International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 21 (2017) pp. 10834-10840

Analysis of the Application of Deterministic Interpolation Methods for Land Cadastral Valuation of Low-Rise Residential Development of Localities

Alina Mikhailovna Rybkina[1], Polina Mikhailovna Demidova[2],* and Vladimir Alekseevich Kiselev[3]

Saint-Petersburg Mining University, Saint-Petersburg, 22nd line, d.2, 199106, Russia.
**Correponding author*
[1]Orcid: 0000-0001-8718-5260, [2]Orcid: 0000-0002-8488-512X, [3]Orcid: 0000-0002-3328-7779

Abstract

The article defines the problem of the method of mass cadastral valuation of land of low-rise residential development of localities, effective in the Russian Federation, for which solution the application of spatial interpolation methods of initial data is suggested. In order to define expediency of deterministic methods use, the analysis of their results was carried and defects were detected. Impossibility of application of deterministic methods for mass cadastral valuation and the need for the use of more correct, under these circumstances, spatial interpolation method, free from the identified defects, was substantiated.

Keywords: Cadastral valuation, cadastral value, regression, spatial interpolation, deterministic methods, geostatistics.

INTRODUCTION

At the present time, regression analysis methods are used for state cadastral valuation of the land plots of low-rise residential development of localities. However, statistic data on challenging the results of cadastral valuation and the results of studies in this area show incorrect application of this approach [1].

This is due to the fact that market price of the land plots is a variable with coordinate referencing (spatial variable), which in return doesn't satisfy such requirements of regression model creation as: existing of at least theoretical possibility of unlimited repetition of realization; initial data independence [2].

According to the existing methodology of cadastral valuation, the main task of regressional dependence receipt is determining cost values beyond the points, where the initial data are located. It is possible to solute the same task by using the spatial interpolation method, which is more acceptable in the conditions of space-distributed initial data, herewith, there are two main groups of interpolation methods: deterministic and geostatistical.

Deterministic interpolation methods suppose existing of the required analytical dependence between function values in the space. These methods are popular due to computational efficiency of valuation with the specified parametric formula.

There are four main approaches to deterministic interpolation: linear models, polynomial models, reverse distance models and basis function models [3].

LINEAR INTERPOLATION MODEL

The underlying assumption of the model here is that spatial variable values between measurement points are changed under the law, which is approximated by a straight line. In two-dimensional space, linear interpolation is performed inside the triangle formed by three noncollinear observation points.

Equation of plane is made according to the data in triangle corners (1) [3]:

$$\varphi = ax + by + c, \tag{1}$$

where φ – the measured value of spatial variable; x, y – observation point coordinates; a, b, c – coefficients.

The equation allows to calculate the interpolated value in any point with specified coordinates x and y in the triangle. If there are many observation points, the area covered by them is divided into several triangles, and each of them has its own interpolational equation being calculated (1).

It should be noted that applying this approach for the purposes of cadastral valuation of the land plots of low-rise residential development of localities implies that cadastral value of theland plots located in every triangle formed is not influenced by market prices of the land plots located outside the triangle. One of the drawbacks of this method is that its applying doesn't imply the detection of searching area size, therefore length of the triangle sides can go beyond the area of initial data influence on valuation value [2].

As an illustration of the defects of linear interpolation model, analysis of dependence for the specific indicator of cadastral value (SICV) of the land plots from their location. To this effect, three optional land plots had been selected from the sample of the land plots of low-rise residential development of Volgograd city [4] and the triangle was constructed, under which the values of SICV of the land plots of low-rise residential development were valuated, the center of which gets inside the formed polygon (Fig. 1).

