UDK 528.06+528.1

# SOME RESULTS OF RESEARCHES OF CORRELATION COEFFICIENT INFLUENCE OF COORDINATES IN THE CALCULATION OF THE MEAN SQUARE ERRORS OF LAND PARCELS AREAS

**V.A. Riabchii, V.V. Riabchii*, M.V. Trehub, A.G. Sovgirenko**
State Higher Institution «National Mining University »
* Lviv Polytechnic National University

**Key words**: correlation coefficient, mean square error, land parcel, boundaries marks, partial derivatives.

**Statement of the problem**

Today, the definition of areas of land parcels according to the results of geodetic measurements is important for surveyors, land managers and all owners or users of land. The most important parameter is the accuracy of the area of land parcels because it concerns the financial aspects.

**Analysis of recent research and publications related to the solution of this problem**

In scientific publications [8, 9 and others.] studied the determination of the mean square errors of areas of land parcels but it is not taken into account the correlation of boundaries marks coordinates. And there are valid reasons: a small number of coordinates and the inability to reliably determine the correlation coefficient among the calculated coordinates.

In studies of geodetic measurements is widely used regression analysis due to significant achievements of famous scientists: V.D. Bolshakov, M.G. Viduyev, S.P. Voytenko, G.S. Kondra, Y.S. Marcuse, K.K. Skydanenka, A.S. Chebotarev etc. But it Viduyev M.G. and Kondra G.S. precisely interpreted the differences in the statistics of the average and the average of geodetic measurements [3].

In many scientific publications, textbooks and textbooks on mathematical processing of the results of measurements [1 - 7, 10, 11, and others.] is defined the formula of calculation the mean square error of function of uncorrelated and correlated arguments.

The question is how the neglecting of the correlation coefficient changes the values of the mean square error of the land parcels area, and right or wrong not to take into account the correlation coefficients.

**Unresolved parts of the general problem**

Consider the applying of formulas to determine the mean square error of the area. Here is the formula of Gauss for calculating the area of land, such as the following:

$$S = \frac{1}{2}\sum_{i=1}^{n} x_i (y_{i+1} - y_{i-1}) . \qquad (1)$$

According to formula (1) the mean square error of the land parcel area in the case of not correlated and correlated coordinates in general would be the following:

$$m_s = \sqrt{\left(\frac{\partial S}{\partial x_1}\right)^2 m_{x_1}{}^2 + \left(\frac{\partial S}{\partial x_2}\right)^2 m_{x_2}{}^2 + \ldots + \left(\frac{\partial S}{\partial x_n}\right)^2 m_{x_n}{}^2 + \left(\frac{\partial S}{\partial y_1}\right)^2 m_{y_1}{}^2 + \left(\frac{\partial S}{\partial y_2}\right)^2 m_{y_2}{}^2 + \ldots + \left(\frac{\partial S}{\partial y_n}\right)^2 m_{y_n}{}^2} , \qquad (2)$$

$$m_s = \sqrt{\begin{array}{l}\left(\frac{\partial S}{\partial x_1}\right)^2 m_{x_1}{}^2 + \left(\frac{\partial S}{\partial x_2}\right)^2 m_{x_2}{}^2 + \ldots + \left(\frac{\partial S}{\partial x_n}\right)^2 m_{x_n}{}^2 + \\ + \left(\frac{\partial S}{\partial y_1}\right)^2 m_{y_1}{}^2 + \left(\frac{\partial S}{\partial y_2}\right)^2 m_{y_2}{}^2 + \ldots + \left(\frac{\partial S}{\partial y_n}\right)^2 m_{y_n}{}^2 + \\ + 2r_{x_1,x_2} m_{x_1} m_{x_2} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial x_2} + \ldots + 2r_{x_{n-1},x_n} m_{x_{n-1}} m_{x_n} \frac{\partial S}{\partial x_{n-1}}\frac{\partial S}{\partial x_n} + \\ + 2r_{y_1,y_2} m_{y_1} m_{y_2} \frac{\partial S}{\partial y_1}\frac{\partial S}{\partial y_2} + \ldots + 2r_{y_{n-1},y_n} m_{y_{n-1}} m_{y_n} \frac{\partial S}{\partial y_{n-1}}\frac{\partial S}{\partial y_n} + \\ + 2r_{x_1,y_1} m_{x_1} m_{y_1} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial y_1} + \ldots + 2r_{x_n,y_n} m_{x_n} m_{y_n} \frac{\partial S}{\partial x_n}\frac{\partial S}{\partial y_n}\end{array}} , \qquad (3)$$

$\left(\frac{\partial S}{\partial x}\right)$, $\left(\frac{\partial S}{\partial y}\right)$ – partial derivative by the abscissa and ordinate of points of boundaries marks; $m_x$, $m_y$ – mean square errors of the abscissa and ordinate; $r_{x,y}$ – correlation coefficients between the coordinates.

