Advances in Intelligent Systems Research, volume 174
Proceedings of the 8th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2020)

Accuracy Analysis of Remote Measurement of Thermokarst Lakes Parameters for Field Dynamics Modeling Problems

Matvey Kupriyanov*
Space Services Centre
Ugra Research Institute of Information Technologies
Khanty-Mansiysk, Russia
kupr@uriit.ru

Ildar Muratov
Space Services Centre
Ugra Research Institute of Information Technologies
Khanty-Mansiysk, Russia
ildarmur@gmail.com

Gleb Kochergin
Space Services Centre
Ugra Research Institute of Information Technologies
Khanty-Mansiysk, Russia
kocherginga@uriit.ru

Yury Polishchuk
Space Services Centre
Ugra Research Institute of Information Technologies
Khanty-Mansiysk, Russia
yupolishchuk@gmail.com

***Abstract*—The article is devoted to the issues of evaluating errors of remote measurement of thermokarst lakes characteristics in the Arctic zone using high resolution Kanopus-V images. Remote investigations were carried out at 98 test sites, located in the territory of the cryolithic zone of Western Siberia and north-east of the European part of Russia. To describe the tortuosity of lake boundaries were used the spatial object shape index, accepted in cartography, which is calculated based on the results of lake area and perimeter measurements using satellite images. A histogram of the distribution of the lakes coastal borders tortuosity degree in a wide range of their sizes from 50 m^2 to 20 ha was built. The analysis of the histogram showed that the distribution of tortuosity degree corresponds to the theoretical log-normal law. For investigation of errors of remote measurement of lake parameters the technique of construction of random flat geometrical figures (model objects) with given value of tortuosity degree is offered. Relative error of object area measurement using space image is studied. Influence of object boundaries smoothing on errors of remote sensing of areas and tortuosity of boundaries is studied. It is shown that smoothing practically does not affect the errors of area measurement and has a significant impact on perimeter and tortuosity measurements.**

***Keywords*—satellite images, geoinformation systems, thermokarst lakes, lake area measurement errors, lake boundary smoothing**

I. INTRODUCTION

Due to the ongoing global warming, the problem of modeling the volume of greenhouse gas emissions from the territory of the Arctic zone, as the area of concentration of thermokarst lakes - intensive sources of methane, is relevant. It is known that the mass of methane accumulated in lakes, and hence its emission into the atmosphere, is determined by the spatial properties of lake fields. For the formation of predicted estimates of the dynamics of the accumulation of lake methane, it is necessary to use methods and tools to model the dynamics of the fields of thermokarst lakes.

Recently, numerous remote studies of the dynamics of the number and size of thermokarst lakes in North Eurasia and North America [1-4] have made it possible to create a geo-imitation model of the dynamics of thermokarst lake fields [5], suitable for the formation of predicted estimates of the dynamics of the accumulation of lake methane. The spatial structure of this model is determined by a set of circles with random sizes and random locations [5, 6]. Remote studies of the parameters of real lakes conducted in different permafrost zones (continuous, discontinuous, and sporadic) in Western Siberia have shown [7, 8] that the fields of thermokarst lakes can be considered as a set of random flat figures with significantly different sizes (areas) and with different degrees of tortuosity of the boundaries.

Lakes and other surface water bodies as important sources of greenhouse gas emissions into the atmosphere in permafrost zones are subject to remote studies because of the difficult accessibility and high degree of marshland in the northern territories. Climate warming in recent decades has led to a significant acceleration of thermokarst processes in the permafrost zone, which may cause an increase in methane concentration in thermokarst lakes.

Fig. 1 shows a fragment of the Kanopus-V space image, demonstrating abundance of thermokarst lakes in a relatively small area of the north-east of the European part of Russia.

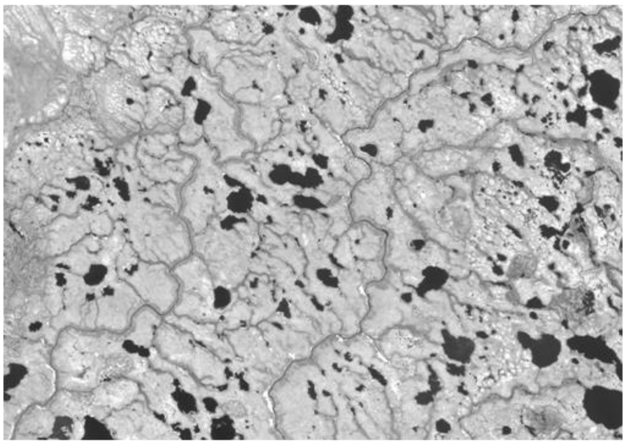

Fig. 1. A fragment of the space image of Kanopus-V.

