

СОЗДАНИЕ ИНФОРМАЦИОННЫХ ЭЛЕКТРОННЫХ ГИДРОГЕОЛОГИЧЕСКИХ КАРТ С ПРИМЕНЕНИЕМ СОВРЕМЕННЫХ ГИС-ТЕХНОЛОГИЙ, МОНИТОРИНГ И АНАЛИЗ ГЕОДАНЫХ

*Черкашин В.И., Мамаев С.А., Ибаев Д. Г., Маммаев А.О.
Институт геологии ДНЦ РАН*

Аннотация

В работе рассматриваются вопросы, связанные с проектированием информационных систем при создании базы данных гидрогеологических параметров артезианских скважин, наблюдений и их анализа, часто называемых системами мониторинга. Данные в таких системах часто слабо или частично структурированы. Их структура может зависеть от конкретного наблюдения и меняться со временем. В общих чертах строится модель данных, ориентированная на задачи мониторинга.

Введение

Проблема питьевого водоснабжения районов Северного Дагестана повлекла за собой принятие мер по выявлению и исследованию многих скважин на наличие канцерогенных элементов. Возникает задача создания эффективного инструмента, дающего возможность не только накапливать информацию об изменениях гидрогеологических параметров артезианских скважин, но и использовать ее для принятия управленческих решений. Традиционные реляционные СУБД не обеспечивают их решения, поскольку отсутствуют средства обработки пространственной информации. Для работы с такого рода информацией используются геоинформационные системы (ГИС), обеспечивающие сбор, обработку, хранение, анализ, отображение пространственно распределенных геоданных и мониторинг, позволяющие осуществлять их интегрирование по территориям с целью эффективного использования и решения прикладных задач, связанных с анализом, прогнозированием, экспертизой и управлением окружающей средой, качеством подземных вод. Методы математического моделирования наиболее эффективны и экономичны при подсчете запасов подземных вод, проектировании водозаборов, оценке экологической обстановки прилегающих территорий и т.д. Решение задач моделирования с помощью компьютера начинают с того, что составляют несколько расчетных гидродинамических схем. Изучают каждую из них с точки зрения экономичности и в итоге выбирают самую оптимальную. Рассчитывают оптимальную длину, диаметр и число скважин, которые будут подавать наибольшее количество воды при наименьших затратах. Расчеты покажут, насколько снизится напор подземных вод в будущем, на какое расстояние распространяется влияние водозабора и как это отразится на работе других скважин, колодцев и экологии окружающей среды. Определенную роль при математическом моделировании играют и методы Монте-Карло. Они заключаются в моделировании случайного процесса, наблюдения и нахождении подхода к решению задачи об оптимальном расположении скважин, их числа, расстояния между ними, сводя эту задачу к принятию решения при многокритериальных условиях (влияние на окружающую среду, стоимость бурения, потребность в воде, и т.д.).

Доминирующей моделью данных в настоящее время является реляционная модель. Тем не менее, было бы рискованным предполагать, что конечный пользователь смотрит на мир с позиций реляционной модели. Для каждого рода деятельности существует свой взгляд на то, как устроена их предметная область. Цель разработки новых моделей данных состоит в том, чтобы приблизить информационную систему, имеющую дело с некоторой предметной областью, к представлениям пользователей об этой предметной области. Более того, если проанализировать современные системы, то окажется, что реляционная модель данных используется в них, главным образом, для обеспечения хранения данных, а большая часть логики, связанной с предметной областью, остается вне компетенции системы управления базами данных.

Компьютеризованные базы данных превратились в широко распространенные рабочие инструменты. Однако базы данных на плоских таблицах не являются эффективным способом хранения большого объема информации. Поэтому, мы применили идею реляционной базы данных для гидрогеологических параметров артезианских скважин в Северо-Дагестанском артезианском бассейне. В реляционной базе данных используется несколько таблиц и между ними устанавливается связь, что позволяет ввести информацию в одну таблицу и связать ее с записью в другой через идентификатор.

