

УДК 004.896

ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ВОССТАНОВЛЕНИЯ ДВУМЕРНЫХ ГЕОПОЛЕЙ

Каковкин Павел Александрович,

студент кафедры вычислительной техники Института кибернетики,
Национального исследовательского Томского политехнического университета,
Россия, 634050, г. Томск, пр. Ленина, 30. E-mail: nitrokot@mail.ru

Друки Алексей Алексеевич,

ассистент кафедры вычислительной техники Института кибернетики,
Национального исследовательского Томского политехнического университета,
Россия, 634050, г. Томск, пр. Ленина, 30. E-mail: druki2008@yandex.ru

Актуальность исследования обусловлена необходимостью разработки алгоритмов и программных средств, позволяющих повысить эффективность пространственного анализа и восстановления двумерных геополей.

Цель работы: разработка алгоритмов и программных средств, позволяющих повысить эффективность восстановления двумерных геополей по не регулярной сети точечных данных; проведение экспериментальных исследований эффективности разработанных алгоритмов и сравнение с существующими аналогами.

Методы исследования. Для решения поставленной задачи применяются методы искусственного интеллекта, методы реализации искусственных нейронных сетей и генетических алгоритмов, комитетные методы – для решения задач классификации, математическое моделирование, теория вероятностей и математическая статистика с использованием программного обеспечения VisualStudio и Surfer.

Результаты. Для восстановления геополей по нерегулярной сети точечных данных было решено использовать методы искусственного интеллекта, так как данная область исследований на сегодняшний день является одной из наиболее интенсивно развивающихся. Таким образом, для решения поставленной задачи был разработан алгоритм, основанный на искусственных нейронных сетях. Алгоритм представляет собой последовательность действий, состоящую из семи шагов. Для восстановления геополей применяются несколько нейронных сетей прямого распределения типа персептрон, работающих по методу бэггинга (bagging). На основе предложенного алгоритма разработано программное приложение, позволяющее выполнять восстановление геополей по нерегулярной сети точечных данных. Выполнено экспериментальное исследование эффективности разработанного алгоритма и проведено сравнение с результатами работы метода обратно-взвешенных расстояний. Проведенные исследования показали, что результаты работы предложенного алгоритма превышают эффективность работы метода обратно-взвешенных расстояний.

Ключевые слова:

Геоинформатика, восстановление геополей, методы геостатистики, кросс-валидация, искусственные нейронные сети.

Восстановление геополя

На сегодняшний день одной из наиболее интенсивно развивающихся и актуальных областей в геоинформатике является анализ пространственных данных, описывающих непрерывные поверхности. Данные поверхности могут описывать распределение в пространстве высот рельефа местности, температуры, давления и т. д. Поверхности, которые описываются функцией от двух пространственных координат, в геоинформатике называются двумерными геополями [1–2].

Цифровая модель геополя представляет собой набор пространственных данных, описывающих непрерывные поверхности. Пространственный анализ геополей предполагает анализ и сопоставление разнородных пространственных данных и выявление взаимосвязи между ними [3]. Для этого применяются различные методы и алгоритмы, позволяющие формировать карты признаков геополей, определять различные пространственные закономерности исследуемых поверхностей.

Важным классом задач в данной области исследований являются задачи восстановления поверхности геополей, которые сводятся к поиску значе-

ний поверхности в точках, где измерения не проводились по каким-либо причинам [3]. Их решением занимались многие российские и зарубежные ученые, среди которых можно отметить М.Ф. Каневского [4], В.И. Аронова, О.Р. Мусина, В.В. Демьянова [4], D. Vatson, V. Franklin [5], D. Douglas и др. Однако многие предложенные ими методы и алгоритмы восстановления оказываются недостаточно эффективными в реальных условиях. Поэтому задача разработки методов и алгоритмов, позволяющих повысить эффективность восстановления геополей, остается актуальной на сегодняшний день.

Методы восстановления поверхности геополей

Задача восстановления поверхности геополя по нерегулярной сети точечных данных предполагает наличие исходных данных в виде сети точек с координатами $s(x_i, y_i)$, $i=1, \dots, n$ в двумерном пространстве и значение геополя f для каждой точки, которое необходимо восстановить [6].

Основная сложность решения данной задачи заключается в том, что обычно такая поверхность представляет собой набор точечных данных в виде нерегулярной сети, т. е. необходимая информация

об исследуемых объектах известна не во всех областях данной поверхности (рис. 1).