As can be seen from Fig. 1, almost all lakes have a curved form of coastal boundaries that differs significantly from the circle that was used in the geo-imitation model. Therefore, the problem of evaluating errors in remote sensing of lake areas due to the tortuosity of their coastal boundaries is important.

The issues of evaluating the error of remote measurement of the object area were considered in [9-18]. In [9] an error of area determination using Landsat space images was studied. In [15] the issues of lake area measurement accuracy using optical Landsat images and ERS-2 radar images have been considered, and the empirical dependence of the lake area remote sensing error on its size using optical and radar images of medium spatial resolution has been established. Empirical dependence of the average relative error of the object area estimation on the actually measured area using Landsat 4, 5 and GeoEye data has been established in [16]. In [17] empirical equations of dependence of average and maximum relative error of water body area measurement on the size of water bodies, separated by space image with different spatial resolution, are established. In [18] it is shown that the estimation of object area determination in different thematic tasks of mapping and change monitoring using Landsat 8 data can be done with sufficient accuracy with the area of observed objects over 10 ha.

It should be noted that in [15-17] the error of remote sensing of the object area using space images is estimated by comparing with the size of the same object area, determined from the images of much higher spatial resolution, which is chosen as the true (reference) object area. In our opinion, it is more universal to get an estimate of the error of remote sensing of the object area using space images depending on the ratio of the object area to the value of the image spatial resolution. Such approach would allow estimating the accuracy of area determination from any images, indicating only their spatial resolution. However, in general case the error under consideration depends not only on the ratio between the object area and the image resolution, but also on the shape of the object (tortuosity of its borders), which makes it much more difficult to study the error of measuring the object area using space images.

In tasks of modeling and prediction of greenhouse gas dynamics from thermokarst lakes, caused by climatic changes, there is a necessity to measure geometric characteristics of lakes (area, tortuosity of coastal borders, etc.) using space images. When using high and ultra-high spatial resolution images the tortuosity of the shore borders can influence the measurement of the lake areas. Therefore different smoothing methods are used in practice for such measurements. However, the issues of evaluation of errors in lake parameters measurement, arising from the use of smoothing procedures, have not been studied so far, therefore the relevance of this paper is beyond doubt.

As it has been established in [18] the discrete nature of raster images which consist of pixels leads to considerable errors at measurement of geometrical characteristics of object - perimeter, area, and also derivatives from these characteristics, for example, form factors. For decrease in an error of measurement of geometrical characteristics of objects digitized on raster images, for example, on space images, it is possible to use various methods of smoothing. Comparison of different smoothing methods is a rather laborious task. That is why we have chosen one sufficiently known smoothing method - Savitsky-Goley method, described in [19], the essence of which consists in building by the method of least squares of an approximating polynomial of a given order. However, the issues of evaluation of errors of remote measurements of thermokarst lake parameters from smoothed images have not been considered at present.

In this regard, the purpose of this paper is to model the influence of smoothing the boundaries of objects vectorized on the raster image by the Savitsky-Goley method on the error of measuring the area, perimeter and the coefficient of form (degree of tortuosity of the boundaries) of the object.

II. Methodological issues

The information basis for the experimental study of properties of thermokarst lake fields is the data of remote measurement of areas and perimeters of lakes using satellite images of the investigated territory. The studies were conducted on the territory of all three permafrost zones (sporadic, discontinuous and continuous) in Western Siberia and in the north-east of the European part of Russia using high resolution satellite images. All images were selected within a rather short period of summer season (late June - August) to minimize the impact of seasonal fluctuations in water level in the lakes. During this period the ice cover on the lakes completely disappears.

Experimental study of tortuosity properties of coastlines of real lakes has been conducted. Satellite-based measurements have been used to calculate the values of the lake shape (degree of tortuosity of borders) according to the formula [20]:

$$z = \frac{p^2}{4\pi s}, \quad (1)$$

where s is the area; p is the perimeter of the lake.