In the formula (3) used the correlation coefficient that describes the closeness of correlation in this case between the coordinates of points of boundaries marks of land parcels.

**The aim of the article**

According to the above defined problems the aim of the article is the following: to establish the values of the correlation coefficient influence on the definition of mean square errors of the areas, if it could be determined.

**Main part**

Correlation coefficient was established to define the interrelation between different statistic values. The example is the following: height and thickness (diameter) of any tree species, height of a person and the size of the feet etc. According to such analysis with some probability economic or social forecasts can be made.

The number of measurements in the statistics, of course, quite significant. When performing ordinary geodetic works for the land management the number of measurements on the contrary – not significant, especially considering the final result – the coordinates of boundaries marks of land parcels.

The interdependence of measurements is very complex and depends on many factors and measurement conditions (weather, instrumental, human, etc.). This list and its detail can be continued for a long time and correctly noted in [3] that the important thing is to determine the conditions of measurement, but it is a never ending process. In statistics, mainly to determine the correlation coefficient used the final results, but the conditions of measurement and accuracy of obtaining the source data is not considered.

Determination of the correlation coefficient of the results of geodetic measurements or coordinates of boundaries marks of land parcels can be defined in two ways. The first – to study measuring conditions and take into account all the factors that affect the measurements and the final result, as well as performed mathematical treatment results. If you go this route, you would not be able to stop and take into account all factors of influence on present understanding of this issue will be difficult or even impossible. As every reader (researcher) from your experience can add a factor that was not taken into account.

The second way to determine the coefficient of correlation – is to define only dependences of outcome values – coordinates. If you go this route, there is a major problem – the number of final results (coordinates of boundaries marks of land parcels). This raises the question: "Is it correct in general examine calculate and determine the correlation coefficient between the coordinates of the boundaries marks?" But is true or not, it must be proven. And here remember the phrase of A.S. Chebotarev: "Requires to compare the yield of wheat in Orlovska region to a number of years and the number of infants begotten in Patahonyy during the same years" [11]. No comments.

It is known that the mean square error of simple arithmetic middle reliably determined by the number of measurements at least eight [1 et al.]. But simple arithmetic formula is applied also to define the middle of two measured values.

Therefore, we consider the influence of the correlation coefficient on the mean square error with a minimum number of the boundaries marks of land parcel. In general, the minimum number is four boundaries marks (Fig. 1).

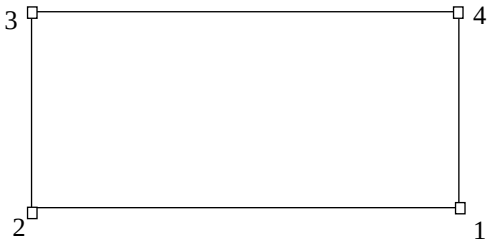


*Fig. 1 Land parcel with four boundaries marks*

The calculation of area for the land parcel with four boundaries marks can be made using the formula (1).

$$S = \frac{1}{2}(x_1(y_2 - y_4) + x_2(y_3 - y_1) + x_3(y_4 - y_2) + x_4(y_1 - y_3)). \quad (4)$$

