НАУЧНЫЕ СТАТЬИ

Гидрометеорология и экология
№ 1 2000

УДК 556.3

К ТЕХНОЛОГИИ ПОСТРОЕНИЯ МОДЕЛЕЙ ПРОГНОЗА ИЗМЕНЕНИЙ ЭКОГЕОСИСТЕМ

Канд.техн.наук Л.М.Павличенко

Метод вращения многомерных осей (матрицы факторных или компонентных нагрузок) предлагается использовать для построения новых моделей, учитывающих динамику и широкий круг исходных данных по параметрам экогеосистем, при этом модели, описывающие взаимосвязи на разные временные или разные качественные уровни, объединяются в систему матричных уравнений, а вращение осей осуществляет их многоуровневую связь. Рассматривается технология построения моделей восполнения исходной информации, комплексных моделей, формализации принципа проблемной организации экологического мониторинга.

Осознание высокой уязвимости экосистемы нашей планеты, понимание острой необходимости принятия действенных мер по реализации программы устойчивого развития биосферы в глобальном масштабе выдвинуло экологические проблемы всем мировым сообществом в ряд наиболее важных, приоритетных глобальных проблем. Это подтвердило своими решениями конференция ООН по окружающей среде и развитию (КООНОСР) в Рио-де-Жанейро в июне 1992 года. На конференции был принят всемирный план действия - Повестка дня на XXI век, направленный на достижение устойчивого развития и принят свод принципов для последующего развития, определяющих право народов на развитие и их обязанности по сохранению окружающей среды. Определены и стратегические задачи в области охраны и рационального использования природных ресурсов, поэтому тщательное изучение, разработка научных основ оценки состояния и прогнозов поведения геоэкосистем в условиях техногенеза приобретают особую актуальность и приоритетность.

Характерной особенностью геоэкологической информации является ее разнородность и огромное число параметров, характеризующих качество всех компонентов окружающей среды. В связи с этим вопрос прогнозирования ожидаемого состояния окружающей среды в хозяйственной деятельности решается не на должном уровне. Одной из важнейших причин такой ситуации является отсутствие совершенной методики прогнозирования состояния сложных объектов, как в Казахстане, так и за рубежом. Прогнозированию частично поддаются воздействия на отдельные компоненты природной среды, например, загрязнение воздушного бассейна, развитие процессов подтопления.

Используемые методы комплексного прогнозирования не совершенны по разным причинам, основными из них являются практически полное отсутствие определений параметров, характеризующих среду, и небольшая продолжительность (чаще всего отсутствие) режимных наблюдений. Так, используемый традиционно метод экстраполяции, заключающийся в ретроспективном изучении процессов и распространении выявленных закономерностей на будущее, не всегда применим из-за недостатка информации. Методы экспертных оценок имеют существенный недостаток, присущий субъективным оценкам.

Но поскольку для реализации экологических мероприятий требуются очень большие средства, наличие методики прогнозирования, приспособленной к реальному состоянию исходного фактического материала, крайне необходимо.

Традиционно окружающая среда, как экосистема, изучается с применением двух методологических аспектов: исторического и системного. Исторический включает два подхода: принцип актуализма (Ч.Лаель), заключающийся в изучении эволюции экосистемы по схеме - знание современных нам феноменов дает ключ к пониманию всех прошедших периодов, т.н. эпигнозный подход, и принцип униформизма - настоящее есть ключ к познанию будущего. В эпигнозной формулировке принцип униформизма широко используется в различных методах прогнозирования. Пространственно-временную трактовку этих принципов можно представить следующим обобщением количественной теории информации: если в одной точке проводить наблюдения бесконечно долго, в ней можно отметить все возможные процессы и явления в системе, или же рассматривая бесконечно большую территорию в один и тот же момент времени можно наблюдать все возможные изменения в системе.

Системная парадигма в качестве основных свойств природных систем выделяет иерархичность, наличие взаимосвязей и самоорганизацию. Построение модели сложного объекта или процесса принципиально не возможно с помощью одной модели вследствие большого числа исследуемых параметров и вероятностного характера динамических изменений в экогеосистеме и должно осуществляться системами моделей. Вторая рекомендация системного анализа – самого лучшего результата при изучении системы можно добиться при изучении трехуровневой иерархии: сама система, над- и подсистема, (т.е. входы и выходы системы). Еще одна очень важная рекомендация – необходимо учитывать соотношение уровня иерархии и количества параметров, описывающих систему [17].

При выборе системы моделей, максимально полно учитывающих исторический и системный подходы, необходимо ориентироваться в возможностях различных типов существующих математических моделей, применяемых для описания сложных природных объектов и процессов. Приведем краткий обзор таких моделей.

Традиционно исследование сложных систем базируется на методах и моделях, основанных на анализе и математической обработке показателей связи между отдельными элементами и параметрами. Чаще всего в регрессионном, компонентном и факторном анализе такими показателями служат коэффициенты корреляции. Меры связи, используемые в этих моделях – вероятностные статистические характеристики, и главной проблемой их применения является вопрос о линейности связей [21].

Моделирование в географии, как и разработка системного подхода, развивалось совершенно самостоятельно, и даже типизация географических (имеются в виду территориальные - ландшафтные) моделей в корне отличается от типизации моделей в других областях знания. Здесь на первое место выдвигается задача выделения функции ГС, которая устанавливается на основе изучения либо взаимосвязей компонентов ГС, либо миграции вещественно-энергетических потоков [20,23]. С математической точки зрения, практически все эти модели относятся к графовым, причем характер взаимосвязей и их количественное выражение задавались на уровне экспертных оценок.

Школа П.К. Анохина разработала математический аппарат информационной технологии, позволяющей обрабатывать естественные распределения переменных, неявно отображенных в массивах экспериментальных данных. В информационной технологии на основе нетрадиционной системологики – технике анализа системы обработки коэффициентов связи – связь понимается как мера упорядоченности элементов в системе по сравнению с их полной беспорядочностью и энергетических или информационных затрат при вхождении элементов в систему. Методы, основанные на этой технологии, хорошо зарекомендовали себя при исследовании функциональных систем живых организмов. Использование логики анализа функциональных систем организмов в сочетании с новейшим методом расчета связей и логикой графов позволяет достаточно точно определять иерархическую структуру, эмерджентные свойства и механизм поведения исследуемой системы [21].

В основу технологии П.К.Анохина положен принцип доминанты, реализуемый для базисных подсистем, где два результата взаимодействия порождают третий эмерджентный результат. Принцип доминанты – один из фундаментальных в организации живых систем, однако, он означает, что подсистема, в которой эмерджентный результат образуется более чем двумя другими результатами, неустойчива, а, следовательно, не может стать базисной и включенной в характеристическую цель последовательных результатов. Таким образом, для исследования надорганизменных систем произвольной природы принцип доминанты не может считаться единственно приемлемым, а потому использование технологии системологики для геоэкосистем (в частности, эколого-гидрогеологических) пока не представляется возможным [21].

Другим распространенным подходом к моделированию сложных систем являются динамические модели и теория катастроф. Модели и теоретические приложения теории катастроф использовались для решения разнообразных задач экологии, геоэкологии, биологии и экономики. Тем не менее, практическое применение теории катастроф, базирующейся на математической теории дифференциальных уравнений, решениями которых являются разрывные функции, на практике трудно осуществимо. В первую очередь, из-за необходимости большого числа наблюдений по всем используемым показателям для определения численных значений коэффициентов дифференциальных уравнений. Очень часто сбор таких данных требует постановки дорогостоящих и длительных по времени натурных экспериментов и исследований. Весьма сложной задачей является обоснование дифференциальных функций, наиболее адекватно связывающих переменные исследуемой системы. Главную же трудность представляет сложность решения систем дифференциальных уравнений при большом их числе (даже при использовании современных быстродействующих ЭВМ). Вследствие этого, т.е. вследствие, по-видимому, недостаточной теоретической разработанности и сложностей математического характера в практических приложениях используется преимущественно одна модель катастрофы – "складка" [21].

Популярность динамических моделей [6] обязана большой гибкости методов, применяемых для описания динамики систем и включающих нелинейные реакции компонентов на регулирующие переменные, а также положительные и отрицательные обратные связи. Но обычно невозможно учесть уравнения для всех компонентов системы, так как даже при наличии суперкомпьютеров имитация реальных процессов для крупных объектов быстро становится слишком сложной. Поэтому возникает необходимость в некотором абстрагировании, основанном на "здравом смысле" и на допущениях относительно того, какие из многих компонентов экосистемы в действительности управляют ее функционированием. Итак, главными недостатками динамических моделей являются принципиальная невозможность построения полной системы уравнений и субъективизм, неизбежно вносимый ограничениями при упрощении системы. Это можно проследить на примере гидрогеологии.

