

ПРИМЕРЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА И ВИЗУАЛИЗАЦИИ ГЕОДАННЫХ

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 примеры использования программного комплекса составителя цифровых геологических моделей. Описаны возможности интерактивной графической визуализации геоданных, подготовки представительных наборов данных для возможных интерпретаций экспертов. Приведены алгоритмы получения оценок, которые позволяют подтвердить обоснованность интерпретаций, анализировать разными методами интеллектуальной обработки данные компьютерных геологических моделей.

ВВЕДЕНИЕ

Цифровые геологические, геоэкологические модели в настоящее время являются обязательной составляющей экспертизы во многих сферах деятельности. Основными этапами информационного обеспечения геологических моделей являются загрузка из различных источников и предварительная обработка данных, корреляция, накопление и систематизация информации по характеристикам среды, интерактивный анализ данных, визуализация с помощью графики, картографирование.

Используемые в геологических и геоэкологических моделях данные являются представительной частью геоданных [1], технологическая особенность которых состоит в том, что их не получают на основе непосредственных измерений, а они формируются в результате постобработки измеренной информации. Системная особенность заключается в том, что после формирования геоданные представляют собой интегрированную в единый комплекс совокупность параметров и описаний разных типов и структур, в которых учитываются пространственные отношения, временной и тематический факторы. Информационная особенность обусловлена тем, что геоданные представляют собой новый информационный ресурс, при этом данные группируют по трём характеристикам: месту, времени, теме. Также особенностью является реализуемое автоматически взаимовлияние графических и атрибутивных данных, что обеспечивает основу для пространственного визуального анализа и управления.

Объемы геоданных растут с очень большой скоростью. Соответственно, естественным является применение технологий «больших данных» (конкретика для геоданных в [2]), автоматизированного интеллектуального анализа данных (ИАД). В [3] акцентируется одна из главных целей ИАД – обнаружение в «сырых» (первичных) массивах данных ранее неизвестных, нетривиальных, практически полезных и понятных интерпретации знаний. В формулировке автора [3]

«интеллектуальный анализ данных не исключает человеческое участие в обработке и анализе, но значительно упрощает процесс поиска необходимых данных из сырых данных, делая его доступным для широкого круга аналитиков, не являющихся специалистами в статистике, математике или программировании.»

Различных действующих и применяемых программных средств интеллектуального анализа данных много, например, в [3] выделены 7 классов систем ИАД, перспективы развития можно проследить по [4]. Разнообразие предлагаемых методик и программных средств обуславливает необходимость оценки качества геоданных, определения их основных характеристик. Критерии определения свойств геоданных обсуждаются в [5]; отмечено, что разнообразие геоданных порождает проблему их согласования, эффективной эксплуатации баз данных географических информационных систем и связанных с ними внешних баз данных; значение оценки качества существенно возрастает при переходе от информационных к интеллектуальным технологиям. Ряд вопросов анализа, оценок качества имеющихся и формируемых пространственных данных можно решать с использованием компьютерного комплекса «Генератор геологической модели залежи (ГГМЗ)» [6–15].

ОСНОВНЫЕ ПОЛОЖЕНИЯ

Интегрированный программный комплекс ГГМЗ, разработанный путем интеграции возможностей системы компьютерной алгебры и геоинформационной системы, алгоритмов и «умных» методов адаптации моделей в процессе их эксплуатации, предназначен и позволяет выполнять тестирование и оценки точности геологических моделей, обоснование интерпретаций исходных геоданных. В комплексе ГГМЗ реализованы методические решения и соответствующие программные инструменты, которые позволяют имитировать замеры, подтвердить обоснованность интерпретаций, рассчитать числовые значения погрешностей получаемых разными методами результатов интеллектуальной предоб-

работки данных, включаемых и используемых в компьютерных геологических моделях. Программные компоненты, предложенные и реализованные методические и технические решения, которые включены ГГМЗ, описаны в [6–8]. Основная идея и цель разработки комплекса ГГМЗ состоит в предоставлении эксперту инструментов выбора и обоснования метода обработки пространственных данных путем сопоставления эталонного цифрового поля и восстановленного по «наблюдениям». Эталонное цифровое поле рассматривается, как явно задаваемая аналитическая функция, определяемая в области с двумя независимыми переменными. Примеры подготовки разных эталонных цифровых полей и представительных моделей, варианты графической визуализации результатов даны в [6–10]. Отдельное внимание уделяется алгоритмам и программным инструментам получения результатов сопоставления, оценок точности, способов визуализации погрешностей при разных вариантах формирования цифровых описаний методами аппроксимации по рассеянному множеству данных замеров [10–13]. Несколько разных алгоритмов получения представительных данных, имитации наблюдений путем «искажения» эталонной модели добавлением «шумов», используя разные генераторы случайных чисел, изложены в [11–16]. В [14–16] приведены стандартные, но достаточно эффективные методы анализа данных с помощью искусственных нейронных сетей. Примерами иллюстрируются разные возможности рассматриваемых алгоритмов и программных средств при обработке данных, предназначенных для использования в специфических моделях микробиологии (например, [17]), промышленных объектов типа [18].