Then according to formula (3) the formula for the mean square error of the area with correlated coordinates of the boundaries marks by the formula (4) can be written as follows:

$$m_S = \sqrt{\left(\frac{\partial S}{\partial x_1}\right)^2 m_{x_1}^2 + \left(\frac{\partial S}{\partial x_2}\right)^2 m_{x_2}^2 + \left(\frac{\partial S}{\partial x_3}\right)^2 m_{x_3}^2 + \left(\frac{\partial S}{\partial x_4}\right)^2 m_{x_4}^2 +$$

$$+ \left(\frac{\partial S}{\partial y_1}\right)^2 m_{y_1}^2 + \left(\frac{\partial S}{\partial y_2}\right)^2 m_{y_2}^2 + \left(\frac{\partial S}{\partial y_3}\right)^2 m_{y_3}^2 + \left(\frac{\partial S}{\partial y_4}\right)^2 m_{y_4}^2 +$$

$$+ 2r_{x_1x_2} m_{x_1} m_{x_2} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial x_2} + 2r_{x_1x_3} m_{x_1} m_{x_3} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial x_3} +$$

$$+ 2r_{x_1x_4} m_{x_1} m_{x_4} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial x_4} + 2r_{x_2x_3} m_{x_2} m_{x_3} \frac{\partial S}{\partial x_2}\frac{\partial S}{\partial x_3} +$$

$$+ 2r_{x_2x_4} m_{x_2} m_{x_4} \frac{\partial S}{\partial x_2}\frac{\partial S}{\partial x_4} + 2r_{x_3x_4} m_{x_3} m_{x_4} \frac{\partial S}{\partial x_3}\frac{\partial S}{\partial x_4} +$$

$$+ 2r_{y_1y_2} m_{y_1} m_{y_2} \frac{\partial S}{\partial y_1}\frac{\partial S}{\partial y_2} + 2r_{y_1y_3} m_{y_1} m_{y_3} \frac{\partial S}{\partial y_1}\frac{\partial S}{\partial y_3} +$$

$$+ 2r_{y_1y_4} m_{y_1} m_{y_4} \frac{\partial S}{\partial y_1}\frac{\partial S}{\partial y_4} + 2r_{y_2y_3} m_{y_2} m_{y_3} \frac{\partial S}{\partial y_2}\frac{\partial S}{\partial y_3} +$$

$$+ 2r_{y_2y_4} m_{y_2} m_{y_4} \frac{\partial S}{\partial y_2}\frac{\partial S}{\partial y_4} + 2r_{y_3y_4} m_{y_3} m_{y_4} \frac{\partial S}{\partial y_3}\frac{\partial S}{\partial y_4} +$$

$$+ 2r_{x_1y_1} m_{x_1} m_{y_1} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial y_1} + 2r_{x_1y_2} m_{x_1} m_{y_2} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial y_2} +$$

$$+ 2r_{x_1y_3} m_{x_1} m_{y_3} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial y_3} + 2r_{x_1y_4} m_{x_1} m_{y_4} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial y_4} +$$

$$+ 2r_{x_2y_1} m_{x_2} m_{y_1} \frac{\partial S}{\partial x_1}\frac{\partial S}{\partial y_1} + 2r_{x_2y_2} m_{x_2} m_{y_2} \frac{\partial S}{\partial x_2}\frac{\partial S}{\partial y_2} +$$

$$+ 2r_{x_2y_3} m_{x_2} m_{y_3} \frac{\partial S}{\partial x_2}\frac{\partial S}{\partial y_3} + 2r_{x_2y_4} m_{x_2} m_{y_4} \frac{\partial S}{\partial x_2}\frac{\partial S}{\partial y_4} +$$

$$+ 2r_{x_3y_1} m_{x_3} m_{y_1} \frac{\partial S}{\partial x_3}\frac{\partial S}{\partial y_1} + 2r_{x_3y_2} m_{x_3} m_{y_2} \frac{\partial S}{\partial x_3}\frac{\partial S}{\partial y_2} +$$

$$+2r_{x_3y_3}m_{x_3}m_{y_3}\frac{\partial S}{\partial x_3}\frac{\partial S}{\partial y_3}+2r_{x_3y_4}m_{x_3}m_{y_4}\frac{\partial S}{\partial x_3}\frac{\partial S}{\partial y_4}+$$

$$+2r_{x_4y_1}m_{x_4}m_{y_1}\frac{\partial S}{\partial x_4}\frac{\partial S}{\partial y_1}+2r_{x_4y_2}m_{x_4}m_{y_2}\frac{\partial S}{\partial x_4}\frac{\partial S}{\partial y_2}+$$

$$+2r_{x_4y_3}m_{x_4}m_{y_3}\frac{\partial S}{\partial x_4}\frac{\partial S}{\partial y_3}+2r_{x_4y_4}m_{x_4}m_{y_4}\frac{\partial S}{\partial x_4}\frac{\partial S}{\partial y_4}}. \quad (5)$$