Как известно, полная система уравнений гидродинамики, диффузии и массообмена в подземных водах содержит несколько уравнений с неизвестными давлением, плотностью, пористостью, концентрациями вещества в жидкой и твердой фазах, компонентами скорости ассоциированного жидкостью вещества. Входящие в эту систему начальные плотность жидкости, вязкость, пористость и проницаемость, вообще говоря, не являются постоянными, а зависят от концентраций вещества в жидкой и твердой фазах. Так, например, в глинистых и торфяных грунтах под воздействием ионообменных процессов появляется способность к коагуляции (или пептизации), что немедленно сказывается на повышении (или понижении) параметров водопроводимости. Заметные изменения происходят также при растворении или кристаллизации солей [4].

Значительные трудности возникают и при выборе гидрогеохимических параметров (коэффициенты молекулярной и конвективной диффузии, гидродинамической дисперсии, постоянных линейной изотермы, констант скорости адсорбции или десорбции и др.), где отмечается зависимость значений не только от степени неоднородности пород, но от несоответствия численных значений параметров, получаемых в лабораторных и полевых условиях.

Все это привело к тому, что в практике гидрогеологических исследований многофакторные модели массопереноса в подземных водах, включающие нелинейные уравнения равновесия и кинетики взаимодействия твердой и жидкой фаз, не нашли в настоящее время достаточно широкого применения, хотя существует довольно большое количество программ, реализующих какие-то конкретные физические или химические модели. Чаще всего применяются они для относительно небольших объектов, а при попытках решения больших региональных задач, которые как раз и необходимо решать при прогнозах изменения эколого-гидрогеологических условий, ошибка прогноза зачастую достигает 100% при сохранении направленности процесса. Такая постановка задачи, как правило, не является удовлетворительной, поскольку оценка направленности процесса может реализоваться более простыми моделями.

Среди других математических методов, разрабатываемых для реализации системного подхода, известны матричные модели, используемые для прогноза состояния системы через заданный интервал времени с помощью, так называемой главной матрицы, характеризующей вероятность перехода системы из одного состояния в другое. Строится эта матрица на основе накопленных за длительный промежуток времени наблюдений за характером этих переходов и при условии постоянства этого характера во времени. Довольно часто имеющихся данных недостаточно, чтобы достоверно оценить вероятности или скорости переходов системы, особенно если эти переходы редки. Как и в других моделях, ее адекватность реальной системе зависит от длительности наблюдений за состоянием системы. Однако для процессов, охватывающих достаточно длительные промежутки времени, возможность предсказания поведения системы весьма проблематична: по литературным данным, даже тридцатилетние ряды наблюдений за уровнями подземных вод не дают гарантии выделения всех имеющихся циклов [см., например, 7].

Естественно, что матричные модели успешно реализуют прогнозы эволюционного развития геоэкологических систем, любое же значительное техногенное вмешательство внесет весьма заметные изменения главной матрицы, предсказать которые вряд ли можно точно. Выходом из этого положения может послужить переход к многомерным моделям, учитывающим не только временные (вертикальные) взаимосвязи, но и горизонтальные, т.е. взаимосвязи между параметрами на каждом временном интервале.

Такие многомерные статистические модели успешно применяются для описания и прогноза состояния сложных систем. К прогнозным относятся всевозможные разновидности корреляционно-регрессионного, дискриминантного, канонического анализов, марковские модели, часть методов теории распознавания образов, метод группового учета аргументов и др. Как и для упомянутых выше моделей, главным критерием адекватности прогнозных статистических моделей является наличие длительного ряда наблюдений за большим набором параметров [1-3, 5, 6, 8, 9 и др.].

К описательным статистическим моделям относятся компонентный, факторный, кластерный анализы, метод многомерного шкалирования, некоторые методы теории распознавания образов. Описательные модели лишены главного недостатка прогнозных – значительной доли субъективности, выражающейся в выборе значащих параметров, необходимости априорного задания опорных функций, собственных чисел переходной матрицы и других предположений. Однако неоднозначность классификаций (в зависимости от критерия объекты могут принадлежать нескольким классам) и возможность лишь качественных (по принадлежности) прогнозов сдерживает их применение. К недостаткам этих моделей относится невозможность численного сравнения результатов их применения для разных исходных матриц (например, на разные моменты времени) вследствие нестабильности нагрузок на исходные признаки, которые как раз и определяют классификационные критерии.

Из приведенного обзора методов моделирования сложных систем видно, что все они имею свои преимущества и недостатки. Поэтому актуальной задачей остается разработка новых методов, по возможности лишенных указанных недостатков и максимально использующих сильные стороны известных моделей.

В работе [14] рассмотрено соотношение статистических и детерминистических моделей и показано, что статистические и детерминированные модели очень тесно взаимосвязаны и взаимообусловлены. Фактически они образуют двойственную противоположную пару, в которой статистические модели определяют область во времени, пространстве свойств и физическом, определяющую концепцию, лежащую в основе детерминированной модели, и в пределах которой "работают" эти детерминированные модели.

В случаях, реализующих ситуацию одна причина – одно следствие, наиболее широко распространены детерминированные модели. Если существует множество причин и множество порождаемых ими следствий, ситуация меняется в пользу методов многомерной статистики, хотя и здесь имеются удачные попытки построения детерминированных моделей.

Моделирование сложных эколого-гидрогеологических объектов и процессов возможно системами, состоящими, как только из детерминированных, так и только статистических моделей. Однако наибольшего успеха для достижения максимальной степени адекватности, видимо, можно добиться при их комплексировании с использованием сильных сторон как детерминированных моделей (большая точность в узких пределах), так и статистических методов (выделение однородных совокупностей и возможность обработки большого количества разнородной информации в случае многомерных моделей).

При создании системы моделей необходимо всегда учитывать, что точность детерминированных моделей во многом определяется правильным выбором краевых условий, а точность статистических – правильным набором исходных признаков [14].

Многолетний опыт автора в использовании классической модели компонентного анализа (КА) в гидрогеологии, гидрогеохимии, гидрохимии и географии показал, что он отвечает требованиям системности. Так, он дает возможность выделять функции системы и ранжировать их по вкладу в суммарную дисперсию системы (эмерджентность), учитывает взаимосвязи, и самоорганизацию системы (наличие прямых и обратных связей определяется соотношением знаков взаимосвязей внутри компоненты). Новым результатом модели является не только выявление функции системы на основе интерпретации системы взаимосвязей исходных признаков, но и районирование территории по интенсивности проявления этой функции.

Например, при анализе процессов формирования гидрогеохимических условий древней дельты р. Или как выход модели КА получен результат, показывающий отсутствие взаимодействия водоносных горизонтов, противоречащий взглядам гидрогеологов на момент защиты кандидатской диссертации (1984 г.) [10] и установленный в 1989 г. натурными исследованиями ИГГиГФ и Института озероведения [18]. При изучении влияния пруда-накопителя очищенных сточных вод г. Жанатаса на подземные воды КА позволил установить характер процессов доочистки в пруде-накопителе, отсутствие влияния на подземные воды и местные источники загрязнения колодцев в прилегающих поселках. При выявлении источников загрязнения р. Сырдарьи пестицидами на основе анализа химического состава поверхностных и поземных вод в районах расположения массивов орошения установлено в основном местное загрязнение пестицидами при незначительном влиянии трансграничного переноса из Узбекистана [15, 16, 19].

Как уже упоминалось, классическая модель КА предусматривает возможность качественного прогнозирования геоэкологических процессов. Примером такого качественного прогноза может служить выявление тенденций загрязнения р. Илек бором. КА позволил осуществить прогноз зоны формирования вторичного очага загрязнения подземных вод бором в зоне расположения питьевого водозабора, не картируемого традиционными методами: был закартирован участок накопления транзитного бора, осаждаемого на механическом геохимическом барьере [12].

Возможность качественного геоэкологического прогнозирования при обработке разнородной информации демонстрирует задача районирования интенсивности проявлений техногенных процессов на основе компонентного анализа карты экзогенно-геологических процессов в Казахстанском Приаралье: КА позволил выделить зоны влияния продвижения границы Арала, самоизлива скважин, массивов орошения, зоны возможных пыльных бурь.

Для количественных прогнозов классическая модель КА непригодна, и потому приходится использовать уравнение регрессии исходных признаков на компонентах. Примером такого способа применения КА является прогноз минерализации подземных вод на основе обработки гидрогеологической карты (Юго-Западное Прибалхашье). Были построены две модели территории на основе КА: первая – на ненарушенный (до строительства ЮКГРЭС), в ней получено уравнение для минерализации по главным компонентам. Вторая – с учетом результатов детерминированной модели фильтрации получены новые значения главных компонент. При прогнозировании использован метод аналогии, результат сопоставлялся с аналитическими расчетами минерализации в блоках методом поршневого вытеснения. К основным недостаткам такого метода относится отсутствие учета изменения компонентных нагрузок, характеризующих гидрогеологическую и гидрогеохимическую ситуацию после техногенного воздействия [13].