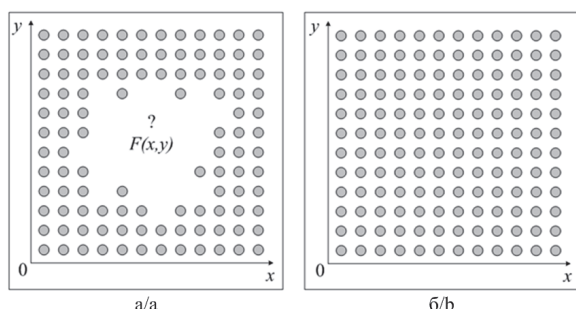


Рис. 1. Точечные данные, описывающие поверхность геополья: а) регулярная сеть точек; б) нерегулярная сеть точек

Fig. 1. Data points describing the surface of a geofield: а) regular network of points; б) irregular network of points

Данную задачу невозможно решить, если не иметь дополнительную информацию о природе и параметрах исследуемого геополья. На сегодняшний день существуют различные методы, применяемые в задачах анализа и восстановления геополья, основные из них: детерминистические и геостатистические.

Среди детерминистических методов, которые также называются интерполяторами, особенно популярна сплайн-интерполяция. Данный вид относится к классу глобальной интерполяции и применяется там, где исследуемая поверхность может быть описана детерминистической функцией на всей исследуемой области.

Однако эффективность данных методов снижается при наличии большого количества исходных данных (десятки и сотни тысяч точек) и если время вычисления детерминистической функции квадратично зависит от количества исходных данных. Поверхность, восстановленная с помощью глобальных интерполяторов, как правило, является заметно сглаженной, что также является недостатком. Также в данном случае возникает необходимость обоснования выбора интерполятора, что требует дополнительной информации о природе восстанавливаемой поверхности, а очень часто такие сведения слабо формализованы. Кроме того, в данных методах используется ряд параметров, которые требуют тщательной и осмысленной настройки.

Как показывает анализ предметной области, наиболее востребованными и эффективными являются геостатистические методы [7].

Геостатистические методы основаны на статистическом анализе пространственных данных и предназначены для построения вариограммы и восстановления неизвестного значения точки в исследуемой области.

Вариограмма описывает степень различия точечных данных в пространстве, основываясь на расстоянии между ними. С учетом пространственной непрерывности геополья два более удаленных друг от друга значения геополья будут менее коррелированы, чем два значения, находящиеся на малом расстоянии, т. е. чем ближе значения геополья (чем меньше разность между ними), тем больше значение вариограммы. Конечная цель применения метода вариографии заключается в построении теоретической модели вариограммы, т. е. в аналитической функции, которая описывает пространственные данные [7–8].

Далее следует этап восстановления геополья методом кригинга (от фамилии южноафриканского геолога Д. Криге). В этом геостатистическом методе интерполяции используются статистические параметры, которые были получены при построении вариограммы.

Для тестирования результатов восстановления геополья используется процедура кросс-валидации, которая заключается в следующем. Из исходного набора пространственных данных удаляется точка. С использованием оставшихся точек восстанавливается значение геополья в исключенной точке [9] (рис. 2).

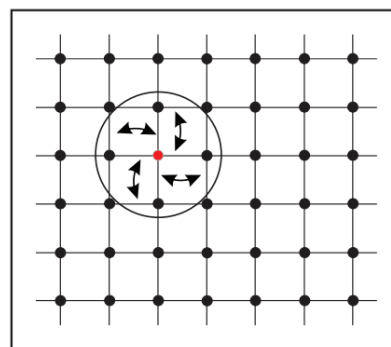


Рис. 2. Восстановление геополья посредством выполнения кросс-валидации

Fig. 2. Geofield restoring by cross-validation

После этого вычисляется величина среднеквадратичного отклонения между полученным значением геополья и исходным. Затем данная процедура повторяется для другой точки и т. д. Итоговая оценка кросс-валидации вычисляется как среднее значение ошибки по всем точкам исследуемого пространства [10]. Кросс-валидация является процессом, требующим больших временных затрат.

На основании экспериментальных исследований был сделан вывод, что при увеличении количества исходных точек более 700 время обработки данных при кросс-валидации стремительно увеличивается и становится недопустимо большим. Данные результаты получены на основе проведения 30 вычислительных экспериментов.

Процесс восстановления геополья методом кригинга реализуется следующим образом:

- 1) методом вариографии определяются входные параметры для метода кригинга: ранг, порог, эффект самородка и теоретическая модель вариограммы;
- 2) вычисляется значение геополья;
- 3) выполняется проверка результатов восстановления с использованием кросс-валидации;

4) если полученные результаты не удовлетворяют требованиям и ошибка восстановления слишком велика, то выполняется переход к шагу 1 [11].

Недостаток данного подхода заключается в длительности, рутинности и трудоемкости всего процесса. Значительная часть временных затрат приходится на проведение кросс-валидации [12].

На рис. 3 представлены результаты 20 численных экспериментов времени выполнения кросс-валидации в зависимости от количества исходных точек. При проведении вычислительных экспериментов не использовались методы параллельных вычислений. По рисунку видно, что при увеличении количества исходных точек до 1000 время выполнения одной итерации кросс-валидации резко увеличивается до не приемлемого уровня.