Модель данных, т.е. набор средств описания концептуальной схемы базы данных (БД), используемой СУБД в той или иной степени интегрируется в информационную систему, построенную на ее основе. Действительно, все основные компоненты модели данных – структуры данных, операции и способы описания ограничений целостности, – должны учитываться при проектировании системы с самого начала, а декомпозиция предметной области на объекты, сущности, отношения, атрибуты, факты и т.д. практически полностью определяются моделью данных, которую привык использовать системный аналитик. Использование понятий и операций, отражающих природу предметной области, значительно снижает затраты на проектирование системы и часто позволяет повысить ее эффективность за счет того, что удастся избежать ошибок проектирования, связанных с неадекватным представлением данных. Таким образом, информационная система в целом наследует многие черты модели данных СУБД, главным образом касающиеся подхода к представлению данных.

Построение информационной системы для мониторинга простой предметной области (однородные объекты наблюдения, небольшое фиксированное количество независимых параметров) не составляет особого труда – для этого достаточно использовать какую-либо стандартную СУБД. Дело осложняется, если требуется фиксировать результаты наблюдений над большим количеством объектов разного типа, которые проводятся множеством независимых пользователей, и сохранять данные для последующего анализа.

Данные, регистрируемые разными пользователями для одного и того же объекта, могут противоречить и вообще никак не соотноситься друг с другом. В этом случае использование СУБД, поддерживающих традиционные модели данных, не намного облегчает задачу, поскольку большинство из них ориентированы на фиксацию единой непротиворечивой картины предметной области. Особенности подходов к моделированию данных и классификация моделей данных представлены в [1, 2, 3].

Другой особенностью круга задач, рассматриваемых в настоящей работе, является потребность в накоплении сведений о наблюдаемых объектах для их анализа. Современные информационные системы, помимо хранения информации о предметной области, предоставляют пользователям широкий набор средств анализа данных. В последнее десятилетие в теории СУБД появилось целое направление – хранилища данных. Особенностью данной технологии является то, что помимо обычной базы данных, которая рассматривается как моментальный снимок предметной области, существует специальное хранилище, куда время от времени сбрасываются данные. Для того, чтобы сделать возможным анализ данных, в соответствии с этой технологией требуется по существу заново спроектировать схему хранилища и постоянно загружать туда информацию из базы данных корпоративной системы.

1. Основные понятия модели

Наблюдение (Измерение). В реальном мире мы можем измерять или наблюдать какие-либо характеристики объектов. Для каждого наблюдения определено множество объектов, над которыми оно производится. В качестве значений параметров в контексте какого-либо наблюдения выступают идентификаторы объектов. В этом смысле наблюдение – это просто отношение между объектами. При этом любая связь "Параметр–Объект–Значение" имеет смысл только в контексте какого-либо наблюдения над объектом.

Для того, чтобы можно было обращаться к результатам наблюдения, это отношение также получает при создании идентификатор, как и любой другой объект. В некотором смысле понятие объекта – это вырожденное наблюдение, т.е. наблюдается различимость объектов. С другой стороны, наблюдение само может выступать в качестве объекта. Например, можно оценивать корректность того или иного наблюдения, делать выводы из значений параметров и т.п., т.е. приписывать наблюдениям какие-либо свойства. Таким образом, наблюдение позволяет присвоить одному или нескольким параметрам некоторого объекта какие-либо значения (рис.1). На рисунке наблюдение представлено прямоугольником, объекты – черными кружками, связь между наблюдением и объектом, над которым оно производится – жирной стрелкой, связь между параметрами и их значениями – пунктиром.

