

МОДЕЛИ В ИНЖЕНЕРНОЙ ГЕОЛОГИИ И ГИДРОГЕОЛОГИИ

УДК 624.131

ОЦЕНКА ТОЧНОСТИ И ДОСТОВЕРНОСТИ ИНЖЕНЕРНО-ГЕОЛОГИЧЕСКИХ МОДЕЛЕЙ НА ОСНОВЕ ПРИНЦИПОВ МАШИННОГО ОБУЧЕНИЯ

© 2023 г. Р. Ю. Жидков^{1,*}, Н. В. Абакумова^{2,**}, Н. Н. Ракитина¹, Г. А. Лесников¹,
В. С. Рекун¹, А. К. Петров¹

¹Московский городской трест геолого-геофизических и картографических работ,
Ленинградский пр. 11, Москва, 123040 Россия

²Московский государственный университет имени М.В. Ломоносова,
Ленинские горы 1, Москва, 119234 Россия

*E-mail: rzhidkov@mgt.ru

**E-mail: abakumova.nv@mail.ru

Поступила в редакцию 09.10.2023 г.

После доработки 29.10.2023 г.

Принята к публикации 20.11.2023 г.

Вопрос оценки достоверности инженерно-геологических моделей до сих пор недостаточно проработан в российской практике. Существенно большее внимание уделено методологии оценки достоверности и качества моделей в области геологии нефти и газа. Сопоставление этих методик с методологией оценки качества моделей машинного обучения показало сходство принципов и подходов, а следовательно, калибровка алгоритмов инженерно-геологического моделирования может выполняться с использованием инструментов оценки качества моделей машинного обучения. В статье описан опыт построения стратиграфо-генетической модели с использованием различных алгоритмических подходов с проведением количественной оценки достоверности моделирования на основе метрик машинного обучения. Показано, что задача построения трехмерной геологической модели может быть рассмотрена и как задача мультиклассовой классификации, и как задача регрессии, а предложенный подход сводит задачу разработки и калибровки алгоритмов к минимизации функции ошибок и позволяет отойти от субъективной оценки качества. Предложен показатель погрешности стратиграфо-генетической модели на основе расчета средних абсолютных ошибок и сопоставления с данными контрольной выборки. Отмечено, что предложенные подходы применимы при тестировании методологий инженерно-геологического моделирования в широком смысле, при этом наиболее сложна проверка прогнозных моделей любого вида. Сделан вывод о необходимости интенсификации разработки и наполнения баз инженерно-геологических данных, что позволит осуществить переход от субъективной экспертной оценки к применению современных подходов к моделированию сложно формализуемых сущностей и явлений.

Ключевые слова: машинное обучение, инженерно-геологическая модель, оценка качества моделей, кросс-валидация, трехмерное инженерно-геологическое моделирование

DOI: 10.31857/S0869780923060115, EDN: CURLAY

ВВЕДЕНИЕ

В 1979 г. британский статистик Дж. Бокс сформулировал ставший афористичным тезис “*все модели ошибочны, но некоторые из них полезны*” [13]. Изначально употребленный в отношении статистических моделей, к настоящему времени он стал общеупотребимым. Признавая его верность в отношении геологических моделей в широком смысле этого понятия, сделаем вывод о том, что качество модели должно определяться полнотой и достоверностью ее содержания. Для того, чтобы приносить пользу, модель должна, как минимум, не вводить своего потребителя в заблуждение.

Трехмерное инженерно-геологическое моделирование в современной постановке – относительно новое направление исследований и изысканий. При этом можно констатировать, что в российской практике не разработана адекватная система оценки точности и достоверности моделей и не выработаны методические обоснования для ее проведения, как, впрочем, не раскрыт вопрос содержания инженерно-геологических моделей.

Методы машинного обучения, в противовес большинству традиционных эмпирических и статистических методов, основываются на “обучении” по выборкам данных для выявления функ-

циональных закономерностей, в том числе в случаях, когда фундаментальные отношения между объектами неизвестны или их физический смысл труднообъясним. В отличие от явных алгоритмов, большинство методов машинного обучения не позволяют проследить ход “принятия решения” в процессе своей работы. Калибровка моделей машинного обучения происходит путем автоматизированного перебора параметров и минимизации функции ошибки, а проверка выполняется на контрольной выборке. Поэтому вопросы оценки качества моделей и работы алгоритмов можно признать основополагающими, а их проработка должна быть выполнена детальным образом.

В рамках этой статьи планируется решение следующих задач:

- определить понятие “инженерно-геологическая модель”;
- рассмотреть подходы к оценке инженерно-геологических моделей в зарубежной практике;
- рассмотреть подходы к качественной и количественной оценке достоверности и адекватности геологических моделей, применяемых в смежных геологических специальностях, и оценить их применимость в практике инженерно-геологического моделирования;
- оценить применимость при создании инженерно-геологических моделей подходов к оценке адекватности моделей, метрик и критериев, используемых в области машинного обучения;
- выполнить оценку достоверности стратиграфо-генетической модели на репрезентативном примере.

Понятия “инженерно-геологическая модель” и “оценка качества” в российской и зарубежной практике

Прежде чем говорить о качестве и достоверности, определимся с понятием “инженерно-геологическая модель”. В СП 47.13330.2012 “Инженерные изыскания для строительства. Основные положения”¹ инженерно-геологическая модель определяется как *совокупность информации о пространственном положении инженерно-геологических элементов в сфере взаимодействия объекта и геологической среды*. Однако в действующем в настоящее время СП 47.13330.2016² данное определение отсутствует.

В соответствии с СП 23.13330.2018 “Основания гидротехнических сооружений”³ *инженерно-геологическая модель – схематизированное отображение размещения в области влияния сооружения инженерно-геологических элементов, наделенных*

постоянными нормативными и расчетными значениями характеристик.

СП 333.1325800.2020 “Информационное моделирование в строительстве”⁴ содержит термин “Цифровая модель местности”, которая должна, в числе прочего, включать результаты инженерно-геологических изысканий. Вопросы о том, должна ли эта модель содержать результаты пространственной интерпретации данных в межскважинном пространстве, в каких цифровых форматах она должна быть предоставлена, какими должны быть иерархия и структура элементов модели, остаются дискуссионными. При этом в соответствии с Постановлением Правительства РФ № 331 от 5 марта 2021 г. (с изменениями от 1 сентября 2023 г.), с 1.07.2024 г. такие модели должны стать обязательным объектом государственной экспертизы для объектов, финансируемых с привлечением средств бюджетов бюджетной системы РФ.

