verification of compliance of its distribution with theoretical models of errors were performed. It highly influenced the statistical estimation of accuracy of control objects, since – in the case when compliance between the distribution of both sets of true errors ε_x , ε_y of the given control objects with the normal distribution is stated – the random variable ε_L^2 ($\varepsilon_L^2 = (\varepsilon_x^2 + \varepsilon_y^2)$) – the square of the length of the shift vector of the control point – would have the χ^2 distribution with two degrees of freedom.

For the needs of surveying, the ε_L type variable is often used. That variable of two degrees of freedom, expresses the length of the vector, the components of which represent two independent random variables of normal distribution and of equal standard deviations. Some transformations of the χ^2 distribution of practical value are known from literature (Cramér, 1958). An example of such a variable is the linear error of point location (Ney, 1976).

The strategy of estimation of accuracy of large-scale digital maps with use of statistical analysis of true errors is shown in Fig. 1.

Statistical analysis of the sets of true errors $\varepsilon_X, \varepsilon_Y$	
I stage: Determination of parameters of the random variable	
• Calculation of empirical parameters of the random variable „true error”	
II stage: Analysis of distribution of the sets of true errors	
• Determination of the type of the random variable „true error” (continuous or discrete type)	
• Formulation of initial assumptions on the distribution of true errors (basing on empirical parameters and determined type of the random variable)	
• Verification of non-parametric hypotheses with respect to compliance of distribution of the sets of true errors $\varepsilon_X, \varepsilon_Y$ with the normal distribution	
III stage: Estimation of accuracy of control objects (independence of $\varepsilon_X, \varepsilon_Y$ errors assumed)	
<i>In the case when compliance of distribution of the sets of true errors $\varepsilon_X, \varepsilon_Y$ of the given control object with the normal distribution is stated</i>	<i>In the case when no compliance of distribution of at least one of the sets of true errors ε_X or ε_Y of the given control object with the normal distribution is stated</i>
Estimation of accuracy of the control object using the χ^2 distribution (with two degrees of freedom) for the random variable ε_L^2 (the square of length of the shift vector of the control point)	Estimation of accuracy of the control object with the use of statistical definition of probability, where the probability of occurrence is expressed as the relative frequency of occurrence – occurrence of a specified length of the shift vector (ε_L) in the set of control points of the given object (for the sufficient number of sample population)

Fig. 1. The strategy of estimation of accuracy of large-scale digital maps by means of statistical analysis of sets of true errors

3.1. Determination of the maximum elements in empirical sets

Basing on own experience gained in the field of testing accuracy of digital maps, produced from existing materials (Dąbrowski et al., 1999; Dąbrowski and Doskocz, 2000), as well as on results and comments presented by other authors (Augustynowicz, 1993; Latoś and Maślanka, 1998; Salzmann et al., 1998; Choi et al., 2001) it has been stated that investigated digital data do not fully meet accuracy requirements. Besides, the accuracy of existing geometric databases (in the final base scale of 1:500) may be often characterised by means of the following quantities (related to geometric details of the 1st accuracy group, i.e. the, so-called, “well defined points”):

- the mean error of point location of 0.30÷0.50 m,
- numerous cases of point location errors of 1÷2 m,
- exceptional point location errors of 3÷5 m.

Geometric databases investigated are heterogeneous in terms of accuracy of point positions. It was confirmed by experiences of surveying administration and by surveyors practitioners producing digital map data of cities being the subject of investigation in this paper. Therefore, accuracy of objects B and D have been investigated within separate fragments of a city (sub-objects – control objects), using a guideline of surveying administration responsible for the state surveying resources. Control objects were reviewed in terms of genealogy of investigated digital maps; it enabled to determine the probable range of their real accuracy. Each control object was marked with a capital letter (assigned to an object) and a number, marking the successive sub-object.

In order to estimate real accuracy of digital maps, the range of the length of the vector of the control point shift ε_L has been determined; it became the basis for calculation of the accuracy of objects B and D. The vectors of length exceeding the limit value of ε_L should be considered as outliers. It has been assumed that vectors ε_L of the length exceeding the limit value do not represent random quantities. Including of outlying elements to the set may lead to deformation of the empirical expected value and the empirical standard deviation (e.g. Ney, 1976). In the case of the problem discussed in the paper, that might affect results of calculations, and thus, make impossible the determination of real accuracy of investigated digital maps.

Two approaches were applied for the determination of the maximum length of ε_L . The first approach refers to the positional accuracy standards assumed for maps in Poland, which states that the error of location of a terrain detail of the 1st accuracy group should not exceed 0.3 mm at the map scale, with respect to the closest point of the horizontal control network. Considering that the objects B and D are digital base maps of urban areas and their geometric databases should ensure the accuracy at the base scale of 1:500 – it has been assumed that the error of location of details of the 1st group included in those databases should not exceed 0.15 m.

The maximum elements in empirical sets (ε_L) were determined using Tschebyshev inequality:

$$P[|X - E(X)| \geq t \cdot \sigma] \leq \frac{1}{t^2} \quad (1)$$

where $E(X)$ is the expected value, and t is multiplication coefficient of the standard deviation σ . Probability of occurrence of the error of $10 \cdot \sigma$ according to (1) equals to 0.01. Referring to the presented investigations of accuracy of large-scale digital maps it confirms the possibility of occurrence of $\varepsilon_L = 1.50$ m in control sets with the probability of 1%.

In the second approach, the outliers were estimated with the use of the test for gross errors of observations (Smirnow and Dunin-Barkowski, 1969), also called the Hrebbs test (Ney, 1976), in which the value of the indicator of the outlying element (t variable) is given as follows

$$t = \frac{x_{\max} - \bar{x}}{m} \quad (2)$$

where x_{\max} is the maximum element in a set, \bar{x} is the empirical expected value (arithmetic mean), m is the empirical standard deviation (mean square error).

The critical area has been determined using the relation

$$P(t \leq t_{\gamma, N}) = \gamma \quad (3)$$

The limit values $t_{\gamma, N}$ correspond to $\gamma = 0.99$ confidence level. Estimation of outliers in particular sets of control points of the objects B and D is presented in Table 1.

