

СРЕДСТВА И ПРИМЕРЫ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ ДАННЫХ ГЕОЛОГИЧЕСКИХ МОДЕЛЕЙ

Таранчук В. Б.

Кафедра компьютерных технологий и систем, Факультет прикладной математики и информатики,
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: taranchuk@bsu.by

Обсуждаются вопросы использования интегрированного программного комплекса составителя цифровых геологических, геоэкологических моделей. Примерами иллюстрируются предоставляемые системой Mathematica 11 инструменты нейронных сетей для обработки исходных данных цифровых моделей, сопоставления результатов.

ВВЕДЕНИЕ

Моделирование объектов геологии предполагает использование сложных математических методов; программирование, реализация алгоритмов построения и адаптации геологических, геоэкологических моделей трудоемко; для поддержания подобных программных продуктов в актуальном состоянии, отвечающем постоянно возрастающим аппаратным возможностям, нужен коллектив опытных программистов. Учитывая текущие возможности по обеспеченности системами компьютерной математики, эффективным является подход, основанный на интеграции в единый программный комплекс модулей современных версий систем компьютерной алгебры (СКА) и географических информационных систем (ГИС). При этом следует понимать, что для решения задач обработки исходных данных, моделирования никакая конкретная ГИС не является полным набором пространственно-аналитических методов и средств анализа. Во многих случаях приходится комбинировать инструменты ГИС с программами статистического анализа данных, средствами для математически сложных вычислений, алгоритмов анализа и интерпретации пространственных данных.

1. ОСНОВНЫЕ ПОЛОЖЕНИЯ

В [1], [2] описаны предложенные и реализованные методические и технические решения, программные компоненты, которые включены в разработанный интегрированный компьютерный комплекс «Генератор геологической модели залежи» – ГГМЗ. Назначение комплекса тестирование, оценки точности настраиваемых геологических моделей на основе применения СКА, ГИС, «умных» методов адаптации моделей в процессе их эксплуатации, «самонастройки» моделей с учётом дополняемых данных фактического развития процессов. Разработка и функциональное наполнение комплекса ГГМЗ осуществлены с использованием системы компьютерной алгебры Wolfram Mathematica ([3] – [5]) и геоинформационной системы Golden Software Surfer [6]. При

программировании в системе Mathematica модулей графики реализованы технические решения, описанные в [7] – [9]. Предусмотрена возможность, когда программный комплекс в конкретной конфигурации может эксплуатироваться после сборки и сохранения в формате вычисляемых документов CDF [10], [11]. Расчеты и работа пользователя с CDF версией приложения возможны на любом персональном компьютере. При просмотре CDF версии, размещенной на вебсервере, программа просмотра автоматически подгружается в виде плагина браузера. Автономная работа с ПК возможна после установки свободно распространяемого компонента CDF Player. Варианты дополнительных настроек, обеспечивающих интерактивность CDF версии, изложены в [7], [8].

Система Mathematica обеспечивает пользователей широким набором встроенных функций численного анализа [12], методов и функций статистического анализа [13]. Требуемые функции подключаются в ГГМЗ простыми обращениями. Все этапы работы с данными в комплексе включают разнообразные варианты графической визуализации, протоколирование и сопоставление получаемых или помещенных в архив результатов [14], [7], [8]. Инструменты комплекса обеспечивают пользователя возможностями «поиграть» исходными данными и сопоставить результаты с подготовленными эталонами, причем, с возможностями импорта и экспорта данных и изображений, масштабирования. Широкие возможности обмена данными важны при необходимости работы в нескольких программных средах с целью использования уникальных алгоритмов.

II. ИНСТРУМЕНТЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Следует понимать, что создание и сопровождение геологической модели не предполагают нахождения единственного решения некоторой математической задачи. Субъективное мнение, квалификация эксперта – факторы, всегда присутствующие в подобной деятельности. При создании моделей приходится оперировать с дан-

ными, точность которых различна, а некоторые исходные данные вовсе противоречивы; на разных участках плотность данных с замерами значительно различается. Поэтому для построения цифровых моделей важно иметь набор инструментов статистической, интеллектуальной обработки данных, коррекции исходной информации.

В комплексе ГГМЗ для визуализации моделей и их сопоставления реализованы более десяти вариантов 1D, 2D и 3D графики, в том числе модули получения и оформления карт и вставок на них, графиков на профилях, 3D визуализации возможных разрезов (вертикальные, горизонтальные), простых и сложных отсечений. Основные возможности программных модулей визуализации для нескольких типичных примеров описаны в упомянутых статьях [1], [2].