Figure 1: Building polygon for applying the method of linear interpolation

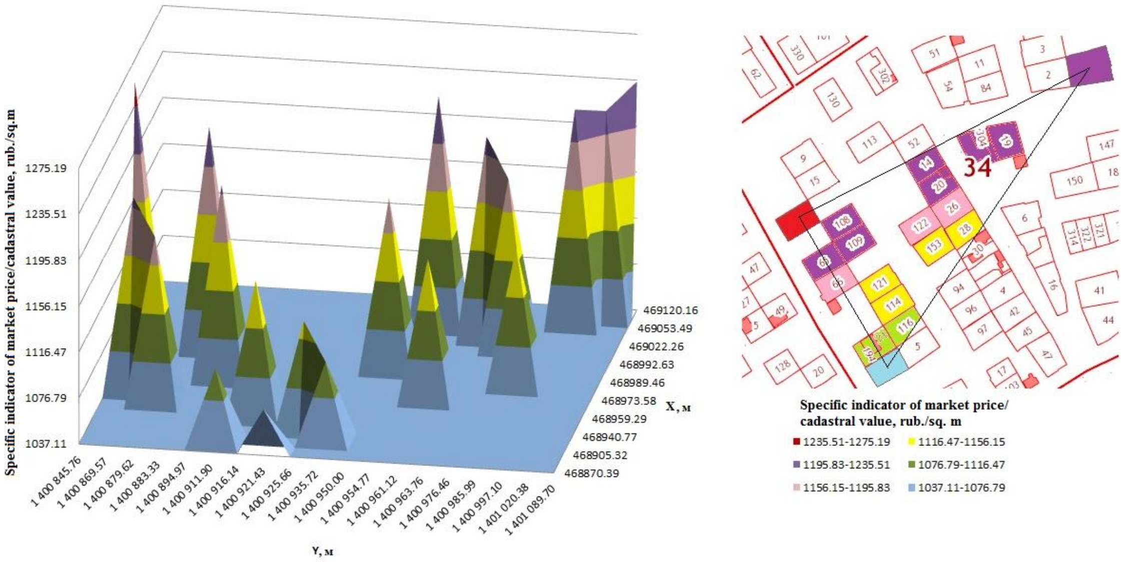

Figure 2. The linear interpolation model of cadastral valuation of land plots of low-rise residential development of Volgograd city: a) - in space b) - on a plane.

In order to determine coefficient values of the equation (1) for this plane, equation system (2) was solved:

$$\begin{cases} 468992.63 \cdot a + 1400845.76 \cdot b + c = 1255.54 \\ 469120.16 \cdot a + 1401089.70 \cdot b + c = 1234.26 \\ 468870.39 \cdot a + 1400921.43 \cdot b + c = 1076.79 \end{cases} \quad (2)$$

It should be noted that the coordinates of the land plots with certain values of market price are presented in MSK-34 (Volgograd region). Solution of the system results in the coefficients: a =1.06396336;b = −0.64346662;c = 403662.05157.. Therefore, interpolation model is described by the formula (3):

$$\begin{aligned} SICV = 1.06396336 \cdot x - 0.64346662 \cdot y + \\ + 403662.05157, \end{aligned} \quad (3)$$

As a result of calculations, linear interpolation model was built (Fig. 2).

Analyzing simulation results allowed detecting a number of defects in the method considered:

- Maximal and minimal values of SICV are reached only in the points with known values of market price while measurement points cannot be the points of function minimum and maximum at some distance.
- In support of this assertion, valuation of the land plots in Volgograd was carried by the linear interpolation method with known values of specific indicator of market price (Fig. 3).