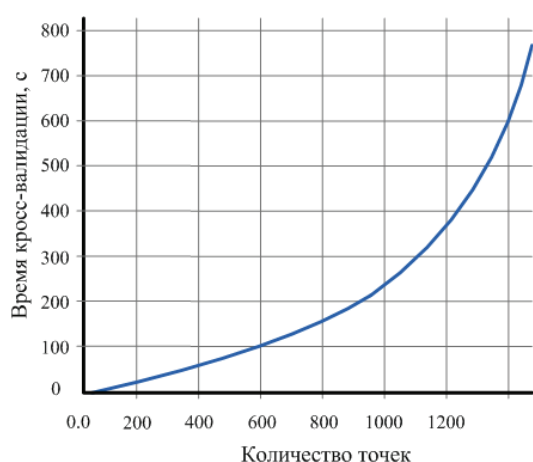


Рис. 3. График зависимости времени выполнения кросс-валидации от количества точечных данных

Fig. 3. Graph of the cross-validation time on the point data number

В результате проведенного анализа и выполнения вычислительных экспериментов можно сделать вывод, что одними из главных недостатков восстановления геополей методами геостатистики являются невысокая точность и большая трудоемкость процесса восстановления.

Исходя из недостатков вышеописанных методов, для решения задачи восстановления поверхности геополей было решено использовать методы искусственного интеллекта, а именно искусственные нейронные сети (ИНС). Данный выбор был сделан исходя из возможности настройки оптимальных параметров ИНС (процесс обучения), что может сделать процесс восстановления геополя более автоматизированным, снизить трудоемкость и повысить его точность.

Разработка алгоритма восстановления геополей на основе искусственных нейронных сетей

На рис. 4 представлена обобщенная схема предложенного алгоритма восстановления геополей.

Рассмотрим каждый этап предложенного алгоритма.

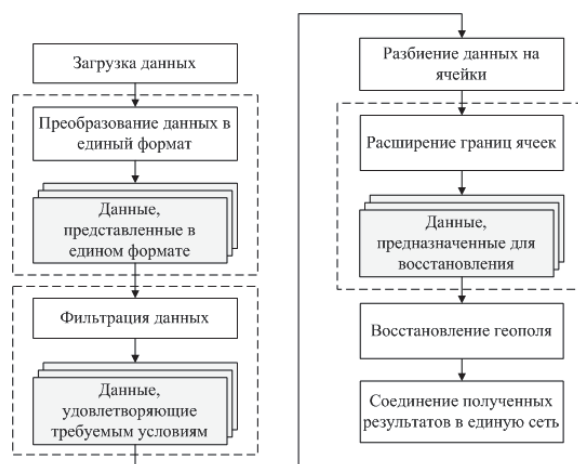


Рис. 4. Обобщенная схема предложенного алгоритма

Fig. 4. The generalized scheme of the proposed algorithm

1. **Загрузка исходных данных.**
2. **Преобразование данных.** Данные, представленные в различных форматах, преобразовываются в единый формат. Такой формат предполагает представление каждого элемента исходного набора данных в виде параметров s и f , которые обозначают координаты точки (x, y) в двумерном пространстве и значение геополя для данной точки [13].
3. **Фильтрация данных.** Из набора полученных данных удаляются точки, которые расположены вне исследуемой области и параметры которых не соответствуют заданным условиям. Далее находятся группы точек, расстояние между которыми меньше заданной величины и заменяются одной точкой. Значение геополя в данной области вычисляется в соответствии с заданным правилом.
4. **Разбиение пространственных данных на ячейки с равномерным заданным шагом** (рис. 5). Данное действие осуществляется для более тщательного последующего анализа исходных данных в каждой отдельной ячейке и сокращения вычислительных затрат [14]. Такие параметры, как количество и величина ячеек, в значительной степени могут влиять на качество восстановления геополя. Данные параметры задаются в интерфейсе пользователя.

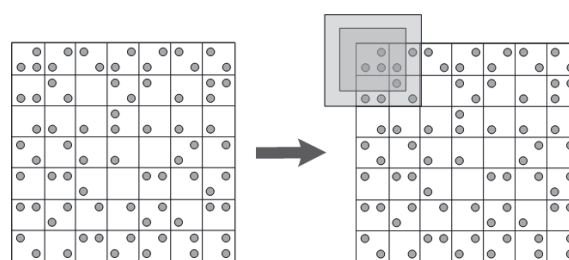


Рис. 5. Разбиение сети точечных данных на ячейки и последующее расширение границ ячеек

Fig. 5. Splitting a point data network on cells and subsequent expansion of the cells boundaries

5. *Расширение границ ячеек.* Данный этап предназначен для учета точек, лежащих в соседних ячейках. Это осуществляется для расширения анализа исследуемой области и получения более сглаженных результатов. Если в ячейке находится менее трех точек, то выполняется расширение границ ячейки до тех пор, пока исследуемая область не будет охватывать необходимое количество точек. Данные параметры также задаются в интерфейсе пользователя.
6. *Восстановление геополя.* С помощью ИНС вычисляется значение регулярной сети точечных данных в каждой ячейке.
7. После вычисления точечных данных во всех ячейках, выполняется соединение полученных результатов в единую сеть.

