

Рапаков Георгий Германович

Кандидат технических наук, доцент,
Вологодский государственный университет
(Вологда, Россия)
E-mail: grapakov@yandex.ru

Rapakov Georgiy Germanovich

PhD in Technical Sciences, Associate
Professor, Vologda State University
(Vologda, Russia)
E-mail: grapakov@yandex.ru

Лебедева Елена Александровна

Кандидат технических наук, доцент,
Вологодский государственный университет
(Вологда, Россия)
E-mail: lebedeva.elena@inbox.ru

Lebedeva Elena Aleksandrovna

PhD in Technical Sciences, Associate
Professor, Vologda State University
(Vologda, Russia)
E-mail: lebedeva.elena@inbox.ru

Горбунов Вячеслав Алексеевич

Доктор физико-математических наук,
профессор, Вологодский государственный
университет
(Вологда, Россия)
E-mail: gorbunov1945@inbox.ru

Gorbunov Vyacheslav Alekseevich

Doctor of Physico-mathematical Sciences,
professor, Vologda State University
(Vologda, Russia)
E-mail: gorbunov1945@inbox.ru

Кузьминов Александр Леонидович

Доктор технических наук, профессор,
Череповецкий государственный университет
(Череповец, Россия)
E-mail: TSiTB@chsu.ru

Kuzminov Aleksandr Leonidovich

Doctor of Technical Sciences, Professor,
Cherepovets State University
(Cherepovets, Russia)
E-mail: TSiTB@chsu.ru

Абдалов Кенгес Абдалович

Доктор медицинских наук, заведующий
хирургическим отделением №2,
Вологодская городская больница №1
(Вологда, Россия)
E-mail: muzgb1@visp.ru

Abdalov Kenges Abdalovich

Doctor of Medicine, M.D. Head of the surgical
department no. 2,
Vologda municipal hospital no. 1
(Vologda, Russia)
E-mail: muzgb1@visp.ru

**ГЕОИНФОРМАЦИОННОЕ
КАРТОГРАФИРОВАНИЕ
НА ОСНОВЕ МАШИННОГО
МОДЕЛИРОВАНИЯ
ВАРИОГРАММЫ**

**GEOINFORMATION
MAPPING USING VARIOGRAM
MACHINE MODELING**

Аннотация. Авторы публикации выполнили компьютерное моделирование теоретических вариограмм в ходе геоэкологического исследования. Оценка качества пространственного распределения удельной электропроводности снежного покрова получена при помощи перекрестной проверки грида. Выполнена визуализация результатов кросс-валидации для метода кригинга. Рас-

Abstract. The authors presented computer simulations results of the comparison of theoretical variograms during implementation of the Vologda city ecological monitoring. To study the spatial interpolation quality of the snow cover electrical conductivity, researches used cross-validation for the kriging method. Visualization the isoline maps and calculation of statistical parameters helps to apply the comparative results

считаны геостатистические показатели моделей вариограмм. Сравнительные результаты вариограммного анализа использованы при поддержке управленческих решений в региональных программах рационального природопользования.

Ключевые слова: компьютерное моделирование, кригинг, модель вариограммы, перекрестная проверка, геоинформационные системы, геостатистика, геоэкологический мониторинг, поддержка решений

of variogram analysis for decision support of regional water conservation project.

Keywords: computer modeling, kriging, variogram method, cross-validation, geographic information systems, geostatistics, geoecological monitoring, decision support

Введение

Одной из проблем машинного моделирования геоинформационных данных, в отличие от численных моделей физических процессов в технологических системах, является пространственная интерполяция по ограниченному набору точечных проб [4], [2]. К перечню социально значимых региональных задач относится снижение рисков здоровью населения и создание безопасной среды обитания [5]. В работе [4] из 12 методов сеточной интерполяции (гридинга) к моделям наилучшего выбора отнесен кригинг. Результаты геоинформационного картографирования при помощи метода кригинга зависят от подбора параметров вариограммы, определяющей меру пространственной корреляции. Для проверки качества моделей при построении регулярной сетки (грида) на основе неструктурированных исходных данных в геостатистике активно используется перекрестная проверка (кросс-валидация) [1], [9]. Актуальной является задача моделирования пространственного распределения химического загрязнения среды обитания на основе кригинга с подгонкой вариограммы и оценкой качества грида методом кросс-валидации. Практическая значимость исследования обусловлена использованием результатов при поддержке управленческих решений в региональном проекте улучшения качества водоснабжения. Целью настоящей работы является исследование методов вариограммного анализа на основе машинного моделирования. Обоснованный выбор теоретической вариограммы оказывает значимое влияние на представление пространственной структуры геоэкологических данных, что обуславливает новизну работы.

