

ИСПОЛЬЗОВАНИЕ АЛГОРИТМА «ДЕРЕВА РЕШЕНИЙ» В СИСТЕМАТИЗАЦИИ ОПРЕДЕЛЯЮЩИХ УРАВНЕНИЙ ДЛЯ ГРУНТОВ

Л.А. Строкова

Томский политехнический университет

E-mail: geyer@tpu.ru

Рассматриваются понятие «определяющие уравнения» и предложения по систематизации определяющих уравнений для грунтов. Как инструмент поддержки принятия решения по выбору определяющих уравнений для грунтов в геотехнических расчетах разработана их классификация при помощи алгоритма дерева решений. Приводятся значимые признаки определяющих уравнений для разных классов грунтов. Обосновывается необходимость составления базы данных по определяющим уравнениям для грунтов и указаны атрибуты такой базы данных.

Ключевые слова:

Численное моделирование, определяющие уравнения, дерево решений.

В последние десятилетия сделаны значительные успехи в области математического моделирования поведения горных пород и разработки компьютерных программ для этих целей. Тем не менее, при использовании компьютерных программ в геотехнических расчетах встречаются трудности, такие как, определение исходных параметров, граничных и начальных условий, выбор математической модели поведения грунта, интерпретация полученных результатов. Данная работа призвана найти практический инструмент для выбора модели поведения грунта при разработке расчетных схем оснований.

Математическая модель механических свойств среды задается уравнением, связывающим напряжения, имеющиеся в окрестности некоторой точки среды, и деформации, возникающие вследствие этого. Это уравнение называется *определяющим уравнением* или *уравнением состояния*. Знакомство с литературными источниками показало, что в настоящее время интерес к разработке определяющих уравнений, отражающих сложное поведение грунта, достаточно велик. Этому способствовало три фактора: потребности современной науки, техники и экономики; наличие уже хорошо разработанных элементов теории и численных методов; возможности современной вычислительной техники.

В отечественной монографической и учебной литературе теория определяющих уравнений представлена отдельными работами [1–3].

За рубежом положения теории определяющих уравнений геоматериалов представлены журнальными статьями и докладами на конференциях, которые проходят практически ежегодно, начиная с 1984 г. О популярности определяющих уравнений для грунтов говорит и тот факт, что результат поиска термина *constitutive relations (laws, equations, models) for soils, rocks, geomaterials* в поисковой системе Google (на июнь 2008 г.) предлагает к ознакомлению более 11000 электронных документов, вероятно, подобное число содержится в журналах и трудах конференций за последние полвека. Более 350 документов в виде монографий и статей в ведущих

мировых журналах по моделям грунтов хранится на образовательном сайте [4]. В некоторых американских университетах на постоянной основе проводятся семинары, например, под руководством профессоров P.V. Lade (Washington), V.N. Kaliakin (Delaware).

Большое разнообразие моделей поведения грунтов вызвало необходимость их систематизации на основе классификации. Попытки классификации предпринимались давно. Одна из первых классификаций представлена в работе G. Gudehus [5]. В этой классификации произведена разбивка объема понятия «определяющие уравнения» на различные разряды по видоизменению признака: выделены три класса – упругость, пластичность, вязкость по наличию и характеру деформаций при изменении напряжений с указанием характерных особенностей каждого класса.

Darve, Lambert [6] указывают на необходимость классификации определяющих уравнений для геоматериалов и предлагают классификацию определяющих уравнений невязких материалов по количеству «тензорных зон». Они понимают под «тензорной зоной» – домен инкрементного пространства нагружения, в котором отношение между приращением деформаций $d\varepsilon$ и приращением напряжений $d\sigma$ есть линейная функция. Предложено различать 5 классов определяющих уравнений: с одной тензорной зоной (закон Гука); двумя тензорными зонами – зоной нагружения и зоной разгрузки: модели Guelin (1980), Mroz (1967), Prevost (1978), Dafalias, Hermann (1980); четырьмя тензорными зонами: модели Davis, Mullenger (1978), Lade (1977), Loret (1981), Nova, Wood (1979), Vermeer (1978); восемью тензорными зонами: модели Darve, Labanieh (1982), Gudehus (1979) и с бесконечным числом тензорных зон, среди которых различают три различных типа: а) эндохронические (endochronoc) модели Valanis (1971), Basant (1978); б) гипопластические модели Kolymbas (1977), Chambon (1994); в) пластической теории ограниченной поверхности (Bounding surface theory of plasticity) – модель Dafalias (1986).

