

УДК 550.8.053:519.2

ВЛИЯНИЕ РЕПРЕЗЕНТАТИВНОСТИ ИСХОДНЫХ ДАННЫХ НА РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ МЕТОДОМ ДВОЙНОГО КРАГИНГА

Шестаков Валерий Владимирович¹,
valeriy.shestakov@inbox.ru

Степанов Дмитрий Юрьевич²,
stepanovdj70@gmail.com

¹ Национальный исследовательский Томский политехнический университет,
Россия, 634050, г. Томск, пр. Ленина, 30.

² ООО Научно-аналитический центр «Недра»,
Россия, 634029, г. Томск, ул. Советская, 26.

Актуальность. При решении обратных геофизических задач важное положение занимает задача построения объемных моделей петрофизических параметров. Наибольшие затруднения при разработке методов решения этой задачи определяются неоднородностью реальной геологической среды, а их точность – недостаточностью сетки скважинных наблюдений. Приведен новый метод, отличительной особенностью которого является совместное использование данных наземной сейсморазведки и геофизических исследований скважин. Он опирается на опыт геостатистического подхода и, решая описанные проблемы, использует предположение о том, что сейсмические и скважинные данные, измеренные в пределах одного геологического объекта, могут иметь схожие ковариационные свойства. Вопросы надежности и эффективности метода двойного крагинга ранее не были опубликованы, хотя их исследование требуется для практического применения метода. Одним из первых вопросов здесь является оценка влияния репрезентативности исходных данных.

Цель работы: рассмотреть влияние репрезентативности исходных данных на качество моделирования методом двойного крагинга и возможные пути по разработке количественной меры оценивания репрезентативности.

Объекты: синтезированная модель куба сейсмического атрибута; данные 3D МОГТ и ГИС Конторовичского месторождения Томской области.

Методы: методы теории случайных функций, методы линейной алгебры, статистическое моделирование и вычислительный эксперимент.

Результаты. Аналитически доказано, что при наличии репрезентативной выборки данных ошибка моделирования равна нулю. Аналитически и численно доказано, что при наличии нерепрезентативной выборки скважин ошибка моделирования и множитель Лагранжа прямо пропорциональны весовым коэффициентам данных, недостающих для репрезентативности выборки. На основании этого факта было выдвинуто предположение о том, что множитель Лагранжа может быть использован в качестве меры репрезентативности используемой выборки. Это предположение было проверено на материалах Конторовичского месторождения Томской области, в рамках которого методом двойного крагинга последовательно осуществлялось моделирование с участием трех, пяти и семи скважин. При увеличении выборки от трех до семи скважин наблюдалось снижение величины множителя Лагранжа, что подтвердило правильность предположения.

Ключевые слова:

Статистическое моделирование, сейсмическая разведка, крагинг, петрофизическое моделирование.

Введение

Наиболее эффективным и часто применяемым инструментом исследования земной коры на сегодняшний день является сейсморазведка. Её лидерство обусловлено большой глубиной, густой сетью изученности, а также сравнительно невысокой стоимостью [1, 2]. Все сейсморазведочные методы основаны на регистрации и изучении упругих волн, искусственно возбуждаемых на поверхности земли. Распространяясь в земной коре во всех направлениях, эти волны преодолевают различные неоднородности и претерпевают процессы отражения, преломления и дифракции. В свою очередь, отраженные и преломленные волны, вернувшись к поверхности земли, регистрируются высокочувствительными датчиками – сейсмоприемниками, в виде сейсмограмм. Набор сейсмограмм, зарегистрированных в различных точках поверхности земли, обрабатывается, в результате чего формируются так называемые 2D или 3D сейсмические (геофизи-

ческие) поля [1, 3]. На их основе рассчитываются различные сейсмические атрибуты, необходимые для решения *обратных задач* сейсморазведки [4]: исследование строения земной коры, оценка перспектив нефтегазонасыщенности, картирование мелких ловушек нефти и газа, прогнозирование распространения коллекторов и др.

В число задач, решаемых с помощью сейсморазведочного подхода, также входит задача прогноза петрофизических параметров геологической среды. Эти параметры несут в себе чрезвычайно важную информацию о наличии или отсутствии нефтяных коллекторов и миграции флюидов в геологической среде, ввиду чего измеряются с помощью высокоточных и дорогостоящих каротажных методов, имеющих обобщенное название *Геофизические исследования скважин* (ГИС). ГИС направлены на изучение геологического строения земных недр, а также на измерение некоторых физических параметров пород вдоль стволов скважин

(таких как удельное электрическое сопротивление, радиоактивность, скорость распространения упругих волн и т. д.) [5]. В отличие от сейсмических полей, данные ГИС обычно измеряются на редкой, неравномерной сети и хранятся в виде набора кривых, привязанных к координатам стволов скважин. Однако, т. к. геологическая среда неоднородна и анизотропна, полное описание её петрофизических свойств требует наличия более информативной выборки данных, которую можно получить двумя способами. Первый способ – это бурение скважин по густой сети с дальнейшим проведением геофизических исследований. Его очевидные недостатки: дороговизна и высокая сложность реализации. Второй способ – решение задачи моделирования петрофизических свойств с применением математических методов интерполяции или аппроксимации.

Методы решения данной задачи известны и условно делятся на две группы: первые используются для проведения расчетов только данные ГИС [6], вторые основаны на совместном использовании сейсмических данных и данных ГИС. Эффективность первой группы методов в основном зависит от плотности сети скважин. Вторая группа методов дает существенно лучшие результаты для редкой сети скважин и на текущий момент развивается наиболее активно [7]. Среди методов второй группы выделяют:

- 1) методы прямого пересчета сейсмических атрибутов в параметры ГИС;
- 2) методы, основанные на машинном обучении и (или) нейронных сетях;
- 3) методы геостатистики.

В данной работе приведен краткий обзор преимуществ и недостатков перечисленных методов и анализ результатов моделирования алгоритмом «двойного крагинга», разработанного авторами статьи [8]. В условиях малого количества исходных данных ГИС одним из важных и еще не исследованных вопросов для подобных алгоритмов становится вопрос о репрезентативности исходных данных: её влиянии на точность моделирования и возможности оценки репрезентативности реальных материалов.