Хотя классический КА реализует системный подход и позволяет решать широкий круг научных и практических задач, прогнозные его возможности ограничены. Это либо качественные прогнозы – устанавливаются новые зоны проявления действующих процессов, либо строится уравнение регрессии на компонентах, однако в результате изменения количества исходных признаков на прогнозный момент времени возникают сомнения в полноте отражения ситуации (детерминированные модели могут дать ограниченный набор прогнозных параметров, а чаще всего - один).

Выходом из этого положения может послужить переход к многомерным моделям, учитывающим не только временные (вертикальные) взаимосвязи, но и горизонтальные, т.е. взаимосвязи между параметрами на каждом временном интервале. Таким образом, методология моделирования сложных экогеосистем требует перехода к системам детерминированных и статистических моделей, что позволит повысить степень адекватности системы моделей изучаемым объектам и процессам.

Отправной точкой при разработке новых моделей для нас является формальное сходство записей матричной модели [9]:

$$\mathbf{A} \cdot \mathbf{a}_t = \mathbf{a}_{t+1}, \qquad (1)$$

где: **A** – главная матрица, характеризующая вероятности перехода составляющих экосистемы из одного состояния в другое, $\mathbf{a}_t$ и $\mathbf{a}_{t+1}$ – векторы состояний на моменты времени t и t+1, и матричного уравнения вращения компонентных осей для упрощения структуры:

$$\mathbf{A}_2 = \mathbf{B} \cdot \mathbf{A}_1, \mathbf{F}_2 = \mathbf{F}_1 \cdot \mathbf{B}, \qquad (2)$$

где: **A** – матрица компонентных (факторных) нагрузок, **F** – матрица значений компонент (факторов) для каждого объекта, **B** – матрица вращения. То есть если индексы 1 и 2 заменить на t_1 и t_2, матрица **B**, не изменяя структуры признаковых осей, будет представлять собой матрицу вращения вокруг временной оси. Эта матрица является аналогом главной матрицы в матричной модели. Вычислить ее можно традиционными методами матричной алгебры с использованием свойств симметричных ортогональных матриц, если известны $\mathbf{A}_1$ и $\mathbf{A}_2$ или $\mathbf{F}_1$ и $\mathbf{F}_2$, т.е. если полностью известны нормированные матрицы $\mathbf{Y}_1$ и $\mathbf{Y}_2$ исходных данных на начальный и прогнозный момент времени $\mathbf{X}_1$ и $\mathbf{X}_2$:

$$\mathbf{Y}_1 = \mathbf{A}_1 \cdot \mathbf{F}_1, \mathbf{Y}_2 = \mathbf{A}_2 \cdot \mathbf{F}_2. \qquad (3)$$

Непосредственный перевод $\mathbf{Y}_1 \rightarrow \mathbf{Y}_2$ или $\mathbf{X}_1 \rightarrow \mathbf{X}_2$ не возможен, поскольку мы предполагаем взаимозависимости между параметрами, а в этом случае ранги матриц исходных данных окажутся меньше числа признаков и при промежуточных построениях квадратные матрицы окажутся вырожденными, их дискриминанты обращаются в 0, и задача не имеет решения в полном пространстве признаков. Следовательно, нужно искать невырожденные матрицы с размерностью, соответствующей рангу (свертывание пространства признаков), что и делает компонентный анализ при вычислении матрицы компонентных (факторных – в случае использования факторного анализа) нагрузок. Можно, конечно, использовать и другие способы свертывания пространства, но для компонентного и факторного анализов хорошо разработан математический аппарат.

Итак, традиционную процедуру вращения компонентных осей (матрицы компонентных или факторных нагрузок) можно использовать для построения новых моделей, учитывающих динамику и широкий круг исходных данных. Технология построения системы моделей заключается в том, что модели, описывающие взаимосвязи на разных временных или качественных уровнях, объединяются в систему матричных уравнений, а вращение осей осуществляет их многоуровневую связь.

В работе автора [11] предложен ряд алгоритмов, разработанных на основе вращения компонентных осей, для восстановления пропущенных значений в матрице исходных данных, сравнения результатов описательных многомерных моделей (анализ динамики экогеосистем). Для примера реализации идеи остановимся на задаче восстановления пропущенных значений в матрице исходных данных.

Большой класс алгоритмов прикладной статистики посвящен решению задачи восстановления "стертых" (пропущенных) наблюдений в матрицах исходных данных, т.е. восстановлению пропущенных отдельных элементов или частей строк (столбцов). Эта задача весьма актуальна в случаях недостаточного количества информации или при большой дороговизне получения исходных данных, когда каждый признак (столбец с пропущенными значениями) или объект (строка, в которой нет замеров некоторых признаков) могут дать полезную информацию, и было бы слишком расточительно исключать их из рассмотрения.

Решение такой задачи можно построить на огромном количестве алгоритмов - от самых простых (замена пропущенных значений признака его средним арифметическим значением, которое оценивается по имеющимся реализациям) до самых сложных алгоритмов теории распознавания образов [см., например, 2, 3, 8, 9 и др.]. Общим требованием всех этих алгоритмов для исходной (экспериментальной) матрицы является условие ее однородности, т.е. если возникает необходимость объединения нескольких частных матриц в одну общую, то необходимо проверить с помощью существующих статистических критериев, являются ли эти частные матрицы выборками из одной и той же генеральной совокупности.

Если же полная матрица составлена из нескольких частных, отстоящих друг от друга на некоторые временные интервалы, условие однородности сводится к проверке стационарности временных рядов [5]. Когда такие ряды не удовлетворяют требованиям стационарности, для восстановления пробелов их значений применяют различные методы регрессионного и тренд- анализов, авторегрессии, сериальной корреляции, спектрального анализа и др. [5]. Естественным требованием для повышения точности оценок пропущенных значений временных рядов является их длина.

Существующие способы восстановления, основанные на теории распознавания образов позволяют дополнить матрицу исходных данных при наличии полного количества объектов и отсутствии у них замеров по части (чаще всего 1/3 объема) признаков [3]. В настоящее время это самые распространенные методы.

Метод группового учета аргументов [8] может восстанавливать информацию, превосходящую по объему исходную матрицу, но при условии априорного задания опорных функций и задания собственных чисел полной матрицы, что означает субъективное распределение ролей между процессами, формирующими значения признаков. Так, последние алгоритмы этого метода – AELITA (многорядный вариант метода группового учета аргументов), RULA и TAIS (шаговой регрессии) – осуществляют предварительное функциональное преобразование переменных и строят в качестве частных моделей полиномы двух переменных. Как известно [22], ввод функциональных зависимостей вносит ложные корреляции, которые повысят величину множественного коэффициента корреляции, а, следовательно, понизится точность всех последующих статистических моделей.

Рассмотрим новую технологию на примере восстановления пропущенных значений, причем здесь вращением многомерных осей реализуется традиционный подход к восполнению. Такая задача может возникнуть при единовременном опробовании территории, при котором достоверность части анализов вызывает сомнения, либо по части территории выполнен неполный комплекс определений.

Из матрицы с пропусками $\mathbf{Y}[m\times n]$ можно составить следующие матрицы: $\mathbf{Y}_1[m\times n_1]$ – матрица, из которой исключены все строки с пропущенными значениями (неполная по объектам); $\mathbf{Y}_2[m_1\times n]$ – матрица, из которой исключены все столбцы с пропусками (неполная по признакам); и, наконец, $\mathbf{Y}_3[m_1\times n_1]$ – матрица, из которой исключены как строки, так и столбцы с пропущенными значениями (неполная по признакам и по объектам). Здесь и далее рассматриваются уже нормированные матрицы исходных данных.

