ЭКОНОМИКО-СТАТИСТИЧЕСКИЕ МОДЕЛИ В ПРОГНОЗИРОВАНИИ И ПЛАНИРОВАНИИ ПРОМЫШЛЕННОГО ПРОИЗВОДСТВА

СТАТИСТИЧЕСКИЕ
МЕТОДЫ
В ЭКОНОМИЧЕСКИХ
ИССЛЕДОВАНИЯХ

ЭКОНОМИКО-СТАТИСТИЧЕСКИЕ МОДЕЛИ В ПРОГНОЗИРОВАНИИ И ПЛАНИРОВАНИИ ПРОМЫШЛЕННОГО ПРОИЗВОДСТВА

Ответственные редакторы
канд. экон. наук *М. Л. Лукацкая*,
д-р экон. наук *Б. Б. Розин*

ИЗДАТЕЛЬСТВО «НАУКА»
СИБИРСКОЕ ОТДЕЛЕНИЕ
Новосибирск·1978

В сборнике освещаются методические и прикладные вопросы как собственно статистического моделирования, так и включения экономико-статистических моделей в общую систему моделей территориально-производственного планирования. Рассматриваются проблемы прогнозирования вероятностных характеристик управляемых экономических процессов, использования статистических процедур и методов теории распознавания образов в постановке и анализе решений оптимизационных задач, а также при агрегировании экономической информации, применения ЭСМ для экономической оценки природных факторов в добывающих отраслях.

Книга предназначена для работников НИИ, плановых органов и предприятий, для студентов и аспирантов экономических специальностей.

Э $\frac{10807\text{-}811}{042(02)\text{-}78}$ 325-78 © Издательство „Наука“, 1978.

ПРЕДИСЛОВИЕ

Настоящий сборник включает в себя статьи по проблемам экономико-статистического моделирования. В нем получили отражение методологические, методические и прикладные аспекты продолжающихся исследований по вопросам как собственно статистического моделирования экономических процессов, так и включения экономико-статистических моделей (ЭСМ) в общую систему моделей территориально-производственного планирования.

Первая статья сборника посвящена общетеоретическим проблемам экономико-статистического анализа, в ней дается методическое обобщение проблем моделирования статистических зависимостей в неоднородных совокупностях. В двух статьях И. Г. Багирова рассмотрены постановка и пути реализации задачи описания и прогнозирования вероятностных характеристик управляемых экономических процессов. Во второй статье общий подход конкретизируется применительно к задаче определения потребности городов в газе.

Несколько статей (Л. Я. Шкрабина, В. Г. Ионина, Б. Б. Розина и М. А. Ягольницера) посвящены проблемам использования статистических процедур и методов теории распознавания образов в постановке и анализе решений оптимизационных задач.

В статьях по вопросам агрегирования экономической информации на основе статистических методов и методов теории распознавания образов (Н. М. Журавель, В. Г. Ионина и Н. П. Иониной, Н. Г. Шишацкого) обсуждается критерий агрегирования, описывается возможная

процедура оптимальной классификации, приводятся некоторые результаты расчетов по лесопромышленному комплексу. В работе Н. Г. Шишацкого предложены способы использования экономико-статистических моделей предприятий в качестве инструмента агрегирования информации в процедурах согласования решений в двухуровневой системе.

Часть работ посвящена методике решения некоторых типичных экономических задач. В частности, в них рассматривается проблема использования ЭСМ для экономической оценки природных факторов в угольной (Н. Е. Ачелашвили) и железорудной промышленности (Г. Т. Берсеневой, Г. П. Бондаренко и Н. В. Овчинниковой); исследуется динамика нефтедобычи (Л. А. Сергеевой, Г. Л. Чудновского, М. А. Ягольницера; Г. Т. Берсеневой, Т. Р. Сазоновой, Л. А. Сергеевой и М. А. Ягольницера) и т. д.

В двух статьях сборника рассмотрены своеобразные области применения методов многомерной классификации. И. Д. Мандель и А. Н. Никитин предлагают использовать их для классификации разных определений одного и того же явления с целью выбора наиболее употребительного. Г. П. Гвоздевой, Е. Е. Горяченко, И. Б. Мучником и Е. Э. Синицыным предпринята попытка использования достаточно общей методики классификации объектов для построения выборки при проведении экономико-статистических исследований, в частности для стратификации совокупности, формирования и тестирования выборки, оценивания параметров, полученных в результате выборочного обследования.

Б. Б. РОЗИН

ПРИРОДА НЕОДНОРОДНОСТИ СТАТИСТИЧЕСКИХ СОВОКУПНОСТЕЙ И МЕТОДЫ ЕЕ УЧЕТА ПРИ МОДЕЛИРОВАНИИ

Одна из важнейших проблем статистического анализа вообще и статистического моделирования в частности заключается в структурной неоднородности изучаемых совокупностей. С этой проблемой статистики сталкиваются при решении самых разнообразных задач обработки информации (расчет различного рода средних показателей, изучение связей методами корреляционно-регрессионного анализа и др.). Она относится к числу самых дискуссионных проблем в статистической литературе. Существует множество подходов к анализу и оценке степени однородности исследуемой совокупности. Однако до сих пор не определено само понятие однородности, а также способы ее содержательного, качественного анализа, не установлены общепризнанные количественные, формальные критерии оценки степени однородности совокупности.

В то же время большинство исследователей признают, что экономические совокупности, как правило, обладают структурой и это обстоятельство должно учитываться в различных областях статистического анализа. В настоящей статье сделана попытка рассмотреть названную проблему применительно к задаче моделирования статистических связей в совокупности промышленных объектов отрасли. Нам представляется, что найденные при решении этого класса прикладных задач методические подходы имеют достаточно общий характер.

ВНУТРЕННЯЯ ПРОТИВОРЕЧИВОСТЬ СТАТИСТИЧЕСКИХ СОВОКУПНОСТЕЙ. ОТНОСИТЕЛЬНОСТЬ ПОНЯТИЯ ОДНОРОДНОСТИ

Ставя задачу экономико-статистического моделирования применительно к неоднородным совокупностям, необходимо прежде всего кратко охарактеризовать специфические черты самого предмета изучения. Статистическая совокупность — это множество объективно существующих в пространстве и во времени явлений, однокачественных в определенной связи [1]. Экономическим явлениям (объектам) присуще множество свойств и измеряющих их признаков. Поэтому понятие однокачественности в зависимости от конкретных целей и условий исследования может рассматриваться в различных аспектах. Поскольку экономические явления обладают множеством свойств и разнообразием форм и стадий развития, одна и та же совокупность явлений может быть однородной в одном отношении и разнородной в другом.

При экономико-статистическом исследовании важно учитывать взаимоотношение совокупности и образующих ее индивидуальных явлений. Хотя статистическая совокупность формируется из однокачественных объектов, ее отдельные единицы значительно различаются по своим признакам. Эти различия, как правило, далеко выходят за рамки индивидуального варьирования. Следовательно, статистическая совокупность по своей природе сложна и противоречива, она выступает и как множество однотипных явлений, и как множество варьирующих, изменяющихся явлений. В процессе развития такая изменчивость ведет к выделению в пределах совокупности отдельных частей, представляющих разные стадии, этапы общего процесса развития. Следует считать твердо установленным фактом подразделение статистической совокупности на своеобразные частные подсовокупности (типы) явлений [2].

Итак, общая статистическая совокупность как предмет экономико-статистического исследования состоит из нескольких частных подсовокупностей явлений с присущими каждой из них специфическими закономерностями. Следовательно, без выделения типов явлений (без типологической группировки) нельзя достаточно точно изучить и описать статистические связи и закономерности, объективно существующие в общей совокупности.

Необходимым условием выделения однородных типов явлений (объектов) является конкретизация понятия однородности. При этом необходимо учитывать, что всякое определение условно, относительно. За последнее десятилетие предложено немало определений понятия однородности. Однако общепризнанного, устоявшегося определения нет. Большинство определений, имеющихся в литературе, в том числе и приводимое в данной статье, следует трактовать как «рабочие», «условные». Эта проблема весьма сложна, и в ней, как, может быть, ни в какой другой задаче экономико-статистического исследования, важно соблюдать принцип примата содержательного анализа над формальными построениями.

По нашему мнению, лучше говорить не об однородности вообще, а о степени, уровне этой однородности. В экономико-статистическом исследовании *мы оперируем с понятием относительной однородности*, и речь может идти лишь о достижении разумной степени однородности совокупности, такой, при которой обеспечивалась бы достаточная для данного исследования точность статистических выводов. В качестве показателя степени однородности может выступать показатель «близости», «похожести» объектов совокупности по признакам их описания. Весь комплекс признаков, описывающих единицы наблюдения, можно разделить на два вида: *общие* (близкие по значению), определяющие сходство, однородность изучаемых единиц наблюдения, и *варьирующие*, влияние которых оценивается в результате статистического исследования. При этом степень однородности характеризуется не столько долей общих признаков во всем их комплексе, сколько их ролью в формировании характерных свойств объектов.

В зависимости от назначения и стадии исследования следует различать *широкое* и *узкое* понятия однородности. Первое предназначено для выделения исходной совокупности из всей массы экономических явлений и используется на стадии постановки задачи исследования, второе — для выделения однородных групп (типов) объектов уже после сбора информации, на стадии собственно экономико-статистического моделирования. Существует иерархия этих двух понятий: широкое понятие — более высокого ранга и составляет необходимое, но недостаточное для формирования однородных в узком смысле этого слова типов объектов условие.

Широкое понятие однородности можно ассоциировать с общеупотребительным в литературе по статистике термином однокачественности элементов экономической совокупности, наличием у них общности в существенных, главных чертах и свойствах. Для выделения таких однокачественных исходных совокупностей, как правило, достаточно небольшого числа признаков. Это объясняется тем, что между признаками, входящими в описание объектов, существуют причинно-следственные связи: одни признаки, такие как, например, вид продукции, способ производства и т. п., определяют значение или диапазон значений других признаков и тем самым вносят большой вклад в обеспечение однокачественности элементов совокупности.

Для выбора признаков на первой, постановочной стадии исследования весьма полезен причинный анализ структуры их связей [3]. На основе такого анализа можно уже при постановке задачи формировать относительно однородные совокупности объектов по ограниченному числу признаков. Вполне понятно, что подбор объектов, у которых совпадает большое число последовательно влияющих друг на друга признаков, обеспечивает высокую однородность совокупности, но резко уменьшает ее объем. Поэтому удовлетворительная для стадии постановки задачи степень однородности обеспечивается подбором объектов, у которых совпадают признаки, находящиеся на входе графа связи. Например, в структурных схемах связи факторов, обусловливающих вариацию показателей работы угольных шахт, входными являются горно-геологические условия, в значительной мере определяющие значения других варьирующих элементов (способ вскрытия, метод подготовки шахтного поля, система разработки угольных месторождений, способы механизации и т. д.). В этом случае достаточная начальная однородность совокупности обеспечивается группировкой шахт по горно-геологическим условиям — в некоторой степени она близка к группировке шахт по бассейнам.

Цель введения узкого понятия однородности заключается в выделении таких типов объектов, внутри которых взаимосвязь показателей и факторов производства стабильна и может быть описана простыми аппроксимирующими функциями (линейные или функции-константы). Уровень этого показателя однородности оценивается апо-

стериорно, после сбора информации на основе анализа многомерного распределения объектов исходной совокупности и выделения «вершин» («холмов») этого распределения. Последние трактуются как однородные в узком смысле этого слова группы объектов.

Следует заметить, что наиболее сложным и дискуссионным является именно вопрос о способах и критериях выделения таких однородных типов объектов в пределах однокачественной исходной совокупности. В последние годы он достаточно широко обсуждался в литературе, в частности на страницах журнала «Вестник статистики» [4, 5, 6]. Предложен ряд возможных критериев оценки однородности. Например, В. И. Сиськов [4] предлагает использовать в качестве такого критерия коэффициент вариации (V) и указывает, что для нормального распределения его критическая величина составляет 0,33, а для равномерного — 0,58 и т. д.

Следует согласиться с мнением В. И. Сиськова, что для любого теоретического закона распределения существует критическая величина коэффициентов вариации, которая может выступать в качестве границы между однородностью и неоднородностью. Однако предложенная им и другими авторами методика основана на двух достаточно сильных посылках: во-первых, априори известен теоретический закон распределения единиц совокупности и, во-вторых, исследуются преимущественно одномерные объекты. Обе эти посылки, как правило, не соблюдаются в прикладных исследованиях.

В. И. Сиськов проблему многомерности предлагает решать двумя путями: либо с помощью комбинационной группировки, в которой интервалы по отдельным признакам определены исходя из величины критерия V, либо методом последовательного уточнения состава групп по отдельным признакам. В обоих случаях набор учитываемых признаков при обычных объемах экономических совокупностей будет сравнительно невелик (3—5 признаков). Между тем развитие экономических объектов — процесс достаточно сложный, и его особенности не могут быть описаны несколькими показателями. Необходим многомерный анализ по значительному числу признаков.

Более радикальный, с нашей точки зрения, выход заключается в использовании непараметрических методов оценивания плотности вероятности и применении методов

распознавания образов (многомерной классификации) для формирования однородных групп явлений. Если априори постулируемый закон распределения не соответствует природе изучаемого явления, то оценивание непараметрическими методами по сравнению с параметрическими дает меньшую ошибку. Ориентировка на методы распознавания образов связана с тем, что этот аппарат позволяет выделять однородные группы одновременно по большому числу признаков.

Следует отметить, что вообще экономические совокупности имеют ряд общих отличий от природных (малочисленность наблюдений, неоднородность единиц и т. д.), что затрудняет применение классического аппарата математической статистики. Теория вероятностей и математическая статистика разрабатывались применительно к задачам исследования природных явлений, и ряд их постулатов и посылок (независимость и однородность единиц наблюдения, большие размеры исходных совокупностей и т. д.) не выполняется в реальных экономических задачах. Отдельные элементы природных совокупностей, например в области статистической физики, чрезвычайно однородны и даже могут быть отождествлены, что совершенно недопустимо в отношении элементов экономических совокупностей [1]. Использование методов распознавания образов в сочетании с традиционными статистическими методами позволяет в значительной мере преодолеть указанные трудности.

В наших исследованиях при построении плотности вероятности по выборке использовались непараметрические оценки Парзена [7] с экспоненциальными ядрами. Выбор «нормальных» ядер обусловлен ориентацией на дальнейшее использование при построении экономико-статистической модели методов кусочно-линейной аппроксимации. Иначе говоря, мы считаем подсовокупность однородной, если многомерное распределение моделируемого показателя и существенно влияющих на него факторов близко к нормальному или к суперпозиции нормальных распределений. Такую подсовокупность можно также назвать линейно-однородной, так как, согласно теории вероятности, если многомерные случайные величины имеют нормальное распределение, то теоретическая линия регрессии любой из этих величин относительно всех остальных является линейной функцией.

СТРУКТУРНАЯ НЕОДНОРОДНОСТЬ ОТРАСЛЕВЫХ СИСТЕМ

Совокупности промышленных объектов одной отрасли (подотрасли) обладают определенной структурой, т. е. общая совокупность предприятий распадается на ряд типов, имеющих свой, специфический характер зависимости экономических показателей от факторов производства.

Тип промышленного объекта — это *определенный диапазон значений* уровня структурных элементов производственного процесса (рабочая сила, орудия, предметы и продукты труда; технология и организация производства) и *определенная их комбинация*. В пределах одного типа объектов имеется значительная свобода вариации значений и комбинаций элементов, но только такая, которая не изменяет характера их воздействия на величину экономических показателей.

Это положение геометрически можно интерпретировать следующим образом. Если признаковое пространство представить в виде n-мерного многогранника, то точки, соответствующие объектам отрасли, будут занимать не весь его объем, а располагаться определенными сгущениями, между которыми имеются разрывы. Каждое такое четко выраженное сгущение можно ассоциировать с типом объектов.

Таким образом, предметом экономико-статистических исследований оказывается не просто совокупность предприятий определенной отраслевой системы, а совокупность, состоящая из нескольких типов предприятий. Этот факт оказывает большое влияние на характер связей экономических показателей и факторов производства. Изменение одного и того же фактора по-разному влияет на экономические показатели при различных сочетаниях структурных элементов производства, для разных типов организации производственного процесса. Так, например, увеличение расхода чугуна оказывает противоположное влияние на экономику мартеновского производства при скрап-процессе на твердом чугуне и скрап-рудном процессе на жидком чугуне. В первом случае расходы по переделу возрастают, во втором — снижаются при увеличении доли чугуна в шихте. Число подобных примеров можно было бы умножить. Более того, даже совокупность

наблюдений за функционированием одного объекта во времени обычно неоднородна, в ней можно обнаружить определенную структуру.

В целом ряде прикладных исследований [8, 9 и др.] показано, что значение и уровень важнейших экономических показателей работы предприятий определяются не только природными и организационно-техническими условиями, но и положением данного предприятия на траектории развития. Весьма укрупненно можно выделить три основных этапа развития объекта: первый этап — освоение предприятия, второй — стационарное функционирование, третий — период «старения». Следует подчеркнуть, что границы отдельных этапов — это не четкие разграничительные линии, а размытые, смазанные рубежи. Переход предприятия с одного этапа траектории развития на другой осуществляется плавно, постепенно. Однако каждому из трех выделенных этапов развития промышленного объекта соответствует своя динамика экономических показателей и свои специфические функции преобразования «затраты — выпуск».

Информационной основой экономико-статистического моделирования обычно служат сведения о функционировании совокупности предприятий на каком-то ограниченном отрезке времени $[t_0, t_n]$. Соответственно можно выделить две задачи анализа — базисной структуры совокупности и процесса ее изменения.

Деление совокупности исследуемых предприятий отрасли в начальный момент t_0 на отдельные типы объективно обусловлено рядом причин. Одна из них — неравномерность развития отдельных единиц совокупности во времени. В любой пространственной совокупности, принятой за базис, неявно присутствуют элементы динамики. В частности, объекты, входящие в нее, находятся на разных точках траектории своего развития, имеют различную предысторию. Предприятия, различающиеся по «возрасту», принадлежат к разным «поколениям». Каждому периоду создания предприятий, как правило, соответствует особая модификация исследуемого процесса, свой уровень развития техники и технологии производства. Объекты, построенные в один или близкие календарные периоды, как правило, сходны по уровню техники и технологии производства и находятся на примерно одинаковых точках траектории своего развития. И наоборот,

объекты разных «поколений» обычно существенно различаются между собой по характеру влияния факторов производства на экономические показатели.

Но и объекты одного «поколения» могут различаться по фактическому состоянию, уровню потенциальных возможностей. «Календарный возраст», конечно, не единственная и не основная характеристика состояния предприятия. Помимо него необходимо учитывать комплекс других факторов, приводящих к неоднородности предприятий одного «поколения». В качестве типообразующих признаков выступают: мощность объектов, назначение и характер продукции, в добывающих отраслях — природные условия. Различие этих признаков в конечном итоге приводит к нестабильности связей факторов и экономических показателей в пределах изучаемой совокупности.

Общее изменение начальной типологической структуры совокупности во времени может быть разложено на следующие элементарные процессы:

дрейф характеристик выделенных типов (классов) объектов;

изменение численности классов за счет перемещения объектов из класса в класс и появления новых объектов;

появление новых классов и исчезновение отдельных старых классов.

Первые два элементарных процесса носят плавный характер, но скорость их протекания переменная. Характер третьего элементарного процесса — изменения набора классов — дискретный.

Изменение базисной типологической структуры совокупности происходит за счет двух источников: развития действующих на момент t_0 предприятий и ввода в эксплуатацию новых предприятий. Влияние первого из них зависит от того, на каком участке своей траектории развития находятся предприятия. Отдельное предприятие на большей части своей траектории развивается сравнительно медленно, плавно, оставаясь похожим на себя. Следовательно, его классовая принадлежность не меняется.

Однако в истории развития предприятия есть ряд периодов, когда происходит резкое изменение его состояния. Это прежде всего реконструкция, после которой предприятие может перейти в другой из существующих классов и даже образовать новый класс. Относительно высокой скоростью изменения характеристик отличается также пе-

риод освоения производственной мощности. В ряде случаев радикальные изменения технологии, характера сырья и выпускаемой готовой продукции могут определить новую классовую принадлежность действующего предприятия. В связи с тем, что предприятия, образующие базисную совокупность, как правило, находятся на разных этапах своей траектории и изменение их способа производства происходит в различные периоды времени, траектории развития отдельных объектов оказываются не параллельными, а пересекающимися. Это приводит к изменению распределения объектов между классами: численность одних классов уменьшается, других — возрастает. При этом набор классов остается стабильным в течение сравнительно большого промежутка времени, так как радикальные изменения способа производства в истории предприятия сравнительно редки.

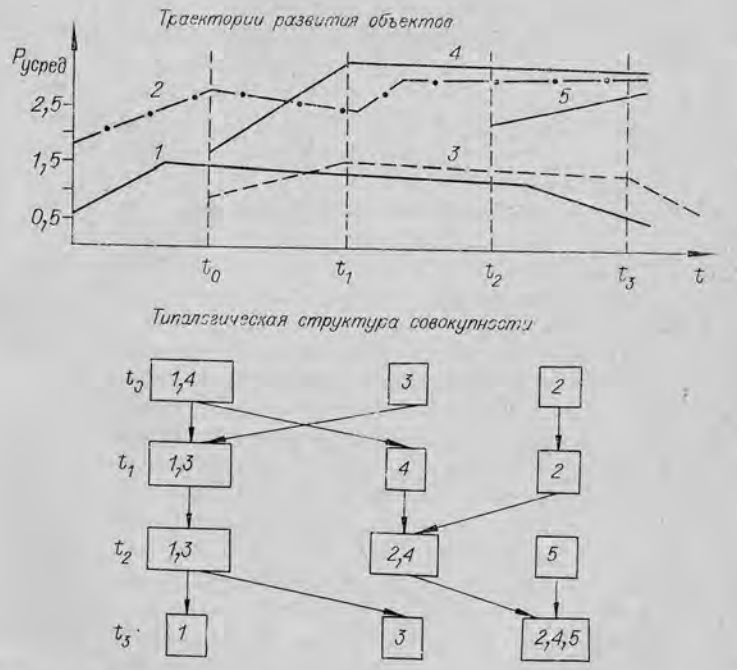


Схема формирования типологической структуры совокупности и ее динамика.

Ввод в эксплуатацию новых предприятий всегда изменяет объем совокупности и элементы ее структуры. За счет новых предприятий увеличивается удельный вес прогрессивных классов и могут появляться новые классы, но последнее, как уже указывалось, происходит сравнительно редко, так как формирование нового класса означает появление принципиально нового способа производства.

На рисунке схематически представлен процесс изменения базисной типологической структуры совокупности под воздействием указанных факторов для пяти предприятий, введенных в действие в три срока (за пределами анализируемого периода $[t_0, t_3]$ и в моменты t_0 и t_2). По каждому из объектов в верхней половине рисунка показана динамика производственной мощности в разные периоды их функционирования (освоение, стационарный режим, реконструкция, «старение»). В нижней части рисунка дана типологическая структура совокупности по показателю производственной мощности на моменты t_0, t_1, t_2, t_3. Видно, что при стабильном числе классов (три) из-за асинхронности процесса развития предприятий имеет место изменение численности классов и значительная миграция объектов между ними.

Основные, глубинные причины вышеописанного процесса изменения структуры совокупности предприятий во времени — политика распределения капитальных вложений и реализация достижений технического прогресса.

МОДЕЛИРОВАНИЕ СТРУКТУРНО-НЕОДНОРОДНЫХ СОВОКУПНОСТЕЙ

В практике экономико-статистического моделирования наиболее употребительны непрерывные регрессионные модели. Линейные непрерывные модели определяют жесткие требования к исходной совокупности: многомерное распределение признаков, описывающих объекты, должно быть нормальным или близким к нормальному. При этом предполагается стабильным характер зависимости и параметров модели во всех областях признакового пространства.

В действительности это далеко не так, и использование непрерывных моделей в таких случаях может привести

к существенным ошибкам, к нарушению принципа адекватности модели и описываемого ею явления, процесса.

Опыт многих исследований свидетельствует о том, что при переходе от одного типа объектов к другому изменяется и характер статистических зависимостей, причем в отдельных случаях достаточно резко. Если попытаться описать такую структурно-неоднородную совокупность объектов одним достаточно простым оператором, например линейной регрессионной моделью, то неизбежны значительные погрешности в индивидуальных оценках значений моделируемого показателя. В том случае, когда совокупность состоит из одного многочисленного типа и нескольких мелких, погрешность на основной массе наблюдений сравнительно невелика. Если же совокупность включает несколько «далеко» расположенных в признаковом пространстве типов с примерно равной численностью и с различным характером зависимостей, то найденный по определенному критерию единый оператор будет лишен содержательного смысла: мы получим лишь какую-то формальную «среднюю» зависимость «входа» и «выхода».

Возможно, что в некоторых сравнительно простых случаях структурной неоднородности совокупности вопрос построения адекватной экономико-статистической модели (ЭСМ) может быть решен с помощью использования нелинейных непрерывных моделей. Однако нам представляется, что такой способ по ряду причин не является кардинальным решением поставленной проблемы.

Во-первых, поиск необходимой формы уравнения связи, по существу, ведется вслепую. Никогда не может быть полной уверенности, что нелинейная форма связи, дающая хорошую интерполяцию, пригодна для экстраполяции, выхода за рамки исследуемой совокупности.

Во-вторых, значительное увеличение числа параметров в нелинейных уравнениях приводит к снижению точности оценок при ограниченных объемах совокупности (наиболее частый случай при исследовании отраслевых систем) и усложняет содержательную интерпретацию.

В-третьих, нелинейные непрерывные модели непригодны для описания негладких функций, т. е. в том случае, когда реальные зависимости имеют скачки и разрывы.

В-четвертых, нелинейные модели вызывают сложности при оптимизации.

Наконец, в-пятых, в условиях структурной неоднородности совокупности большинство статистических критериев не работает.

Таким образом, нелинейные непрерывные модели могут использоваться для описания зависимостей в структурно-неоднородных совокупностях лишь в ограниченных случаях. Основным направлением учета структурной неоднородности следует считать комбинацию, синтез методов многомерной группировки и методов корреляционно-регрессионного анализа, а также использование специальных типов моделей.

Совместное применение методов многомерной классификации и корреляционно-регрессионного анализа приводит к тому, что при экономико-статистическом моделировании могут использоваться три принципиально различных способа описания зависимостей: два так называемых «чистых» способа — при помощи классификационных (дискретных) или непрерывных моделей — и один комбинированный— при помощи дискретно-непрерывных моделей.

Дадим краткую характеристику[1] двух сравнительно новых для практики экономико-статистических исследований типов моделей: классификационных (дискретных) и дискретно-непрерывных, позволяющих в определенной мере решить проблему учета типологической структуры совокупности.

Классификационная модель промышленного объекта представляет собой две классификации объектов — в пространстве факторов и показателей производства — и описание связи этих классификаций. Такие модели могут широко использоваться в экономических исследованиях, что объясняется, во-первых, пороговым характером реакции экономических систем на те или иные воздействия, во-вторых, необходимостью в ряде случаев представления данных в дискретной форме, в-третьих, эффективностью использования классификационных моделей для анализа и прогнозирования комплекса взаимосвязанных экономических показателей.

Дискретно-непрерывная модель (ДНМ) представляет собой комбинацию моделей двух типов — классификационной (дискретной) и непрерывной. Исследуемая зависи-

[1] Подробнее эти вопросы изложены в работе [10].

мость здесь представлена набором уравнений, каждое из которых действительно в определенной области признакового пространства. Если признаковое пространство разбить на непересекающиеся области $A_1, A_2, \ldots, A_k$, то дискретно-непрерывная модель имеет вид

$$y = \begin{cases} f_1(\vec{x}) \text{ при } \vec{x} \in A_1, \\ f_2(\vec{x}) \text{ при } \vec{x} \in A_2, \\ \ldots\ldots\ldots\ldots\ldots \\ f_k(\vec{x}) \text{ при } \vec{x} \in A_k. \end{cases}$$

Построение ДНМ включает: разбиение пространства признаков — построение дискретной части модели, построение непрерывной части модели, оценку аппроксимирующей и прогнозирующей способности ДНМ.

Выбор одного из указанных типов моделей в конкретном исследовании должен базироваться на учете как общих требований к статистической модели, так и специфических условий данного исследования. К числу общих требований следует отнести:

а) наиболее полное отображение реальной структуры исследуемой совокупности объектов в соответствии с принципом структурного подобия;

б) максимально возможное использование информации, содержащейся в исходных данных;

в) максимально возможную простоту модели, обеспечивающую решение задачи с достаточной точностью.

Эти общие требования к типу модели зачастую оказываются противоречивыми, и возникает задача поиска наиболее рационального компромисса между ними.

Специфические условия конкретного исследования — это назначение будущей модели, наличие качественных признаков в описании объекта, объем исследуемой совокупности и т. п.

Разные типы модели в разной степени обеспечивают выполнение указанных требований.

Регрессионная непрерывная модель для всей совокупности не вызывает потерь в информации, связанных с разбиением совокупности на части; при том же объеме наблюдений дает более надежные оценки параметров, позволяет в явной форме оценить влияние отдельных производственных факторов. В то же время она непригодна для неоднородных совокупностей, так как в этом случае не обеспечивается принцип структурного подобия. Часто возникают затруднения в использовании этого типа моделей при нелинейных зависимостях и наличии качественных признаков в описании объектов.

Классификационные модели позволяют учитывать общее влияние комплекса признаков, исключают затруднения с учетом качественных признаков, а также ограничения, связанные с формой связи и однородностью совокупности, и т. д. В то же время их применение сопряжено с существенными потерями информации за счет усреднения.

Анализ материалов по ряду отраслевых систем показывает, что наиболее универсальна дискретно-непрерывная модель. При рациональном сочетании составных частей она позволяет получить лучшие результаты в смысле адекватности модели исследуемому процессу, чем при использовании каждого из «чистых» способов моделирования в отдельности. Большим преимуществом этого типа модели перед непрерывными регрессионными является также возможность более точного отражения процесса развития объектов отрасли под влиянием технического прогресса и других факторов.

В наших исследованиях рациональный тип модели выбирался на основе специальных схем. В качестве иллюстрации приведем такую схему для наиболее простого случая — выбора рационального типа модели для пространственной совокупности.

Схема включает три основные операции.

Операция 1. Оценка степени структурной однородности (неоднородности) совокупности. Здесь возможны две ситуации:

а) если совокупность (или ее основная часть) имеет достаточно высокую степень однородности, то наиболее целесообразным типом модели считается линейная регрессионная функция;

б) если совокупность неоднородна, т. е. в ней выделяются отдельные четко выраженные подсовокупности (классы), то переходят к операции 2.

Операция 2. Классификация совокупности, выделение однородных классов, анализ характера внутриклассовых связей. В зависимости от результата указанного анализа — наличия или отсутствия существенной

Схема выбора рационального типа модели для пространственной совокупности.

внутриклассовой вариации показателей и факторов производства — определяется направление дальнейшего движения. При отсутствии существенных внутриклассовых связей строится классификационная модель, в противном случае процесс выбора рационального типа модели продолжается. При выборе типа модели однозначное решение совершенно не обязательно: может оказаться, что для отдельных классов общей совокупности целесообразно строить внутриклассовые уравнения регрессии, а для других — нет.

Операция 3. Исследование характера расположения классов в пространстве признаков. В случае «далеких», «изолированных» классов строится дискретно-непрерывная модель без учета весов принадлежности объектов к классам (нерандомизированная модель). При «смазанных», «близких» классах, наличии переходных типов объектов предпочтительнее построение взвешенной дискретно-непрерывной модели (рандомизированной), которая учитывает вероятность принадлежности отдельных объектов к сформированным классам. В такой ситуации этот тип модели дает значительный выигрыш в использовании информации, так как оценка параметров модели выполняется не по каждому классу в отдельности, а по всей совокупности в целом.

Выбор рационального типа модели существенно усложняется при построении динамических моделей. В этом случае необходимо учесть изменения структуры выборки во времени и выделить «однородные» пространственно-временны́е блоки.

Изменение базисной структуры совокупности во времени вызывается не просто развитием входящих в нее объектов, а *неравномерностью их развития*. Представим себе такой гипотетический случай: начальная точка (год пуска в эксплуатацию) у всех объектов одна и та же, развитие идет с одинаковой (близкой) скоростью и по одинаковой (близкой) траектории. Структура такой совокупности оставалась бы стабильной во времени, и характер взаимосвязей показателей и факторов производства изменялся бы у всех объектов в один и тот же момент под влиянием: а) перехода предприятий с одного участка траектории развития на другой (например, с периода освоения на стационарный режим); б) внедрения технико-технологических и организационных новшеств; в) дей-

ствия эволюторных факторов, опосредованных фактором времени.

В действительности предприятия, образующие базисную совокупность, как правило, находятся на разных этапах своей траектории развития, и изменение способа производства у них происходит в различные периоды времени. Поэтому задача формирования однородных пространственно-временны́х блоков включает два аспекта: выделяются предприятия одного «поколения» и одного характера развития в рассматриваемом интервале времени, а для этих однородных (в смысле истории развития) предприятий — участки траектории со стабильным характером взаимосвязи показателей и факторов производства.

Для решения задачи в общем случае необходимо было бы провести разбиение всех объектов за все T лет и для каждой из полученных групп построить динамическую модель. Однако (помимо вычислительных трудностей) такое одновременное пространственно-временно́е разбиение может привести к получению большого числа малочисленных классов, в которых невозможно применение статистических методов анализа. Поэтому, как правило, приходится использовать однородные группы объектов за каждый год исследуемого периода и анализировать характер изменения элементов пространственной классификации во времени. Такой анализ — один из наиболее важных этапов, принципиально отличающих схему построения динамической ЭСМ от статической. Коротко содержание этого этапа может быть сформулировано как установление соответствия между классами объектов разных «поколений» и выделение устойчивых классов. Особую роль в установлении соответствия между классами играет матрица связи между ними. Обозначим ее через λ:

$$\Lambda = \| \lambda_{\tau t} \|, \qquad \tau, t = \overline{1, T},$$

где $\lambda_{\tau t}$ есть мера классификации τ-го и t-го годов. В качестве таких мер могут использоваться, например, коэффициенты взаимной сопряженности Пирсона или Чупрова, коэффициент близости двух классификаций [11] и другие аналогичные показатели.

