

Tamara Myslyva¹, Dr. Sci. (Agric.), Professor

Olesya Kutsaeva¹, Senior Lecturer

Yurij Bilyavskij², Cand. Sci. (Agric.), Associate Professor

¹Belarus State Agrarian Academy, e-mail: byrty41@yahoo.com

²Zhytomyr National Agroecological University

EFFECTIVENESS OF GIS-BASED INTERPOLATION METHODS FOR ESTIMATION OF SPATIAL DISTRIBUTION OF pH IN SOIL

Analysis of spatial variation of soil properties is fundamental to sustainable agricultural and rural development and environmental protection. The spatial variability of soil property is often measured using various interpolation methods resulting in map generation. Selecting a proper spatial interpolation method is crucial in surface analysis, since different methods of interpolation can lead to different surface results. Among statistical methods, geo-statistical kriging-based techniques have been frequently used for spatial analysis and surface mapping. The main objective of this study is to review and evaluate the three common interpolation methods namely: Inverse Distance Weighting (IDW), Radial Base Function (RBF) and Ordinary Kriging (OK), and generate maps of soil pH property using these methods. The accuracy and efficiency of the generated maps have been examined and the most fitting technique for the soil pH in the study area was identified. For the analysis were used data on pH_{KCl} of soil solution obtained from materials of agrochemical survey of the territory of RUE "Uchkhoz BGSMA" (Republic of Belarus, Mogilev region, Goretzky district), executed in 2014. The total area of the surveyed territory is 6361.67 ha. Statistical analyses were performed in three stages. First, the frequency distributions of data were analyzed and normality was tested using the Kolmogorov-Smirnov test. Secondly, the distribution of data was described using conventional statistics such as mean, maximum, minimum, median, Standard Deviation (S.D), Coefficient of Variation (CV), skewness and kurtosis. Thirdly, Global trend analysis was performed to examine the existence of trend in pH data. A three-dimensional perspective of the data was created using the trend analysis tool of ArcMap 9.2 software package. In this investigation, three common interpolation methods were used to study the spatial distributions of soil pH in an arable land. Interpolation techniques were used to estimate the pH measurement in unsampled points and create a continuous dataset that could be represented over a map of the entire study area. The method investigated includes; Inverse Distance Weighting (IDW), Radial base Function (RBF) and Ordinary Kriging (OK). The performance of conventional statistics showed that soil pH had a low variation in this study. Experimental anisotropic semivariograms were fitted with the Spherical, Exponential and Gaussian Exponential models and the Exponential model was found as the best fitted

model using the cross-validation method. The performances of interpolation methods were evaluated and compared using the cross-validation. The results showed that RBF method performed better than IDW and OK for prediction of the spatial distribution of topsoil pH.

Keywords: *pH, forecasting, geoinformation systems, deterministic interpolation, geostatistical interpolation.*

УДК 631.472.51

Т. М. Мислива¹, д-р с.-г. наук, професор

О. О. Куцаєва¹, старший викладач

Ю. А. Білявський², канд. с.-г. наук, доцент

¹*Belarus State Agrarian Academy, e-mail: byrty41@yahoo.com*

²*Zhytomyr National Agroecological University*

EFFECTIVENESS OF GIS-BASED INTERPOLATION METHODS FOR ESTIMATION OF SPATIAL DISTRIBUTION OF pH IN SOIL

Проаналізовано різні методи інтерполяції в середовищі географічної інформаційної системи ArcGIS і оцінено їх придатність для прогнозування просторового розподілу рН в 0-20-сантиметровому шарі ґрунту на території РУП «Учгосп БДСГА». Зроблено аналіз стратифіковань випадкової вибірки з 64 значень рН. Для генерації просторового розподілу досліджуваного показника використовувалися три методи інтерполяції: метод зворотних зважених відстаней (IDW), метод радіальних базисних функцій (RBF) і ординарний крігінг (OK). Для оцінки точності методів інтерполяції застосовували перехресну перевірку шляхом порівняння величин середньої помилки (ME) і середньоквадратичної помилки (RMSE). Установлено, що метод радіальних базисних функцій є найбільш придатним для інтерполяції просторового розподілу рН ґрунту з найменшим значенням RMSE, рівним 0,383.