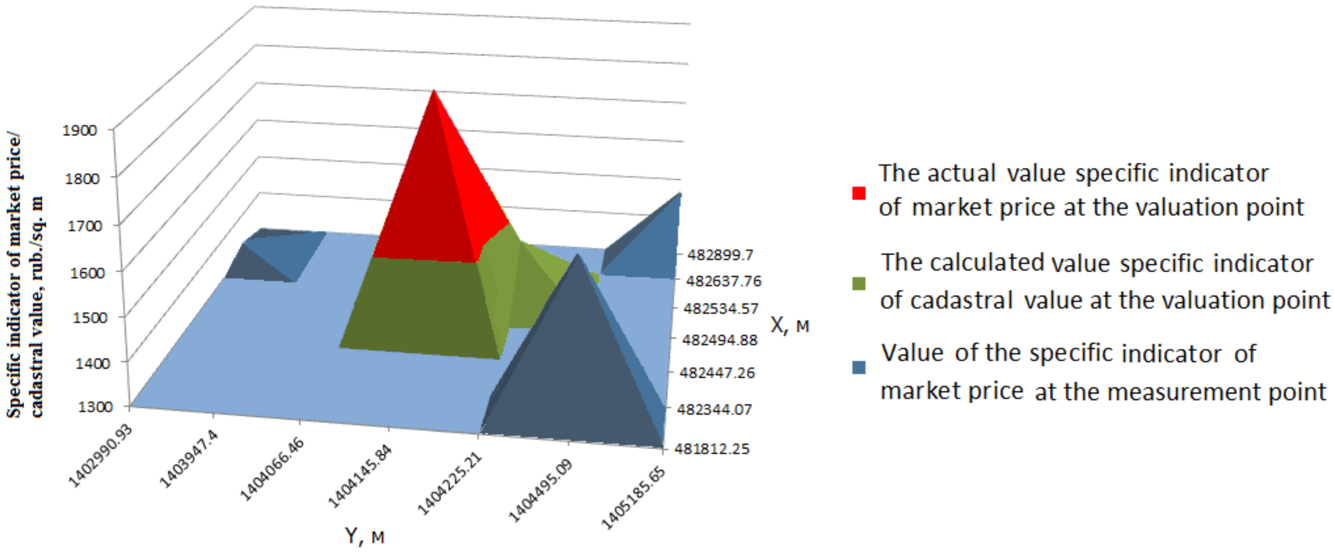

Figure 3. Comparison of the simulation results using linear interpolation method with actual data

So, it was found that the interpolation model built inadequately reflects the real situation of distribution of cadastral value of the land plots of low-rise residential development in Volgograd.

- The linear interpolation accurately reproduces the values in sample points, i.e. it is an accurate interpolator, for which reason there is no possibility to valuate accuracy of the built model on learning sample.
- This method allows to consider the coordinates of the land plots in simulation but herewith, doesn't address the existing space autocorrelation in initial data.

POLYNOMIAL MODELS

This includes two groups of methods: global polynomial method and local polynomial method, which in their turn are mild interpolators. Mild interpolator in a supporting point gives the value distinct from the measured and allows to avoid sharp peaks and dishes on resulting surface [5].

- Global polynomial interpolation chooses smooth surface specified for mathematical function (polynomial) to input supporting points. Global polynomial surface gradually changes and characterizes coarse structure in data [5]. In practice, this method is used for valuation of spatial trend in data, it doesn't try to prognose unknown variable values and loses detail local information, included in data. The equation of global polynomial model is built with least square method on the basis of all initial data (search neighborhood is not applied), and the method is deemed to be global and smoothing interpolator [6]. Thus, global polynomial method rather refers to approximation methods.

In practice, one of the following polynomial types is used for two dimensional system:

1. Plane. Described by formula (4) [2]:

$$P_1(x, y) = a + bx + cy\,, \quad (4)$$

where P_1 – 1st degree polynomial; x, y – observation point coordinates; a, b, c – coordinates.

2. Squared. Determined by equation (5) [2]:

$$P_2(x, y) = a + bx + cy + dxy + ex^2 + fy^2\,, \quad (5)$$

where P_2 – 2nd degree polynomial; a, b, c, d, e, f – coefficients.

3. Cubic. Expressed with formula (6) [2]:

$$P_3(x, y) = a + bx + cy + dxy + ex^2 + fy^2 + gx^2y + hxy^2 + ix^3 + jy^3\,, \quad (6)$$

where P_3 – 3rd degree polynomial; a, b, c, d, e, f, g, h, i, j – coefficients.