Можно всегда предположить, что $\lambda_{ii} = \max \lambda_{ij} = 1$. Следовательно, $\lambda_{ij} \leqslant 1$, $i, j = \overline{1, T}$. Для получения критерия устойчивости (преемственности) классификаций по

годам удобно перестроить матрицу λ по какому-либо алгоритму диагонализации (см., например, [12, 13]).

Разумеется, принятые решения о соответствии классов разных лет (т. е. о том, что тот или иной класс $(t + 1)$-го года есть развитие определенного класса t-го года) связывается с некоторым порогом, задаваемым в зависимости от содержания конкретных задач.

* * *

Резюмируя, можно сказать, что при экономико-статистическом моделировании мы практически почти всегда имеем дело со структурно-неоднородными совокупностями. Обнаруженный во множестве эмпирических исследований факт существования естественных типов объектов обусловливает эффективность методов многомерной классификации как инструмента экономико-статистического исследования. Применение методов распознавания образов в комбинации с корреляционно-регрессионным анализом позволяет найти приемлемое решение сложной проблемы структурной неоднородности совокупности.

ЛИТЕРАТУРА

1. **Суслов И. П.** Теория статистических показателей. М., «Статистика», 1975. 246 с.
2. **Плошко Б. Г.** Группировка и системы статистических показателей. М., «Статистика», 1971. 176 с.
3. **Бородкин Ф. М., Лукацкая М. Л.** К вопросу о статистической оценке причинных связей.— В кн.: Статистические методы в экономическом анализе производства. Новосибирск, «Наука», 1968, с. 3—9.
4. **Сиськов В. И.** Об определении величины интервалов при группировках.— «Вестник статистики», 1971, № 12, с. 4—11.
5. **Пасхавер Б.** Проблема интервалов в группировках.— «Вестник статистики», 1972, № 6, с. 24—28.
6. **Дружинин Н. И.** К вопросу о группировках.— «Вестник статистики», 1973, № 2, с. 43—47.
7. **Parsen E.** On the Estimation of a Probability Function and Mode.— «Ann. Math. Stat.», 1962, v. 33, N 3.
8. **Кулешов В. В., Лукацкая М. Л., Талышева Л. П.** Исследование некоторых особенностей динамики функционирования ле-

созаготовительного предприятия.— В кн.: Распознавание образов и регрессионный анализ в экономических исследованиях. Новосибирск, 1972, с. 171—178.
9. **Розин Б. Б., Сергеева Л. А.** Прогнозирование сроков и кривых освоения производственной мощности угольных шахт.— В кн.: Вопросы экономико-статистического моделирования и прогнозирования в промышленности. Новосибирск, 1970, с. 90—127.
10. **Розин Б. Б.** Теория распознавания образов в экономических исследованиях. М., «Статистика», 1973. 223 с.
11. **Миркин Б. Г.** Об одном подходе к обработке нечисловых данных.— В кн.: Математические методы моделирования и решения экономических задач. Новосибирск, 1969, с. 141—150.
12. **Беккер А. В., Лукацкая М. Л.** Об анализе структуры матрицы коэффициентов связи.— В кн.: Вопросы экономико-статистического моделирования и прогнозирования в промышленности. Новосибирск, 1970, с. 273—284.
13. **Лумельский В. Я.** Группировка параметров на основе квадратной матрицы.— «Автоматика и телемеханика», 1970, № 1, с. 133—143.

Б. Б. РОЗИН, В. В. КУЛЕШОВ,
М. Л. ЛУКАЦКАЯ

ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ ЭКОНОМИКО-МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ В ТЕХНОЛОГИИ ПЛАНОВЫХ РАСЧЕТОВ

Повышение качества планирования — одна из важнейших задач экономической науки и практики. Выступая на XXV съезде КПСС, Л. И. Брежнев сказал: «Необходимо в первую очередь обеспечить серьезное совершенствование планирования... Здесь широкое поле для приложения усилий экономической науки, для внедрения современных научных методов, в том числе экономико-математических, для использования автоматизированных систем управления»[1].

Принципиально новым направлением совершенствования планирования является внедрение в технологию планирования автоматизированной системы плановых расчетов (АСПР). АСПР определяется как система разработки народнохозяйственных планов и контроля за их выполнением. Определяющее значение для эффективного функционирования АСПР имеет широкое использование в разработке народнохозяйственного плана (в различных режимах и на разных стадиях его формирования) системы экономико-математических моделей.

РОЛЬ И ЗНАЧЕНИЕ ЭКОНОМИКО-МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ В АСПР

При использовании аппарата экономико-математического моделирования процесс разработки плана должен рассматриваться как объект моделирования высокой сте-

[1] **Брежнев Л. И.** Отчет Центрального Комитета КПСС и очередные задачи партии в области внутренней и внешней политики. Доклад XXV съезду КПСС 24 февр. 1976 г. М., Политиздат, 1976, с. 71.

пени сложности. Модель технологии плановых расчетов (МТПР) может быть представлена в виде системы сводно-задающих и структурно-функциональных схем АСПР. Такую систему следует рассматривать как первоначальный вариант МТПР. В дальнейшем эта модель будет совершенствоваться, что найдет отражение в сводно-задающих и структурно-функциональных схемах, в наборах плановых функций, последовательности их выполнения, агрегировании плановых функций в блоки и т. д.

Модель технологии плановых расчетов необходимо ориентировать на отражение организационно-методических принципов планирования. В ней должны быть представлены все подсистемы АСПР, а также отражена возможность обращения к внешним источникам информации, последовательность прохождения последней, время ее переработки в каждой подсистеме и периодичность обращения информации между подсистемами. Естественно, что МТПР должна в качестве элементов включать и весь арсенал используемых экономико-математических моделей. Критериями выбора форм, схем организации плановых расчетов могут служить время разработки плана и оперативность его корректировки при изменении целевых установок или политико-экономической ситуации. В ходе отбора сводно-задающих схем необходимо ориентироваться на качественную сторону планирования: степень сбалансированности и близость к оптимальному значению принятого критерия.

В ходе реализации МТПР могут варьироваться структурные звенья АСПР. Изложенный тезис основан на следующей посылке: изменение целей функционирования экономической системы может потребовать не только изменения плановых показателей, но и установления новых функций для подсистем и даже изменения всей системы плановых расчетов. В частности, ориентация внутренней политики на повышение уровня жизни народа закономерно приводит к наделению соответствующей сводно-функциональной подсистемы свойствами целеполагания, такими как «материальный уровень жизни», «культура», «экология» и др. В связи с необходимостью повышения эффективности территориального разделения труда правомерным представляется включить в число сводных подсистем народнохозяйственного плана подсистему пространственной организации экономики.

Использование экономико-математических моделей (ЭММ) в АСПР наиболее эффективно при их объединении в единую систему. Увязка отдельных моделей, их объединение в блоки взаимосвязанных моделей и создание в конечном счете системы ЭММ АСПР должны осуществляться на основе схем, отражающих технологию разработки народнохозяйственного плана.

Система ЭММ АСПР включает модели, при помощи которых в диалоговом режиме решаются предусмотренные технологией планирования задачи; эти модели объединяются сначала через неформализованные процедуры, а затем своими входами и выходами, информационно и алгоритмически. Процесс внедрения ЭММ в АСПР — постадийный, отдельные стадии различаются степенью соответствия требованиям технологии планирования, согласованности входов и выходов отдельных моделей системы, их информационной и алгоритмической увязки.

Включение ЭММ в АСПР — длительный многоэтапный процесс, его реализация предусматривает адаптацию ЭММ к сложившейся технологии планирования, с одной стороны, и последующее уточнение (совершенствование) последней под обратным влиянием модельных процедур — с другой.

Уже сейчас внедрение модельной технологии, требования системно-программного подхода предполагают некоторые изменения технологии планирования, например выделение подсистемы отраслевых комплексов, включение в сводный народнохозяйственный план территориального разреза, более полный учет научно-технического прогресса.

На первом этапе в качестве отправной точки можно принять сложившуюся технологию планирования с некоторыми уточнениями. Начальный шаг в процессе сближения ЭММ и плановых функций структурно-функциональных схем предполагает «понятийное приближение» этих двух категорий. Оно должно выражаться в формулировании набора требований и условий со стороны плановых функций к моделям и в изучении возможностей («потребительских свойств») различных типов моделей. Это, в свою очередь, предполагает классификацию всех плановых функций по характеру их требований, а также описание и классификацию различного типа моделей по возможностям выполнения тех или иных требований.

Поэтому целесообразно провести типизацию функций по характеру требований и моделей по их возможностям с последующим наложением для сравнительного анализа и выявления взаимосопряженности. Это позволит по крайней мере установить последовательность встраивания (включения) моделей различных классов в технологический процесс планирования.

АНАЛИЗ ПЛАНОВЫХ ФУНКЦИЙ

Инструментом реализации начального шага может быть экспертный опрос ведущих специалистов плановых органов и научно-исследовательских учреждений, занимающихся разработкой и применением ЭММ.

Основополагающие признаки классификации плановых функций следующие: режим, стадия, итеративный цикл внутри стадии и уровень плановой функции. Считаем, что плановые функции могут быть подразделены на такие классы (названия ориентировочны):

А — формирование целей и задач экономической политики;

Б — аналитико-оценочные и прогнозные функции, формирование возможных альтернатив развития;

В — итеративные сводно-балансовые функции, формирование сбалансированных вариантов развития;

Г — оптимизационно-балансовые функции, выбор оптимального варианта развития;

Д — принятие решения по окончательному варианту плана.

Коротко охарактеризуем каждый из выделенных классов.

Содержанием плановых функций класса А является определение целей экономического развития, их уточнение и детализация в различных аспектах (социально-экономический, технический прогресс и др.), в территориальном и программном разрезах; разработка системы показателей (критериев), характеризующих степень достижения целей. Плановые функции данного класса в значительной степени неформализуемы, однако в последние годы разработаны методы, позволяющие вносить элементы алгоритмизации в процедуры их выполнения. Здесь также может быть полезно выявление (опросы) общественного мнения по вопросам уровня жизни и т. п.

Плановые функции класса Б служат для прогнозирования возможных путей развития и оценки их социально-экономических последствий. Отличительная черта функций, включаемых в этот класс, состоит в том, что большое внимание уделяется анализу прошлых тенденций, оценке положения на начальный момент планирования и укрупненным прогнозам, выявлению узких мест, диспропорций. Все это определяет большую роль специалистов при разработке данных функций и, как следствие, возможность использования автономных моделей [2]. Такие аналитические оценочные процедуры присутствуют на всех стадиях и во всех режимах планирования, но их роль возрастает по мере увеличения горизонта планирования и убывает по мере приближения к конечным стадиям разработки планового документа в определенном режиме. Для процедур этого типа могут быть использованы, например, экспертно-статистические подходы и укрупненные статистические модели автономного характера.

По направленности плановые функции класса Б могут быть разделены на три подкласса:

$Б_1$ — оценка и прогноз альтернатив развития исходя из научно-технического прогресса, сдвигов в отраслевой и территориальной структуре производства с детализацией их по отраслям материального производства и целевым программам;

$Б_2$ — оценка и прогноз возможностей удовлетворения альтернатив развития основными фондами, трудовыми и материальными ресурсами;

$Б_3$ — оценка и прогноз нормативов, формирование показателей затрат, их уточнение в разрезе важнейших материалов, ресурсов, машин. При окончательном формировании системы нормативов должны учитываться балансовые и оптимизационные расчеты.

Назначение плановых функций класса В — выработка первичных сбалансированных заданий и вариантов с дифференциацией по направлениям, в отраслевом

[2] Под автономной понимается модель, которая функционирует вне связи с другими моделями, использует свою специфическую информацию. Результаты расчетов по таким моделям обычно не требуют согласования ни между собой, ни с более высокими уровнями.

и территориальном разрезах. Процесс сводки носит двойственный итеративный характер: с одной стороны, по временны́м циклам внутри одного уровня, с другой — по уровням планирования. Эти функции включают, как правило, первичную сводку (для постепенно возрастающего уровня агрегирования), выявление невязок (дисбалансов), уточнение нормативов, новую сводку и т. д. до получения допустимой точности решения. Допустимая погрешность балансовых процедур возрастает по мере увеличения горизонта планирования и убывает при приближении к конечным стадиям разработки планового документа в определенном режиме. Соответственно в противоположном направлении изменяется и роль балансовых процедур, и необходимое количество итераций. Характер задач и методов, применяемых для выполнения балансовых процедур, неодинаков и зависит от уровня планирования — отраслевого, комплекса отраслей, макроуровня. На начальных итерациях балансовых процедур для отраслевых подсистем большую роль будут играть методы прямого счета и модели оценки нормативов различного рода в зависимости от технического уровня производства, природно-географических и организационных факторов. Для последующих итераций и более высоких уровней основное значение будут иметь матрично-балансовые модели и обеспечивающие их модели агрегированных норм и нормативов расходов и потребностей.

По степени точности и разагрегированности (дифференциации) показателей функции класса В могут быть разделены на три подкласса:

B_1 — определение масштабов, темпов и пропорций расширенного воспроизводства, объема и структуры капиталовложений, финансовых пропорций, ресурсов научно-технического прогресса, количественных характеристик народного благосостояния; дифференциация темпов и пропорций в территориальном разрезе, в привязке к отраслям и отраслевым комплексам; анализ и согласование отраслевых проектировок;

B_2 — дальнейшая дифференциация и уточнение по объемам капиталовложений, финансовым пропорциям, трудовым ресурсам, темпам и пропорциям расширенного воспроизводства, а также в территориальном разрезе;

B_3 — разработка финансового плана и системы материальных балансов.

Содержанием плановых функций класса Г является разработка оптимального варианта основных направлений по территории, капиталовложениям, трудовым ресурсам, основным показателям, по программному разрезу плана. К этому этапу должны быть разработаны плановые нормативы, достигнута основная балансировка и сформулированы критерии предпочтения вариантов. В разработке данного класса функций могут быть широко применены оптимизационные модели или их системы.

Наконец, назначение плановых функций класса Д состоит в подготовке системы показателей, общеэкономических пропорций, в формировании проекта основных направлений и доведении его до исполнителей. Здесь необходимо отметить определяющую роль специалистов.

КЛАССИФИКАЦИЯ ЭММ ПО СПОСОБУ ПОСТРОЕНИЯ

С точки зрения использования ЭММ в АСПР основной является классификация моделей по назначению, т. е. по их способности обеспечивать выполнение определенного класса плановых функций (экономического заказа). Но «потребительские» свойства модели, ее возможности в значительной мере определяются и способами (методами) ее построения. Эти две системы классификации взаимосвязаны, но не жестко: один и тот же заказ может быть выполнен моделями, построенными разными способами. Поэтому необходимо рассмотреть обе классификации.

Разработанные и прошедшие к настоящему времени экспериментальную и практическую проверку ЭММ по способам построения условно можно подразделить на пять типов.

I — *балансовые модели*, в основе которых лежат соотношения, связывающие входные и выходные переменные. Основные цели этих моделей состоят в проверке реализуемости фиксированных состояний входных переменных, в анализе влияния уровня входных переменных на состояние и выход системы, в сравнении и анализе различных альтернатив.

II — *оптимизационные модели*, где на множестве допустимых состояний системы вводится тем или иным спо-

собом отношение предпочтения. В зависимости от способа введения этого предпочтения оптимизационные модели разделяются на одноцелевые и многоцелевые.

III — *экономико-статистические модели*, реализация которых основывается на использовании методов математической статистики и теории распознавания образов.

IV — *сетевые модели*. Рассматриваемый процесс в них — это последовательность взаимосвязанных этапов с временны́ми и ресурсными характеристиками. Сетевые модели могут использоваться в разных режимах разработки планов. Они в удобной форме отражают динамический аспект процесса и к настоящему времени располагают программным и математическим обеспечением. Применение этих моделей в АСПР распространяется на широкий класс задач планирования и прогнозирования, включая разработку крупных научно-технических проектов, инвестиционных программ, программ развития регионов и т. п. С помощью сетевых моделей можно осуществлять распределение ресурсов в сводно-функциональных подсистемах в режиме краткосрочного и среднесрочного планирования.

V — *модели прямых расчетов*, реализация которых заключается в выполнении на ЭВМ по заданному алгоритму простейших арифметических операций и операций сводки.

Кроме перечисленных пяти «чистых» способов построения моделей необходимо выделить (и ориентироваться на него) новый, гибридный, комбинированный тип модели. Каждый тип модели обладает своими специфическими особенностями и возможностями, удовлетворяет определенным требованиям. Комбинированная модель позволяет использовать достоинства нескольких типов моделей и тем самым расширить сферу их применения. Путь создания и использования таких моделей становится в настоящее время одним из основных направлений совершенствования ЭММ. Кроме внутримодельной гибридизации весьма перспективно и объединение в систему моделей различных типов, например сочетание сетевых моделей с линейно-программными оптимизационными или линейно-программных моделей верхнего уровня с экономико-статистическими моделями нижнего уровня и т. д.

Следует также особо выделить метод коллективных экспертных оценок со сложившейся в настоящий момент

времени первичной методикой организации экспертных опросов и обработки их результатов. Это объясняется следующими обстоятельствами: во-первых, указанный метод применим в слабоформализуемых сферах плановых функций (особенно для прогнозирования научно-технического прогресса и его последствий, прогнозирования и управления социальными процессами и т. п.); во-вторых, он позволяет работать с минимумом информации и характеризовать качественно новые процессы и явления для долгосрочной перспективы; в-третьих, метод коллективных экспертных оценок открывает возможности для использования в процессе прогнозирования и планирования качественной информации.

К признакам классификации моделей по способам построения помимо перечисленных примыкают и такие, как способ отражения динамики, количество уровней моделирования, характер учета неопределенности и т. п.

ОСОБЕННОСТИ ЭКОНОМИКО-СТАТИСТИЧЕСКИХ МОДЕЛЕЙ И ИХ МЕСТО В СИСТЕМЕ ЭММ В АСПР

Экономико-статистические модели (ЭСМ) в отличие от различных типов балансовых и оптимизационных моделей имеют скорее исследовательский, прогнозный характер и могут быть названы моделями описания экономических процессов. В ЭСМ главным предметом изучения являются не способы достижения определенного состояния, а сам процесс развития объектов в пространстве и во времени. Такие модели приспособлены для выявления тенденций, имевших место в прошлом и ожидаемых в будущем.

Эта особенность обусловливает место и значение ЭСМ в системе моделей АСПР. Они особенно важны на начальных стадиях составления плана: при анализе существующих состояний, прогнозировании, оценке и формировании альтернатив развития и могут применяться на всех уровнях планирования, начиная от производственного объекта и кончая уровнем народного хозяйства.

Экономико-статистическая модель — особая форма ЭММ, она содержит элементы неопределенности и описывает стохастические связи, закономерности под воздей-

ствием множества причин и следствий в массовых повторяющихся явлениях. Такая модель может быть построена только на основе обобщения уже имеющейся информации о функционировании экономических объектов и систем, т. е. индуктивно-эмпирическим путем.

ЭСМ не только форма (способ) описания определенных объективных закономерностей (что является функцией любой модели), но в то же время и метод анализа, агрегации, сжатия исходной уже существующей информации. Прочие модели требуют, как правило, поступления вторичной, обработанной, специально приспособленной к к задаче информации, во многих случаях носящей прогнозный, плановый, проектный характер, короче говоря, не существующей в начальный период планирования информации. Поэтому, имея самостоятельную аналитико-прогнозную функцию, ЭСМ может выступать в качестве преобразователя информации (уже подготовленной соответствующим образом) для моделей других типов и быть средством представления первоначальной исходной информации в свернутом, агрегированном виде. Основной способ агрегирования — это описание главных (важных) закономерностей, содержащихся в информации в неявном виде, с помощью статистических моделей.

Информационные свойства ЭСМ определяют в то же время и ее недостатки: она не может отразить принципиально новых способов производства, пассивна, т. е. в ней нет цели, она описывает процесс, но не управляет им, срок прогнозирования по ЭСМ ограничен и зависит от степени инерционности процесса. Именно этим объясняется тот факт, что статистические модели функционирования объектов хорошо дополняют и стыкуются с нормативными оптимизационными моделями поиска оптимальных состояний. Их синтез, взаимоувязка позволяет преодолеть узость каждого отдельного подхода: с одной стороны, в нормативные модели ввести процесс движения к оптимальному состоянию, учесть возможности его практического достижения существующими механизмами производственно-хозяйственной деятельности; с другой стороны, включить элемент цели, нормативности в модели функционирования, т. е. наметить цели процесса развития, направить его в желаемую сторону.

Необходимо отметить, что элементы статистического подхода (в явной или неявной форме) имеются во всех

других типах модели. В частности, при построении сетевых моделей ЭСМ используются для оценки длительности работ и величины затрат; применяемые в матричном планировании нормативы во многом имеют статистическую природу; при имитации экономических процессов не только информация, но и сама процедура решения задачи носит статистический характер.

Итак, экономико-статистические модели могут выполнять самостоятельную роль на начальных стадиях планового процесса и вспомогательную — на стадии балансировки и оптимизации.

ЭСМ КАК БАЗА ФОРМИРОВАНИЯ ПЛАНОВЫХ НОРМ И НОРМАТИВОВ В АСПР

Одна из наиболее сложных и малоразработанных проблем модельной технологии планирования состоит в применении ЭММ для разработки агрегированных нормативов различных уровней планирования. Наибольшую трудность в решении этой проблемы для реального планирования, как указывает академик В. М. Глушков [3], представляет выбор правильных значений нормативов. На макроэкономическом уровне обычно используются агрегированные нормативы, в то время как научно-технические идеи, приводящие к изменению нормативов, возникают на уровне отдельных процессов, агрегатов, машин. Без специальной системы, связывающей с макроэкономическими моделями этот нижний уровень, на котором протекает вся реальная деятельность по совершенствованию и созданию новых технологий, такие модели не могут быть использованы для планирования реального развития экономики. Все вышесказанное вполне справедливо можно отнести не только к сводному планированию на народнохозяйственном уровне, но и к другим уровням планирования и соответственно агрегирования нормативов: комплексу отраслей, отрасли, региону, программе и т. п.

Нормы и нормативы — это и начальное, и конечное звено технологического процесса разработки плана. В ходе реального планирования последовательно происходит итеративный процесс агрегирования, уточнения и оптимизации норм и нормативов. В связи с этим можно предложить следующий первоначальный вариант процедуры

применения ЭММ для разработки агрегированных нормативов, включающий четыре этапа.

Первый этап заключается в формировании библиотеки экономико-статистических моделей объектов, описывающих зависимость реальных норм расходов сырья, электроэнергии, капиталовложений и т. д. от комплекса производственных факторов. Такая библиотека дает дифференцированные закономерности формирования удельных расходов на нижнем уровне процесса производства на начало планового периода.

Второй этап состоит в прогнозировании дифференцированных нормативов на плановый период. Для этого осуществляется корректировка как самих моделей, полученных на первом этапе, так и значений производственных факторов на плановый период. Корректировка должна учитывать научно-технический прогресс, изменение природных и внешних условий и т. д. Для решения задач второго этапа помимо содержательного анализа могут быть использованы различные методы статистической экстраполяции и коллективные экспертные оценки.

Третий этап — агрегирование дифференцированных нормативов на основе учета запланированных структурных соотношений различных способов производства. Здесь могут быть использованы результаты расчетов по балансовым моделям и методы прямого счета (например, средневзвешенные показатели и др.).

Четвертый этап включает задачи уточнения и оптимизации нормативов с учетом их взаимосвязей и взаимозаменяемости. Тут могут быть использованы результаты расчета по оптимизационным моделям. Полученные агрегированные нормативы после анализа могут приниматься как директивные.

Основой предложенной процедуры разработки агрегированных нормативов является библиотека ЭСМ производственных объектов. С ее помощью госплан может автономно, самостоятельно, вне зависимости от отраслевых АСУ осуществлять прикидочные расчеты различных нормативов (расходы и потребности в сырье, энергии, материалах, труде и др.), в связи с намечающимися изменениями условий производства, а также таких экономических показателей, как производительность, фондоотдача, себестоимость и др. Это вполне реализуемо, так как для описания процесса формирования какого-либо норматива

по объектам отрасли достаточно иметь небольшое число моделей (порядка 10). Эти модели позволяют учесть влияние технического прогресса, природных факторов, изменения номенклатуры продукции и др. организационно-экономических факторов на величину индивидуальных и средних нормативов.

Информационная библиотека должна состоять из моделей формирования показателей работы предприятий. Предприятие — один из важнейших уровней организации материального производства: оно и производитель материальных ценностей, и источник информации об условиях и результатах их производства. Показатели предприятий — основа построения агрегированных показателей всех более высоких уровней планирования народного хозяйства. Реализация всех направлений технического прогресса осуществляется на предприятиях.

КЛАССИФИКАЦИЯ ЭММ ПО «ПОТРЕБИТЕЛЬСКИМ» СВОЙСТВАМ

Одним из наиболее важных признаков классификации моделей по их «потребительским» свойствам является уровень подсистемы АСПР, на которую они рассчитаны. Как и сама АСПР, система ее ЭММ имеет иерархическую структуру, ориентированную на функциональную и организационную структуру системы планирования. Она взаимодействует с ЭММ внешних по отношению к АСПР автоматизированных систем и опирается на результаты расчетов, выполняемых в территориальных, отраслевых и ведомственных АСУ, в научно-исследовательских и проектных организациях. Эта иерархическая структура может быть рассмотрена по крайней мере в трех аспектах: функциональном (производственном), территориальном и программном. Функциональный аспект включает следующие уровни: народнохозяйственный, комплекс отраслей, отрасль, производственное объединение (предприятие)[3]. В территориальном аспекте различаются следующие

[3] Хотя АСПР и не включает в себя подсистему производственных объектов, но тем не менее необходимо предусмотреть возможность обращения АСПР к моделям этого уровня как к источнику информации для всех более высоких уровней планирования.

уровни: СССР в целом, республика (крупный регион), область (район), предприятие, в программном аспекте — народнохозяйственный уровень (программа повышения уровня жизни, обороноспособности страны, обеспечения внешних экономических связей и т. п.), межотраслевой (аграрно-промышленный комплекс), многоотраслевой (топливно-энергетический комплекс), территориальный (БАМ, Нечерноземье) и территориально-производственный уровень (ТПК).

Модели каждого типа характеризуются целым спектром «потребительских» свойств (признаков), которые могут быть объединены в следующие группы.

I. Характеристика выхода:

1) цель применения модели (анализ, прогноз, сжатие информации, балансировка, оптимизация);

2) территориальная дифференциация (точечная, по республикам, по регионам (объектам));

3) горизонт планирования (долгосрочный, среднесрочный, текущий);

4) временна́я дифференциация (одно состояние, несколько состояний, вся траектория развития);

5) форма выхода модели (точечная оценка (детерминистская), диапазон значений, интервал с вероятностными оценками, закон распределения).

II. Характеристика входа:

1) необходимая информация (полное совпадение с действующей отчетностью, относительно несложные формальные преобразования, необходимы специальные разработки);

2) источник информации (отчетная, плановая);

3) природа информации (объективная, субъективная (экспертная)).

III. Уровень реализуемости модели:

1) методическая отработанность (теоретическая модель, эксплуатируемая модель, проект методики, утвержденная методика);

2) математическая обеспеченность (стандартная программа, сервис-программы, библиотека программ, пакет программ);

3) техническая обеспеченность (М-222, БЭСМ-6, ЕС-1020, ЕС-1033, ЕС-1050, «Эллиот»);

4) способы расчетов по модели (разовые, диалоговый режим, смешанный);

5) время расчетов на машине;

6) время подготовки информации (начальная стадия);

7) время подготовки информации (в рабочем состоянии).

IV. Степень освоенности:

1) экспериментальная проверка (расчеты условные, по информации, близкой к фактической, по фактической информации, по формам госплана);

2) практическое внедрение (использование в научно-аналитических целях, выработка рекомендаций и предложений по модельным расчетам, разовые расчеты в управляющих органах, включение в технологию планирования).

Соответственно выделенным иерархическим уровням и способам построения ЭММ ниже приводится типизация моделей, разработанных в Институте экономики и организации промышленного производства СО АН СССР.

I. Модели производственного (функционального) планирования

А. Народнохозяйственный уровень

1. Динамическая модель межотраслевого баланса (ДММБ) — 35 отраслей.

2. Оптимизационная динамическая модель межотраслевого баланса (ОДММБ) — 35 отраслей.

3. Натурально-стоимостной межотраслевой баланс производственных мощностей (НСБПМ): а) статический, б) динамический.

4. Укрупненная балансовая динамическая модель народного хозяйства (УДМ).

5. Межотраслевая модель оптимизации структуры потребления (ММОСП) — 35 отраслей.

6. Оптимизационная модель структуры потребления с целевой функцией благосостояния (ОМЦФБ) — 15 отраслей.

7. Модель оптимизации соотношения потребления и производственного накопления в национальном доходе (МОПиН).

8. Оптимизационная балансовая модель нормативов относительной эффективности капитальных вложений в экономию сырья и материалов (МОЭКВ).

9. Экономико-статистическая балансовая модель системы макроэкономических показателей (ММЭП) — 55 показателей.

10. Экономико-статистическая балансовая адаптационная межотраслевая динамическая модель с элементами оптимизации (МДМ): а) СССР, б) Японии, в) США.

Б. Уровень комплекса отраслей

1. Натурально-стоимостной межотраслевой баланс производственных мощностей (НСБПМ): а) статический, б) динамический.

2. Координирующая модель оптимизации развития аграрно-промышленного комплекса (МОК) — 10 отраслей, 17 районов.

В. Отраслевой уровень

1. Оптимизационная модель развития и размещения группы отраслей (ОМГО) — союзная: а) лесной промышленности; б) нефтедобывающей, нефтеперерабатывающей и нефтехимической промышленности, в) электротехнической промышленности, г) местных строительных материалов.

2. Оптимизационная модель развития и размещения отдельных отраслей промышленности (ООМ) — союзная: а) производства чугуна, стали, проката, б) угледобывающей промышленности, в) комбайностроения, г) дорожного машиностроения.

3. Адаптивная оптимизационная модель развития и размещения объектов газовой промышленности (АООМ)— союзная.

Г. Уровень предприятия

1. Информационная экономико-статистическая модель производственных объектов (ИЭСМ объектов): а) лесозаготовительной промышленности, б) угледобывающей, в) нефтедобывающей, г) цементной промышленности, д) производства стали, е) промышленности стройматериалов.

2. Оптимизационная экономико-статистическая модель производственных объектов (ОЭСМ объектов) лесозаготовительной промышленности.

Тип модели		Хара						
		Цель применения модели					Террит диффер	
шифр	сокращенное наименование	анализ	прогноз	сжатие информации	балансировка	оптимизация	точечная	
		1	2	3	4	5	6	
IA1	ДММБ	+	+		+		+	
IA2	ОДММБ	+	+			+	+	
IA3а	НСБПМ стат.	+	+		+	+	+	
IA3б	НСБПМ дин.	+	+		+	+	+	
IA4	УДМ	+	+				+	
IA5	ММОСП	+	+			+	+	
IA6	ОМЦФБ	+	+			+		
IA7	МОПиН	+	+		+	+	+	
IA8	МОЗКВ		+			+	+	
IA9	ММЭП	+	+				+	
IA10а	МДМ СССР	+	+		+	+	+	
IA10б	МДМ Японии	+	+		+	+	+	
IA10в	МДМ США	+	+		+	+	+	
IБ1а	НСБПМ стат.	+	+		+	+	+	
IБ1б	НСБПМ дин.	+	+		+	+	+	
IБ2	МОК					+		
IВ1а	ОМГО лес.		+			+		
IВ1б	ОМГО нефть	+			+	+		
IВ1в	ОМГО эл.-тех.	+			+	+		

II. Модели территориального планирования

А. Народнохозяйственный уровень

1. Оптимизационная межотраслевая межрайонная модель (ОМММ): а) укрупненная — 16 отраслей, 11 районов; б) детализированная — 50 отраслей, 16 районов; в) двухстадийная (взаимодействие производственного и транспортного блоков) с фиксированной структурой выпуска продукции в дробной номенклатуре — 1000 наименований продукции, 16 и более районов.

Б. Уровень республики, региона

1. Оптимизационная модель производственной структуры республики, региона (ОМР) — порядка 100 отраслей, 15 областей.
2. Оптимизационная модель пространственной структуры системы ТПК района (МСТПК) — 30 отраслей, до 10 ТПК.

В. Уровень области (района)

1. Оптимизационная модель формирования основных параметров ТПК (МПТПК) — 10—15 отраслей.
2. Оптимизационная модель пространственной структуры ТПК (МТПК) — 15—20 отраслей, 10—16 промышленных узлов.

III. Модели программного планирования

А. Межотраслевые программы

1. Координирующая модель оптимизации развития аграрно-промышленного комплекса (МОК).

Б. Территориальные программы

1. Сетевая модель реализации крупной экономической программы типа БАМ (СМЭП).