The z-value was determined based on the data from Kanopus-V images of 20 TS in the north-east of the European part of Russia (EPR) and 78 test sites of the West Siberian investigated territory. Fig. 1 shows empirical histograms of lake boundaries tortuosity distribution obtained as a result of remote studies in permafrost zone of the European part of Russia and Western Siberia. The z value was calculated in accordance with (1) the results of remote measurements of areas and perimeters.

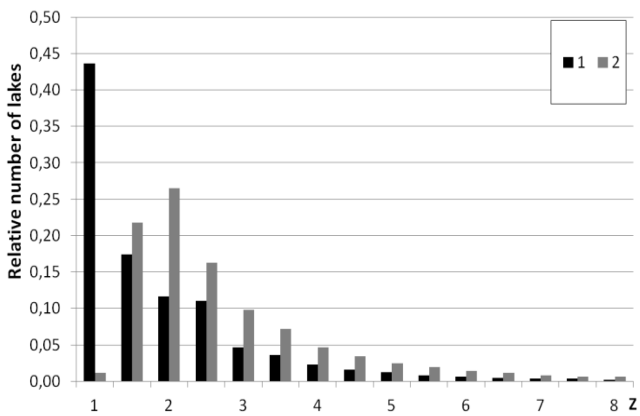

Fig. 2. Histograms of lake boundaries tortuosity distribution based on the Kanopus-V images in the study areas in the European part of Russia (1) and Western Siberia (2)

It is found that the z value changes at different test sites in the range for the most part from 1 to 10 (the value "1" corresponds to a circle). Average z values in different permafrost zones in West Siberian territory take the following values: in the sporadic zone 3.5; in the discontinuous zone 3.38 and in the continuous zone 2.95.

For the tasks of modeling the fields of thermokarst lakes and forecasting the emission of lake methane, knowledge of the laws of distribution of the tortuosity degree of the coastal borders of lakes is important. Fig. 2 shows empirical histograms of the distribution of the degree of tortuosity of the lake boundaries based on the data on the value of z calculated in accordance with (1) on base of the results of remote measurements on 20 TS on the investigated territory of the north-east of the EPR and on 78 TS in Western Siberia.

The histogram graph in Fig. 2 let us to assume that the distribution of the tortuosity degree will correspond to the theoretical log-normal law. To substantiate the possibility of using the lognormal law of the distribution of the tortuosity degree of the lake boundaries, we have checked whether this law corresponds to the empirical histogram. The results of this conformity check according to the data in Western Siberia (according to χ^2 criterion) are given below. According to [21], the probability density for the lognormal law of the tortuosity degree distribution (z) is determined by the equation:

$$f(z) = \frac{1}{\sqrt{2\pi} z\sigma} \exp\left(-\frac{(\ln z - \ln a)^2}{2\sigma^2}\right),$$

where a and σ are the parameters of the lognormal distribution law.

Parameters of the log-normal law of distribution of tortuosity of lake boundaries in Western Siberia were estimated on the basis of sample data. The conformity of the experimentally obtained high spatial resolution images of lake distribution with the theoretical lognormal law was estimated using the standard procedure of checking statistical hypotheses about the type of distribution law [21], which showed that the value of χ2 calculated from empirical data is 5.32. Therefore, it can be assumed that the empirical histogram of the distribution of the tortuosity degree of lake contours in the Arctic cryolithozone with 0.95 probability corresponds to the lognormal law according to Pearson's criterion.

Let us consider methodological issues of analysis of errors in remote measurement of thermokarst lake parameters. For carrying out the analysis we used the sets of model objects (MO), the method of their formation is described in [18].

For modeling of measurement process on raster image of geometrical characteristics of real object having smooth and continuous boundary, rasterization procedure, i.e. transformation of vector data type into raster one, was applied. After rasterization there was a reverse conversion to a vector data type with preservation of distortions introduced at rasterization.

Rasterization was carried out using ArcGIS Desktop software by one of the simplest rasterization methods - "maximum combined area method" described in detail in [22]. The main idea of this rasterization method is that the pixel of the raster image is filled in (drawn) in case if the total area of the polygon contour covers at least 50% of the pixel area, otherwise the pixel is not filled in.

Rasterization was carried out in such a way that after conversion the calculated number of filled pixels would be equal to the specified value N (1, 2, 5, 10, 20, 50, etc. up to 100 000), which was determined by the formula:

$$N = \frac{S}{R^2},$$

where N is the calculated number of pixels in the boundaries of the model object at fixed spatial resolution of the image;

S - area of the model object (MO);

R - the spatial resolution of the image.