Ключові слова: *рН, прогнозування, геоінформаційні системи, детермінована інтерполяція, геостатистична інтерполяція.*

УДК 631.472.51

Т. Н. Мыслыва¹, д-р с.-х. наук, профессор**О. А. Куцаева¹, старший преподаватель****Ю. А. Белявский², канд. с.-х. наук, доцент**

¹УО «Белорусская государственная сельскохозяйственная академия»,
e-mail: byrty41@yahoo.com

²Житомирский национальный агроэкологический университет

ЭФФЕКТИВНОСТЬ МЕТОДОВ ИНТЕРПОЛЯЦИИ НА ОСНОВЕ ГИС ДЛЯ ОЦЕНКИ ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ pH В ПОЧВЕ

Проанализированы различные методы интерполяции в среде географической информационной системы ArcGIS и оценена их пригодность для прогнозирования пространственного распределения pH в 0-20-сантиметровом слое почвы на территории РУП «Учхоз БГСХА». Анализировалась стратифицированная случайная выборка из 64 значений pH. Для генерации пространственного распределения исследуемого показателя использовались три метода интерполяции: метод обратных взвешенных расстояний (IDW), метод радиальных базисных функций (RBF) и ординарный кригинг (OK). Для оценки точности методов интерполяции применялась перекрестная проверка путем сравнения величин средней ошибки (ME) и среднеквадратичной ошибки (RMSE). Установлено, что метод радиальных базисных функций является наиболее пригодным для интерполяции пространственного распределения pH почвы с наименьшим значением RMSE, равным 0,383.

Ключевые слова: pH, прогнозирование, геоинформационные системы, детерминированная интерполяция, геостатистическая интерполяция.

Введение. Устойчивость и стабильность агроэкосистем в значительной мере предопределена буферностью почвы и ее способностью противостоять внешним негативным воздействиям. Кислотно-основное состояние обуславливает многочисленные особенности поведения химических элементов в почвенной экосистеме, с ним связаны режимы органического вещества и элементов минерального питания, подвижность различных соединений, в том числе и токсичных для растений. В кислых малобуферных почвах усиливается растворимость Mn, Fe, B, Cu, Zn, избыток которых негативно влияет на рост, развитие, а в конечном счете и на продуктивность сельскохозяйственных культур (Надточий, Мыслыва, 2014). В основе научной концепции координатного или точного земледелия лежат представления о существовании неоднородностей в пределах одного поля. Для оценки и детектирования этих неоднородностей используются новейшие технологии, в т.ч. и созданные на

базе геоинформационных систем (ГИС) (Якушев, 2016). Управление пространственной изменчивостью, которая широко известна как точное земледелие, имеет важное значение для достижения двойной цели: повышения продуктивности агроэкосистем и снижения темпов их экологической деградации (Zandi et al, 2011). Показатели-константы, главным образом относящиеся к характеристикам почвы, к которым принадлежит и рН почвенного раствора, дают информацию о базовых экологических постоянных.

Важнейшей составляющей системы точного земледелия является измерение пространственной изменчивости свойств почвы с использованием возможностей геопрограммного анализа и дальнейшая визуализация его результатов в виде карт. Геоистатистика является эффективным методом изучения пространственного распределения характеристик почвы и их несогласованности (Zandi et al, 2011; Liu et al, 2014; Vehera, 2015). Геопрограммные методы и сравнение их эффективности при оценке пространственной связи агрохимических свойств почв и географической изменчивости почвенных характеристик применялись целым рядом исследователей. В частности, проблемам оценки пригодности различных моделей интерполяции для изучения пространственного распределения свойств почвы посвящены исследования российских (Симбатова и др., 2016; Samsonova et al, 2017) и украинских (Moskalenko, 2012; Кохан и др., 2013) ученых. Однако, в научной литературе имеется недостаточно сведений о разработке данного направления исследований на территории Беларуси. Поэтому основной целью данного исследования стало выполнение сравнения эффективности методов интерполяции – метода обратных взвешенных расстояний (IDW), метода радиальных базисных функций (RBF) и ординарного кригинга (ОК) для оценки пространственного распределения рН почвы, и применение перекрестной проверки для оценки точности пространственного моделирования и идентификации наиболее подходящего метода.