Таблица 2

Адресная привязка моделей ИЭиОПП СО АН СССР к сводно-задающим схемам технологии планирования в терминах АСПР

Тип модели		Режим планирования			Стадия планирования			Классы функций					Структурные звенья															
шифр	сокращенное наименование	текущее	пятилетнее	долгосрочное	концепция	основные направления	проект плана	формирование целей	аналитико-оценочные и прогнозные	итеративные и сводно-балансовые	оптимизационно-балансовые	принятие решения	темпы и пропорции	капиталовложения	труд и кадры	финансы	материальные балансы	наука и техника	уровень жизни	вычислительная техника	внешнеэкономические связи	себестоимость и прибыль	нормы и нормативы	цены и ценообразование	территориальное планирование	подсистема отраслевых комплексов	отраслевая подсистема	АСПР союзных республик
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
IA1	ДММБ		+	+		+				+			+	+														
IA2	ОДММБ		+			+			+		+		+	+														
IA3а	НСБПМ стат.	+	+				+				+		+													+	+	
IA3б	НСБПМ дин.		+			+					+		+													+	+	
IA4	УДМ			+	+				+				+	+														
IA5	ММОСП		+	+	+	+			+		+									+								
IA6	ОМЦФБ		+	+		+			+		+									+								
IA7	МОП и Н		+	+	+	+			+		+		+	+														
IA8	МОЭКВ		+	+		+	+				+	+		+									+					
IA9	ММЭЛ	+	+	+		+			+				+	+														
IA10а	МДМ СССР		+	+	+	+	+	+	+	+	+		+	+	+													
IA10б	МДМ Япония		+	+	+	+	+	+	+	+	+		+	+	+													
IA10в	МДМ США		+	+	+	+	+	+	+	+	+		+	+	+													
IБ1а	НСБОМ стат.	+	+				+				+		+													+	+	
IБ1б	НСБПМ дин.		+			+					+		+													+	+	
IБ2	МОК		+	+	+	+					+															+	+	
IВ1а	ОМГО лес		+	+		+			+		+															+	+	
IВ1б	ОМГО нефть		+	+		+				+	+																+	
IВ1в	ОМГО эл.-тех.		+	+		+			+		+																+	
IВ1г	ОМГО стройматериалы		+	+		+					+																+	
IВ2а	ООМ чермет.		+	+		+					+																+	
IВ2б	ООМ уголь		+	+		+					+																+	
IВ2в	ООМ комбайн		+	+		+					+																+	
IВ2г	ООМ дормаш.		+	+		+					+																+	
IВ3	АООМ газ	+	+			+																					+	
IГ1а	ИЭСМ лес	+	+			+			+		+							+			+	+					+	
IГ1б	ИЭСМ уголь	+	+			+			+		+							+			+	+					+	
IГ1в	ИЭСМ нефть	+	+			+			+		+							+			+	+					+	

1	2	3	4	5	6	7	8	9	10	11
IГ1г	ИЭСМ цемент	+	+			+			+	
IГ1д	ИЭСМ сталь	+	+			+			+	
IГ1е	ИЭСМ строймат.	+	+			+			+	
IГ2	ОЭСМ лес	+	+			+			+	
IIA1а	ОМММ укрупненная			+	+				+	+
IIA1б	ОМММ детализированная			+	+	+			+	+
IIБ1	ОМР			+	+	+			+	
IIБ2	МСТПК		+	+		+		+		
IIВ1	МПТПК			+		+				
IIВ2	МТПК		+	+		+		+		
IIIA1	МОК		+	+	+	+				
IIIБ1	СМЭП			+		+			+	
IIIВ1	СМТПК			+		+				

В. Территориально-производственные программы

1. Сетевая оптимизационная модель создания и функционирования ТПК (СМТПК).

Модель технологии плановых расчетов с использованием ЭММ. Процесс разработки народнохозяйственного плана следует рассматривать как объект моделирования высокой степени сложности. Система сводно-задающих и структурно-функциональных схем АСПР является и базой построения модели технологии плановых расчетов (МТПР). Для создания МТПР, подлежащей реализации в АСПР, необходимо включить в эту систему схем апробированные экономико-математические модели, позволяющие решать соответствующие плановые задачи.

Эффективным путем построения МТПР, на наш взгляд, является сравнительный анализ и взаимоувязка классификации плановых функций с классификацией ЭММ. Анализ характеристик отдельных типов моделей по их «потребительским» свойствам (табл. 1) в сочетании с классификацией плановых функций позволяет определить место отдельных моделей в АСПР, условия и очередность их последовательного включения в технологию планирования (табл. 2).

12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
+								+			+					+	
+								+			+	+				+	
+								+			+	+				+	
+		+	+				+	+			+	+		+		+	
+		+	+				+							+			
+														+			+
+														+			+
+														+	+		
+														+	+		
+															+		

Структурно-функциональные схемы с включенными в них типами ЭММ можно рассматривать как первый вариант МТПР.

ЛИТЕРАТУРА

1. **Безруков В. Б.** О внедрении первой очереди АСПР.— «Плановое хозяйство», 1976, № 5, с. 46—52.
2. **Безруков В. Б., Шеховцов В. В.** Роль АСПР и АСУ в совершенствовании планирования.— «Вопросы экономики», 1975, № 3, с. 107—116.
3. **Глушков В. М.** Макроэкономические модели и принципы построения ОГАС. М., «Статистика», 1975. 160 с.
4. **Методические** указания к разработке государственных планов развития народного хозяйства СССР. М., «Экономика», 1974.
5. **Проскуров В. С.** Информация в АСПР. М., «Экономика», 1975.
6. **Типовое** проектирование блока «Капитальное строительство» отраслевых подсистем АСПР.— В кн.: Научные труды научно-исследовательского института планирования и нормативов. М., 1976. 152 с.
7. **Тихомиров Б. И.** Вопросы типового проектирования функциональных подсистем АСПР.— В кн.: Материалы семинара-совещания «Научно-методические и организационные проблемы АСПР» (Кишинев, 1973), ч. III. М., ГВЦ Госплана СССР, 1973, с. 185—309.
8. **Мацнев Д. А., Райзберг Б. А., Сосков В. Ф., Уриксон Я. М.** Функционально-структурные схемы разработки пятилетнего народнохозяйственного плана.— «Экономика и математические методы», 1977, т. XIII, вып. 2. с. 224—237.

А. В. БЕККЕР, А. Н. БЕККЕР, Б. Б. РОЗИН

СТАТИСТИЧЕСКИЙ ПОДХОД К ОДНОЙ ЗАДАЧЕ УПРАВЛЕНИЯ

Для создания системы управления некоторым объектом, построения ее структуры, выбора параметров и законов управления в последние годы широко применяются математические методы. В частности, в настоящей работе предлагается подход к управлению, основанный на применении методов математической статистики и методов многомерной классификации.

Задача управления возникает на различных уровнях и для разных объектов. Здесь она рассматривается применительно к промышленному предприятию, что вызвано в основном тем вниманием, которое уделяется именно этому объекту. Однако управление промышленным предприятием рассматривается в рамках отрасли или подотрасли.

Управление может быть *тактическим*, направленным на поддержание существующего процесса функционирования предприятия, и *стратегическим*, направленным на перевод объекта в качественно новое состояние. В настоящей работе исследуются некоторые проблемы второго из перечисленных типов управления — стратегического. Кратко охарактеризуем его основные особенности.

Стратегическое управление всегда связано с преодолением некоторого порога, с таким изменением результативных экономических показателей функционирования предприятия, которое существенно значимо для последующих звеньев (уровней) промышленного производства. Следовательно, первая характерная черта стратегического управления заключается в его дискретном, ступенчатом характере.

Далее, изменение качественного уровня функционирования промышленного объекта, как правило, измеряется не одним каким-либо результативным показателем, а целым комплексом таких показателей.

И наконец, перевод предприятия с одного уровня (состояния) на другой осуществляется путем воздействия на множество регулируемых параметров.

Итак, сам характер задачи стратегического управления обусловливает эффективность использования методов многомерной классификации. Соответственно этому предлагаемый в работе подход к решению задачи базируется на идеях обучения, развиваемых в теории распознавания образов. Управление в рассматриваемом нами смысле заключается в сравнении изучаемого объекта с некоторым набором объектов (эталоном) и в определении на основе этого сравнения необходимых изменений комплекса регулируемых параметров для перевода объекта с существующего уровня качества на уровень эталона.

Для формализации идей, определяющих данный подход к задаче управления, используем некоторые определения и положения из работы [1].

Как уже отмечалось, решение задачи управления, по существу, связано не с отдельным предприятием, а с изучением совокупности объектов некоторой отрасли в определенный период времени. Объект управления характеризуется рядом входных и выходных параметров, а процесс развития объекта — их изменением. В пространстве этих параметров можно наблюдать определенные сочетания их значений. Другими словами, если рассматривать множество объектов управления, то можно выделить определенные комбинации параметров, которые объединяют объекты в образы (типы). Под образом в данном случае следует понимать производственную ситуацию, характеризующую такое состояние объекта управления, при котором комбинация регулирующих воздействий предопределяет его выходное состояние.

Понятие образа отражает наличие во всей совокупности однородных подсовокупностей объектов. Разумеется, речь идет не о полной (абсолютной) однородности, а лишь о такой, при которой допустимо приложение установленных статистических закономерностей формирования результативных показателей ко всем элементам, входящим в состав исследуемой подсовокупности. Критерий

однородности здесь формируется исходя из изучаемой связи. В частности, таким критерием может быть требование статистической идентичности уравнений для частей подсовокупности, получаемых при любом произвольном разбиении ее на связанные области.

Отрасль промышленности, как правило, включает предприятия, построенные в разные периоды времени, оснащенные различной техникой, имеющие неодинаковые объемы производства и т. п. Значит, она может быть представлена как совокупность объектов, обладающих определенной структурой, и эта структура изменяется во времени (см. [1]).

Основными элементами типологической структуры отрасли в каждый данный момент времени являются набор классов (образов) и распределение объектов по этим классам. Соответственно можно выделить и два основных процесса развития структуры базисной совокупности: изменяется как распределение объектов по базисному алфавиту классов, так и сам набор классов.

В связи с тем, что предприятия, образующие базисную совокупность, как правило, находятся на разных этапах своей истории развития и изменение их способа производства происходит в разные периоды времени, траектории развития отдельных объектов оказываются не параллельными, а пересекающимися. В результате изменяется распределение объектов между классами: численность одних классов уменьшается, других — возрастает. Однако так как радикальные изменения способа производства в жизни предприятий сравнительно редки, то набор классов, как правило, остается стабильным в течение сравнительно большого времени. Более того, относительно стабильны и статистические характеристики имеющихся классов. Последнее обстоятельство (относительная стабильность алфавита классов) создает хорошие предпосылки для широкого применения методов многомерной классификации с целью прогнозирования и управления.

Развитие объектов управления идет по пути улучшения качества. Это значит, что существующие классы объектов можно упорядочить по степени выражения качества, которое отражается в комбинациях результативных показателей объекта. Можно, например, различать объекты низкого, среднего и высокого уровня. В общем случае число градаций может быть больше или, наоборот, мень-

ше. Однако в задаче, о которой пойдет речь ниже, предполагается наличие не менее двух классов (образов).

Будем считать, что в базисный период совокупность предприятий отрасли распадается на s классов и эти классы уже упорядочены. Обозначим их через $A_1, A_2, \ldots, A_s$. Следовательно, развитие объекта можно представить как его движение по классам A_i, $i = \overline{1, s}$. При этом переход через несколько классов — довольно редкое явление. Обычно это связано с резким изменением регулирующих воздействий и не всегда возможно.

Заметим, что переход объекта из класса в класс в настоящее время не является полностью управляемым процессом, и в этом смысле развитие объекта частично происходит пассивно.

Мы же предлагаем метод активного управления процессом перехода из класса в класс. Задача состоит в нахождении набора значений регулирующих параметров некоторого объекта из класса A_j таким образом, чтобы этот объект перешел в класс A_{j+l}, где l — число, выбираемое исследователем. Переход из класса с низким «качеством» в класс высокого «качества» обеспечить трудно, поэтому число l не может быть очень большим.

Таким образом, считаем, что у нас имеется всего два класса — A_1 и A_2, и будем отыскивать значения регулирующих параметров для каждого из объектов класса A_1, чтобы эти объекты перешли в класс A_2.

Математической моделью, на которой основываются наши выводы, является многомерное нормальное распределение. Современный уровень знаний таков, что строгому многомерному анализу, за редким исключением, поддаются лишь задачи, где рассматривается нормальный случай распределения.

Использование именно нормального распределения оправдывается, с одной стороны, приемлемостью такой математической модели, с другой — применением гипотезы о компактности [2], что позволяет считать классы A_1 и A_2 однородными. Понятие однородной совокупности в том смысле, который нам требуется, состоит в следующем. Статистическую совокупность объектов $B_1, B_2, \ldots, B_N$ считаем однородной, если совокупность их измерения $Z_1, Z_2, \ldots, Z_n$ можно рассматривать как выборку, которая образована в результате N независимых испытаний над некоторым объектом B^*.

Такое определение однородности выборки позволяет предположить, что плотность распределения случайной величины, из генеральной совокупности которой взята данная выборка, является многомерной нормальной плотностью. По крайней мере можно считать, что существующее распределение хорошо описывается нормальным распределением.

Итак, будем считать, что объекты класса A_1 представляют собой выборку из генеральной совокупности с многомерным нормальным распределением с оценками для параметров $\overline{Z_1}$ и Σ_1. Здесь Z_1 — оценка математического ожидания, полученная по выборке A_1; Σ_1 — оценка ковариационной матрицы.

Объекты класса A_2 также образуют выборку из генеральной совокупности с многомерным нормальным распределением. Оценки параметров этого распределения Z_2 и Σ_2 будут получены по выборке из класса A_2.

Обозначим матрицу значений параметров, характеризующую выборку A_1, через Z_1. Эта матрица имеет размерность $N_1 \times L$, где N_1 — число элементов в выборке A_1, а L — количество всех рассматриваемых параметров, которыми характеризуется объект.

Соответственно через Z_2 обозначим такую же матрицу для класса A_2, тогда N_2 — число элементов в выборке A_2. Предполагаемый процесс управления будет заключаться в линейном преобразовании матрицы Z_1 исходных данных для класса A_1 таким образом, чтобы статистические характеристики преобразованной выборки были «близки» к статистическим характеристикам $\overline{Z_2}$ и Σ_2 из класса A_2.

В понятие «близость» в данном случае вкладывается статистический смысл. Другими словами, мы должны принять гипотезу H: исследуемые генеральные совокупности идентичны. Это означает, что данные генеральные совокупности имеют одинаковые средние и ковариации.

Статистическая теория проверки подобной гипотезы довольно хорошо разработана [3]. Однако, прежде чем проверить гипотезу H, надо найти соответствующее преобразование. Здесь мы воспользуемся методом наименьших квадратов, хотя возможны и другие критерии.

Пусть искомым преобразованием будет

$$\hat{Z}_1 = Z_1 \times C + B,$$

где C — диагональная матрица порядка $L \times L$, а B —

матрица размерности $N_1 \times L$ с постоянными столбцами.

Другими словами, для любого объекта j из класса A_1 его параметры преобразуются по формуле

$$\hat{Z}_{1k}^{(j)} = Z_{1k}^{(j)} C_k + b_k, \quad k = \overline{1, L}; \quad j = \overline{1,N},$$

где C_k — элемент, стоящий на k-м месте диагональной матрицы C; b_k — элемент k-го столбца матрицы B.

Подсчитаем оценку параметров, характеризующих выборку, по преобразованным данным. Очевидно

$$\overline{\hat{Z}_1} = \overline{Z_1} C + B.$$

Здесь $\overline{\hat{Z}_1}$ и $\overline{Z_1}$ — это матрицы с L постоянными столбцами, т. е.

$$\overline{\hat{Z}}_{1k} = \overline{Z}_{1k} C_k + b_k, \quad k = \overline{1, L}.$$

Оценкой ковариационной матрицы в данном случае будет следующая:

$$\overline{\Sigma_1} = C \Sigma_1 C' = \| C_k C_l \sigma_{kl} \|, \quad k = \overline{1, L}.$$

Действительно, ковариация параметров k и l равна

$$1/N \sum_{j=1}^{N_1} (\hat{Z}_{1k}^{(j)} - \overline{\hat{Z}}_{1k})(\hat{Z}_{1l}^{(j)} - \overline{\hat{Z}}_{1l}) =$$

$$= 1/N_1 \sum_{j=1}^{N_1} (Z_{1k}^{(j)} C_k - \overline{Z}_{1k} C_k)(Z_{1l}^{(j)} C_l - \overline{Z}_{1l} C_l) = C_k C_l \sigma_{1.kl},$$

где $\sigma_{1.kl}$ —ковариация параметров l и k в непреобразованной выборке.

Обозначим ковариационную матрицу для преобразованной выборки через $\Sigma_1(C)$. Теперь можно выбрать константы b_k для каждого из параметров так, чтобы среднее значение преобразованной выборки совпадало со средним значением для выборки A_2.

Очевидно

$$b_k = \overline{Z}_{2k} - C_k \overline{Z}_{1k};$$

действительно, в этом случае

$$\hat{Z}_{1k}^{(j)} = Z_{1k}^{(j)} C_k + b_k = Z_{1k}^{(j)} C_k + \overline{Z}_{2k} - C_k \overline{Z}_{1k}.$$

Находя среднее этого выражения, получаем

$$\overline{\widehat{Z}_{1k}} = \overline{Z_{2k}}.$$

Величины C_k требуется подобрать таким образом, чтобы матрица $\Sigma_1(C)$ была близка к матрице Σ_2.

Если нам удастся это сделать, то наша цель (управление) будет достигнута, так как в статистическом смысле классы $\widehat{A_1}$ и A_2 не будут отличаться друг от друга. Здесь $\widehat{A_1}$ — преобразованный класс.

Перейдем к определению величин C_k, $k = \overline{1, L}$. Для этого нужно определить понятие близости двух матриц.

Будем считать две матрицы близкими, если их нормы близки. В качестве нормы матрицы Σ возьмем следующее выражение

$$\| \Sigma \| = \| (\sigma_{ij}) \| = \left(\sum_{i,j} \sigma_{ij}^2 \right)^{1/2}.$$

Отсюда критерием выбора величин C_k, $k = \overline{1, L}$ будет

$$\| \widehat{\Sigma}_1(C) - \Sigma_2 \| \to \min_C$$

или

$$\sum_{i,j} (\sigma_{1.ij} C_i C_j - \sigma_{2.ij})^2 \to \min_C.$$

Выбор этого критерия полностью определяется его простотой.

Таким образом, мы приходим к задаче нахождения безусловного экстремума, которую легко решить, например, методом покоординатного спуска.

После получения характеристик $\widehat{\Sigma}_1(C)$ можно применять статистический критерий для сравнения двух выборок. Это дает возможность статистически оценить качество используемого преобразования. В принципе этот критерий можно было бы использовать и в качестве основного для получения величин C_k, $k = \overline{1, L}$.

Решение поставленной задачи осуществлялось нами для случая детерминированного разделения классов A_1 и A_2. Когда же эти классы не разделены строго, т. е. классификация практически состоит в указании для каждого объекта вероятностей его принадлежности к классам A_1 и A_2, постановка задачи должна сводиться к нахождению преобразования, увеличивающего вероятности отнесения объектов к классу A_2.

ЛИТЕРАТУРА

1. **Розин Б. Б.** Природа неоднородности статистических совокупностей и методы ее учета при моделировании.— В настоящем сборнике.
2. **Айзерман М. А., Браверман Э. М., Розоноэр Л. И.** Метод потенциальных функций в теории обучения машин. М., «Наука», 1970. 384 с.
3. **Андерсон Т.** Введение в многомерный статистический анализ. М., Физматгиз, 1963. 500 с.

И. Г. БАГИРОВ

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ УПРАВЛЯЕМЫХ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ С ПОМОЩЬЮ СТОХАСТИЧЕСКИХ УРАВНЕНИЙ ДИФФУЗИОННОГО ТИПА

Разработка формальных методов, позволяющих количественно описывать вероятностные свойства экономических процессов, является одной из центральных задач теории прикладной экономической статистики [1]. Сложность комплекса задач, связанного с оптимизацией управления стохастическими экономическими системами, требует, чтобы эти методы не только позволяли получать количественные оценки тех или иных вероятностных характеристик оптимизируемой системы, но и воспроизводили с достаточной полнотой весь вероятностный механизм, управляющий ее поведением. Описание такого механизма с помощью системы математических выражений называется экономико-статистической моделью (ЭСМ) управляемого процесса.

Существующее множество ЭСМ можно классифицировать двояким образом: во-первых, по задачам, которые ставятся перед ЭСМ общей целью экономического исследования; во-вторых, по степени сложности вероятностных конструкций, которые они описывают. В первом случае множество ЭСМ можно укрупненно разбить на два подмножества. Одно из них составляют ЭСМ, предназначенные для получения статистических оценок будущих

[1] Под экономическим процессом в настоящей работе понимается изменение состояний экономической системы на определенном отрезке времени. Поэтому в тех случаях, когда речь будет идти о статистическом моделировании, в качестве объекта моделирования выступает экономический процесс, когда же речь идет об управлении, то объектом управления выступает экономическая система.

значений рассматриваемых экономических показателей. Модели, принадлежащие этому подмножеству, можно назвать прогнозирующими. Для прогнозирующих ЭСМ характерна относительная независимость их структуры от общей задачи экономического исследования, использующей результаты прогнозирования.

Структура моделей, составляющих другое подмножество, строится в соответствии с целями оптимизационной задачи таким образом, чтобы с их помощью стохастические свойства оптимизируемой системы учитывались непосредственно в целевой функции. Оптимизационные модели, построенные с участием таких ЭСМ, получили название оптимальных экономико-статистических моделей (ОЭСМ) [1]. Использование ОЭСМ позволяет в ряде случаев перейти от исходной оптимизационной задачи в стохастической постановке к эквивалентной ей (в смысле вида получаемого решения) оптимизационной задаче в детерминированной постановке.

Во втором случае множество ЭСМ разбивается на два подмножества, не совпадающие с подмножествами предыдущего разбиения. К одному из них относятся все ЭСМ, при построении которых принимается, что поведение экономического показателя на рассматриваемых отрезках времени описывается одним и тем же случайным процессом. Прогнозирующие ЭСМ, построенные в соответствии с этим предположением, называются экстраполяционными. К другому подмножеству принадлежат ЭСМ, при построении которых необходимо исходить из предположения, что поведение экономического показателя на различных отрезках времени описывается различными случайными процессами. ЭСМ, составляющие это подмножество, реализуют более сложную вероятностную конструкцию по сравнению с обычным случайным процессом, а именно конструкцию с изменяющимися характеристиками. Случайные процессы такого типа получили название управляемых случайных процессов.

Выбор типа ЭСМ для описания поведения изучаемого экономического процесса определяется тем, на какое именно свойство реальных экономических систем делается акцент в предстоящем экономическом исследовании. Одно из таких свойств — управляемость экономических систем, т. е. возможность регулирования в определенной области по крайней мере некоторой части их основных

параметров. В сочетании с объективной стохастичностью поведения экономических систем это свойство приводит к тому, что в результате реализации различных управляющих воздействий меняются вероятностные свойства объекта управления. Иначе говоря, меняется случайный процесс, описывающий его поведение. Следовательно, если на рассматриваемых отрезках времени некоторая экономическая система испытывает различные управляющие воздействия, то адекватной математической моделью ее поведения будет ЭСМ управляемого случайного процесса.

Принципиальной основой корректного применения экстраполяционных ЭСМ служит другое свойство экономических систем — их инерционность. Она означает невозможность существенного изменения состояний экономической системы за малые промежутки времени. Это приводит к тому, что управляющим воздействиям относительно некоторой экономической системы требуется определенное время для своего срабатывания. В пределах времени срабатывания управляющих решений применение ЭСМ случайного процесса будет не только оправданным, но и, по-видимому, наиболее эффективным в настоящее время средством для описания поведения управляемых экономических процессов.

В настоящее время при наличии самых разнообразных видов экстраполяционных ЭСМ почти совершенно отсутствуют прогнозирующие ЭСМ, которые позволяли бы учитывать изменение вероятностных свойств объекта прогнозирования в зависимости от вновь принимаемых относительно его плановых решений. Такое резкое несоответствие в развитии обоих типов ЭСМ в значительной мере объясняется сравнительно поздним возникновением должной математической базы — теории управляемых случайных процессов (примерная дата возникновения 1960—1961 гг.).

Это же обстоятельство объясняет и некоторое несовершенство существующих ОЭСМ. При построении ОЭСМ неявно предполагается, что в результате действия управляющих решений меняется только математическое ожидание случайного процесса, описывающего поведение объекта управления. Управляющие воздействия планируются таким образом, чтобы средняя траектория управляемой экономической системы сближалась с наиболее выгодной для нас по некоторому критерию кривой.

При таком предположении одним из эффективных способов учета стохастических особенностей объекта управления будет применение ЭСМ в форме регрессионных уравнений.

Аппарат регрессионных уравнений наилучшим образом приспособлен именно для описания закономерностей изменения математических ожиданий зависимых переменных как функций изменения независимых переменных. При построении существующих ОЭСМ игнорируется тот факт, что управляющие воздействия, изменяя в нужном нам направлении какую-либо вероятностную характеристику объекта управления, не могут не затронуть других его вероятностных характеристик. В результате если даже при осуществлении управляющих воздействий средняя траектория управляемой системы будет располагаться вблизи наиболее выгодной для нас кривой, то и вероятность ухода с этой траектории в каждый момент времени окажется большой по величине. На наш взгляд, существующие ОЭСМ необходимо перестроить таким образом, чтобы в процессе поиска оптимальных значений входных параметров оптимизируемой системы одновременно учитывались зависимости от них математических ожиданий и дисперсий выходных параметров.

Необходимо также отметить, что моделирование экономических процессов в динамике в виде управляемых случайных процессов дает теоретическую основу для решения некоторых статистических задач, связанных с применением в экономическом анализе алгоритмов распознавания образов без учителя (таксономии). Принципиальное отличие экономических процессов от физических и химических состоит прежде всего в невозможности многократного их воспроизведения с достаточной для исследования точностью. В результате до недавнего времени существовала объективная необходимость делать статистические выводы о вероятностных характеристиках какого-либо экономического процесса на основе единственной реализации его прошлых состояний.

Положение существенно изменилось с использованием в экономическом анализе алгоритмов таксономии. Если известен набор признаков, определяющих степень близости вероятностных характеристик для некоторого множества однотипных экономических процессов, то применение алгоритмов таксономии позволяет разбить множество

реализаций их состояний в пространстве значений признаков на подмножества так, что попавшие в одно подмножество реализации можно считать реализациями одного случайного процесса. Здесь и всюду в настоящей работе однотипными считаются экономические процессы с одинаковыми видами входных и выходных параметров. Это могут быть, например, изменения состояний предприятий одной отрасли, потребляющих одинаковый набор ресурсов и выпускающих один и тот же ассортимент продукции, или совокупности городов одной страны, рассматриваемой с точки зрения потребления какого-либо ресурса (набора ресурсов), зависящего от общих для всех городов факторов.

Алгоритмы таксономии предполагают заданными — в качестве исходной информации — значения признаков, в пространстве которых осуществляется разбиение. Ответ на вопрос, какие именно признаки определяют степень близости вероятностных характеристик однотипных процессов, может быть получен с помощью содержательного экономического анализа параметров управления этими процессами и при условии, если для них определен тип ЭСМ.

Пусть задано некоторое множество однотипных систем. Для простоты изложения ограничимся ситуацией, когда они имеют один выходной параметр. Примем в качестве модели поведения систем управляемый случайный процесс. Необходимо установить, какие именно признаки определяют вероятностные свойства выходного параметра этих систем.

Предположим также, что в результате предварительного исследования были установлены входные параметры систем. Управляемый случайный процесс считается заданным, если определено семейство случайных процессов и задано множество значений параметра (параметра управления), определяющего каждый элемент этого семейства. А именно каждое фиксированное значение параметра управления вместе с начальным состоянием случайного процесса однозначно определяет его вид. В силу принятого выше предположения множество реализаций выходного параметра можно рассматривать как реализации одного семейства случайных процессов, получающихся при различных значениях параметра управления и начальных значениях выходного параметра. Параметр управления может иметь различные виды. Рассмотрим два наиболее важных для нас его вида.

1. Значения параметра управления представляют собой детерминированную функцию времени.

2. Значения параметра управления зависят в каждый данный момент времени от всей предыстории выходного параметра.

Чтобы решить вопрос о признаках разбиения, необходимо определить, какая информация о входных параметрах однотипных систем соответствует значениям двух видов параметра управления. Если рассматривать каждое значение последнего как функцию определенного комплекса управляющих воздействий, то нетрудно заметить, что параметр первого вида моделирует ситуацию перспективного планирования, когда управляющие воздействия относительно некоторой категории экономических систем требуют для своего срабатывания времени, соизмеримого с величиной планового периода.

Для статистического анализа вероятностных свойств выходного параметра экономической системы, определяемых только общей стратегией планового управления без учета ее текущей корректировки, необходимо осуществлять разбиение в пространстве тех значений входного параметра, которые статистически не связаны со значениями выходного параметра. В этом случае в набор признаков разбиения включаются значения выходного параметра на начало рассматриваемого отрезка времени и плановые значения входных параметров на всем протяжении рассматриваемого периода.

Параметр второго вида моделирует ситуацию с непрерывной корректировкой принимаемых решений. В этом случае признаками разбиения множества реализаций выходного параметра должны быть его начальные значения и фактические значения входных параметров на всем протяжении рассматриваемого отрезка времени. Приведенные рассуждения будут дополнены и уточнены ниже с помощью соответствующих формальных выкладок. Задача настоящей работы — дать методические подходы к решению указанных выше проблем экономического прогнозирования и построения ОЭСМ на основе теории управляемого случайного процесса.

Пусть $x(t, \omega)$ — случайный процесс, моделирующий выходной параметр оптимизируемой экономической сис-

темы, где $\omega \in \Omega$ — множество элементарных событий. Для каждого фиксированного ω имеем детерминированную функцию времени $x_\omega(t)$ — одну из возможных реализаций случайного процесса. Входной параметр системы обозначим через y. Ограничимся рассмотрением входного и одного выходного параметра системы, поскольку все полученные в дальнейшем выводы могут быть без особого труда распространены на более общую ситуацию, когда $\vec{x} = (x_1, x_2, \ldots, x_n)$ — вектор зависимых переменных и $\vec{y} = (y_1, y_2, \ldots, y_m)$ — вектор независимых переменных.

Выходной параметр x считается зависящим в вероятностном смысле от входного параметра y. Это значит, что одномерная параметрическая функция распределения случайного процесса $x(t, \omega)$ при фиксированном значении y имеет вид $F(t, \omega, y)$. Обозначим через $\hat{x}'(t)$ наиболее выгодную по некоторому критерию кривую, вид которой нам известен. Изменения входного параметра y в некоторой области его значений $y \in Y$ будем считать доступными нашему воздействию.

На множестве траекторий $x_\omega(t)$ $[0 \leqslant t \leqslant T]$ возможных реализаций случайного процесса $x(t, \omega)$ введем норму $||x_\omega(t)||$. Нормой в нашем случае может быть, например, функционал $||x_\omega(t)|| = \max\limits_{0 \leqslant t \leqslant T} |x_\omega(t)|$. Если теперь возьмем норму разности двух произвольных кривых $x_{\omega'}(t)$ и $x_{\omega''}(t)$ из множества реализаций случайного процесса $x(t, \omega)$, то такая норма $\|x_{\omega'}(t) - x_{\omega''}(t)\|$ будет служить расстоянием между кривыми. Условное математическое ожидание случайного процесса $x(t, \omega)$ при фиксированном y, по определению, равно

$$M[x(t, \omega)/y] = \int_{\Omega} x(t, \omega)\, dF(t, y, \omega) = x(t, y),$$

где M — знак математического ожидания; $x(t, y)$ — динамическое уравнение регрессии зависимости переменной x от независимых переменных t и y.

Меняя y в области значений $y \in Y$, доступных нашему воздействию, можно получать различные усредненные по вероятности траектории движения оптимизируемой системы. Потери для случая, когда $x(t, y)$ отклоняется от оптимальной траектории $\hat{x}'(t)$, будут функцией расстояния

между этими кривыми $f_1[\|\hat{x}'(t) - x(t, y)\|]$. Оптимальное значение $y \in Y$ находится из условия

$$\min_{y \in Y} f_1[\|\hat{x}'(t) - x(t, y)\|]. \qquad (1)$$

Модели такого вида называются оптимальными экономико-статистическими. Название этого класса моделей объясняется той ролью, которую играет в них ЭСМ. Определяя вид регрессионной зависимости $x(t, y)$, ЭСМ в этом случае непосредственно участвует в определении оптимального решения. Непосредственно по виду выражения (1) можно заметить, что существующие ОЭСМ не учитывают влияния изменения входного параметра y на дисперсию случайного процесса $x(t, \omega)$.

По определению, условная дисперсия случайного процесса $x(t, \omega)$ при фиксированном значении y равна

$$M[x(t, \omega) - x(t, y)]^2/y = \int_{\Omega} [x(t, \omega) - x(t, y)]^2 \times$$
$$\times\, dF(t, y, d, \omega) = \sigma^2(t, y).$$

Обозначим через $\hat{x}''(t)$ кривую, определяемую равенством $\hat{x}''(t) = \min\limits_{y \in Y} \|\sigma^2(t, y)\|$, где минимум берется по введенной ранее норме. Значения функции $\hat{x}''(t)$ представляют собой величины условной дисперсии случайного процесса $x(t, \omega)$ для случая, когда единственной целью изменения значений входного параметра является получение случайного процесса с минимальной дисперсией. Через $f_2[\|\hat{x}''(t) - \sigma^2(t, y)\|]$ обозначим потери для случая, когда $\sigma^2(t, y)$ может отклоняться от $\hat{x}''(t)$. Тогда ОЭСМ, одновременно учитывающая зависимость от значений входного параметра y математического ожидания и дисперсии случайного процесса $x(t, \omega)$, в общем виде запишется так:

$$\min_{y \in Y} \{f_1[\|\hat{x}'(t) - x(t, y)\|] + f_2[\|\hat{x}''(t) -$$
$$- \sigma^2(t, y)\|]\}. \qquad (2)$$

Вид выражения (2) показывает, что для решения подобного типа оптимизационных задач необходимо уметь определять вид зависимости от параметра управления y двух вероятностных характеристик — $x(t, y)$ и $\sigma(t, y)$ —

оптимизируемой системы. Для описания предлагаемого в настоящей работе подхода к построению ЭСМ, отражающих зависимость от входного параметра вероятностных характеристик выходного случайного процесса, коротко опишем математический аппарат, используемый в теории управляемых случайных процессов.