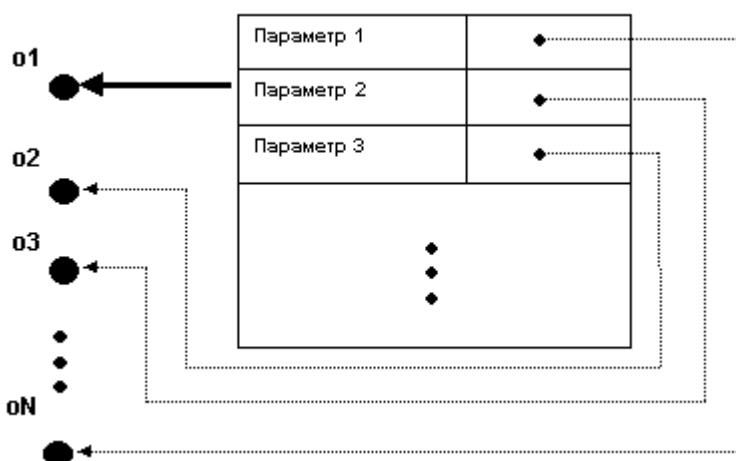


Рис. 1. Связь между параметрами наблюдений и их значениями

Состояние. Объект в каждый момент времени характеризуется своим состоянием, которое представляет собой совокупность значений всех его параметров, зафиксированных при последних наблюдениях этого объекта, предшествующих данному моменту. Будем называть такое состояние актуальным на этот момент, или просто актуальным, если из контекста ясно о каком моменте времени идет речь. Актуальное состояние в текущий момент времени будем называть текущим состоянием. Для любых двух состояний объекта можно всегда определить такое, которое предшествует обоим этим состояниям (или совпадает с одним из них). Еще одно выделенное – терминальное состояние T – это состояние, в котором все параметры объекта имеют значение T. Если все параметры объекта имеют значение T, то это эквивалентно его уничтожению. Это не означает, что с таким объектом больше ничего нельзя сделать. Объект может быть идентифицирован по какому-либо предшествующему состоянию и параметры объекта могут быть изменены. На рис. 2 кружки обозначают состояния, стрелка – переход из одного состояния в другое.

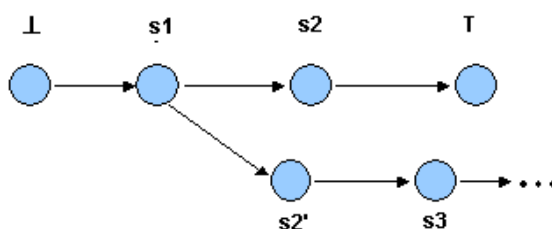


Рис.2. Схема перехода из одного состояния в другое

Наблюдение может быть либо новым, первичным, то есть оно производится над новым, вновь созданным объектом, и тогда объект идентифицируется в силу того, что он создан. Либо наблюдение может быть повторным, то есть проводится над объектами, над которыми ранее уже проводились какие-то наблюдения, и тогда объект или объекты выбираются на основе ранее зафиксированных параметров или через явное обозначение объекта. При каждом первичном наблюдении возникает не менее двух объектов: тот, над которым производится наблюдение, и само это наблюдение (его представитель), которое можно рассматривать как объект.

Источником информации для наблюдения могут быть вводимые пользователем данные, либо данные, автоматически поступающие в вычислительную систему по каким-либо каналам. При каждом наблюдении состояние соответствующего объекта изменяется. Исходное состояние определяется указанием наблюдения или нескольких наблюдений, которые соответствуют этому состоянию. Если в качестве исходного указывается не последнее наблюдение (т.е. не текущее состояние), то возникает "расщепление" состояний.

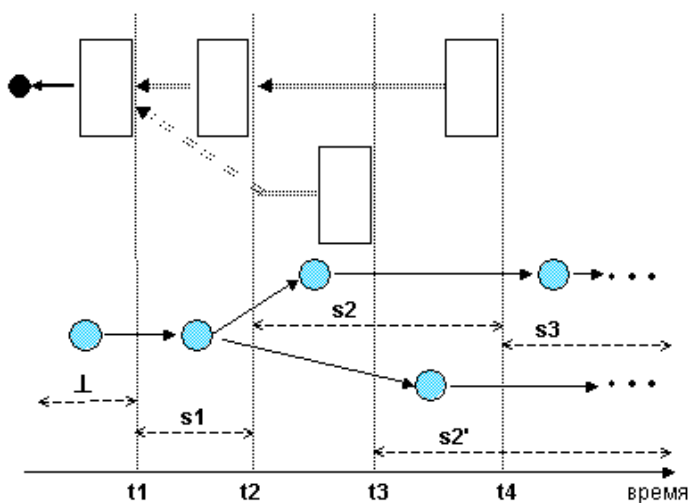


Рис. 3. Соотношение между наблюдениями и состояниями

На рис. 3 показано соотношение между наблюдениями и состояниями (s_i) для одного объекта. Прямоугольники обозначают наблюдения, двойная линия со стрелкой указывает на наблюдения, определяющие исходные для наблюдений состояния. Начиная с момента t_3 , наряду с состоянием s_2 возникает s_2' и состояние объекта расщепляется. Некоторые параметры объектов могут не измеряться непосредственно, а выводиться из ранее наблюденных значений. Для таких параметров в типе наблюдения определяются формулы, которые вычисляются при создании нового или повторного наблюдения, после того, как получили значения все параметры, участвующие в формуле.