Применив модель компонентного анализа к каждой из этих матриц, получим:

$$\mathbf{Y}_1[m\times n_1]=\mathbf{A}_1[m\times q_1]\cdot\mathbf{F}_1[q_1\times n_1]; \tag{4}$$

$$\mathbf{Y}_2[m_1\times n]=\mathbf{A}_2[m_1\times q_2]\cdot\mathbf{F}_2[q_2\times n]; \tag{5}$$

$$\mathbf{Y}_3[m_1\times n_1]=\mathbf{A}_3[m_1\times q_3]\cdot\mathbf{F}_3[q_3\times n_1]. \tag{6}$$

Матрицы компонентных нагрузок $\mathbf{A}_2$ и $\mathbf{A}_3$ описывают одно и то же пространство признаков, и при условии устойчивости компонент, справедливом при наличии большого числа объектов, представляющих генеральную совокупность, должно выполняться условие $q_2=q_3$ и, следовательно, $\mathbf{A}_2=\mathbf{A}_3$. Однако на практике такое условие практически никогда не выполняется, поэтому матрица $\mathbf{A}_3$ является искаженной матрицей A_2, причем степень искажения определяется разницей в количестве объектов и геометрией их расположения. Безусловно, если пропуски равномерны, искажение минимально, если же они сосредоточены в одном месте – максимально (следует учитывать, что замечание относится лишь к случаю восполнения по одной матрице, в случае же расширения сети опробования при наличии разновременных замеров это условие значительно смягчается). В понятие "геометрия", естественно, входят и локальные изменения значений признаков. Именно возможность искажения и обусловливает применение модели компонентного анализа, где одновременно анализируются не только закономерности изменения каждого признака в отдельности (среднее значение, дисперсия, распределение нормированных значений признаков), но и взаимосвязи между всеми признаками, что позволяет оценить вклад всех признаков в дисперсию каждого.

Используя преобразования ортогонального вращения, получаем для преобразования осей $A_2 \to A_3$:

$$\mathbf{A}_3[m_1 \times q_3] = \mathbf{A}_2[m_1 \times q_2] \cdot \mathbf{B}[q_2 \times q_3]; \qquad (7)$$

Умножив (4) слева на транспонированную матрицу $\mathbf{A}_2'$, имеем:

$$\mathbf{A}_2'[q_2 \times m_1]\mathbf{A}_3[m_1 \times q_3] =$$
$$= \mathbf{A}_2'[q_2 \times m_1] \cdot \mathbf{A}_2[m_1 \times q_2] \cdot \mathbf{B}[q_2 \times q_3].$$

Произведение $\mathbf{A}_2'[q_2 \times m_1] \cdot \mathbf{A}_2[m_1 \times q_2]$ представляет собой малый момент ортогональных матриц компонентных нагрузок $\mathbf{A}$ [9], поэтому последнее уравнение можно переписать в виде:

$$\mathbf{A}_2'[q_2 \times m_1] \cdot \mathbf{A}_3[m_1 \times q_3] = \Lambda_2 \cdot \mathbf{B}[q_2 \times q_3],$$

откуда находим алгоритм вычисления матрицы $\mathbf{B}$, характеризующей искажение компонентных нагрузок в результате неполноты количества объектов:

$$\mathbf{B}[q_2 \times q_3] = \Lambda_2^{-1} \cdot \mathbf{A}_2'[q_2 \times m_1] \cdot \mathbf{A}_3[m_1 \times q_3]. \qquad (8)$$

Повернув на те же углы матрицу значений компонент $\mathbf{F}_3$, получим значения компонент в исправленной системе координатных осей уже для полного числа объектов $\mathbf{F}_3(2)$:

$$\mathbf{F}_3(2)[q_3 \times n] = \mathbf{B}' \quad [q_3 \times q_2] \cdot \mathbf{F}_2[q_2 \times n] \tag{9}$$

и восстановленную по объектам нормированную матрицу $\mathbf{Y}_3(2)[m \times n]$ (по-прежнему сохраняется неполное число признаков):

$$\begin{aligned} \mathbf{Y}_3(2)[m_1 \times n] &= \mathbf{A}_2[m_1 \times q_2] \times \\ &\times \mathbf{B}[q_2 \times q_3] \cdot \mathbf{B}'[q_3 \times q_2] \cdot \mathbf{F}_2[q_2 \times n]. \end{aligned} \tag{10}$$

Определив матрицу поворота $\mathbf{A}_3 \rightarrow \mathbf{A}_1$, обеспечим восстановление $\mathbf{Y}_3$ по числу признаков:

$$\mathbf{Y}_3(1)[m_1 \times n_1] = \mathbf{A}_1[m_1 \times q_1] \cdot \mathbf{C}[q_1 \times q_3] \cdot \mathbf{C}'[q_3 \times q_1] \cdot \mathbf{F}_1[q_1 \times n_1]. \tag{11}$$

Матрицу поворота в данном случае удобнее искать из условия вращения матрицы значений компонент $\mathbf{F}_3 \quad [q_3 \times n_1] \rightarrow \mathbf{F}_1 \quad [q_1 \times n_1]$:

$$\mathbf{F}_3 \quad [q_3 \times n_1] = \mathbf{C}' \quad [q_3 \times q_1] \cdot \mathbf{F}_1 \quad [q_1 \times n_1]. \tag{12}$$

После транспонирования последнее уравнение запишется в виде:

$$\mathbf{F}_3' \quad [n_1 \times q_3] = \mathbf{F}_1' \quad [n_1 \times q_1] \cdot \mathbf{C} \quad [q_1 \times q_3],$$

а после умножения слева на $\mathbf{F}_1 \quad [q_1 \times n_1]$ имеем:

$$\mathbf{F}_1[q_1 \times n_1] \cdot \mathbf{F}_3' \quad [n_1 \times q_3] = \mathbf{F}_1[q_1 \times n_1] \cdot \mathbf{F}_1' \quad [n_1 \times q_1] \cdot \mathbf{C} \quad [q_1 \times q_3].$$

Произведение ортогональных матриц $\mathbf{F}_1[q_1 \times n_1] \cdot \mathbf{F}_1' \quad [n_1 \times q_1]$ представляет собой малый момент ортогональной матрицы значений компонент (факторов) для каждого объекта $\mathbf{F}$ [9] и равно числу объектов, входящих в матрицу $\mathbf{F}_1(n_1)$, поэтому:

$$\mathbf{F}_1[q_1 \times n_1] \cdot \mathbf{F}_3' \quad [n_1 \times q_3] = n_1 \cdot \mathbf{C} \quad [q_1 \times q_3],$$

т.е. получили алгоритмы вращения компонентных осей для восстановления числа признаков:

$$\mathbf{C} \quad [q_1 \times q_3] = n_1^{-1} \cdot \mathbf{F}_1[q_1 \times n_1] \cdot \mathbf{F}_3' \quad [n_1 \times q_3]. \tag{13}$$

Повернув теперь на те же углы матрицу $\mathbf{A}_1[m \times q_1]$, отражающую полный набор признаков, получим восстановленную (исправленную) по количеству признаков $\mathbf{A}_3$ в единицах нагрузок $\mathbf{A}_1$:

$$\mathbf{A}_3(1)[m \times q_3] = \mathbf{A}_1[m \times q_1] \cdot \mathbf{C} \quad [q_1 \times q_3]. \qquad (14)$$

Имея матрицы восстановления по количеству объектов (8) и признаков (13), можно найти полностью восстановленную матрицу нормированных исходных данных из системы уравнений (10) и (11), которые можно переписать в виде:

$$\begin{aligned} \mathbf{Y}_3(2)[m_1 \times n] &= \mathbf{A}_3(2)[m_1 \times q_3] \cdot \mathbf{F}_3(2)[q_3 \times n]; \\ \mathbf{Y}_3(1)[m \times n_1] &= \mathbf{A}_3(1)[m \times q_3] \cdot \mathbf{F}_3(1)[q_3 \times n_1]. \end{aligned}$$

Матрица $\mathbf{A}_3(1)[m \times q_3]$ учитывает полное количество признаков, но не полное количество объектов, поэтому необходимо совершить переход $\mathbf{A}_3(1)[m \times q_3] \to \mathbf{A}_3(1{,}2)[m \times q_2]$ через $\mathbf{B}'$ $[q_3 \times q_2]$:

$$\mathbf{A}_3(1{,}2)[m \times q_2] = \mathbf{A}_3(1)[m \times q_3] \cdot \mathbf{B}' \quad [q_3 \times q_2]. \qquad (15)$$

Повернув на те же углы, как и в (4.9), матрицу значений компонент с полным набором объектов (12), получим:

$$\begin{aligned} \mathbf{F}_3(2{,}1)[q_2 \times n] &= \mathbf{B} \quad [q_2 \times q_3] \cdot \mathbf{F}_3(2)[q_3 \times n] = \\ &= \mathbf{B} \quad [q_2 \times q_3] \cdot \mathbf{B}' \quad [q_3 \times q_2] \mathbf{F}_2[q_2 \times n] \end{aligned} \qquad (16)$$

Тогда полностью восстановленная нормированная матрица исходных данных запишется как

$$\mathbf{Y}_n[m \times n] = \mathbf{A}_3(1{,}2)[m \times q_2] \cdot \mathbf{F}_3(2{,}1)[q_2 \times n]. \qquad (17)$$