Прежде всего, это стохастические дифференциальные уравнения. Управляемый случайный процесс рассматривается как совокупность различных решений стохастического дифференциального уравнения, параметры которого изменяются в некоторой области. Первые стохастические дифференциальные уравнения были разработаны в процессе поисков математической модели для описания явления диффузии броуновских частиц. По названию первоначального объекта исследования — процесса диффузии — такие уравнения впоследствии получили название диффузионных. В общем виде диффузионное уравнение записывается так:

$$\frac{dx_t}{dt} = a(t, x_t) + b(t, x_t) W_t,$$

где x_t — координата диффундируемой частицы; W_t — винеровский процесс, т. е. случайный процесс с $MW_t = 0$, $MW_t^2 = 1$ и независимыми нормально распределенными приращениями.

Содержательный смысл этого стохастического уравнения состоит в следующем: скорость движения диффундируемой частицы dx_t/dt в момент времени $t + dt$ определяется средней скоростью ее движения $a(t, x_t)$ (зависящей только от положения частицы в момент времени t), на которую накладывается «шум» $b(t, x_t)$, вызывающий разброс относительно среднего движения. Величина «шума» $b(t, x_t)$ также зависит только от положения частицы в момент времени t. По аналогии с процессом диффузии коэффициенты $a(t, x_t)$ и $b(t, x_t)$ получили название коэффициентов переноса и диффузии соответственно.

Диффузионное стохастическое уравнение — удобная модель для описания явлений, в которых события, происходящие в некоторый момент времени, зависят лишь от событий в предыдущем моменте времени. Случайные процессы такого типа называются марковскими случайными процессами. А. И. Колмогоров показал, что марковские случайные процессы описываются уравнениями в частных

производных относительно условных плотностей распределения в некоторый момент времени t при заданных краевых условиях.

Уравнение для плотности распределения марковского процесса в некоторый момент времени t с краевым условием для плотности распределения в момент времени $s < t$ получило название прямого уравнения Колмогорова. Его также называют уравнением Фоккера — Планка — Колмогорова. Уравнение для плотности распределения марковского процесса в момент времени t с краевым условием в момент времени $s > t$ называется обратным уравнением Колмогорова. В дальнейшем И. И. Гихманом было установлено соответствие между уравнениями Колмогорова и видом отвечающего этим уравнениям диффузионного статистического уравнения [3]. Этот результат, помимо чисто теоретического значения, имеет широкую область практического использования. В частности, его применение может быть полезным для задач статистического анализа тех экономических процессов, которые могут рассматриваться как марковские случайные процессы.

Возможность использования установленного соответствия между видами коэффициентов переноса и диффузии диффузионного уравнения и уравнениями Колмогорова основывается на том, что в ряде случаев по имеющейся априорной информации о процессе удается приблизительно оценить характер зависимости его среднего значения от состояния процесса в предыдущий момент времени, т. е. установить $a(t, x_t)$ и оценить накладываемый на это среднее движение «шум» — $b(t, x_t)$. Тогда, воспользовавшись уравнениями Колмогорова, можно получить достаточно полное вероятностное описание поведения экономической системы. Введем в коэффициенты переноса и диффузии параметр управления y:

$$dx_t = a(t, x_t, y)dt + b(t, x_t, y)dW_t. \qquad (3)$$

Полученное стохастическое уравнение является довольно простой моделью управляемого случайного процесса. В тех случаях, когда значения y в каждый момент времени t не зависят от предыстории процесса $x(t, \omega)$ до момента t, решениями уравнения (3) будут марковские случайные процессы.

Несмотря на удобство использования диффузионных стохастических уравнений, применение их для описания

поведения реальных экономических процессов наталкивается на серьезные трудности. Это объясняется тем обстоятельством, что экономические системы, как правило, довольно инерционны и состояния их в каждый момент времени зависят от всей предыстории системы до этого момента времени. Естественно, что, с одной стороны, их описания в виде марковского случайного процесса оказываются слишком грубым приближением к реальному ходу экономического процесса. С другой стороны, полнота вероятностного описания диффузионным стохастическим уравнением побуждает к поиску более сложных форм стохастических уравнений, не связанных ограничением марковости их решений и в то же время позволяющих определять вероятностные характеристики этого процесса.

В настоящей работе для описания поведения управляемых экономических процессов предлагается использовать следующее обобщение диффузионного уравнения:

$$dx_t = a(t, x)dt + b(t, x)dW_t,$$

где $a(t, x)$ и $b(t, x)$ — функционалы, при каждом t зависящие от всей предыстории процесса $x(t, \omega)$.

Уравнения такого вида относятся к классу стохастических диффузионного типа. Они могут описывать довольно широкий круг реальных экономических процессов. Введем в написанное уравнение параметр управления y. Получим выражение

$$dx_t = a(t, x, y)dt + b(t, x, y)dW_t. \qquad (4)$$

В отличие от (3) решениями уравнения (4) для любых видов параметра управления будут немарковские случайные процессы. Характер таких процессов зависит от того, какая информация используется при выборе значений параметра управления y. При этом возможно несколько различных стратегий:

1) значение y фиксируется на одном уровне в течение всего рассматриваемого отрезка времени;

2) значения y меняются в течение всего отрезка времени и представляют собой детерминированную функцию времени;

3) при выборе значения y используется вновь поступающая информация о текущем состоянии экономического процесса;

4) при выборе значения y в каждый момент времени используется вся информация о предыстории процесса до этого момента времени.

Первые два вида параметра управления моделируют ситуацию перспективного планирования, два других его вида характеризуют ситуацию текущего, оперативного управления относительно малоинерционными системами, когда требуется непрерывная корректировка принимаемых решений. Такой тип управления называется управлением с обратной связью. В процессе статистического анализа на основе уравнения (4) для ситуации перспективного планирования необходимо брать такие значения y, которые не связаны с текущими значениями выходного параметра, т. е., по нашему мнению, плановые значения y. Для случая управления с обратной связью в процессе статистического анализа должны использоваться его фактические значения.

Статистическая информация о поведении экономической системы, которую мы получаем в каждом конкретном случае, имеет всегда дискретный характер. На основании этой информации можно судить о динамике значений экономического показателя лишь в некоторые, как правило, равноотстоящие друг от друга моменты времени. В связи с этим целесообразно перейти от уравнений, описывающих непрерывное изменение экономического показателя, к соответствующим рекуррентным уравнениям, определяющим управляемый случайный процесс в дискретном времени. Случайные процессы в дискретном времени называют также случайными последовательностями. Рекуррентным аналогом уравнения (4) будет уравнение вида

$$x(t + 1, y) = a(t, x, y) + b(t, x, y)\varepsilon(t + 1), \qquad (4').$$

где $\varepsilon(t + 1)$ — независимые от процесса $x(t + 1)$ нормально распределенные случайные величины с $M\varepsilon(t + 1) = 0$ и $M\varepsilon^2(t + 1) = 1$. Для любых различных моментов времени $t + 1$ и $t' + 1$ случайные величины $\varepsilon(t' + 1)$ и $\varepsilon(t + 1)$ являются некоррелированными.

Пусть теперь нам дано множество однотипных экономических систем с одним входным параметром y и одним выходным параметром x. Поведение выходного параметра каждой системы из рассматриваемого множества подчиняется уравнению (4'). Для определенности будем считать,

что значение параметра управления y представляет собой детерминированную функцию времени.

Предположим теперь, что нам известны реализации выходного параметра процессов $x(t, \omega)$ на отрезке времени $[0 \leqslant t \leqslant T]$. Обозначим их через $\vec{x}_T^k = (x_T^1, \ldots, x_T^k)$, где k — номер системы в рассматриваемом множестве, $k = \overline{1, W}$. Пусть также известны плановые значения входного параметра системы для всех $t \in [0, T]$. Обозначим их через $\vec{\hat{y}}_T^k = (\hat{y}_T^1, \ldots, \hat{y}_T^k)$. Значения выходного параметра в момент времени $t = 0$ обозначим через x_0^k. При подстановке в уравнение (4′) набора значений $(\vec{\hat{y}}_T^k, x_0^k)$, характеризующих k-й процесс, получаем случайный процесс, описывающий поведение выходного параметра $\vec{x}_T^k$. Следовательно, если мы разобьем множество реализаций $\{\vec{x}_T^k\} = X_T$ в пространстве признаков $(\vec{y}_T^k, x_0^k)$ с помощью какого-либо алгоритма таксономии, то попавшие в один класс реализации можно считать реализациями одного случайного процесса.

Выбор в пользу определенного алгоритма таксономии не является в настоящее время строго обусловленным. Объясняется это отсутствием в теории распознавания образов общепринятого формализованного в статистическом смысле определения понятия однородности. Между тем каждый алгоритм таксономии в своей основе является конкретизацией понятия однородности как количественной меры сходства между исследуемыми объектами. В настоящей работе для разбиения объектов на однородные классы предлагается использовать алгоритм «Рельеф». С его помощью выделяются такие однородные группы, для которых совместная плотность вероятностного распределения признаков с удовлетворяющей нас точностью аппроксимируется гауссовским распределением [4].

Предположим теперь, что с помощью алгоритма «Рельеф» исходное множество реализаций X_T разбилось на однородные классы. Если бы был известен вид уравнения (4′), описывающего поведение выходных параметров рассматриваемого множества однотипных процессов, то мы могли бы с помощью стандартных статистических процедур конкретизировать вид случайных процессов, отвечающих набору реализаций в каждом классе разбиения.

К сожалению, для большинства встречающихся на практике ситуаций на основе имеющейся априорной информации о поведении экономического показателя невозможно определить вид адекватного ему стохастического уравнения.

В этом случае, по-видимому, единственно возможным выходом является поиск в некотором выбранном заранее классе стохастических уравнений. Причем отыскивается такое уравнение, которое бы наиболее точно описывало реализации в каждом классе разбиения. Более точно это означает, что мы должны найти вид стохастического уравнения с параметром управления y, решения которого мало отличаются от решений исходного уравнения. В данной ситуации решения исходного стохастического уравнения с параметром управления y представлены классами реализаций нашего разбиения. Для решения поставленной задачи необходимо иметь формальный критерий, показывающий точность описания каким-либо стохастическим уравнением имеющегося набора классов реализаций. Таким критерием является статистический критерий минимума различающей информации [5]. Опишем вкратце сущность его применения.

В математической статистике к оценкам неизвестных параметров вероятностного распределения предъявляются следующие требования: состоятельности, несмещенности и эффективности. Требование эффективности выражает стремление получать в некотором смысле оптимальные оценки неизвестного параметра. Обычно в качестве критерия оптимальности оценки выбирается $\min\limits_{\hat{a}} M(\hat{a} - a)^2$, где a — оцениваемый параметр в плотности распределения $\psi(x, a)$ и $\hat{a}$ — его статистическая оценка. Оценки, удовлетворяющие этому условию, называются эффективными.

Пусть $\vec{x} = (x_1, x_2, \ldots, x_n)$ — реализации многомерной случайной величины с плотностью совместного распределения $\psi(\vec{x}, a)$, где a — неизвестный параметр. Эффективная оценка в предположении ее существования отыскивается методом максимума правдоподобия. Он заключается в том, что в известную функцию плотности совместного распределения вместо переменных $\vec{x}$ под-

ставляются значения реализаций случайной величины и полученная функция максимизируется по параметру a. Значение параметра a, образующее эту функцию в максимум, является его эффективной оценкой $\hat{a}$. Оказывается, что среди всех статистических оценок параметра a только эффективные оценки обращают в минимум выражение

$$\sum_{ij} r_{ij}(a-\hat{a}_i)(a-\hat{a}_j), \qquad (4'')$$

где $\hat{a}_i$ — статистическая оценка параметра a только по i-й реализации; r_{ij} — элементы матрицы $R=\{r_{ij}\}$, обратной матрице ковариаций $\sum=\{\sigma_{ij}^2\}$ многомерной случайной величины, т. е. $R=\sum^{-1}$.

Выражение (4″) называется критерием минимума различающей информации. Если мы минимизируем его по a, то найденное значение $\hat{a}$ будет эффективной оценкой по всем n реализациям случайной величины. По своему содержанию критерий минимума различающей информации можно рассматривать как расстояние между набором реализаций случайной величины и плотностью распределения, взятой для ее описания. Если у нас имеется несколько плотностей распределения, то для описания реализаций случайной величины необходимо взять ту из них, для которой критерий минимума различающей информации имеет меньшую величину. В том случае, когда плотность распределения зависит от нескольких параметров, выражение для критерия минимума различающей информации принимает вид

$$\sum_k \sum_{tt'} r_{tt'}\left(a^k-\hat{a}_t^k\right)\left(a^k-\hat{a}_{t'}^k\right), \qquad (5)$$

где $k=\overline{1,\,m}$ — число параметров; $\hat{a}_t^k$ — эффективная статистическая оценка k-го параметра по t-й реализации; a^k — k-й оцениваемый параметр.

Можно доказать, что минимум выражения (5), взятый по значениям a^k, равен выражению

$$\sum_k \sum_{tt'} r_{tt'}\left(a_t^k-\hat{a}_t^k\right)\left(a_{t'}^k-\hat{a}_{t'}^k\right), \qquad (5')$$

где a_t^k — оценки максимума правдоподобия параметра a^k по t-й реализации.

Возьмем произвольный класс полученного разбиения с числом реализаций r. Перенумеруем попавшие в один класс реализации от 1 до r. Массив значений реализаций $\vec{x}_T^1, \vec{x}_T^2, \ldots, \vec{x}_T^r$ представляет собой матрицу $X=\{x_{ti}\}$, где $t=\overline{0,\,T}$, $i=\overline{1,\,r}$. Доказано, что статистическая оценка математического ожидания в виде выборочного среднего для широкого класса вероятностных распределений является асимптотически эффективной. Тогда для больших r можно рассматривать оценки

$$\hat{x}_t=\sum_{i=1}^{r} x_{ti}/r \text{ и } \hat{\sigma}_t=\sum_{i=1}^{r}(x_{ti}-\hat{x}_t)^2/r$$

как приблизительно эффективные. Обозначим через $\psi(\vec{x}, \vec{a})$, где $\vec{x}=(x_0, x_1, \ldots, x_T)$; $\vec{a}=(a^1, a^2, \ldots, a^m)$, плотность распределения, найденную по виду стохастического уравнения, взятого для описания полученных классов разбиения.

Считая реализации одного класса разбиения независимыми, найдем вид плотности распределения, описывающей весь массив реализаций:

$$\prod_{i=1}^{r}\psi_i(\vec{x},\vec{a})=\psi_1(\vec{x},\vec{a})\,\psi_2(\vec{x},\vec{a})\ldots\psi_r(\vec{x},\vec{a}). \qquad (6)$$

Подставив в выражении (6) в каждую i-ю функцию вместо переменных $\vec{x}$ значения реализаций $\vec{x}_T^1, \ldots, \vec{x}_T^r$ и максимизируя полученную функцию по параметрам $\vec{a}$, найдем их статистические оценки $\vec{\hat{a}}=\hat{a}^1, \ldots, \hat{a}^m$.

Полученные оценки $\vec{\hat{a}}$ подставим в исходную функцию $\psi(\vec{x}, \vec{a})$ вместо параметров $\vec{a}$. Интегрируя функцию $\psi(\vec{x}, \vec{a})$ по всем переменным $\vec{x}$, за исключением переменной x_t, $t=\overline{0,\,T}$, будем последовательно получать одномерные плотности вероятностных распределений $\psi(x_t)$. Найдем по этим функциям математические ожидания $Mx_t=\int_{x_t} x_t\psi(x_t)\,dx_t$. Интегрируя функцию $\psi(\vec{x}, \vec{a})$ по всем переменным $\vec{x}$, за исключением x_t и $x_{t'}$, $t'=\overline{0,\,T}$, будем получать двумерные плотности вероятностных распреде-

лений $\psi(x_t, x_{t'})$. Зная Mx_t и вид функций $\psi(x_t, x_{t'})$, найдем матрицу ковариаций

$$\Sigma = \{\sigma^2_{tt'}\}, \text{ где } \sigma^2_{tt'} = \int_{t'} \int_{t} (x_t - Mx_t)(x_{t'} - Mx_{t'})\,\psi(x_t, x_{t'})\,dx_t dx_{t'}.$$

По элементам матрицы Σ найдем обратную матрицу $R = \Sigma^{-1} = \{r_{tt'}\}$. Подставим в выражение (5′) для $k = 1, 2$ вместо $a^1_t, a^2_t, \hat{a}^1_t$ и $\hat{a}^2_t$ соответственно Mx_t, σ^2_t, x_t и σ^2_t и подсчитаем его значение.

Сравнивая его затем с аналогичными значениями, полученными по другим плотностям распределений, соответствующих выбираемым стохастическим уравнениям, в качестве описывающего заданный класс реализаций возьмем то из них, для которого значение выражения (5′) наименьшее. Поскольку точность описания должна рассматриваться по нескольким классам разбиения, в качестве критерия отбора возьмем сумму значений выражения (5′) по всем классам разбиения.

Определим теперь общий вид стохастического уравнения с параметром управления, решения которого мы будем сопоставлять наборам реализаций в каждом классе разбиения. Это уравнение должно выбираться таким, чтобы незначительными его модификациями можно было получить представительную выборку управляемых случайных процессов, из которых затем отбирался бы тот, для которого сумма значений выражения (5′) по всем классам разбиения будет наименьшей. С этой целью рассмотрим систему стохастических уравнений вида

$$\begin{cases} x(t+1, y) = \theta(t, y) + b(t, x_t, y)\,\varepsilon(t+1), \\ \theta(t+1, y) = \theta(t, y) + \varepsilon(t+1), \end{cases} \tag{7}$$

где $\theta(t)$ — случайный процесс, определяемый предысторией процесса $x(t, \omega)$, есть аналог коэффициента переноса $a(t, x)$.

Система уравнений (7) решается при начальных условиях $a_0 = M\theta_0/x_0$; $\gamma_0 = M(\theta_0 - x_0)^2/x_0$. Для того чтобы иметь возможность получить на основе такой системы уравнений вид семейства плотностей распределения выходного параметра x с параметром управления y, необходимо уточнить вид зависимостей от y процессов $\theta(t, y)$

и $b(t, x_t, y)$. Другими словами, необходимо конкретизировать вид уравнений, входящих в систему (7). Эти зависимости зададим следующим образом. Заметим, что при изменении параметра y в области значений, доступных нашему воздействию, одновременно меняются процессы $\theta(t, y)$ и $b(t, x_t, y)$. Наличие в них общего параметра управления y приводит к тому, что определенному виду коэффициента диффузии $b(t, x_t, y)$ для некоторого фиксированного значения y соответствует вполне определенный вид процесса $\theta(t, y)$.

Соответствие, которое устанавливает общий параметр y между видами процессов $\theta(t, y)$ и $b(t, x_t, y)$, говорит о том, что мы можем использовать для задания управляемого случайного процесса вместо системы (7), определяющей в явном виде зависимости от y, описание процесса $\theta(t)$ в зависимости от различных видов процесса $b(t, x_t)$. С этой целью уточним вначале вид зависимости от y коэффициента диффузии $b(t, x_t, y)$. Рассмотрим отдельно две формы коэффициента диффузии:

$$b(t, x_t, y) = \sum_{i=1}^{m} \alpha_i(y)\,t^i \tag{8}$$

и

$$b(t, x_t, y) = \sum_{j=1}^{n} l_j(y)\,x_t^j. \tag{8′}$$

В обоих случаях предполагается, что при изменении параметра управления y меняются только коэффициенты $\alpha_i = \alpha_i(y)$, $i = \overline{1, m}$ и $l_j = l_j(y)$, $j = \overline{1, n}$. Общий вид коэффициента диффузии остается неизменным.

Зависимость видов процессов $\theta(t)$ от различных видов коэффициента диффузии определим следующим образом: при изменении коэффициентов $\alpha_i = \alpha_i(y)$ или $l_j = l_j(y)$ (в зависимости от того, в какой форме берется коэффициент диффузии), процесс $\theta(t)$ меняется таким образом, что при этом условная плотность распределения $\psi[\theta(t)/x_0^t]$, где $x_0^t = (x_0, x_2, \ldots, x_t)$ — предыстория процесса $x(t)$, остается все время гауссовской. В силу сделанных предположений система уравнений (7) конкретизируется в систему уравнений вида

$$\begin{cases} x(t+1) = \theta(t) + b(t, x_t)\,\varepsilon(t+1), \\ \theta(t+1) = \theta(t) + \varepsilon(t), \end{cases} \tag{7′}$$

где $b(t, x_t)$ определяется видом выражения (8) или (8′). Причем при изменении его коэффициентов в некоторой области их значений характер условного распределения $\theta(t)$ при фиксированной предыстории процесса $x(t, \omega)$ не изменяется, оставаясь все время гауссовским.

Это условие вместе с системой уравнений (7′) дает возможность получать статистические оценки процесса (θt) в каждый данный момент времени t как функции наблюдаемой предыстории x_0^t процесса $x(t, \omega)$. Задачи подобного рода, в которых по известным значениям одного случайного процесса находятся статистические оценки другого, связанного с ним случайного процесса, относятся к классу так называемых задач оценивания, или фильтрации.

В общем случае к ним могут быть отнесены различные задачи статистического анализа, постановка которых укладывается в следующую схему: имеется двумерный случайный процесс $[\theta(t), x(t),]$ компоненты которого зависимы. Наблюдению по каким-либо причинам доступна только одна компонента, допустим, $x(t)$. Требуется по известной реализации процесса $x(t)$ до момента t оценить значение процесса $\theta(t)$ в момент времени t так, чтобы полученная оценка $\widehat{\theta}(t)$ была оптимальной по некоторому критерию. Обычно в качестве критерия оптимальности оценки $\widehat{\theta}(t)$ берется минимум $M\,[\widehat{\theta}\,(t) - \theta\,(t)]^2/x_0^t$ математического ожидания квадрата отклонения оценки $\widehat{\theta}(t)$ от истинного значения $\theta(t)$ при фиксированной реализации x_0^t процесса $x(t)$. Доказано, что оптимальной по этому критерию оценкой $\widehat{\theta}(t)$ будет оценка, удовлетворяющая условию

$$\hat{\theta}\,(t) = M\theta\,(t)/x_0^t.$$

Для решения задач фильтрации необходимо уточнить вид зависимости между процессами $x(t)$ и $\theta(t)$. Первые задачи фильтрации были решены в предположении гауссовости совместного распределения процессов $\theta(t)$ и $x(t)$. Затем были найдены уравнения фильтрации для случая, когда двумерный процесс $[\theta(t), x(t)]$ является марковским случайным процессом. Дальнейшее развитие методов решения задач фильтрации шло в направлении ослабления условий, налагаемых на совместное распределение процессов $\theta(t)$ и $x(t)$. Р. М. Липцером и А. П. Ширяевым найдено уравнение фильтрации для нахождения оценок $\widehat{\theta}(t)$ по предыс-

тории процесса x_0^t в предположении гауссовости только условного распределения ненаблюдаемого процесса $\theta(t)$ при фиксированной предыстории x_0^t [6].

Таким образом, предположив гауссовость условного распределения $\psi\left[\theta\,(t)/x_0^t\right]$, можно применить к случайным процессам $\theta(t)$ и $x(t)$, удовлетворяющим системе (7′), уравнения оптимальной фильтрации для нахождения оценки $\hat{\theta}\,(t) = M\theta\,(t)/x_0^t$. Обозначим $\widehat{\theta}(t)$ через a_t. Из уравнений оптимальной фильтрации получим систему рекуррентных уравнений для a_t:

$$\begin{cases} a_{t+1} = a_t + \dfrac{b\,(t,\ x_t) + \gamma_t}{b^2\,(t,\ x_t) + \gamma_t}\,(x_{t+1} - a_t), \\ \gamma_{t+1} = \gamma_{t+1} - \dfrac{[b\,(t,\ x_t) + \gamma_t]^2}{b^2\,(t,\ x_t) + \gamma_t}. \end{cases} \tag{9}$$

Система (9) решается при начальных условиях (a_0, γ_0), которые будем считать неизвестными. Причем относительно (a_0, γ_0) введем следующее дополнительное ограничение: если $b(t, x_t)$ берется в форме $b\,(t, x_t) = \sum\limits_i \alpha_i\,(y)\,t^i$, то $a_0 \geqslant 0,\ \gamma_0 > 0$. Если же $b\,(t, x_t) = \sum\limits_j l_j\,(y)\,x_t^j$, то $a_0 \geqslant 0$, $\gamma_0 = 0$. Это означает, что условная начальная плотность распределения $\psi(\theta_0/x_0)$ для $\gamma_0 = 0$ вырождается в дельта-функцию $\psi(\theta_0/x_0) = c\,\delta(\theta_0 - a_0)$. Если подставить в нижнее уравнение системы (9) начальное значение $\gamma_0 = 0$, то сразу же можно убедиться, что в этом случае $\gamma_t \equiv 0$ и система уравнений (9) переходит в

$$\begin{cases} a_{t+1} = a_t + \dfrac{1}{b\,(t,\ x_t)}\,[x_{t+1} - a_t], \\ \gamma_t = 0. \end{cases} \tag{9′}$$

Таким образом, при выбранных начальных условиях (a_0, γ_0) получаем систему уравнений (9) для $b\,(t, x_t) = \sum\limits_i \alpha_i t^i$ и (9′) для $b\,(t, x_t) = \sum\limits_j l_j x_t^j$. Обе они определяют в явном виде зависимость процесса $\theta(t)$ от различных видов коэффициентов $b(t, x_t)$. Система уравнений (7′) вместе с предположением гауссовости условного распределения

$\psi\left[\theta(t)/x_0^t\right]$ определяет связь между процессами $\theta(t)$ и $x(t)$.

Возникает вопрос, не может ли эта связь быть выражена в одном стохастическом уравнении. Другими словами, нельзя ли информацию о процессах $\theta(t)$ и $x(t)$, которую несут в себе система уравнений (7′) и предположение гауссовости условного распределения $\psi\left[\theta(t)/x_0^t\right]$, т. е. системы уравнений (9) или (9′), записать более компактно в виде одного стохастического уравнения.

Ответ на этот вопрос является частным случаем решения следующей более общей задачи: по заданному случайному процессу $x(t, \omega)$ найти максимально простой по структуре случайный процесс $\tilde{\varepsilon}(t)$, который нес бы в себе ту же информацию о процессе $x(t, \omega)$, что и сам этот процесс. Поясним возможность построения такого процесса на следующем примере. Пусть $x(t, \omega)$ — дискретный случайный процесс с условными математическими ожиданиями $Mx_{t+1}/x_1^t = \varphi(x_1, x_2, \ldots, x_t)$ и условными дисперсиями $M[x_{t+1} - \varphi(x_1, \ldots, x_t)]^2/x_1^t = \sigma^2(x_1, \ldots, x_t)$. Случайную последовательность $\tilde{\varepsilon}(t)$ построим следующим образом:

$$\tilde{\varepsilon}_1 = Mx_1; \quad \tilde{\varepsilon}_2 = \frac{x_2 - \varphi(x_1)}{\sigma^2(x_1)}; \quad \tilde{\varepsilon}_3 = \frac{x_3 - \varphi(x_1, x_2)}{\sigma^2(x_1, x_2)} \text{ и т.д.}$$

Полученная последовательность довольно простая в том смысле, что $M\tilde{\varepsilon}_t = 0$, $M\tilde{\varepsilon}_t^2 = 1$, и в то же время для каждого момента времени t она несет в себе всю новую информацию о величине x_t по сравнению с ее предысторией x_0^t. По этому свойству нести в себе всю вновь возникающую информацию о некотором процессе $x(t, \omega)$ такие последовательности называются обновляющими. Применительно к нашей системе уравнений (7′) с учетом (9) и (9′) можно получить два вида обновляющих последовательностей $\tilde{\varepsilon}(t)$, для которых $M\tilde{\varepsilon}(t) = 0$ и $M\tilde{\varepsilon}^2(t) = 1$. Они имеют следующий вид:

$$\tilde{\varepsilon}(t+1) = \frac{x(t+1) - a_t}{\sqrt{b^2(t, x_t) + \gamma_t}} \text{ для случая } b(t, x_t) = \sum_i \alpha_i t^i$$

и $\tilde{\varepsilon}(t+1) = \dfrac{x(t+1) - a_t}{b(t, x_t)}$ для случая $b(t, x_t) = \sum_j l_j x_t^j$.

Решая эти равенства относительно $x(t+1)$, получим следующие стохастические уравнения диффузионного типа:

$$\begin{cases} x(t+1) = a_t + \sqrt{b^2(t, x_t) + \gamma_t}\,\tilde{\varepsilon}(t+1), \\ x(t+1) = a_t + b(t, x_t)\,\tilde{\varepsilon}(t+1). \end{cases} \tag{10}$$

В теории вероятностей доказывается, что для плотностей распределения двух случайных величин или двух случайных процессов в дискретном времени $\psi(\vec{x}_t)$ и $\psi(\vec{y}_t)$, где $\vec{x}_t = (x_1, \ldots, x_t)$; $\vec{y}_t = (y_1, \ldots, y_t)$, когда переменные $\vec{x}_t$ и $\vec{y}_t$ заданы на одной и той же области, при некоторых условиях существует их отношение в виде функции, заданной на той же области:

$$\psi(\vec{x}_t)/\psi(\vec{y}_t) = s(\vec{x}_t).$$

Если мы теперь найдем плотность распределения $\psi(\vec{z}_t)$ — знаменателя в производной Радона — Никодима, то плотность распределения $\psi(\vec{x}_t)$ отыщется в результате умножения $\psi(\vec{z}_t)$ на выражение производной Радона — Никодима. Плотности распределений процессов $z(t+1) = \sqrt{b^2(t, z_t) + \gamma_t} \cdot \tilde{\varepsilon}(t+1)$ и $z(t+1) = b(t, z_t)\tilde{\varepsilon}(t+1)$ легко находятся, так как в силу выбора коэффициентов диффузии $b(t, z_t)$ оба процесса оказываются в общем случае марковскими. Их плотности распределения имеют вид

$$\psi(\vec{t}) = \prod_{j=1}^{t} \left[2\pi\left(b_{j-1}^2 + \gamma_{j-1}\right)\right]^{-1/2} \exp\left\{-\sum_{j=1}^{t} x_j^2/2\left(b_{j-1}^2 + \gamma_{j-1}\right)\right\}$$

и

$$\psi(\vec{z}_t) = \prod_{j=1}^{t} \left[2\pi b_{j-1}^2\right]^{-1/2} \exp\left\{-\sum_{j=1}^{t} x_j^2/2b_{j-1}^2\right\}.$$

Умножая $\psi(\vec{z}_t)$ на выражения производной Радона — Никодима, получим соответственно два семейства плотностей распределений:

$$\psi(\vec{x}_t) = \prod_{j=1}^{t} \left[2\pi\left(b_{j-1}^2 + \gamma_{j-1}\right)\right]^{-1/2} \times$$

$$\times \exp\left\{-\sum_{j=1}^{t} \frac{(x_j - a_{j-1})^2}{2\left(b_{j-1}^2 + \gamma_{j-1}\right)}\right\}, \tag{11}$$

где $b_{j-1} = \sum_i \alpha_i (j-1)^i$ и a_j выражается через b_{j-1} по формуле (9),
и

$$\psi(\vec{x_t}) = \prod_{j=1}^{t} [2\pi(b_{j-1}^2)]^{-1/2} \exp\left\{-\sum_{j=1}^{t} \frac{(x_j - a_{j-1})^2}{2b_{j-1}^2}\right\}. \quad (11')$$

Отношение это называется производной Радона — Никодима, а условия, определяющие ее существование, заключаются в следующем: области значений переменных $\vec{x_t}$ и $\vec{y_t}$, для которых эти плотности вероятностного распределения обращаются в нуль, должны совпадать. Плотности распределения, для которых выполняется данное условие, называются эквивалентными. Надо отметить, что приведенное определение производной Радона — Никодима является весьма упрощенным, так как в общем случае понятие производной Радона — Никодима определено и для случайных процессов в непрерывном времени. Но для объяснения смысла дальнейших выкладок таким определением можно ограничиться. В работе [6] приводится выражение этой производной для двух случайных процессов, связанных стохастическими уравнениями типа (10).

Обозначим через $z(t+1)$ члены в уравнениях (10), определяющие разброс относительно среднего движения a_t: $z(t+1) = \sqrt{b^2(t, z_t) + \gamma_t}\,\widetilde{\varepsilon}(t+1)$ для случая, когда $b(t, z_t) = \sum_i \alpha_i t^i$ и $z(t+1) = b(t, z)\,\widetilde{\varepsilon}(t+1)$, когда $b(t, z_t) = \sum_j l_j z_t^j$. Производная Радона — Никодима для обоих случаев имеет вид

$$\frac{\psi(\vec{x_t})}{\psi(\vec{z_t})} = \exp\left\{\sum_{j=1}^{t} \frac{a_{j-1}x_j}{b_{j-1}^2 + \gamma_{j-1}} - \frac{1}{2}\frac{a_{j-1}^2}{b_{j-1}^2 + \gamma_{j-1}}\right\},$$

где $b_{j-1} = b(j-1,\ x_{j-1}) = \sum_i \alpha_i (j-1)^i$;

$$\frac{\psi(\vec{x_t})}{\psi(\vec{z_t})} = \exp\left\{\sum_{j=1}^{t} \frac{a_{j-1}x_j}{b_{j-1}^2} - \frac{1}{2}\frac{a_{j-1}^2}{b_{j-1}^2}\right\},$$

где $b_{j-1} = \sum_k l_k x_{j-1}^k$

и a_j выражается через b_{j-1} по формуле (9).