Quite complicated formula. And that's just the four boundaries marks. But this formula can be reduced for further consideration. To do this, consider the formula calculating the mean square error of correlated coordinates (3) and (5). All multiplications of the root of the mean square error are coordinates. Assume that they are equal, i.e., the condition must be true:

$$m_{x_1} = m_{x_2} = \cdots = m_{x_n} = m_{y_1} = m_{y_2} = \cdots m_{y_n} = m_{x,y}. \quad (6)$$

This condition (6) you can accept when the allowable mean square error is calculated in the case of non-correlated and correlated coordinates.

Then the mean square error can be taken from the root, that will allow shorten the expression (5).

Consider the well-known formula for calculating the correlation coefficient, which is as follows:

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y}, \quad (7)$$

$\bar{x}$, $\bar{y}$ – simple arithmetic middle of the variables *x*, *y*, respectively; $\sigma_x$, $\sigma_y$ – standard deviation of the values *x*, *y*, calculated by the formulas:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}, \quad (8)$$

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n}}. \quad (9)$$

We can assume that to set the correlation coefficient for each pair of the coordinates is not possible and especially because there are only two values of the coordinates (only one pair). And for a specific group or for all coordinates is possible. Assuming the principle of the medium, we can assume that all correlation coefficients are equal and they can be taken out of the brackets of doubled multiplications and the overall correlation coefficient can be presented without indexes using the letter *r*:

$$r_{x_1,x_2} = \cdots = r_{x_{n-1},x_n} = r_{y_1,y_2} = \cdots = r_{y_{n-1},y_n} = r_{x_1,y_1} = \cdots = r_{x_n,y_n} = r. \quad (10)$$

In the formulas (3) and (5) are defined the following groups of amounts. The first – the squares of the partial derivatives of abscissa of boundaries marks. Their number equals the number of boundaries marks. Second – the squares of partial derivatives by the ordinate of boundaries marks. Their number is also equal to the number of boundaries marks. Third – doubled multiplications of partial derivatives abscissa together. Fourth – doubled multiplications of partial derivatives ordinate with each other. Fifth – doubled multiplications of partial derivatives of abscissas and ordinates together. Number of multiplications doubled the third and fourth groups of $n_3$ and $n_4$ can be calculated using the formula:

$$n_3 = n_4 = \frac{n!}{2(n-2)!}, \quad (11)$$

*n* – number of boundaries marks.

Number of doubled multiplications between abscissas and ordinates $n_5$ of the fifth group will be:

$$n_5 = n^2. \quad (12)$$

General number of doubled multiplications of the third, fourth and fifth group, i.e. of all doubled multiplications can be calculated according to the following formula:

$$n_{3,4,5} = \frac{n!}{(n-2)!} + n^2. \quad (13)$$

General number of all members under the root will be the following:

$$n_{заг} = 2n + \frac{n!}{(n-2)!} + n^2. \quad (14)$$

For example, for any quadrangle $n_{заг} = 36$.

Define the partial derivatives of the square by the abscissa and ordinate using formula (4):

$$\frac{\partial S}{\partial x_1} = \frac{1}{2}(x_2 - y_4); \quad \frac{\partial S}{\partial x_2} = \frac{1}{2}(x_3 - y_1);$$

$$\frac{\partial S}{\partial x_3} = \frac{1}{2}(x_4 - y_2); \quad \frac{\partial S}{\partial x_4} = \frac{1}{2}(x_1 - y_3); \quad (15)$$

$$\frac{\partial S}{\partial y_1} = \frac{1}{2}(x_4 - x_2); \quad \frac{\partial S}{\partial y_2} = \frac{1}{2}(x_1 - x_3);$$

$$\frac{\partial S}{\partial y_3} = \frac{1}{2}(x_2 - x_4); \quad \frac{\partial S}{\partial y_4} = \frac{1}{2}(x_3 - x_1). \quad (16)$$

After substituting the partial derivatives (15) and (16) into (5) and as a result of some changes (to reduce the amount of paper they are presented) determined that doubling sum of multiplications of the fifth group will be zero, provided that the mean square error of coordinates and correlation coefficients are equal in their value. In the third and fourth groups sum of multiplications of partial derivatives by interactions with each other except for the sum derived from indices (*i* + 1) and (*i* – 1) will also be zero.