Анализируя особенности различных моделей грунтов, типы, компоненты определяющих уравнений P.V. Lade [7] предложил классифицировать определяющие уравнения в механике грунтов на 3 категории, по совершенству описания различных эффектов поведения грунта при нагрузках, возможности определения параметров традиционными экспериментами.

В работе T. Schanz [8] даны характеристики основных определяющих уравнений, даются рекомендации по их использованию для разных видов расчетов.

Существующие обзоры нацелены на теоретическое осмысление многообразия моделей, но им присущ существенный недостаток — отсутствие инструментария для поддержки решения инженера — расчетчика о выборе определяющего уравнения. Отсутствие инструментария обусловлено особенностями моделируемой среды, как грунтового массива, так и возникающих предельных состояний, необходимостью учета множества факторов и целей, имеющих сложную структуру взаимосвязи.

Для решения задач подобного рода в последние годы широко используются методы интеллектуального анализа данных (Data Mining). Существует множество определений Data Mining, но в целом они совпадают в выделении 4-х основных признаков. Вот определение, которое дал Г.И. Пятецкий-Шапино [9], один из ведущих мировых экспертов в области Data Mining: *«Data Mining (интеллектуальный анализ данных) — это процесс обнаружения в «сырых» данных ранее неизвестных, нетривиальных, практически полезных и доступных для интерпретации знаний (закономерностей), необходимых для принятия решений в различных сферах человеческой деятельности»*. В основу современной технологии Data Mining положена концепция шаблонов (паттернов), отражающих фрагменты многоаспектных взаимоотношений в данных. Эти шаблоны представляют собой закономерности, свойственные подвыборкам данных, которые могут быть компактно выражены в понятной человеку форме. Поиск шаблонов производится методами, не ограниченными рамками априорных предположений о структуре выборке и виде распределений значений анализируемых показателей. Иными словами, Data Mining — это анализ данных с целью отыскания в них типовых образцов или стереотипных изменений, скрытых от нас по причине невозможности осмысления и анализа такого количества данных и взаимосвязей между ними [9].

Одной из особенностей процесса разработки математической модели является отсутствие в настоящий момент «тренировочного» набора данных — данных, для которых был бы априори известен тип определяющих уравнений. Это существенно ограничивает круг алгоритмов, которые можно было использовать при классификации: например, классификаторы Байеса, нейронные сети и т. п. методы, требующие предварительного обучения на

тренировочном наборе данных. В нашем распоряжении остались лишь деревья решений и алгоритмы иерархической кластеризации.

Дерево решений — это способ представления иерархической структуры классифицирующих правил типа «ЕСЛИ... ТО...» (*if — then*), имеющей вид дерева. Для принятия решения, к какому классу отнести некоторый объект или ситуацию, требуется ответить на вопросы, стоящие в узлах этого дерева, начиная с его корня. Вопросы имеют вид «значение параметра *A* больше *x*?». Если ответ положительный, осуществляется переход к правому узлу следующего уровня, если отрицательный — то к левому узлу; затем снова следует вопрос, связанный с соответствующим узлом.

Данный метод является одним из наиболее популярных для решения задач классификации благодаря наглядности, понятной интерпретации правил, отсутствия требования определить входные атрибуты, возможности работать с категориальными типами данных, способности решать задачи, в которых отсутствует априорная информация о виде зависимости между исследуемыми данными [10].