Выражения (11) и (11′) определяют для различных значений параметров $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m)$ и $\vec{l} = (l_1, \ldots, l_n)$ два семейства случайных процессов, т. е. два управляемых случайных процесса. Оставим без изменения вид выражений (11) и (11′), заменив только вид коэффициента диффузии так, чтобы он в одном случае оставался детерминированной функцией времени, а в другом был в каждый момент времени функцией только текущего состояния процесса x_t. Тогда можно получать различные виды управляемых случайных процессов. Отметим, что при таком выборе коэффициента диффузии выражение (11) всегда будет описывать семейство гауссовских случайных процессов, а (11′) — семейство существенно негауссовских процессов. Из полученного множества управляемых случайных процессов выберем тот, для которого сумма значений выражений (5′) по всем классам разбиения минимальна.

Определив вид управляемого случайного процесса, наилучшим образом описывающего совокупности реализаций в каждом классе разбиения, мы тем самым устанавливаем соответствие между значениями признаков разбиения для каждого класса и значениями параметров семейства плотностей распределений. В этом заключается последний этап, завершающий построение ЭСМ управляемого случайного процесса. Для определения вида этой закономерности могут использоваться различные формальные процедуры. В частности, рассматривая полученное соответствие как обучающую выборку, для ее распространения на некоторую область значений признаков разбиения и параметров управляемого случайного процесса можно применить байессовскую теорию принятия решений [7].

Подводя итог проведенному поэтапному построению ЭСМ управляемого случайного процесса, можно дать следующее (одно из возможных) ее определение. Под ЭСМ управляемого случайного процесса понимается формальное задание:

1) семейства плотностей совместных распределений компонент вектора $\vec{x_t}$, вид которого определяется выражениями, аналогичными (11) и (11′);

2) зависимостей коэффициентов переноса и диффузии от параметра управления;

3) соответствия между наборами признаков разбиения (характеризующих каждый класс разбиения) и значе-

ниями параметров семейства плотностей распределений;

4) статистической аппроксимации полученного соответствия на случай произвольных допустимых значений признаков разбиения.

ЛИТЕРАТУРА

1. **Аганбегян А. Г., Козлов Л. А., Кулешов В. В.** и др. Статистические модели в оптимальном отраслевом планировании. М., «Статистика», 1975. 191 с.
2. **Кушнер Г. Дж.** Статистическая устойчивость и управление. М., «Мир», 1969. 198 с.
3. **Гихман И. И., Скороход А. В.** Теория случайных процессов. Т. III. М., «Наука», 1975. 496 с.
4. **Розин Б. Б.** Проблемы экономико-статистического моделирования промышленного производства. Дис. на соиск. учен. степени докт. экон. наук. Новосибирск, 1975. 419 с.
5. **Кульбак С.** Теория информации и статистика. М., «Наука», 1967. 408 с.
6. **Липцер Р. М., Ширяев Л. Н.** Статистика случайных процессов. М., «Наука», 1974. 696 с.
7. **Дуда Р. О., Харт П. Е.** Распознавание образов и анализ сцен. М., «Мир», 1976. 511 с.

И. Г. БАГИРОВ

ПРОГНОЗНАЯ ОЦЕНКА ВЕРОЯТНОСТНЫХ ХАРАКТЕРИСТИК ПОТРЕБНОСТИ В ДЕФИЦИТНОМ РЕСУРСЕ

(на примере систем газоснабжения городов СССР)

Современная единая газоснабжающая система (ЕГС) СССР — типичный пример широкого класса экономических объектов, рассматриваемых как сложные и большие системы. В то же время поведение подсистем на всех уровнях ЕГС резко отличается от поведения прочих больших систем энергетики и народного хозяйства своей исключительной динамичностью. Это объясняется той ролью, которую играет газ как ресурс в общем энергоснабжении страны.

Удельный вес газовой промышленности в общем топливно-энергетическом балансе страны непрерывно повышается, и по крайней мере в обозримом будущем газ имеет тенденцию занять в этом балансе главное место. Поэтому проблема надежного функционирования всех звеньев ЕГС приобретает особую остроту, стимулируя поиск методов, позволяющих определять и учитывать степень возможных отклонений состояний подсистем ЕГС от запланированных траекторий.

Одним из таких методов является построение ЭСМ, обеспечивающих получение статистической оценки возможных отклонений прогнозируемого экономического показателя от запланированных его значений.

В нашей предыдущей статье приводится общая схема построения ЭСМ, аппроксимирующей с некоторой погрешностью управляемую случайную последовательность. Она, в частности, применима для описания изменений величины потребляемого ресурса в каждом планируемом отрезке времени. Непосредственное использование предложенной схемы для прогнозирования размеров потребле-

ния рассматриваемого ресурса предполагает сопоставимость плановых величин потребления ресурса с фактическими в смысле учета при принятии плановых решений всех основных факторов, влияющих на величину потребления. Применительно к дефицитному ресурсу это условие зачастую не выполняется.

Дефицитность газа обусловливается опережающим по сравнению с темпами его добычи ростом спроса на этот вид топлива, обладающий значительным преимуществом по сравнению с другими видами энергоресурсов. В настоящее время дефицитность газа — объективно сложившийся факт, и он влияет на распределение газа на всех уровнях ЕГС. Несмотря на то, что дефицитность может вызвать довольно значительные отклонения фактически потребляемых объемов газа от запланированных, это совершенно не учитывается в существующей методике проектирования газоснабжающих систем для городов СССР.

Специфика задач управления указанными системами состоит в том, что перспективные планы их развития разрабатываются исключительно в составе генеральных планов развития городов. Современный город является очень сложной и динамичной системой, изменения состояний которой носят ярко выраженный стохастический характер. Игнорирование этого факта также необходимо отнести к основным недостаткам сложившейся методики плановых расчетов для газоснабжающих систем городов СССР, делающим во многом несостоятельными предварительные расчеты, полученные без учета возможных отклонений траектории развития города от запланированной в генеральном плане. В результате важнейшая характеристика системы газоснабжения — надежность — полностью зависит от уровня профессионального опыта разработчиков проекта такой системы.

Настоящая работа преследует две цели: во-первых, распространить описанный в предыдущей статье метод построения управляемой случайной последовательности на ситуацию, когда прогнозируемый ресурс оказывается дефицитным; во-вторых, продемонстрировать применение предлагаемого в той же статье подхода на конкретном статистическом материале, собранном по проектам газоснабжающих систем городов СССР.

Выходным параметром систем городского газоснабжения является величина фактического потребления городом

газа. Обозначим его значение в некоторый момент времени t через $x^*(t)$; входные параметры, определяющие уровень потребления городом газа,— через $\vec{y}_{m+n} = (y_1, y_2, \ldots, y_m, \ldots, y_{m+n})$; величину газа, необходимую для полного удовлетворения потребности города в газе в момент времени t,— через $x(t)$. Дефицитность газа означает, что $x^*(t) \leqslant x(t)$ для всех t.

В территориальном планировании подсистема городского газоснабжения выступает одним из низовых звеньев иерархической структуры ЕГС. Верхним уровнем для этой подсистемы будет область. Пусть $N^*_j(t)$ — величина фактического потребления газа j-й областью в момент времени t; $N_j(t)$ — величина потребности в газе j-й области. Теперь можно записать:

$$N^*_j(t) = \sum_{i=1}^{c_j} x^*_{ij}(t); \quad N_j(t) = \sum_{i=1}^{c_j} x_{ij}(t),$$

где c_j — число городов в j-й области; x^*_{ij} и x_{ij} — величины фактического потребления газа и потребности в нем i-го города j-й области.

Распределение фактически потребляемого количества газа управляющей системой верхнего уровня (область) между подсистемами нижнего уровня (городами, входящими в область) в формальном отношении можно рассматривать как преобразование некоторой системой значений входных параметров $[x_{ij}(t), x_{i_1 j}(t), \ldots, x_{i_k j}(t), N_j(t), N^*_j(t)]$, в значения $x^*_{ij}(t)$ выходного параметра, т. е.

$$x^*_{ij}(t) = \hat{L}[x_{ij}(t), x_{i_1 j}(t), \ldots, x_{i_k j}(t), N_j(t), N^*_j(t)], \qquad (1)$$

где $\hat{L}$ — оператор преобразования; $k = k(i)$ — функция от i, показывающая число городов, потребности которых в газе влияют на величину отпускаемого i-му городу газа; $(i_1, i_2, \ldots, i_k)$ — индексы этих городов.

Множества индексов $(i_1, i_2, \ldots, i_k)$ определяют для разных i множество подсистем j-й системы (область), элементы которых связаны между собой. Для определения вероятностных характеристик выходного параметра $x^*(t)$ необходимо знать вид выражения (1).

В настоящей работе будет рассмотрен следующий вид распределения количества газа $N^*_j(t)$ между городами,

входящими в j-ю область. Предполагается, что все города, входящие в j-ю область, можно разделить на две однородные группы по степени удовлетворения их потребностей в газе. Одну группу составляют города, потребности которых в газе удовлетворяются полностью. Оставшееся в распоряжении верхней подсистемы количество газа распределяется между городами второй группы пропорционально их потребностям.

В силу такого предположения выражение (1) примет вид

$$x^*_{ij}(t) = \frac{N^*_j(t) - \sum\limits_{i_k} x_{i_k j}(t)}{N_j(t) - \sum\limits_{i_k} x_{i_k j}(t)},$$

где i_k — индексы городов, потребности в газе которых удовлетворяются полностью; i — номер города, принадлежащего ко второй группе городов.

Обозначим $\sum\limits_{i_k} x_{i_k j}(t)$ через $x_j(t)$ и $\dfrac{N^*_j(t) - \sum\limits_{i_k} x_{i_k j}(t)}{N_j(t) - \sum\limits_{i_k} x_{i_k j}(t)}$ — через $\varkappa_j(t)$. Очевидно, $0 \leqslant \varkappa_j(t) \leqslant 1$.
Теперь можно записать:

$$x^*_{ij}(t) = \varkappa_j(t)\, x_{ij}(t). \qquad (2)$$

Для удобства изложения опустим в дальнейших выкладках индекс области j.

Предположим также, что в территориальном отношении иерархическую структуру ЕГС составляют два уровня — область и город. Сделанное предположение не меняет существа дела, так как все дальнейшие рассуждения легко обобщаются на любое число уровней. Тогда случайные последовательности $N^*_j(t)$ и $N_j(t)$ будут слабо зависимыми, и их можно считать некоррелированными с плотностями совместного распределения, принадлежащими одному из вероятностных семейств (11)* или (11′)*, для случая, когда функции $b(t, x_t)$ имеют вид (8)* или (8′)*. (Здесь и далее номер выражения со знаком * означает ссылку на соответствующую формулу нашей предыдущей статьи.)

Пусть $\psi_{x_i}(x_1, x_2, \ldots, x_t) = \psi_x(\vec{x}_t)$; $\psi_N(N_1, N_2, \ldots, N_t) = \psi_N(\vec{N}_t)$; $\psi_{N^*}(N^*_1, N^*_2, \ldots, N^*_t) = \psi_{N^*}(\vec{N}^*_t)$ — плотности совместных распределений компонент случайных векторов $\vec{x}_{it}$, $\vec{N}_t$, $\vec{N}^*_t$ соответственно, вид которых нам известен. (Во избежание путаницы, чтобы не вводить новые переменные значений рассматриваемых в дальнейшем случайных величин, в качестве их обозначений взяты символы самих случайных величин.)

Найдем вначале выражение для плотности совместного распределения компонент вектора $\vec{\varkappa}_t = (\vec{\varkappa}_1, \varkappa_2, \ldots, \vec{\varkappa}_t)$, считая известными $\psi_{x_i}(\vec{x}_t)$; $\psi_N(\vec{N}_t)$; $\psi_{N^*}(\vec{N}^*_t)$. Обозначим через $\vec{x}^0_t$ произвольное фиксированное значение вектора $\vec{x}_t$, а разности векторов $\vec{N}^*_t - \vec{x}^0_t$ и $\vec{N}_t - \vec{x}^0_t$ — через $\vec{s}^1_t$ и $\vec{s}^2_t$ соответственно. Под разностью двух векторов понимается вектор, компонента которого равна разности соответствующих компонент этих векторов.

Функция распределения $P_{\varkappa}(\vec{\varkappa}_t)$ равна

$$P_{\varkappa}(\vec{\varkappa}_t) = \int\limits_{\vec{\varkappa}_t} P\left(\frac{s^1_1}{s^2_1} < \varkappa_1, \ldots, \frac{s^1_t}{s^2_t} < \varkappa_t\right) \times \psi_x(\vec{x}^0_t)\, dx^0_1, \ldots, dx^0_t.$$

Функция распределения $P\left(\dfrac{\vec{s}^1_t}{\vec{s}^2_t} < \vec{\varkappa}_t\right)$ имеет вид

$$P\left(\frac{\vec{s}^1_t}{\vec{s}^1_t} < \vec{\varkappa}_t\right) = \int\limits_{\vec{s}^2_t} P\left(s^1_1 < \varkappa_1 s^2_1, \ldots, s^1_t < \varkappa_t s^2_t\right) \times \psi_N(\vec{s}^2_t + \vec{x}^0_t)\, ds^1_1, \ldots, ds^2_t,$$

где под $\vec{s}^2_t$ понимается произвольное фиксированное значение случайного вектора $\vec{s}^2_t$. Тогда для функции $\psi_{\varkappa}(\vec{\varkappa}_t/\vec{x}^0_t)$ имеем выражение

$$\psi_{\varkappa}(\vec{\varkappa}_t/\vec{x}^0_t) = \int\limits_{\vec{s}^2_t} \psi_{N^*}(s^2_1\varkappa_1 + x^0_1, \ldots, s^2_t\varkappa_t + x^0_t) \times s^2_1, \ldots, s^2_t \psi_N(\vec{s}^2_t + \vec{x}^0_t)\, ds^2_1, \ldots, ds^2_t.$$

Принимая во внимание, что $\vec{x}_t^0$ и $\vec{s}_t^2$ — произвольные фиксированные значения случайных векторов $\vec{x}_+$ и $\vec{s}_t^1$, окончательно получим

$$\psi_{\varkappa}(\vec{\varkappa}_t) = \int\limits_{\vec{x}_t} \psi_x(\vec{x}_t)\, dx_1, \ldots, dx_t \int\limits_{\vec{s}_t^2} \psi_{N^*}(s_1^2 \varkappa_1 + x_2, \ldots, s_t^2 \varkappa_t + x_t), \ldots, s_1^2, \ldots, s_t^1 \times \psi_N(\vec{s}_t^2 + \vec{x}_t)\, ds_1^2, \ldots, ds_t^2. \quad (3)$$

Зная $\psi_{\varkappa} = (\vec{\varkappa}_t)$, найдем зависимость между плотностями совместного распределения $\psi_{x_i^*}(\vec{x}_t^*)$ и $\psi_{x_i}(\vec{x}_t)$, компонент векторов $\vec{x}_{it}^*$ и $\vec{x}_{it}$, связанных соотношением (2):

$$\psi_{x_i^*}(\vec{x}_t^*) = \int\limits_{\vec{\varkappa}_t} \psi_{\varkappa}(\vec{\varkappa}_t)\, \psi_{x_i}\left(\frac{x_1}{\varkappa_1}, \ldots, \frac{x_t}{\varkappa_t}\right) \times \frac{1}{\varkappa_1}, \ldots, \frac{1}{\varkappa_t}\, d\varkappa_1, \ldots, d\varkappa_t, \quad (4)$$

причем $\varkappa(t)$ предполагаются строго положительными для всех t. При выводе формул (3) и (4) мы исходили из того, что вид функций $\psi_{\varkappa}(\vec{x}_t)$, $\psi_N(\vec{N}_t)$, $\psi_{N^*}(\vec{N}_t^*)$ нам известен. Покажем теперь, как находить вид этих функций с тем, чтобы выполнить вторую цель настоящей работы — поэтапно продемонстрировать получение обучающей выборки для построения ЭСМ управляемой случайной последовательности. Предварительно нам необходимо описать факторы $\vec{y}_{m+n}$, влияющие на величину потребности городов СССР в газе.

В системе городского газоснабжения отчетливо прослеживаются три подсистемы, несущие различные целевые нагрузки по потребителям: промышленное, индивидуально-бытовое и коммунально-бытовое газоснабжение. Все три подсистемы имеют различные технические характеристики, обусловленные спецификой их целевого назначения. Наибольший удельный вес в общей системе городского газоснабжения принадлежит подсистеме промышленного газоснабжения, составляющей, как правило, не менее 80% от общегородского потребления газа. Подсистемы индивидуально-бытового и коммунально-бытового газоснабжения целесообразно рассматривать как одну подсистему, обслуживающую все городское население. Следовательно, множество факторов $\vec{y}_{m+n}$ можно разбить на две группы $\vec{y}_m$ и $\vec{y}_n$, одна из которых определяет количество газа, необходимое населению города для его индивидуальных и коммунально-бытовых нужд, другая — для удовлетворения нужд промышленных предприятий города.

Анализ существующей методики проектирования городских систем газоснабжения позволяет выделить следующие факторы $\vec{y}_m$ (где $m = 5$ в нашем случае), определяющие потребность населения города в газе на его индивидуально-бытовые и коммунально-бытовые нужды:

y_1 — численность населения города;

y_2 — структура этажности городских зданий;

y_3 — доля участия ТЭЦ и ГРЭС в общегородском горячем водоснабжении;

y_4 — уровень развития города, выражающийся в количестве коммунально-бытовых услуг на душу населения;

y_5 — количество отопительных дней в году или средняя температура отопительного сезона.

Описание ведущей подсистемы городского газоснабжения — промышленного газоснабжения — представляет собой совокупность регрессионных уравнений зависимости объема потребляемого газа от влияющих на него факторов на предприятиях газоемких отраслей, представленных в данном городе. Построение каждого такого уравнения является довольно сложной и трудоемкой задачей, имеющей значительный самостоятельный интерес, в первую очередь для широкого круга проблем оперативного управления технологическими процессами на промышленном предприятии.

В данном случае задача сильно упрощается, так как с небольшой по отношению к общей величине потребления газа погрешностью можно ограничиться, как правило, двумя-тремя ведущими факторами из всей совокупности факторов, входящих в регрессионное уравнение. Например, для предприятий взятой в качестве иллюстрации типичной газоемкой отрасли — цементной промышлен-

ности — такими ведущими факторами $\vec{y}_n$ ($n = 3$ в нашем случае) будут:

$\vec{y}_6$ — объемы выпуска различных марок цемента;
y_7 — способ производства (мокрый или сухой);
y_8 — влажность шлама.

Наконец, если t пробегает достаточно длинные отрезки времени внутри планового периода, то можно ограничиться объемами продукции, выпускаемой на предприятиях газоемких отраслей.

Определим плотность вероятностного распределения $\psi_{x_i}(\vec{x}_t)$ компонент вектора $\vec{x}_{it} = (x_{i1}, \ldots, x_{it})$ для случаев, когда $\vec{x}_{it}$ — вектор величин потребностей населения i-го города в газе и $\vec{x}_{it}$ — вектор величин потребностей цементной промышленности i-го города в газе. С этой целью разобьем оба множества реализаций $\vec{x}_{it}$, $i = \overline{1, c_j}$; $j = \overline{1, R}$ по всем городам страны на подмножества с помощью алгоритма таксономии «Рельеф» отдельно для каждой подсистемы городского газоснабжения в пространстве признаков

$$[x_{0,}\vec{y}_{1t}, \ldots, \vec{y}_{mt}]$$

и $[x_{0,}\vec{y}_{m+1,t}, \ldots, \vec{y}_{m+n,t}]$,

где $(\vec{y}_{1t}, \ldots, \vec{y}_{mt})$ и $(\vec{y}_{m+1,\ t}, \ldots, \vec{y}_{m+n,\ t})$ — плановые показатели описанных выше входных параметров обеих подсистем на всем протяжении рассматриваемого отрезка времени. Реализации x_{it}, попавшие в один класс, можно считать реализациями одного случайного процесса.

В качестве возможных вариантов вероятностного описания полученного набора классов рассматриваются семейства плотностей вероятностного распределения типа (11)* и (11′)* при условии, что вид коэффициента диффузии в этих формулах определяется выражениями типа (8)* и (8′)*. Для выбора из них семейства вероятностных распределений, описывающего полученные классы, сравниваются суммы выражений (5′)* для каждого семейства вероятностных распределений, найденных по всем классам разбиения. Семейство вероятностных распределений, для которого эта величина наименьшая, выбирается в качестве описания наборов классов разбиения.

Л. Я. ШКРАБИН

ОБ ИНТЕРПРЕТАЦИИ ВЕКТОРНОЙ ОПТИМИЗАЦИИ В ТЕРМИНАХ МНОГОМЕРНОЙ КЛАССИФИКАЦИИ НАБЛЮДЕНИЙ

На возможность применения статистического подхода к решению методических вопросов, связанных с векторной оптимизацией, указывает существенная стохастичность внешних воздействий на производственную систему, ее реакций на фиксированные входы и неопределенный характер ее внутренних элементов при фиксированных входе и выходе.

До утверждения плановых заданий и лимитов на ограниченные ресурсы отраслевым органам управления необходимо иметь не один вариант оптимизированного плана, а «библиотеку» вариантов, отвечающую достаточно широкой и глубоко дифференцированной по различным показателям области вариации состояний и траекторий развития отрасли. Такая необходимость обусловлена ориентировочным характером целей, ограничений и критериев оценки эффективности производственной деятельности, устанавливаемых на народнохозяйственном уровне в форме контрольных цифр, систем показателей учета и отчетности. В процессе согласования отраслевых и региональных проектировок на народнохозяйственном уровне, как правило, контрольные цифры становятся более конкретными за счет учета и снижения противоречий между отраслями и регионами и более точного обоснования глобальных целей, реализуемых народным хозяйством.

В процессе итеративного, по существу, согласования уровней отраслевых показателей с народнохозяйственными целями и средствами конкретизируются цели отрасли, ее возможности и необходимые для их реализации

уровни общесистемных ресурсов. При этом отраслевая система исследуется в различных аспектах, выясняется противоречивость отраслевой самоориентации и внешних ориентиров ее функционирования и развития. В предположении полного и всестороннего согласования входов и выходов отраслевой системы в утвержденном отрасли плане эффективность деятельности отрасли может быть измерена по степени превышения или недовыполнения «эталонных» (плановых) уровней выходных показателей, а мера объективных благоприятствований или препятствий выполнению плана — по степени достижения уровней входных показателей.

Утверждаемый отрасли план, как правило, не полностью соответствует действительно необходимым уровням народнохозяйственных потребностей и выделяемым отрасли ресурсам. Итеративный процесс их согласования обрывается на лучшем на данную дату варианте, а не по достижении наилучшего плана (в силу дефицита времени, недостаточно высокой скорости процесса согласования и ограниченности средств, выделяемых на предплановые разработки, а также ограниченности методологических, методических и технических средств исследования). Поэтому согласование деятельности отрасли и народнохозяйственных целей продолжается по мере раскрытия в процессе реализации плана не полностью согласованных взаимных требований, неопределенных моментов плана, а также по мере обнаружения неточностей и технических погрешностей в расчетах.

Это согласование в ходе реализации планов выступает, с одной стороны, в форме заявок отраслевых организаций в вышестоящие органы управления на дополнительные объемы трудовых, материальных и финансовых ресурсов и просьб (с соответствующими обоснованиями) об уменьшении объемов выпуска продукции, изменении номенклатуры выпуска и т. д. С другой стороны, органы управления народным хозяйством корректируют лимиты на ресурсы (как правило, в сторону их уменьшения), объемы выпуска (в сторону увеличения), качественные характеристики продукции и услуг отрасли (повышение их уровней), номенклатуру (увеличение доли высококачественной и трудоемкой продукции). Для такой корректировки необходимо иметь на народнохозяйственном уровне информацию о возможных объемах и вариантах номенклатуры выпуска отраслевой системы, соответствующих изменившимся условиям ее обеспечения общесистемными ресурсами. Причем эта информация должна поступать в оперативном режиме или наличествовать в форме «библиотеки» вариантов (функций или номограмм), показывающих реакцию отрасли на изменение ее входов.

С позиций органов управления отраслевой системой доплановое и послеплановое согласование — это источники вариации экзогенных переменных, неопределенности воздействий внешней среды. Поэтому необходимость увеличения числа и разнообразия вариантов оптимизированных планов отрасли, различающихся уровнями и структурами капиталовложений, трудовых ресурсов, народнохозяйственной потребности, обусловлена не только задачей совершенствования качества управления народным хозяйством в целом. Отраслевые органы управления должны быть готовы к изменению внешних условий функционирования и развития отрасли в информационном аспекте, чтобы иметь возможность практически сразу после сигнала об изменении этих условий приступить к осуществлению изменений в отраслевой системе, адекватных новым условиям.

Полагаем, что не требует доказательства утверждение о том, что поведение отраслевой системы в принципе относится к высшему классу целенаправленных поведений в классификации А. Розенблюта, Н. Винера и Д. Бигелоу [17], а именно: оно должно быть активным, с обратной связью, с высшим порядком предсказания, обусловленным целенаправленным поведением. Данное утверждение имеет практическое обоснование. Отраслевые организации-заказчики оптимизационных расчетов многократно обращаются к организациям-исполнителям с просьбами о расчете дополнительных оптимизированных вариантов плана развития и размещения предприятий отрасли при различных сочетаниях входных и выходных параметров отрасли, с измененным или другого рода критерием оптимизации. Совместно с сотрудниками брянского ГПИстроймаша В. А. Бажановым разработана «библиотека» вариантов развития предприятий отрасли строительно-дорожного и коммунального машиностроения. Ее опробование показало, что диалог по согласованию вариантов строительства или реконструкции предприятий между проектным институтом и министерством стал существенно менее

трудоемким, более плодотворным (так как сравнению могло быть подвергнуто большее число вариантов) и требует меньших затрат календарного времени. На наш взгляд, «библиотека» вариантов развития отрасли могла бы, как минимум, дать такие же преимущества в диалоге с народнохозяйственными органами управления. Еще более плодотворным представляется нам подход, заключающийся в статистической аппроксимации уровней выходных показателей системы нижнего уровня как функций входных показателей [22, гл. 12, 13]. Как считают авторы работы [22], данный подход может быть более полно реализован при следующем условии. Связь между входами и выходами определяется не только по ретроспективной информации, но и по данным о вероятных состояниях системы, дополняемым требованием оптимального отклика на данную структуру и интенсивность входа.

Действительно, активность отраслевой системы двоякая. Поэтому тщательному исследованию и экстраполяции подлежат не только вероятные сочетания и уровни входных показателей (целевые и ресурсные), но и наилучшие формы организации отраслевой системы, позволяющие реализовывать задаваемые уровни выходных показателей при минимальных уровнях входных (затраты лимитированных ресурсов). Очевидно, что число способов реализации данных уровней выходных показателей при фиксированных уровнях входных показателей (если реализуем хотя бы один способ) с увеличением сложности системы возрастает.

На стадии планирования, а также в процессе обоснования необходимых корректировок реализуемого плана все чаще используются статистические методы. Соответственно растет число работ о методах оптимизации систем. В них рассматриваются проблемы оптимального и субоптимального анализа и синтеза нелинейных систем [15], учета неопределенности при принятии решений в больших системах [4, 8, 10, 19, 20] и особых поведенческих характеристик управляемой и управляющей систем (надежность, устойчивость, маневренность и т. д.) [20], рациональной организации информационного обмена при согласовании многоуровневых систем [16, 22], а также оптимальной аппроксимации сложных систем более простыми с целью обеспечения информационной и вычислительной разрешимости задач планирования [2, 12, 16 и др.].

В работах [5—7, 9, 14, 16, 21], посвященных проблеме векторной оптимизации, методы оценки эффективности альтернативных вариантов плана рассматриваются как методы принятия оптимального решения в условиях компромисса. В данной работе предлагается указанную проблему интерпретировать следующим образом:

на первом этапе, исследуя субоптимальные уровни целевых показателей-компонент вектора цели $\{L_k^*\}$, формировать эталонное выражение вектора $\{L_1^*, \ldots, L_k^*, \ldots, L_K^*\}$;

на втором формировать скалярное обращение вектора-цели;

на третьем выбор оптимального плана предлагается осуществлять при заданной структуре ограничений и потребности;

на четвертом этапе проводить анализ и обобщение результатов расчетов, выполненных для различных вариантов структуры и уровней ограниченных ресурсов и потребностей.

Основное внимание уделяется второму этапу. Предполагается, что на первом этапе найдены субоптимальные уровни по каждому целевому показателю и из них сформирован недостижимый для реализации идеал вектора-цели, все компоненты которого равны численным значениям субоптимальных уровней компонент. При построении множества эквивалентных в смысле какой-либо статистической меры многомерно-оптимальных планов предполагается использовать методы многомерной классификации наблюдений, например распознавания образов [1, 18]. В случае евклидовой меры близости допустимого варианта плана эталону векторная оптимизация может быть заменена скалярной со следующей формой функционала:

$$F_1(x) = \sum_{k=1}^{K} \left(\frac{L_k(x) - L_k^*}{L_k^*} \right)^2 \to \min. \qquad (1)$$

При линейных ограничениях на переменные в неотрицательном квадранте с функционалом в форме (1) задача векторной оптимизации преобразуется в задачу квадратического программирования. Для того чтобы остаться в рамках линейного программирования, предлагается использовать меру близости, аналогичную средней ошибке

аппроксимации [1]:

$$F_2(x) = \sum_{k=1}^{\widehat{K}} \frac{L_k^* - L_k(x)}{L_k^*} + \sum_{k=K+1}^{K} \frac{L_k(x) - L_k^*}{L_k^*} \to \min. \qquad (2)$$

Она не вполне симметрична по отношению минимизируемых компонент к максимизируемым. В максимизируемых компонентах абсолютные отклонения отнесены к максимальному в своем ряду значению, а в минимизируемых — к минимальному. Это ведет к систематическому увеличению веса суммы относительных отклонений по минимизируемым компонентам по сравнению с максимизируемыми. При одном порядке в среднем минимизируемых и максимизируемых компонент указанное различие сглаживается. В случаях, когда среди абсолютных величин компонент различия в среднем велики, необходимо в знаменателях брать оптимальные, а не субоптимальные уровни целевых показателей. Решения по таким функционалам могут быть получены методом последовательных приближений. Для этого на второй и последующих итерациях необходимо в знаменателях поставить вместо субоптимальных уровней оптимальные, соответствующие предыдущему шагу.

Очевидно, что в двух оптимальных планах, имеющих одинаковые уровни меры F_1 или F_2, индивидуальные отклонения, соответствующие одной и той же компоненте, могут значительно различаться. Кроме того, в одном и том же оптимальном плане разным компонентам $L_k(x)$ могут соответствовать различные уровни относительных отклонений от принятого эталона. От этих различий избавляются методами параметрического программирования: введением переменных параметров в целевую функцию или в ограничения.

Дополнительным источником информации, позволяющим уточнить меру близости с точки зрения ее селективности, являются данные о поведении целевых показателей в допустимой области или в окрестностях субоптимальных состояний. Если уровень целевого показателя относительно мало убывает или возрастает в одном из этих множеств,

[1] $\{1, \ldots, \widehat{K}\}$ — индексы максимизируемых и $\{\widehat{K}+1, \ldots, K\}$ — минимизируемых целевых показателей.

то он полог и в меньшей степени препятствует переходу к другим альтернативам, более эффективным с точки зрения других целевых показателей, являющихся менее пологими. Когда мера пологости целевых показателей построена, величины, обратные к пологостям, можно ввести штрафными весами в меры (1) и (2).

Мерой пологости может служить среднеквадратическое отклонение уровней целевого показателя по всем элементам множества от субоптимального уровня (хотя теоретически точнее было бы измерять пологость в точке многомерного оптимума и в его окрестности). С учетом нормировки на величину субоптимального уровня и различий между максимизируемыми и минимизируемыми уровнями указанные величины таковы:

для минимизируемых компонент вектора-цели

$$\mu L_k^* = \frac{\sigma\left[L_k^*(x)\right]}{\max\limits_{x \in X}\{L_k(x)\} - L_k^*}; \qquad (3)$$

для максимизируемых компонент вектора-цели

$$\mu L_k^* = \frac{\sigma\left[L_k^*(x)\right]}{L_k^* - \min\limits_{x \in X}\{L_k(x)\}}, \qquad (4)$$

где x — элемент множества вариантов плана в исследуемой окрестности субоптимума.

Среднеквадратические отклонения в выражениях (3) и (4) могут быть заменены средним арифметическим отклонением, так как нет необходимости заботиться о знаках отдельных отклонений.

Как показано О. Девисом и Э. Пирсоном [3], при малом числе наблюдений среднеквадратическое отклонение достаточно хорошо оценивается размахом. Замена в (3) и (4) среднеквадратических отклонений размахами, очевидно, устраняет эффект взвешивания по пологостям, так как все μL_k^* становятся равными единице. Этот факт настораживает и в отношении применимости среднеквадратического отклонения: оно действительно тяготеет к более строгому учету бо́льших индивидуальных отклонений, а нас интересует поведение меньших индивидуальных отклонений в окрестности субоптимума. Последнее в боль-

шей мере удовлетворяется применением средней взвешенной гармонической $\Delta_{\Gamma}[L_k^*(x)]$. И более правильно в качестве μL_k^* брать величины, прямо пропорциональные средним взвешенным гармоническим и обратно пропорциональные среднеквадратическим отклонений от соответствующих субоптимумов.