Further transformation allow to obtain the following formula:

$$m_S = \frac{m_{x,y}}{2}\sqrt{D_{2-4}^2 + D_{3-1}^2 + D_{4-2}^2 + D_{1-3}^2 - r(D_{2-4}^2 + D_{3-1}^2 + D_{4-2}^2 + D_{1-3}^2)}, \quad (17)$$

*D* – distance between the boundaries marks (*i* + 1) i (*i* – 1).

For land parcel with four boundaries marks, this distances are the diagonals of the quadrilateral, i.e.

$$m_S = \frac{m_{x,y}}{2}\sqrt{2\left(D_{2-4}^2 + D_{1-3}^2\right) - 2r\left(D_{2-4}^2 + D_{1-3}^2\right)}. \quad (18)$$

For land parcel with any number of boundaries marks the formula (17) can be written in the general form:

$$m_S = \frac{m_{x,y}}{2}\sqrt{\sum_{i=1}^{n} D_{(i+1)-(i-1)}^2 - r\sum_{i=1}^{n} D_{(i+1)-(i-1)}^2}. \quad (19)$$

Now let us analyze the possible values of correlation coefficients. In this case, the attentive reader has already identified himself options, but we need to give them.

If the correlation coefficient is equal or close to plus one, then in this case under the root is received zero or value close to zero. And when the correlation coefficient is equal to minus one, or close to minus one, then the mean square error of area is twice more or close to doubled value if the correlation coefficient would not be taken into account.

Of course, it's taken extreme cases. Correlation coefficient for each doubled multiplications may be different (individual). But it is possible to choose the average value of the correlation coefficient, which can be used to obtain the same value of mean square error of the area, as well as using the correlation coefficients of each pair of coordinates.

Suppose that the correlation coefficient is equal or close to unity. Then we can define a formula and build the regression line. And what does that give? Forecast of some new boundaries marks? But this is not correct!

According to the results of calculations the correlation coefficient using the formula (7) for different areas and configurations of land parcels was found that the correlation coefficient will affect the orientation of land parcels, relative position of boundaries marks of the boundary, congestion and greatly the coefficient of elongation of land parcel.

Now consider the geometric essence of values that are included in calculation of the correlation coefficient the formula (7). Average values of the abscissa and ordinate – are actually the coordinates of the geometric center (weight center) of land parcel. Differences in the numerator which are multiplied are deviation of abscissa and ordinate of coordinates of the geometric center respectively. In fact they are the increments of coordinates of boundaries marks of the land parcel (vertices of the polygon) relatively to the geometric center and their sum is zero. Value of multiplies are the areas of rectangles of abnormalities that can dock with each other completely or partially (Fig. 2 and 3). Number of rectangles of deviations is equal to the number of boundaries marks of the land parcel. In Fig. 2 and 3, each rectangle of deviation shaded in different directions and different colors.