Theoretically, the higher order polynominals may be used. However, it should be noted that the more complex polynominal is the harder to add it the physical sense.

- Fundamental difference of local polynomial method is that the coefficients are searched only based on the data got in the search area. Search neighborhood can be identified by size and form, number of neighbors and sector configuration.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 21 (2017) pp. 10834-10840

However, local polynomial interpolation is founded on the following assumptions:

– samples are selected under regular mesokurtic i.e. through equal distances;

– values of data in search neighborhood are normally distributed.

In practice, the most of data sets don't comply with these assumptions.

In the study, polynomial models were built for the sample of land plots market prices of low-rise residential development of Volgograd city [4], since it has more land plots in a sample. To determine degree of a polynomial, the trend analysis has been carried and it was found that the 3rd order polynominal describes the distribution in the best way (Fig. 4).

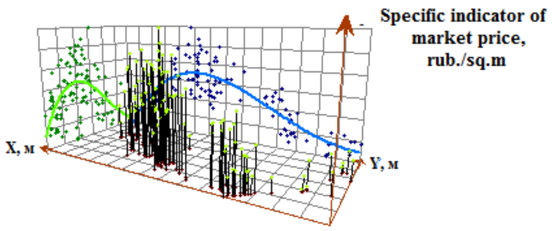

Figure 4: Trend analysis

The analysis of the simulation results presented on figure 5 has showed that two models have bad quality at an average error of approximation of 14 per cent.

Approximation error of the interpolation model built by global polynomial method reached 40-41 per cent at some points, by local polynomial method it reached 50-57 per cent.

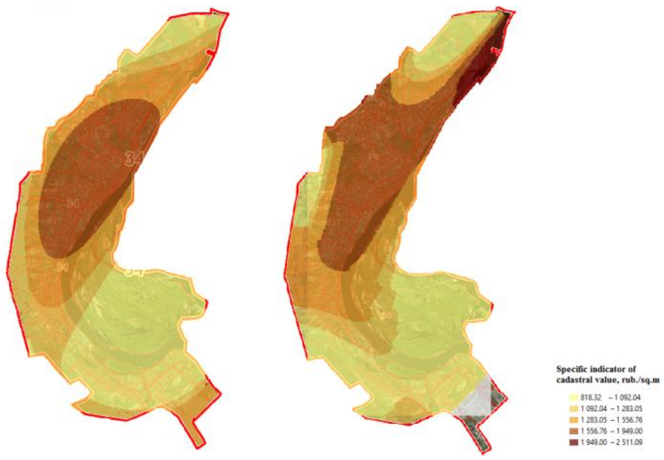

Figure 5. Map of interpolated values a) - global polynomial method b) local polynomial method

Quite large value of approximation error in the first event has been caused by the fact that global polynomial surface characterizes coarse structure in data; valuation is done without regard to the influence radius across the surface under study. However, application of this method allowed finding the global trend. It existence is explained by the fact that in Volgograd, the factor of object remoteness from historical and cultural center of the locality has the greatest impact on cadastral value of the land plots of low-rise residential development [4]. Correspondingly, extensiveness of different price category areas is explained not only by locality configuration but as well by the fact that historical and cultural center of Volgograd has a similar structure and is located in the area with maximal values of SICV (Fig. 6) [7].

Inadequacy of the interpolation model built by local polynomial method sample was due to the fact that it hasn't complied with the requirements specified to the initial sample (Fig. 6).