Статические параметры вычисляются при проведении наблюдения один раз с использованием некоторого запроса. В дальнейшем их значение не изменяется. Подобные запросы должны поставлять значение и их исполнение не должно вызывать побочных эффектов. Это означает, что процедура вычисления статического параметра не может инициировать проведения других наблюдений, либо требовать активности со стороны пользователя. Статический параметр не может также зависеть от динамических параметров. Если параметр определен как статический, то в системе сохраняется его значение, а не формула, по которой он вычислялся.

Динамический параметр вычисляется каждый раз при обращении к этому параметру. Значением динамического параметра является запрос. В частности, это может быть просто ссылка на текущее значение некоторого параметра другого объекта.

Таким образом, при описании параметра определяются следующие его свойства:

- Тип значения
- Инициализация
- Процедура, определяющая действия при присвоении параметру какого-либо значения, исполняющаяся после проверки соответствия типа
- Стандартный формат или другая процедура получения изображения
- Обязательный / необязательный
- Права доступа
- Период, в течение которого данные действительны.

2. Практические задачи

Основным свойством геологической информации является пространственная и временная приуроченность данных. В соответствии с множеством изучаемых процессов проводится комплекс периодических исследований геологических процессов и явлений на постоянных полигонах и в отдельно взятых точках некоторого региона. Разные процессы имеют разную скорость развития, в связи с чем некоторые из них рассматриваются как стационарный фон, а другие нуждаются в исследованиях разной (от столетий до долей суток) периодичности [4, 5]. Результаты исследований накапливаются в единой информационной системе для последующего анализа.

К настоящему времени создано большое количество программных систем для решения таких задач. Но каждая из них имеет свои преимущества и недостатки. Поэтому для получения необходимого результата можно попытаться использовать эти системы в комплексе. В частности, некоторые из них можно применять для более надежного хранения информации, другие – для статистической обработки информации, третьи – для наглядной визуализации. Среди многочисленных средств обработки информации можно перечислить более практичные системы – Microsoft Access, ArcView и 3D Analyst и др. Первая из них используется для хранения и статистической обработки информации, вторая – для наглядного отображения данных по артезианским скважинам на географических картах, третья – для получения трехмерных графических разрезов, которые дают более наглядное представление о внутренней структуре земли.

Используя весь этот комплекс программных систем, разрабатывается постоянно действующая гидрогеолого-математическая модель формирования и управления ресурсами и качеством подземных вод крупнейшего на юге России Теркско-Кумского артезианского бассейна, и, в первую очередь, его Дагестанской части, где функционирует свыше трех тысяч эксплуатационных скважин для водоснабжения населения.

Эта модель разрабатывается на базе программного обеспечения ArcGis и MS Access и, наряду со стандартными формами работы с пространственными данными, содержит дополнительные инструменты для сбора, анализа, отображения и хранения результатов лабораторных и полевых исследований.

Модель реализуется как электронная карта, отвечающая требованиям ArcGIS в части организации данных, методов их ввода, редактирования, анализа и отображения. Топографической основой модели является топографическая цифровая карта Дагестана масштаба 1:200000. Данные организованы в объектно-ориентированную базу геоданных в формате Dbase. Доступ к данным, а также просмотр и редактирование электронных таблиц осуществляется функциями MS Access.

Функциональность, предназначенная для решения специфических задач гидрогеологии, реализована в виде дополнительных модулей в приложении MS Access и обеспечивает:

- ввод данных (ввод и редактирование результатов измерений с возможностью выбора наблюдательных скважин и т.д.);
- проверка текущего состояния скважин;
- построение диаграмм химсостава воды для различных регионов;
- построение геологических разрезов;

- систематизация, упорядочение всех геолого-гидрогеологических, гидродинамических, литофациальных, фильтрационных, геохимических, геоэкологических и прочих параметров.