Г.Г. Болдырев определяет процесс трехмерного инженерно-геологического моделирования как *представление данных топологии (связности), геометрии и свойств геологической среды*. При этом автор отмечает, что процесс инженерно-геологического моделирования не требует восстановления геологической истории района изысканий, а скорее представляет собой интерполяцию между данными из горных выработок при соблюдении основных правил и принципов [1].

В 2022 г. Международной ассоциацией по инженерной геологии и окружающей среде (IAEG) опубликован обзор, содержащий теоретические и практические вопросы создания и применения инженерно-геологических моделей. В этом обзоре инженерно-геологическая модель определена как *комплексный набор данных, позволяющий логически оценивать и интерпретировать геологические, геоморфологические и гидрогеологические условия, которые могут повлиять на проекты сооружений и их инженерные характеристики*. Она может включать в себя как концептуальные компоненты, так и результаты фактических наблюдений и испытаний, и может состоять из ряда взаимосвязанных моделей [14]. В нашей работе мы будем придерживаться этого определения как более полного и широкого.

В [14] приведен раздел по оценке качества инженерно-геологических моделей, где отмечается, что геологическая модель должна обеспечиваться достаточной точностью и надежностью, разумно приближаясь к фактическим условиям и характеристикам геологической среды. Повышение точности модели происходит за счет накопления инженерно-геологической информации на различных этапах исследований, которые проводятся по мере реализации проекта, в рамках которого про-

¹ <https://docs.cntd.ru/document/1200096789>

² <https://minstroyrf.gov.ru/docs/13629/>

³ <https://docs.cntd.ru/document/553863425>

⁴ <https://docs.cntd.ru/document/573514520>

изводится моделирование. Достоверность трехмерной геологической модели при этом определяется качеством и объемом данных, используемых при ее создании, методикой моделирования, а также квалификацией специалистов, осуществляющих ее построение.

Оценка качества построения модели зависит от уровня разработки трехмерной геологической модели:

- 1 уровень – упрощенная геологическая модель. Оценка качества осуществляется внутри подразделения организации, в котором построена модель (внутренний контроль);
- 2 уровень – базовая инженерно-геологическая модель. Оценка качества осуществляется специалистом из другого подразделения той же организации (внешний контроль);
- 3 уровень – детально проработанная инженерно-геологическая модель. Оценка качества осуществляется группой сторонних экспертов (независимая экспертиза).

Все результаты качественных и количественных оценок должны быть задокументированы. **Качественная оценка** модели осуществляется оценкой исходных данных моделирования, проверкой согласованности геологических границ, а также методом визуальной экспертной оценки построенной модели. **Полуколичественный метод** заключается в оценке компонентов модели с начислением баллов за качество построения каждого компонента. **Количественная оценка** включает в себя ряд методов, в том числе моделирование случайных полей и метод Монте-Карло, геостатистические методы, например, методы кригинга, стохастическое моделирование [14].

Принципы оценки качества геологических моделей, применяемые в области геологии нефти и газа

Исторически и технологически сложилось так, что сфера трехмерного геологического моделирования наиболее широкое развитие имеет в области нефтяной и газовой разведки. Начиная с конца 1980-х гг. в этой сфере разрабатываются и используются программные пакеты, предназначенные для трехмерного моделирования геологических структур [6]. В области инженерной геологии трехмерные модели на тот момент представляли собой концептуальные иллюстрации [14]. Современные программы, применяемые в нефтегазовой отрасли, обрабатывают результаты бурения скважин, сейсмической, гравитационной, магнитной, электрической разведки и радиометрии. Существующие методы и технологии работ позволяют в ходе геофизических исследований создавать трехмерные модели, которые отражают геологическую структуру и свойства подземных формаций и могут использоваться для проведе-

ния различных оценок и прогнозов. Разработаны основные методологические подходы к построению моделей, определены этапы их построения. Кроме того, выработан ряд критериев, как количественных, так и качественных, позволяющих судить о достоверности и адекватности моделей [5].

Качественные критерии основаны на субъективной экспертной оценке при рассмотрении трехмерной визуализации моделей, горизонтальных срезов и вертикальных сечений, вариограмм, гистограмм распределения показателей и т.д., а количественные – на вычислении параметров, численно отражающих степень сходства построенной модели и результатов расчета по фактическим данным. Построенная модель может быть оценена как адекватная в случае, если результаты оценки попадают в установленный интервал расхождения. В частности, для сейсморазведочных работ приняты следующие количественные характеристики ошибок структурных построений [8]:

– способ вычисления стандартного отклонения глубин, определенных по сейсмической карте, в точках скважины от скважинных отметок глубин горизонта (фактически – метод сравнения с эталонной поверхностью) [8];

– метод “выколов” или эталонной скважины. Оценка погрешности прогноза, основанная на вычислении ошибки при последовательном исключении одной скважины [8];

– метод скользящего экзамена, предусматривающий разделение всего набора данных на обучающую (70–90%) и контрольную (10–30%) выборки. Проводятся построение глубинно-скоростной модели среды с использованием только обучающей выборки и последующая оценка погрешности построений в точках контрольной выборки. Далее все повторяется на следующих контрольных выборках, так что, в конечном счете, охватываются все скважины [8];

– метод сопоставления распределения мощностей пород “куба литологии” со статистикой по данным буровых скважин [5].

Принципы оценки качества моделей машинного обучения

Машинное обучение находит широкое применение при решении задач классификации и регрессии. Классификация – это разбиение множества наблюдений на группы по двум (бинарная классификация) или более признакам (мультиклассовая классификация). Регрессия, в свою очередь, представляет собой зависимость математического ожидания случайной величины от одной или нескольких других случайных величин и может быть описана как линейной функцией, так и нелинейной.