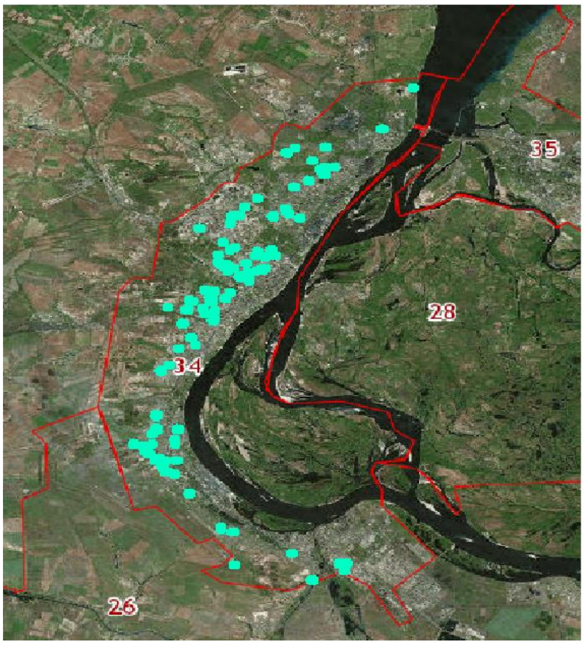

Figure 6. Distribution of the 1st subgroup land plots on the territory of Volgograd

Moreover, the disadvantages of the considered approach can include the fact that it doesn't consider interdependence in the market price values of the land plots of low-rise residential development of localities in Volgograd region.

INVERSE (WEIGHTED) DISTANCE METHOD

In the basis of the model was the idea that measurement impact motion under inverse-square-law attraction (r) from the measurement point, therefore the model is often called potential. Interpolated value φ at every point is found as weighted average from the measured values in neighbor points n and calculated by formula (7) [3]:

$$\varphi = \sum_{i=1}^{n} \frac{\varphi_i}{r^2} : \sum_{i=1}^{n} \frac{1}{r^2}, \qquad (7)$$

where φ_i – measured value in the point i, rub./sq.m; n – number of measurement points; r – distance from measurement point, m.

For predicting three or five nearest measurement points are used or they are restricted by an arbitrary radius R. All measurement points shall be taken into account within this radius. Impact of the measured values is not considered beyond the radius.

Seven land plots of low-rise residential development located in Volgograd were valuated to check usage options of inverse (weighted) distance method for interpolation. For this purpose, calculations of SICV considered 4 measurement points located in triangle corners belonged to the learning sample of market prices of low-rise residential development of Volgograd [4].

The analysis of results showed that when applying inverse (weighted) distance method, minimal and maximal values of SICV for built model are solely reached in measurement points, which is the main disadvantage of inverse distance method [3].

In order to evaluate the adequacy of the built model, calculated and actual values of the specific indicators of market price/ cadastral value of the evaluating land plots were compared. For that purpose the diagram of value characteristics dependence of land plots from the distance to north-western corner of the triangle (Fig. 7).

Analysis of the diagram showed that all calculated values of SICV of the land plots range from 1532.19 to1702.43 rub./sq.m, which complies with the extent between minimal and maximal value of specific indicator of market price in measurement points. However, actual values of specific indicator of market price at evaluated points fall outside the specified range, which shows inadequacy of the built model.

Apart from the detected defects, inverse (weighted) distance method has other defects. In particular: failure to characterize quality of valuation, spatial correlation is not considered [2].

MODELS OF BASIS FUNCTIONS

In this method, valuation of variable φ in an arbitrary point of the studied area is a linear combination of the values of radial basis functions (RBF) [6]:

$$\hat{\varphi}_0 = \sum_{i=1}^{n} \lambda_i B(d_{i0}), \qquad (8)$$

where λ_i is a coefficient for the 1st selective point; $B(d_{i0})$ is radial basis function, where an argument is distance d_{i0}; d_{i0} – distance between the point, where the valuation is calculated, and the 1st measurement point.