Объекты и методы исследований. Геопрограммный анализ данных рН_{КС1} почвенного раствора выполнялся с помощью модуля Geostatistical Analyst программного продукта ArcGIS версии 10.2.1. Для анализа использовались данные о рН_{КС1} почвенного раствора, полученные из материалов агрохимического обследования территории РУП «Учхоз БГСХА» (Республика Беларусь, Могилевская область, Горецкий район), выполненного в 2014 г. УКПП «Могилевская областная проектно-исследовательская станция агрохимизации». Общая площадь обследованной территории составляет 6361,67 га. Почвенный покров территории исследований представлен преимущественно дерново-подзолистыми супесчаными на водно-ледниковых супесях и дерново-подзолистыми суглинистыми на лессовидных суглинках почвами.

В настоящем исследовании использованы детерминированные (т.е., создающие поверхности из измеренных значений) и геоистатистические (т.е.,

использующие для построения поверхности статистические свойства измеренных точек) методы интерполяции. Среди детерминированных методов интерполяции использовались метод взвешенных расстояний (Inverse Distance Weighting, IDW) и метод радиальных базисных функций (Radial Basis Functions, RBF), а среди геостатистических – ординарный кригинг (Ordinary Kriging, ОК).

IDW является одним из наиболее часто применяемых детерминированных методов интерполяции в области почвоведения. Его целесообразно использовать, когда набор точек достаточно плотный, чтобы фиксировать степень локального изменения поверхности, необходимую для анализа. IDW определяет значения ячейки с использованием линейно-взвешенного набора комбинаций точек выборки. Назначенный вес является функцией расстояния от входной точки до местоположения выходных ячеек. Чем больше расстояние, тем меньшее влияние на интерполированную ячейку оказывает выходное значение (Childs, 2004). Поскольку IDW не предусматривает вычисление стандартных ошибок интерполяции, обоснование использования этой модели является проблематичным. Интерполированное значение при применении метода IDW определяется по формуле (1):

$$z(x_0) = \frac{\sum_{i=1}^n \frac{x_i}{h_{ij}^\beta}}{\sum_{i=1}^n \frac{1}{h_{ij}^\beta}} \quad (1),$$

где: $z(x_0)$ – интерполированное значение; n – общее количество значений данных выборки; x_i – i -е значение данных; h_{ij} – расстояние между интерполированным значением и значением данных выборки; β – весовое значение.

Радиальные базисные функции (RBF) представляют собой жесткие интерполяторы, которые создают сглаженные поверхности. Они дают хорошие результаты для плавно меняющихся значений. Поскольку интерполяторы являются жесткими, радиальные базисные функции могут быть локально чувствительны к выпадающим значениям (т.е. поверхность будет содержать локально высокие или низкие значения). Метод RBF представляет собой семейство из пяти методов детерминированной точной интерполяции: тонкопленочный сплайн (TPS), сплайн с напряжением (SPT), полностью регуляризованный сплайн (CRS), многокватричная функция (MQ) и обратная многокватричная функция (IMQ). В данном исследовании был выбран наиболее широко используемый метод – CRS.

Ординарный кригинг (ОК) предполагает, что по крайней мере часть пространственных вариаций, наблюдаемых в природных явлениях, можно смоделировать при помощи случайных процессов с использованием пространственной автокорреляции. Приемы работы кригинга могут быть использованы для описания и моделирования пространственных структурных закономерностей, предсказания значений в неизмеренных местоположениях, оценки погрешности, связанной с прогнозируемым значением в неизмеренных

местоположениях. Оценка кригинга $z^*(x_0)$ и дисперсия оценки ошибок $\sigma_k^2(x_0)$ в любой точке x_0 вычисляются следующим образом (2) и (3):

$$z^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i) \quad (2);$$

$$\sigma_k^2(x_0) = \mu + \sum_{i=1}^n \lambda_i \gamma(x_0 - x_i) \quad (3),$$

где: λ_i – веса; μ – постоянная Лагранжа; $\gamma(x_0 - x_i)$ – значение вариограммы, соответствующее расстоянию между x_0 и x_i .

В качестве основного инструмента для изучения структуры пространственного распределения содержания гумуса в почве использовались семивариограммы. Основываясь на региональной теории вариаций и внутренних гипотезах (Gouri et al., 2016), семивариограмма выражается следующим образом (4):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (4),$$

где: $\gamma(h)$ – полувариантность, h – расстояние запаздывания, Z – параметр свойства почвы, $N(h)$ – количество пар мест, разделенных расстоянием лага h , $Z(x_i)$ и $Z(x_i + h)$ – значения Z в положениях x_i и $x_i + h$ (Wang, Shao, 2013).