III. Results

For creation of sets of the model data the technique of formation of model vector objects [18] has been used. With use of the technique 10 sequences of model objects were generated. Each sample contained 1 000 model objects, the area of each of them was equal to 10 000 m2. The model objects inside one sample had the same value of tortuosity, which was determined by formula (1).

The area of the rasterized MO was determined by the formula:

$$S_{mes} = N_{mes} \times L^2,$$

where S_{mes} is the measured area of the rasterized MO, m2;

N_{mes} - measured number of pixels in the rasterized MO;

L - raster image spatial resolution, m.

For each sample average values of relative error of measurement of the rasterized MO areas according to the formula were calculated:

$$\delta_m = \frac{1}{n}\sum_{i=1}^{n} \delta_i ,$$

where δ_m is the average relative error of measuring the areas of rasterized MO; n - number of MO in the sample; δ_i - relative error of measuring the area of the i-th rasterized MO.

Fig. 3 shows graphs of dependence of δm on N in double logarithmic scale with different values of the degree of tortuosity.

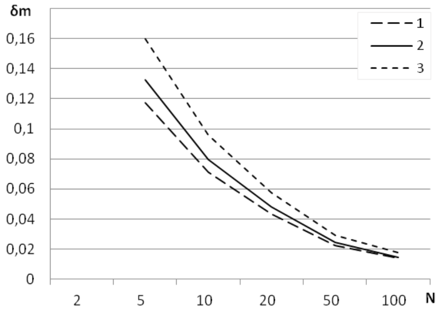

Fig. 3. Dependence of the area measurement error on the number of pixels at different tortuosity values: lower quartile (1), median (2), upper quartile (3)

As you can see from the Fig. 3, the error of remote measurement of area for objects with N > 10 does not exceed 8 % on average.

Let's consider the results of error estimation connected with smoothing of objects' borders. To determine the size of smoothing parameters by Savitsky-Goley method was created a limited sample, which included 10 random model objects (and corresponding to them rasterized representations in the number of 160) with different size z. Further, using the programming language Python, library SciPy [23], was made smoothing of rasterized model objects by Savitsky-Goley method. At the same time, the parameters of the method by complete search (taking into account the limitations of this method) were varied in the following limits: polynomial order (p) from 1 to 11; filtration window size (w) from 5 to 49. After smoothing the rasterized model objects, their geometrical characteristics (perimeter, area, shape coefficient) were determined and recorded in the table for further analysis.

For selection of smoothing parameters p, w the average relative deviation of perimeter measurement error after smoothing by the formula was calculated:

$$\omega_{p,w} = \sum_{i=1}^{m} \left| \frac{L'_i - L^0_i}{L^0_i} \right|,$$

where $\omega_{p,w}$ is the average relative deviation of perimeter measurement error after smoothing of rasterized MO in a limited sample with smoothing parameters p, w;

m – number of objects in the sample;

L'_i – perimeter of the i-th rasterized MO after smoothing;

L^0_i – perimeter of the i-th MO.

The optimal values of Savitsky-Goley smoothing parameters are selected: p = 6, w = 9. Further smoothing was carried out using the specified parameter values, which provide almost the same low value $\omega_{p,w}$, equal to approximately 10 %.

To evaluate the smoothing results, the relative error of measurement of the area and tortuosity of the rasterized object (before smoothing) was calculated by the formula:

$$\delta_X = \left| \frac{X - X^0}{X^0} \right|,$$

where δ_X is the relative deviation of the measurement error of area or tortuosity before smoothing;

X - the area or tortuosity of the object before smoothing;

X^0 - area or tortuosity of the MO.

Then the relative error of measurement of the area and tortuosity of the smoothed object was calculated by the formula:

$$\delta'_X = \left| \frac{X' - X^0}{X^0} \right|,$$

where δ'_X is the relative deviation of the measurement error of the area or tortuosity of the smoothed object;

X' is the area or tortuosity of the smoothed object;

X^0 is the area or tortuosity of the MO.

In fig. 4 and 5 shows the results of calculation of errors of measurement of the degree of tortuosity and area of rasterized and smoothed objects. As can be seen from the figure, with the calculated number of pixels in the measured object N ⩾ 10, the average absolute deviation of the error of measurement of tortuosity becomes positive, with the absolute deviation of the error of measurement of the area fluctuates around zero. This indicates a decrease in the average error of measurement of the tortuosity of objects after smoothing.