Обзор методов построения моделей геологических сред

Методы прямого пересчета сейсмических атрибутов в параметры ГИС в общем случае сводятся к установлению функциональной зависимости между петрофизическим параметром и одним или несколькими сейсмическими атрибутами [9]. Проблема этого подхода заключается в том, что на текущий момент на практике преимущественно используются линейные функции, эффективные в достаточно небольших пределах одного геологического образования. Для моделирования в рамках целого месторождения они неприменимы в связи с тем, что для отдельно взятой глубины и отдельно взятого объекта зависимости между сейсмическими

атрибутами и петрофизическими параметрами разнятся. Вдобавок любую взаимосвязь, как линейную, так и нелинейную, можно установить только в условиях хорошей изученности объекта. В противном случае о качественном моделировании не может быть и речи.

Методы машинного обучения и нейронные сети [1, 7, 9, 10], в отличие от методов прямого пересчета, основаны на минимизации ошибки осуществляемых расчетов, ввиду чего дают лучший результат и позволяют строить более правдоподобные модели. Как алгоритмы машинного обучения, так и нейронные сети для решения задачи прогноза петрофизических параметров обучаются по выборке сейсмических атрибутов. Обучение при этом организовано таким образом, чтобы прогнозные значения петрофизического параметра максимально совпадали с известными реальными аналогами. Основными преимуществами нейронных сетей и машинного обучения являются [11, 12]:

1. Гибкость. Обучение может осуществляться как для отдельно взятого пласта, так и для определенного интервала; с использованием одного или нескольких сейсмических атрибутов и петрофизических параметров.
2. Высокая скорость расчета прогнозных параметров при наличии уже обученной сети (алгоритма).

К недостаткам этих методов можно отнести:

- 1) необходимость переобучения для каждого отдельного месторождения, пласта, петрофизического параметра и т. д.;
- 2) возможность несходности алгоритма обучения для определенного набора входных данных;
- 3) реализацию аппроксимации, а не интерполяции скважинных данных.

Говоря о недостатках более подробно: первый – несущественен, т. к. влечет за собой лишь увеличение времени расчета и не влияет на качество моделей; второй – может быть компенсирован предварительным отбором исходных данных; третий – приводит к несоответствию реальных и прогнозных значений петрофизических параметров.

Последний недостаток является значимым и может быть устранен лишь с помощью интерполяционных методов, оптимальными из которых считаются *методы геостатистики* [13, 14]. Геостатистика, в общем, представляет собой технологию интерполяции пространственно-распределенных данных, использующую статистические методы. Большинство методов геостатистики базируются на применении вариограмм, которые в свою очередь описывают пространственную корреляцию прогнозируемого поля и рассчитываются в предположении о непрерывности и однородности исследуемого объекта [13]. Последнее некорректно по отношению к геологической среде, ввиду чего в условиях редкой сети данных ГИС применение классических геостатистических методов для прогноза петрофизических параметров приводит к грубому, далекому от реальности результату. Для

описания пространственно-неоднородной модели среды необходимо отказаться от вариограммного анализа данных ГИС и использовать ковариационный анализ сейсмического атрибута в скользящих пространственно-глубинных окнах. В основе такого подхода должно лежать предположение об идентичности или, по крайней мере, близости ковариационных свойств сейсмического атрибута и прогнозируемого параметра, т. к. они измеряются по одному и тому же физическому объекту. Метод, основанный на данных предположениях, разработан авторами статьи и реализован в виде алгоритма, описание которого излагается ниже.

Алгоритм двойного крагинга

В качестве базового уравнения геостатистика использует уравнение крагинга, которое в рамках рассматриваемой задачи может быть записано следующим образом [13]:

$$\hat{F}(x, y, z) = \sum_{i=1}^N w_i(x, y, z) f_i(z), \quad (1)$$

где x, y, z – пространственные координаты; $w_i(x, y, z)$ – весовая функция i -й скважины; $f_i(z)$ – известные кривые параметра ГИС в i -й скважине; $\hat{F}(x, y, z)$ – прогнозируемое значение параметра ГИС; N – количество скважин.

Минимальный средний квадрат ошибки решения уравнения (1) обеспечивается требованием

$$M[(F(x, y, z) - \hat{F}(x, y, z))^2] \rightarrow \min, \quad (2)$$

а несмещенность оценки – ограничением

$$\sum_{i=1}^N w_i(x, y, z) = 1, \quad (3)$$

где $F(x, y, z)$ – истинное значение параметра ГИС [13]. Анализ (2) в произвольной точке (x, y, z) приводит к системе линейных алгебраических уравнений (СЛАУ) относительно неизвестных весовых функций:

$$C^f(z)W(x, y, z) = C^F(x, y, z), \quad (4)$$

где $C^f(z)$ – матрица ковариационных функций $C_{ij}^f(z) = M[f_i(z)f_j(z)]$ известных кривых ГИС; $W(x, y, z)$ – матрица весовых функций; $C^F(x, y, z)$ – матрица ковариационных функций $C_{ij}^F(x, y, z) = M[f_i(z)F(x, y, z)]$ известных кривых ГИС с параметром ГИС в точке прогноза (x, y, z) . Основная проблема геостатистического подхода заключается в необходимости моделирования $C^F(x, y, z)$ правой части системы (4), априори известной только в точках скважин.

В классическом крагинге предполагается, что прогнозируемое поле удовлетворяет условию однородности второго рода, т. е. ковариационные функции могут быть описаны функциями расстояний, а не координат. Это предположение позволяет проводить статистический анализ пространственной зависимости известных данных и использовать вместо ковариационных функций вариацию разницы значений прогнозируемого параметра,

зависящую только от удаленности точек поля [7, 8]. Однако переход от ковариации к вариации приводит к потере информации о пространственной неоднородности моделируемого объекта.

Такой подход можно не применять, если привлечь к моделированию данные наземной сейсморазведки $S(x, y, z)$, охватывающие густой сетью наблюдений прогнозируемую площадь среды. Аналогично (1) построим уравнением крагинга статистическую модель сейсмического атрибута по данным, измеренным только вблизи скважин:

$$\hat{S}(x, y, z) = \sum_{i=1}^N w_i^0(x, y, z) S_i(z), \quad (5)$$

где $w_i^0(x, y, z)$ – весовая функция сейсмического атрибута в точке i -й скважины; $S_i(z) = S(x_i, y_i, z)$ – кривая сейсмического атрибута вблизи точки i -й скважины; $\hat{S}(x, y, z)$ – прогнозируемое значение атрибута.