Table 1. Estimation of outliers in particular sets of ε_L of control points of the objects B and D

No	Control object	Size of a set N	Empirical parameters		Application of Hrebbs test		
			\bar{x} [m]	m [m]	$t_{\gamma, N}$	$t_{\gamma, N} \cdot m$ [m]	x_{\max} [m]
1	B-1	257	0.12	0.15	4.00	0.60	0.72
2	B-2	217	0.15	0.21	4.00	0.84	0.99
3	B-3	209	0.24	0.33	4.00	1.32	1.56
4	B-4	241	0.18	0.22	4.00	0.88	1.06
5	B-5	318	0.12	0.20	4.00	0.80	0.92
6	B-6	264	0.10	0.14	4.00	0.56	0.66
7	B-7	72	0.12	0.21	3.60	0.76	0.97
8	D-1	1001	0.31	0.38	4.41	1.68	1.99
9	D-2	549	0.37	0.45	4.24	1.91	2.28
10	D-3	240	0.39	0.46	4.00	1.84	2.23
11	D-4	236	0.26	0.33	4.00	1.32	1.58
12	D-5	134	0.25	0.30	3.77	1.13	1.38
13	D-6	115	0.34	0.39	3.77	1.47	1.81
Sum of the maximum elements estimated by means of Hrebbs test							18.15
Mean value of the maximum element in a set (18.15 m / 13)							1.40

The analysis of the maximum elements in the sets of vectors ε_L estimated by means of Hrebbs test followed by averaging their values and settling the result with the first solution shows that only ε_L not exceeding 1.50 m should be considered in estimation of accuracy of large scale digital maps. Such vectors were used in analyses of objects B and D. Besides, it has been assumed that the occurrence of $\varepsilon_L > 1.50$ m is affected not only by random errors, therefore those vectors were excluded from further investigations. Sets of length of the vector of shift of control points, not larger than 1.50 m, and its components, i.e. true errors ε_x , ε_y of increments of coordinates were the subject of statistical analysis.

However, it should be mentioned, that the assumptions concerning the normal distribution of vectors ε_L in Hrebbs test and use of Tschebyshev inequality for estimation of the maximum lengths of vectors ε_L which characterise the accuracy of the objects B and D are rather simplified; improvement of that part of analysis requires further research.

3.2. Determination of parameters of a random variable

Two groups of parameters, i.e. measures of location and measures of scattering are most frequently applied. The basic numerical characteristics of the distribution of the random variable are particular cases of moments of the random variable (Baran, 1983).

The following parameters describing the basic numerical characteristics are helpful to specify hypothetical reasons of anomalies of distributions of empirical sets (Szacherska and Kurpiewska, 1976; Wiśniewski, 1986):

- existing skewness ($S \neq 0$) in an empirical set proves that the set was created from combination of subsets of various size and of various expected values. This also proves that a deterministic factor occurs (a factor of non-random characteristics) which is systematic with respect to the sign but variable with respect to the value;

- theoretical justification of $e > 0$ coefficient (slender empirical curve) is the occurrence of elements of various accuracy in a set;
- flatness of an empirical curve ($e < 0$) proves that systematic errors occur;
- non-zero expected value of the error ($\bar{x} \neq 0$) indicates the occurrence of permanent errors.

In the discussed analysis of true errors the following values have been empirically determined: the average values $\bar{x} = \sum_{i=1}^N \varepsilon_{x_i} / N$ (expressed by its estimator – the arithmetic mean) and standard deviation of the random variable $m = \sqrt{\sum_{i=1}^N \varepsilon_{x_i}^2 / N}$ (estimated by the mean square error). Asymmetry of the distribution has been specified by asymmetry factor (skewness) $S = \mu_3 / m^3$ and the level of flattening of curves of empirical distribution with respect to theoretical models has also been determined. The factor of flattening (excess) has been calculated as $e = \mu_4 / m^4 - 3$. The following quantities have also been calculated: the mean square error of the average value of the random variable $m_{\bar{x}} = m / \sqrt{N}$, the average error of the random variable $d = \sum_{i=1}^N |\varepsilon_{x_i}| / N$, and the d/m ratio, where N is the size of the set. Parameters of random variables of control sets investigated are given in Table 1 and Table 2.

Table 2. Empirical parameters of the random variables from control objects of digital maps of the City of Olsztyn

Control object (size)	Set of errors	Size of the set	Empirical parameters of random variables						
			\bar{x} [m]	m [m]	$m_{\bar{x}}$ [m]	d [m]	d/m	S	e
A (200 ha)	ε_X	478	0.00	0.05	0.002	0.04	0.80	-0.23	0.04
	ε_Y	478	0.00	0.04	0.002	0.03	0.75	0.01	2.58
C (115 ha)	ε_X	311	-0.01	0.15	0.008	0.11	0.73	0.12	0.64
	ε_Y	311	-0.02	0.15	0.008	0.11	0.73	0.51	1.35
D-1 (46 ha)	ε_X	1001	0.10	0.26	0.008	0.19	0.73	-0.16	2.25
	ε_Y	1001	-0.08	0.28	0.009	0.21	0.75	0.31	1.54
D-2 (154 ha)	ε_X	549	0.10	0.31	0.013	0.23	0.74	-0.32	2.27
	ε_Y	549	-0.09	0.32	0.014	0.24	0.75	0.22	1.34
D-3 (25 ha)	ε_X	240	-0.06	0.31	0.020	0.23	0.74	0.00	0.95
	ε_Y	240	0.06	0.34	0.022	0.27	0.79	-0.03	0.34
D-4 (20 ha)	ε_X	236	0.01	0.22	0.014	0.16	0.73	-0.79	3.25
	ε_Y	236	0.05	0.25	0.016	0.17	0.64	-0.61	4.24
D-5 (100 ha)	ε_X	134	0.11	0.24	0.018	0.18	0.75	-0.60	1.26
	ε_Y	134	-0.02	0.18	0.016	0.14	0.78	0.06	1.17
D-6 (10 ha)	ε_X	115	0.10	0.29	0.027	0.23	0.79	-0.11	-0.77
	ε_Y	115	-0.08	0.25	0.023	0.19	0.76	0.14	0.22

Table 3. Empirical parameters of the random variables from control objects of digital base map of the City of Zielona Góra