Результаты и обсуждение. Применение модуля «Геостатистический анализ» для пространственного моделирования рН_{KCl} почвенного раствора предусматривает предварительную оценку исходных данных на предмет их пригодности для целей моделирования. В результате применения инструмента «Анализ данных» данного модуля создается гистограмма распределения исходных данных и исследуется форма их распределения, а также рассчитываются основные статистические характеристики выборки, сведения о которых содержатся в табл. 1.

1. Статистические показатели выборки данных о рН_{KCl} почвенного раствора, используемой для оценки моделей интерполяции

Название показателя	Число наблюдений	Значение показателя			Sd	Cv, %	Med	Экссесс	Асимметрия
		min	max	mid					
Значение показателя	64	5,2	7,0	6,07	0,47	7,76	6,05	0,34	0,04

Примечание: Sd – среднеквадратическое отклонение; Cv – коэффициент вариации; Med – медиана.

Предварительная оценка данных позволяет установить необходимость проведения их преобразования. В частности, если распределение данных имеет несколько пиков (экстремумов), т.е. данные распределены асимметрично, к ним применяется логарифмическое преобразование, которое приближает распределение к нормальному. В нашем случае выполнение преобразования нецелесообразно, поскольку распределение данных выборки близкое к нормальному, а среднее значение и медиана довольно близки по значениям (рис. 1). В частности, установлено, что фактическое значение λ -критерия Колмогорова-Смирнова для уровня значимости 0,05 составляет 0,083, а его теоретическое значение - 1,358. Величина коэффициента вариации 7,76 %

также свидетельствует о том, что выборка исходных данных достаточно однородная.

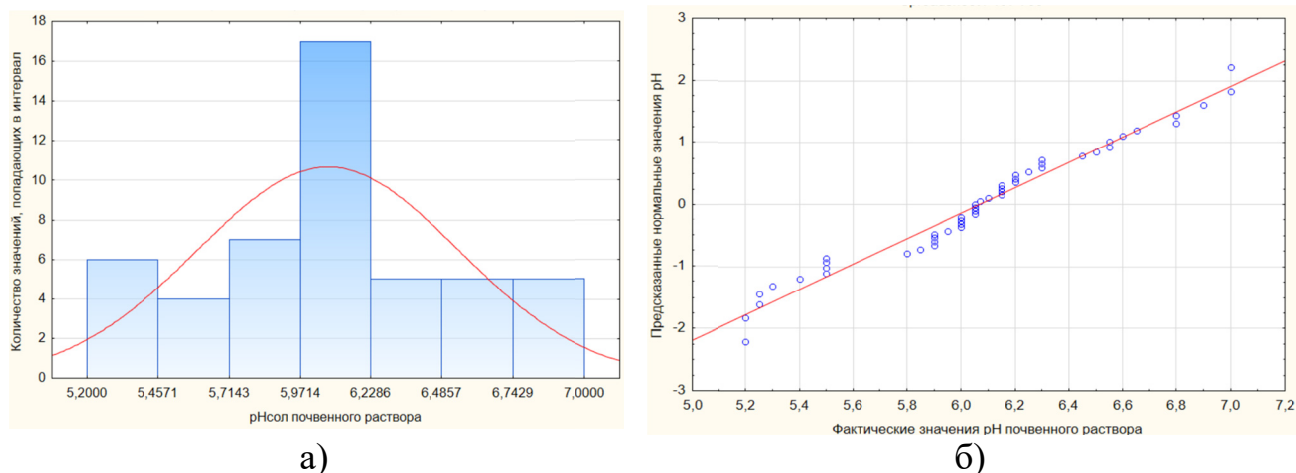


Рис. 1. Гистограмма частот (а) и нормальный график распределения данных (б) о рН почвенного раствора, слой почвы 0-20 см

Инструмент «Анализ тренда» модуля «Геостатистический анализ» позволяет отображать данные в трехмерной перспективе. Местоположения опорных точек наносятся на плоскость x, y . Уникальной особенностью данного инструмента является то, что значения проецируются на перпендикулярные плоскости $x-z$ и $y-z$ в виде диаграмм рассеивания. Затем на проецируемых плоскостях выполняется подгон полиномов с помощью диаграмм рассеивания. Линия наилучшего соответствия (полином), проведенная через проецируемые точки, показывает тренды изменения данных в определенных направлениях (рис. 2). В нашем случае наблюдается определенный тренд как в направлении запад-восток, так и в направлении север – юг. Поскольку тренд имеет U-образную форму, при проведении интерполяции целесообразно использовать полином первого порядка в качестве глобальной модели тренда и использовать опцию «Порядок удаления тренда» при построении моделей методом ординарного кригинга.