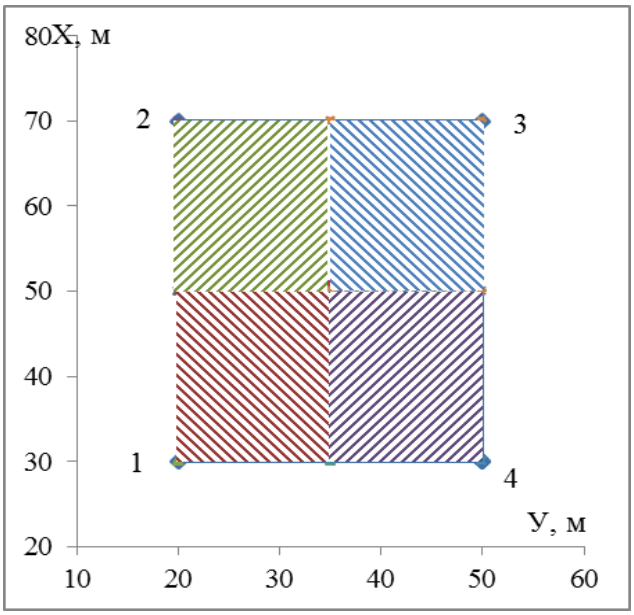


*Fig. 2. Land parcel oriented along the axes x and y*

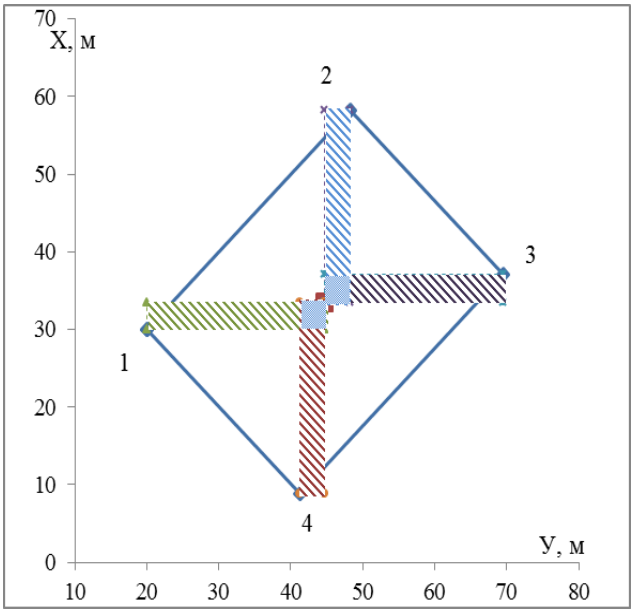


*Fig. 3. Land parcel with random orientation along the axes x and y*

If to sum-up areas of the rectangles, their total area is equal to the area of the parcel (Fig. 2) or will be less (Fig. 3). In fact, the sum of the areas of rectangles defines the value and the sign of the correlation coefficient. If the land parcel is the correct rectangle and its sides are parallel to the axes, these rectangles are not extend beyond the boundaries of land parcel.

Consider the denominator of the formula (7). The multiply of the mean square of deviations on axes of coordinates is also the area of the rectangle. It can be called the mean square rectangle. A multiply of the number of boundaries marks of the land parcel and the area of square mean rectangle gives a square mean area of land parcel.

Now suppose that the sum of the areas of rectangles in the numerator is not zero. Then the correlation

coefficient shows the ratio of the areas sum of rectangles of the numerator to mean square area in the denominator. Of course, this ratio is less than unity and may have a sign as a minus as a plus.

In addition to this interpretation of the ratio of values the numerator and denominator in equation (7), can be considered the following. If the sum of the areas of rectangles of the numerator divided by the number of boundaries marks, we obtain the average value of the area of the rectangle. Then we obtain ratio of the area of the rectangle to the area of average mean square rectangle. But it more refers not to determining the correlation of coordinates but to the characterization of the relative position of boundaries marks and configuration of land parcel. Thus the sum of the multiplies of differences between the coordinates of boundaries marks and the coordinates of the geometric center can be taken without regard to signs.

Consider the proof of the formula calculating the mean square error of function of the correlated arguments. In order to reduce the term (as it is known to most readers) we will present only that part of the formula, where appears the correlation coefficient.

General form of any function $u$ is the following:

$$u = f(x, y, z), \tag{20}$$

$x$, $y$, $z$ – arguments of the function.

The true error of the function $\Delta_u$ is

$$\Delta_u = \frac{\partial u}{\partial x}\Delta_x + \frac{\partial u}{\partial y}\Delta_y + \frac{\partial u}{\partial z}\Delta_z, \tag{21}$$

$\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}$ – partial derivatives of the function by the arguments; $\Delta_x$, $\Delta_y$, $\Delta_z$ – true errors of arguments $x$, $y$, $z$.

Assuming that the function (21) was calculated $n$ times and partial derivatives for each argument are equal to each other, after the raising to the square and summing, we can write:

$$[\Delta_u\Delta_u] = \left(\frac{\partial u}{\partial x}\right)^2 [\Delta_x\Delta_x] + \left(\frac{\partial u}{\partial y}\right)^2 [\Delta_y\Delta_y] + \left(\frac{\partial u}{\partial z}\right)^2 [\Delta_z\Delta_z] + 2\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}[\Delta_x\Delta_y] + \ldots \tag{22}$$