Очевидно, что вышеуказанное взвешивание компонент мер (1) и (2) приближает к достижению равных относительных отклонений оптимальных уровней целевых показателей от субоптимальных и, следовательно, может служить альтернативой их параметрическому уравниванию. Необходимость дополнительного взвешивания, учитывающего иерархию и приоритеты частных целей, нами не отвергается. Однако прежде необходимо учесть разные уровни пологостей частных функционалов, соответствующих этим целям.

В ряде работ [6, 14 и др.] исследован вопрос о недостаточности отнесения абсолютных отклонений в мерах (1) и (2) к оптимальным или субоптимальным уровням целевых показателей и доказывается, что более правомерно относить их к среднеквадратическим отклонениям или к размахам в окрестности субоптимума. Нетрудно показать, что в этом случае будет минимизироваться не величина относительного отклонения оптимума от субоптимумов, а средняя пологость функционалов. При этом компоненты вектора-цели в составном функционале утратят соизмеримость. Действительно, $1/\widehat{K}\sum_{k=1}^{\widehat{K}}(L_k^* - L_k)\Big/\sigma\{L_k^*(x)\}$ измеряет среднюю по всем K показателям пологость максимизируемых показателей в интервале между оптимумом и субоптимумами.

В общем случае для исследования пологости целевых показателей необходимо проводить выборочные статистические исследования, так как сплошное обследование области допустимых планов нецелесообразно и технически малоосуществимо.

По отрасли комбайностроения нами совместно с И. И. Селиным проведен поиск векторно-оптимизированных планов с использованием меры (2), модифицированной штрафными весами, которые учитывают пологость функционалов. Предпочтительность данной меры, на наш взгляд, заключается в геометрической ясности ее конструкции и результатов расчетов. В дальнейшем предполагается исследовать сравнительную ценность описанного метода векторной оптимизации.

После того как данная работа была сдана в печать, опубликована монография О. Д. Проценко и Д. И. Соловейчика [23], которые в анализе долговременных хозяйственных связей поставщиков и потребителей используют подобные же подходы к векторной оптимизации. Надо отметить, что эти авторы не вводят понятия пологости функционала в окрестности оптимального плана, хотя используемое ими понятие ценности различных критериев может быть развито до понятия крутизны частного критерия в окрестности субоптимума или в окрестности плана, равноотстоящего в пространстве, выделяемом частными критериями, от субоптимумов. При учете пологости констатируется меньшая привязанность к данному плану по критерию, который более полог в его окрестности, и необходимость строгого слежения за поведением более крутых критериев в этой же окрестности. С точки зрения учета неопределенности в задачах отраслевого оптимального планирования эти вопросы рассмотрены нами в работе [24]. В ней показано, что ценность, т. е. внешне устанавливаемые приоритеты частных критериев, присваивается в процессе согласования с системами более высокого уровня.

Авторы работы [23] отмечают необходимость анализа функциональной взаимозависимости критериев: «Критерий минимума времени работы оборудования на выполнение портфеля заказов является подчиненным, поскольку высвобождение времени призвано обеспечить получение дополнительной продукции с имеющихся мощностей. Анализ планов прикрепления показывает, что при максимизации по этому критерию высвобождается время, как правило, на малопроизводительных станках, при этом не достигается максимум выпуска продукции. Ввиду этого дальнейшее рассмотрение данного критерия нецелесообразно» (с. 99). Такие взаимосвязи критериев не всегда обнаруживаются методами причинно-следственного анализа, поэтому на выборке планов из окрестностей, близких к субоптимумам, целесообразно исследовать коррелируемость критериев.

В подобном анализе может быть с незначительными изменениями использован аппарат корреляционного ана-

лиза, например выявления групп мультиколлинеарно связанных критериев [2]. В работе [23] для выявления связи между множеством значений критериев и множеством весовых коэффициентов построены уравнения регрессии на случайной выборке планов из допустимой области (методом Монте-Карло) и в схеме планирования эксперимента (под названием метода симплекс-решетчатого планирования).

Ранее Л. Билиньский [25] на основе теоретико-множественных рассуждений классифицировал всевозможные соотношения пар целей: идентичные (реализация цели А влечет реализацию цели В и, напротив, реализация цели В влечет реализацию цели А), находящиеся в отношении доминирования или следования (реализация подчиненной цели А влечет частичную реализацию цели В, в то время как реализация доминирующей цели В влечет реализацию цели А), частично расходящиеся (реализация цели А влечет частичную реализацию цели В, в свою очередь, реализация цели В влечет лишь частичную реализацию цели А), взаимоисключающие. После установления этой классификации Л. Билиньский переходит к рассмотрению согласования расходящихся и взаимоисключающих целей в терминах оптимума Парето. На наш взгляд, важной является проблема выработки практических подходов к классификации компонент вектора-цели в конкретных задачах многомерной оптимизации. На этом пути наиболее плодотворными представляются методы математической статистики и многомерной классификации.

ЛИТЕРАТУРА

1. **Айвазян С. А., Бежаева З. И., Староверов О. В.** Классификация многомерных наблюдений. М., «Статистика», 1974. 240 с.
2. **Бажанов В. А.** Об одном способе формирования исходной информации для задач оптимизации отраслевых планов.— В кн.: Оптимизационные и экономико-статистические методы в перспективном отраслевом планировании. Новосибирск, «Наука», 1975, с. 60—66.
3. **Боярский Э. А.** Порядковые статистики. М., «Статистика», 1973. 109 с.
4. **Вагнер Г.** Основы исследования операций, т. 1. М., «Мир», 1972. 335 с.

[2] Соответствующие преобразования в парном коэффициенте корреляции показаны в работе [24].

5. **Вилкас Э. И.** Векторные критерии оптимальности.— В кн.: Математика и кибернетика в экономике. М., «Экономика», 1975, с. 66—68.
6. **Волкович В. Л.** Многокритериальные задачи и методы их решения.— В кн.: Кибернетика и вычислительная техника. Вып. 1. Киев, 1969, с. 44—52.
7. **Волкович В. Л.** Проблема многокритериальности.— В кн.: Энциклопедия кибернетики, т. 2. Киев, УСЭ, 1975, с. 11—12.
8. **Вопросы** построения информационных систем управления развитием электроэнергетических систем. Вып. 1. Учет неопределенности исходной информации. Иркутск, 1973. 162 с.
9. **Добровольский В. К.** Экономико-математическое моделирование. Вопросы методологии. Киев, «Наукова думка», 1975. 183 с.
10. **Ермольев Ю. М.** Методы стохастического программирования. М., «Наука», 1976. 239 с.
11. **Козырев В. П.** Аксиоматическое определение меры близости матриц некоторого класса.— В кн.: Исследование по дискретной математике. М., «Наука», 1973, с. 200—213.
12. **Лоран П. Ж.** Аппроксимация и оптимизация. М., «Мир», 1975. 496 с.
13. **Макроэкономические** модели для планирования и определения экономической политики. Женева, Секретариат Европейской экономической комиссии, 1967. 234 с.
14. **Онищенко А. М.** Критерии оптимизации сельскохозяйственного производства и методы нахождения наиболее эффективных планов по нескольким критериям. Киев, 1970. 192 с.
15. **Оптимальные** системы. Статистические методы.— В кн.: Труды III Международного конгресса международной федерации по автоматическому управлению. М., «Наука», 1971. 312 с.
16. **Планирование** отраслевых систем. Модели и методы оптимизации. М., «Экономика», 1974. 320 с.
17. **Розенблют А., Винер Н., Бигелоу Д.** Поведение, целенаправленность и телеология.— В кн.: Кибернетика или управление и связь в животном и машине. М., «Сов. радио», 1968, с. 285— 294.
18. **Розин Б. Б.** Теория распознавания образов в экономических исследованиях. М., «Статистика», 1973. 224 с.
19. **Фактор** неопределенности при принятии оптимальных решений в больших системах энергетики, т. 1. Иркутск, 1974. 253 с.
20. **Фактор** неопределенности при принятии оптимальных решений в больших системах энергетики, т. 2. Иркутск, 1974. 184 с.
21. **Щукин В. Н.** Оптимизация производственной структуры отрасли и промышленного предприятия. Новосибирск, «Наука», 1973. 108 с.
22. **Статистические** модели в оптимальном отраслевом планировании. М., «Наука», 1975. 191 с.
23. **Проценко О. Д., Соловейчик Д. И.** Планирование долговременных хозяйственных связей. М., «Экономика», 1976. 143 с.
24. **Шкрабин Л. Я.** Об учете специальных видов неопределенности в задачах отраслевого планирования.— В кн.: Проблемы экономико-статистического моделирования отдельных объектов и отраслевого звена в целом. Новосибирск, 1976, с. 116—127.
25. **Bilinski L.** Optimalne decyzje w przypadki wielorokosci celow.— «Economista», 1969, № 3, S. 717—732.

В. Г. ИОНИН

ПОДБОР ПУНКТОВ РАЗМЕЩЕНИЯ ПРОИЗВОДСТВА НА ПЕРСПЕКТИВУ НА ОСНОВЕ СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Выбор вариантов развития и размещения производства на перспективу, особенно для вновь осваиваемых территорий страны, требует комплексного экономико-статистического анализа возможных пунктов размещения производства, учета динамики развития этих пунктов и их будущих хозяйственных связей, динамики потребления промышленной продукции и экономических показателей деятельности отдельных предприятий. Для экономико-статистического анализа деятельности предприятий, функционирующих в освоенных районах, имеется или может быть собрана обширная статистическая информация, а для предприятий на вновь осваиваемых территориях такая информация, естественно, отсутствует, и основным инструментом статистического исследования выступает метод статистического моделирования. На наш взгляд, построение и предварительное статистическое исследование имитационных моделей развития хозяйственных комплексов и районов должно предшествовать принятию детальных решений, а получаемая в ходе такого исследования информация может быть использована для выбора оптимальных вариантов развития, размещения и концентрации производства.

В данной статье делается попытка решить методами статистического моделирования одну из задач, предшествующих составлению оптимального плана развития и размещения производства, а именно задачу формирования списка вариантов территориального размещения предприятий во вновь осваиваемых районах. Предварительный анализ такого списка позволяет уменьшить размерность задачи оптимального планирования за счет уменьшения размерности транспортного блока. Предварительный отбор пунктов производства весьма полезен и с экономической точки зрения. Он дает возможность установить наиболее устойчивые варианты размещения при возможных изменениях в объеме выпуска продукции и характере производственных связей как в пределах планируемого периода, так и вне его. Последнее особенно важно, если учесть, что фактор территориального размещения производства является более долговременным, чем границы планового периода, задаваемые при решении конкретной плановой задачи.

Принятый ранее вариант территориального размещения производства будет оказывать влияние на все последующие в перспективе плановые решения. Поэтому целью предварительного просмотра пунктов возможного размещения производства является исключение из списка тех вариантов, которые случайно, по условиям производства, сложившимся к моменту расчета, могут попасть в схему развития территории, но окажутся невыгодными при реализации или за пределами планируемого периода.

Предварительный анализ списка пунктов размещения предприятий соответствует принципу развития концентрации производства. В частности, сокращение списка пунктов размещения в задаче оптимального планирования уменьшает возможности распределения объема производства и соответственно капитальных вложений по многочисленным объектам.

Ниже излагается процедура предварительного анализа списка вариантов территориального размещения производства, основанная на использовании статистического моделирования. Алгоритм решения, составленный на α-языке и запрограммированный для ЭВМ типа М-220, предусматривает отбор пунктов производства для однопродуктовой отрасли либо для случая, когда объемы производства и потребления выражены в условных единицах. При наличии агрегированных групп продуктов он может быть применен с учетом возможностей вычислительной техники и для многопродуктового производства.

Обозначим:

n — число пунктов потребления продукции;

Π_j—спрос на продукцию в пункте потребления j ($j = \overline{1, \Pi}$);

m_1— число пунктов размещения производства, выбираемых для включения в задачу оптимального планирования. Параметр m_1 назначается исходя из допустимой размерности плановой задачи и требований принципа развития концентрации производства;

m_2 — общее число пунктов возможного размещения производства;

C_{ij} — производственно-транспортные затраты, связанные с изготовлением единицы продукции в пункте i и доставкой ее до пункта потребления j($i = \overline{1, m_2}$; $j = \overline{1, \pi}$).

Неизвестными в задаче являются номера конкретных m_1 пунктов размещения из заданного списка длиной m_2.

Последовательность решения можно представить следующим образом.

1. Рассматривается n выборок, в каждую из которых включаются m_1 пунктов из списка длиною m_2. Число выборок должно приниматься с тем условием, чтобы с достаточной уверенностью можно было на очередном цикле расчетов исключить из списка несколько самых неэкономичных пунктов размещения. Так как этот вопрос требует специального исследования, то для определенности было принято, что число выборок (n) должно быть в 3 раза больше числа пунктов в списке (m_2). Алгоритм составлен таким образом, что по мере решения число анализируемых пунктов сокращается. Соответственно изменяется и число выборок в очередном цикле расчетов.

2. По каждой выборке формируется матрица производственно-транспортных затрат размерностью $m_1 \times \pi$ (m_1 — строки, π — столбцы).

3. Разыгрываются различные варианты распределения объема производства, равного общему спросу, по m_1 пунктам размещения. Процедура выполняется для каждой из n выборок. Для определенности принято, что количество вариантов распределения объемов производства для каждой выборки равно утроенному числу пунктов возможного размещения предприятий (m_1).

Распределение объемов производства по m_1 пунктам осуществляется таким образом, что предпочтение получают варианты выпуска продукции с объемами, близкими к граничным значениям. С этой целью организуется следующий циклический процесс.

3.1. Определяется случайная величина X, имеющая нормальный закон распределения с параметрами $m_X = 0$ и $\sigma_X = \pi/6$, где

$$\Pi = \sum_{j=1}^{\pi} \Pi_j. \qquad (1)$$

Значение X рассчитывается по формуле (см. [3])

$$X = \sigma_X \sqrt{2} \left(\sum_{i=1}^{6} R_i - 3 \right) + m_X, \qquad (2)$$

где $R_i(i = \overline{1, 6})$ — реализация случайной величины, равномерно распределенной в интервале [0,1].

3.2. Если полученное значение X отрицательно, то выполняется преобразование $X := X + \Pi$. Причем X рассматривается как объем производства продукта на предприятии, номер которого находится случайной выборкой из последовательности $\overline{1, m_1}$.

3.3. Объем производства X суммируется по каждому предприятию с объемами выпуска, найденными на предыдущих этапах (3.1 — 3.2), после чего выполняется преобразование $\Pi := \Pi - X$.

3.4. Повторяется процедура распределения (3.1—3.3) нового значения Π, и так до тех пор, пока не будет размещен весь объем производства продукции. По окончании счета в п. 3.1—3.4 исходное значение Π восстанавливается. Отмечено, что указанный процесс заканчивается в среднем через 5—6 циклов, т. е. весь объем распределяется по относительно небольшому числу предприятий (3—4 объекта). Если требуется отобрать большее число пунктов, то случайная величина X рассматривается как основа для повторения расчетов в п. 3.1—3.2 с математическим ожиданием, равным нулю, и средним квадратическим отклонением в $X/6$.

Несколько большее число предприятий оказывается задействованным, если $m_X \neq 0$. В машинном алгоритме принималось $m_X = \sigma_X$. Это соотношение было найдено опытным путем.

4. Для каждого варианта распределения объема производства по m_1 пунктам находится оптимальный план

мерной классификации и распознавания образов. Это дает основание полагать в качестве исходной предпосылки наших рассуждений наличие однородных группировок предприятий и соответствующих регрессионных зависимостей.

Данный тезис и лежал в основе определения понятия статистического агрегирования.

Агрегирование — это форма осуществления вертикальных информационных связей в многоуровневой системе управления. Поэтому важным приложением указанного выше свойства статистических моделей является рассмотрение их в рамках одной из схем двухуровневой координации. Для наших целей не имеет принципиального значения выбор конкретной схемы.

Рассмотрим двухуровневую систему моделей, где модель первого уровня описывает деятельность группы предприятий (отрасли), а второй уровень образуют экономико-статистические модели предприятий.

Любая схема двухуровневой оптимизации предполагает наличие определенных элементов включаемых в нее моделей: целевой функции, вектора ограничений, векторной функции технологических связей, ограничений на переменные. Стандартная экономико-статистическая модель не соответствует этим условиям. На основе экономико-статистической модели возможно построение линейно-программной модели предприятия, отвечающей перечисленным выше условиям [3].

Пусть $x_1^j, \ldots, x_n^j$ — выделенные независимые признаки деятельности j-го объекта, а $y_1^j, \ldots, y_m^j$ — результирующие показатели этого объекта. Заданные числа $\underline{N}_k^j$, $\underline{M}_i^j$ и $\overline{N}_k^j$, $\overline{M}_i^j$ — соответственно нижние и верхние границы интервалов зависимых и независимых показателей:

$$\underline{M}_i^j \leqslant x_i^j \leqslant \overline{M}_i^j \quad (i = \overline{1, n}); \tag{1}$$

$$\underline{N}_k^j \leqslant y_k^j \leqslant \overline{N}_k^j \quad (k = \overline{1, m}); \tag{2}$$

$$y_k^j = \sum_{i=1}^{n} a_{ik}^j x_i^j + a_{0k}^j \quad (k = \overline{1, m}), \tag{3}$$

где a_{ik}^j — коэффициенты k-го уравнения регрессии по i-му фактору; a_{0k}^j — свободный член k-го уравнения.

В качестве функционала выбирается один из моделируемых показателей.

Именно такие модели вида (1)—(3) мы и подразумеваем в качестве подмоделей второго уровня рассматриваемой системы. Их использование открывает широкие возможности для агрегированного представления объектов.

Во-первых, входящие в некоторую статистическую группировку объекты описываются одним данным набором уравнений регрессии, и поэтому структура моделей этих объектов идентична. Различны лишь значения ограничений на зависимые и независимые переменные.

Во-вторых, возможны достаточно разнообразные группировки объектов по интересующим нас признакам.

Введем понятие агрегированной модели — модели группы предприятий. Согласно (1)—(3), она будет выглядеть следующим образом:

$$\underline{M}_i^I \leqslant x_i^I \leqslant \overline{M}_i^I \quad (i = \overline{1, n}); \tag{4}$$

$$\underline{N}_k^I \leqslant y_k^I \leqslant \overline{N}_k^I \quad (k = \overline{1, m}); \tag{5}$$

$$y_k^I = \sum_{i=1}^{n} a_{ik}^I x_i^I + a_{0k}^I \quad (k = \overline{1, m}), \tag{6}$$

где y_k^I и x_k^I — соответственно значения зависимых и независимых переменных;

$$\overline{M}_i^I = \sum_{j \in S^I} \overline{M}_i^j / l^I, \quad \underline{N}_k^I = \sum_{j \in S^I} \underline{N}_k^j / l^I,$$

$$\underline{M}_i^I = \sum_{j \in S^I} \underline{N}_i^j / l^I, \quad \overline{N}_k^I = \sum \overline{N}_k^j / l^I$$

— нижние и верхние границы интервалов зависимых и независимых переменных; l^I — число объектов, входящих в группировку I; S^I — множество индексов объектов группировки I.

Замена моделей вида (1)—(3) на модель вида (4)—(6) в системе представляет собой операцию агрегирования в явном виде. В самом деле, мы заменяем l^I моделей объектов, принадлежащих группировке I, одной агрегированной моделью.

Решение системы с моделями вида (4)—(6) в качестве подмоделей может дать результат, удовлетворяющий исходным посылкам формирования экономико-статисти-

прикрепления поставщиков к потребителям и соответствующее ему значение производственно-транспортных затрат. В вычислительном отношении оказалась удобной процедура построения плана транспортной задачи по методу Фогеля [5], результаты применения которой можно рассматривать как приближающиеся к оптимальным.

5. Для каждого варианта выборки в m_1 пунктов рассчитываются средние ожидаемые затраты по наилучшим планам прикрепления поставщиков к потребителям. Усреднение проводится по числу вариантов распределения объемов производства среди m_1 пунктов. В результате n испытаний формируется массив $\{T_p\}$ длиной p, где $T_p (p = \overline{1, n})$ — среднее значение производственно-транспортных затрат для выборки с номером p. Подчеркнем, что средние затраты находятся по вариантам наилучших связей производства и потребления.

6. Просмотренные варианты выборки m_1 пунктов из общего количества m_2 фиксируются в виде массива $\{Л_{pi}\}$, где

$$Л_{pi} = \begin{cases} 1, \text{ если для варианта с номером } p \text{ выбран} \\ \text{пункт производства с номером } i, \\ 0 \text{ в противном случае} \\ (p = \overline{1, n}; \quad i = \overline{1, m_2}). \end{cases} \tag{3}$$

7. Переменные в массиве $\{T_p\}$ упорядочиваются по возрастанию. Соответственно располагаются и строки в массиве $\{Л_{pi}\}$.

8. Рассчитывается число повторений пункта производства с номером i ($i = \overline{1, m_2}$) в выборках с наилучшими средними значениями производственно-транспортных затрат. Для определенности на первом цикле решения просматриваются первые m_2 элементов массива $\{T_p\}$. На последующих шагах длина просматриваемого списка уменьшается. Расчет выполняется по формуле

$$\varepsilon_i = \sum_{p=1}^{m_3} Л_{pi}, \tag{4}$$

где ε_i — число повторений пункта с номером i среди лучших m_3 испытаний;

m_3 — длина списка пунктов размещения, анализируемого на данном цикле расчетов. В начале решения $m_3 = m_2$.

9. Для дальнейшего анализа выбираются лишь те пункты размещения, которые повторились наибольшее число раз. Номера исключаемых пунктов определяются по массиву $\{\varepsilon_i\}$. Количество исключаемых пунктов рассчитывается по формуле

$$m_{ф} = 0{,}38(m_3 + 1 - m_1), \tag{5}$$

где $m_{ф}$, m_3, m_1 — соответственно число исключаемых из рассмотрения пунктов размещения, прежняя длина списка пунктов возможного размещения предприятий, количество пунктов размещения, которое требуется выбрать для решения задачи оптимального планирования.

После расчета по формуле (5) определяется новая длина списка

$$m_3 := m_3 - m_{ф}. \tag{6}$$

10. Решение задачи получено, если оказывается, что в исходном списке осталось требуемое число пунктов, т. е. $m_{ф} = 0$. По номерам, исключаемым на каждом шаге, определяются конкретные пункты размещения предприятий, оставляемые для решения задачи оптимального размещения производства. Если $m_{ф} \neq 0$, то цикл расчетов повторяется, начиная с п. 1.

Как видно из описанной выше последовательности предварительного отбора наиболее перспективных пунктов размещения, число таких пунктов вводится в качестве параметра, а результаты решения определяют конкретные номера и средние ожидаемые производственно-транспортные затраты при возможных изменениях объема выпуска продукции. Во всех случаях затраты рассчитываются по наилучшим планам прикрепления поставщиков к потребителям. Схема расчета допускает возможность дальнейшего совершенствования алгоритма.

Первое направление такого совершенствования заключается в привлечении экономико-статистических моделей, связывающих производственно-транспортные затраты с объемом производства. Следующее направление состоит в более детальном исследовании производственных возможностей каждого предприятия и наложении соответствующих ограничений по каждому пункту. Вопрос о задании желаемого числа пунктов размещения может решаться исходя из возможностей реализации конкретной задачи оптимизации на ЭВМ. Но более правильным было

Исходные данные для решения задачи отбора пунктов размещения производства (удельные производственно-транспортные затраты в зависимости от спроса на продукцию, руб/т)

Номер пункта размещения	Спрос на продукцию по пунктам потребления, тыс. т																	
	556	652	163	595	284	538	1055	23	96	52	96	74	200	262	8	10	22
1	9,2	8,4	8,7	8,8	8,5	9,4	10,7	—	11,4	—	12,3	15,6	16,4	18	—	—	—
2	13,3	12,4	12,3	12,4	12,7	12,9	14,2	—	14,9	—	15,7	19,8	20,5	22,2	—	—	—
3	9,3	8,4	8,5	8,2	8,7	9	10,3	15	11	—	11,9	15,9	16,7	18,3	—	—	—
4	7,9	7,1	7,2	7,1	7,5	6,4	7,7	10,3	8,4	—	9,3	12,6	13,4	15	—	—	—
5	24,4	23,5	23,3	23,5	23,9	22,3	21	—	20,3	28,2	19,5	16,7	16,2	17,9	24,3	30,2	30,5
6	23,9	23,1	22,9	23	23,5	21,9	20,6	—	19,9	29,7	19	16,3	14,9	14,3	25,5	25,1	19,1
7	21	20,1	20	20,1	20,6	19,1	17,8	—	17,1	—	16,2	13,9	14,7	16,3	—	—	—
8	25,3	24,5	24,3	24,5	24,9	23,3	22	—	21,3	—	20,4	17,7	16,3	14,8	25,4	25	19

бы подходить к заданию количества пунктов, отбираемых для включения в задачу оптимального планирования, с позиций испытания их различного числа и выбора такой величины, которая в среднем (в данном случае — на перспективу) дает наименьшие производственно-транспортные затраты.

В таблице представлены условные исходные данные (о спросе и удельных производственно-транспортных затратах) для решения задачи отбора трех наиболее экономичных пунктов размещения производства из восьми заданных при числе пунктов потребления, равном 17. На первой итерации по восьми наилучшим выборкам с наименьшими производственно-транспортными затратами суммарные затраты на производство и транспортировку колеблются от 46,2 до 63,7 тыс. руб. В результате первой итерации из списка пунктов оказываются исключенными пункты с номерами 7 и 8, которые даже в лучших выборках повторились только по одному разу, в то время как пункт № 1 повторился шесть раз, № 2 — четыре, № 3 — три, № 4 — пять и № 5 и 6 — по два раза. На следующей итерации средние производственно-транспортные затраты по шести наилучшим выборкам колеблются от 42,3 до 54,6 тыс. руб. Исключенными оказались пункты с № 2 (повторился только один раз) и 5 (в выборку не попал). Пункты № 1 и 3 повторились по четыре раза, № 4 — шесть, № 6 — три раза. На третьей итерации исключенным оказался пункт с № 3; в результате были отобраны пункты с № 1, 4, 6, а общие затраты на производство и транспортировку продукции ожидаются в интервале от 43,7 до 49,4 тыс. руб.

ЛИТЕРАТУРА

1. **Казакевич Д. М.** Производственно-транспортные модели в перспективном отраслевом планировании. М., «Экономика», 1972. 295 с.
2. **Щукин В. И., Ионин В. Г.** Некоторые подходы к решению задач выбора плана развития и размещения комплекса взаимосвязанных подотраслей.— В кн.: Оптимизационные и экономико-статистические методы в перспективном отраслевом планировании. Новосибирск, «Наука», 1975, с. 14—28.
3. **Вентцель Е. С.** Исследование операций. М., «Сов. радио», 1972. 576 с.
4. **Бусленко Н. П., Шрейдер Ю. А.** Метод статистических испытаний. М., Физматгиз, 1961. 226 с.
5. **Рейнфельд Н., Фогель У.** Математическое программирование. Методы решения производственных и транспортных задач. Пер. с англ. М., ИЛ, 1960. 303 с.

В. Г. ИОНИН

ТОЧНОСТЬ РАСЧЕТА ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ВАРИАНТОВ РАЗВИТИЯ ПРЕДПРИЯТИЙ В ЗАДАЧЕ ПЕРСПЕКТИВНОГО ОТРАСЛЕВОГО ПЛАНИРОВАНИЯ [1]

Для расчета экономических показателей варианта развития предприятия необходимо охватить весьма большой круг затрат, связанных с созданием и эксплуатацией нового объекта. Наиболее точно и детально метод подобных расчетов разработан в системе проектных разработок. Но не во всех случаях есть необходимость в проведении таких детальных разработок. Так, варианты развития на отдаленную перспективу можно, по-видимому, рассчитывать с меньшей детализацией и с бóльшим огрублением отдельных статей затрат. Структуру элементов затрат в расчете варианта развития ограничивает, на наш взгляд, требование обеспечения заданной точности решения отраслевой задачи в целом. Рассмотрим в этой связи некоторые вопросы, связанные с точностью решения отраслевой задачи в целом и расчета элементов затрат по вариантам развития предприятий.

С одной стороны, точность расчета экономических показателей варианта развития каждого предприятия отрасли определяет в итоге точность решения отраслевой задачи. С другой стороны, если заданы пределы колебаний величины вложений в основные фонды, себестоимости, приведенных затрат и т. д., то следует предъявить определенные требования к точности расчета соответствующих показателей в вариантах развития отдельных предприятий.

[1] Вопросы подготовки вариантов развития предприятий для решения отраслевой задачи рассматривались нами ранее (в работе [2]).

Рассмотрим один из подходов к расчету точности, с которой формируется вариант развития предприятий, на примере отраслевой модели производственного типа с дискретными переменными [1]. В такой модели оптимальное значение целевой функции представляет собой сумму из n слагаемых затрат по числу предприятий, включаемых в решение:

$$L = \sum_{j=1}^{n} C_j, \tag{1}$$

где C_j — приведенные затраты, соответствующие варианту развития предприятия j, вошедшему в оптимальное решение.

Предположим, что величина затрат по вариантам развития предприятий определена с относительной погрешностью δ_j, понимая под этим, что C_j является реализацией случайной величины, распределенной в интервале $[\overline{C}_j(1 - \delta_j), \overline{C}_j(1 + \delta_j)]$, где $\overline{C}_j$ — среднее значение случайной величины затрат по варианту. Поскольку точность каждого варианта не зависит от точности расчета по остальным вариантам и предприятиям, то, согласно теореме сложения дисперсий (см., например, [3]), имеем

$$D_L = \sum_{j=1}^{n} D(C_j), \tag{2}$$

где D_L — дисперсия значения целевой функции отраслевой задачи;
$D(C_j)$ — дисперсия значений затрат по варианту развития предприятия j.

Можно также предположить, что величина ошибки в определении значения целевой функции подчиняется закону нормального распределения, поскольку представляет сумму из n случайных величин C_j, которые предположительно оказывают в среднем одинаковое воздействие на значение L. Тогда для нормального закона распределения справедливо (с вероятностью 0,98) соотношение

$$D_L = (\overline{L} \times \delta_L / 3)^2. \tag{3}$$

Здесь δ_L — относительная погрешность целевой функции отраслевой задачи.

Закон распределения каждого из слагаемых в формуле (2), по-видимому, нельзя указать однозначно, так как

случайная величина затрат C_j зависит от большого числа неодинаково и даже противоположно действующих факторов. Рассмотрим две гипотезы: о нормальности распределения случайной величины C_j и о равномерном распределении C_j в интервале $[\overline{C_j}(1-\delta_j),\ \overline{C_j}(1+\delta_j)]$.

В первом случае по аналогии с формулой (3) можно записать:

$$D(C_j) = \left(\frac{C_j\,\delta_j}{3}\right)^2. \qquad (4)$$

Тогда соотношение (1) примет вид

$$\left(\frac{\overline{L}\delta_L}{3}\right)^2 = \sum_{j=1}^{n}\left(\frac{C_j\delta_j}{3}\right)^2, \qquad (5)$$

откуда после преобразований и с учетом предположения, что $\delta = \delta_1 = \delta_2 = \ldots = \delta_n$, получаем

$$\delta_L/\delta = \sqrt{\sum_{j=1}^{n} C_j^2/L^2} = \sqrt{\sum_{j=1}^{n} C_j^2/(n^2\overline{C}^2)} =$$

$$= \sqrt{n \times (\overline{C}^2 + D_C)/(n^2\overline{C}^2)} = \sqrt{1+v^2}/\sqrt{n}, \qquad (6)$$

где v — коэффициент вариации. Обозначив $\sqrt{1+v^2}$ через K_v, получаем

$$\delta_L/\delta = K_v/\sqrt{n}. \qquad (7)$$

Аналогично для равномерного распределения C_j имеем

$$\delta_L/\delta = K_v\sqrt{3}/\sqrt{n}. \qquad (8)$$

Из формул (7) и (8) видно, что соотношение между относительными ошибками в определении величины целевой функции и затрат по вариантам развития отдельных объектов связано с размерностью отраслевой задачи, видом распределения ошибки в расчете затрат по вариантам развития объектов и степенью однородности (по величине затрат) оптимальной совокупности вариантов. С помощью этих соотношений можно решать как прямую задачу — определение погрешности целевой функции при известной точности расчета слагаемых (вариантов), так и обратную — при заданной точности решения отраслевой задачи в целом рассчитать максимально допустимую погрешность при формировании вариантов развития отдельных предприятий отрасли.

Нетрудно убедиться, что даже в случае неоднородной совокупности вариантов значение K_v близко́ к единице. С учетом вероятностного характера расчетов ниже приведены соотношения между точностью расчета затрат по вариантам развития предприятий (δ) и затрат по реализации отраслевого плана (δ_L) при $K_v = 1$ и различных предположениях относительно закона распределения величины δ (%):

$n=10$				$n=40$			
Равномерное		Нормальное		Равномерное		Нормальное	
δ_L	δ	δ_L	δ	δ_L	δ	δ_L	δ
10	18	10	32	10	36	10	63
9	16	9	28	9	33	9	57
8	15	8	25	8	29	8	50
7	13	7	22	7	26	7	44
6	11	6	19	6	22	6	38
5	9	5	16	5	18	5	32
4	7	4	13	4	15	4	25
3	6	3	10	3	11	3	19
2	4	2	6	2	7	2	13
1	2	1	3	1	4	1	6

(Заметим, что выбор того или иного вида распределения должен определяться результатами статистического исследования соотношения между проектными и фактическими затратами.)