Прежде чем подставить (15) и (16) в (17), упростим (16), для чего рассмотрим произведение $\mathbf{B} \cdot \mathbf{B}'$. Из (8) имеем:

$$\begin{aligned} \mathbf{B} \cdot \mathbf{B}' = \mathbf{\Lambda}_2^{-1} \cdot \mathbf{A}_2{}'[q_2 \cdot m_1] \cdot \mathbf{A}_3[m_1 \cdot q_3] \times \\ \times \mathbf{A}_3{}'[q_3 \cdot m_1] \cdot \mathbf{A}_2[m_1 \cdot q_2] \cdot \mathbf{\Lambda}_2^{-1}. \end{aligned}$$

Используем свойство ортогональных матриц и учтем, что главный момент матрицы компонентных нагрузок равен единичной матрице [9]:

$$\mathbf{A}_3[m_1 \times q_3] \cdot \mathbf{A}_3{}'[q_3 \times m_1] = \mathbf{I}[m_1 \times m_1].$$

Тогда:

$$\mathbf{B}\cdot\mathbf{B}'=\mathbf{\Lambda}_2^{-1}\cdot\mathbf{A}_2'[q_2\times m_1]\cdot\mathbf{A}_2[m_1\times q_2]\cdot\mathbf{\Lambda}_2^{-1},$$

поскольку малый момент ортогональной матрицы $\mathbf{A}_2$ есть $\mathbf{\Lambda}_2$. Теперь (16) можно записать как:

$$\mathbf{F}_3(2,1)[q_2\times n]=\mathbf{\Lambda}_2^{-1}\cdot\mathbf{F}_2[q_2\times n]$$

и, следовательно, с учетом (15) и (17) преобразовать к окончательному виду:

$$\mathbf{Y}_n[m\times n]=n_1^{-1}\cdot\mathbf{Y}_1[m\times n_1]\cdot\mathbf{Y}_3'[n_1\times m_1]\cdot\mathbf{A}_2[m_1\times q_2]\cdot\mathbf{\Lambda}_2^{-1}\cdot\mathbf{F}_2[q_2\times n]$$

Возможности метода вращения многомерных осей не ограничиваются разработкой только системы статистических моделей – ниже рассматривается технология разработки комплексной модели для экогеосистем, в которых детерминированные модели учитываются результатами прогнозов и новыми граничными условиями, описывающими результаты предполагаемых (проектных) воздействий.

Постановка задачи для детерминированной модели требует значительного количества исходных данных, характеризующих моделируемую среду, т.е. для комплексной модели полностью известной может быть лишь матрица исходных данных на начальный момент времени (модель начального состояния). Детерминированные модели могут дать лишь небольшое число прогнозных параметров, сюда же можно отнести и параметры, отражающие измененные граничные условия, поэтому прогнозная матрица никогда не может быть полной. Именно на этом этапе необходимо подключение статистических моделей: установленные на начальный момент времени взаимосвязи между признаками с учетом временного вращения помогут восстановить отсутствующие параметры на прогнозный момент времени.

Формализованные исходные данные для детерминированной модели после этапа схематизации и аппроксимации сеточной областью можно представить в виде матрицы, для которой усредненные значения параметров привязаны к центрам блоков сеточной разбивки. Для статистической модели (компонентного анализа) эти данные также являются исходной матрицей, хотя компонентный анализ можно осуществлять и по данным замеров непосредственно в скважинах (реальных точках наблюдения).

Итак, на начальный момент времени имеем две полные матрицы исходных данных в центрах блоков и в реальных точках наблюдения (здесь и далее под исходными данными понимаются нормированные значения каждого признака): $\mathbf{Y}_{сф0}[m\times n]$ и $\mathbf{Y}_{t'0}[m\times n_{t'0}]$, которые максимально полно отражают область фильтрации и химический состав подземных вод. В результате обработки матрицы $\mathbf{Y}_{f\cdot 0}\left[m\times n_{f\cdot 0}\right]$ детерминированными моделями получим один (возможно, несколько) прогнозный параметр. Новые значения части параметров, характеризующих измененные граничные условия, задаются условиями проектируемых воздействий (изменениями в мощностях водоносных горизонтов, площадями, занимаемыми гидротехническими сооружениями, и т.д.), формируя неполную матрицу прогнозных значений – модель конечного (результирующего) состояния.

Часть исходных параметров, для которых пока еще не существует эффективных детерминированных моделей, в прогнозной матрице будет отсутствовать. В частности, это химический состав подземных вод, измененный в результате проектируемых воздействий и более всего влияющий на состояние биотической составляющей экосистемы.

Таким образом, в результате применения детерминированных моделей получаем неполную (выборочную) прогнозную матрицу $\mathbf{Y}_{f\cdot r}\left[m\times n_{f\cdot r}\right]$со значениями в центрах блоков. Восполнить недостающие параметры прогнозной матрицы можно применением многомерных статистических моделей (модифицированным компонентным анализом). Иными словами, комплексная модель необходима для восполнения модели конечного состояния.

Прогнозные значения в центрах блоков можно путем интерполяции пересчитать в прогнозные значения в реальных точках наблюдения (водозабор, наблюдательная сеть, дренажные сооружения и др.). Эту же задачу можно решить и методом вращения многомерных осей. Как и в предыдущем случае, из полных исходных матриц можно составить выборочные с неполным количеством признаков, совпадающим с таковым в прогнозной матрице: $\mathbf{Y}_{-\hbar 0‰}\left[m_1\times n_{-\hbar}\right]$и $\mathbf{Y}_{f\cdot 0‰}\left[m_1\times n_{f\cdot}\right]$.Ко всем пяти матрицам применим модель компонентного анализа:

$$\mathbf{Y}_{-\hbar 0}\left[m\times n_{-\hbar}\right]=\mathbf{A}_{-\hbar 0}\left[m\times q_{-\hbar 0}\right]\cdot\mathbf{F}_{-\hbar 0}\left[q_{-\hbar 0}\times n_{-\hbar}\right], \qquad (18)$$

$$\mathbf{Y}_{f\cdot 0}\left[m\times n_{f\cdot}\right]=\mathbf{A}_{f\cdot 0}\left[m\times q_{f\cdot 0}\right]\cdot\mathbf{F}_{f\cdot 0}\left[q_{f\cdot 0}\times n_{f\cdot}\right], \qquad (19)$$

$$\mathbf{Y}_{f\cdot r‰}\left[m_1\times n_{f\cdot}\right]=\mathbf{A}_{f\cdot r‰}\left[m_1\times q_{f\cdot r‰}\right]\cdot\mathbf{F}_{f\cdot r‰}\left[q_{f\cdot r‰}\times n_{f\cdot}\right], \qquad (20)$$

$$\mathbf{Y}_{-\hbar 0‰}\left[m_1\times n_{-\hbar}\right]=\mathbf{A}_{-\hbar 0}\left[m_1\times q_{-\hbar 0‰}\right]\cdot\mathbf{F}_{-\hbar 0‰}\left[q_{-\hbar 0‰}\times n_{-\hbar}\right], \qquad (21)$$

$$\mathbf{Y}_{f\cdot 0‰}\left[m_1\times n_{f\cdot}\right]=\mathbf{A}_{f\cdot 0‰}\left[m_1\times q_{f\cdot 0‰}\right]\cdot\mathbf{F}_{f\cdot 0‰}\left[q_{f\cdot 0‰}\times n_{f\cdot}\right]. \qquad (22)$$

Используя преобразование ортогонального вращения, из уравнений (19) и (21) запишем преобразование осей $\mathbf{A}_{f\cdot r‰}\rightarrow\mathbf{A}^{-}_{f\cdot 0‰}$:

$$\mathbf{A}_{f\cdot r‰}\left[m_1\times q_{f\cdot r‰}\right]=\mathbf{A}_{f\cdot 0‰}\left[m\times q_{f\cdot 0‰}\right]\cdot\mathbf{T}\left[q_{f\cdot 0‰}\times q_{f\cdot r‰}\right]. \qquad (23)$$

Запись уравнения (23) показывает, что матрица $\mathbf{A}_{ţ'\text{t}\%_о}$ может быть выражена через матрицу $\mathbf{A}_{ţ'0\%_о}$: $\mathbf{A}_{ţ'\text{t}\%_о}(\mathbf{A}_{ţ'0\%_о})$. Умножим (23) слева на транспонированную матрицу $\mathbf{A}_{ţ'0\%_о}'\left[q_{ţ'0\%_о} \times m_1\right]$, имеем:

$$\begin{aligned}&\mathbf{A}_{\text{б10в}}'\left[q_{\text{б10в}} \times m_1\right] \cdot \mathbf{A}_{\text{бив}}\left[m_1 \times q_{\text{бив}}\right] = \\ &= \mathbf{A}_{\text{б10в}}'\left[q_{\text{б10в}} \times m_1\right] \cdot \mathbf{A}_{\text{б10в}}\left[m_1 \times q_{\text{б10в}}\right] \cdot \mathbf{T}\left[q_{\text{б10}} \times q_{\text{бив}}\right]\end{aligned}$$