Разведочный анализ данных был осуществлен при помощи академической версии аналитической платформы «Deductor» компании BaseGroup Labs, специализирующей на разработке систем для глубокого анализа данных, охватывающих вопросы сбора, консолидации, очистки данных, построения моделей и визуализации [11]. В качестве основы для разведочного анализа были взяты данные P.V. Lade [7] для 31 модели грунта. Каждая модель охарактеризована 18 атрибутами, отражающими различные аспекты поведения пород при разных типах нагружения, способы оценки параметров грунтов.

Анализ данных позволил отобрать следующие значимые для классификации признаки: класс грунта; изменение знака нагружения; реальное поровое давление; реальное 3D поведение; дренированное разупрочнение; циклическое нагружение.

Вначале разведочного анализа для устранения незначимых факторов был проведен корреляционный анализ данных. В табл. 1 указаны важные атрибуты и их значимость для классификации моделей.

Таблица 1. Данные корреляционного анализа исходных данных

Номер атрибута	Атрибут	Значимость, %
2	Изменение знака напряжений	31,53
4	Реальное поровое давление	23,19
5	Реальное 3-мерное поведение	19,88
1	Дренированное разупрочнение	18,63
6	Циклическое нагружение	6,77
3	Затруднения при лабораторных испытаниях	0,0

Поскольку дальнейшая классификация по полному множеству признаков и моделей давала низкие результаты классификации (поддержка и достоверность решения составляла менее 20 %) было принято решение отфильтровать имеющиеся моде-

ли по классу грунтов и определить значимость тех или иных атрибутов моделей внутри класса грунтов.

Фильтр для несвязных дисперсных грунтов позволил классифицировать только 4 модели: 1 – закон Гука, 5 – закон дилатансии Rowe (1962), 18 – модель Darve (1982) и 19 – гипопластические модели Kolymbus (1991), Gudehus (1996), Bauer (1996) из 15. Дерево решений для песчаных пород приведено на рис. 1.

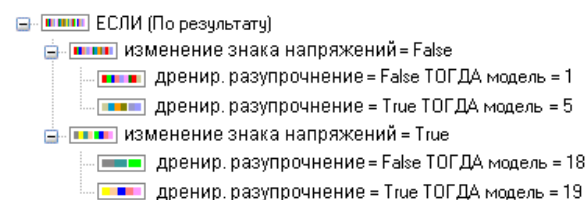


Рис. 1. Дерево решений для песчаных грунтов

Было отобрано 4 решающих правила и 2 значимых фактора с высокой значимостью порядка 50 % (табл. 2, 3) – это «изменение знака нагружения» и «дренированное разупрочнение».

Таблица 2. Решающие правила для несвязных грунтов

№	Условие	Следствие (модель)	Поддержка		Достоверность	
			%	Кол-во	%	Кол-во
1	Изменение знака напряжений = False И дренированное разупрочнение = False	1	39,13	9	11,11	1
2	Изменение знака напряжений = True И дренированное разупрочнение = False	5	26,09	6	16,67	1
3	Изменение знака напряжений = False И дренированное разупрочнение = True	18	13,04	3	33,33	1
4	Изменение знака напряжений = True И дренированное разупрочнение = True	19	21,74	5	20,00	1