Control object (size)	Set of errors	Size of the set	Empirical parameters of random variables						
			\bar{x} [m]	m [m]	$m_{\bar{x}}$ [m]	d [m]	d/m	S	e
B-1 (92 ha)	ε_X	257	-0.04	0.11	0.007	0.08	0.73	0.14	0.70
	ε_Y	257	-0.01	0.10	0.006	0.07	0.70	0.56	1.73
B-2 (13 ha)	ε_X	217	0.00	0.14	0.010	0.09	0.64	0.77	3.46
	ε_Y	217	0.00	0.16	0.011	0.10	0.62	-0.77	9.91
B-3 (46 ha)	ε_X	209	-0.02	0.20	0.013	0.14	0.70	-0.80	2.38
	ε_Y	209	0.06	0.27	0.019	0.17	0.63	1.47	4.61
B-4 (60 ha)	ε_X	241	0.07	0.15	0.010	0.12	0.80	0.08	-0.15
	ε_Y	241	-0.06	0.15	0.010	0.11	0.73	0.15	1.30
B-5 (90 ha)	ε_X	318	0.00	0.12	0.007	0.08	0.67	1.09	13.90
	ε_Y	318	0.00	0.15	0.008	0.07	0.47	1.11	39.34
B-6 (18 ha)	ε_X	264	-0.01	0.11	0.007	0.08	0.73	1.31	4.34
	ε_Y	264	-0.01	0.09	0.006	0.06	0.67	0.03	9.59
B-7 total (7 ha)	ε_X	96	0.05	0.26	0.026	0.16	0.62	1.58	2.57
	ε_Y	96	0.15	0.31	0.032	0.20	0.64	1.15	0.39
B-7 sheet 2(2) (5 ha)	ε_X	72	0.04	0.13	0.015	0.09	0.69	0.43	0.59
	ε_Y	72	-0.06	0.16	0.019	0.12	0.75	-0.34	-0.88

3.3. Analysis of distribution of true errors

The second stage of analysis was focused on the determination of the distribution of empirical sets. Initially the type of the random variables “true error of increment of coordinates” ε_X and ε_Y was visually recognised as the continuous one. Finally, the continuous type of the random variable was verified with the use of the hypothesis of wave structure of errors (Adamczewski, 1961).

The specified type of random variable and calculated empirical parameters of the random variable (Tables 2 and 3) allowed to assume that distributions of investigated sets of true errors may belong to “a family” of normal distributions. Besides, basing on empirical parameters, initial specification of distribution of true errors was performed. It was noted that the majority of sets are asymmetric ($S \neq 0$) and are affected by a systematic factor ($\bar{x} \neq 0$). It was also found that, for the considerable part of sets, the obtained d/m ratio vary from its theoretical value in the normal distribution (for Gaussian distribution that ratio equals to $d/m = \sqrt{2/\pi} \cong 0.80$ (Ney, 1976)). The excess values in empirical sets are almost all positive, what indicates more distinctive slenderness of empirical curves of probability density of errors than a normal curve of the variance determined in the given empirical set. Therefore, the use of modified normal distribution, in particular the Modulated Normal Distribution (MND) (Romanowski, 1974; 1979) was considered. The analysis of relative frequency of true errors ε_X , ε_Y in the investigated sets indicates the similarity of their distribution with the MND distribution of the 1st type.

The empirical sets of observation errors frequently substantially differ from the Gaussian model; it concerns asymmetry and excess (Szacherska, 1974; Łyszkowicz, 1975b; Wiśniewski, 1984; 1986). Such differences were also exhibited by the empirical sets investigated. In order to determine the empirical distribution density curves, Pearson

method was applied, that requires calculating numerical coefficients β_1, β_2, k (basing on values of central moments of empirical sets, calculated earlier): $\beta_1 = \mu_3^2/\mu_2^3, \beta_2 = \mu_4/\mu_2^2, k = \beta_1 \cdot (\beta_2 + 3)^2/4 \cdot (2 \cdot \beta_2 - 3 \cdot \beta_1 - 6) \cdot (4 \cdot \beta_2 - 3 \cdot \beta_1)$. Those coefficients become the basis for the determination of the curve, out of 12 Pearson curves, that might be used for approximation of the given empirical set. Pearson distributions of I, II, IV and VII types were identified as suitable for typical surveying data (Łyszkowicz, 1975a; 1975b).

The search for Pearson distributions that might be used for approximation of empirical sets was undertaken. Besides, two statistical tests, i.e. Pearson χ^2 test and Kołmogorow λ test were used for verification of non-parametric hypotheses concerning the compliance of distributions of empirical sets with theoretical distributions (Gaussian and MND 1st type distribution characterized by the leptokurtosis index ω). Investigations of compliance of distribution of true errors of $\varepsilon_X, \varepsilon_Y$ with theoretical models was performed for the standardized random variable (e.g. $\varepsilon_X' = (\varepsilon_X - \bar{x})/m$).

The χ^2 test is of higher sensitivity than λ test, but, at the same time, it requires larger set (at least above one hundred elements) (Ney, 1970). The χ^2 test better characterizes the difference between the empirical distribution and the theoretical model, since it uses all class intervals of a distributive series. In the λ test the level of compliance between the empirical and theoretical distributions is expressed by the point value D_i that is the difference between empirical and theoretical distribution function. The argument that supports the use of the Kołmogorow λ method is that the λ test is not sensitive on discrepancies which occur for final small class intervals of the distributive series (which often prevail in empirical sets) (Hellwig, 1998).

The sets $\varepsilon_X, \varepsilon_Y$ analysed in the presented work are slender (leptokurtic) in majority of cases, with respect to Gaussian distribution (except of ε_X and ε_Y sets of the object A, the ε_Y set of the object D-5, and the ε_Y set of the object B-7); prevailing discrepancies of their distribution occur in central class intervals of the distributive series.

Results of investigations of compliance between distributions of sets of true errors with theoretical distributions are presented in Tables 4, 5, 6 and 7. The coefficient ω is the measure of leptokurtosis of the probability density distribution curve of the empirical set with respect to the normal distribution curve. The leptokurtosis index ω is the ratio of relative frequency of occurrence of the error of zero value in empirical distribution to the maximum probability of occurrence of the error of zero value in the normal distribution. The $r = p - (s + 1)$ coefficient is the number of degrees of freedom in the χ^2 distribution of the χ_0^2 variable (calculated in the Pearson χ^2 test of compliance), where p is the number of class intervals of the distributive series and s is the number of parameters of empirical distribution, determined using the random sample.

The Kołmogorow λ method was applied as an auxiliary test when it was necessary.