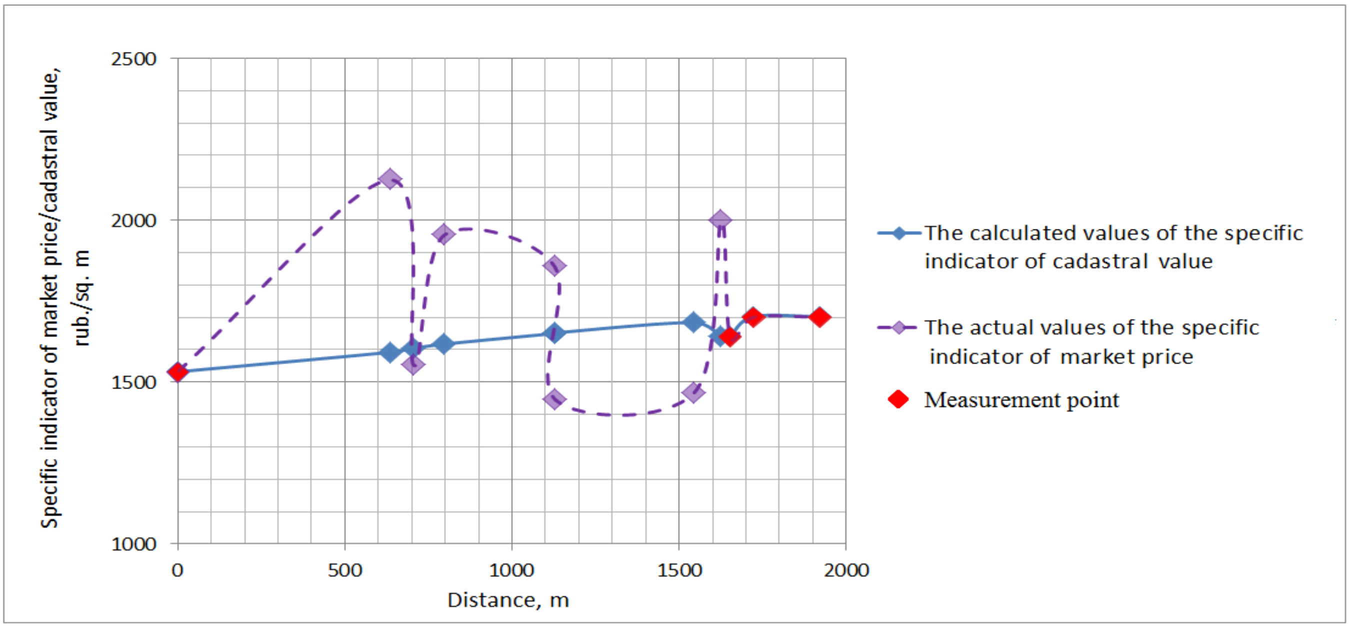

Figure 7. Comparison of the results of simulation by inverse (weighted) distance method with actual data

RBF is the kernel function, which determines optimal weights applied to data points during interpolation. There are many variations of RBF, however, traditionally the following types of basis kernel functions are used [2,6]:

- Inverse multiquadric function:

$$B(d)=\frac{1}{\sqrt{d^2+\delta^2}}, \qquad (9)$$

where d – distance from interpolation point to selective point; δ – smoothing parameter, which rational values range from an average distance between sample points to the half of this average.

- Multilogarithm function:

$$B(d)=\log(d^2+\delta^2); \qquad (10)$$

- Multiquadric function:

$$B(d)=\sqrt{d^2+\delta^2}; \qquad (11)$$

- Cubic spline:

$$B(d)=(d^2+\delta^2)^{\frac{3}{2}}; \qquad (12)$$

- Plane spline:

$$B(d)=(d^2+\delta^2)\log(d^2+\delta^2). \qquad (13)$$

Multiquadric function is the most commonly used, which, as many scientists believe, is the best from the point of view of smooth surface building [6].

The coefficients λ_i are determined from the condition of evaluation precision in the known points of output surface passing through values φ in n selective points (determination of weights is made under $\delta = 0$):

$$\sum_{i=1}^{n}\lambda_i B(d_{ij})=\varphi_j , \qquad (14)$$

where $j = 1, \ldots, n$.

n coefficients λ_i. are unknowns in the equation. Sequence of actions for their determination includes: Calculation of distances between all sample points (d_{ij}), calculation of RBF values ($B(d_{ij})$), using them, solution of equation system.