Основная часть

Сеточная интерполяция позволяет перейти от неструктурированной сетки с неравномерно распределенными точками исходных данных к регулярной решетке. При этом необходимо определить значения показателя, интересующего исследователя, в узлах правильной сети. Кригинг выполняет исследование и моделирование пространственной корреляционной структуры данных, в качестве меры которой выступает статистический момент второго порядка – вариограмма. Реализация кригинга предполагает расчет опытной $\gamma^*(h)$ вариограммы на основе данных экспериментальных исследований. Выборочную вариограмму аппроксимируют модельной функцией вариограммы $\gamma(h)$. Основные функции теоретической вариограммы для геостатистического программного обеспечения (ПО) Surfer:

экспоненциальная [7]

$$\gamma(h) = C(1 - e^{-h})$$

линейная [8]

$$\gamma(h) = C(h)$$

степенная [11]

$$\gamma(h) = C(h^n), \quad 0 < n < 2;$$

рациональная квадратичная [7]

$$\gamma(h) = C \left(\frac{h^2}{1 + h^2} \right);$$

периодическая [7]

$$\gamma(h) = C \left(1 - \frac{\sin(h)}{h} \right);$$

пентасферическая [10]

$$\gamma(h) = C(1,875 h - 1,25 h^3 + 0,375 h^5),$$

гауссова [6]

$$\gamma(h) = C(1 - e^{-h^2})$$

логарифмическая [8]

$$\gamma(h) = C(\log(h)), \quad h > 0$$

квадратическая [6]

$$\gamma(h) = \begin{cases} C(2h - h^2), & h < 1; \\ C, & h \geq 1 \end{cases};$$

сферическая [11]

$$\gamma(h) = \begin{cases} C(1,5h - 0,5h^3), & h < 1; \\ C, & h \geq 1 \end{cases};$$

кубическая [10]

$$\gamma(h) = C(7h^2 - 8,75h^3 + 3,5h^5 - 0,75h^7);$$

где h – разность расстояний между парами точек, C – параметр масштаба для определения порогового значения $\gamma(h)$. В алгоритмы интерполяции геостатистического программного обеспечения (ПО) встраиваются различные формулы вариограмм. Если размер набора входных данных менее 250 наблюдений, разработчики ПО Surfer рекомендует использовать для пространственной интерполяции кригинг с линейной вариограммой, назначаемой по умолчанию. В качестве альтернативы рассматривается применение искусственных нейронных сетей: метод радиальных базисных функций с мультиквадратической функцией, что подтверждается исследованиями [4]. Наилучшее приближение обеспечивает наименьшую дисперсию отклонений между эмпирическими и расчетными значениями. На основе отобранной модельной вариограммы составляют системы уравнений кригинга и выполняют расчет.

Для оценки качества непрерывного поля, используемого при описании пространственной изменчивости показателя, применяется кросс-валидация. При подборе наилучшей вариограммы ее оценка выполняется по величине средней квадратичной ошибки RMSE [1]:

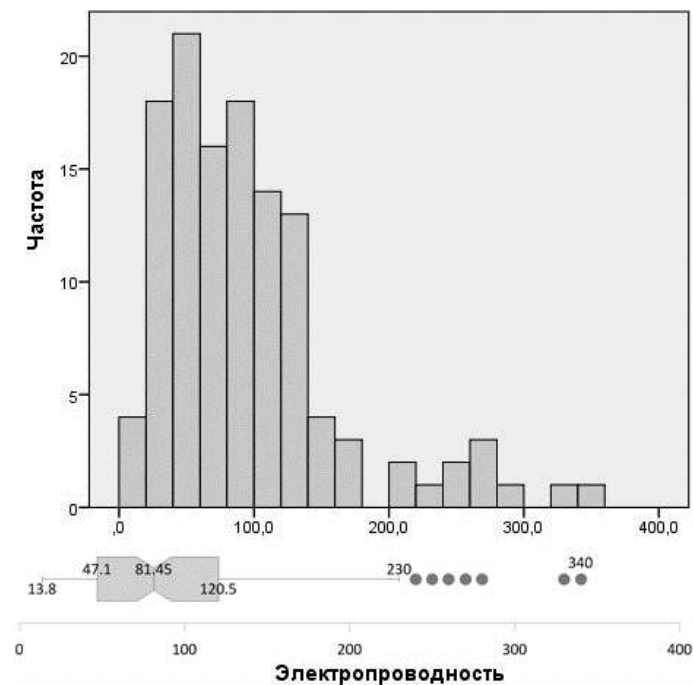
$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n [Z(x_i) - Z^*(x_i)]^2}.$$

Анализ состояния исследований проблемы основан на обзоре научных публикаций. В работах [6]–[8], [10], [11] обсуждаются типы моделей вариограмм. Авторы [3], анализируя гидрофизические поля, рассмотрели обычный и универсальный кригинг. Последний метод применялся для случая геоморфологических процессов, не являющихся стационарными, когда в данных присутствовал значимый тренд. Для моделирования использовалась открытая ГИС QGIS. Оценка погрешности процедур интерполяции была выполнена методом кросс-валидации. Карты изолиний темпера-

турных полей построены при помощи 4 функций модельной вариограммы: линейной, квадратической, кубической и экспоненциальной. В работе [1] выполнен сравнительный валидационный анализ 6 методов пространственного моделирования по данным экологического мониторинга радиоактивного загрязнения почвы Am^{241} . Были использованы вариограммы по всем направлениям, зависящие от модуля вектора, разделяющего точки пары и анизотропные, учитывающие его направление.



а)



б)

Рис. 1. Карта сбора точечных проб (а);
гистограмма и ящичная диаграмма электропроводности (б)

Обеспечение геоэкологической безопасности требует системного подхода и проведения комплексных мероприятий по снижению воздушного загрязнения, утилиза-

ции снега, обработке и доочистке сточных вод [5]. Пробы снега были отобраны в период начала весеннего снеготаяния в 2017 г. Координаты точек определялись путем моделирования геопространственной сетки с учетом данных об общественном здоровье (рис. 1а). Объем выборки обеспечивает точность оценки не ниже 5 % с доверительной вероятностью $\alpha = 0,95$. Гистограмма и ящичная диаграмма электропроводности фильтрата представлены на рис. 1б.

Сравнительная верификация в ходе вариограммного анализа выполняется как качественно – на основе визуального оценивания (рис. 2, 3), так и количественно – при помощи статистик моделей вариограмм (см. таблицу). Худшие результаты демонстрируют следующие теоретические вариограммы: периодическая, гауссова и кубическая. Значения глобальной характеристики меры ошибки RMSE для них максимальны, а построения полей формируют артефакты. Среднее RMSE для всего набора моделей без учета трех выпадающих значений составляет 70,344. Линейная модель вариограммы демонстрирует устойчивые средние значения и является предпочтительной с точки зрения предотвращения эффекта избыточного моделирования. Кригинг на ее основе строит пространственное распределение высокой визуальной привлекательности с четкой равномерной структурой и малой зашумленностью. Для визуализации результатов моделирования применяется точечный кригинг, отвечающий нулевому размеру сегмента оценки для пробы электропроводности. Из 11 теоретических вариограмм к моделям наилучшего выбора отнесены: степенная, логарифмическая, экспоненциальная и линейная. Рациональная квадратичная, квадратичная, пентасферическая и сферическая модели – ближайшие к ним.