Разработка нейронной сети для восстановления точечных данных

Искусственная нейронная сеть (ИНС) представляет собой математическую, программную или аппаратную модель, функционирующую по принципу биологических нервных клеток человеческого организма. ИНС имеют иерархическую структуру и состоят из простых взаимодействующих между собой процессоров, называемых нейронами, которые сгруппированы в слои. Каждый нейрон имеет входные и выходные связи, ведущие к нейронам соседних слоев. Данные связи называются синаптическими и характеризуются определенными весовыми коэффициентами [15–16].

Значение нейрона при прямом проходе по ИНС определяется как взвешенная сумма его входных значений:

$$S = \sum_{i=1}^N x_i w_{ij},$$

где x_i – текущий выход i -го нейрона; w_{ij} – весовой коэффициент между i -м и j -м нейронами [16].

Каждый нейрон определяется функцией активации, которая предназначена для определения зависимости сигнала на выходе нейрона от взвешенной суммы сигналов на его входах [16]. Пример стандартной логистической функции активации:

$$f(x) = \frac{1}{(1 + e^{-x})},$$

где s – взвешенная сумма входных значений нейрона.

Для всех данных из обучающей выборки, поступающих на вход нейронной сети, определяются их требуемые значения, которые затем сравниваются с текущими значениями, и вычисляется ошибка. Основным принцип обучения ИНС заключается в сведении к минимуму эмпирической ошибки классификации [16]:

$$E = \frac{1}{2} \sum_{i=0}^N (t_i - x_i)^2,$$

где x_i – действительный выход i -го нейрона; t_i – желаемый выход i -го нейрона.

После этого на основе метода градиентного спуска вычисляется величина коррекции синаптических коэффициентов между нейронами [16]:

$$\Delta w_{ij} = -\eta \delta_i x_j,$$

где δ_i – значение ошибки нейрона i ; x_j – текущий выход j -го нейрона; η – коэффициент скорости обучения, $0 < \eta < 1$.

Весовые коэффициенты изменяются до тех пор, пока ошибка не сведется к минимуму и станет меньше заданного порога.

При разработке оптимальной структуры нейронной сети и проведении вычислительных экспериментов было решено реализовать ансамбль нейронных сетей, работающих по методу бэггинга (bagging) [17]. Данный метод представляет собой совокупность отдельных базовых классификаторов, когда комитетное решение принимается на основе голосования или взвешенным усреднением решений отдельных классификаторов.

Метод бэггинга предполагает обучение классификаторов по принципу бутстреп (bootstrap), когда обучающая выборка разделяется на M равных непересекающихся частей, по количеству используемых классификаторов. Таким образом, каждый классификатор обучается на различных множествах данных. Также в данном подходе целесообразно применять классификаторы, имеющие различную архитектуру и конфигурацию.

Для решения поставленной задачи были разработаны пять нейронных сетей прямого распространения класса персептрон (рис. 6).

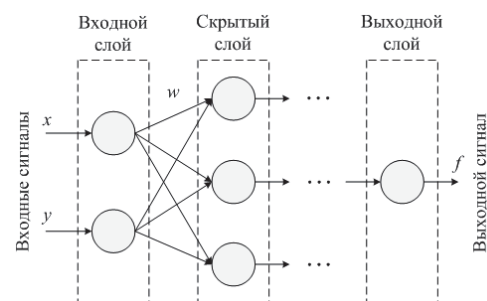


Рис. 6. Нейронная сеть прямого распространения

Fig. 6. Neural network of direct propagation

Разработанные нейронные сети имеют различную конфигурацию (таблица).

Таблица. Конфигурация нейронных сетей

Table. Configuration of neural networks

ИНС ANN	Количество нейронов в слоях Amount of neurons in the network					Функция активации Activation function
	1	2	3	4	5	
1	2	3	1	–	–	Лог-сигмоидная Log-sigmoid
2	2	3	3	3	1	
3	2	5	10	1	–	
4	2	5	5	10	1	Гиперболический тангенс Hyperbolic tangent
5	2	10	20	20	1	

ИНС – ANN (artificial neural network).

Входной слой нейронных сетей содержит два входа, на которые поступают значения (x, y) в двумерном пространстве. Выходной слой содержит один выход, на котором высчитывается значение геополя f в каждой точке пространства. Исходные значения синаптических весовых коэффициентов инициализированы случайным образом из диапазона от $-0,1$ до $+0,1$. Начальная скорость обучения выбиралась в диапазоне от 0 до 1. В качестве функции ошибки ИНС использовалось среднеквадратичное отклонение.