In the study, the land plots considered in the previous section have been evaluated by multiquadric function. Thus, in order to find coefficients λ_i the system of four equations should be soluted (15):

$$\begin{cases}\lambda_2\cdot 1721.19+\lambda_3\cdot 1920.30+\lambda_4\cdot 1651.04=1532.19\\ \lambda_1\cdot 1721.19+\lambda_3\cdot 1826.92+\lambda_4\cdot 2351.27=1702.43\\ \lambda_1\cdot 1920.30+\lambda_2\cdot 1826.92+\lambda_4\cdot 916.88=1702.43\\ \lambda_1\cdot 1651.04+\lambda_2\cdot 2351.27+\lambda_3\cdot 916.88=1641.63\end{cases}, \qquad (15)$$

Solution of the system results in the coefficients: λ_1= 0.3556297; λ_2= 0.3868966; λ_3= 0.1578973; λ_4= 0.3410319.

Application of the method of radial basis functions showed that it can predict the values, which are above maximal and below minimal measured values (Fig. 8).

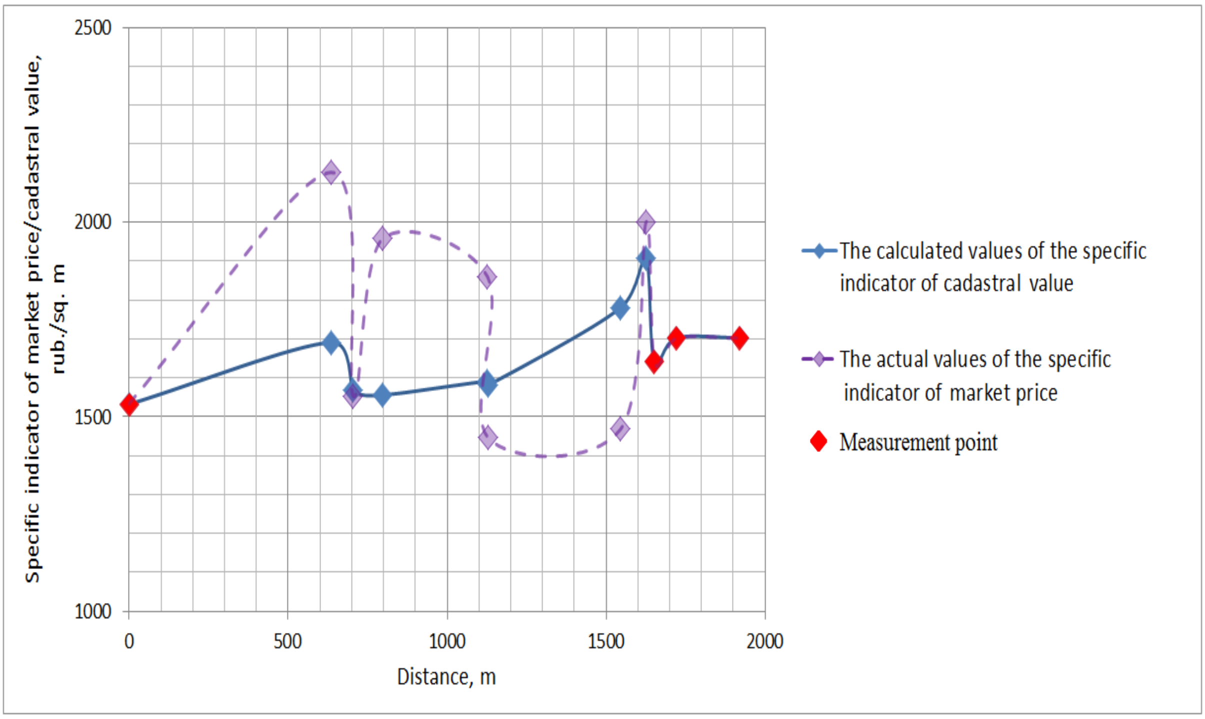

Figure 8. Comparison of the simulation results by the method of radial basis functions with actual data

The main disadvantage of the considered method is the absence of stringent requirements for determining of smoothing parameter value (iterative approach), the method of radial basis functions as well doesn't take into account existence of spatial correlation in initial data and gives no possibility to evaluate accuracy of the built models according to the learning sample.