Table 4. Analysis of compliance of distribution of true errors with theoretical Gaussian distribution for control objects of digital maps of the City of Olsztyn

Control object	Set of errors	Size of the set	Numerical coefficients					Verification of compliance of empirical distribution with normal distribution				
			β_1	β_2	k	ω	r	χ_0^2 and, possibly D_{max}, λ_0	Level of significance of the test			
									$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	
									Mutual relation of values read-out from tables $\chi_a^2(\lambda_a)$ and calculated in the test $\chi_0^2(\lambda_0)$			
A	ε_X'	478	0.05	3.04	0	0.94	6	38.36	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$	
	ε_Y'	478	0	5.58	0	0.91	8	22.27	$\chi_{(a,r)}^2 = 1.646; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.733; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 3.490; \chi_{(a,r)}^2 < \chi_0^2$	
C	ε_X'	311	0.01	3.64	0	1.50	6	26.60	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$	
	ε_Y'	311	0.26	4.35	-0.01	1.40	7	41.52	$\chi_{(a,r)}^2 = 1.239; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.167; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.833; \chi_{(a,r)}^2 < \chi_0^2$	
D-1	ε_X'	1001	0.02	5.25	0	1.47	7	126.43	$\chi_{(a,r)}^2 = 1.239; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.167; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.833; \chi_{(a,r)}^2 < \chi_0^2$	
	ε_Y'	1001	0.10	4.54	0	1.31	8	60.28	$\chi_{(a,r)}^2 = 1.646; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.733; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 3.490; \chi_{(a,r)}^2 < \chi_0^2$	
D-2	ε_X'	549	0.10	5.27	0	1.37	6	43.85	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$	
	ε_Y'	549	0.05	4.34	0	1.33	6	24.34	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$	
D-3	ε_X'	240	0	3.95	0	1.42	4	12.96	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$	
	ε_Y'	240	0	3.34	0	1.21	6	12.83	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$	
D-4	ε_X'	236	0.70	6.25	-0.03	1.43	6	27.58	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$	
	ε_Y'	236	0.38	7.24	-0.0	1.69	5	39.53	$\chi_{(a,r)}^2 = 0.554; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.145; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.610; \chi_{(a,r)}^2 < \chi_0^2$	
D-5	ε_X'	134	0.36	4.26	-0.02	1.69	3	22.72	$\chi_{(a,r)}^2 = 0.115; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.352; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.584; \chi_{(a,r)}^2 < \chi_0^2$	
	ε_Y'	134	0	4.17	0	0.80	4	6.25	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$	
D-6	ε_X'	115	0.01	2.23	0	1.04	4	0.99	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 > \chi_0^2$	
	ε_Y'	115	0.02	3.22	0	1.45	2	14.64	$\lambda_\alpha = 0.44; \lambda_\alpha > \lambda_0$	–	–	
									$\chi_{(a,r)}^2 = 0.020; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.103; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.211; \chi_{(a,r)}^2 < \chi_0^2$	

Table 5. Analysis of compliance of distribution of true errors with theoretical models of errors for control objects of digital maps of the City of Olsztyn

Control object	Set of errors	Size of the set	Numerical coefficients					Verification of compliance of empirical distribution with theoretical distributions (Pearson and MND)							
			β_1	β_2	k	ω	r	Type of Pearson distribution	ω, χ_0^2 and, possibly D_{\max}, λ_0	Level of significance of the test					
										$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
										Mutual relation of values read-out from tables $\chi_\alpha^2(\lambda_\alpha)$ and calculated in the test (λ_0)					
A	ε_X'	478	0.05	3.04	0	0.94	6	–	–	–	–	–	–	–	
	ε_Y'	478	0	5.58	0	0.91	8	VII	–	–	–	–	–	–	
C	ε_X'	311	0.01	3.64	0	1.50	6	–	1.41; 6.79	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$			
									0.0190; 0.34	$\lambda_\alpha = 0.44; \lambda_\alpha > \lambda_0$	–	–			
	ε_X'	311	0.26	4.35	-0.01	1.40	7	–	1.41; 14.86	$\chi_{(a,r)}^2 = 1.239; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.167; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.833; \chi_{(a,r)}^2 < \chi_0^2$			
									0.0209; 0.37	$\lambda_\alpha = 0.44; \lambda_\alpha > \lambda_0$	–	–			
D-1	ε_X'	1001	0.02	5.25	0	1.47	7	–	1.41; 49.02	$\chi_{(a,r)}^2 = 1.239; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.167; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.833; \chi_{(a,r)}^2 < \chi_0^2$			
	ε_Y'	1001	0.10	4.54	0	1.31	8	–	1.24; 20.67	$\chi_{(a,r)}^2 = 1.646; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.733; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 3.490; \chi_{(a,r)}^2 < \chi_0^2$			
D-2	ε_X'	549	0.10	5.27	0	1.37	6	–	1.41; 18.56	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$			
	ε_Y'	549	0.05	4.34	0	1.33	6	–	1.41; 6.41	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$			
									0.0190; 0.44	$\lambda_\alpha = 0.44; \lambda_\alpha = \lambda_0$	$\lambda_\alpha = 0.52; \lambda_\alpha > \lambda_0$	–			
D-3	ε_X'	240	0	3.95	0	1.42	4	VII	1.41; 2.78	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$			
									0.0193; 0.30	$\lambda_\alpha = 0.44; \lambda_\alpha > \lambda_0$	–	–			
	ε_X'	240	0	3.34	0	1.21	6	VII	1.24; 15.34	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$			
D-4	ε_X'	236	0.70	6.25	-0.03	1.43	6	–	1.41; 12.61	$\chi_{(a,r)}^2 = 0.872; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.635; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 2.204; \chi_{(a,r)}^2 < \chi_0^2$			
	ε_Y'	236	0.38	7.24	-0.0	1.69	5	–	1.96	–	–	–			
D-5	ε_X'	134	0.36	4.26	-0.02	1.69	3	–	1.96	–	–	–			
	ε_Y'	134	0	4.17	0	0.80	4	VII	–	–	–	–			
D-6	ε_X'	115	0.01	2.23	0	1.04	4	–	1.04	–	–	–			
	ε_Y'	115	0.02	3.22	0	1.45	2	–	1.41; 8.93	$\chi_{(a,r)}^2 = 0.020; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.103; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.211; \chi_{(a,r)}^2 < \chi_0^2$			

Table 6. Analysis of compliance of distribution of true errors with theoretical Gaussian distribution for control objects of digital base map of the City of Zielona Góra