Dividing all the members of the last equation by the number of measured arguments and considering the Gauss formula, equation (22) will be the following:

$$m_u^2 = \left(\frac{\partial u}{\partial x}\right)^2 m_x^2 + \left(\frac{\partial u}{\partial y}\right)^2 m_y^2 + \left(\frac{\partial u}{\partial z}\right)^2 m_z^2 + 2\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\frac{[\Delta_x\Delta_y]}{n} + \ldots \tag{23}$$

Then, in addition to the formula of calculating the correlation coefficient, the following assumptions are made:

$$\Delta_x = x_i - X \cong x_i - \bar{x},$$
$$\Delta_y = y_i - Y \cong y_i - \bar{y}, \tag{24}$$
$$\Delta_z = z_i - Z \cong z_i - \bar{z}$$

and also:

$$m_x = \sqrt{\frac{\Delta_x\Delta_x}{n}} \cong \sigma_x = \sqrt{\frac{(x_i - \bar{x})^2}{n}},$$
$$m_y = \sqrt{\frac{\Delta_y\Delta_y}{n}} \cong \sigma_y = \sqrt{\frac{(y_i - \bar{y})^2}{n}}, \tag{25}$$
$$m_z = \sqrt{\frac{\Delta_z\Delta_z}{n}} \cong \sigma_z = \sqrt{\frac{(z_i - \bar{z})^2}{n}}.$$

Considering assumptions (24) i (25) we can get the equation (3).

If the deviations from the mean values of these arguments are not significant, then this assumption is possible. But to equate the deflection of abscissa and ordinate of coordinates of land parcels boundaries marks from their average values (coordinates of the geometric center) to errors of these coordinates taking into consideration their values is incorrect. Since the average value of the abscissa or ordinate (coordinates of the geometric center) is not the value that is closest to the true value of the boundaries marks. Note that these proofs and arguments concerning only the above case.

All that above considered under the assumption that condition (6) is true, i.e. the values of abscissa and ordinate of the boundaries marks of land parcel are defined with equal precision. But in most cases the values of the mean square errors of coordinates are different and this should be taken into account. To define correlation coefficient, in the case of unequal precision of values of abscissa and ordinate must be considered their weight.

As you know, in addition to the classical definition of measurement scales is also important the following expression: these are the variables (coefficients) that allow you to perform common mathematical operations with the unequal precision results of measurements [7].

If you in general use the correlation coefficient and calculate it using coordinates, then you must use the general formula of arithmetic middle and the corresponding deviation. Here the question may arise: "In general, if the standard deviations in equation (7) are different in precision, how it should be taken into account?". It can be the following cases:

1) all values $x$ and $y$ are defined with the same precision;

2) values $x$ and $y$ each separately identified with equal precision but among themselves – not;

3) values $x$ and $y$ identified with equal precision among themselves but separately – not;

4) all values $x$ and $y$ are identified not equal.

If it is so, you have to define the weight of the obtained values $x$ and $y$. Then the formula for calculating the correlation coefficient in general is as follows:

$$r_{x,y} = \frac{\sum_{i=1}^{n} \sqrt{P_{x_i}}(x_i - \bar{x}_з)\sqrt{P_{y_i}}(y_i - \bar{y}_з)}{n\sigma_x\sigma_y}, \qquad (26)$$

$\sigma_x$ and $\sigma_y$ – are calculated using the equation of general arithmetic mean:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n} P_{x_i}(x_i - \bar{x}_з)^2}{n}}, \qquad (27)$$

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{n} P_{y_i}(y_i - \bar{y}_з)^2}{n}}. \qquad (28)$$

The authors in [5] write: "The most important application of correlation analysis is solving a problem of scientific predictions...". But what prediction that can be done if the correlation coefficient for the calculation results will not be zero? It's a rhetorical question, the answer is probably: "No predictions".

Therefore, we can assume that in the calculation of the mean square errors of areas of land parcels is not required to take into account the doubled multiplies and the correlation coefficients. An example of this approach is the formula of Gauss to compute the mean square errors of measurement for the true errors. Gauss refused to take into account the doubled multiplies.

But this does not mean that the correlation coefficient is not necessary to consider otherwise. In any case, it is necessary to solve individually and the studies of our eminent scientists and researchers show that in many cases the determination of the coefficient of correlation and its consideration is important for the forecasting, development of methods, improving efficiency of accuracy estimation and so on.