Предпосылками послужило то, что для поиска необходимых параметров тратится довольно много времени и порой можно сильно запутаться в этом огромном количестве гидрогеологической информации.

Подобное развитие структуры информационного сопровождения является наглядной иллюстрацией неограниченных возможностей развития ГИС-систем практически в любых направлениях, на любой стадии функционирования структур обеспечения информацией. Использование созданной модели позволит на территории ТКАБ автоматизировать ввод и накопление информации и обеспечить более обоснованный выбор и принятие оперативных управленческих решений.

Кроме того, эксплуатация модели даст возможность получить необходимые данные для отслеживания тенденций изменения химического состава воды, изменения экологии и прогнозирования ситуации на долгосрочный период с учетом широкого спектра сопутствующих факторов (социальных, медико-биологических, экономических и т.д.).

Оперативность представления данных, интеграция разноплановой информации на единой географической основе в соединении с инструментами математического моделирования выделяет ГИС-системы из ряда существующих прикладных программ, объясняет привлекательность ГИС-технологий для специалистов практически любых отраслей с целью повышения информативной обеспеченности принятия решений [6].

База гидрогеологических параметров артезианских скважин состоит из совокупности объектов – таблиц, форм, страниц доступа к данным, запросов и отчетов, используемых для представления данных и управления ими; макросов и модулей, используемых для программирования базы данных в диапазоне от автоматизации стандартных задач до создания изошренных, настраиваемых интерфейсов базы данных с совершенно новыми функциями и возможностями. В приложении Microsoft Access происходит управление объектами базы данных с помощью окна базы данных.

Общая характеристика скважины

№ п/п	Отм. устья	Глуб.	дата бурения	дата экспл.	назначение скважины
1	1040	144	1932	1971	разведочная на газ
2	1600	500	1933	1975	разведочная на газ
3	5	453	1967	1978	водоснабжение
4	4	450	1964	1965	водоснабжение
5	5,7	478	1973	1974	водоснабжение

Привязка

№ п/п	1-й №	обл	район	населенный пункт	СШ г	СШ м	ВД г
1	2	Р.Д.	Ахтынский	с. Ахты	43	40	46
2	3	Р.Д.	Ахтынский	с. Ахты	43	40	46
3	193	Р.Д.	Бабаюртовский	с. Адильянгиюрт	43	34	46
4	195	Р.Д.	Бабаюртовский	с. Адильянгиюрт	43	34	36
5	949	Р.Д.	Бабаюртовский	с. Янгиюрт	43	35	46

Формула солевого состава

№ п/п	номер ВГ	формула
1	1	$\begin{array}{l} \text{HCO}_3 \text{ 48 SO}_4 \text{ 24 CO}_3 \text{ 22.1 Cl 5.9} \\ \text{M 0.77} \quad \text{-----} \\ \text{Na 83 Ca10.2 Mg 8.5} \end{array}$
2	1	$\begin{array}{l} \text{HCO}_3 \text{ 48 SO}_4 \text{ 24 CO}_3 \text{ 22.1 Cl 5.9} \\ \text{M 0.77} \quad \text{-----} \\ \text{Na 83 Ca10.2 Mg 8.5} \end{array}$

Геологическая характеристика

№ п/п	№ стр	возраст	кровля	подошва	Мощн.	наименование породы (наименование)
1	1	Q		4.8	4.8	суглинок
1	2	J	4,8	8,3	3,5	сланец темно-серый серый, глинистый
1	3	J	8.3	10.2	1.9	брекчии
1	4	J	10.2	10.7	0.5	песчаник темно-серый тонкозернистый, слюдястый

Опробование				
№ п/п	номер ВГ	понижение, м	дебит, л/с	продолжительность, сут
1		самоизлив	2,3	
2		св нет	0,6	
3	2	21,5	14,5	
4		11	10	

Характеристика ВГ			
порядковый но	номер ВГ	глубина появления воды,	установленный уровень воды, м
1	1	75,7	фонт
2	1	24,8	св нет
3	1	375	22
4	1	192	3,8
5	1	395	18

Качество воды					
№ п/п	температура, г	РН	Жест. общая,	сух остаток, мг/л	окисл перман
3	21	7	354	617	
4	24	6,5			
6		8	2	680	
9	22	7,8	7,38	995	
12	18	7	0,29	1992	