Оценка сходимости прогнозируемых и фактических результатов выполняется с применением

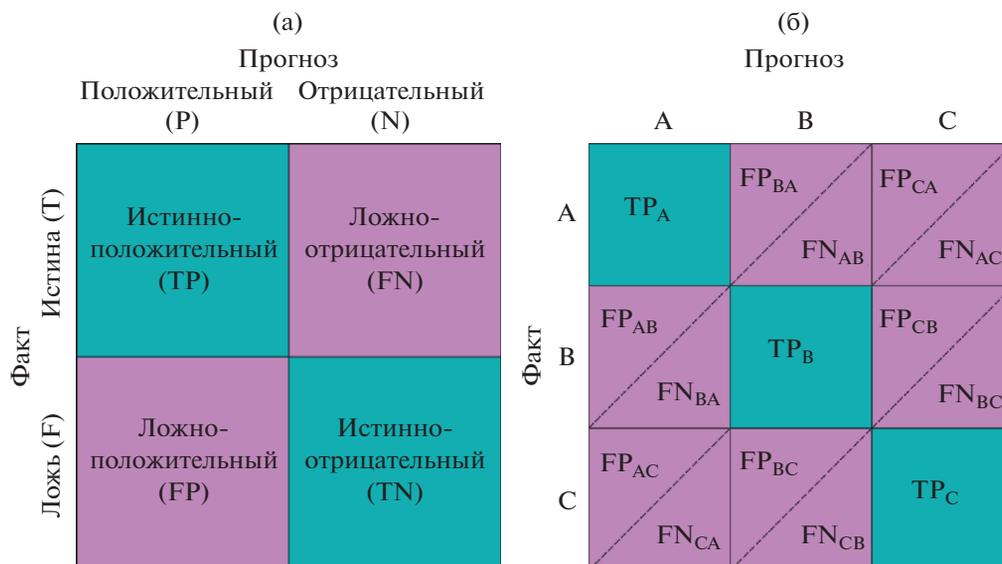


Рис. 1. Матрица ошибок для бинарной (а) и мультиклассовой (б) классификации.

критериев качества работы – метрик, которые отличаются в зависимости от типа решаемой задачи. Метрики для оценки решения задачи классификации основываются на матрицах ошибок (рис. 1). Прогнозируемые значения описываются в случае бинарной классификации как положительные (positive) и отрицательные (negative), а фактически – как истинные (true) и ложные (false). В таком случае получается четыре комбинации проверки прогнозируемых результатов:

- истинно-положительные (TP) – фактически положительные значения были распознаны как положительные;
- истинно-отрицательные (TN) – фактически отрицательные значения были распознаны как отрицательные;
- ложно-положительные (FP) – отрицательные значения были ошибочно распознаны как положительные;
- ложно-отрицательные (FN) – положительные значения были ошибочно распознаны как отрицательные.

Для мультиклассификации результат может быть расценен как положительный, когда спрогнозированный класс совпал с фактическим (TP_A, TP_B, TP_C), либо как отрицательный с указанием ошибочно присвоенного класса (рис. 1б). Например, результат FP_{CB} означает, что при работе алгоритма машинного обучения объекту выборки был ошибочно присвоен класс С при фактическом В.

Выбор метрик определяется задачами, которые должна решать разработанная модель машинного обучения и характером данных. Многие метрики дают некорректный результат в случае, если набор данных характеризуется неравномер-

ным распределением, т.е. имеет место дисбаланс классов. В табл. 1 приведены наиболее часто используемые метрики для бинарной классификации и их аналоги для мультиклассовой

Метрика *точность* (accuracy) хорошо работает при равномерном распределении классов и сообщает долю корректно классифицированных примеров. *Положительная прогностическая значимость*, *ППЗ* (precision), представляет собой вероятность того, что результат действительно положительный при предсказании положительного класса. *Полнота* (recall) – это доля примеров положительного класса, которые были корректно распознаны. *F-мера* (f-score) рассчитывается как среднее гармоническое ППЗ и полноты и позволяет сглаживать расчеты за счет исключения экстремальных значений. Остальные метрики, такие как индекс Юдена (Youden's index), сбалансированная точность (Balanced accuracy), коэффициент Коэна (Cohen's Kappa) и др., используются существенно реже при решении специфических задач [9, 15].

Для мультиклассификации аналогичные метрики вычисляются различными методами [12] (см. табл. 1).

Макро-усредненное (macro-averaged) определяется как среднее рассчитанных индивидуально для каждого класса метрик. И малые, и большие по количеству событий классы входят в расчет метрики с одинаковым весом. То есть одно событие в малом классе более значимо, чем одно событие в большом.

Микро-усредненное (micro-averaged) – рассчитывается метрика для всех событий, безотносительно классов. При таком подходе самый большой класс

Таблица 1. Основные метрики оценки качества классификаций

Бинарная классификация	Мультиклассификация	
<p>Точность (accuracy)</p> $AC = \frac{TP + TN}{FP + FN + TP + TN}$	<p>Микро-усредненная точность (micro-averaged accuracy)</p> $AC_{\text{micro}} = \frac{\sum_{i=1}^C TP_i}{\sum_{i=1}^C TP_i + \sum_{i=1}^C FN_i + \sum_{i=1}^C FP_i}$	<p>Макро-усредненная точность (macro-averaged accuracy)</p> $AC_{\text{macro}} = \frac{1}{C} \sum_{i=1}^C AC_i,$ <p>где $AC_i = \frac{TP_i}{TP_i + FN_i + FP_i}$</p>
<p>ППЗ (precision)</p> $PR = \frac{TP}{TP + FP}$	<p>Микро-усредненная ППЗ (micro-averaged precision)</p> $PR_{\text{micro}} = \frac{\sum_{i=1}^C TP_i}{\sum_{i=1}^C TP_i + \sum_{i=1}^C FP_i}$	<p>Макро-усредненная ППЗ (macro-averaged precision)</p> $PR_{\text{macro}} = \frac{1}{C} \sum_{i=1}^C PR_i, \text{ где } PR_i = \frac{TP_i}{TP_i + FP_i}$
<p>Полнота (recall)</p> $R = \frac{TP}{TP + FN}$	<p>Микро-усредненная полнота (micro-averaged recall)</p> $R_{\text{micro}} = \frac{\sum_{i=1}^C TP_i}{\sum_{i=1}^C TP_i + \sum_{i=1}^C FN_i}$	<p>Макро-усредненная полнота (macro-averaged recall)</p> $R_{\text{macro}} = \frac{1}{C} \sum_{i=1}^C R_i, \text{ где } R_i = \frac{TP_i}{TP_i + FN_i}$
<p>F-мера (f-score)</p> $F_1 = 2 \cdot \frac{PR \cdot R}{PR + R}$	<p>Микро-усредненная f-мера (micro-averaged f-score)</p> $F_{\text{micro}} = 2 \cdot \frac{Pr_{\text{micro}} \cdot R_{\text{micro}}}{Pr_{\text{micro}} + R_{\text{micro}}}$	<p>Макро-усредненная f-мера (macro-averaged f-score)</p> $F_{\text{macro}} = 2 \cdot \frac{Pr_{\text{macro}} \cdot R_{\text{macro}}}{Pr_{\text{macro}} + R_{\text{macro}}}$

Примечание: C – количество классов.