CONCLUSIONS

The analysis of the results for application of deterministic interpolation methods for purposes of cadastral valuation of land of low-rise residential development of localities revealed a number of serious defects in this group of methods: failure to characterize quality of valuation, neglecting spatial correlation in data. Moreover, deterministic interpolation methods similarly to regression analysis methods allow valuation on the samples with enough data on land transactions.

Thus, deterministic interpolation methods allow building the surface of interpolated values, including only the values of coordinates but not including spatial correlation of valuation objects. All above mentioned allows concluding on

impossibility of using the considered methods to determine cadastral value of the land plots of low-rise residential development in Volgograd region and results in choosing spatial interpolation method, which is free from the disadvantages determined. In this case, it is useful to consider the possibility of applying geostatistical interpolation methods, which allow taking account of a spatial position of the land plots and spatial correlation of valuation objects. Application of geostatistical interpolation methods proved themselves both abroad and in the Russian Federation [1, 8, 9, 10].

REFERENCES

[1] Rybkina A. M. (2016) Justification of the use of geostatistical interpolation methods when conducting cadastral valuation of urban lands suitable for individual residential development. *Imushchestvennye otnosheniya v Rossiyskoy Federatsii*, No. 11 (182), pp. 29-40. (in Russian)

[2] Demyanov V. V. and Saveleva E. A. (2010). *Geostatistika: teoriya i praktika* [Geostatic: theory and practice]. Moskva: Nauka. (in Russian)

[3] Porotov G. S. (2006) *Matematicheskie metody modelirovaniya v geologii* [Mathematical methods of simulation in geology]. Saint-Petersburg: St. Petersburg State Mining Institute (technical university). (in Russian)

[4] On determining cadastral value of the land plots as part of the land of localities within the territory of Volgograd region. Report of Federal State Unitary Enterprise Rostechinventarizatsya – Federal BTI No. 34-GKOSNP-2015 dated 10.09.2015. Available at: https://rosreestr.ru/wps/portal/p/cc_ib_portal_services/cc_ib_ais_fdgko/ (accessed 18.10.2016).

[5] ArcGIS for Desktop. Available at: http://desktop.arcgis.com/ru/arcmap/10.3/guide-books/extensions/geostatistical-analyst/deterministic-methods-for-spatial-interpolation.htm (accessed 22.01.2017).

[6] Maltsev K. A. , Mukharamova S. S. (2014) *Postroenie modeley prostranstvennykh peremennykh (s primeneniem paketa SURFER)* [Building of spatial variable models (using Surfer package)]. Kazan: Kazan. Universe, 103 p. (in Russian)

[7] Volgograd Cartographic fund. Available at: http://www.volgmap.ru/pzzvlg.map/?layers=Terr_zone (accessed 22.01.2017).

[8] Rybkina A.M., Demidova P.M., Kiselev V.A. (2016) Working-Out of the Geostatistical Model of Mass Cadastral Valuation of Urban Lands Evidence from the City Vsevolozhsk (Russia). *International Journal of Applied Engineering Research*, 2016, Vol. 11, no. 24, pp. 11631-11638.

[9] Kuntz M., Helbich M. (2014) Geostatistical mapping of real estate prices: An empirical comparison of kriging and Cokriging. *International Journal of Geographical Information Science*, 2014, 28 (9), pp. 1904–1921.

[10] Jorge Chica-Olmo (2007) Prediction of Housing Location Price by a Multivariate Spatial Method: Cokriging. *JRER,* 2007, Vol. 29, no. 1, pp. 93-112.