Характеристика фильтра					
порядковый номер	номер ВГ	интервал устан	интервал устан	интервал устан	интервал устан
3	1	386	393	443	449
4	1	192	198	408	421
4	2	207	213	434	442

Рис.4. Образцы информационных таблиц гидрогеологических параметров в базе данных скважин ТКАБ

До начала работы с базой данных необходимо определить способ хранения информации. В приложении Microsoft Access данные хранятся в таблицах. Структуризация баз данных основывается на группировании показателей информации по видам гидрогеологических, геоэкологических работ первичного уровня. Своды данных формируются с использованием традиционных форм геологической документации. Отдельно группируются собственно геоданные и привязочные показатели. Таблицу составляют строки и столбцы, а некоторые отдельные данные вносятся в ячейки, создаваемые сеткой. В Microsoft Access столбцы таблицы представляют собой поле, которое является категорией информации, в то время как каждая строка содержит определенную запись, хранящую всю информацию по определенному вопросу. Структура баз данных определяется группированием показателей в табличной форме, а в составе таблиц – тематическими разделами.

В каждой базе данных содержится одна или несколько таблиц для хранения связанной информации. Например, в одной таблице может храниться информация о привязке скважин, в другой – о геологической характеристике, в третьей – о химическом составе воды и т. д. (рис.4.). Между этими таблицами должны быть определены связи, чтобы все таблицы могли работать слаженно. Такая комбинация всех таблиц и их взаимная связь составляют интегральную основу реляционной базы данных. Реляционная база данных имеет несомненные преимущества, перед плоскими, поскольку мы избавляемся от многократного ввода одних и тех же данных в одну таблицу, тем самым уменьшается дисковое пространство и резко уменьшаются ошибки.

На рис.5. показана схема связи между тематическими таблицами базы данных региональных гидрогеологических, инженерно-геологических и геоэкологических параметров артезианских скважин.

При работе с информацией в базе данных оператор зачастую имеет дело с единственной записью. В частности, при просмотре данных таблицы с помощью формы представляется возможным одновременно видеть только одну запись, и при печатании отчета каждая напечатанная строчка обычно представляет единственную запись (При

проектировании формы и отчета оператор определяет, какие поля из этой записи будут включены).

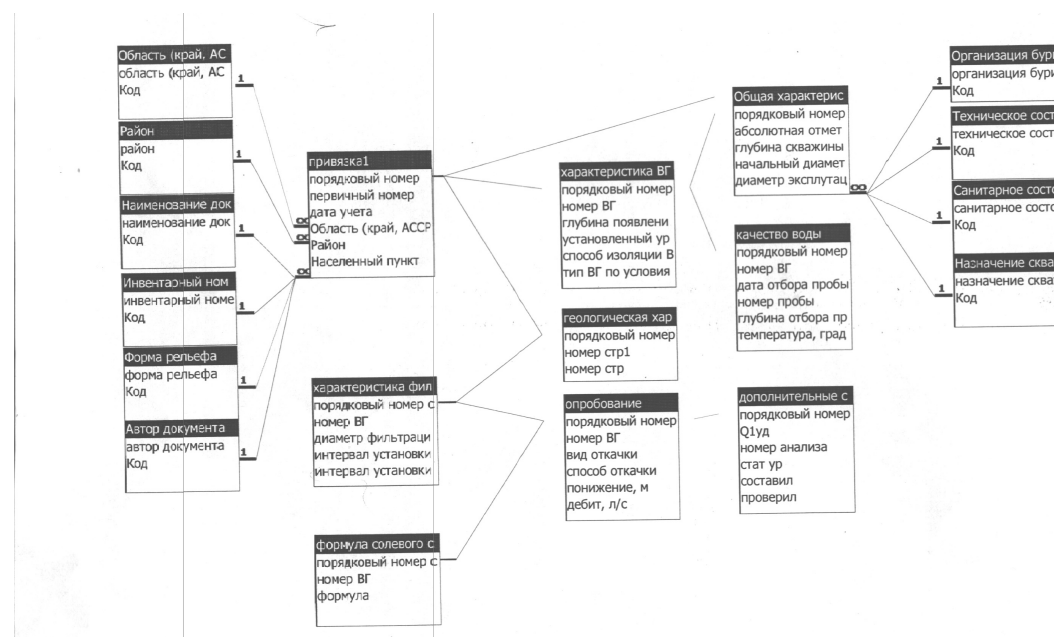


Рис.5. Схематическая связь таблиц гидрогеологических параметров артезианских скважин в базе данных