Control object	Set of errors	Size of the set	Numerical coefficients					Verification of compliance of empirical distribution with normal distribution			
			β_1	β_2	k	ω	r	χ_0^2 and, possibly $D_{\max}; \lambda_0$	Level of significance of the test		
									$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
									Mutual relation of values read-out from tables $\chi_a^2(\lambda_a)$ and calculated in the test $\chi_0^2(\lambda_0)$		
B-1	ϵ_X'	257	0.02	3.70	0	1.46	5	24.12	$\chi_{(a,r)}^2 = 0.554; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.145; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.610; \chi_{(a,r)}^2 < \chi_0^2$
	ϵ_Y'	257	0.31	4.73	-0.01	1.30	5	29.23	$\chi_{(a,r)}^2 = 0.554; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.145; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.610; \chi_{(a,r)}^2 < \chi_0^2$
B-2	ϵ_X'	217	0.59	6.46	-0.02	1.66	4	33.80	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$
	ϵ_Y'	217	0.60	12.91	-0.02	1.80	4	49.54	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$
B-3	ϵ_X'	209	0.64	5.38	-0.03	1.48	4	25.75	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$
	ϵ_Y'	209	2.15	7.61	-0.10	1.60	3	53.58	$\chi_{(a,r)}^2 = 0.115; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.352; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.584; \chi_{(a,r)}^2 < \chi_0^2$
B-4	ϵ_X'	241	0.01	2.85	0	1.38	4	14.09	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$
								0.06065; 0.94	$\lambda_a = 0.44; \lambda_a < \lambda_0$	$\lambda_a = 0.52; \lambda_a < \lambda_0$	$\lambda_a = 0.58; \lambda_a < \lambda_0$
	ϵ_X'	241	0.02	4.30	0	1.23	4	12.08	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$
B-5	ϵ_X'	318	1.19	16.90	-0.03	1.64	4	45.42	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$
	ϵ_Y'	318	1.23	42.34	-0.03	2.43	3	195.69	$\chi_{(a,r)}^2 = 0.115; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.352; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.584; \chi_{(a,r)}^2 < \chi_0^2$
B-6	ϵ_X'	264	1.71	7.34	-0.08	1.90	5	78.67	$\chi_{(a,r)}^2 = 0.554; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.145; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.610; \chi_{(a,r)}^2 < \chi_0^2$
	ϵ_Y'	264	0	12.59	0	1.65	4	46.46	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$
B-7 total	ϵ_X'	96	2.48	5.57	-0.17	1.88	1	50.46	$\chi_{(a,r)}^2 = 0.000; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.004; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.016; \chi_{(a,r)}^2 < \chi_0^2$
	ϵ_Y'	96	1.32	3.39	-0.10	1.55	2	16.75	$\chi_{(a,r)}^2 = 0.020; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.103; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.211; \chi_{(a,r)}^2 < \chi_0^2$
B-7 sheet 2(2)	ϵ_X'	72	0.19	3.59	-0.01	1.42	2	9.20	$\chi_{(a,r)}^2 = 0.020; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.103; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.211; \chi_{(a,r)}^2 < \chi_0^2$
								3.06	$\chi_{(a,r)}^2 = 0.020; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.103; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.211; \chi_{(a,r)}^2 < \chi_0^2$
	ϵ_Y'	72	0.11	2.12	-0.01	0.92	2	0.05703; 0.48	$\lambda_a = 0.44; \lambda_a < \lambda_0$	$\lambda_a = 0.52; \lambda_a > \lambda_0$	–

Table 7. Analysis of compliance of distribution of true errors with theoretical models of errors for control objects of digital base map of the City of Zielona Góra

Control object	Set of errors	Size of the set	Numerical coefficients					Verification of compliance of empirical distribution with theoretical distributions (Pearson and MND)						
			β_1	β_2	k	ω	r	Type of Pearson distribution	$\omega_i; \chi_0^2$ and, possibly $D_{\max}; \lambda_0$	Level of significance of the test				
										$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
										Mutual relation of values read-out from tables $\chi_a^2(\lambda_a)$ and calculated in the test $\chi_0^2(\lambda_0)$				
B-1	ε_X'	257	0.02	3.70	0	1.46	5	–	1.41; 10.34	$\chi_{(a,r)}^2 = 0.554; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.145; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.610; \chi_{(a,r)}^2 < \chi_0^2$		
	ε_X'	257	0.31	4.73	-0.01	1.30	5	–	0.0396; 0.63	$\lambda_\alpha = 0.44; \lambda_\alpha < \lambda_0$	$\lambda_\alpha = 0.52; \lambda_\alpha < \lambda_0$	$\lambda_\alpha = 0.58; \lambda_\alpha < \lambda_0$		
B-2	ε_X'	217	0.59	6.46	-0.02	1.66	4	–	1.41; 14.99	$\chi_{(a,r)}^2 = 0.554; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.145; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.610; \chi_{(a,r)}^2 < \chi_0^2$		
	ε_Y'	217	0.60	12.91	-0.02	1.80	4	–	1.96	–	–	–		
B-3	ε_X'	209	0.64	5.38	-0.03	1.48	4	–	1.41; 11.33	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$		
	ε_Y'	209	2.15	7.61	-0.10	1.60	3	–	1.41; 32.20	$\chi_{(a,r)}^2 = 0.115; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.352; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.584; \chi_{(a,r)}^2 < \chi_0^2$		
B-4	ε_X'	241	0.01	2.85	0	1.38	4	–	1.41; 3.31	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$		
	ε_X'	241	0.02	4.30	0	1.23	4	–	0.0294; 0.46	$\lambda_\alpha = 0.44; \lambda_\alpha < \lambda_0$	$\lambda_\alpha = 0.52; \lambda_\alpha > \lambda_0$	–		
B-5	ε_X'	318	1.19	16.90	-0.03	1.64	4	–	1.24; 5.25	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$		
	ε_Y'	318	1.23	42.34	-0.03	2.43	3	–	0.0361; 0.56	$\lambda_\alpha = 0.44; \lambda_\alpha < \lambda_0$	$\lambda_\alpha = 0.52; \lambda_\alpha < \lambda_0$	$\lambda_\alpha = 0.58; \lambda_\alpha > \lambda_0$		
B-6	ε_X'	264	1.71	7.34	-0.08	1.90	5	–	1.41; 16.31	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$		
	ε_Y'	264	0	12.59	0	1.65	4	–	1.96	–	–	–		
B-7 total	ε_X'	96	2.48	5.57	-0.17	1.88	1	–	1.41; 19.61	$\chi_{(a,r)}^2 = 0.297; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.711; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 1.064; \chi_{(a,r)}^2 < \chi_0^2$		
	ε_Y'	96	1.32	3.39	-0.10	1.55	2	I(J)	1.96	–	–	–		
B-7 sheet 2(2)	ε_X'	72	0.19	3.59	-0.01	1.42	2	–	1.41; 14.23	$\chi_{(a,r)}^2 = 0.020; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.103; \chi_{(a,r)}^2 < \chi_0^2$	$\chi_{(a,r)}^2 = 0.211; \chi_{(a,r)}^2 < \chi_0^2$		
	ε_X'	72	0.11	2.12	-0.01	0.92	2	I	0.0778; 0.66	$\lambda_\alpha = 0.44; \lambda_\alpha < \lambda_0$	$\lambda_\alpha = 0.52; \lambda_\alpha < \lambda_0$	$\lambda_\alpha = 0.58; \lambda_\alpha < \lambda_0$		