Таблица 2. Основные метрики оценки качества регрессионных моделей

Название метрики	Формула	Описание/особенности	Единицы измерения
Среднеквадратичная ошибка (Mean Squared Error, MSE)	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$	Применяются в случаях, когда важно избежать частных ошибок. Метрики “штрафуют” за большие в абсолютном выражении ошибки и, наоборот, снижают влияние малых ошибок. В метрике MSE за счет квадратичной зависимости аномальные значения непропорционально влияют на подсчет ошибки, поэтому метрика более чувствительна к ним	x^2
Корень из среднеквадратичной ошибки (Root Mean Squared Error, RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$		x
Средняя абсолютная ошибка (Mean Absolute Error, MAE)	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $	Рассчитывается как среднее абсолютных разностей между фактическим и предсказанным значением, поэтому является линейной оценкой. Это значит, что все ошибки в среднем взвешены одинаково	x
Коэффициент детерминации (Coefficient of determination, R^2)	$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2}$	Показывает долю дисперсии зависимой переменной, объясненной с помощью регрессионной модели	ед.

Примечание: x – единицы измерения значений выборки.

оказывает наибольшее влияние на расчет метрики. Метрики малых классов могут затеряться.

При определении *средневзвешенного* (*weighted-averaged*) малые и большие классы входят в расчет метрики с весом пропорционально вкладу в выборку. Влияние каждого события зависит от размера выборки и соотношения классов.

Допускается использовать произвольный подход к осреднению метрик в зависимости от решаемых задач, если классовый состав выборки сбалансирован [3].

В задачах регрессии самыми популярными метриками являются среднеквадратичная MSE, корень из среднеквадратичной ошибки RMSE, средняя абсолютная ошибка MAE и коэффициент детерминации R^2 (табл. 2).

Коэффициент детерминации R^2 изменяется в диапазоне от $-\infty$ до 1. Если он равен 1, это означает, что сумма квадратов их отклонений равна 0. Такая ситуация соответствует идеальной модели, когда все точки наблюдений лежат точно на линии регрессии. Если коэффициент детерминации равен 0, то связь между переменными регрессионной модели отсутствует. При отрицательном значении коэффициента детерминации ошибка модели простого среднего становится меньше ошибки регрессионной модели. Таким образом, добавление в модель с константой некоторой переменной только ухудшает ее.

Кросс-валидация, или скользящий контроль – метод, предназначенный для оценки качества работы алгоритмов, широко применяемый в машинном обучении. Он помогает сравнить между собой различные алгоритмы построения модели (в самом широком смысле понятия) и выбрать наилучший для конкретной задачи. При создании инженерно-геологических моделей этот подход позволяет статистически обосновать методологию моделирования.

Метод заключается в последовательном многократном разбиении выборки исходных данных на две группы (обучающую и контрольную) и вычислении средней оценки качества моделей, построенных на разных обучающих выборках. В зависимости от способа разбиения могут быть следующие варианты кросс-валидации (список не исчерпывающий):

– *полный скользящий контроль*, когда выборка разбивается всеми возможными способами на две части $T' = T \setminus T^{l-t}$, где t и l – количество элементов в тестовой и обучающей выборках соответственно;

– *контроль по k -блокам*, когда обучающая выборка случайным образом разбивается на k непересекающихся блоков одинакового размера, затем обучает и оценивает модель дерева решений k раз, каждый раз выбирая новый блок для оценки и тренируясь на остальных $k-1$ блоках (как

правило, $k = 10$ или 5 в случае малого размера выборки);

– *кросс-валидация по отдельным объектам*, когда выборка разбивается на $n-1$ и 1 объект n раз;

– *случайные разбиения*, когда выборка разбивается в случайной пропорции [15].

Выбор метода кросс-валидации зависит от типа поставленной задачи, количества и соотношения классов, размера выборки, вычислительных мощностей.

При решении задач машинного обучения оценка качества работы алгоритмов выполняется не только и не столько при оценке результатов работы алгоритмов, но и в процессе обучения – при автоматизированном подборе параметров модели. При этом происходит минимизация функции ошибки, которая вычисляется на обучающей выборке с использованием выбранной метрики. Пользователь может управлять процессом, задавая *гиперпараметры*, которые не изменяются в процессе обучения. В широком смысле в качестве гиперпараметров можно рассматривать алгоритмы и их комбинации, выбор метрики для оценки ошибки, заданные диапазоны изменения параметров. В статье, опубликованной в 2021 г. [16], отмечается, что в обозримом будущем при анализе больших массивов данных процесс калибровки моделей в сфере наук о Земле должен уступить место технологиям глубокого обучения. Отметим, что этот подход не может быть реализован на локальном уровне, а требует разработки региональных баз данных и существенных вычислительных мощностей. Также признаем, что определенным ограничением для реализации этого подхода в настоящее время можно считать дефицит специалистов, компетентных как в области инженерной геологии, так и анализа данных и программирования, что должно учитываться при подготовке новых и обновлении существующих учебных планов в профильных высших учебных заведениях [7].

Оценка качества стратиграфо-генетической модели на примере участка в Москве

В соответствии с методологией, приведенной в руководстве [14], процесс создания инженерно-геологической модели участка должен начинаться с разработки концептуальной геологической (стратиграфо-генетической) модели. Эта методология хорошо согласуется с принципами генетического подхода к изучению грунтов, ставшего методологической основой современного грунтоведения [11]. С практической точки зрения при создании инженерно-геологической модели стратиграфо-генетическая модель представляет собой структурный каркас для оценки распределения инженерно-геологических характеристик в пространстве, которая в зависимости от задачи



Рис. 2. Геоморфологическая схема участка исследования (автор А.К. Петров).