Представляет определенный интерес исследование таких соотношений между фактическими и плановыми затратами, когда наблюдается систематическое превышение расходов по реализации вариантов развития по сравнению с затратами, принятыми при проектировании предприятия или при формировании схемы развития отрасли. Предположим, что относительное отклонение фактических затрат от проектных по варианту развития предприятия имеет центр распределения не в нулевой точке, а сдвинуто вправо на величину Δ. Таким образом, математическое ожидание затрат по реализации отраслевого плана возрастает на величину $\Delta L_{\text{пр}}$, а максимально возможная сумма затрат на реализацию отраслевого плана будет равна

$$L_{\max} = (1+\delta_L')(1+\Delta)\,L_{\text{пр}}, \qquad (9)$$

где δ_L' — относительное отклонение затрат от среднего значения;

$L_{пр}$ — проектная величина затрат на реализацию отраслевого плана.

Следовательно, наличие положительного смещения затрат по вариантам развития предприятий вызывает соответствующее увеличение области возможного возрастания затрат в целом по отрасли. Относительная величина его составит в таком случае

$$\delta_L = \delta'_L + \Delta + \delta'_L \Delta. \qquad (10)$$

Предположим теперь, что постоянный сдвиг выражается через ошибку в расчете показателей затрат. Пусть, например, $\Delta = \delta/3$. Тогда при заданной точности решения отраслевой задачи в целом, известном числе предприятий и виде распределения ошибки в расчете затрат по каждому варианту предприятия получим следующие допустимые показатели относительной ошибки исчисления δ и δ_L при систематическом отклонении фактических затрат от проектных (%):

$n=10$				$n=40$			
Равномерное		Нормальное		Равномерное		Нормальное	
δ_L	δ	δ_L	δ	δ_L	δ	δ_L	δ
16	18	21	32	22	36	31	63
14	16	16	28	20	33	28	57
13	15	16	25	18	29	25	50
11	13	14	22	14	26	22	44
10	11	12	19	13	22	19	38
8	9	10	16	11	18	10	32
6	7	8	13	9	15	12	25
5	6	6	10	7	11	9	19
3	4	2	6	4	7	6	13
2	2	2	3	2	4	3	6

Сравнение приведенных показателей относительной ошибки показывает, что при систематическом смещении затрат в сторону увеличения возрастают требования к точности формирования вариантов. Требуемая точность расчета возрастает в 2—3 раза.

Остановимся на некоторых особенностях, связанных с динамической постановкой задачи отраслевого планирования [1]. Для динамической постановки характерно, что варианты, которые могут быть реализованы в начале планового периода, должны задаваться с большей точностью, чем предполагаемые на конец периода. Таким образом, относительная ошибка в расчете показателей по каждому варианту может возрастать по мере удаления от начального момента.

Введем обозначения:

T — длительность периода планирования;

δ_t — допустимая относительная ошибка при расчете вариантов развития предприятия для года t ($t = \overline{1, T}$);

δ_L^T — относительная ошибка целевой функции в задаче динамического типа с числом лет T, возникающая в результате неточности исходных данных (затрат) по вариантам развития.

Сделаем следующие предположения:

а) ошибка в расчете затрат по каждому из вариантов подчиняется закону равномерного распределения;

б) ввод предприятий в действие осуществляется в среднем равномерно на протяжении периода T;

в) возрастание относительной ошибки в расчете показателей затрат по вариантам развития происходит внутри планового периода с постоянным темпом α,

$$\alpha = \delta_t/\delta_{t-1}. \qquad (11)$$

Тогда для динамической задачи соотношения (2)—(7) запишутся так:

$$D_L^T = \sum_{j=1}^{n} D(C_j) \qquad (12)$$

или

$$(\bar{L}\delta_L^T/3)^2 = \sum_{j=1}^{n} (C_j\delta_j)^2/3. \qquad (13)$$

Считая, что для всех вводимых в t-м году g (где $g = n/T$) предприятий погрешность δ_t одинакова, и учитывая, согласно предположению в), что

$$\delta_t = \alpha^{(t-1)}\delta_1, \ t = \overline{1, T}, \qquad (14)$$

имеем

$$n\delta_L^T/(\sqrt{3}\delta_1) = \sqrt{\sum_{t=1}^{T} \alpha^{2(t-1)} \sum_{i=1}^{g} C_{it}^2/\bar{C}^2} \qquad (15)$$

или

$$n\delta_L^T/(\sqrt{3}\sqrt{g}\delta_1) = \sqrt{\sum_{t=1}^{T} \alpha^{2(t-1)} (\bar{C}_t^2 + D_{C_t})/\bar{C}^2}. \qquad (16)$$

Здесь $\overline{C_t}$ — средние затраты по вариантам, вводимым в году t; D_{C_t} — дисперсия затрат по вариантам в году t.

Если принять, что затраты по вариантам развития предприятий должны распределяться в среднем равномерно по годам планируемого периода, т. е. $\overline{C}_t \cong \overline{C}(t=\overline{1,T})$, то получим

$$n \cdot \delta_L^T / (\sqrt{3} \cdot \sqrt{g}\delta_1) = \sqrt{\sum_{t=1}^{T} \alpha^{2(t-1)} (1 + v_t^2)}. \quad (17)$$

Так как $(1 + v_t^2)$ близко́ к единице, то формулу (17) можно переписать в виде

$$(n\delta_L^T)/(\sqrt{3} \cdot \sqrt{g}\delta_1) = \sqrt{\sum_{t=1}^{T} \alpha^{2(t-1)}}. \quad (18)$$

Таким образом, при заданной размерности отраслевой задачи (n и T), требуемой точности расчета целевой функции (δ_L^T) и вариантов развития предприятий в начале планового периода определяется α-допустимый темп роста погрешности расчета показателей по вариантам развития предприятий в плановом периоде T. При этом удобно пользоваться таблицей, составленной для различной длительности периода T и фиксированных значений левой части формулы (18) (табл. 1).

Для иллюстрации рассмотрим пример, когда $T = 15$ (лет), а $n = 45$ (предприятий), следовательно, $g = n/T = 3$. Потребуем, чтобы фактическое значение целевой функции отличалось от расчетного не более чем на $\pm 5\%$ $(\delta_L^{15} = 5)$ и чтобы варианты развития предприятий для первого года подготавливались с точностью $\delta_1 = \pm 3\%$.

Тогда $n\delta_L^{15}/(\sqrt{3} \cdot \sqrt{g}\delta_1) = 25$. Из табл. 1 находим, что α заключено в пределах от 1,204 до 1,216. Интерполируя, принимаем $\alpha = 1{,}208$ и получаем следующий ряд допустимых погрешностей при формировании вариантов развития предприятий:

1	2	3	4	5	6	7
3	3,6	4,5	5,3	6,4	7,7	9,3

8	9	10	11	12	13	14	15
11,3	13,6	16,5	19,8	24,0	29,0	35,0	42,5

Для ориентировочных расчетов в этом примере можно принять, что варианты развития предприятий должны подготавливаться с точностью (%):

для $t = 1$—5 лет . . . . 3—5
» $t = 6$—10 » . . . . . 10—15
» $t = 11$—15 » . . . . . 20—40

Наряду с задачей выбора допустимой ошибки результирующих показателей варианта развития предприятия существует и другая — выбора предельно допустимой ошибки в системе исходных данных, которые используются при формировании вариантов. Поскольку итоговая характеристика варианта развития (например, приведенные затраты) является результатом действия комплекса различных факторов с разной степенью влияния, то использование аналитического метода определения допустимой погрешности исходных данных оказывается невозможным из-за сложности расчетов.

Для исследования устойчивости результатов расчета варианта развития предприятия при изменении исходных данных был применен метод статистического моделирования. Согласно этому методу, выполнялась серия испытаний, в каждом из которых по методике проектных разработок [2] рассчи-

Таблица 1

Темпы роста допустимой погрешности формирования вариантов развития предприятий в динамической отраслевой задаче

Длительность планового периода T, лет	$\left(n\sigma_L^T\right)/\left(\sqrt{3} \cdot \sqrt{g}\delta_1\right)$														
	11	12	13	14	15	16	17	18	20	22	24	27			
10	1,229	1,244	1,257	1,271	1,283	1,294	1,300	—	1,289	—	—	—			
11	1,198	1,211	1,223	1,234	1,245	1,255	1,264	1,273	1,267	1,300	—	—			
12	1,173	1,185	1,196	1,206	1,215	1,224	1,232	1,240	1,279	1,279	1,295	—			
13	1,153	1,164	1,174	1,183	1,191	1,199	1,207	1,214	1,227	1,238	1,249	1,263			
14	—	1,150	1,156	1,164	1,172	1,179	1,186	1,192	1,204	1,215	1,224	1,237			
15	—	—	—	1,150	1,156	1,162	1,168	1,174	1,185	1,195	1,204	1,216			

Анализ устойчивости результатов расчета показателей варианта

Характеристика расчета	σ=±5%			
	C	K	$C_{пр}$	C
Среднее значение, тыс. руб.	24 859	40 185	30 887	24 859
Минимальное значение, тыс. руб.	23 603	38 669	29 419	22 098
Максимальное значение, тыс. руб.	26 268	42 048	32 273	27 492
Разброс (±), %	5,4	4,2	4,6	9,0
Среднеквадратическое отклонение, тыс. руб.	634	803	698	1201
Коэффициент эксцесса	−0,794	−0,624	−0,773	−0,655
Отношение максимального значения к минимальному	1,11	1,09	1,10	1,24

тывались показатели варианта развития предприятия. В процессе расчета каждая из величин, используемых в качестве исходных данных, бралась в ее случайной реализации. Пределы возможных изменений каждой величины ограничивались заданной точностью. Для упрощения расчетов точность задавалась постоянной в каждой серии испытаний и изменялась лишь в новой серии. Случайные величины исходных данных выбирались по закону равномерного распределения в интервале $\pm\delta$ (задаваемая точность) от среднего значения. Исключение было сделано для показателей объемов работ по видам оборудования. Здесь, чтобы не моделировать случайную загрузку по каждой операции, принималось нормальное распределение итоговой загрузки каждого вида оборудования с математическим ожиданием, равным количеству единиц оборудования в группе, исчисленному по среднему объему работ, и средним квадратическим отклонением, рассчитываемым по формуле

$$\sigma_q = (\delta \times K_q)/3, \quad q = \overline{1, Q}, \qquad (19)$$

где q — номер вида оборудования;

σ_q — среднеквадратическое отклонение количества оборудования в группе q от среднего значения;

K_q — количество оборудования в группе q, рассчитанное по средней загрузке;

Таблица 2

развития предприятия при изменении точности исходных данных

σ=±10%		σ=±15%			σ=±30%		
K	$C_{пр}$	C	K	$C_{пр}$	C	K	$C_{пр}$
40 185	30 887	24 859	40 185	30 887	24 859	40 185	30 887
36 295	27 065	21 152	34 433	26 366	17 441	30 444	20 607
44 477	34 249	29 330	45 664	35 448	34 107	61 163	45 571
10,2	11,6	16,4	12,7	14,7	33,5	38,2	38,0
1 844	1 454	1 821	2 552	1 945	3 529	5 281	4 232
−0,285	−0,1482	−0,580	−0,840	−0,420	−0,204	0,014	−0,436
1,22	1,27	1,39	1,29	1,35	1,95	2,02	2,21

δ — задаваемая точность исходных данных (коэффициент).

В результате расчета, проведенного на ЭВМ М-220, получены статистические характеристики распределения годовой себестоимости (C), капитальных вложений (K) и приведенных затрат ($C_{пр}$). Моделирование осуществлялось на основе информации, полученной по одному из машиностроительных предприятий Новосибирска. Предприятие выпускает 44 вида продукции и имеет 54 группы оборудования. Результаты расчетов приведены в табл. 2.

Из полученных данных видно, что характерно плосковершинное распределение результирующих показателей варианта развития предприятия, поскольку коэффициент эксцесса отрицательный (−0,8). Относительный разброс значений результирующих показателей в целом совпадает с точностью исходных данных, а в отдельных случаях может даже превысить их относительную погрешность. Так, при $\delta = \pm 5\%$ показатель себестоимости колеблется в пределах $\pm 5{,}4\%$; при $\delta = \pm 30\%$ величина приведенных затрат по варианту может находиться в пределах $\pm 38\%$ от среднего значения, что в абсолютном выражении дает разброс в 10—15 млн. руб. Следовательно, при формировании варианта развития машиностроительного предприятия точность исходных данных должна быть не меньше точности, которая требуется от итоговых показа-

телей. Предположение, что допустима меньшая точность исходных данных (за счет «гашения» ошибок при расчете затрат по варианту развития предприятия), расчетами не подтверждается.

Представляют интерес соотношения максимальных и минимальных значений величин затрат по вариантам развития предприятия (см. табл. 2). Это отношение растет с увеличением погрешности в системе исходных данных. Так, если при расчете используется относительно точная информация ($\delta = 5—10\%$), то различие между наибольшими и наименьшими затратами по одному и тому же варианту составляет 10—30%. Но уже при $\delta = 30—50\%$ максимально возможная величина затрат может в 2—3 раза превзойти минимальную. Таким образом, если информация, используемая при расчете варианта развития предприятия, недостоверна и для обоснования эффективности принята ориентация на нормативные показатели, ведущие к преуменьшению затрат, то фактические затраты по принятому варианту могут в несколько раз превысить проектные показатели себестоимости и капитальных вложений.

ЛИТЕРАТУРА

1. **Методические** положения оптимального отраслевого планирования в промышленности. Новосибирск, «Наука», 1972. 312 с.
2. **Щукин В. Н., Ионин В. Г.** Формирование вариантов развития предприятия в задаче перспективного отраслевого планирования.— В кн.: Оптимизационные отраслевые модели. Новосибирск, 1973, с. 69—89.
3. **Вентцель Е. С.** Теория вероятностей. М., «Наука», 1969. 576 с.

Б. Б. РОЗИН, М. А. ЯГОЛЬНИЦЕР

СТАТИСТИЧЕСКИЕ ИМИТАЦИОННЫЕ МОДЕЛИ КАК ИНСТРУМЕНТ АНАЛИЗА НЕОПРЕДЕЛЕННОСТИ ОПТИМАЛЬНЫХ ПЛАНОВ

Концепция оптимального функционирования экономики постоянно уточняется, развивается и еще далека от своего теоретического завершения. В настоящее время идет дискуссия о тех предпосылках, на которых она должна строиться. Разработанные научные методы планирования сложных экономических систем позволяют рассчитывать план, согласованный на всех уровнях хозяйственной иерархии. Основу всех этих методов составляет принцип декомпозиции — расчленения задачи на ряд подзадач и дальнейшей взаимоувязки их решений [4].

Однако применение таких методов планирования еще не разрешает всех тех проблем, которые возникают в процессе реализации плана. Действительно, процесс функционирования сложной экономической системы не есть точная копия планового расчета в силу неопределенного характера параметров модели и агрегированного описания объектов более низкого уровня.

Первое обстоятельство вызвано тем, что суждения относительно значений параметров модели строятся исходя из прогнозов о спросе, научно-техническом прогрессе, ожидаемых метеорологических условиях, результатах геологоразведки и т. д. Точно предсказать значения тех или иных параметров модели в плановом периоде на основе имеющейся информации невозможно, и это обусловливает их неопределенность. Таким образом, первое обстоятельство связано с тем, что расчеты, получаемые по моделям перспективного территориально-производственного планирования, относятся к будущему.

Можно различать два основных типа неопределенности входных параметров экономико-математических моделей [1] — *случайные* и *расплывчатые*. Под случайностью понимается неопределенность, касающаяся принадлежности некоторого объекта к нерасплывчатому множеству. Понятие же расплывчатости относится к классам, в которых могут иметься различные градации степени принадлежности, промежуточные между полной принадлежностью и непринадлежностью объектов к данному классу [1].

В экономико-математических задачах мы сталкиваемся с обоими типами неопределенности входных параметров. Например, метеорологические условия можно отнести к классу случайных явлений, а результаты научных или геологических открытий — к классу расплывчатых явлений. При этом необходимо подчеркнуть, что основным источником погрешности плановых решений является второй тип неопределенности.

Отсюда следует, что один из развиваемых радикальных подходов к учету неопределенности, ориентирующийся только на вероятностное описание моделируемых процессов (методы стохастического программирования), принципиально не может полностью решить проблему неопределенности, не говоря уже о технических и информационных трудностях его применения.

Нам представляется достаточно эффективным в данной ситуации использование методов имитационного моделирования. Это новое направление исследования природы упомянутых явлений возникло в результате развития математических методов в экономических и социальных явлениях и появления современных быстродействующих ЭВМ. «Имитационная система — это совокупность модели, имитирующей изучаемое явление и закодированной в ЭВМ, системы внешнего математического обеспечения и системы внутреннего обеспечения»— такова наиболее удачная, на наш взгляд, формулировка имитационной системы, предложенная Н. Н. Моисеевым в предисловии к работе [7].

В данном определении содержится несколько основных положений. Во-первых, имитация — это *численный метод*,

[1] Расплывчатое множество — класс, в котором нельзя указать резкую границу, отделяющую элементы, принадлежащие к данному классу и не принадлежащие к нему [10].

и он используется в тех случаях, когда аналитические методы исследования не существуют или крайне затруднительны. Во-вторых, имитация представляет собой *эксперимент*, но в отличие от обычного этот эксперимент проводится с моделью реальной системы, а не с самой системой. Наконец, в-третьих, имитация открывает широкие возможности для *исследования систем в условиях неопределенности*.

Эти характерные особенности метода имитационного моделирования делают его весьма универсальным инструментом исследования сложных систем. Как очень удачно сказано в работе [9], «вне зависимости от сложности построенной модели имитационные методы позволяют сделать вытекающие из нее выводы» (с. 82).

Следует отметить глубокую аналогию между статистическими и имитационными методами исследования. При имитационном моделировании строится оператор, связывающий входы системы и ее выходы на основе машинных экспериментов. Подобные же зависимости строятся на множестве фактических реализаций и при статистическом моделировании. Основная идея использования статистических моделей для имитации поведения сложных систем (моделей) заключается в замене прошлой фактической информации данными, генерируемыми в машинном эксперименте. В качестве генератора информации выступают оптимальные экономико-математические модели исследуемого уровня.

Итак, сущность задачи статистической имитации заключается в установлении закономерных связей между возможными изменениями входных переменных и результатами решений оптимизационных задач (значение целевой функции, набор способов). Тем самым будет получен оператор, закономерно описывающий область возможных изменений оптимального решения, что важно, с одной стороны, для аналитических целей (установление набора существенных переменных, характера их взаимосвязи, влияния и т. д.), а с другой — для целей управления процессом принятия решений, т. е. для корректировки плановых решений в зависимости от появления новых, ранее не рассматривавшихся ситуаций.

Главное отличие предлагаемого подхода к учету неопределенности от широко известного метода, разработанного в Сибирском энергетическом институте СО АН СССР [5], заключается в том, что помимо формирования самой

зоны неопределенности оптимальных планов строится оператор, связывающий изменения входных (экзогенных) воздействий и оптимальных решений.

Процесс построения имитационной модели включает ряд операций. Первая из них — *формулирование проблемы и выбор объекта имитации*. В частности, объектами имитации могут выступать детерминированные модели отраслевой оптимизации. Естественно, что на начальном этапе освоения метода следует рассмотреть одноуровневые системы, а затем проводить обобщение на случай многоуровневых систем. Можно ставить задачу построения оператора, связывающего изменения оптимального решения и части (всех) входных переменных как при единственном критерии оптимизации, так и при наличии нескольких критериев.

Вторая операция — *исследование характера входных воздействий (переменных) и способов их формирования*. В случае известной входной информации следует подразделить ее на случайную [2] и расплывчатую. Для первой необходимо изучить многомерные законы распределения параметрическими или непараметрическими методами, а для расплывчатой информации можно воспользоваться концепциями субъективной вероятности [3], развиваемыми в рамках теории принятия решений [2]. Возникновение последней характеризует новый этап в развитии исследования операций. Субъективные вероятности назначаются экспертами и постоянно уточняются в процессе эксперимента. Таким образом, байессовский подход к моделированию систем в условиях неопределенности характеризуется непосредственным участием лица, ответственного за принятие решения в определении вероятностных параметров модели.

Весьма важная и острая проблема осуществления второй операции — собственно процесс формирования входных переменных. Здесь, как показывают наши исследования [8], весьма эффективно использование экономико-статистических моделей объектов более низкого уровня как генераторов исходной информации для моделей более высокого уровня. Тем самым решается и проблема агрегирования, когда детальная, изменчивая и недоступная

[2] Детерминированная информация представляет собой частный вариант случайной — при вероятности, равной единице.

точному описанию информация об объектах более низкого уровня заменяется статистической закономерностью более высокого уровня и в то же время частично решается задача учета внутренней неопределенности оптимизируемой системы.

Третья операция — *выбор* из всех возможных сочетаний входных переменных *представительного подмножества*. Эта задача может решаться различными статистическими методами. Если известны распределения входных воздействий, то целесообразно применить метод Монте-Карло или методы многомерной классификации, обеспечивающие типологическую выборку. Здесь могут оказаться весьма полезными методы планирования эксперимента и другие, а также сочетания различных методов в связи с различным характером и природой входной информации.

Четвертая операция — *организация машинного эксперимента*, включающая расчет оптимальных планов и значений целевой функции для выбранного подмножества входных воздействий.

Пятая операция — *построение оператора*, или имитационной статистической модели, на основе выбранного подмножества входных и рассчитанных значений выходных переменных. В качестве такой модели будет выступать либо система регрессионных уравнений (при однородной структуре массива данных), либо система дискретно-непрерывных моделей (при структурной неоднородности массива данных).

Шестая операция — *проверка качества имитатора*, оценка его погрешности, уточнение его, корректировка.

Наконец, седьмая операция — *анализ и интерпретация результатов* имитации, установление закономерностей, корректировка самой имитируемой модели по результатам анализа.

В качестве иллюстративного примера статистической имитации кратко рассмотрим результаты моделирования гидродинамических расчетов [3]. Сущность задачи заключалась в поиске сравнительно простого статистического имитатора очень сложной и трудоемкой программы гид-

[3] Подробно см. статью «Имитационное моделирование динамики нефтедобычи (на примере Западной Сибири)» в настоящем сборнике.

родинамического расчета [6]. Входные переменные программы можно было разделить на два класса: случайные (геологопромысловые характеристики месторождений) и управляющий параметр (плотность сетки скважин). Имитация проводилась отдельно для каждого вида управления.

На основе анализа многомерных распределений первой группы факторов для условий Западной Сибири выявлена структура и сформировано представительное подмножество входных воздействий (геологопромысловые характеристики). Проведен машинный эксперимент для получения результирующего показателя (ежегодная добыча нефти на месторождении в течение 15-летнего периода). В частности, для имитации одного из распространенных управлений (плотность сетки 64 га) потребовалось 30 экспериментов. Поскольку объектом имитации выступал динамический процесс, то имитационные статистические модели строились для каждого года в отдельности. При этом обнаружен интересный факт смещения во времени структуры влияния переменных. Необходимо отметить, что даже простейший линейный имитатор на 2 фактора (в имитируемой программе было 26 переменных, гидродинамический расчет базируется на решении систем разностных уравнений) дает относительно неплохое приближение. Некоторая его модификация (линейная регрессия с 9 факторами) обеспечивает достаточно высокую точность имитатора.

Таким образом, рассмотренный пример свидетельствует о перспективности применения статистических методов имитации для описания функционирования сложных систем.

В заключение рассмотрим место имитационного моделирования на отдельных этапах экономико-математического исследования.

Любое операционалистское исследование сложных систем, в том числе и экономических, включает четыре важных этапа:

1) сбор данных и наблюдение за системой;

2) формулировку и построение математической модели системы;

3) предсказание поведения системы на основе математической модели;

4) оценку качества модели.

Машинная статистическая имитация и предназначена для осуществления перечисленных этапов исследования.

Иногда невыполним или требует больших затрат первый этап исследования. Так, например, при изучении экономической динамики могут отсутствовать данные за какой-то отрезок времени, хотя некоторые вероятностные характеристики их известны в другие периоды. Тогда можно попытаться с помощью имитации воспроизвести недостающую информацию. Если модель используется для прогнозирования, имитация позволяет получить фактически не встречавшиеся, но теоретически возможные сочетания исходных данных. Кроме того, имитированные данные совместно с эмпирическими могут служить для построения математической модели системы.

Наибольшую пользу имитация приносит на третьем этапе исследования. Современные экономические системы настолько сложны, что не поддаются аналитическому изучению с помощью традиционных методов. Трудности эти непомерно возрастают при рассмотрении реальных систем, функционирующих в условиях неопределенности. Все это вынуждает нас обращаться к численному анализу, или имитации, как к альтернативному способу исследования [7].

Итак, основная причина использования имитации — возможность преодолеть указанные трудности, с которыми сталкиваются специалисты при моделировании и управлении сложными реальными системами в условиях неопределенности. Существуют и другие преимущества имитации, способствующие ее широкому применению.

1. Имитация позволяет детально исследовать сложные внутренние взаимодействия в рассматриваемой системе.

2. С помощью имитации можно изучить влияние на систему информационных и организационных изменений.

3. Знания, полученные при разработке имитационной модели, часто становятся источником изменений в имитируемой системе. Это случай, когда опыт построения имитационной модели имеет даже бóльшую ценность, чем имитация сама по себе.

4. Имитация позволяет определить наиболее существенные переменные системы и их взаимодействие.

5. Она может быть использована для изучения новых, слабо изученных ситуаций, с помощью имитации исследуются стратегии поведения в будущем.

ЛИТЕРАТУРА

1. **Беллман Р., Заде Л.** Принятие решений в расплывчатых условиях.— В кн.: Вопросы анализа и процедуры принятия решений. М., «Мир», 1976, с. 172—215.
2. **Вальд А.** Статистические решающие функции.— В кн.: Позиционные игры. М., «Наука», 1967.
3. **Ефимов В. М., Спивак В. А.** О неопределенности и вероятности (обзор).— «Экономика и математические методы», 1972, т. VIII, вып. 5, с. 726—739.
4. **Лэсдон Л.** Оптимизация больших систем. М., «Наука», 1975. 431 с.
5. **Макаров А. А., Макарова А. С., Зейлигер А. Н.** Исследование зоны неопределенности оптимального развития сложных экономических систем.— «Экономика и математические методы», 1970, т. VI. вып. 6, с. 849—863.
6. **Борисов Ю. П., Воинов В. В., Рябинина З. К.** Особенности проектирования разработки нефтяных месторождений с учетом их неоднородности. М., «Недра», 1976. 286 с.
7. **Нейлор Т., Ботон Дж., Бердик Р.** и др. Машинные имитационные эксперименты с моделями экономических систем. М., «Мир», 1975. 500 с.
8. **Розин Б. Б.** Статистическое моделирование экономических показателей. Новосибирск, «Наука». 1976. 135 с.
9. **Kalman J. Cohen.** Computer models of the shoe, leather, hid sequence. Englewood Cliffs. N. J., 1960. 320 p.
10. **Zadeh L. A.** Fuzzy Sets.— «Inform. and Control.», 1965, N 8, p. 338—353.

Н. М. ЖУРАВЕЛЬ, Н. П. ПОНИНА

О КРИТЕРИИ МИНИМУМА ПОТЕРЬ ИНФОРМАЦИИ В ИМИТАЦИОННОМ АГРЕГИРОВАНИИ

Проблема агрегирования имеет солидную историю, но остается актуальной во многих областях знания (экономика, социология, биология). В экономике ее актуальность возрастает в связи с необходимостью системного моделирования процессов планирования и управления народным хозяйством и созданием АСУ и АСПР. Потребность в агрегировании определяется невозможностью переработки управляющими органами всей информации и ограничительными возможностями вычислительной техники.

Среди информации, циркулирующей по каналам хозяйственного организма, особое место занимают данные по отдельным видам продукции, поскольку конечная цель производственных процессов в народном хозяйстве — создание тех или иных видов продукции, служащих удовлетворению разнообразных потребностей общества. Количество видов продукции огромно (порядка 20 млн.), поэтому информационные процессы, питающие сферу планирования и управления производством, неизбежно включают в себя операции по агрегированию продукции в однородные группы.

Анализ наиболее разработанной в теоретическом и прикладном аспектах процедуры итерационного агрегирования показывает, что решены лишь вопросы назначения весов при объединении элементов в группы. Между тем даже не рассмотрены вопросы «о целесообразном числе позиций укрупненной номенклатуры, о целесообразном объединении видов продуктов (или ресурсов) в те или

иные укрупненные позиции, о количестве этапов агрегирования» [1, с. 334]. Однако, как показала практика решения отраслевых задач и их увязки с решениями по моделям более высокого уровня, именно эти вопросы становятся узким звеном в процессе получения согласованных решений и их реализации в планировании.

В работе [2] выдвинуто положение о том, что проблема агрегирования должна решаться комплексно, начиная с процедуры получения иерархии последовательно укрупняющихся классификаций агрегируемых объектов и кончая расчетом агрегатов и их взаимоувязкой. Реализация такого сложного комплекса возможна лишь на базе привлечения практического опыта агрегирования, накопленного в процессе планирования и управления. С этой целью используется имитационный подход, включающий два элемента: имитационную модель и машинный эксперимент.

Имитационная модель — это иерархия смысловых групп, заложенных в нормативные материалы (классификаторы, ГОСТы, ценники), дополненная по каждому объекту максимальным набором классификационных факторов, который формируется исходя из суперцели задачи агрегирования. Это определение применимо к объектам, для которых существуют указанные нормативные материалы. В частности, практически разработаны классификаторы продукции, видов деятельности, природных ресурсов, финансовых средств, организационных, территориальных единиц и др. [3]. Машинный эксперимент заключается в генерировании с помощью оптимальных классификационных процедур серий классификаций, которые впоследствии сравниваются и оцениваются с помощью специально вводимого критерия — минимума потерь информации.

Отсюда вытекает и название всего комплекса исследований по агрегированию — имитационное агрегирование. Итак, общая идея имитационного агрегирования заключается в следующем: *на базе имитационной модели придать процессу агрегирования оптимальный характер по критерию минимума потерь информации* (*МПИ*) с помощью оптимальных классификационных процедур.

Следует отметить в первую очередь актуальность постановки задачи об агрегировании по критерию минимума потерь информации. Насколько можно судить по литературным источникам, это направление лишь осознается как перспективное в рассматриваемой области. Нам известны только два источника, где решение проблемы агрегирования непосредственно связывается с учетом потерь информации.

Один из них — работа французских исследователей [5], выполненная в Национальном институте статистики и экономических исследований (Париж), по агрегированию производств деревообработки. Потери информации измерялись расхождением между индивидуальными и средними по группам значениями признаков.

Вторая работа — доклад чл.-кор. АН СССР Н. Н. Моисеева на конференции по исследованию операций [6], в котором относительно проблемы агрегирования он утверждает следующее: «Все работы, посвященные этой теме, имеют одно общее. Проблема агрегирования рассматривается как некоторый специальный метод решения оптимизационной задачи. Классическим примером подобного подхода является метод декомпозиции Данцига — Вулфа в линейном программировании.

Мне кажется, что решение агрегирования должно лежать совсем в другой плоскости. Необходимо четко постулировать следующий факт: в процессе агрегирования происходит не только сжатие, но (неизбежно) и потеря информации» (с. 162, 163). В этой работе конкретных предложений по способам учета потерь информации не содержится.

Две статьи настоящего сборника посвящены разработке отдельных сторон имитационного агрегирования. В данной статье рассматриваются вопросы формирования критерия минимума потерь информации и имитационной модели. В [4] излагаются алгоритмы для реализации машинного эксперимента.

В качестве объекта, на котором ведется отработка этого подхода, взята продукция химической переработки древесины. На схеме 1 приведена последовательность стадий этой переработки и указано количество видов продукции, функционирующее на каждой стадии.

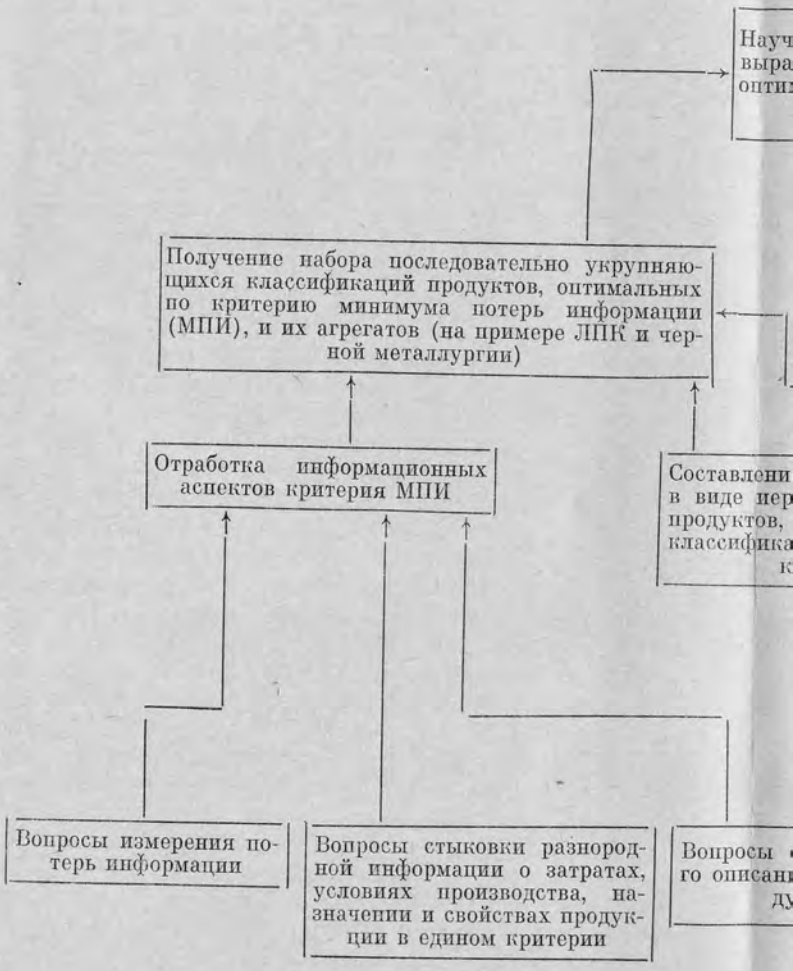
Решение сложной проблемы имитационного агрегирования лежит на стыке нескольких разделов научного знания: конкретной экономики отрасли, экономической статистики, математической статистики, распознавания образов, теории информации, математического программи-

Схема 1. Последовательность стадий химической переработки древесины.