Для использования модели (5) необходимо принять предположение, что ковариационные свойства, как свойства внутренней изменчивости или взаимосвязи, сейсмических данных $S(x, y, z)$ и ковариационные свойства прогнозируемого параметра $F(x, y, z)$ идентичны, или хотя бы близки, т. к. измерены по одному и тому же объекту. Это позволяет утверждать, что система (4) в произвольной точке (x, y, z) будет эквивалентна системе, построенной на данных сейсморазведки [8]:

$$\begin{cases} \sum_{j=1}^N w_j^0(x, y, z) c_{ij}(z) + \mu(x, y, z) = \\ = c_{i0}(x, y, z), \quad i = \overline{1, N}, \\ \sum_{j=1}^N w_j^0(x, y, z) = 1, \end{cases} \quad (6)$$

где $c_{ij}(z) = M[S_i(z)S_j(z)]$ – ковариации сейсмического атрибута в точках скважин с координатами (x_i, y_i) и (x_j, y_j) на глубине z ; $c_{i0}(x, y, z) = M[S_i(z)S(x, y, z)]$ – ковариации сейсмического атрибута в точках скважин и в точке прогноза; $\mu(x, y, z)$ – множитель Лагранжа. Или в матричной форме:

$$C(z)W^0(x, y, z) = C^0(x, y, z),$$

где

$$C(z) = \begin{pmatrix} c_{11}(z) & \cdots & c_{1N}(z) & 1 \\ c_{21}(z) & \cdots & c_{2N}(z) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ c_{N1}(z) & \cdots & c_{NN}(z) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix},$$

$$W^0(x, y, z) = \begin{pmatrix} w_1^0(x, y, z) \\ w_2^0(x, y, z) \\ \vdots \\ w_N^0(x, y, z) \\ \mu(x, y, z) \end{pmatrix},$$

$$C^0(x, y, z) = \begin{pmatrix} c_{10}(x, y, z) \\ c_{20}(x, y, z) \\ \vdots \\ c_{N0}(x, y, z) \\ 1 \end{pmatrix}.$$

Хотя сейсмическое поле является неоднородным и нестационарным процессом, ранее было показано, что $S(x, y, z)$ в любой точке пространства является импульсным случайным процессом, для которого условие локальной стационарности допустимо в пределах длины волны λ [15]. Поэтому ковариацию можно вычислять путем осреднения в скользящем глубинном окне

$$c_{ij}(z) = M[S(x_i, y_i, z)S(x_j, y_j, z)] = \frac{1}{\lambda} \int_{z-\lambda/2}^{z+\lambda/2} S(x_i, y_i, \tau)S(x_j, y_j, \tau) d\tau.$$

Теперь, если сейсмические данные определены в точке прогноза, правая часть системы (6) определяется по реальным данным и моделирование ковариационных свойств не требуется.

Отметим, что область значений весовых коэффициентов в системах (4), (6) ничем не ограничена, что на практике приводит в отдельных областях к появлению «физически нереального» результата: $\hat{F}(x, y, z)$ приобретает значения, не входящие в область допустимых. Например, известно, что кажущаяся удельная проводимость, измеряемая индукционным каротажем (ИК), интенсивность естественного γ -излучения по гамма-каротажу (ГК), коэффициент пористости не могут иметь отрицательные значения [5]. Для учета подобных ограничений необходимо наложить дополнительное ограничение на все весовые коэффициенты:

$$w_i^0(x, y, z) \geq 0. \quad (7)$$

Очевидно, что при одновременном соблюдении (3) и (7) результат суммирования (1) не будет превышать пределов, определенных значениями исходных данных. Ввиду того, что (7) представляет собой неравенство, в явном виде добавить его в систему (6) невозможно. Поэтому поиск решения необходимо осуществлять итерационно с исключением из системы строк и столбцов тех слагаемых, которым соответствуют отрицательные весовые коэффициенты.

В случае нетривиального решения системы (6), согласно методу Крамера [16], неизвестные СЛАУ могут быть найдены с помощью определителя основной матрицы и определителя матрицы C_i , в которой i -й столбец заменен на столбец свободных членов

$$w_i^0(x, y, z) = \frac{\det(C_i)}{\det(C)}, \quad \mu(x, y, z) = \frac{\det(C_{N+1})}{\det(C)}.$$

В случае, когда прогнозная точка совпадает с k -й точкой скважины, столбец свободных членов будет совпадать с k -м столбцом основной матрицы

$c_{i0}(z) = c_{ik}(z)$ и, следовательно, $\det(C_k) = \det(C)$, а $\det(C_i) = 0$ для $\forall i \neq k$. Из чего следует, что решением будут являться $w_k^0(x_k, y_k, z) = 1$ и $w_i^0(x_k, y_k, z) = 0$ для $\forall i \neq k$, $\mu(x_k, y_k, z) = 0$. Подставляя полученное решение в формулу (1), получаем что $\hat{F}(x_k, y_k, z) = w_k^0(x_k, y_k, z)f_k(z) = f_k(z)$ в точках скважин прогнозное значение будет точно совпадать с заданным.

Вне точек скважин расхождение параметров реальной среды и прогнозной модели определяется двумя основными факторами: множество известных данных ГИС не является репрезентативным, и статистические свойства сейсмического атрибута не отражают в полной мере свойства прогнозного параметра. В данной работе рассматривается влияние только первого фактора. Репрезентативным множеством данных будем называть такое множество минимального объема, по которому можно с заданной точностью восстановить параметры среды в любой точке их линейной комбинацией (1).

Исследование алгоритма

По аналогии с анализом ошибок крагинга можно привести дисперсию прогнозного значения [13, 14]:

$$\begin{aligned} \sigma_{\text{ок}}^2(x, y, z) &= M[(S_0(x, y, z) - \hat{S}(x, y, z))^2] = \\ &= \psi_s^2(x, y, z) - 2 \sum_{i=1}^N w_i(x, y, z) c_{0i}(x, y, z) + \\ &+ \sum_{i=1}^N \sum_{j=1}^N w_i(x, y, z) w_j(x, y, z) c_{ij}(x, y, z) = \\ &= \psi_s^2(x, y, z) - 2 \sum_{i=1}^N w_i(x, y, z) c_{0i}(x, y, z) + \\ &+ \sum_{i=1}^N w_i(x, y, z) c_{0i}(x, y, z) - \mu(x, y, z) = \\ &= \psi_s^2(x, y, z) - \sum_{i=1}^N w_i^0(x, y, z) c_{0i}(x, y, z) - \\ &- \mu(x, y, z), \end{aligned} \quad (8)$$

где $\psi_s^2(x, y, z)$ – вариация прогнозируемого поля. Второе слагаемое определяет взвешенную сумму ковариаций прогнозной точки и исходных данных и показывает, что ошибка прогнозирования будет меньше в тех точках, где существует ковариационная зависимость с известными данными. Значение третьего слагаемого – множителя Лагранжа – в литературных источниках не приводится и, как написано в [8], «само по себе интереса не представляет». При использовании геостатистических алгоритмов крагинга ошибка (8) равна нулю только в точках, где определены скважины [14, 17]. В межскважинном пространстве величина $\sigma_{\text{ок}}^2$ определяется выбранной моделью вариограммы и, как правило, растет с удалением от скважин.