3.4. Summary of results of analysis of distribution of true errors

In the first two stages of statistical analysis (Fig. 1) it was found that the distributions of ε_X , ε_Y sets "are governed by their own law" which makes it difficult to present them by means of known, theoretical models of errors. Only in a few cases the compliance between empirical distribution with the normal distribution was stated. The only set which proved the compliance between the distribution of the empirical set and the normal distribution was the empirical set ε_X of D-6 control object (Table 4).

Distributions of investigated sets of true errors are, in general, not compliant with MND Romanowski distributions (Tables 5 and 7), what results from their skewness as it is known from literature "utilisation of modulated normal distributions is not fully consistent in conditions of existing asymmetry" (Wiśniewski, 1986).

Also Pearson curves, which accept the presence of asymmetry and excess in an empirical set, could, to a very limited extend (ε_Y set of A, ε_Y set of B-7, ε_X and ε_Y sets of D-3, ε_Y set of D-5) approximate investigated sets of true errors (Tables 5 and 7).

4. Estimation of accuracy of investigated digital maps

Because of lack of compliance between the random variable "the square of the length of the shift vector" and the theoretical distribution χ^2 , the probability of occurrence of specified values of the point location error m_p (expressed by the length of the vector of shift of the control point ε_L) has been estimated on the basis of the relative frequency W_N of occurrence of particular values ε_L within the control objects. The required probability was determined by means of the method presented in (Smirnow and Dunin-Barkowski, 1969).

The confidence interval for the unknown probability may be determined by that method when the condition $N \cdot P \cdot Q > 9$ is met (N is the size of the set; P is the event probability to be determined; Q is the probability of the complementary event). This condition defines the sufficient size of the sample that is required for estimating probability of an analysed event.

The limits of confidence interval for the estimated probability P are expressed as $P_1(W, N) < P < P_2(W, N)$ (Smirnow and Dunin-Barkowski, 1969), where

$$P_1(W, N) = \frac{2 \cdot N \cdot W_N + t_\gamma^2 - t_\gamma \cdot \sqrt{D}}{2 \cdot (N + t_\gamma^2)}$$

$$P_2(W, N) = \frac{2 \cdot N \cdot W_N + t_\gamma^2 + t_\gamma \cdot \sqrt{D}}{2 \cdot (N + t_\gamma^2)}$$

$$\sqrt{D} = \sqrt{4 \cdot N \cdot W_N \cdot (1 - W_N) + t_\gamma^2}$$

In this work probabilities have been estimated for the confidence level $\gamma = 0.997$ for $t_\gamma = 3$.

Statistical estimation of the positional accuracy of large scale digital maps, in particular control objects, is given in Tables 8 and 9. Columns 6 and 8, present probabilities that the

map error does not exceed stated point location error (m_p) multiplied by factor 3 and 2, respectively, at the given control object. Column 10 presents the probability that the theoretical error will be exceeded (m_t – determined in accordance to the accuracy standard for mapping). The symbol (*) by the confidence interval means that the probability is determined in situation when the relative frequency of the considered event in the empirical set had not ensured that the condition of the sufficient sample size was fulfilled.

Table 8. Statistical estimation of accuracy of control objects on digital maps of the City of Olsztyn

Control object	Analysed variable	Size of the set	m_p [m]	Statistical characteristics of accuracy of control objects determined for the confidence level $\gamma = 0.997$					
				Estimated probabilities of occurrence of the(Δ) error in general population					
				W_N	$\Delta \leq 3m_p$	W_N	$\Delta \leq 2m_p$	W_N	$\Delta > m_t$
1	2	3	4	5	6	7	8	9	10
A	ε_L	478	0.04	0.95	0.91<P<0.97	–	–	0.02	0.01<P<0.05*
C	ε_L	311	0.21	–	–	0.97	0.92<P<0.99	0	0<P<0.03*
D-1	ε_L	1001	0.38	0.99	0.98<P<1*	0.95	0.93<P<0.97	0.19	0.16<P<0.24
D-2	ε_L	549	0.45	0.99	0.97<P<1*	0.95	0.92<P<0.97	0.27	0.22<P<0.33
D-3	ε_L	240	0.46	1	0.96<P<1*	0.95	0.90<P<0.98	0.32	0.24<P<0.42
D-4	ε_L	236	0.33	0.98	0.93<P<0.99*	0.96	0.90<P<0.98*	0.15	0.09<P<0.23
D-5	ε_L	134	0.30	–	–	0.96	0.88<P<0.99*	0.14	0.07<P<0.26
D-6	ε_L	115	0.39	–	–	0.97	0.88<P<0.99*	0.29	0.18<P<0.42

Table 9. Statistical estimation of accuracy of control objects on the digital base map of the City of Zielona Góra

Control object	Analysed variable	Size of the set	m_p [m]	Statistical characteristics of accuracy of control objects determined for the confidence level $\gamma = 0.997$					
				Estimated probabilities of occurrence of the(Δ) error in general population					
				W_N	$\Delta \leq 3m_p$	W_N	$\Delta \leq 2m_p$	W_N	$\Delta > m_t$
1	2	3	4	5	6	7	8	9	10
B-1	ε_L	257	0.15	–	–	0.97	0.92<P<0.99*	0.004	0<P<0.04*
B-2	ε_L	217	0.21	0.99	0.94<P<1*	0.96	0.90<P<0.98*	0.03	0.01<P<0.09*
B-3	ε_L	209	0.33	0.98	0.93<P<1*	0.92	0.85<P<0.96	0.13	0.08<P<0.22
B-4	ε_L	241	0.22	–	–	0.94	0.87<P<0.97	0.05	0.02<P<0.12
B-5	ε_L	318	0.20	0.98	0.94<P<0.99*	0.97	0.93<P<0.99*	0.03	0.01<P<0.07*
B-6	ε_L	264	0.14	0.98	0.94<P<0.99*	0.96	0.90<P<0.98	0.02	0<P<0.06
B-7 total	ε_L	96	0.41	0.98	0.88<P<1*	0.91	0.78<P<0.96*	0.15	0.07<P<0.28
B-7 sheet 2(2)	ε_L	72	0.21	–	–	0.97	0.84<P<1*	0.03	0<P<0.16*