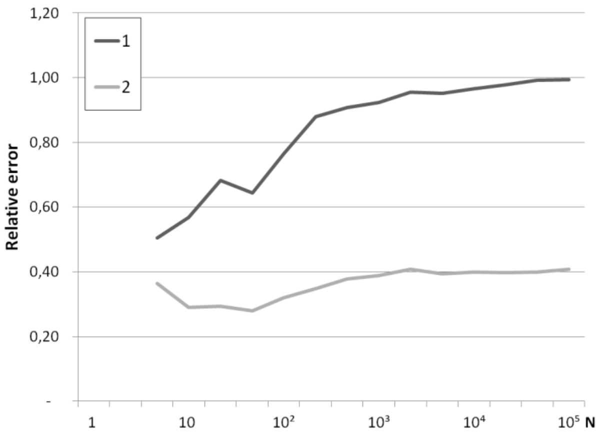

Fig. 4. Relative error in determining the tortuosity of rasterized (1) and smoothed (2) objects

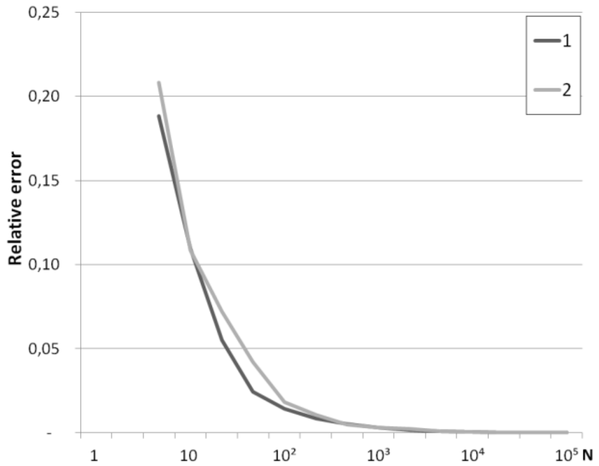

Fig. 5. Relative error in determining the area of rasterized (1) and smoothed (2) objects

IV. CONCLUSION

The results of remote measurements of thermokarst lakes parameters in the Arctic zone in the north-east of the European part of Russia and Western Siberia in a fairly wide range of lake sizes based on space images of high spatial resolution of Kanopus-V are presented. Researches were carried out on 98 test sites of the study area. Images were processed in automatic mode using ArcGIS 10.3 geoinformation system tools.

The obtained experimental data on the degree of tortuosity showed that the degree of tortuosity of the lake boundaries changes mainly from 1 to 10, which requires taking into account the tortuosity when assessing the measurement error of thermokarst lake parameters.

Error analysis was carried out using sequences of model lakes whose parameters corresponded to experimentally measured parameters of real lakes. The discrete nature of raster satellite images of lakes with a significant tortuosity degree of boundaries leads to significant errors in measuring their parameters. Therefore, when processing the results of remote measurements it is necessary to smooth the boundaries of objects. As a result of the model study it is shown that the application of the Savitsky-Goley method with parameters: polynomial order - 6 and filtration window size - 9 for smoothing the boundary of an object whose raster representation consists of 10 or more pixels, on average, reduces the relative error of measurement of the degree of tortuosity by the value from 25 to 55%. The relative error of area measurement does not exceed 2% on average.

ACKNOWLEDGMENT

The work was performed under the state order of the Department of Information Technologies and Digital Development of the Khanty-Mansi Autonomous Okrug - Yugra for 2020.

REFERENCES

[1] S. Kirpotin, Y. Polishchuk, and N. Bryksina, “Abrupt changes of thermokarst lakes in Western Siberia: impacts of climatic warming on permafrost melting,” International Journal of Environmental Studies, vol. 66, no. 4, pp. 423–431, 2009.

[2] M. Luoto, and M. Seppala, “Thermokarst ponds as indicator of the former distribution of palsas in Finnish Lapland,” Permafrost and Periglasial Processes, vol. 14, pp. 19–27, 2003.

[3] B. Riordan, D. Verbyla, and A.D. McGuire, “Shrinking ponds in subarctic Alaska based on 1950-2002 remotely sensed images,” J. Geophys. Res., vol. 111, 2006.