Для анализа ошибок алгоритма двойного крагинга рассмотрим решение системы (6) в фиксированной точке (x_0, y_0, z_0) и примем общую статистическую модель сейсмических наблюдений в виде

линейной комбинации известных N значений, M неизвестных значений и случайного шума:

$$S_0(z) = I(z) + Q(z) + \eta(z) = \sum_{i=1}^N v_i S_i(z) + \sum_{j=1}^M v_j S_j(z) + \eta(z),$$

где $I(z)$ и $Q(z)$ – соответственно, известная и неизвестная составляющие; $\eta(z)$ – случайная составляющая модели. В данной работе примем, что между значениями сейсмического атрибута $S_i(z)$, между $S_j(z)$, между $S_i(z)$ и $S_j(z)$ может существовать корреляционная связь, но их совокупность представляет собой линейно-независимое множество. Актуальным и первым вопросом при исследовании метода является вопрос о репрезентативности исходных данных. В рамках описанной выше модели этот вопрос может быть перефразирован как вопрос о наличии или отсутствии неизвестной составляющей $Q(z)$. Поэтому в качестве первой модели исследования примем абстрактную ситуацию, когда все слагаемые модели известны, а случайная составляющая отсутствует (*модель репрезентативная*):

$$S_r(z) = \sum_{i=1}^N v_i S_i(z), \quad \sum_{i=1}^N v_i = 1,$$

где $S_r(z)$ – значение сейсмического атрибута в прогнозной точке; $S_i(z)$ – значение сейсмического атрибута в точке i -й скважины ($i=1, \dots, N$); v_i – вес i -го значения. Вариация первой модели

$$\psi_r^2 = M[S_r^2(z)] = \sum_{i=1}^N v_i \sum_{j=1}^N v_j M[S_i(z)S_j(z)] = \sum_{i=1}^N v_i \sum_{j=1}^N v_j c_{ij}. \quad (9)$$

Для такой модели элементы матрицы свободных членов C^0 представляются в виде линейной комбинации элементов основной матрицы (8):

$$\begin{aligned} c_{10} &= M[S_1(z)S_r(z)] = v_1 c_{11} + v_2 c_{12} + \dots + v_N c_{1N}, \\ c_{20} &= M[S_2(z)S_r(z)] = v_1 c_{21} + v_2 c_{22} + \dots + v_N c_{2N}, \\ &\vdots \\ c_{N0} &= M[S_N(z)S_r(z)] = v_1 c_{N1} + v_2 c_{N2} + \dots + v_N c_{NN}. \end{aligned} \quad (10)$$

Соответственно, матрицы C_i ($i=1, \dots, N$) с учетом условия нормирования весовых коэффициентов будут иметь вид

$$C_i = \begin{pmatrix} c_{11} & \dots & v_1 c_{11} + \dots + v_N c_{1N} & \dots & c_{1N} & 1 \\ c_{21} & \dots & v_1 c_{21} + \dots + v_N c_{2N} & \dots & c_{2N} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{N1} & \dots & v_1 c_{N1} + \dots + v_N c_{NN} & \dots & c_{NN} & 1 \\ 1 & \dots & v_1 + \dots + v_N & \dots & 1 & 0 \end{pmatrix}. \quad (11)$$

Учитывая, что определитель матрицы – это полилинейная функция строк или столбцов [18], определитель матрицы (11) можно разложить в виде следующей суммы:

$$\det(C_i) = v_1 \cdot 0 + v_2 \cdot 0 + \dots$$

$$+ \dots + v_i \det(C) + \dots + v_N \cdot 0 = v_i \det(C).$$

Анализируя матрицу C_{N+1} аналогичным способом, можно показать, что $\det(C_{N+1})=0$. Следовательно, в соответствии с формулами (10), решением СЛАУ (6) будут являться весовые коэффициенты принятой модели и нулевой множитель Лагранжа:

$$w_i^0(x_0, y_0, z_0) = v_i \frac{\det(C)}{\det(C)}, \quad \mu(x_0, y_0, z_0) = 0. \quad (12)$$

Таким образом, если для прогнозируемой точки (x_0, y_0, z_0) известные данные представляют собой репрезентативную выборку, ошибка оценивания с учетом (9), (10), (12) оказывается нулевой:

$$\begin{aligned} \sigma_{OK}^2(x_0, y_0, z_0) &= \psi_r^2(z_0) - \sum_{i=1}^N v_i c_{i0}(z_0) = \psi_r^2(z_0) - \sum_{i=1}^N v_i \sum_{j=1}^N v_j c_{ij} = 0. \end{aligned}$$

Во второй модели исследования предположим, что сейсмический атрибут является линейной комбинацией не только известных значений, но и одного неизвестного слагаемого (*модель нерепрезентативная*):

$$S_u(z) = \sum_{i=1}^N v_i S_i(z) + v_{N+1} S_{N+1}(z), \quad \sum_{i=1}^{N+1} v_i = 1,$$

где $S_u(z)$ – значение сейсмического атрибута в прогнозной точке; $S_{N+1}(z)$ и v_{N+1} – неизвестное значение сейсмического атрибута и его вес. Вариация второй модели

$$\begin{aligned} \psi_u^2 &= M[S_u^2(z)] = \sum_{i=1}^N v_i \sum_{j=1}^N v_j M[S_i(z)S_j(z)] + \\ &+ v_{N+1} \sum_{i=1}^N v_i M[S_i(z)S_{N+1}(z)] = \\ &= \sum_{i=1}^N v_i \left(\sum_{j=1}^N v_j c_{ij} + v_{N+1} c_{iN+1} \right). \end{aligned}$$

Проводя далее аналогичный анализ матриц C_i , C_{N+1} и C_0 получим следующее решение:

$$\begin{aligned} w_i^0(x_0, y_0, z_0) &= v_i + v_{N+1} \frac{\det(C_i)}{\det(C)}, \\ \mu(x_0, y_0, z_0) &= v_{N+1} \frac{\det(C_{N+1})}{\det(C)}. \end{aligned} \quad (13)$$

Таким образом, если для прогнозируемой точки (x_0, y_0, t_0) известные данные представляют собой нерепрезентативную выборку:

- найденные весовые функции отличаются от модельных аддитивной добавкой, пропорциональной весовому коэффициенту неизвестного значения;
- множитель Лагранжа отличен от нуля и прямо пропорционален весовому коэффициенту неизвестного значения.