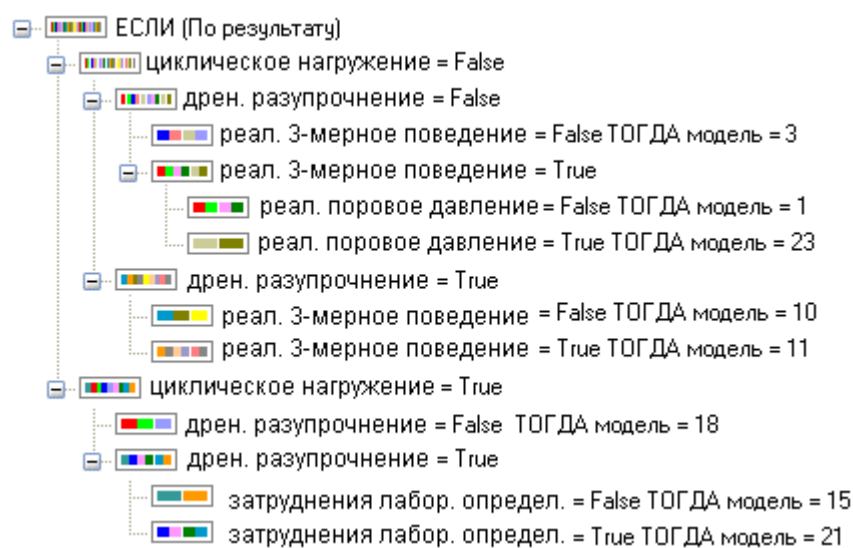


Рис. 2. Дерево решений для связных грунтов

Таблица 3. Значимость атрибутов моделей для несвязных грунтов

Номер атрибута	Атрибут	Значимость, %
1	Дренированное разупрочнение	50,87
2	Изменение знака напряжений	49,13
3	Циклическое нагружение	0,0

Фильтр для связных дисперсных грунтов позволил классифицировать только 8 моделей из 15 (рис. 2, табл. 4, 5); это следующие: 3 – модель Drucker-Prager (1952); 1 – закон Гука; 23 – модель Plaxis Hardening, предложенную Brinkgreve и Vermeer (1997); 10 – Modified Cam Clay, авторы Roscoe и Burland (1968); 11 – модель Elasto-Viscoplastic, предложенную Adachi, Oka (1982); 18 – модель Darve (1982); 15 – модель Nor-Sand, предложенную Jefferies (1993); 21 – модель Fuzzy Set Plasticity, предложенную Klisinski и др. (1987, 1988).

Таблица 4. Значимость атрибутов моделей для связных грунтов

Номер атрибута	Атрибут	Значимость, %
1	Дренированное разупрочнение	33,38
3	Циклическое нагружение	31,09
5	Реальное 3-мерное поведение	22,03
4	Реальное поровое давление	6,75
6	Затруднения при лабораторных испытаниях	6,75
2	Изменение знака напряжений	0,0

Фильтр для скальных и полускальных грунтов позволил классифицировать еще 5 моделей из 16 моделей (рис. 3, табл. 6, 7). А именно: 3 – модель Drucker-Prager (1952); 1 – закон Гука; 23 – модель Plaxis Hardening, предложенную Brinkgreve и Vermeer (1997); 18 – модель Darve (1982); 21 – модель Fuzzy Set Plasticity, предложенную Klisinski и др. (1987, 1988).

То обстоятельство, что не все охарактеризованные P.V. Lade модели были отобраны этим алгоритмом, объясняется специфичностью и уникально-

стью условий, при которых конкретная модель была сформулирована. И наоборот, не случайно, алгоритм «дерево решений» наиболее часто отбирал закон Гука, упруго-пластические модели с изотропным упрочнением Plaxis Hardening и Cam Clay. И в практике эти модели широко используются из-за их универсальности, достоверного отображения условий работы грунта в основании.

Таблица 5. Решающие правила для связных грунтов

№	Условие	Следствие (модель)	Поддержка		Достоверность	
			%	Кол-во	%	Кол-во
1	Циклическое нагружение = False И дренированное разупрочнение = False И реальное 3-мерное поведение = False	3	14,29	4	25,00	1
2	Циклическое нагружение = False И дренированное разупрочнение = False И 3-мерное поведение = True И поров. давл. = False	1	14,29	4	25,00	1
3	Циклическое нагружение = False И дренированное разупрочнение = False И 3-мерное поведение = True И поров. давл. = True	23	7,14	2	50,00	1
4	Циклическое нагружение = False И дренированное разупрочнение = True И реал. 3-мерное поведение = False	10	10,71	3	33,33	1
5	Циклическое нагружение = False И дренированное разупрочнение = True	11	21,43	6	16,67	1
6	Циклическое нагружение = True И дренированное разупрочнение = False	18	10,71	3	33,33	1
7	Циклическое нагружение = True И дренированное разупрочнение = True И затруднения при лабораторных испытаниях = False	15	7,14	2	50,00	1
8	Циклическое нагружение = True И дренированное разупрочнение = True И затруднения при лабораторных испытаниях = True	21	14,29	1	25,00	1