Таблица

Статистики перекрестной проверки моделей вариограмм

Статистические показатели для моделей вариограмм	Минимум	Максимум	Среднее значение	RMSE	Абсолютное отклонение от медианы	Сумма квадратов остатков
Экспоненциальная	-234,786	151,854	1,408	69,242	35,115	580 367
Гауссова	-3574,328	2778,683	0,967	680,051	192,323	55 958 867
Линейная	-233,546	167,067	-0,008	73,775	39,897	658 566
Логарифмическая	-227,568	111,558	0,772	64,811	29,684	508 332
Степенная	-228,626	99,118	0,526	64,205	29,930	498 830
Квадратичная	-239,253	168,605	1,140	74,708	39,778	675 492
Рациональная квадратичная	-238,761	122,747	1,269	65,958	34,852	526 607
Сферическая	-236,856	169,111	0,884	77,096	35,606	719 302
Периодическая	-7267,460	17535,584	62,814	2635,146	681,734	840 704 680
Кубическая	-1002,862	554,063	-7,617	176,696	72,519	3 784 895
Пентасферическая	-238,167	169,736	1,002	75,222	39,258	684 783

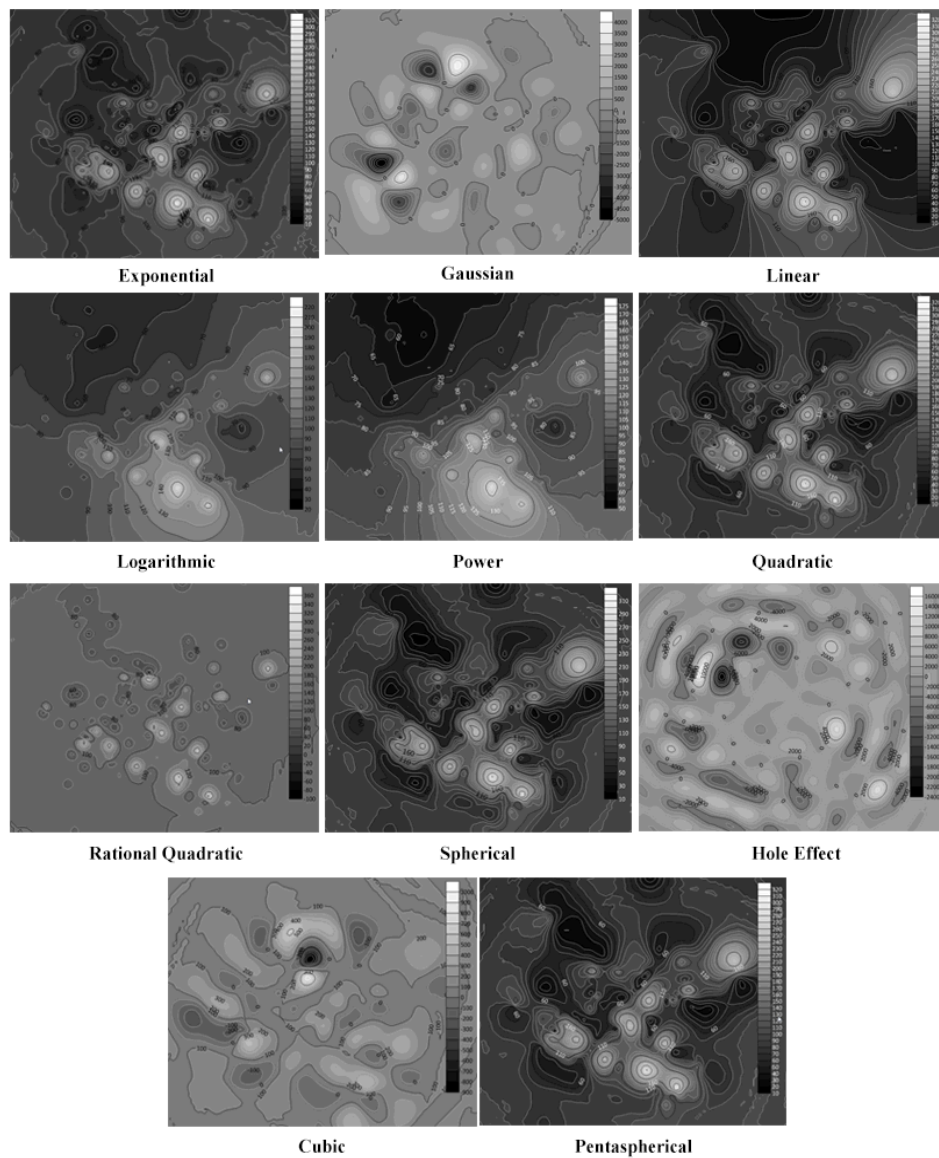


Рис. 2. Исследование моделей вариограмм электропроводности для метода кригинга

Карты пространственного распределения удельной электропроводности снегового фильтрата с интервалом между изолиниями 10 мкСм/см, полученные методом кригинга с применением исследуемых моделей вариограмм, представлены на рис. 2. Соответствующие поверхности невязок кросс-валидации, построенные при помощи точного интерполятора – метода обратных расстояний без сглаживания, позволяют оценить достоверность результатов подгонки с учетом пространственной структуры и значений ошибок (рис. 3).