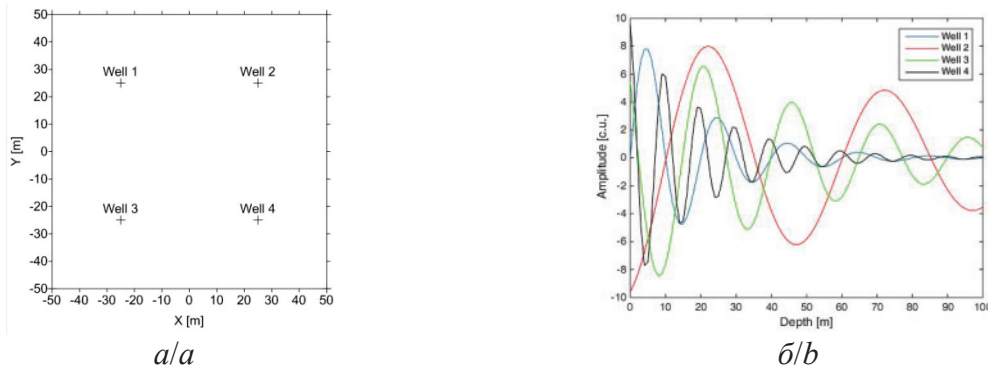


Рис. 1. Расположение модельных скважин на плоскости (а) и модельные значения сейсмического атрибута $S_i(z)$ в точках скважин (б)

Fig. 1. Model wells location in the plane (a) and seismic attribute model values $S_i(z)$ at the wells points (b)

Ошибка оценивания также становится отличной от нуля

$$\begin{aligned} \sigma_{OK}^2(x_0, y_0, t_0) &= \sum_{i=1}^N v_i \left(\sum_{j=1}^N v_j c_{ij} + v_{N+1} c_{iN+1} \right) - \\ &- \sum_{i=1}^N \left(v_i + v_{N+1} \frac{\det(U_i)}{\det(C)} \right) c_{i0}(t_0) - \mu(x_0, y_0, z_0) = \\ &= \sum_{i=1}^N v_i \left(\sum_{j=1}^N v_j c_{ij} + v_{N+1} c_{iN+1} \right) - \\ &- \sum_{i=1}^N \left(v_i + v_{N+1} \frac{\det(U_i)}{\det(C)} \right) \left(\sum_{j=1}^N v_j c_{ij} + v_{N+1} c_{iN+1} \right) - \\ &- \mu(x_0, y_0, z_0) = -\mu(x_0, y_0, z_0) - \\ &- \frac{v_{N+1}}{\det(C)} \sum_{i=1}^N \det(C_i) \left(\sum_{j=1}^N v_j c_{ij} + v_{N+1} c_{iN+1} \right) \end{aligned}$$

и прямо пропорциональна весовому коэффициенту неизвестного значения, определяется множителем Лагранжа и ковариациями не только известных значений, но и неизвестного значения с известными.

Численное исследование полученных соотношений проводилось на синтезированной модели сейсмических данных. В ограниченной пространственной области ($100 \times 100 \times 100$ м) было определено положение четырех вертикальных скважин (рис. 1, а). Значение сейсмического атрибута в скважинах моделировалось импульсом Берлаге с различными параметрами (рис. 1, б) [19]. Весовые функции каждой скважины $v_i(x, y)$ определялись численно из условий равенства единице в точке скважины и существования области нулевых весов – областей, где i -я скважина не участвует в моделировании (пример на рис. 3, а, в и 4, а). Куб сейсмических данных моделировался путем весового суммирования модельных значений атрибута всех четырех скважин на сетке 1×1 м (рис. 2):

$$S(x, y, z) = \sum_{i=1}^4 v_i(x, y) S_i(z).$$

Решение обратной задачи опиралось на предположение о наличии только трех скважин: № 2–4.

Скважина № 1 с координатами $[-25; 25]$, расположенная в левой верхней части модели, считалась за неизвестную. Для каждой точки пространства решалась система (6). Примеры результатов расчета весовых коэффициентов представлены на рис. 3, б, в. Как и следовало ожидать, согласно (13), искажения оценок весовых функций наблюдаются в области, где наибольшее влияние имела первая скважина, и наоборот – в области, где первая скважина имела нулевой вес, весовые функции восстановлены точно.

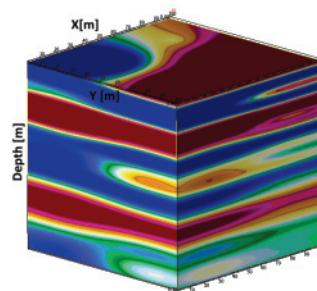


Рис. 2. Экспериментальная модель куба сейсмического атрибута

Fig. 2. Experimental model of the seismic attribute cube

Далее для определения ошибки моделирования проводилось создание куба сейсмического атрибута по оценкам весовых функций согласно (6):

$$\hat{S}(x, y, z) = \sum_{i=2}^4 w_i^0(x, y) S_i(z)$$

и вычисление дисперсии моделирования [20]

$$\sigma_{OK}(x, y) = \int_0^{100} (S(x, y, z) - \hat{S}(x, y, z))^2 dz.$$

Для сопоставления полученные распределения вынесены на рис. 4. Из рисунка видно, что в зонах, где вес $v_1(x, y) = 0$ (зона репрезентативности выборки из трех скважин № 2, 3, 4) множитель Лагранжа и ошибка моделирования равны нулю. В зонах, где вес $v_1(x, y) \neq 0$ (зона нерепрезентативности выборки из трех скважин), вес $v_1(x, y)$, множитель Лагранжа и ошибка моделирования имеют одинаковый характер распределения.