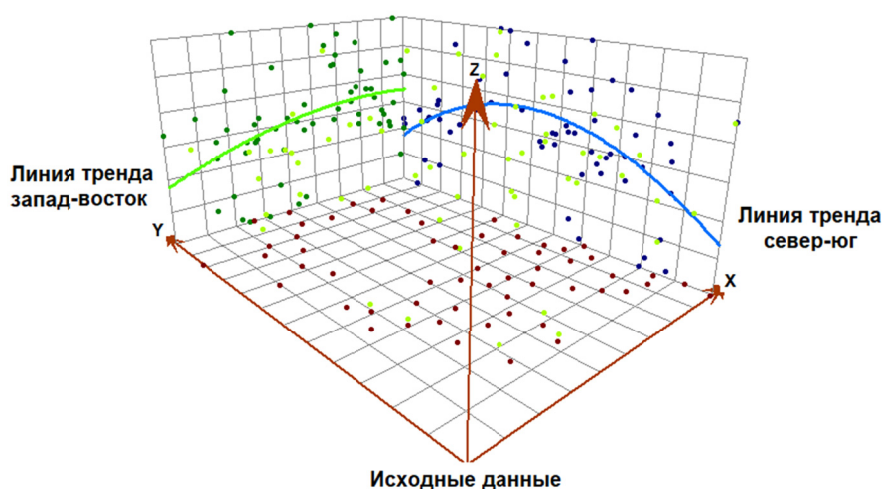


Рис. 2. Тренд пространственного распределения рН почвенного раствора

Для определения возможной пространственной структуры рН почвы были рассчитаны экспериментальные анизотропные вариограммы. По результатам кросс-валидации, результаты которой представлены в табл. 2, в качестве лучшей модели вариограммы была идентифицирована экспоненциальная функция.

2. Результаты кросс-валидации моделей вариограммы

Модель вариограммы	ME	RMSE	RMSS	AN	ASE
Сферическая	0,022	0,402	1,155	0,055	0,345
Экспоненциальная	0,017	0,400	1,151	0,041	0,346
Гауссова	0,023	0,403	1,156	0,059	0,346

Лучшая модель была выбрана на основе пяти критериев: средней ошибки (ME), среднеквадратичной ошибки (RMSE), среднеквадратичной нормированной ошибки (RMSS), средней нормированной ошибки (AN) и средней стандартной ошибки (ASE) В этом исследовании экспоненциальная модель вариограммы в большинстве случаев удовлетворяет требованиям, предъявляемым к наилучшей модели.

Исследуемые методы интерполяции были реализованы для оценки значений неизмеренных данных об уровне рН почвенного раствора и создания поверхностей из проинтерполированных точек. Параметры, используемые в моделях интерполяции, примененных для создания карт прогнозирования рН почвы, представлены в табл. 3, а визуализация результатов прогнозирования – на рис. 3.

3. Параметры методов интерполяции, используемых для создания карт прогноза рН почвы (область поиска: $\min = 2$; $\max = 15$)

Метод обратных взвешенных расстояний (IDW)	Метод радиальных базисных функций (RBF)	Ординарный кригинг (OK)
Параметр мощности: 2,5335	Функция ядра: полностью регуляризованный сплайн; параметр ядра: 0,020	Модель вариограммы: экспоненциальная