Произведение $\mathbf{A}_{ţ'0\%_о}'\left[q_{ţ'0\%_о} \times m_1\right] \cdot \mathbf{A}_{ţ'0\%_о}\left[m_1 \times q_{ţ'0\%_о}\right]$ представляет собой малый момент ортогональных матриц компонентных нагрузок [9], и поэтому последнее уравнение можно переписать в виде:

$$\mathbf{A}_{\text{б10в}}'\left[q_{\text{б10в}} \times m_1\right] \cdot \mathbf{A}_{\text{б10в}}\left[m_1 \times q_{\text{б10в}}\right] = \Lambda_{\text{б10в}} \cdot \mathbf{T}\left[q_{\text{б10в}} \times q_{\text{бtв}}\right],$$

откуда находим алгоритм вычисления матрицы $\mathbf{T}$, характеризующей чисто временное вращение осей:

$$\mathbf{T}\left[q_{\text{б10в}} \times q_{\text{бив}}\right] = \Lambda_{\text{б10в}}^{-1} \cdot \mathbf{A}_{\text{б10в}}'\left[q_{\text{б10в}} \times m_1\right] \cdot \mathbf{A}_{\text{бив}}\left[m_1 \times q_{\text{бив}}\right]. \quad (24)$$

Повернув на те же углы матрицу значений компонент $\mathbf{F}_{\text{б10в}}$, получим прогнозные значения матрицы $\mathbf{F}_{\text{бtв}}$, выраженные через начальные значения:

$$\mathbf{F}_{\text{бив}}(\mathbf{A}_{\text{бив}})\left[q_{\text{бив}} \times n_{\text{бт}}\right] = \mathbf{T}'\left[q_{\text{бив}} \times q_{\text{б10в}}\right] \cdot \mathbf{F}_{\text{б10в}}'\left[q_{\text{б10в}} \times n_{\text{бт}}\right]. \quad (25)$$

Перемножив (23) и (25), получим аналог уравнения (19):

$$\begin{aligned}&\mathbf{Y}_{\text{бив}}(\mathbf{A}_{\text{б10в}})\left[m_1 \times n_{\text{бт}}\right] == \mathbf{A}_{\text{бtв}}(\mathbf{A}_{\text{б10в}})\left[m_1 \times q_{\text{б10в}}\right] \times \\ &\times \mathbf{F}_{\text{бив}}(\mathbf{A}_{\text{б10в}})\left[q_{\text{б10в}} \times n_{\text{бт}}\right] = \mathbf{A}_{\text{б10в}}\left[m_1 \times q_{\text{б10в}}\right] \cdot \mathbf{T}\left[q_{\text{б10в}} \times q_{\text{бtв}}\right] \times \\ &\times \mathbf{T}'\left[q_{\text{бtв}} \times q_{\text{б10в}}\right] \cdot \mathbf{F}_{\text{б10в}}\left[q_{\text{б10в}} \times n_{\text{бт}}\right] = \\ &= \mathbf{A}_{\text{бив}}\left[m_1 \times q_{\text{бив}}\right] \cdot \mathbf{T}'\left[q_{\text{бtв}} \times q_{\text{б10в}}\right] \cdot \mathbf{F}_{\text{б10в}}\left[q_{\text{б10в}} \times n_{\text{бт}}\right]\end{aligned} \quad (26)$$

Таким образом, чтобы получить нормированные прогнозные значения исходных признаков через начальные, можно оставить старые оси, но умножить слева на матрицу временного вращения $\mathbf{T}\left[q_{ţ'0\%_о} \times q_{ţ'\text{t}\%_о}\right]$матрицу прогнозных значений компонент в центрах блоков $\cdot\, \mathbf{F}_{ţ'\text{t}\%_о}\left[q_{ţ'\text{t}\%_о} \times n_{ţ'}\right]$, или оставить старые значения компонент, но изменить оси – умножить $\mathbf{A}_{ţ'\text{t}\%_о}\left[m_1 \times q_{ţ'\text{t}\%_о}\right]$ на транспонированную матрицу временного вращения справа.

Из уравнений (23) и (21) можно найти матрицу $\mathbf{C}\left[q_{ţ\cdot 0}\times q_{ţ\cdot 0‰}\right]$, восстанавливающую число признаков. Ее удобнее искать из вращения $\mathbf{F}_{ţ\cdot 0}\left[q_{ţ\cdot 0}\times n_{ţ\cdot}\right]\rightarrow \mathbf{F}_{ţ\cdot 0‰}\left[q_{ţ\cdot 0‰}\times n_{ţ\cdot}\right]$:

$$\mathbf{F}_{ţ\cdot 0}\left[q_{ţ\cdot 0}\times n_{ţ\cdot}\right]=\mathbf{C}\left[q_{ţ\cdot 0}\times q_{ţ\cdot 0‰}\right]\mathbf{F}_{ţ\cdot 0‰}\left[q_{ţ\cdot 0‰}\times n_{ţ\cdot}\right], \qquad (27)$$

откуда, используя свойства малого момента ортогональной матрицы значений компонент $\mathbf{F}_{ţ\cdot 0‰}\left[q_{ţ\cdot 0‰}\times n_{ţ\cdot}\right]$ [9], имеем:

$$\mathbf{C}\left[q_{ţ\cdot 0}\times q_{ţ\cdot 0‰}\right]=n_{ţ}^{-1}\cdot \mathbf{F}_{ţ\cdot 0}\left[q_{ţ\cdot 0}\times n_{ţ\cdot}\right]\cdot \mathbf{F}_{ţ\cdot 0‰}'\left[n_{ţ\cdot}\times q_{ţ\cdot 0‰}\right]. \qquad (28)$$

Выражение (28) – алгоритм вычисления матрицы восполнения числа признаков. Повернув на эти же углы матрицу $\mathbf{A}_{\bar{o}\imath 0}$, получим $\mathbf{A}_{\bar{o}\imath 0 в}$ с восстановленным числом признаков:

$$\mathbf{A}_{ţ\cdot 0‰}\left(\mathbf{A}_{ţ\cdot 0}\right)\left[m\times q_{ţ\cdot 0‰}\right]=\mathbf{A}_{ţ\cdot 0}\left[m\times q_{ţ\cdot 0}\right]\cdot \mathbf{C}'\ \left[q_{ţ\cdot 0‰}\times q_{ţ\cdot 0}\right]. \qquad (29)$$

Теперь $\mathbf{Y}_{\bar{o}\imath 0 в}$ с восстановленным числом признаков запишется в виде:

$$\begin{aligned}
&\mathbf{Y}_{\bar{o}\imath 0 в}\left(\mathbf{A}_{\bar{o}\imath 0}\right)\left[m\times n_{\bar{o}\imath}\right]=\mathbf{A}_{\bar{o}\imath 0}\left(\mathbf{A}_{\bar{o}\imath 0}\right)\left[m\times q_{\bar{o}\imath 0}\right]\times\\
&\times \mathbf{F}_{\bar{o}\imath 0 в}\left(\mathbf{A}_{\bar{o}\imath 0}\right)\left[q_{\bar{o}\imath 0}\times n_{\bar{o}\imath}\right]=\\
&=\mathbf{A}_{\bar{o}\imath 0}\left[m\times q_{\bar{o}\imath 0}\right]\cdot \mathbf{C}'\left[q_{\bar{o}\imath 0}\times q_{\bar{o}\imath 0 в}\right]\cdot \mathbf{C}\left[q_{\bar{o}\imath 0 в}\times q_{\bar{o}\imath 0}\right]\times\\
&\times \mathbf{F}_{\bar{o}\imath 0 в}\left[q_{\bar{o}\imath 0 в}\times n_{\bar{o}\imath}\right]=\\
&=\mathbf{A}_{\bar{o}\imath 0}\left[m\times q_{\bar{o}\imath 0}\right]\cdot \mathbf{C}\left[q_{\bar{o}\imath 0}\times q_{\bar{o}\imath 0 в}\right]\cdot \mathbf{F}_{\bar{o}\imath 0 в}\left[q_{\bar{o}\imath 0 в}\times n_{\bar{o}\imath}\right]
\end{aligned} \qquad (30)$$

С учетом (26) и (30) можно записать прогнозные значения в центрах блоков с восстановленным набором признаков:

$$\begin{aligned}
&\mathbf{Y}_{\bar{o}\imath п}\left[m\times n_{\bar{o}\imath}\right]=\mathbf{A}_{\bar{o}\imath 0}\left[m\times q_{\bar{o}\imath 0}\right]\cdot \mathbf{C}\left[q_{\bar{o}\imath 0}\times q_{\bar{o}\imath 0 в}\right]\times\\
&\times \mathbf{T}\left[q_{\bar{o}\imath 0 в}\times q_{\bar{o}\imath п в}\right]\cdot \mathbf{F}_{\bar{o}\imath п в}\left[q_{\bar{o}\imath п в}\times n_{\bar{o}\imath}\right]
\end{aligned} \qquad (31)$$

Это выражение – алгоритм вычисления прогнозной матрицы в центрах блоков с полным набором признаков, т.е. решена задача восстановления числа параметров в прогнозной матрице для центров блоков.