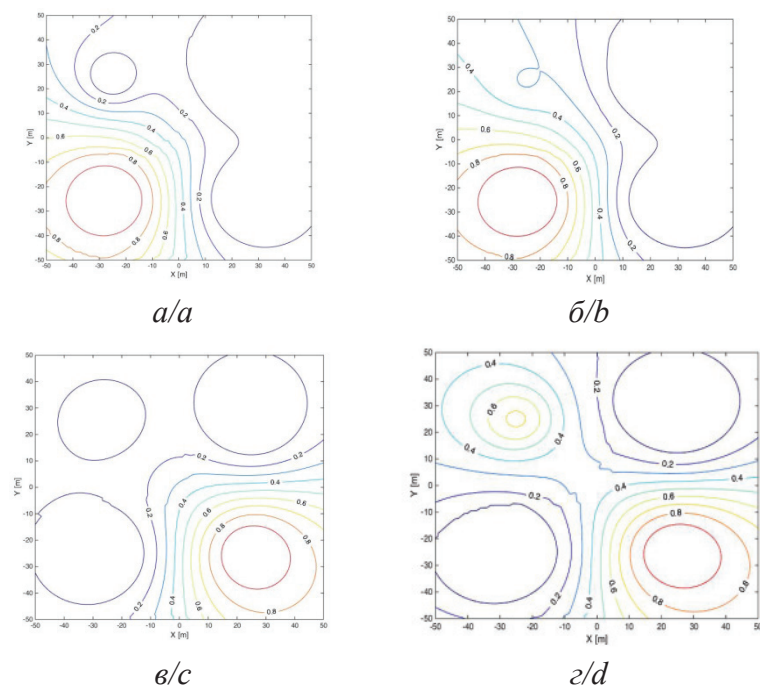


Рис. 3. Модельные весовые функции $v_3(x,y)$ (а), $v_4(x,y)$ (в) и оценки весовых функций $v_3^0(x,y)$ (б), $v_4^0(x,y)$ (г)

Fig. 3. Model weight functions $v_3(x,y)$ (a), $v_4(x,y)$ (c) and estimates of weight functions $v_3^0(x,y)$ (b), $v_4^0(x,y)$ (d)

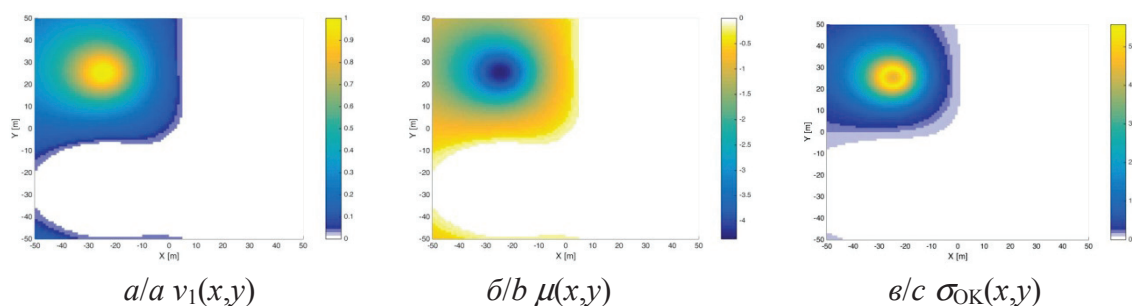


Рис. 4. Модельная весовая функция неизвестных данных (а), множитель Лагранжа (б) и дисперсия моделируемого куба (в)

Fig. 4. Model weight function of unknown data (a), the Lagrange multiplier (b) and the variance of the simulated cube (c)

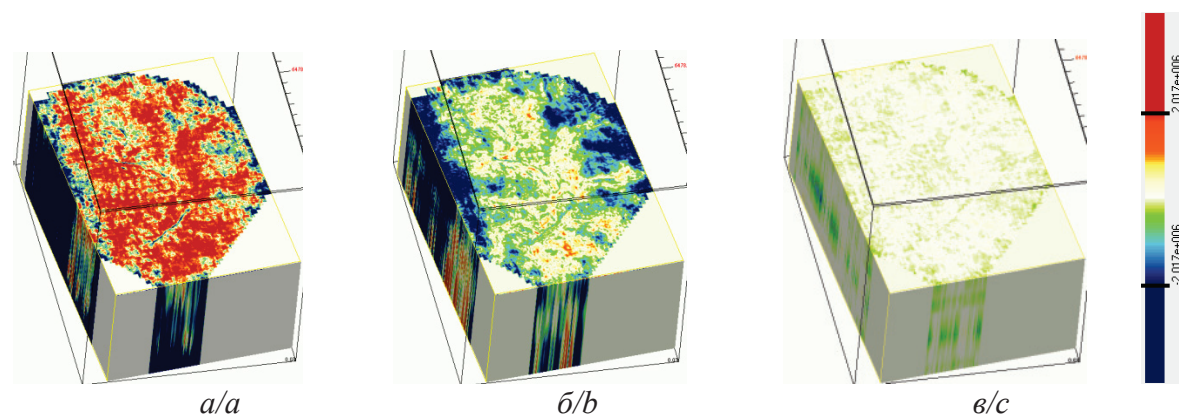


Рис. 5. Срез куба множителя Лагранжа $\mu(x,y,z)$ на уровне подошвы Баженовской свиты при моделировании по трем скважинам (а), пяти скважинам (б) и семи скважинам (в). Конторовичское месторождение Томской области (с разрешения ООО «Томскгеонефтегаз»)

Fig. 5. Cut of the Lagrange multiplier cube at the bottom level of the Bazhenovskaya formation when simulating three wells (a), five wells (b) and seven wells (c). Kontorovichskoe deposit of the Tomsk region (with the permission of ООО Tomskgeoneftegaz)

Таким образом, множитель Лагранжа можно считать мерой оценки репрезентативности исходного набора данных, т. е. их достаточности для моделирования в конкретной точке. Данное заключение было проверено на материалах 3D МОГТ Конторовичского месторождения Томской области, на территории которого пробурено семь скважин. В качестве сейсмического атрибута использовалась форма импульса волны, отраженной от подошвы Баженовской свиты, после специальных процедур повышения разрешенности сейсмической записи. Решение системы (6) искалось три раза. На первом этапе использовались данные только трех скважин (рис. 5, а). Множитель Лагранжа оказался близок по модулю к дисперсии исходных данных практически во всей анализируемой области, за исключением небольших узких зон, и показал нерепрезентативность такого набора. На втором этапе в систему (6) были добавлены две скважины, расположенные в центральной части площади (рис. 5, б), что привело к существенному снижению значения множителя Лагранжа не только вблизи скважин, но и на значительном удалении от них. На третьем этапе при моделировании использовались все семь скважин площади (рис. 5, в) – значение множителя Лагранжа по всей площади оказалось по модулю намного ниже дисперсии исходных данных и свидетельствует о репрезентативности этого набора скважин и достижении удо-

влетворительного качества моделирования практически во всем анализируемом объеме.