Для обучения нейронных сетей использовался алгоритм обратного распространения ошибки (Back Propagation) [18], его модификации (Resilient propagation и Quick propagation) [19], а также несколько вариантов генетических алгоритмов (Genetic algorithm) [20] с различными параметрами.

Вычисление итогового значения осуществляется на основе сравнения результатов, полученных каждым классификатором. Если результаты классификаторов отличаются друг от друга на величину меньше заданного порога, то выполняется их взвешенное усреднение. Иначе осуществляется комитетный принцип голосования, когда решение принимается на основе большинства классификаторов, показавших схожий результат (т. е. результаты, отличающиеся друг от друга на величину меньше заданного порога).

Исследование результатов работы алгоритма

В процессе исследований выполнялось сравнение показателей точности восстановления геополей предложенным алгоритмом с классическим методом обратно-взвешенных расстояний. При проведении вычислительных экспериментов были сформированы различные наборы нерегулярных точечных данных. Некоторые наборы являются реальными данными, которые были получены из различных источников, другие были сформированы случайным образом. Для каждого сформированного набора выполнялось восстановление геополя. Регулярная сеть имеет размер 50×50 ячеек. Для каждой восстановленной точки регулярной сети выполнялся расчет ошибки для двух сравниваемых методов.

На рис. 7 представлена динамика изменения ошибки восстановления геополя в зависимости от количества исходных точечных данных для сравниваемых методов.

Результаты получены на основе 30 численных экспериментов. Точность правильного восстановления геополя составляет около 93 %. Из рис. 7 следует, что ошибка восстановления геополя предложенным методом меньше ошибки, показанной методом обратно-взвешенных расстояний. Однако при использовании предложенного метода, если количество исходных точечных данных уменьша-

ется, ошибка стремительно возрастает, но при этом её значение остается ниже ошибки сравниваемого алгоритма.

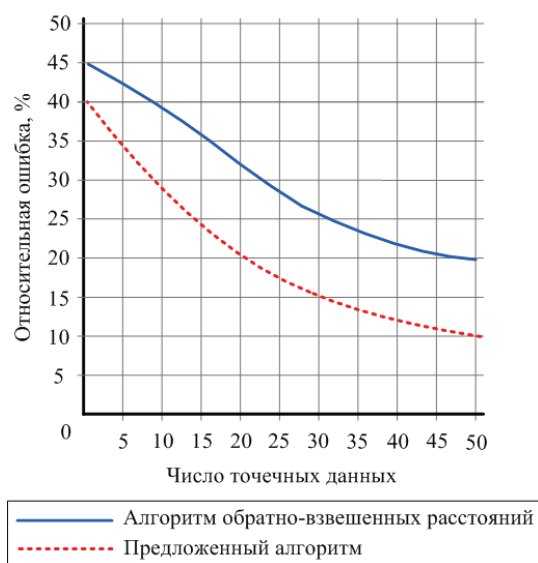


Рис. 7. Зависимость относительной ошибки восстановления геополя от количества точечных данных

Fig. 7. Dependence of the recovery geofield error on the number of data points

На рис. 8 представлен эталонный пример геополя и примеры восстановления данного геополя с помощью предложенного алгоритма и метода обратно-взвешенных расстояний.

По представленным выше рисункам можно сделать вывод, что оба алгоритма по качеству восстановления геополей примерно равны. Однако при более тщательном анализе можно заметить, что предложенный алгоритм показал более близкий к оригиналу вариант с достаточно подробным восстановлением мелких областей пространственных данных. Метод обратно-взвешенных расстояний показал более сглаженный результат с большей степенью усреднения некоторых областей пространственных данных. Однако скорость работы предложенного алгоритма в несколько раз ниже, чем у метода обратно-взвешенных расстояний. Средняя скорость работы предложенного алгоритма составила 5,2 с, скорость работы метода обратно-взвешенных расстояний — 0,47 с. При проведении вычислительных экспериментов не использовались методы параллельных вычислений.

Технические характеристики персонального компьютера, использованного для вычислительных экспериментов: процессор — Intel Pentium Quad Core 2.16 ГГц; ОЗУ — 2048 МБ; видеосистема — Intel GMA HD 512 Мб; операционная система — Microsoft Windows 8, 64-разрядная.