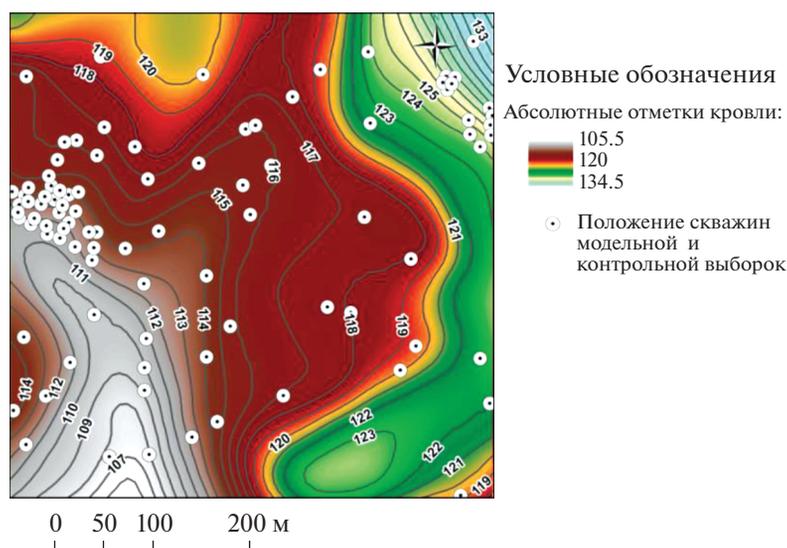


Рис. 3. Карта рельефа кровли дочетвертичных отложений (по материалам геологического атласа [2]).

может быть выполнена в детерминированном или стохастическом варианте. Кроме того, стратиграфо-генетическая модель служит основой для выполнения геофильтрационного моделирования. Таким образом, конфигурация стратиграфо-генетической модели может существенным образом повлиять на результаты дальнейших инженерно-геологических и гидрогеологических исследований, и оценка ее качества — основополагающий этап в процессе инженерно-геологического моделирования.

Рассмотрим процедуру оценки качества с применением метода кросс-валидации на примере участка в центральной части Москвы. Участок ре-

презентативен для оценки, потому что характеризуется высокой степенью инженерно-геологической изученности и изменчивостью инженерно-геологических условий. На участке выделяется 7 геоморфологических элементов, включая пойму р. Москвы и ее погребенного притока (рис. 2), а рельеф дочетвертичных отложений характеризуется высокой расчлененностью, так как участок приурочен к борту доледниковой долины (рис. 3).

Моделирование и оценка качества моделей выполнялись на основе 95 буровых скважин глубиной от 5 до 140 м (в среднем 21 м). Средняя обеспеченность скважинами составила 4 выработки на гектар участка, однако, как видно на

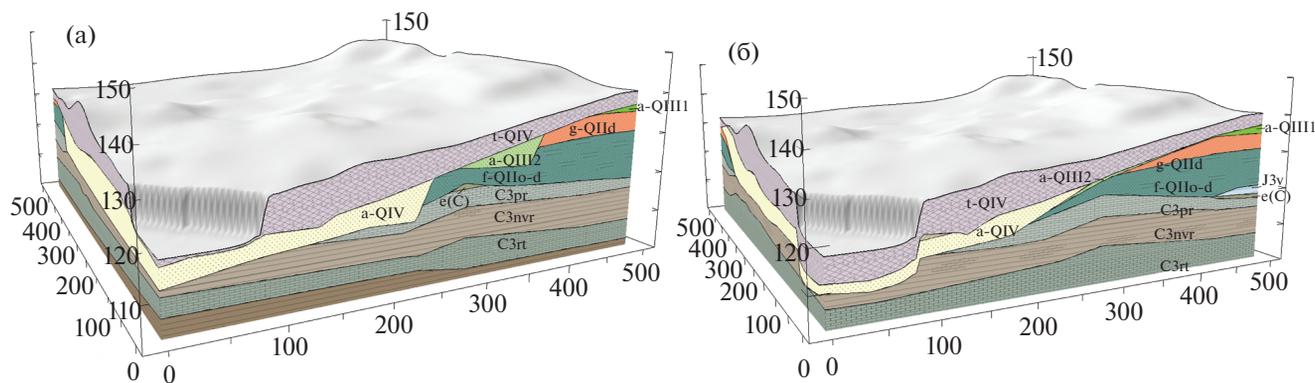


Рис. 4. Внешний вид трехмерной стратиграфо-генетической модели (вариант с соотношением модельной и контрольной выборки 1:1), построенной на основе идентичной выборки скважин: а – по авторской методике, б – послойно с применением триангуляционных построений.

рис. 2, 3, распределение скважин по площади неравномерно. Данные были получены в процессе проведения инженерно-геологических изысканий разных лет, обобщены и сведены между собой в рамках работы над проектами крупномасштабного геологического картирования г. Москвы и разработки региональной трехмерной геологической модели [2, 10].

Моделирование выполнялось в трех альтернативных алгоритмических реализациях:

- последовательное моделирование напластования с применением триангуляционного построения поверхностей;

- последовательное моделирование напластования с применением ординарного кригинга с автоматическим подбором вариограммы;

- по комплексной авторской методике, разработанной для построения региональной трехмерной геологической модели г. Москвы, учитывающей историю техногенной трансформации и условия осадконакопления в московском регионе [4, 10]. Методология разрабатывалась в 2017–2019 гг., а калибровка модельных параметров осуществлялась в основном путем визуальной экспертной оценки корректности модельных разрезов в программе “Геонавигатор”.

Во всех трех случаях построение модели осуществлялось в автоматизированном режиме: в качестве программной среды для моделирования использовался пакет Safe Software FME с интеграцией модуля RBF (радиальная базисная функция) в составе свободно распространяемой версии библиотеки Alglib и пакета geoR для языка программирования R.

Оценка качества выполнялась с использованием метода кросс-валидации с пятикратным разбиением выборки на модельную и контрольную. Для оценки влияния плотности входных данных на точность моделирования разделение

на выборки выполнялось в трех соотношениях: 1:1, 7:3, 9:1. Таким образом, суммарно было построено 45 вариантов модели геологического строения участка с применением различных алгоритмов и на разных наборах данных.