Заключение

Рассмотрены вопросы применения алгоритма двойного крагинга при моделировании петрофизических свойств среды по данным ГИС и наземной сейсморазведки. Показано, что данный алгоритм относится к классу интерполяционных и, в отличие от известных геостатистических методов, опирается на неоднородную модель среды. Анализ ошибок прогнозирования параметра среды на моделях репрезентативной и нерепрезентативной выборки показал, что:

- в зонах, где исходные данные являются репрезентативными, ошибка моделирования методом двойного крагинга будет равна нулю вне зависимости от удаления от скважин;
- «неинтересный» в теории геостатистики множитель Лагранжа прямо пропорционален весовой функции неизвестных данных и может быть использован в качестве меры оценки репрезентативности исходных данных.

Полученные выводы проверены на синтезированной модели сейсмического атрибута и материалах 3D МОГТ Конторовичского месторождения Томской области. Точность моделирования алгоритма двойного крагинга будет определяться не столько плотностью и объемом исходных данных, сколько их репрезентативностью.

СПИСОК ЛИТЕРАТУРЫ

1. Воскресенский Ю.Н. Полевая геофизика – М.: Недра, 2010. – 479 с.
2. Ермаков А.П. Введение в сейсморазведку. – Тверь: Изд-во ГЕРС, 2012. – 160 с.
3. Сейсморазведка. Справочник геофизика / под ред. И.И. Гурвича, В.П. Номоконова. – М.: Недра, 1981. – 464 с.
4. Кузнецов В.И. Элементы объемной (3D) сейсморазведки. 2-е изд. с изм. – Уфа: Инфореклама, 2012. – 272 с.
5. Косков В.Н., Косков Б.В. Геофизические исследования скважин и интерпретация данных ГИС. – Пермь, Перм. гос. техн. ун-т., 2007. – 304 с.
6. Беляков Е.О., Мухидинов Ш.В. Использование обобщенных зависимостей для построения петрофизических моделей фильтрационно-емкостных свойств с оценкой граничных параметров выделения коллекторов и определения их характера насыщенности // Петрофизика сложных коллекторов: проблемы и перспективы 2015; Сборник статей / сост. Б.Н. Еникеев. – М.: ООО «EAGE Геомодель», 2015. – 383 с.
7. Seismic-guided estimation of log properties (P. 3: A controlled study) / Ph.S. Shultz, Sh. Ronen, M. Hattori, Ch. Corbett // The leading Edge. – 1994. – V. 13. – Iss. 5. – P. 305–310.
8. Shestakov V.V., Sysolyatina G.A., Stepanov D.Ju. Three-dimensional models of geoenvironmental parameters // Advances in computer science research. – 2016. – V. 51: Information technologies in science, management, social sphere and medicine (ITSMSSM 2016). – P. 126–129.
9. Hampson D., Schuelke J., Quirein J. Use of Multi-Attribute Transforms to Predict Log Properties from Seismic Data // Geophysics. – 2001. – V. 66 (1). – P. 220–236.
10. Лаврик А.С., Логинов Д.В., Ванярхо М.А. Применение нейросетей для прогнозирования пористости в 3D кубах (расчет по сейсмическим атрибутам и каротажным данным) // Приборы и системы разведочной геофизики. – 2012. – № 01 (39). – С. 54–57.
11. Samui P., Roy S.S., Balas V.E. Handbook of Neural Computation. – London: Academic press, 2017. – 631 p.
12. Bonaccorso G. Machine Learning Algorithms. – United Kingdom, Birmingham: Packt Publishing Ltd, 2017. – 333 p.
13. Matheron G. Traité de géostatistique appliquée. – Paris, France: Editions BGRM, 1962. – 460 p.
14. Демьянов В.В., Савельева Е.А. Геостатистика: теория и практика. – М.: Наука, 2010. – 328 с.
15. Степанов Д.Ю., Яппарова Е.А. Разрешающая способность и параметры веерной фильтрации при обработке сейсмических волновых полей // Известия Томского политехнического университета. – 2008. – Т. 312. – № 5. – С. 17–22.
16. Demmel J.W. Applied Numerical Linear Algebra. – California, Berkeley: University of California, 1997. – 435 p.
17. Ковалевский Е.В. Геологическое моделирование на основе геостатистики. – М.: EAGE, 2011. – 117 с.
18. Ильин В.А., Позняк Э.Г. Линейная алгебра. – М.: Наука, 1999. – 297 с.
19. Митрофанов Г.М. Обработка и интерпретация геофизических данных. – Новосибирск, Изд-во НГТУ, 2017. – 168 с.
20. Rasch D., Schott D. Mathematical Statistics. – USA: John Wiley & Sons, 2017. – 672 p.

Поступила 10.11.2018 г.

Информация об авторах

Шестаков В.В., аспирант Национального исследовательского Томского политехнического университета.

Степанов Д.Ю., кандидат технических наук, заместитель директора по науке ООО Научно-аналитический центр «Недра».

UDC 550.8.053:519.2

INFLUENCE OF INITIAL DATA REPRESENTATION ON THE RESULTS OF SIMULATION BY DOUBLE KRIGING METHOD

Valery V. Shestakov¹,
valeriy.shestakov@inbox.ru

Dmitry Yu. Stepanov²,
stepanovdj70@gmail.com

¹ National Research Tomsk Polytechnic University,
30, Lenin avenue, Tomsk, 634050, Russia.

² Scientific Analytic Center «Nedra»,
26, Sovetskaya street, Tomsk, 634029, Russia.