5. Summary of performed research

Results of statistical analysis proved the lack of compliance of considered sets of true errors with the normal distribution (except the ε_x set of the control object D-6). That is why the

estimation of positional accuracy of investigated digital maps was based on statistical definition of probability where the probability of an event is identified with the relative frequency of this event. Obtained results of accuracy estimation using statistical analysis are coherent with conclusions developed on the basis of classical estimation of accuracy (Dąbrowski and Doskocz, 2004).

Statistical analysis confirmed high accuracy of the digital map produced on the basis of survey with a total station (object A). Probability that the theoretical error (0.3 mm at the map scale) will be exceeded in the object A does not exceed 5% (Table 8, col. 10).

The probability of occurrence of the point location error which does not meet the accuracy standards in the case of a digital map produced on the basis of the past field surveys is relatively low (object B); in general it does not exceed 10% (except control objects B-3 and B-7) (Table 9, col. 10).

Statistical analysis confirmed the high accuracy (in relation to well identified details of the 1st group, i.e. characteristic points of technical utilities of a digital orthophotomap (object C). In the case of that object the probability of occurrence of errors larger than the theoretical error does not exceed 3% (Table 8, col. 10).

The lowest accuracy was confirmed for a digital map produced by means of graphical-and-digital processing of analogue maps (object D). The estimated probability of occurrence of the point location error that exceeds the theoretical error value (0.3 mm at the map scale) within the object D equals to 20÷30% (Table 8, col. 10).

The more detailed discussion of results of statistical analysis of sets of true errors ε_x , ε_y was presented in (Doskocz, 2002). That elaboration also contains an attempt to confirm estimated positional accuracy of investigated digital maps. However, precise specification of factors that could explain the stated accuracies requires the detailed investigations of genealogy of particular maps (in the case of objects B and D covering the period of the past 30 years).

6. Conclusions

Results of performed investigations allow to draw the following conclusions:

- 1) Large-scale digital maps based on data acquired from various methods do not always meet the requirements of accuracy standards specified in Technical Guidelines K-1.2.
- 2) In the case of data collected as the result of field measurements (objects A and B) or contemporary photogrammetric works (object C) the probability of exceeding the theoretical values of the point location error (0.3 mm at the map scale) is not larger than 5%–10%.
- 3) In the case of digital map data produced by means of graphical and digital processing of analogue maps (object D) the probability of occurrence of the point location error which exceeds the value of the theoretical error equals to 20%–30%.
- 4) There is a need for accuracy evaluation of large-scale digital maps to ensure an appropriate quality of the state surveying resources.

Acknowledgments

This work is the result of own research under scientific leadership of Prof. Wł. Dąbrowski within the frames of the Project No.522-0309-0213 "Creation and updating of databases of digital maps for the needs of surveying administration and for educational purposes".

The author wishes to thank to Prof. L.W. Baran from the Institute of Geodesy of the University of Warmia and Mazury in Olsztyn for fruitful discussions and kind suggestions.

References

- Adamczewski Z., (1961): *Hipoteza falowej struktury błędów*, Geodezja i Kartografia, T. X, Z. 2, pp. 151-159.
- Augustynowicz A., (1993): *Opracowanie mapy numerycznej fragmentu Elbląga z wykorzystaniem przetwarzania graficzno-numerycznego w świetle wyników pomiarów bezpośrednich*, MSc Thesis, University of Warmia and Mazury in Olsztyn (112 pp.).
- Baran L.W., (1983): *Teoretyczne podstawy opracowania wyników pomiarów geodezyjnych*, PWN, Warsaw.
- Bolstad P.V., Gessler P., Lillesand T.M., (1990): *Positional uncertainty in manually digitized map data*, Int. J. Geogr. Inf. Systems, Vol. 4, No 4, pp. 399-412.
- Choi W., Lee M., Sim W., (2001): *Cadastral maps – how to make digital from graphical*, International Conference "New technology for a new century", FIG Working Week 2001, Seoul, Korea, 6-11 May 2001. <http://www.fig.net/pub/proceedings/korea/full-papers/pdf/session7/choi-lee-sim.pdf> (10 pp.).
- Cramér H., (1958): *Metody matematyczne w statystyce*, PWN, Warsaw.
- Dąbrowski W., Dąbrowska D., Doskocz A., Lubarski J., (1999): *Czy numerycznie znaczy dokładniej?*, Magazyn Geoinformacyjny GEODETA, Nr 4 (47), pp. 26-27, <http://www.atomnet.pl/~geodeta/1999/47text2.htm>.
- Dąbrowski W., Doskocz A., (2000): *Badanie dokładności mapy numerycznej miasta Zielona Góra*, in: J. Gaździcki, E. Musiał (eds.), *Systemy Informacji Przestrzennej*, X Konferencja Naukowo-Techniczna, Wydawnictwo Wieś Jutra, Warsaw, pp. 365-371.
- Dąbrowski W., Doskocz A., (2004): *Ocena dokładności położenia szczegółów sytuacyjnych pozyskanych z map numerycznych*, XI Sesja Naukowo-Techniczna "Aktualne problemy naukowe i techniczne prac geodezyjnych", Olsztyn, Poland, 16-17 September 2004, CD-ROM, (9 pp.).
- Doskocz A., (2002): *Badanie dokładności wielkoskalowych map numerycznych wykonanych różnymi metodami*, PhD Dissertation, University of Warmia and Mazury in Olsztyn (174 pp.).
- Doskocz A., (2003): *Ocena dokładności ortofotomapy cyfrowej*, Przegląd Geodezyjny, Nr 4, pp. 9-11.
- Gaździcki J., (1995): *Systemy katastralne*, PPWK, Warsaw-Wrocław.
- Hellwig Z., (1998): *Elementy rachunku prawdopodobieństwa i statystyki matematycznej*, PWN, Warsaw.
- Husár K., (1996): *Presnost' digitálnych priestorových údajov*, Kartografické Listy, No 4, pp. 69-78.
- Latoś S., Maślanka J., (1998): *Analiza dokładności map numerycznych i cyfrowych*, Zeszyty Naukowe AGH, Seria Geodezja, T. 4, Z. 2, pp. 163-177.
- Łyszkowicz A., (1975a): *Określenie typów krzywych Pearson'a aproksymujących zbiory błędów pomiarów geodezyjnych*, Geodezja i Kartografia, T. XXIV, Z. 4, pp. 275-282.
- Łyszkowicz A., (1975b): *Wybór matematycznych modeli dla rozkładów empirycznych wyników pomiarów geodezyjnych*, PhD Dissertation, University of Warmia and Mazury in Olsztyn (84 pp.).
- Ney B., (1970): *Kryteria zgodności rozkładów empirycznych z modelami*, Polska Akademia Nauk – Oddział w Krakowie, Prace Komisji Górniczo-Geodezyjnej, Seria Geodezja, Nr 7, pp. 31-46.
- Ney B., (1976): *Metody statystyczne w geodezji*, AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Nr 497, Kraków.
- Podobnikar T., (1999): *Modelling and Visualisation of Spatial Data Error. Monte Carlo Simulations in Slovenia*, GIM International, Vol. 13, No 7, pp. 47-49.
- Romanowski M., (1974): *Teoria modulacji przypadkowych błędów spostrzeżeń*, Polska Akademia Nauk – Oddział w Krakowie, Prace Komisji Górniczo-Geodezyjnej, Seria Geodezja, Nr 18, pp. 19-46.
- Romanowski M., (1979): *Random Errors in Observations and the Influence of Modulation on their Distribution*, Verlag Konrad Wittwer, Stuttgart.