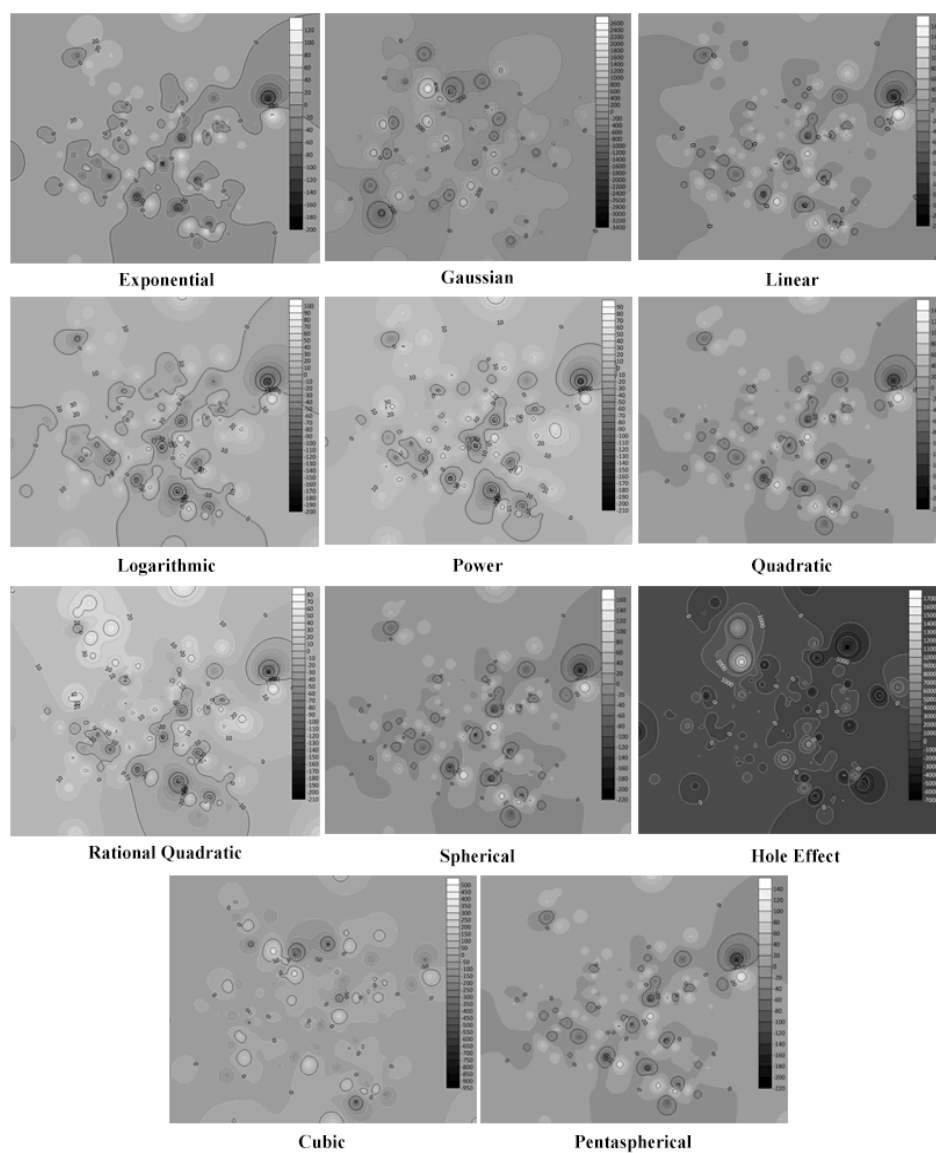


Рис. 3. Распределения остатков при перекрестной проверке моделей вариограмм

Выводы

Выполнено компьютерное моделирование пространственного распределения удельной электропроводности в ходе обработки данных геоэкологического мониторинга для 11 теоретических вариограмм. При оценке качества применяется метод кросс-валидации. Рассчитаны статистики и визуализированы невязки перекрестной проверки моделей. Геоинформационное картографирование выполнено в виде карт изолиний методом кригинга. Перспективы исследований связаны с изучением пространственной анизотропии.

Литература

1. Демьянов В.В., Савельева Е.А., Арутюнян Р.В. Геостатистика: теория и практика. М.: Наука, 2010. 327 с.
2. Лукин С.В., Кибардин А.Н., Антонова Ю.В. Алгоритм оценки параметров теплового состояния сляба на линии «МНЛЗ – термостат – нагревательная печь» // Вестник Череповецкого государственного университета. 2017. № 5 (80). С. 25–36. DOI 10.23859/1994-0637-2017-5-80-3
3. Новикова А.М., Каширина Е.С., Новиков А.А., Полонский А.Б., Панкеева Т.В. ГИС в морских исследованиях: мировой опыт и возможности его применения на примере Черноморского региона // Труды Карадагской научной станции им. Т.И. Вяземского природного заповедника РАН. 2017. № 1 (3). С. 54–66.
4. Рапаков Г.Г., Лебедева Е.А., Горбунов В.А., Кузьминов А.Л., Абдалов К.А. Компьютерное моделирование и анализ методов сеточной интерполяции при обработке геоинформационных данных // Вестник Череповецкого государственного университета. 2018. №2 (83). С. 31–39. DOI 10.23859/1994-0637-2018-1-83-4