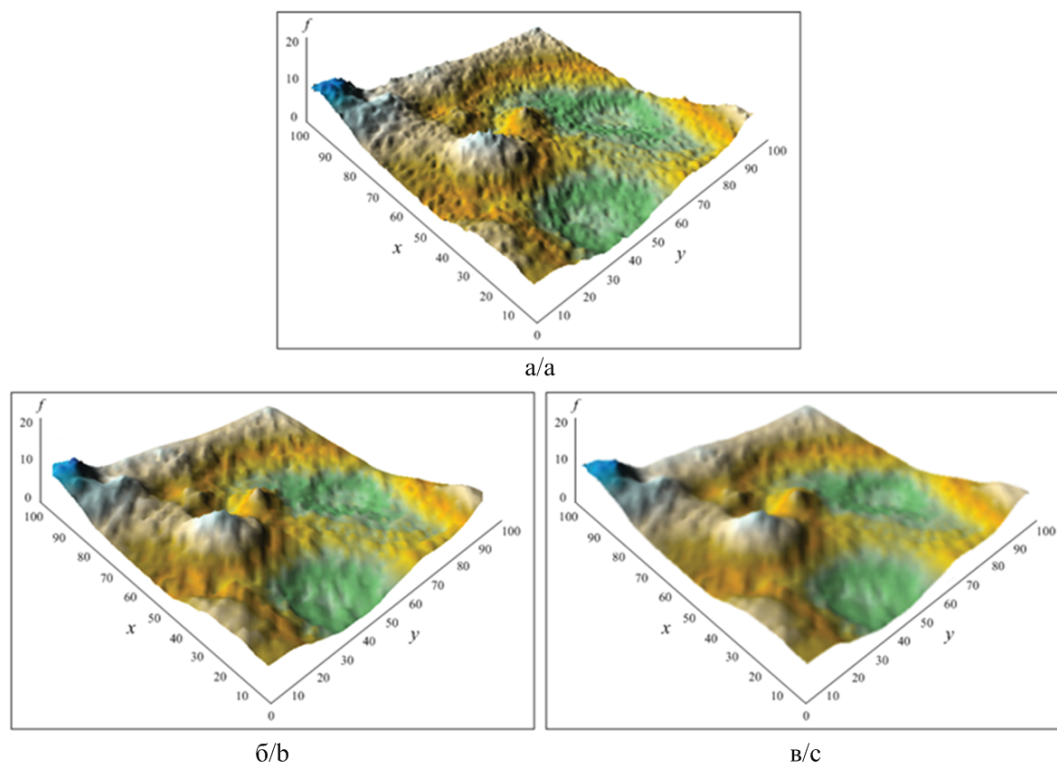


Рис. 8. Сравнение эффективности алгоритмов: а) эталонный пример геополя; б) предложенный алгоритм; в) метод обратно-взвешенных расстояний

Fig. 8. Comparison of algorithms efficiency: a) geofield benchmark; b) proposed algorithm; c) method of inverse-distance weighted

Заключение

На сегодняшний день существуют различные методы и алгоритмы восстановления геополей по точечным данным. Наибольшее распространение и популярность получили детерминистические и геостатистические методы пространственной интерполяции. Данные методы показывают высокие результаты при восстановлении геополей, требуют значительную трудоемкость процесса восстановления. Для решения этих проблем был предложен алгоритм восстановления геополей, основанный на искусственных нейронных сетях.

Результаты вычислительных экспериментов показали, что предложенный алгоритм обеспечи-

вает более высокую точность восстановления геополей, в отличие от метода обратно-взвешенных расстояний. К недостаткам предложенного алгоритма можно отнести низкую скорость работы, сложность выбора оптимальной конфигурации ИНС и необходимость обеспечения эффективного процесса их обучения. Несмотря на имеющиеся недостатки реализованного алгоритма, можно сделать вывод, что применение ИНС и методов искусственного интеллекта в данной области исследований является актуальным и перспективным направлением, а также требует проведения дальнейших исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Капралов Е.Г., Кошкарёв А.В., Тикунов В.С. Геоинформатика. – М.: Академия, 2010. – 480 с.
2. Журкин И.Г., Шайтура С.В. Геоинформационные системы. – М.: КУДИЦ-Пресс, 2009. – 272 с.
3. Ковин Р.В., Марков Н.Г. Геоинформационные системы и технологии. – Томск: Изд-во ТПУ, 2009. – 267 с.
4. Основные понятия и элементы геостатистики / М.Ф. Каневский, В.В. Демьянов, Е.А. Савельева, С.Ю. Чернов // Проблемы окружающей среды и природных ресурсов. – 1999. – № 11. – С. 15–21.
5. Gousie M.B., Franklin W.R. Augmenting Grid-Based Contours to Improve Thin Plate DEM Generation // Photogrammetric Engineering & Remote Sensing. – 2005. – V. 71 (1). – P. 69–79.
6. Копнов М.В., Марков Н.Г. Восстановление двумерных геополей методами геостатистики // Проблемы информатики. – 2011. – № 2 (10). – С. 36–43.
7. Ковин Р.В., Копнов М.В., Марков Н.Г. Особенности поиска значений параметров при автоматизированном построении прогнозных карт // Современные техника и технологии: сборник трудов XVI Международной научно-практической конференции студентов, аспирантов и молодых ученых. – Томск, 2010. – Т. 2 – С. 342–343.
8. Ермолаев В.А. Интерполяционные восстанавливающие фильтры: метод динамической интерполяции // Приборы и системы. Управление, контроль, диагностика. – 2002. – № 2. – С. 39–42.
9. Новаковский Б.А., Прасолов С.В., Прасолова А.И. Цифровые модели рельефа реальных и абстрактных геополей. – М.: Научный мир. – 2003. – 64 с.