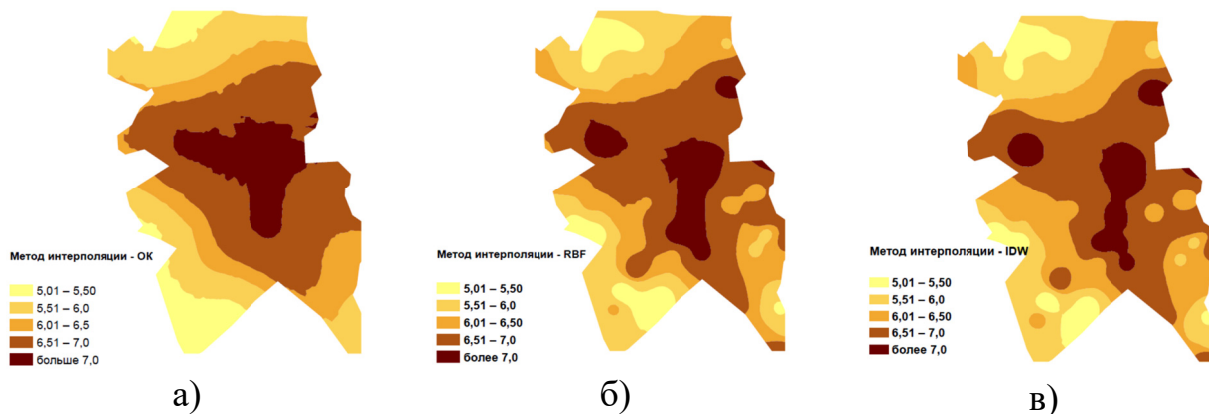


Рис. 3. Карты прогноза рН_{KCl} почвы, созданные: (а) ОК; (б) RBF и (в) IDW

На рис. 4 показаны результаты перекрестной проверки прогнозных

моделей, генерируемых методами ординарного кригинга, радиальных базисных функций и обратных взвешенных расстояний.

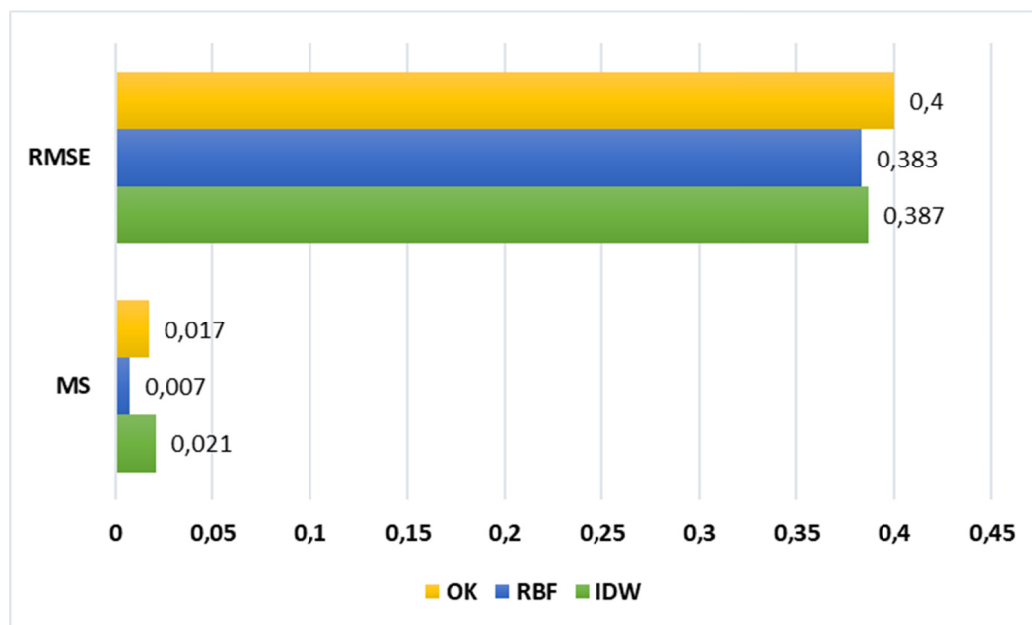


Рис. 4. Результаты кросс-валидации прогнозных моделей

Поскольку для методов интерполяции RBF и IDW при выполнении процедуры кросс-валидации определяются только два критерия – средняя ошибка (ME) и среднеквадратичная ошибка (RMSE), сравнение эффективности методов для целей прогнозирования рН почвы проводили только по этим показателям. Наиболее пригодной для целей прогнозирования оказалась модель, созданная методом радиальных базисных функций, значение среднеквадратичной ошибки которой составило 0,383. Метод RBF подходит к поверхности через измеренные значения проб при минимизации общей кривизны поверхности и неэффективен, когда происходит резкое изменение значений поверхности на коротких расстояниях. В нашем случае такое явление отсутствует, а изменение исследуемого показателя носит достаточно плавный характер (см. табл. 1 и рис. 1), что и обуславливает высокую эффективность метода RBF при интерполяции. По точности интерполяции исследованные детерминированные и геостатистические методы расположились в следующий убывающий ряд RBF > IDW > ОК.