ЗАКЛЮЧЕНИЕ

В докладе будут конспективно изложены и проиллюстрированы представительными примерами основные из перечисленных выше методов анализа исходных данных для компьютерных геологических моделей. Основное внимание будет уделено сопоставлению результатов, получаемых с использованием инструментов классической статистической обработки данных и аппарата искусственных нейронных сетей.

СПИСОК ЛИТЕРАТУРЫ

1. Савиных, В. П. Геоданные как системный информационный ресурс / В. П. Савиных, В. Я. Цветков // Вестник Российской академии наук. –2014. –Т. 84, № 9. –С. 826–829.
2. Tsvetkov, V. Ya. Big Data as Information Barrier / V. Ya. Tsvetkov, A. A. Lobanov // European Researcher. –2014. V. 78, № 7-1. –P. 1237–1242.
3. Шайтура, С. В. Интеллектуальный анализ данных геоданных / С. В. Шайтура // Перспективы науки и образования. –2015. № 6 (18). –С. 24–30.
4. Golenkov, V. V. Ontology-based design of intelligent systems / V. V. Golenkov // Open semantic technologies for intelligent systems (OSTIS-2017) : materials of VII Intern. sc.-tech. conf., Minsk, 16–18 Febr. 2017 / Belarus. State Univ. of Informatics a. Radioelectronics. –Minsk, –2017. –P. 37–56.
5. Дышленко, С. Г. Анализ и разработка характеристик качества геоданных / С. Г. Дышленко // Перспективы науки и образования. –2016. № 2 (20). –С. 23–27.
6. Таранчук, В. Б. Программный комплекс адаптации геологических моделей. Концепция, решения, примеры реализации / В. Б. Таранчук // Проблемы физики, математики и техники. –2017. –№ 3 (32). –С. 81–90.
7. 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.
8. 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.
9. Taranchuk, V. The integrated program complex of the composer of geological models. The concept, solutions / V. Taranchuk // Computer Algebra Systems in Teaching and Research. –2017. –Volume VI. –P. 186–203.
10. Таранчук, В. Б. Методы и инструментарий оценки точности компьютерных геологических моделей / В. Б. Таранчук // Вестник БрГТУ, № 2, –2018. –С. 64–70.
11. Таранчук, В. Б. Средства и примеры интеллектуальной обработки данных для геологических моделей / В. Б. Таранчук // Проблемы физики, математики и техники. –2019. –№ 3 (40). –С. 117–122.
12. Таранчук, В. Б. Методы и примеры интеллектуальной обработки данных для геологических моделей / В. Б. Таранчук // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. –2019. –№ 3 (46). –С. 511–522. DOI 10.18413/2411-3808-2019-46-3-511-522.
13. Taranchuk, V. Methods and Tools of Intelligent Data Processing for Geological Models / V. Taranchuk // Computer Algebra Systems in Teaching and Research. –2019. –Volume VIII. –P. 147–158.
14. Taranchuk, V. New computer technologies, analysis and interpretation of geodata / V. Taranchuk // MATEC Web of Conferences IPICSE-2018. –V. 251, 04059. –P. 1–8.
15. Taranchuk, V. B. Examples of the use of artificial neural networks in the analysis of geodata / V. B. Taranchuk // Open Semantic Technologies for Intelligent Systems : Research Papers Collection. –2019. –Issue 3. –P. 225–230.
16. Таранчук, В. Б. Интеллектуальные вычисления, анализ, визуализация больших данных / В. Б. Таранчук // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов V Междунар. науч.-практ. конф. В 2 ч. Ч. 1. –Минск : БГУИР, –2019. –С. 337–346.
17. Chichurin, A. Computer simulation of two chemostat models for one nutrient resource / A. Chichurin, H. Shvychkina // Mathematical Biosciences, 278, 2016, –P. 30–36.
18. Orlov, V. Mathematical modeling of emergency situations at objects of production and gas transportation / V. Orlov, E. Detina, O. Kovalchuk // MATEC Web of Conferences IPICSE-2018. –V. 251, 04059.