В настоящей работе обсуждаются возможности системы Mathematica 11 и поясняются примерами приемы использования и настройки нейронной сети при обработке типичных наборов геологических данных. Иллюстрации даются для комплектов данных, которые извлекаются из формируемых, как в [1] или [2] эталонных распределений. К таковым применяются имеющиеся в комплексе ГГМЗ средства и модули «искажения» эталонной модели – пользователь интерактивно формирует эталонную модель, вносит в нее возмущения. В частности, данные на профилях наблюдений искажаются с использованием функций Mathematica: RandomChoice (случайный выбор), RandomVariate (реализация случайной переменной), SeedRandom (инициализация генератора псевдослучайных чисел), RandomSample (случайная выборка).

Для анализа данных, их интеллектуальной обработки с помощью нейронной сети используются и поясняются варианты настройки функций системы Mathematica 11: NetChain (NetEvaluationMode, TargetDevice, BatchSize, DotPlusLayer, ElementwiseLayer), NetModel (EvaluationFunction, EvaluationNet, SourceMetadata, TrainingSetData, TaskType), NetTrain (Method, BatchSize, MaxTrainingRounds, LearningRateMultipliers, LossFunction, Quantity, ValidationSet, TrainingProgressFunction). Специально подобранными примерами поясняются эффекты изменения параметров сети ValidationSet, Scaled, NetEvaluationMode, BatchSize, TakeDrop.

ЗАКЛЮЧЕНИЕ

В докладе будут изложены и проиллюстрированы примерами возможности системы Mathematica версии 11 по настройке и вариантам применения аппарата искусственных нейронных сетей при анализе исходных данных геологического компьютерного моделирования.

III. СПИСОК ЛИТЕРАТУРЫ

1. Taranchuk, V. B. The integrated computer complex of an estimation and adapting of digital geological models / V. B. Taranchuk // *Studia i Materiały*. Volume 14. – 2017. – № 2. – С. 73–86 (in Engl.)
2. Таранчук, В. Б. Интегрированный программный комплекс тестировщика геологических моделей / В. Б. Таранчук // *Научные ведомости Белгородского государственного университета*. Серия: Экономика. Информатика. – 2017. – № 16 (265). – С. 148–159.
3. Wolfram *MATHEMATICA*. Наиболее полная система для современных технических вычислений в мире [Электронный ресурс] / Режим доступа: <http://www.wolfram.com/mathematica>. – Дата доступа: 23.09.2018.
4. Таранчук, В. Б. Введение в язык Wolfram : учеб. материалы для студентов фак. прикладной математики и информатики спец. 1-31 03 04 «Информатика» / В. Б. Таранчук // Минск: БГУ, 2015. – 51 с.
5. Таранчук, В. Б. Основы программирования на языке Wolfram : учеб. материалы для студентов фак. прикладной математики и информатики спец. 1-31 03 04 «Информатика» / В. Б. Таранчук // Минск: БГУ, 2015. – 49 с.
6. Transform data into knowledge [Электронный ресурс] / Режим доступа: <http://www.goldensoftware.com>. – Дата доступа: 23.09.2018.
7. Таранчук, В. Б. Особенности функционального программирования интерактивных графических приложений / В. Б. Таранчук // *Вестник Самарского государственного университета*. Естественнонаучная серия, раздел Математика: – 2015. – № 6 (128). – С. 178–189.
8. Таранчук, В. Б. Возможности и средства Wolfram Mathematica для разработки интеллектуальных обучающих систем / В. Б. Таранчук // *Научные ведомости Белгородского государственного университета*. Серия: Экономика. Информатика: – 2015. – № 1 (198) том 33. – С. 102–110.
9. Taranchuk, V. B. Methods and tools of development of interactive teaching materials on the computer mechanics / V. B. Taranchuk, M. A. Zhuravkov // *Studia i Materiały*. – 2016. – № 11. – С. 69–83 (in Engl.).
10. CDF. Формат вычисляемых документов – Документы оживают благодаря возможностям вычислений. [Электронный ресурс] / Режим доступа: <http://www.wolfram.com/cdf>. – Дата доступа: 23.09.2018.
11. Таранчук, В. Б. О подготовке и распространении на базе системы Mathematica интерактивных графических приложений / В. Б. Таранчук, В. А. Куликович // *Информатизация образования*: – 2015. – № 1 (75). – С. 3–13.
12. Морозов, А. А. Программирование задач численного анализа в системе Mathematica : учеб. пособие / А. А. Морозов, В. Б. Таранчук // Минск : БГПУ – 2015. – 145 с.
13. Таранчук, В. Б. О возможностях и инструментах реализации вероятностно-статистических методов в системе Mathematica / В. Б. Таранчук // *Теория вероятностей, случайные процессы, математическая статистика и приложения : материалы Междунар. науч. конф., посвящ. 80-летию проф., д-ра физ.-мат. наук Г.А. Медведева*, Минск : РИВШ, 2015. – С. 320–325.
14. Таранчук, В. Б. Графический сервис вычислительного эксперимента : учеб.-метод. пособие / В. Б. Таранчук // Минск : БГУ – 2009. – 124 с.