Визуальная оценка результатов моделирования показывает общее структурное сходство построенных моделей, хотя и с локальными расхождениями. Например, на моделях, представленных на рис. 4, расхождения могут быть выявлены в части пространственной характеристики толщи техногенных отложений, смыкания речных террас, выявленных выветрелых зон в кровле каменноугольных отложений и конфигурации толщи верхнеюрских отложений. Отметим, что все результаты прошли проверку на соответствие входным данным, и расхождения проявляются в не обеспеченных входными данными блоках пространства моделей.

Построение геологической модели может быть рассмотрено и как задача классификации (определение принадлежности каждого элементарного блока модели к тому или иному стратиграфо-генетическому комплексу), и как задача регрессии (предсказание отметок и мощностей залегания пластов в каждой точке пространства).

Выполним оценку точности моделирования с применением классификационных и регрессионных метрик. Для анализа данных используем пакет Scikit-learn для языка программирования Python, содержащий инструменты для работы с методами машинного обучения.

Для оценки качества классификации выполним сопоставление классов (меток стратиграфо-генетических комплексов) в геологических скважинах контрольной выборки с классами в “виртуальных” скважинах, полученных на основе модели. Для этого проведем дискретизацию описа-

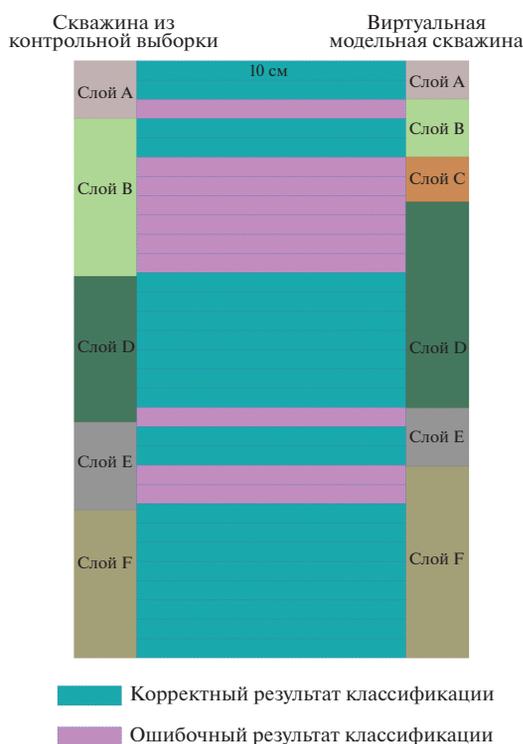


Рис. 5. Пример дискретизации контрольных и модельных скважин для оценки качества классификации.

ния как контрольных, так и виртуальных скважин с шагом 0.1 м (рис. 5).

В результате получим массив данных, содержащий координатную привязку и метки предсказанного и фактического класса. Отметим, что в выборке присутствует существенный дисбаланс классов (рис. 6). В данной задаче будем использовать микроусредненную f -меру, поскольку для общей оценки работы алгоритма моделирования

важна каждая ошибка, независимо от размера класса (табл. 3).

Для оценки качества моделирования с применением регрессионных метрик выполним сопоставление модельных и фактических значений абсолютных отметок кровли (z) и мощности каждого геологического слоя (H). Вычислим среднюю абсолютную погрешность MAE для этих значений и возьмем их среднее арифметическое значение. В результате получим абсолютное значение показателя погрешности стратиграфо-генетической модели в метрах (табл. 4):

$$MAE_{mdl} = (MAE_z + MAE_H)/2,$$

где MAE_z – средняя абсолютная ошибка значений кровли геологических слоев в скважинах; MAE_H – средняя абсолютная ошибка значений мощности геологических слоев в скважинах.

ВЫВОДЫ

Предложенная методология оценки достоверности инженерно-геологических моделей основана на применении технологии кросс-валидации и метрик, регрессии и/или классификации. На наш взгляд, при оценке достоверности стратиграфо-генетических моделей, абсолютный показатель погрешности моделирования удобнее, так как не зависит от выбранной дискретизации и выражается в метрических единицах. В качестве преимуществ рассмотрения процесса моделирования, как задачи классификации, можно назвать возможность построения матриц ошибок, которые могут использоваться для тонкой отладки алгоритмов и в процессе инженерно-геологической схематизации.

Отметим, что оценка погрешности по вышеизложенной методике может быть выполнена как в отношении работы алгоритмов при автоматизи-

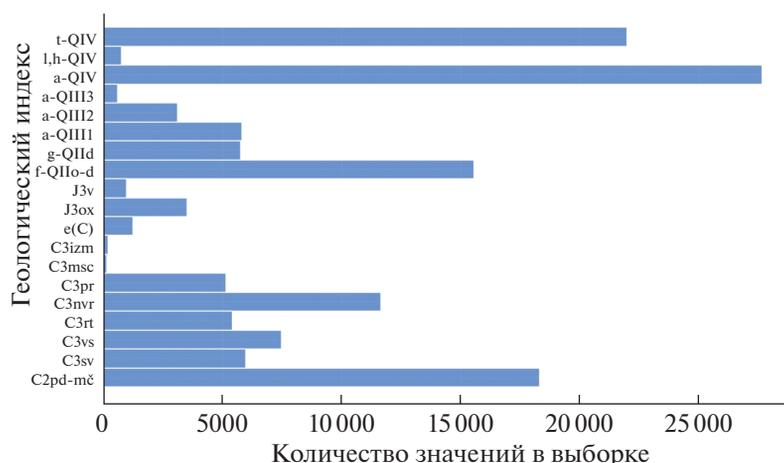


Рис. 6. Распределение значений классов в выборке.

Таблица 3. Результаты расчета микроусредненной f -меры для стратиграфо-генетических моделей

Алгоритм моделирования	Соотношение модельной и тестовой выборок		
	1:1	7:3	9:1
Последовательное моделирование напластования (триангуляция)	0.72	0.76	0.77
Последовательное моделирование напластования (кригинг)	0.74	0.76	0.79
Авторский алгоритм, учитывающий последовательность осадконакопления	0.75	0.78	0.81

Таблица 4. Результаты расчета показателя погрешности стратиграфо-генетической модели в метрах

Алгоритм моделирования	Соотношение модельной и тестовой выборок		
	1:1	7:3	9:1
Последовательное моделирование напластования (триангуляция)	1.53	1.46	1.45
Последовательное моделирование напластования (кригинг)	1.50	1.42	1.36
Авторский алгоритм, учитывающий последовательность осадконакопления	1.39	1.34	1.23

рованном построении моделей, так и при моделировании в специализированных программных пакетах в ручном режиме, при условии соблюдения разделения выборки на рабочую и контрольную. Однако такой подход применим либо в научных целях, либо при оценке квалификации специалистов, поскольку в практике изыскательской работы построение модели на неполном наборе данных вряд ли имеет смысл.