**Conclusions and suggestions**

1. Since the sum of squares and doubled multiplies of partial derivatives of the coordinates by the areas in (3) is zero (excluding the values of mean square errors of abscissa and ordinate and the corresponding correlation coefficients), the study of correlation coefficient, which can be used for the calculation of permissible mean square errors of area, today is impossible.

2. To determine the values for correlation coefficient of abscissa and ordinate of coordinates of boundaries marks of land parcel is inappropriate and incorrect, because its values will depend on the orientation of the land parcel, its configuration, congestion of boundaries marks and the elongation coefficient.

3. Considering two conclusions above we can assume that for all calculations of fact and permissible mean square error of land parcels areas should be used the equation (2).

**References**


1. Bolshakov V.D. Teoryia oshybok nabliudenyi: uchebnyk dlia vuzov. – 2-e yzd., pererab. y dop. / V.D. Bolshakov. – M.: Nedra, 1983. – 223 p.
2. Vyduev N.H. Veroiatnostno-statystycheskyi analyz pohreshnostei yzmerenyi / N.H. Vyduev, H.S. Kondra. – M.: Nedra, 1969. – 320 p.
3. Voitenko S.P. Matematychna obrobka heodezychnykh vymiriv. Teoriia pokhybok vymiriv: navch. posibnyk / S.P. Voitenko. – K.: – KNUBA, 2003. – 216 p.
4. Haidaev P.A. Teoryia matematycheskoi obrabotky heodezycheskyi yzmerenyi / P.A. Haidaev, V.D. Bolshakov. – M.: Nedra, 1969. – 400 p.
5. Zazuliak P.M. Osnovy matematychnoho opratsiuvannia heodezychnykh vymiriuvan: navch. posibnyk / P.M. Zazuliak, V.I. Havrysh, E.M. Yevsieieva, M.D. Yosypchuk. – Lviv: Vydavnytstvo «Rastr-7», 2007. – 408 p.
6. Mazmyshvyly A.Y. Teoryia oshybok y metod naymenshykh kvadratov / A.Y. Mazmyshvyly. – M.: Nedra, 1978. – 311 p.
7. Riabchii V.A. Teoriia pokhybok vymiriuvan: navch. posibnyk / V.A. Riabchii, V.V. Riabchii. – D.: Natsionalnyi hirnychyi universytet, 2006. – 166 p.
8. Riabchii V.A. Vyznachennia dopustymykh znachen serednikh kvadratychnykh pokhybok obchyslennia ploshch zemelnykh dilianok u riznykh typakh naselenykh punktiv / V.A. Riabchii, V.V. Riabchii, M.V. Trehub [Text] // Heodeziia, kartohrafiia ta aerofotoznimannia. – 2011. – Vyp. 75. – P. 157-167.
9. Riabchii V.V. Doslidzhennia ta aproksymatsiia funktsii vyznachennia dopustymykh serednikh kvadratychnykh pokhybok ploshch zemelnykh dilianok / V.V. Riabchii, M.V. Trehub [Text] // Heodeziia, kartohrafiia ta aerofotoznimannia. – 2012. – Vyp. 76. – P. 117-126.
10. Smyrnov N.V. Teoryia veroiatnostei y matematycheskaia statystyka v prylozhenyy k heodezyy / N.V. Smyrnov, D.A. Beluhyn. – M.: Nedra, 1969. – 379 p.
11. Chebotarev A.S. Sposob naymenshykh kvadratov s osnovamy teoryy veroiatnostei: Uchebnyk dlia heodezycheskykh vuzov y fakultetov / A.S. Chebotarev. – M.: Heodezyzdat, 1958. – 606 p.


**Some results of researches of correlation coefficient influence of coordinates in the calculation of the mean square errors of land parcels areas**


V.A. Riabchii, V.V. Riabchii,
M.V. Trehub, A.G. Sovgirenko



Defined the problem of determination the coefficient of correlation for the calculation of mean square errors of land parcels areas. Found that nowadays the definition of the correlation coefficient of the boundaries marks coordinates is almost impossible. In addition it's calculation according to the values of coordinates abscissa and ordinate of the boundaries marks of land parcels is inappropriate and incorrect. Made suggestions to ignore the correlation coefficient of coordinates for calculations of the mean square errors of land parcels areas.