The relevance. When solving inverse geophysical problems, the important task is to construct petrophysical parameters volumetric models. The greatest difficulties in developing the methods for solving this problem are determined by the inhomogeneity of the real geological environment, and their accuracy is defined by the inadequacy of the borehole observations grid. This article presents a new method, the distinguishing feature of which is the joint use of seismic exploration and well logging data. It is based on the geostatistical approach experience and solving the described problems uses the assumption that seismic and borehole data measured within a single geological object can have similar covariance properties. The reliability and effectiveness of the double kriging method have not been previously published, although their research is required for application of the method. One of the first questions here is to assess the impact of the input data representativeness.

The main aim is to consider the influence of the input data representativeness on the modeling quality by the double kriging method and possible ways of developing a quantitative measure of the representativeness estimation.

Objects: model of the seismic attribute cube; 3D CDP and GWL data of the Kontorovichsky Tomsk region deposit.

Methods: theory of random processes, methods of linear algebra, statistical modelling and simulation experiment.

Results. The authors have proved analytically that using the representative data sample entails a simulation error equal to zero. They proved analytically and practically that in the presence of an unrepresentative data sample, the simulation error and the Lagrange multiplier are directly proportional to the weight coefficients of the lacking data. On the basis of this fact, it was suggested that the Lagrange multiplier can be used as a measure of the used sample representativeness. This suggestion was practically verified on the Kontorovichskoye field materials of the Tomsk region, within the framework of which a simulation by double kriging method was carried out using three, five and seven wells. With the increase in the sample from three to seven wells, the value of the Lagrange multiplier was decreased, which confirmed the correctness of the proposed assumption.

Key words:

Statistical modelling, seismic exploration, kriging, petrophysical modeling.

REFERENCES

1. Voskresensky Yu.N. *Polevaya geofizika* [Field geophysics]. Moscow, Nedra Publ., 2010. 479 p.
2. Ermakov. A.P. *Vvedenie v seysmorazvedku* [Introduction to seismic exploration]. Tver, GERS Publ. house, 2012. 160 p.
3. Gurvich I.I., Nomokonov V.P. *Seysmorazvedka. Spravochnik geofizika* [Seismic exploration. Reference book of geophysics]. Moscow, Nedra Publ., 1981. 464 p.
4. Kuznetsov V.I. *Elementy obemnoy (3D) seysmorazvedki* [Elements of three-dimensional (3D) seismic survey]. 2nd ed. Ufa, Infoklady Publ., 2012. 272 p.
5. Koskov V.N., Koskov B.V. *Geofizicheskie issledovaniya skvazhin i interpretatsiya dannykh GIS* [Geophysical well logging and interpretation of GWL data]. Perm, Perm State University Publ., 2007. 304 p.
6. Belyakov E.O., Muhidinov. Sh.V. *Ispolzovanie obobshchennykh zavisimostey dlya postroyeniya petrofizicheskikh modeley filtratsionno-emkostnykh svoystv s otsenkoy granichnykh parametrov vydeleniya kollektorov i opredeleniya ikh kharaktera nasyshtchenosti* [Use of generalized dependencies for petrophysical models construction of filtration-capacitive properties with estimation of the boundary parameters of the reservoirs separation and their saturation character determination]. *Petrofizika slozhnykh kolektorov: problem i perspektivy* [Petrophysics of complex reservoirs: problems and prospects]. Compiler B.N. Enikeev. Moscow, EAGE Geomodel Publ., 2015. 383 p.
7. Shultz Ph.S., Ronen Sh., Hattori M., Corbett Ch. Seismic-guided estimation of log properties (P. 3: A controlled study). *The leading Edge*, 1994, vol. 13, Iss. 5, pp. 305–310.
8. Shestakov V.V., Sysolyatina G.A., Stepanov D.Ju. Three-dimensional models of geoenvironmental parameters. *Advances in computer science research*, 2016, vol. 51: Information technologies in science, management, social sphere and medicine (ITSMSM 2016), pp. 126–129.
9. Hampson D., Schuelke J., Quirein J. Use of Multi-Attribute Transforms to Predict Log Properties from Seismic Data. *Geophysics*, 2001, vol. 66 (1), pp. 220–236.
10. Lavrik A.S., Loginov D.V., Vanyarho M.A. *Primenenie neyrosetey dlya prognozirovaniya poristosti v 3D kubakh (raschet po seysmicheskim atributam i karotazhnykh dannym)* [Application of neural networks to predict porosity in 3D cubes (calculation by seismic attributes and log data)]. *Instruments and systems of exploration geophysics*, 2012, vol. 1 (39), pp. 54–57.
11. Samui P., Roy S.S., Balas V.E. *Handbook of Neural Computation*. London, Academic press, 2017. 631 p.
12. Bonaccorso G. *Machine Learning Algorithms*. United Kingdom, Birmingham, Packt Publishing Ltd, 2017. 333 p.
13. Matheron G. *Traité de géostatistique appliquée* [Treatise of applied geostatistical]. Paris, France, BGRM Publ., 1962. 460 p.
14. Demyanov V.V., Saveleva E.A. *Geostatistika: teoriya i praktika* [Geostatistics: theory and practice]. Moscow, Nauka Publ., 2010. 328 p.

15. Stepanov D.Yu., Yapparova E.A. Resolving power and the parameters of fan filtration during processing of seismic wave fields. *Bulletin of the Tomsk Polytechnic University*, 2008, vol. 5, no. 5, pp. 17–22.
16. Demmel J.W. *Applied Numerical Linear Algebra*. California, Berkeley, University of California, 1997. 435 p.
17. Kovalevskiy E.V. *Geologicheskoe modelirovanie na osnove geostatistiki* [Geological modeling based on geostatistics]. Moscow, EA-GE Publ., 2011. 117 p.
18. Ilin V.A., Poznyak E.G. *Lineynaya algebra* [Linear algebra]. Moscow, Nauka Publ., 1999. 297 p.
19. Mitrofanov G.M. Obrabotka i interpretatsiya geofizicheskikh dannykh [Processing and interpretation of geophysical data]. Novosibirsk, NSTU Publ., 2017. 168 p.
20. Rasch D., Schott D. *Mathematical Statistics*. USA, John Wiley & Sons, 2017. 672 p.

Received: 10 November 2018.

Information about the authors

Valery V. Shestakov, graduate student, National Research Tomsk Polytechnic University.

Dmitry Yu. Stepanov, Cand. Sc., deputy director for Science, Scientific Analytic Center «Nedra».