10. Almeida J.A., Lopes M. Stochastic Simulation of Rainfall Using a Space-time Geostatistical Algorithm // *Geostatistics for Environmental Applications*. – 2005. – P. 455–466.
11. Wim C.M. Kriging Metamodeling in Discrete-Event Simulation // *Proceedings of the 2005 Winter Simulation Conference*. – Orlando, USA, 2005. – P. 345–348.
12. Hendrick W., Dhaene T. Sequential Design and Rational Metamodeling // *Proceedings of the 2005 Winter Simulation Conference*. – Orlando, USA, 2005. – P. 368–373.
13. Russo A., Soares A., Pereira M.J. Joint Space-Time Geostatistical Model for Air // *Proceedings of the VII Geostatistics for Environmental Applications Conference*. – Netherlands, 2007. – P. 173–185.
14. Березовский В.Е., Друки А.А. Разработка и анализ приложений, позволяющих решать задачу восстановления геопольа по нерегулярной сети точек // *Технологии Microsoft в теории и практике программирования: сборник трудов IX Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых*. – Томск: Изд-во ТПУ, 2012. – С. 135–137.
15. Друки А.А., Спицын В.Г. Распознавание структурированных символов на изображениях с использованием гистограмм средней интенсивности и сверточной нейронной сети // *Известия Томского политехнического университета*. – 2013. – Т. 322. – № 5. – С. 120–125.
16. Хайкин С. Нейронные сети: полный курс. – М.: Вильямс, 2006. – 1103 с.
17. Aslam J.A., Popa R.A., Rivest R.L. On Estimating the Size and Confidence of a Statistical Audit // *Proceedings of the Electronic Voting Technology Workshop (EVT '07)*. – Boston, 2007. – P. 254–259.
18. Rumelhart D., Hinton G.E., Williams R.J. Learning Representations of Back-Propagation Errors // *Nature*. – 1986. – № 323. – P. 533–536.
19. Riedmiller M., Braun H. A Direct Adaptive Method for Faster Backpropagation Learning: the RPROP Algorithm // *International Conference of Neural Networks*. – San Francisco, USA, 1993. – V. 1. – P. 586–591.
20. Holland J.H. Genetic Algorithms and Adaptation // *Adaptive Control of Ill-Defined Systems*. – New York, USA, 1984. – P. 317–333.

Поступила 20.12.2015 г.

APPLICATION OF ARTIFICIAL NEURAL NETWORKS FOR SOLVING THE PROBLEM OF TWO-DIMENSIONAL GEOFIELDS RECONSTRUCTION

Pavel A. Kakovkin,

National Research Tomsk Polytechnic University,
30, Lenin Avenue, Tomsk, 634050, Russia. E-mail: nitrokot@mail.ru

Aleksey A. Druki,

National Research Tomsk Polytechnic University,
30, Lenin Avenue, Tomsk, 634050, Russia. E-mail: druki2008@yandex.ru

The relevance of the research is caused by the necessity to develop algorithms and software to increase the efficiency of the spatial analysis and two-dimensional geofields recovery.

The main aim of the study is to develop the algorithms and software to increase the efficiency of two-dimensional geofields recovery on irregular grid point data; to carry out the experimental studies to determine the effectiveness of the developed algorithms and to compare them with the existing analogues.

The methods used in the study. To solve the task the authors have applied the methods of artificial intelligence, methods of implementation of artificial neural networks and genetic algorithms; the committee methods are applied to solve the classification problems, mathematical modeling, probability theory and mathematical statistics with the help of software Visual Studio and MatLab.

The results. The artificial intelligence methods were used for restoring geofields on irregular grid point data, as this area of research is one of the most intensively developing now. The algorithm based on artificial neural networks was developed to solve the problem. The algorithm is a sequence of actions. It consists of seven steps. The multiple neural networks of direct distribution, such as perceptron, which operate according to the bagging method, are used to restore geofield. The software application that allows solving the assigned task is developed on the basis of the proposed algorithm. The authors carried out the experimental study of the algorithm effectiveness and compared the results obtained with the results of the inverse distance-weighted method. The carried out studies shown that the results of the proposed algorithm operation are higher than the operating efficiency of the inverse-distance weighted method.

Key words:

Geoinformatics, geofields reconstruction, geostatistical methods, cross-validation, artificial neural networks.