5. Шестакова Е.А., Калугина Н.Ф., Шестаков Н.И., Аншелес В.Р. Об особенностях применения технологии микрофильтрации для доочистки сточных вод МУП «Водоканал» г. Череповца // Череповецкие научные чтения – 2016: материалы всероссийской научно-практической конференции. Череповец, 2017. С. 221–222.
6. Alfaro M. The Random Coin Method: Solution of the Problem of the Simulation of a Random Function in the Plane. *Mathematical Geology*, 1980. vol. 12 (1), pp. 25–32. DOI: 10.1007/bf01039901
7. Cressie N.A.C. *Statistics for Spatial Data Revised Edition* (Wiley series in probability and mathematical statistics. Applied probability and statistics section). A Wiley – Interscience Publication: John Wiley & Sons, Inc., 1991. 917 p.
8. Kitanidis P. K. *Introduction to Geostatistics: Applications in Hydrogeology*. Cambridge University Press, 1997. 271 p.
9. Montero J-M., Fernandez-Aviles G., Mateu J. *Spatial and Spatio – Temporal Geostatistical Modeling and Kriging*. John Wiley & Sons, Ltd., 2015. 385 p.
10. Olea R. A. *Geostatistics for Engineers and Earth Scientists*. Springer Science+Business Media, LLC, 1999. 309 p. DOI: 10.1007/978-1-4615-5001-3
11. Pannatier Y. *Variowin Software for Spatial Data Analysis in 2D*. Springer-Verlag New York, Inc., 1996. 91 p. DOI: 10.1007/978-1-4612-2392-4