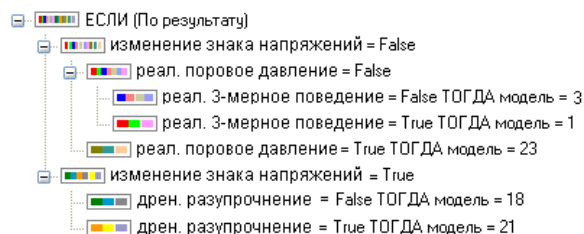


Рис. 3. Дерево решений для скальных и полускальных грунтов

Предложенная классификация — первая попытка сформулировать практический инструмент для выбора модели поведения грунта из некоторого нечеткого множества моделей. Слабым местом пред-

ложенной классификации является малый объем доступных для анализа данных о моделях. Разработка инструментария поддержки решения средствами Data Mining при выборе модели поведения грунта позволяет достичь нового качества проектирования сооружений и получить экономический эффект в результате: 1) повышения точности решения, по сравнению с ранее применяемыми способами на основе экспертных суждений; 2) снижение субъективности решения; 3) снижение объема необходимых для принятия решения данных по сравнению с объемом, используемым экспертом; 4) ускорение процесса принятия решения; 5) возможность решения ранее нерешаемых задач при появлении новых моделей поведения грунта.

Таблица 6. Значимость атрибутов моделей для скальных и полускальных грунтов

Номер атрибута	Атрибут	Значимость, %
2	Изменение знака напряжений	41,29
4	Реальное поровое давление	23,83
5	Реальное 3-мерное поведение	18,65
1	Дренированное разупрочнение	16,23
6	Затруднения при лабораторных испытаниях	0,0
3	Циклическое нагружение	0,0

Таблица 7. Решающие правила для скальных и полускальных грунтов

№	Условие	Следствие (модель)	Поддержка		Достоверность	
			%	Кол-во	%	Кол-во
1	Изменение знака напряжений = False И поровое давление = False И реал. 3-мерное поведение = False	3	25,00	4	25,00	1
2	Изменение знака напряжений = False И поровое давление = False И реал. 3-мерное поведение = True	1	18,75	3	33,33	1
3	Изменение знака напряжений = False И поровое давление = True	23	18,75	3	33,33	1
4	Изменение знака напряжений = True И дренированное разупрочнение = False	18	18,75	3	33,33	1
5	Изменение знака напряжений = True И дренированное разупрочнение = True	21	18,75	3	33,33	1

Учитывая темпы применения численных методов в решении геотехнических задач, структурированность информации о моделях грунта становятся не просто желательным, а необходимым безальтернативным условием. Реализация современных требований сокращения сроков и стоимости проектирования, повторного использования накопленной

информации при проектировании новых зданий и сооружений, обеспечения необходимой информационной поддержки проекта на протяжении всего его жизненного цикла уже невозможна без создания специальной базы данных об имеющихся моделях грунтов. В базе данных необходимо представить экспертные описания каждой физико-математической модели определенного класса, с характеристикой достоинств и недостатков каждой модели, основных физических допущений и ограни-

чений, сделанных авторами при разработке модели, с указанием необходимых уравнений и входящих в них параметров, с указанием примеров реализации модели, ее адекватности моделируемому процессу и области применения в расчетах.

Выполнение разведочного анализа данных показало принципиальную возможность использования предложенного подхода при выборе определяющих уравнений для решения различных геотехнических задач в зависимости от типа грунтов.