Но, наряду с этими преимуществами, Microsoft Access имеет и несколько недостатков, такие как: ограниченные графические возможности, невозможность ввода и редактирования формул, которые часто встречаются в научной информации. Несмотря на это, опыт составления базы данных путем систематизации геолого-гидрогеологических параметров (местоположение, координаты, глубина скважины, водоносного горизонта, дебит, давление, мощность, температура, химический состав, водопроводимость, режим и т.д.) Северо-Дагестанского артезианского бассейна показал, что приложение Microsoft Access успешно можно применять и для составления базы данных научной информации в последующем применении как приложение к ArcView.

Для того чтобы осуществить привязку базы данных в ArcView необходимо сначала нанести все скважины, по которым имеется информация, на электронную карту. Это можно сделать привязав предварительно карту, на которой указаны скважины, а затем нанести по ней скважины в виде точек. Другой способ – визуальный метод нанесения точек. Точнее конечно метод, основанный на привязке карты, однако из-за отсутствия карт со всеми скважинами нам приходится совмещать оба эти способа.

Привязка карты выполняется следующим образом. Карту-основу на бумажном носителе импортируют в проект в виде растрового рисунка. Далее привязка карты к координатам осуществляется с помощью реперных точек, которыми могут быть точки координатной сетки исходной карты или объекты с известными координатами. Дальнейшее создание карты в ArcView выполняется инструментальными средствами, предоставляемыми программой. После нанесения всех точек, соответствующих скважинам, на карту программа автоматически создает таблицу атрибутов для вновь созданного слоя. В этой таблице нам достаточно указать номера скважин, соответствующие им в базе данных. Далее необходимо выполнить соединение атрибутивной таблицы слоя в ArcView и таблицы из базы данных, составленной в MS Access. Соединение таблиц выполняется по ключевым полям. После соединения таблиц основная работа по созданию базы геоданных в ГИС системе завершена. Пользователи данной системы могут выполнять все действия, которые им предоставляет система ArcView. Они имеют возможность выполнять разнообразные статистические вычисления, построение отчетов, диаграмм, выполнение запросов. При наличии такого приложения как 3D Analyst имеется

возможность построения трехмерных разрезов, наглядно отображающих состав Земной коры, на основе геофизических и других данных. Ранее, до использования ГИС-систем, построение разрезов выполнялось вручную, что является достаточно трудоемким и не самым точным методом.

Построение разрезов в программе 3D Analyst выполняется следующим образом. По имеющимся скважинам составляется таблица, в которой указывается, на какой высоте находятся различные породы. Таким образом, скважины представляют собой опорные точки. По этим точкам программа экстраполирует поверхность разреза.

Несомненным преимуществом ГИС-систем является то, что они позволяют представить данные, хранящиеся в различных таблицах и в описательном виде – более наглядном для человека виде изображения.

Литература

1. Klein H.K., Hirschheim R.A. A Comparative Framework of Data Modelling Paradigms and Approaches. The Computer Journal. Vol.30. № 1. 1987. Pp. 8-15.
2. Марков Б.Л. Проектирование систем регистрации и анализа данных //Сборник научных трудов ИМВС РАН за 2002 г. – М., 2003.
3. Марков Б.Л. Организация данных в системах мониторинга //Высокопроизводительные вычислительные системы и микропроцессоры. Сборник научных трудов ИМВС РАН за 2000г. – М., 2000.
4. IDEF1X. "FIPS Integration Definition for Information Modeling (IDEF1X)," Federal Information Processing Standards Publication 184, Computer Systems Laboratory, National Institute of Standards and Technology. – 1993.
5. Manna, Z., Pnueli A.: The Temporal Logic of Reactive and Concurrent Systems. Springer-Verlag, 1992.
6. Создание гидрогеологических карт с применением компьютерных технологий. – М.:ВСЕГИНГЕО МПР РФ, 2001. 196 С.