Выводы. Проведенными исследованиями установлено следующее: 1) среди исследованных методов интерполяции наиболее оптимальным для прогнозирования и картографирования пространственного распределения рН почвы является метод радиальных базисных функций, обеспечивающий среднеквадратическую ошибку среднего прогнозируемых значений на уровне 0,383; 2) как метод интерполяции IDW, так и метод RBF требуют настройки параметра мощности и определения оптимального радиуса поиска для повышения точности прогноза; 3) хотя метод IDW относительно прост в использовании и не предусматривает вычисления стандартных ошибок

интерполяции, в нашем случае он обеспечил более высокую точность при прогнозе пространственного распределения рН почвы по сравнению с ординарным кригингом.

Дальнейшие исследования целесообразно сосредоточить в направлении сравнения результатов интерполяции, полученных для рН пахотного и подпахотного горизонтов почвы.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

Кохан С. С. Геоінформаційне забезпечення якісної оцінки ґрунтів / С. С. Кохан, А. А. Москаленко, Л. Г. Шило // Східно-Європейський журнал передових технологій. – 2013. – № 6. – С. 18-25.

Kohan S. S., Moskalenko A. A., Shylo L. G., 2013, "Geoinformation providing of qualitative estimation of soils", East European Magazine of Advanced Technology, No. 6, pp. 18-25.

Надточий П. П. Эталонные величины кислотно-основной буферности дерново-подзолистых почв для фонового мониторинга / П. П. Надточий, Т. Н. Мысльва // Агрохимия. – 2014. – № 3. – С. 83-89.

Nadtochij P. P., Myslyva T. N., 2014, "Reference values of acid-base buffering of sod-podzolic soils for background monitoring", Agrochemistry, No. 3, pp. 83-89.

Симбатова А. Т. Моделирование пространственного распределения органического вещества почв: обзор современных подходов / А. Т. Симбатова, С. С. Рязанов, И. А. Сахабиев // Рос. журнал прикладной экологии. – 2016. – № 2. – С. 48-54.

Simbatova A. N., Ryazanov S. S., Sahabiev I. A., 2016, "Modeling of spatial distribution of soil organic matter: an overview of modern approaches", Russian Journal of Applied Ecology, No. 2, pp. 48-54.

Якушев В. В. Точное земледелие: теория и практика / В. В. Якушев. – СПб.: ФГБНУ АФИ, 2016. – 364 с.

Yakushev V. V., 2016, "Precise farming: theory and practice", St. Petersburg, FGBNU AFI, 364 p.

Behera S. K., Shukla A. K., 2015, "Spatial distribution of surface soil acidity, electrical Conductivity, soil organic carbon content and exchangeable Potassium, calcium and magnesium in some cropped acid Soils of India, Land Degrad. Dev, Vol. 16, pp. 71-79.

Childs C., 2004, "Interpolation surfaces in ArcGIS Spatial Analyst", ArcUser, Vol. 3, pp. 32-35.

Gouri S. B., Shit P. K., Maiti R., 2016, "Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC)", Journal of the Saudi Society of Agricultural Sciences, Vol. 2, pp. 1-13.

Moskalenko A. A. Justification of geoinformation system on land soils quality monitoring / A. A. Moskalenko // Землеустрій, кадастр і моніторинг земель. – 2012. – № 3-4. – С. 108-112.

Moskalenko A. A., 2012, "Justification of geoinformation system on land soils quality monitoring", Land system, cadastre and land monitoring, No. 3-4, pp. 108-112.

Samsonova V. P., Blagoveshchenskii Y. N., Meshalkina Y. L., 2017, "Use of empirical bayesian kriging for revealing heterogeneities in the distribution of organic carbon on agricultural lands", Eurasian Soil Science, Vol. 50, No. 3, pp. 305-311.

Liu L., Wang H., Dai W. and all, 2014, "Spatial variability of soil organic carbon in the forestlands of northeast China", J. Forest. Res., Vol. 25(4), pp. 867-876.

Wang Y. Q., Shao M. A., 2013, "Spatial variability of soil physical properties in a region of the Loess Plateau of PR China subject to wind and water erosion", Land Degrad. Dev., Vol. 24(3), pp. 296-304.

Zandi S., Ghobakhlou A., Sallis P., 2011, "Evaluation of spatial interpolation techniques for mapping soil pH", 19th International Congress on Modelling and Simulation, Perth, Australia, 12-16 December 2011, pp. 1153-1159.