REFERENCES

1. Kapralov E.G., Koshkarev A.V., Tikunov V.S. *Geoinformatika* [Geoinformatics]. Moscow, Akademia Publ. house, 2010. 480 p.
2. Zhurkin I.G., Shaytura S.V. *Geoinformatsionnye sistemy* [Geoinformation systems]. Moscow, KUDITS-Press, 2009. 272 p.
3. Kovin R.V., Markov N.G. *Geoinformatsionnye sistemy i tekhnologii* [Geoinformation Systems and Technologies]. Tomsk, TPU Publ. house, 2009. 267 p.
4. Kanevskiy M.F., Demyanov V.V., Saveleva E.A., Chernov S.Yu. Osnovnye ponyatiya i elementy geostatistiki [The basic concepts and elements of geostatistics]. *Environment and Natural Resources*, 1999, no. 11, pp. 15–21.
5. Gousie M.B., Franklin W.R. Augmenting Grid-Based Contours to Improve Thin Plate DEM Generation. *Photogrammetric Engineering & Remote Sensing*, 2005, vol. 71 (1), pp. 69–79.
6. Kopnov M.V., Markov N.G. Vosstanovlenie dvumernykh geopoly metodami geostatistiki [Restoration of two-dimensional geofields by geostatistical methods]. *The Problems of Informatics*, 2011, no. 2 (10), pp. 36–43.
7. Kovin R.V., Kopnov M.V., Markov N.G. Osobennosti poiska znacheniy parametrov pri avtomatizirovannom postroenii prognoznykh kart [Features of searching for parameters values at automated constructing of forecast maps]. *Sovremennye tekhnika i tekhnologii. Sbornik trudov XVI Mezhdunarodnoy nauchno-prakticheskoy konferentsii studentov, aspirantov i molodykh uchenykh* [Proc. XVI Modern equipment and technology Conference]. Tomsk, April 12–16, 2010. Vol. 2, pp. 342–343.
8. Ermolaev V.A. Interpolyatsionnye vosstanavlivayushchie filtry: metod dinamicheskoy interpolatsii [Interpolation reducing filters: method of dynamic interpolation]. *Devices and systems. Management, monitoring, diagnostics*, 2002, no. 2. pp. 39–42.
9. Novakovskiy B.A., Prasolov S.V., Prasolova A.I. *Tsifrovye modeli reliefa realnykh i abstraktnykh geopoly* [Digital elevation models of real and abstract geofields]. Moscow, Nauchny mir Publ., 2003. 64 p.
10. Almeida J.A., Lopes M. Stochastic Simulation of Rainfall Using a Space-time Geostatistical Algorithm. *Geostatistics for Environmental Applications*, 2005, pp. 455–466.
11. Wim C.M. Kriging Metamodeling in Discrete-Event Simulation. *Proc. of the 2005 Winter Simulation Conference*. Orlando, USA, 2005. pp. 345–348.
12. Hendrick W., Dhaene T. Sequential Design and Rational Metamodelling. *Proc. of the 2005 Winter Simulation Conference*. Orlando, USA, 2005. pp. 368–373.
13. Russo A., Soares A., Pereira M.J. Trigo Joint Space-Time Geostatistical Model for Air. *Proc. of the VII Geostatistics for Environmental Applications Conference*. Netherlands, 2007. pp. 173–185.
14. Berezovskiy V.E., Druki A.A. Razrabotka i analiz prilozheniy, pozvolyayushchikh reshat zadachu vosstanovleniya geopolya po neregulyarnoy seti tochek [Development and analysis of applications to solve the problem of geofield recovery by irregular point network]. *Tekhnologii Microsoft v teorii i praktike programmirovaniya. Sbornik trudov IX Vserossiyskoy nauchno-prakticheskoy konferentsii studentov, aspirantov i molodykh uchenykh* [Proc. IX Microsoft technologies in theory and practice of programming Conference]. Tomsk, 2012, March 21–22. pp. 135–137.
15. Druki A.A. Raspoznavanie strukturirovannykh simvolov na izobrazheniyakh s ispolzovaniem gistogramm sredney intensivnosti i svertochnoy neyronnoy seti [Recognition of structured characters in the images using the histograms of average intensity and a convolutional neural network]. *Bulletin of the Tomsk Polytechnic University*, 2013, vol. 322, no. 5, pp. 120–125.

16. Haykin, S. *Neyronnye seti* [Neural networks]. Moscow, Viliams Publ., 2016. 1103 p.
17. Aslam J.A., Popa R.A., Rivest R.L. On Estimating the Size and Confidence of a Statistical Audit. *Proc. of the Electronic Voting Technology Workshop (EVT '07)*. Boston, 2007. pp. 254–259.
18. Rumelhart D., Hinton G.E., Williams R.J Learning Representations of Back-Propagation Errors. *Nature*, 1986, no. 323, pp. 533–536.
19. Riedmiller M., Braun H. A Direct Adaptive Method for Faster Backpropagation Learning: the RPROP Algorithm. *International Conference of Neural Networks*. San Francisco, USA, 1993. Vol. 1, pp. 586–591.
20. Holland J.H. Genetic Algorithms and Adaptation. *Adaptive Control of Ill-Defined Systems*, New York, USA, 1984. pp. 317–333.

Received: 20 December 2015.