[4] F.S. Zuidhoff, and E. Kolstrup, “Changes in palsa distribution in relation to climate change in Laivadalen, Northern Sweden, especially 1960-1997,” Permafrost and Periglacial Processes, vol. 11, pp. 55–69, 2000.

[5] V.Y. Polishchuk, and Y.M. Polishchuk, “Modeling of thermokarst lake dynamics in West-Siberian permafrost,” In: Permafrost: Distribution, Composition and Impacts on Infrastructure and Ecosystems, vol. VI, O. Pokrovsky, Ed. NY: Nova Science Publishers, 2014, pp. 205-234.

[6] V.Y. Polishchuk, and Y.M. Polishchuk, Geo-simulation modeling fields of thermokarst lakes in zone of permafrost, Khanty-Mansiysk: Yugra State University Press, 2013, 219 p.

[7] V.Y. Polishchuk, and Y.M. Polishchuk, “Remote studies of variability of the shape of coastal boundaries of thermokarst lakes in the permafrost of West Siberia,” Study of Earth from space, vol 1, pp. 61-64, 2012.

[8] Y.M. Polishchuk, I.N.Muratov, and V.Y.Polishchuk, “Study of the tortuosity of the coastal boundaries of the thermokarst lakes of Western Siberia using high-resolution images of Kanopus-V,” Current problems in remote sensing of the Earth from space, vol. 16, no. 5, pp. 130-137, 2019.

[9] P.F. Crapper, “Errors incurred in estimating an area of uniform land cover using Landsat,” Photogrammetric Engineering and Remote Sensing, vol. 46, no. 10, pp. 1295–1301, 1980.

[10] P.F. Crapper, “Geometric properties of regions with homogeneous biophysical characteristics,” Australian Geographical Studies, vol. 19, no. 1, pp. 117–124, 1981.

[11] P.F. Crapper, “An estimate of the number of boundary cells in a mapped landscape coded to grid cells,” Photogrammetric Engineering and Remote Sensing,” vol. 50, no. 10, pp. 1497–1503, 1984.

[12] P.F. Crapper, P.A. Walker, and P.M. Manninga, “Theoretical prediction of the effect of aggregation on grid cell data sets,” Geo-Processing, vol. 3, no. 2, pp. 155–166, 1986.

[13] Y.S. Frolov, and D.H. Maling, “The accuracy of area measurement by point counting techniques,” The Cartographic Journal, vol. 6, no. 1, pp. 2–35, 1969.

[14] P.R. Lloyd, “Quantisation error in area measurement,” The Cartographic Journal, vol. 13, no. 1, pp. 22–25, 1976.

[15] N.A. Bryksina, and Y.M. Polishchuk, “Research of remote measurement accuracy of lake areas using space images,” Geoinformatika, no. 1, pp. 64–68, 2013.

[16] N.V. Koroleva, and D.V. Ershov, “Assessment of an error of determination of areas of windfalls using space images of high spatial resolution of Landsat-TM,” Current problems in remote sensing of the Earth from space, vol. 9, no. 1, pp. 80–86, 2012.

[17] S.G. Kornienko, “Assessment accuracy of measurement of the water body area in the permafrost using different spatial resolution satellite imagery,” Earth's Cryosphere, vol. 18, no. 4, pp. 86–93, 2014.

[18] M.A. Kupriyanov, G.A. Kochergin, and Y.M. Polishchuk, “Accuracy analysis of remote measurement of spatial objects area based on modeling ,” Bulletin of Ugra state university, iss. 3, pp. 25-34, 2018.

[19] A. Savitzky, and M.J.E. Golay, “Smoothing and differentiation of data by simplified least squares procedures,” Analytical chemistry, vol. 36, no. 8, pp. 1627-1639, 1964.

[20] A.M. Berlyant, Cartographic research method. Moscow: Moscow University Press, 1978.

[21] N.Sh. Kremer, The theory of probability and mathematical statistics. Moscow: YuNITI-DANA, 2004.

[22] How Polygon To Raster works, ESRI: http://desktop.arcgis.com/en/arcmap/10.3/tools/conversion-toolbox/how-polygon-to-raster-works.htm

[23] Description of function call scipy.signal.savgol_filter at the SciPy library, The Scipy community: https://docs.scipy.org/doc/scipy-0.16.1/reference/generated/scipy.signal.savgol_filter.html