Рассмотренный в статье пример не претендует на обоснование методологии стратиграфо-генетического моделирования даже в условиях геологии Москвы, поскольку охарактеризован конкретный участок с определенным геологическим строением. Рассмотрение результатов, полученных в разных геологических условиях, — одно из направлений дальнейших исследований в этой области. Более того, мы рассчитываем, что выработанная методология оценки качества алгоритмов позволит усовершенствовать технологию моделирования, используемую при построении региональной трехмерной геологической модели Москвы с тем, чтобы повысить качество моделирования и оптимизировать работу алгоритмов. Предложенный подход сводит задачу разработки и калибровки алгоритмов к минимизации функции ошибок и позволяет отойти от субъективной оценки корректности геологического строения, хотя и не отменяет необходимость визуального контроля результатов моделирования.

В рамках данного исследования никак не рассматривался вопрос качества входных данных для модели. Использовалась выверенная и сведенная выборка данных, что позволило получить близкие характеристики точности вне зависимости от алгоритма моделирования. Мы предполагаем,

что качество и характер данных определяют результат в гораздо большей степени, чем алгоритмы интерполяции и моделирования, как бы откалиброваны они ни были, и планируем изучить этот вопрос в дальнейших исследованиях.

Предложенные подходы применимы при тестировании методологии инженерно-геологического моделирования в самом широком смысле. Граница применимости методологии проходит там, где исключается возможность проверяемости результатов моделирования — сопоставления модельного предположения с реальностью. В этом отношении наиболее сложна проверка прогнозных моделей любого характера. Поэтому крайне важно сейчас, когда для этого есть все технические возможности, организовывать сбор баз данных разнообразных инженерно-геологических данных, включая полевые и лабораторные испытания, результаты их обработки, выполненные на их основе прогнозные оценки и выводы, мониторинговые измерения разного рода, данные дистанционного зондирования Земли и т.д. Только при наличии больших данных в инженерной геологии будет возможен переход от субъективной экспертной оценки к применению современных подходов к моделированию сложно формализуемых сущностей и явлений.

СПИСОК ЛИТЕРАТУРЫ

1. *Болдырев Г.Г.* Трехмерное моделирование и визуализация данных инженерно-геологических изысканий. Состояние вопроса и практическая реализация // *Инженерные изыскания*. 2022. Т. XVI. № 1. С. 8–26.
2. *Геологический атлас Москвы* (в 10 томах с пояснительной запиской). Масштаб 1:10 000 / Под ред.

- А.В. Антипова. М: Изд-во ГУП “Мосгоргеотрест”, 2012.
3. Гулин В.В. Методы снижения размерности признакового описания документов в задаче классификации текстов // Вестник Московского энергетического института. 2013. № 2. С. 115–121.
 4. Жидков Р.Ю., Абакумова Н.В., Рекун В.С. Применение комплексного ретроспективного анализа при определении конфигурации массивов техногенных грунтов на примере г. Москвы // Инженерная геология. 2023. Т. XVIII. № 1. С. 18–34.
 5. Закревский К.Е., Майсюк Д.М., Сыртланов В.Р. Оценка качества 3D моделей. М.: ООО “ИПЦ “Маска”, 2008. 272 с.
 6. Закревский К.Е., Попов В.Л. История развития трехмерного геологического моделирования как метода изучения залежей нефти и газа // Известия Томского политехнического университета. Инжиниринг георесурсов. 2021. № 5. С. 89–100.
 7. Королев В.А. О задачах цифровизации и искусственного интеллекта в инженерной геологии // Инженерная геология. 2021. Т. XVI. № 1. С. 10–23.
 8. Левянт В.Б., Ампилов Ю.П., Глоговский В.М. и др. Методические рекомендации по использованию данных сейсморазведки (2D, 3D) для подсчета запасов нефти и газа. М.: Министерство природных ресурсов. ОАО “Центральная геофизическая экспедиция”, 2006. 39 с.
 9. Михайличенко А.А. Аналитический обзор методов оценки качества алгоритмов классификации в задачах машинного обучения // Вестник Адыгейского государственного университета. Сер. 4: Естественно-математические и технические науки. 2022. № 4 (311). С. 52–59.
 10. Технология трехмерного моделирования геологической среды и ее апробация на объектах жилищного фонда города Москвы по программе реновации // 75 лет инженерных изысканий в Москве / Под ред. А.Ю. Серова и др. Тверь: Талан Групп, 2019. С. 185–189.
 11. Трофимов В.Т., Королев В.А. О фундаментальных аспектах генетического подхода к изучению грунтов // Инженерная геология. 2019. Т. XIV. № 1. С. 8–19.
 12. Шенета Д.А., Головин Е.Н., Иванова М.С. Метрики качества алгоритмов машинного обучения в задачах мультиклассификации // Научная сессия ГУАП: сб. докладов научной сессии, посв. Всемирному дню авиации и космонавтики. В 3-х частях, Санкт-Петербург, 08–12 апреля 2019 г. Часть II. СПб: ГУАП, 2019. С. 278–281.
 13. Box G.E.P. Robustness in the Strategy of Scientific Model Building / R.L. Launer, G.N. Wilkinson (ed.) // Robustness in Statistics – Academic Press. 1979. P. 201–236.
 14. Baynes F.J., Parry S. Guidelines for the development and application of engineering geological models on projects. International Association for Engineering Geology and the Environment (IAEG) Commission 25 Publication № 1. 2022. 129 pp.
 15. Geron A. Hands-On Machine Learning with Scikit-Learn and TensorFlow, Sebastopol: OREILLY. 2019. 548 pp.
 16. Tsai W.P., Feng D., Pan M., et al. From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling // Nature Communications. 2021. V. 12 (1). P. 1–13.