Аналогичная задача решается и для значений в реальных точках: из (20) и (19) имеем:

$$\mathbf{A}_{-ђђ}\ \left[m_1\times q_{ţ\cdot 0‰}\right]=\mathbf{A}_{-ђ0‰}\left[m_1\times q_{-ђ0‰}\right]\cdot \mathbf{e}\ \ \left[q_{-ђ0‰}\times q_{ţ\cdot 0‰}^{\cdot}\right], \qquad (32)$$

откуда находим матрицу $\mathbf{B}$, учитывающую переход от центров блоков сетки к точкам наблюдения (скважинам) и временное изменение осей (ход решения аналогичен вычислению матрицы временного вращения):

$$\mathbf{B}\left[q_{ск0в} \times q_{бпв}\right] = \mathbf{\Lambda}^{-1}_{ск0в} \cdot \mathbf{A}_{ск0в}{}'\left[q_{ск0в} \times m_1\right] \times \\ \times \mathbf{A}_{ск0в}\left[m_1 \times q_{ск0в}\right] \tag{33}$$

Повернув на те же углы матрицу $\mathbf{F}_{блв}$, получим значения компонент в блоках, выраженные через значения в скважинах на начальный момент времени, но с сокращенным набором признаков:

$$\mathbf{F}_{-ббв}\left(\mathbf{A}_{-б0\%в}\right)\left[q_{ft\%в} \times n_{-б}\right] = \mathbf{e}'\left[q_{ft\%в} \times q_{-б0\%в}\right] \cdot \mathbf{F}_{-б0\%в}\left[q_{-б0\%в} \times n_{-б}\right] \tag{34}$$

Перемножим $\mathbf{F}_{бпв}$ и $\mathbf{A}_{бпв}$, записанные в виде (31) и (32), получаем нормированные прогнозные значения сокращенного набора признаков через начальные значения в скважинах:

$$\mathbf{Y}_{ft}\left(\mathbf{A}_{-б0\%в}\right)\left[m_1 \times n_{-б}\right] = \mathbf{A}_{-б0\%в}\left[m_1 \times q_{-б0\%в}\right] \cdot \mathbf{e}\left[q_{-б0\%в} \times q_{ft\%в}\right] \times \\ \times \mathbf{B}'\left[q_{ft\%в} \times q_{-б0\%в}\right] \cdot \mathbf{F}_{-б0\%в}\left[q_{-б0\%в} \times n_{-б}\right] \tag{35}$$

Теперь необходимо восстановить полный набор признаков в прогнозной матрице для реальных объектов. Матрицу восстановления числа признаков удобнее искать из вращения $\mathbf{F}_{-б0\%в} \rightarrow \mathbf{F}_{-б0}$:

$$\mathbf{F}_{-б0\%в}\left[q_{-б0\%в} \times n_{-б}\right] = \mathbf{D}\left[q_{-б0\%в} \times q_{-б0}\right] \cdot \mathbf{F}_{-б0}\left[q_{-б0} \times n_{-б}\right], \tag{36}$$

откуда по аналогии с решением для $\mathbf{C}$ находим:

$$\mathbf{D}\left[q_{-б0\%в} \times q_{-б0}\right] = n_1^{-1} \cdot \mathbf{F}_{-б0\%в}\left[q_{-б0\%в} \times n_{-б}\right] \cdot \mathbf{F}_{-б0}{}'\left[n_{-б} \times q_{-б0}\right]. \tag{37}$$

Умножив на $\mathbf{D}'$ справа матрицу $\mathbf{A}_{ск0}$, получим с учетом (36) выражение для $\mathbf{A}_{ск0в}$ с восстановленным признаков числом:

$$\mathbf{A}_{б0\%в}\left(\mathbf{A}_{-б0}\right)\left[m \times q_{-б0\%в}\right] = \mathbf{A}_{-б0}\left[m \times q_{-б0}\right] \cdot \mathbf{D}'\left[q_{-б0} \times q_{-б0\%в}\right] \tag{38}$$

Если в уравнении (34) $\mathbf{A}_{-б0\%в}$ выразить через $\mathbf{A}_{-б0\%в}\left(\mathbf{A}_{-б0}\right)$, то получим решение для вычисления прогнозных значений всех признаков в скважинах:

$$\begin{aligned}&\mathbf{Y}_{скпр}\left[m \times n_{ск}\right] = \mathbf{Y}_{бпв}\left(\mathbf{A}_{ск0в}\left(\mathbf{A}_{ск0}\right)\right) = \\ &= \mathbf{A}_{ск0}\left[m \times q_{ск0}\right] \cdot \mathbf{D}'\left[q_{ск0} \times q_{ск0в}\right] \times \\ &\times \mathbf{B}\left[q_{ск0в} \times q_{бпв}\right] \cdot \mathbf{B}'\left[q_{бпв} \times q_{ск0в}\right] \cdot \mathbf{F}_{ск0в}\left[q_{ск0в} \times n_{ск}\right]\end{aligned} \qquad (39)$$

Уравнение (39) позволяет с учетом (32) и (37) вычислить значения полного набора прогнозных признаков для реальных объектов. Возвращение к исходным единицам измерения возможно при условии задания законов изменения средних значений и дисперсий всех исходных признаков. Поиск окончательного варианта еще не завершен. Пока же $\bar{x}_i$ и σ_i для имеющихся прогнозных признаков взяты из прогнозной матрицы, а для пропущенных – из матрицы начального состояния. Здесь желательно использовать разработанные вероятностные методы оценки смещения среднего значения и дисперсии.

С точки зрения практического применения информационных технологий наиболее перспективен мониторинг окружающей среды. Геоэкологический мониторинг представляет собой систему наблюдения, изучения и контроля для управления состоянием окружающей среды [23]. Поскольку окружающая среда – многокомпонентная система, характеризуемая огромным количеством параметров и их взаимосвязей, их изучение контролируется ограниченными финансовыми вложениями в систему мониторинга. Комплексный подход к изучению этой системы реализуется рядом принципов экологического мониторинга. В качестве основного международный опыт изучения экогеосистем разного иерархического уровня рекомендует принцип проблемной организации, когда на короткое время разворачивается широкая программа наблюдений под определенную проблему. После детального изучения этой проблемы, заключающегося в выявлении взаимосвязей наблюдаемых параметров, программа сворачивается, а наблюдения продолжаются по сокращенной сети и/или ограниченному набору параметров [23].

Такая постановка справедлива для сохраняющейся структуры взаимосвязей, однако, известно, что техногенная эволюция геоэкосистем по своей скоротечности на многие порядки превышает естественную, а потому является не просто системным процессом, но системно-динамическим. Мало того, современная теория техногенной эволюции геоэкосистем должна иметь в своем методологическом арсенале возможности прогнозирования, предвидения хода явлений-возможно никогда ранее не существовавших в естественных условиях. Традиционно результаты мониторинга используются для создания моделей сложных объектов и процессов. С другой стороны, проблемная постановка геоэкологического мониторинга обязательно предполагает использование моделей сложных систем и оптимизационных моделей.

Остановимся кратко на формальной постановке задачи реализации принципа проблемной организации экологического мониторинга, которая может служить задачей оптимизации режимной сети. Предположим, мы имеем одну полную исходную матрицу данных $\mathbf{Y}_1[m \times n_1]$ и неполную матрицу на другой момент(t_2) времени $\mathbf{Y}_2[m \times n_2]$. Задачей ставим нахождение полной матрицы исходных данных на другой момент времени $\mathbf{Y}_{2n}[m \times n_1]$. Здесь m – число признаков, приписываемых каждому объекту; n_1 – число объектов в полной матрице исходных данных; n_2 – число объектов в неполной матрице исходных данных на момент времени t_2

Из полной матрицы $\mathbf{Y}_1[m \times n_1]$ на момент t_1 составляется выборочная с полным набором признаков, но сокращенным количеством объектов, совпадающих с объектами на момент времени t_2, которые описываются матрицей $\mathbf{Y}_2[m \times n_2]$ – наборы признаков и объектов на момент t_2 совпадают с таковыми на момент времени t_1 для выборочной матрицы. Далее применяется обычная технология построения системы матричных уравнений и вычисления матриц вращения компонентных осей и $\mathbf{Y}_{2n}[m \times n_1]$.