References

1. Alfaro M. The Random Coin Method: Solution of the Problem of the Simulation of a Random Function in the Plane. *Mathematical Geology*, 1980, vol. 12 (1), pp. 25–32. DOI: 10.1007/bf01039901
2. Cressie N.A.C. *Statistics for Spatial Data Revised Edition* (Wiley series in probability and mathematical statistics. Applied probability and statistics section). A Wiley – Interscience Publication: John Wiley & Sons, Inc., 1991. 917 p.
3. Dem'ianov V.V., Savel'eva E.A., Arutiunian R.V. *Geostatistika: teoriia i praktika* [Geostatistics: the theory and practice]. Moscow: Science, 2010. 327 p.
4. Kitanidis P.K. *Introduction to Geostatistics: Applications in Hydrogeology*. Cambridge University Press, 1997. 271 p.
5. Lukin S.V., Kibardin A.N., Antonova Iu.V. Algoritm otsenki parametrov teplovogo sostoiianiia sliaba na linii “MNLZ – termostat – nagrevatel'naia pech” [Algorithm of slab heat state parameters evaluation on the line «CCM – thermostat – heating stove»]. *Vestnik Cherepovetskogo gosudarstvennogo universiteta* [Cherepoves State University bulletin], 2017, no. 5 (80), pp. 25–36. DOI 10.23859/1994-0637-2017-5-80-3
6. Montero J-M., Fernandez-Aviles G., Mateu J. *Spatial and Spatio – Temporal Geostatistical Modeling and Kriging*. John Wiley & Sons, Ltd., 2015. 385 p.
7. Novikova A.M., Kashirina E.S., Novikov A.A., Polonskii A.B., Pankeeva T.V. GIS v morskikh issledovaniakh: mirovoi opyt i vozmozhnosti ego primeneniia na primere Chernomorskogo

regiona. *Trudy Karadagskoi nauchnoi stantsii im. T.I. Viazemskogo prirodnogo zapovednika RAN*, 2017, no. 1 (3), pp. 54–66.

8. Olea R.A. *Geostatistics for Engineers and Earth Scientists*. Springer Science+Business Media, LLC, 1999. 309 p. DOI: 10.1007/978-1-4615-5001-3

9. Pannatier Y. *Variowin Software for Spatial Data Analysis in 2D*. Springer-Verlag New York, Inc., 1996. 91 p. DOI: 10.1007/978-1-4612-2392-4

10. Rapakov G.G., Lebedeva E.A., Gorbunov V.A., Kuz'minov A.L., Abdalov K.A. Komp'yuternoe modelirovanie i analiz metodov setochnoi interpolatsii pri obrabotke geoinformatsionnykh dannykh. *Vestnik Cherepovetskogo gosudarstvennogo universiteta* [Cherepoves State University bulletin], 2018, no. 2 (83), pp. 31–39. DOI 10.23859/1994-0637-2018-1-83-4

11. Shestakova E.A., Kalugina N.F., Shestakov N.I., Ansheles V.R. Ob osobennostiakh primeneniia tekhnologii mikrofil'tratsii dlia doochistki stochnykh vod MUP "Vodokanal" g. Cherepovtsa [About features of application the microfiltration technology for additional cleaning of sewage MUP "Vodokanal" Cherepovets city]. *Cherepovetskie nauchnye chteniia – 2016: materialy vserossiiskoi nauchno-prakticheskoi konferentsii* [Cherepovets scientific readings – 2016: materials of the all-Russian scientific – practical conference]. Cherepovets: Cherepoves State University, 2017, pp. 221–222.

Для цитирования: Рапаков Г.Г., Лебедева Е.А., Горбунов В.А., Кузьминов А.Л., Абдалов К.А. Геоинформационное картографирование на основе машинного моделирования вариограммы // Вестник Череповецкого государственного университета. 2018. №3(84). С. 46–54. DOI: 10.23859/1994-0637-2018-3-84-6.

For citation: Rapakov G.G., Lebedeva E.A., Gorbunov V.A., Kuzminov A.L., Abdalov K.A. Geoinformation mapping using variogram machine modeling. *Bulletin of the Cherepovets State University*, 2018, no. 3 (84), pp. 46–54. DOI: 10.23859/1994-0637-2018-3-84-6.