ACCURACY AND RELIABILITY ASSESSMENT OF ENGINEERING GEOLOGICAL MODELS BASED ON MACHINE LEARNING

R. Yu. Zhidkov^{a,#}, N. V. Abakumova^{b,##}, N. N. Rakitina^a, G. A. Lesnikov^a, V. S. Rekun^a, and A. K. Petrov^a

^aState Budgetary Institution “Mosgorgeotrest”,
Leningradsii pr. 11, Moscow, 123040 Russia

^bLomonosov Moscow State University,
Leninskie gory 1, Moscow, 119234 Russia

[#]E-mail: rzhidkov@mggt.ru

^{##}E-mail: abakumova.nv@mail.ru

Assessing reliability of engineering geological models still requires further studies. Much more attention is paid to the methodology for assessing the reliability and quality of models in oil and gas geology. A comparison of the methods used in this field with the methodology for assessing the quality of machine learning models showed the similarity of principles and approaches. Therefore, the algorithms for engineering geological modeling can be justified and calibrated using tools for evaluating the quality of machine learning models. The article systematizes and analyzes the metrics used in solving problems and classification, and describes the methodology of using cross-validation techniques to assess the quality of algorithms. The practical experience of constructing a computer stratigraphic-genetic model using various algorithmic approaches is described: on the basis of triangulation constructions, using ordinary kriging with automated variogram composition and using the custom heuristic algorithm that considers the history of sedimentation and technogenic transformation of the territory. It is shown that the problem of three-dimensional geological modeling can be considered both as a classification and regression problem. The error index of the stratigraphic-genetic

model is proposed based on the calculation of the average absolute errors in determining the spatial position of geological layers. The proposed approaches are applicable to testing methodologies of engineering geological modeling in a broad sense, the verification of predictive models of any kind being the most difficult issue. It is emphasized that the development and filling of databases of various engineering and geological data should be intensified, including field and laboratory data, the results of their processing, forecast estimates and conclusions based on them, monitoring measurements of various kinds, remote sensing data, etc. The possibilities of processing and analyzing big data in engineering geology will allow us to move from subjective expert estimations to the application of modern approaches to modeling complexly formalized objects and phenomena using the capabilities of machine learning and artificial intelligence.

Keywords: *machine learning, engineering geological model, model precision measurement, cross-validation, three-dimensional engineering geological modeling*

REFERENCES

1. Boldyrev, G.G. [Three-dimensional modeling and visualization of engineering geological survey data. State-of-the-art and practical implementation]. *Inzhenernye izyskaniya*, 2022, vol. XVI, no. 1, pp. 8–26. (in Russian)
2. [Geological atlas of Moscow (in 10 volumes with an explanatory note). Scale 1:10 000]. A.V. Antipov, Ed. Moscow, GUP “Mosgorgeotrest” Publ, 2012. (in Russian)
3. Gulin, V.V. [Methods of reducing the dimension of the characteristic description of documents in the problem of text classification]. *Vestnik Moskovskogo energeticheskogo instituta*, 2013, no. 2, pp. 115–121. (in Russian)
4. Zhidkov, R.Yu., Abakumova, N.V., Rekun, V.S. [Application of complex retrospective analysis in determining the configuration of man-made soil massifs on the example of Moscow]. *Inzhenernaya geologiya*, 2023, vol. XVIII, no. 1, pp. 18–34. (in Russian)
5. Zakrevskii, K.E., Maisyuk, D.M., Syrtlanov, V.R. [Evaluation of the quality of 3D models]. Moscow, ООО “ИПТС “Маска” Publ., 2008, 272 p. (in Russian)
6. Zakrevskii, K.E., Popov, V.L. [The history of the development of three-dimensional geological modeling as a method of studying oil and gas deposits]. *Izvestiya Tomskogo politekhnicheskogo universiteta. Inzhiniring georesursov*, 2021, no. 5, pp. 89–100. (in Russian)
7. Korolev, V.A. [On the problems of digitalization and artificial intelligence in engineering geology]. *Inzhenernaya geologiya*, 2021, vol. XVI, no. 1, pp. 10–23. (in Russian)
8. Levyant, V.B., Ampilov, Yu.P., Glogovskii, V.M. et al. [Methodological recommendations on the use of seismic data (2D, 3D) for calculating oil and gas reserves]. Moscow, Ministry of Natural Resources. Central Geophysical Expedition, 2006, 39 p. (in Russian)
9. Mikhailechenko, A.A. [Analytical review of methods for assessing the quality of classification algorithms]. *Vestnik Adygeiskogo gosudarstvennogo universiteta. Ser. 4: Estestvenno-matematicheskie i tekhnicheskie nauki*, 2022, no. 4 (311), pp. 52–59. (in Russian)
10. [Technology of three-dimensional modeling of the geological environment and its approbation at the objects of the Moscow housing stock under the renovation program]. In: [75 years of engineering surveys in Moscow]. A.Yu. Serov et al, Eds. Tver, Talan Group, 2019, pp. 185–189. (in Russian)
11. Trofimov, V.T., Korolev, V.A. [About the fundamental aspects of the genetic approach to the study of soils]. *Inzhenernaya geologiya*, 2019, vol. XIV, no. 1, pp. 8–19. (in Russian)
12. Shepeta, D.A., Golovin, E.N., Ivanova, M.S. [Quality metrics of machine learning algorithms in multiclassification tasks]. In: [Proc. of the GUAP scientific session on World Aviation and Cosmonautics Day]. St. Petersburg, 2019, Part II, pp. 278–281. (in Russian)
13. Box, G.E.P. Robustness in the strategy of scientific model building. In: Robustness in statistics, R.L. Launer, G.N. Wilkinson, Eds. Academic Press, 1979, pp. 201–236.
14. Baynes, F.J., Parry, S. Guidelines for the development and application of engineering geological models on projects. IAEG Commission 25 Publication, 2022, no. 1, 129 p.
15. Geron, A. Hands-on machine learning with Scikit-Learn and TensorFlow. Sebastopol, OREILLY, 2019, 548 p.
16. Tsai, W.P., Feng, D., Pan, M. et al. From calibration to parameter learning: harnessing the scaling effects of big data in geoscientific modeling. *Nature Communications*, 2021, vol. 12 (1), pp. 1–13. <https://doi.org/10.1038/s41467-021-26107-z>