СПИСОК ЛИТЕРАТУРЫ

1. Ахмеров Р.Р., Хакимзянов Г.С., Чубаров Л.Б. Математическое моделирование [Электронный ресурс]. — режим доступа: <http://www.ict.nsk.su/lab2.3/ru/stuff/listMM.htm>. — 25.04.2008.
2. Болдырев Г.Г. Устойчивость и деформируемость оснований анкерных фундаментов. — М.: Стройиздат, 1987. — 80 с. [Электронный ресурс]. — режим доступа: <http://geotek.ru/publications/public1/>. — 25.04.2008.
3. Ключников В.Д. Математическая теория пластичности. — М.: Изд-во МГУ, 1979. — 208 с. [Электронный ресурс]. — режим доступа: <http://sci-lib.com/book000486.html>. — 25.04.2008.
4. Информационные материалы образовательного портала [Электронный ресурс]. — режим доступа: www.geotechlinks.com. — 25.04.2008.
5. Gudehus G. Bodenmechanik. — Stuttgart: Enke, 1981. — 268 S.
6. Darve F., Labanieh S. Incremental constitutive law for sands and clays, simulation of monotonic and cyclic test // Int. J. Num. Anal. Meth. Geomech. — 1982. — № 6. — P. 243–273.
7. Lade P.V. Overview and evaluation of constitutive models // Soil Constitutive Models: Evaluation, Selection, and Calibration. Ed. J.A. Yamamuro, V.N. Kaliakin. — American Society of Civil Engineers, 2005. — V. 128. — P. 69–98.
8. Schanz T. Aktuelle Entwicklungen bei Standsicherheits- und Verformungs-berechnungen in der Geotechnik // Geotechnik. — 2006. — № 1 (29). — S. 13–28.
9. Дюк В.А., Самойленко А.П. Data Mining: учебный курс. — СПб.: Питер, 2001. — 368 с.
10. Чубукова И.А. Data Mining. Учебный курс Интернет-Университета Информационных Технологий [Электронный ресурс]. — режим доступа: <http://www.intuit.ru/departments/database/damining/9/2.html>. — 25.04.2008.
11. Учебные материалы об аналитической платформе «Deductor» компании BaseGroup Labs [Электронный ресурс]. — режим доступа: www.basegroup.ru. — 25.04.2008.

Поступила 17.11.2008 г.

УДК 681.518:622.276

МЕТОД И АЛГОРИТМ ОЦЕНКИ КОЭФФИЦИЕНТОВ ОХВАТА ВЫТЕСНЕНИЕМ И ЗАВОДНЕНИЕМ

А.А. Захарова

Томский политехнический университет
E-mail: alen@cc.tpu.edu.ru

Представлен метод и реализующий его алгоритм оценки коэффициентов охвата вытеснением и заводнением для выбора оптимальных вариантов разработки нефтегазовых месторождений, реализуемых в рамках проектных технологических документов. Апробация осуществлялась на примере ряда месторождений при выполнении проектов пробной эксплуатации и технологических схем разработки.

Ключевые слова:

Геологическая модель, гидродинамическая модель, нефтегазовое месторождение, методы, алгоритмы, программное обеспечение, оптимизация.

Моделирование месторождений нефти и газа сложный трудоемкий процесс, поэтому создание методов для экспресс-оценки геолого-технологических параметров с целью оптимизации указанного процесса весьма актуально. При оценке эффективности разработки нефтегазовых месторождений принято оперировать таким показателем, как коэффициент охвата $k_{\text{охв}}$, который характеризует охват залежи при разработке вытеснением (отбор из добывающих

скважин) и заводнением (вытеснение флюида от нагнетательных скважин). $k_{\text{охв}}$ принято определять как $k_{\text{охв}} = \text{КИН} / k_{\text{выт}}$, где КИН — коэффициент извлечения нефти, а $k_{\text{выт}}$ определяется по результатам капиллярметрии на образцах керн. Однако такой расчет можно считать грубым и при этом не учитываются геометрическая сложность контура водонефтяного контакта (ВНК) и истинные площадные характеристики зон, не охваченных разработкой. Поэтому