Принципиальная возможность оптимизации мониторинговой сети на основе установленного пространственно-временного изменения внутрисистемных взаимосвязей показана на восстановлении полного набора данных гидрогеохимического опробования для разряженной более чем вдвое сети скважин мониторинга угольного разреза "Восточный" (Экибастуз).

Модель компонентного анализа, положенная в основу разработанной комплексной модели, по своей сути ориентирована на выявление групповых взаимосвязей, фиксирующих состояние объекта исследований. Процедура временного вращения позволяет оценить динамику этих взаимосвязей. Таким образом, разработанная технология построения системы моделей решает задачи, сходные с задачами информационной технологии на основе системологики П.К. Анохина, не выделяя явно базисные подсистемы. Сложный вопрос учета нелинейности взаимосвязей снимается многомерностью задачи (фактически, используются их линейные разложения).

Разработанную методику можно применять при построении системы мониторинга, моделей управления и охраны качеством среды региона, а также при выполнении ОВОС, экологического аудита, в градостроительной практике и пр. Главными проблемами в применении методики является обоснование изменений средних значений и дисперсий прогнозных признаков и жесткая привязанность временным шагам детерминированной модели.

Основными выводами по работе являются следующие:

- разработана методика для оценки состояния и прогноза поведения сложных природных объектов. Она базируется на многомерных статистических описательных моделях, объективно фиксирующих состояние объектов на определенный момент времени. Детерминированная составляющая, в первую очередь фактор времени и/или закономерности территориального распределения, учитывается построением новых алгоритмов, основанных на вращении многомерных компонентных осей. При этом модели, описывающие взаимосвязи на разные временные или разные качественные уровни, объединяются в систему матричных уравнений, а вращение осей осуществляет их многоуровневую связь;
- разработанная комплексная модель решает задачи, сходные с задачами информационной технологии на основе системологики П.К. Анохина, не выделяя явно базисные подсистемы. Сложный вопрос учета нелинейности взаимосвязей снимается многомерностью задачи (фактически, используются их линейные разложения). Использование в качестве фактического материала результатов детерминированных моделей, новых граничных условий и большого количества объективных косвенных показателей, свидетельствующих о происходящих процессах, повышает степень адекватности модели реальной геоэкосистеме;
- анализ результатов моделирования не требует специальной математической подготовки, т.к. они представлены в тех же единицах измерения и в том же наборе параметров, что и исходные. Оценка точности модели осуществляется автоматически проверкой качества восстановления известных характеристик объекта. Разработанную методику можно применять для исследования при построении и оптимизации системы экологического мониторинга, моделей управления и охраны сложных объектов региона, а также при выполнении ОВОС гидротехнических сооружений, в градостроительной практике и пр.;
- разработанная модель позволяет реализовать принцип организации экологического мониторинга принцип проблемной организации, когда на короткое время разворачивается широкая программа наблюдений под определенную проблему. После детального изучения этой проблемы, заключающегося в выявлении взаимосвязей наблюдаемых параметров, программа сворачивается, а наблюдения продолжаются по сокращенной сети и/или ограниченному набору параметров. Предложенный метод не уступает методам теории распознавания образов по возможности восстановления количества информации. Дополнительно к возможностям этой теории восстанавливает объекты целиком (по полному набору параметров) в условиях изменяющегося "обучающего образа".

Литература

1. Айвазян С.А. Модельно- и методоориентированные интеллектуализированные программные комплексы по статистическому анализу данных // Многомерный статистический анализ и вероятностное моделирование реальных процессов. – М.: Наука, 1990. – С. 6-30.
2. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. – М.: Финансы и статистика, 1983. – 471 с.
3. Алгоритмы и программы восстановления зависимостей / В.Н. Вапник, Т.Г. Гладкова, В.А. Кащеев и др. – М.: Наука, 1984. – 815 с.
4. Алексеев В.С., Коммунар Г.М., Шержуков Б.С. Массоперенос в водонасыщенных горных породах // Итоги науки и техники. Сер.Гидрогеология. Инженерная геология. –М.: ВИНИТИ. 1989. -Т.11. –143 с.
5. Андерсон Т. Статистический анализ временных рядов. – М.: Мир, 1976. – 755 с.
6. Джефферс Дж. Введение в системный анализ: применение в экологии. – М.: Мир. 1981. – 213 с.
7. Зальцберг Э.А., Деч В.Н. Исследование и прогноз режима подземных вод с помощью многомерного спектрального и компонентного анализов // Разведка и охрана недр. – 1979. – № 6. – С. 40-46.
8. Ивахненко А.Г. Долгосрочное прогнозирование и управление сложными объектами. – Киев: Техника, 1975. – 311 с.
9. Йёреског К.Г., Клован Д.И., Реймент Р.А. Геологический факторный анализ. – М.: Недра, 1980. – 223 с.
10. Павличенко Л.М. Исследование региональных гидрогеохимических процессов на основе компонентного анализа. – Дис. ... канд. техн. наук. – Алма-Ата, 1984. – 189 с.
11. Павличенко Л.М. Система многомерных статистических моделей анализа неполных эколого-гидрогеологических данных // Вопросы изучения водных ресурсов Центральной Азии. – Алматы: Гылым, 1993. – С. 89-103.
12. Павличенко Л.М. Контроль процесса загрязнения подземных вод бором на основе метода стандартной матрицы нагрузок // Мат. III Международного симпозиума "Освоение месторождений минеральных ресурсов и подземное строительство в сложных гидрогеологических условиях". – Белгород, ВИОГЕМ, 1995. – С. 141-147.
13. Павличенко Л.М. Прогноз изменения минерализации подземных вод в результате воздействия ЮК ГРЭС на основе компонентной модели // Географические основы устойчивого развития Республики Казахстан. – Алматы: Гылым, 1998. – С. 307-311.
14. Павличенко Л.М., Веселов В.В. Системный подход к построению моделей сложных природных процессов // Вестник КазГУ. Серия географическая. – 1997. – № 4. – С. 25-29.
15. Павличенко Л.М., Шапиро С.М. Применение методов многомерной статистики для прогноза процессов загрязнения подземных вод на основе оценки изменения их химического состава // Теоретические основы и методика гидрогеологического прогноза загрязнения под-

земных вод (Современные проблемы биосферы). – М.: Наука, 1990. – С. 137-141.
16. Павличенко Л.М., Шапиро С.М., Джумагулов М.Т. Техногенные гидрогеологические процессы в пределах городских агломераций и их экологическое значение // Экологические проблемы Казахстана (Доклады республиканского совещания. Алма-Ата, 1991). – Алма-Ата: ООП РИИЦ Госкомстата КазССР, 1991. – С. 150-159.
17. Перегудов Ф.И., Тарасенко Ф.П. Введение в системный анализ. – М.: Наука, 1989. – 477 с.
18. Подольный О.В. Подземные воды как компонент экосистем аридной зоны в условиях техногенеза. Автореф. дис. ... доктор геол.-минер. наук. – Ташкент, 1991. – 44 с.
19. Попов Ю.М., Павличенко Л.М., Богачев В.П. Исследование загрязненности реки Сырдарьи для построения комплексной оценки качества воды // Гидрология и экология. – 1996. – № 2. – С.207-223.
20. Сочава В.Б. Введение в учение о геосистемах. – Новосибирск: Наука, 1978. – 398 с.
21. Чепурных Н.В., Новоселов А.Л. Планирование и прогнозирование природопользования:Учебное пособие.–М.:Интерпракс,1995.–288 с.
22. Четвериков Н.С. О ложной корреляции // Статистические исследования (теория и практика). – М.: Наука, 1975. – С. 298-318.
23. Экоинформатика: Теория. Практика. Методы и системы / Под ред. В.Е. Соколова. – СПБ.: Гидрометеоиздат, 1992. – 495 с.

Казахский государственный национальный университет им.аль-Фараби
Институт гидрогеологии и гидрофизики им.У.М.Ахмедсафина МОН РК

ЭКОГЕОЖҮЙЕЛЕРІНІҢ ӨЗГЕРУІН БОЛЖАМДАУ МОДЕЛЬДЕРІНІҢ ТЕХНОЛОГИЯСЫН ЖАСАУ МӘСЕЛЕСІНЕ

Ғыл.техн.канд. Л.М.Павличенко

Экогеожүйелерінің параметрлеріне сай кең шеңберде алғашқы берілгендерін және динамикасын есепке алатын көп өлшемді ось айналым тәсілін пайдалану ұсынылған. Әртүрлі уақыт аралығында, бірбірімен байланыстылығы, әртүрлі сапалы деңгейлері матриціі тендеулер жүйесіне топтасады, ал ось айналымы көп деңгейлі байланысты орындыйды. Экологиялық мониторингті мәселелі ұйымдастыру принциптеріне сай, алғашқы берілген информацияларды толықтыру модельдерін құру технологиясы қарастырылған.