- Salzmann M., Hoekstra A., Schut T., (1998): *Cadastral map renovation: a Dutch perspective*, The XXI International Congress of Surveyors FIG'98, Brighton, England, 19-25 July 1998, <http://www.sli.unimelb.edu.au/fig7/Brighton98/Comm7Papers/SS34-Salzmann.html> (8 pp.).
- Smirnow N.W., Dunin-Barkowski I.W., (1969): *Kurs rachunku prawdopodobieństwa i statystyki matematycznej dla zastosowań technicznych*, PWN, Warsaw.
- Szacherska M.K., (1974): *Model kompozycji błędów pomiarów geodezyjnych*, Geodezja i Kartografia, T. XXIII, Z. 1, pp. 21-51.
- Szacherska M.K., Kurpiewska A., (1976): *Dokładność niwelacji precyzyjnej w świetle badań statystycznych*, Geodezja i Kartografia, T. XXV, Z. 1, pp. 17-33.
- Urbański J., (1997): *Zrozumieć GIS. Analiza informacji przestrzennej*, PWN, Warsaw.
- Wiśniewski Z., (1984): *Zastosowanie rozkładów Pearsona typu II i VII w wyrównaniu sieci geodezyjnych*, Geodezja i Kartografia, T. XXXIII, Z. 3, pp. 85-104.
- Wiśniewski Z., (1986): *Wyrównanie sieci geodezyjnych z zastosowaniem probabilistycznych modeli błędów pomiaru*, Acta Acad. Agricult. Techn. Olst. Geodaesia et Ruris Regulatio, No 15, Supplementum C., University of Warmia and Mazury in Olsztyn (105 pp.).

Legal regulations and sectoral standards

- Instruction, (1965): *Instrukcja B-III, Poligonizacja techniczna (B-III Instruction to Technical Traversing)*, Head Office of Geodesy and Cartography in Poland.
- Instruction, (1983): *Instrukcja techniczna G-4, Pomiary sytuacyjne i wysokościowe (G-4 Instruction. Measures of locations and altitudes)*, Head Office of Geodesy and Cartography in Poland.
- Instruction, (1998): *Instrukcja techniczna K-1, Mapa zasadnicza (Technical Instruction K-1. The base map)*, Head Office of Geodesy and Cartography in Poland.
- Decree, (2001): *Rozporządzenie Ministra Rozwoju Regionalnego i Budownictwa z dnia 29 marca 2001 roku w sprawie ewidencji gruntów i budynków (Decree of the Minister of Regional Development and Building Industry dated 29 March 2001 on register of lands and buildings)*, Dz. U. nr 38. poz. 454, Polish Law's Server <http://bap-psp.lex.pl/cgi-bin/demo.cgi?id=&comm=jn&akt=nr16896641&ver=-1&jedn=-1>.
- Guidelines, (1981): *Wytyczne techniczne K-1.2, Mapa zasadnicza – aktualizacja i modernizacja (Technical Guidelines K-1.2. The base map – updating and modernization)*, Head Office of Geodesy and Cartography in Poland.

Wykorzystanie analizy statystycznej do oceny dokładności opracowania sytuacyjnego wielkoskalowych map numerycznych

Adam Duskocz

Uniwersytet Warmińsko-Mazurski w Olsztynie
Katedra Geodezji Szczegółowej
ul. Heweliusza 12, 10-724 Olsztyn
e-mail: adam.duskocz@uwm.edu.pl

Streszczenie

W pracy przedstawiono ocenę dokładności opracowania sytuacyjnego map numerycznych poprzez analizę statystyczną. Badania zrealizowano na przykładzie czterech wielkoskalowych map numerycznych wykonanych w oparciu o różne metody pozyskiwania danych numerycznych: nowy pomiar tachimetrem elektronicznym (obiekt A), przeliczenie wyników wcześniejszych pomiarów bezpośrednich (ortogonalnych i biegunowych) (obiekt B), manualną wektoryzację rastrowego obrazu ortofotomapy (obiekt C) oraz przetworzenie graficzno-numeryczne map analogowych (obiekt D).

Analizę wykonano na dużych próbach statystycznych zbiorów długości wektorów przesunięcia punktów kontrolnych ε_L oraz jego składowych, tj. błędach prawdziwych przyrostów współrzędnych $\varepsilon_x, \varepsilon_y$. W przypadku mapy wykonanej z nowych pomiarów tachimetrem elektronicznym jako błędy prawdziwe przyjęto różnice dwukrotnie wyznaczonych współrzędnych punktów kontrolnych. Natomiast w odniesieniu do pozostałych metod pozyskiwania danych do tworzenia map numerycznych, błędy prawdziwe uzyskano z różnic współrzędnych pozyskanych z badanej mapy i współrzędnych wyznaczonych z nowego pomiaru bezpośredniego.