ровани. Пяредставление о том, как они могут взаимодействовать в ходе исследования, может быть получено из начального дерева целей (схема 2), которое отражает основные смысловые узлы проблемы. Дерево целей названо начальным, поскольку по мере развязки отдельных узлов оно будет меняться в отдельных частях и достраиваться вниз путем детализации целей.

О СУЩНОСТИ КРИТЕРИЯ МИНИМУМА ПОТЕРЬ ИНФОРМАЦИИ

Решение проблемы агрегирования связано с развитием общей теории информации, учитывающей три аспекта последней: *синтактический, прагматический и семан-*



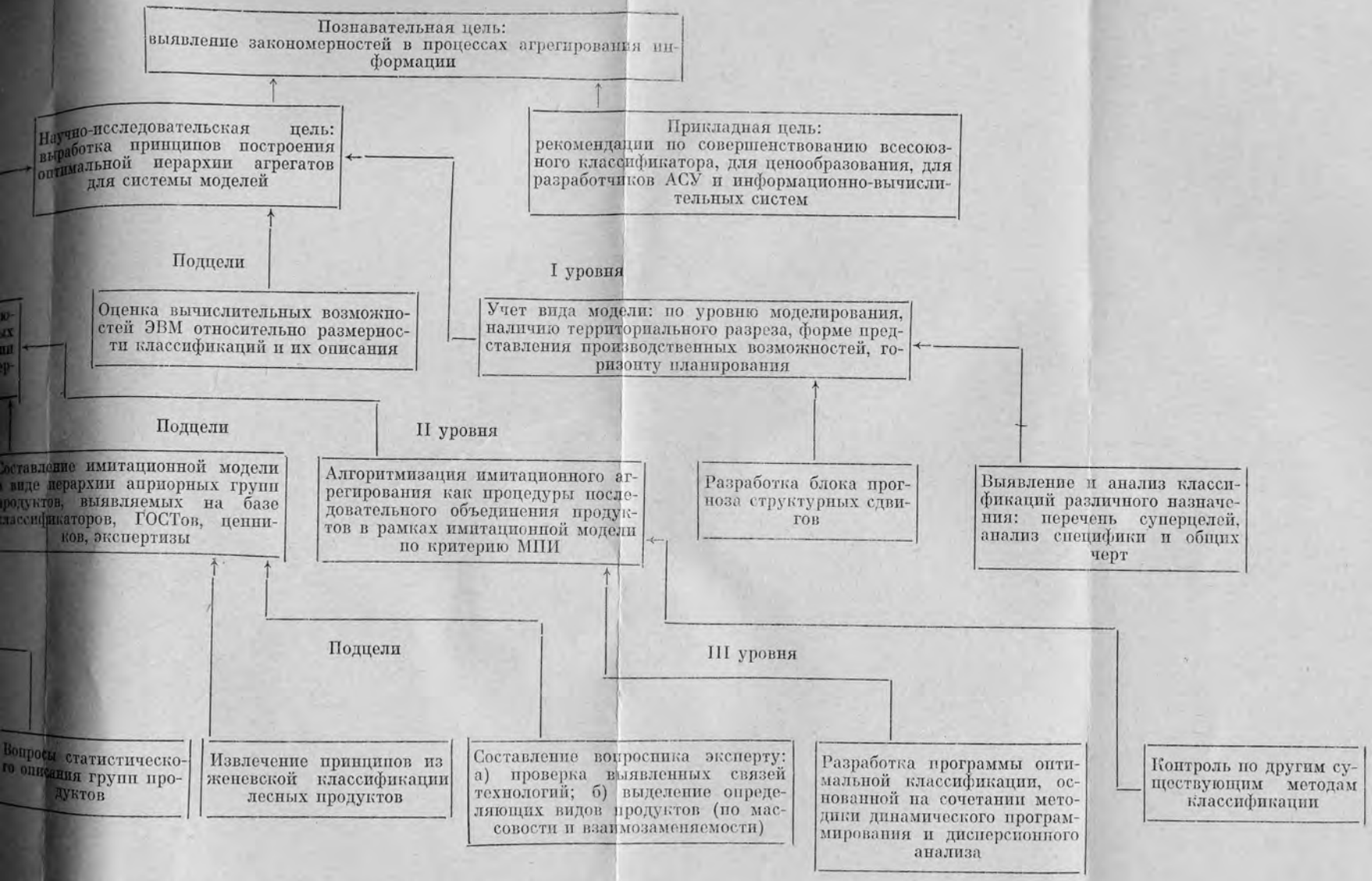


Схема 2. Дерево целей исследования по имитационному агрегированию.

тический. Классическая теория изучает формальные свойства потоков информации на синтактическом уровне, т. е. с позиций техники ее передачи. Она базируется на энтропийном определении количества информации К. Шеннона и единой размерности в битах.

Как показала практика разработки АСУ, экономическая информация не может характеризоваться только с точки зрения количества сигналов, с помощью которых она передается. Необходимо рассмотреть еще два аспекта: прагматический (ценность информации с точки зрения достижения конечных целей) и семантический (содержание и способы выражения информации) [3, 7]. Оба они исследованы мало, во всяком случае универсального метода получения непротиворечивых количественных оценок на этих уровнях не создано. Нами формирование критерия минимума потерь информации в имитационном агрегировании рассматривается именно под углом зрения прагматики и семантики.

Агрегирование информации всегда осуществляется для решения конкретной задачи, т. е. под определенную цель. В данном исследовании речь идет об агрегировании информации, относящейся к продукции, при системном моделировании территориально-производственного планирования. Наиболее крупным агрегатом в указанной системе моделей выступает отрасль или отраслевой комплекс. Поэтому задача агрегирования продукции на этапе классификации в данном случае конкретизируется так: *получить оптимальный набор последовательно укрупняющихся групп продукции отдельных отраслей или их комплексов* для моделей разного уровня детализации. Оптимум понимается здесь в двух указанных аспектах: во-первых, классификацию следует осуществлять по набору признаков, которые важны для системы моделей, учитывая тем самым ценность информации для конечной цели; во-вторых, в группу объединяемых должны включаться наиболее похожие по выделенному набору признаков продукты, так чтобы их усредненное представление содержало минимум различий в этих признаках. Здесь уместна аналогия с распознавательной задачей, которую решает наш мозг, классифицируя предметы окружающего мира по информации, доставляемой ему органами чувств: по наиболее важным признакам (прагматический аспект существенности информации) отыскивается аналог (семан-

тический аспект минимизации различий) из прошлого опыта. Эта аналогия становится особенно близкой, когда ставится задача классификации продукции будущих периодов с учетом появления новых ее видов.

Аспект существенности информации. Понятийный аппарат, используемый в системе моделей территориально-производственного планирования и разрабатываемый в Институте экономики и организации промышленного производства СО АН СССР [8], в сжатом виде дан в [9]. Анализ этого аппарата показывает, что в нем фигурируют три основных вида данных о продукции:

1) затраты на производство и транспортировку продукции (трудоемкость, материалоемкость, фондоемкость, транспортные затраты);

2) назначение, объемы и свойства продукции (виды продукции, удовлетворяющие определенную потребность, объемы производства, оценки);

3) условия производства (технологические способы, районы размещения).

Из первого вида все перечисленные составляющие затрат, за исключением транспортной,— это существенные элементы классификационного набора, поскольку через них формируется распределение основных ресурсов во всех моделях системы. При сравнении информационной ценности отдельных составляющих затрат нужно отталкиваться от приоритета тех задач, которые предстоит решать в плановом периоде на базе модельных проработок. Например, для 9-й пятилетки наиболее существенным фактором была материалоемкость. Каждый процент ее снижения экономил 4 млрд. руб. материальных затрат, что в 2,4 и в 4 раза больше результатов экономии за счет такого же снижения фонда заработной платы и капитальных затрат соответственно. Транспортные затраты, как по мнению специалистов, так и по результатам экспериментальных расчетов в системе моделей, не оказывают существенного влияния на результаты моделирования и поэтому вообще могут быть исключены из набора классификационных признаков.

Информация второго вида (о назначении и свойствах продукции) представлена в системе моделей косвенным путем — через набор потребностей в продукции определенного вида и соответствующие объемы ее производства. Между тем в нормативных документах (ГОСТы, ТУ, прейскуранты) реально существует обширная по объему непосредственная информация о назначении и свойствах продукции. Привлечение этой информации в процессе решения по системе моделей позволит лучше отразить разнообразие, взаимозаменяемость и качество продукции в исходных условиях моделирования. С этой целью, по-видимому, возможно как включение указанной информации прямо в параметры модели (через трансформацию существующих или введение новых), так и создание более совершенной схемы агрегирования продукции.

Третий вид информации — об условиях производства — характеризует технологию производства продукции и территориальный фактор. Причем технология представляется в моделях обычно в виде списка наименований конкурирующих вариантов, а различия вариантов фиксируются только через затратные элементы, которые уже включены в набор классификационных признаков. На материалах Братского ЛПК нами опробована возможность учета при классификации продукции непосредственных технологических характеристик производственного процесса. Основной вывод состоит в том, что прямое включение этих характеристик в классификационный набор возможно только для отдельных, близких по технологии, т. е. с одинаковым или близким описанием, продуктов. Для продуктов с принципиально различной технологией производства возможно включение в общий классификационный набор только самых общих характеристик (экстенсивный и интенсивный коэффициенты использования оборудования, средний разряд основных производственных рабочих, жесткость выдерживания технологического регламента).

Нужно заметить, что удовлетворительное решение вопроса о включении ГОСТовских характеристик в классификационный набор практически снимает подобный вопрос о технологии, поскольку она почти однозначно отражается в показателях назначения и качества продукции. Что касается территориального фактора, то его включение в классификационный набор полностью определяется наличием или отсутствием территориального аспекта в модели.

Предварительная проработка вопроса об информационном содержании критерия минимума потерь информации показывает, что в максимальный классификационный набор следует включить следующие факторы:

1) затраты на производство (сырьевые, трудовые, основных фондов);

2) назначение и свойства продукции (вид продукта и степень его готовности к потреблению, качественные характеристики по ГОСТу);

3) условия производства (технологический способ, средний разряд основных производственных рабочих, коэффициент использования оборудования, территориальный фактор).

Аспект минимизации различий. Рассмотрим, с помощью каких методов можно сделать «работающим» критерий минимума потерь информации. Методические трудности его применения обусловливаются двумя обстоятельствами: необходимостью свертывания информации в двух направлениях — по факторам и объектам (продуктам), а также многомерностью сжатого классификационного набора и вытекающей из этого необходимостью вычисления интегральных потерь по всему набору.

При свертывании информации используется как универсальная методика, применимая и к описанию в факторном пространстве, и к объектам, так и совокупность дифференцированных методов сжатия. Универсальная методика, основанная на сочетании идей динамического программирования и дисперсионного анализа, описывается в следующей статье данного сборника. Она специально разрабатывается для задачи оптимального агрегирования продукции, и с ее помощью можно оптимальным образом группировать и факторы, и объекты. Но группировка в каждом случае осуществляется автономно, т. е. сначала на базе классификации объектов, полученной по максимальному набору, нужно сжать описание, а затем по сжатому описанию окончательно группировать объекты.

Иной подход к разработке универсальной методики предлагается в работах московской школы по теории распознавания образов [10]: матрица исходных данных аппроксимируется некоторым набором ее подматриц. Среди них специальным алгоритмом отыскивается лучший набор, который экстремизирует функционал, характеризующий качество аппроксимации матрицы.

Любая универсальная методика в принципе позволяет полностью автоматизировать все расчеты по агрегированию информации. В этом ее преимущество, но при одном, весьма жестком условии: нужна единая исходная матрица задачи. Такое условие выполнимо только для простых задач. В нашем случае это задачи агрегирования продуктов с близкой технологией, по которым может быть получено единое описание в факторном пространстве. Расширение номенклатуры агрегируемой продукции ведет к структуризации исходной матрицы. Для универсальных методик это связано или с возникновением математически неразрешимых проблем, вызванных появлением нулевых блоков в структуре матрицы, или с вычислительными трудностями из-за размерности задачи. Решение проблемы, видимо, кроется в самом факте наличия структуры сложной задачи, которая распадается на три подзадачи:

1. *Выявление структуры*. Эта подзадача чаще всего решается неформально, на базе знаний и методов конкретных наук, с которыми связано практическое приложение решаемой задачи. В нашем исследовании для выявления структуры строится имитационная модель объекта агрегирования.

2. *Агрегирование информации внутри структурных составляющих* — с помощью универсальных методик или дифференцированных методов.

3. *Стыковка результатов.* Она может носить механический характер, т. е. представлять собой простое соединение результатов агрегирования по отдельным блокам структуры, а может быть поставлена и как самостоятельная оптимальная задача агрегирования, в которой элементарными объектами выступают структурные блоки первоначальной матрицы, т. е. осуществляется возврат к использованию методов п. 2.

Несколько слов о специфике дифференцированных методов. Количество и разнообразие их велико. Подробную характеристику этих методов можно найти в [11]. Специфика дифференцированных методов заключается в том, что их применение требует, как правило, предварительного знания о характере несжимаемой компоненты. Если отыскивается информативный набор факторов, то желательно иметь представление о типах объектов, различия в которых должен отобразить информативный набор. И наоборот, при необходимости в классификации объектов дифференцированный метод сработает тем лучше, чем существеннее классификационные факторы отражают внешнюю цель классификации.

Надо заметить, что применение методов сжатия информации требует безусловного знания материала, поскольку результаты сжатия нуждаются в содержательной интерпретации и оценке качества сжатия.

Применение информационного критерия нацелено именно на то, чтобы через количественное измерение потерь информации сделать процесс оценки качества сжатия информации более управляемым. Остановимся теперь на методических вопросах количественного измерения потерь информации.

Наш базовый тезис, от которого отталкиваются все последующие рассуждения, состоит в том, что количественно потери информации при сжатии в любом направлении (по объектам или по факторам) могут быть измерены через показатели разброса индивидуальных значений группируемых величин от среднего значения по группе.

Вид выбираемого показателя разброса и среднего значения определяется исследователями в зависимости от цели и специфики решаемой задачи сжатия информации. Наиболее употребительны дисперсия (σ^2) и среднеарифметическое значение ($\overline{x}$).

Для удобства последующего изложения введем некоторую условную матрицу $\|a_{ij}\|$, элемент которой a_{ij} соответствует нормированному значению j-го фактора (x_j) у i-го объекта (количество факторов $j = \overline{1, m}$, количество объектов $i = \overline{1, n}$).

Задача агрегирования продуктов чаще всего многомерна, т. е. формирование родственных групп продуктов происходит по нескольким общим факторам. Это положение полностью соответствует целям нашего исследования.

Пусть исходная информация, которой мы располагаем до проведения агрегирования, представлена в виде матрицы $\|a_{ij}\|$. В результате агрегирования будем иметь l групп продуктов вместо n первоначальных (горизонтальное сжатие), которые получены на основе k факторов, наиболее отвечающих целям агрегирования (вертикальное сжатие, $k < m$). Матрица размерностью $n \times m$ сжата до $l \times k$. Чтобы оценить количественно потери информации при горизонтальном сжатии, воспользуемся соотношениями, выведенными для случая сложения дисперсий зависимых и независимых величин [1], определяемых

по формулам:

$$D_i^2(x_1 + \ldots + x_k) = \sigma^2(x_1) + \ldots + \sigma^2(x_k) + 2r_{12}\sigma(x_1)\sigma(x_2) + \ldots + 2r_{1k}\sigma(x_1)\sigma(x_k) + 2r_{23}\sigma(x_2) \times \sigma(x_3) + \ldots + 2r_{2k}\sigma(x_2)\sigma(x_k) + \ldots + 2r_{(k-1)k} \times \sigma(x_{k-1})\sigma(x_k); \quad (1)$$

$$D_{\text{гор}}^2 = \sum_{i=1}^{l} D_i^2, \quad i = \overline{1, l}. \quad (2)$$

Использование формул (1) и (2) для измерения потерь информации основывается на двух посылках: базовом тезисе о принципиальной возможности представления потерь через показатели типа «дисперсия» и положении о необходимости учета взаимосвязей в многомерном векторе классификационных признаков.

По формуле (1) рассчитываются групповые потери информации как сумма дисперсий факторов сжатого набора в одной группе продуктов с учетом взаимосвязей факторов через корреляционную матрицу. Формула (2) позволяет найти суммарные потери горизонтального сжатия. Показатель $D_{\text{гор}}^2$ служит основой для расчетов по критерию на минимум потерь информации, выражая потери в абсолютном размере.

Чтобы иметь возможность сравнивать потери для группировок объектов, получаемых при разных информативных наборах, необходим показатель относительной потери информации, измеряющий величину потерь по отношению к общей информации И, содержащейся в несгруппированной матрице. В качестве такого показателя нами использована величина

где
$$\text{П} = \frac{D_{\text{гор}}^2}{\text{И}}, \quad (3)$$

$$\text{И} = \sqrt{\sum_{j=1}^{k} \left(\sum_{i=1}^{l} a_{ij} \right)^2}. \quad (4)$$

Иными словами, мерой общей информации в показателе П выступает сумма векторов описания отдельных (не сгруппированных) продуктов в сжатом факторном пространстве. Если важно учесть потери информации при вертикальном сжатии исходной матрицы $m \times n$, то вели-

чина И рассчитывается по формуле

$$И = \sqrt{\sum_{j=1}^{k}\left(\sum_{i=1}^{n} a_{ij}\right)^2}. \qquad (5)$$

ЭКСПЕРИМЕНТАЛЬНАЯ ПРОВЕРКА ДЕЙСТВЕННОСТИ КРИТЕРИЯ МПИ

Изложенные выше методические подходы к формированию и измерению функционала в критерии МПИ проверялись на фактических материалах функционирования Братского лесопромышленного комплекса. Для характеристики продукции выделены две группы входных факторов: затраты на производство продукции и условия ее производства и две группы выходных показателей: назначение и свойства продукции, объемно-ценностные показатели.

В первую группу входных факторов включены: затраты на основное сырье и материалы, энергетические затраты, заработная плата основных рабочих с начислениями, расходы на содержание и эксплуатацию оборудования, цеховые расходы, полная трудоемкость продукции.

Во вторую группу входных факторов (условия производства) включены показатели качества сырья (технологическая щепа, талловое масло, скипидар-сырец), технологические параметры и степень использования оборудования (среднесуточная выработка, скорость машин, коэффициент использования календарного времени и среднесуточное время работы).

Выходные показатели представлены качественными характеристиками конечных и побочных продуктов, а также себестоимостью продукции и объемами ее производства. Приведем сводный перечень показателей, общий по целлюлозе и картону, на основе которых осуществлялись расчеты по критерию МПИ:

Итого сырья, материалов и полуфабрикатов, руб/т . . . . . x_1
Итого энергии, руб/т . . . . . x_2
Заработная плата, руб/т . . . . . x_3
Расходы на содержание и эксплуатацию оборудования, руб/т x_4
Цеховые расходы, руб/т . . . . . x_5
Полная трудоемкость, чел.-ч/т . . . . . x_6

Содержание в щепе (%): влажности . . . . . x_7
лиственницы . . . . . x_8
крупной щепы (30×30) . . . . . x_9
нормальной » (15×15). . . . . x_{10}
мелкой » . . . . . x_{11}
опилок . . . . . x_{12}
коры . . . . . x_{13}
сучков . . . . . x_{14}
гнили. . . . . x_{15}
Общая щелочность белого щелока, г/л . . . . . x_{16}
Активная » » » г/л . . . . . x_{17}
Сульфидность белого щелока, % . . . . . x_{18}
Каустизация » » , ... x_{19}
Плотность слабого черного щелока, г/см³ . . . . . x_{20}
Среднесуточная выработка варной целлюлозы, т . . . . . x_{21}
Коэффициент использования календарного времени варочного оборудования, % . . . . . x_{22}
Среднесуточная выработка сушильных агрегатов, т. . . . . x_{23}
Скорость машин, м/мин . . . . . x_{24}
Среднесуточное время работы сушильных агрегатов, ч . . . . x_{25}
Коэффициент использования календарного времени сушильных агрегатов, % . . . . . x_{26}

Цель эксперимента. Главная цель экспериментальных расчетов заключается в проверке истинности критерия МПИ, т. е. в установлении возможности выявления с помощью этого критерия из множества группировок объектов такой, которая отличается минимальными потерями информации.

Эксперимент основывается на априорных знаниях о лучшей группировке и использовании этих знаний при сравнении результатов расчета функционала МПИ для всех группировок. Суть эксперимента состоит в следующем. Для рабочей матрицы «факторы-продукты» устанавливается стабильная схема сжатия в каком-либо одном направлении, например по продуктам ЛПК. Для Братского ЛПК закреплено неподвижно именно это направление, поскольку число продуктов невелико и заранее известна их общность:

Продукт	Группа
Целлюлоза кордная » ролевая	1
Картон всех марок	2
Скипидарные продукты талловые Канифольные » »	3

Вариация группировок идет за счет второго направления — сжатия набора классификационных факторов, для

которого априори можно предположить лучший вариант сжатия. Если в расчетах по критерию МПИ в качестве оптимального определен этот же вариант, то главная цель эксперимента достигнута: доказана реальность применения предлагаемого критерия.

Одновременно в процессе эксперимента реализуется побочная цель — разработка рекомендации о сжатом наборе факторов, который может быть использован для расчетов по широкой номенклатуре.

В ходе эксперимента испытано четыре методических подхода к сжатию входной информации:

а) выбор информативного набора факторов по методу случайного поиска с адаптацией (СПА);

б) выбор представителей из групп факторов, сформированных с помощью факторного анализа;

в) выбор представителей из априорной группировки, сформированной на базе содержательных представлений о сущности факторов и их связях;

г) компоновка информации, содержащейся в априорной группировке, в форме модуля вектора факторов каждой группы. Для удобства в дальнейшем этот подход будем называть компоновкой информации.

Указанные методические подходы применялись для сжатия двух подматриц размерностью 26×3 (26 входных факторов, описывающих 3 продукта — целлюлозу сульфатную кордную, целлюлозу ролевую и картон) и 7×5 (7 входных факторов, описывающих все 5 продуктов Братского ЛПК). Анализ выдвинутых подходов и полученных результатов группировки позволяет высказать гипотезу H_0 о том, что по объему теряемой при сжатии информации (с учетом ее представительности) подходы проранжируются следующим образом: 1 — СПА, 2—3 — факторный анализ и априорная группировка, 4 — компоновка информации.

Логическим обоснованием этой гипотезы служит то, что изучаемое факторное пространство обладает определенной структурой, важной с точки зрения выявления существенных факторов. Знание этой структуры при подходах б) и в) используется частично (берутся единичные представители содержательных групп) и при подходе г) — полностью. По методу СПА эта важная дополнительная информация не используется, поскольку наборы информативных признаков формируются случайным об-

разом. И в принципе метод не гарантирует улавливания этой структуры через информативные подсистемы факторов. Надо заметить, что при использовании в качестве результатов факторного анализа не представителей групп, а более информативных обобщенных факторов подход б) должен быть поставлен на третье место.

Результаты расчетов по критерию МПИ. Показатель относительной потери информации, выступающий в качестве функционала в критерии МПИ, рассчитан по формулам (1)—(5) и приводится в табл. 1.

Таблица 1

Подход к выбору информативных факторов	Пятипродуктовая задача		Трехпродуктовая задача	
	$D^2_{\text{гор}}/\Pi_{\text{общ}}$	$D^2_{\text{гор}}/\Pi_{\text{соб}}$	$D^2_{\text{гор}}/\Pi_{\text{общ}}$	$D^2_{\text{гор}}/\Pi_{\text{соб}}$
	1	2	3	4
СПА	1,69	2,47	0,26	1,25
Факторный анализ	1,69	2,47	0,13	0,30
Априорная группировка . .	1,38	2,26	0,12	0,56
Компоновка информации . .	0,38	0,09	0,19	0,16

Анализ полученных результатов позволяет сделать следующие выводы.

1. Проверяемая гипотеза о ранжировке подходов по величине потерь информации в целом подтвердилась (см. графы 2 и 4 табл. 1): потери минимальны при компоновке информации, максимальны — по методу СПА, факторный анализ и априорная группировка занимают промежуточное положение. Конечно, для пятипродуктовой задачи этот вывод подтверждается с известной натяжкой: по СПА и факторному анализу взят один и тот же информативный набор, и поэтому потери информации при обоих подходах равны. Более того, руководствуясь формальным принципом выбора представителя группы при факторном анализе по максимуму факторной нагрузки, нужно было фактору x_6 (полная трудоемкость) предпочесть фактор x_3 (доля заработной платы в себестоимости), и тогда бы потери информации составили 2,87. Но, во-первых, вели-

чины факторных нагрузок у этих факторов отличаются очень незначительно ($r_6 = 0{,}985$; $r_3 = 0{,}990$); во-вторых, полная трудоемкость, несомненно, имеет характер наиболее обобщающего показателя в этой группе: x_3, x_4, x_5, x_6, x_{26}. Таким образом, основная цель эксперимента достигнута: доказана возможность использования критерия МПИ в расчетах по оптимальному агрегированию объектов различной природы (продукты и факторы).

На основе предыдущего вывода можно сделать заключение о том наборе факторов, который рекомендуется в качестве желательного при классификации продукции лесного комплекса: сырьевые, трудовые, энергетические затраты, расходы на содержание и эксплуатацию оборудования. Эти факторы фигурируют в информативных классификационных наборах в более общей пятипродуктовой задаче.

ХАРАКТЕРИСТИКА ИМИТАЦИОННОЙ МОДЕЛИ ДЛЯ АГРЕГИРОВАНИЯ ПРОДУКЦИИ ЛПК

В настоящее время имитационный подход в экономико-математических исследованиях получил широкое распространение и имеет много толкований. Начало ему положено методом случайных испытаний Монте-Карло. Но в дальнейшем использование имитационной техники расширилось за счет включения диалоговых процедур в схемы модельного изучения явлений с помощью ЭВМ. К таковым можно отнести разработанные в Институте экономики и организации промышленного производства СО АН СССР модели динамического межотраслевого баланса [13] и хозрасчетной производственной единицы [14]. Но более всего существу термина «имитация», на наш взгляд, отвечает направление, разрабатываемое в ВЦ АН СССР [15]. Его характерная особенность состоит в усилении эвристического начала в процессах моделирования с целью снятия чрезмерной сложности моделей и придания им характера самосовершенствования.

В начале статьи имитационная модель для объектов, подобных рассматриваемым (продукты ЛПК), определена как иерархия априорных групп продуктов и их связей. Принципы иерархии для таких объектов следующие:

1. Выделение главных уровней, соответствующих

структуре объекта, по связям между производствами (в нашем случае три главных уровня, представленные на схеме 1).

2. Выделение внутри главных уровней Т-подуровней, различающихся технологией.

3. Дробление Т-подуровней на Н-подуровни по признаку назначения продукции и далее на отдельные виды и сорта. Приведем пример такого дробления для II уровня (схема 3).

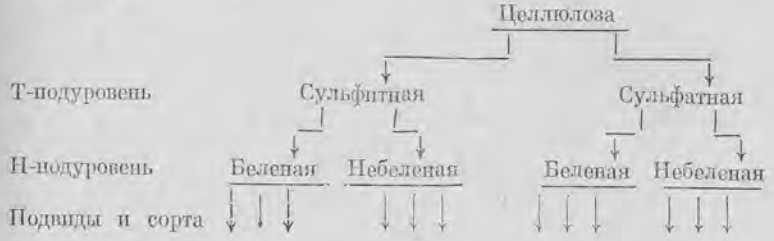


Схема 3. Дробление II уровня (блок целлюлозы) последовательности химической переработки древесины.

Наиболее обобщенным источником сведений для создания имитационной модели, в котором указанные принципы выдерживаются достаточно последовательно, являются классификаторы продукции. В нашей работе использована характеристика международной классификации лесных продуктов, разработанная 9-й рабочей сессией Объединенной рабочей группы по статистике лесного хозяйства и лесных товаров (21—25 мая 1972 г., Женева) экономической комиссии при ООН. В ней выделены 7 групп лесных продуктов, каждая из которых, в свою очередь, подразделяется на подгруппы, позиции, подпозиции (табл. 2).

Может возникнуть сомнение в необходимости привлекать какие-то формальные методы агрегирования, когда сама структура иерархии жестко формирует группы продуктов по сырьевым и технологическим связям. Но это кажущаяся жесткость. Свобода выбора, необходимая для применения формальных методов классификаций, описанных ниже, существует. Области агрегирования лежат внутри структурных подразделений типа «целлюлоза», «картон», «бумага», а также в объединении самих круп-

Таблица 2

Наименование группы	Количество		
	подгрупп	позиций	подпозиций
Необработанная древесина	2	4	22
Древесная щепа и стружка. Прочие древесные отходы	2	6	30
Древесина продольной распиловки, строганая и нестроганая, другие виды древесины простой обработки	4	12	12
Листовые древесные материалы. Однослойная фанера	5	14	28
Масса из древесины и других растительных волокон, макулатура	9	22	23
Бумага и картон	6	21	26
Необработанная, полуобработанная и обработанная пробка	3	7	7

ных подразделений. Конечно, чем дальше от начального уровня иерархии находится структурное подразделение, тем область свободы делается у́же, поскольку объединение определенным образом продуктов нижнего уровня предопределяет в известной мере структуру вышележащего уровня.

Самый сложный вопрос в информационном обеспечении имитационной модели — это учет назначения продукции. Пока видится три возможных способа введения формального описания назначения продукции: минимизированным набором свойств, имеющих количественное описание; списком назначений в виде порядкового признака; матрицей смежности или графом связей.

Начато опробование первого способа по двум структурным составляющим имитационной модели — целлюлозе и картону. Чем привлекает этот способ? Любой продукт производится в соответствии с требованиями государственного стандарта, в котором содержится количественная характеристика основных свойств, определяющих назначение и качество продукта. Через ГОСТы, следовательно, открывается доступ к унифицированной информации. Кроме того, удовлетворительное решение вопроса о включении ГОСТовских характеристик в классификационный набор практически снимает вопрос об информационном

обеспечении факторов, характеризующих технологические условия производства продукции, поскольку они в значительной мере коррелируют с показателями назначения и качества продукции.

Главная трудность при реализации первого способа заключается в разнородности описания. Так, по 19 ГОСТам картона числится в технических требованиях 63 показателя. Однако лишь 11 из них повторились в различных ГОСТах два и более раз. Для преодоления этой трудности возможны, во-первых, специальный анализ наименований свойств для выявления единых по содержанию, но имеющих разные формы представления в технических требованиях; группировка свойств по общности; поиск наиболее адекватных представителей групп; во-вторых, анализ возможностей применения метода В. Г. Гмошинского [16] обработки патентной информации, позволяющего качественные особенности научно-технических документов (патенты, стандарты и т. п.) преобразовывать в безразмерную величину — коэффициент полноты.

В заключение авторы выражают искреннюю признательность и благодарность Б. Г. Миркину за ценные советы и помощь при проведении исследований и написании статьи.

ЛИТЕРАТУРА

1. **Дудкин Л. М.** Система расчетов оптимального народнохозяйственного плана. М., «Экономика», 1972. 382 с.
2. **Журавель Н. М.** О некоторых подходах к агрегированию при системном моделировании производства.— В кн.: Проблемы экономико-статистического моделирования отдельных объектов и отраслевого звена в целом. Новосибирск, 1976, с. 104—115.
3. **Экономическая** информация (методологические проблемы). М., «Статистика», 1974. 238 с.
4. **Журавель Н. М., Ионин В. Г., Ионина Н. П.** Использование процедур оптимизации при классификации объектов и факторов.— Ст. в настоящем сборнике.
5. **Annales de l'INSEE**, 1970, mai — sept., № 4.
6. **Материалы** 1-й Всесоюзной конференции по исследованию операций. Минск, 1975. 363 с.
7. **Информация** и модели структур управления. М., «Наука», 1972. 333 с.
8. **Аганбегян А. Г., Багриновский К. А., Гранберг А. Г.** Система моделей народнохозяйственного планирования. М., «Мысль», 1972. 351 с.

9. **Журавель Н. М.** Статистические методы и сферы их применения в системном моделировании производства.— В кн.: Проблемы экономико-статистического моделирования отдельных объектов и отраслевого звена в целом. Новосибирск, 1976, с. 25—41.
10. **Браверман Э. М.** Лингвистические методы обработки эмпирических данных.— В кн.: VI Всесоюзное совещание по проблемам управления. Ч. 1. М., «Наука», 1974, с. 62—65.
11. **Айвазян С. А., Бежаева З. И., Староверов С. В.** Классификация многомерных наблюдений. М., «Статистика», 1974. 240 с.
12. **Крамер Г.** Математические методы статистики. М., «Мир», 1975. 648 с.
13. **Шатилов Н. Ф.** Анализ зависимостей социалистического расширенного воспроизводства и опыт его моделирования. Новосибирск, «Наука», 1974. 250 с.
14. **Багриновский К. А.** Моделирование процессов управления производством. М., «Знание», 1974. 63 с.
15. **Моисеев Н. Н.** Элементы теории оптимальных систем. М., «Наука», 1975. 517 с.
16. **Гмошинский В. Г.** Практика прогнозирования. Ч. 1. М., «Знание», 1972. 63 с.

Н. М. ЖУРАВЕЛЬ, В. Г. ИОНИН, Н. Н. ИОНИНА

ИСПОЛЬЗОВАНИЕ ПРОЦЕДУР ОПТИМИЗАЦИИ ПРИ КЛАССИФИКАЦИИ ОБЪЕКТОВ И ФАКТОРОВ

При классификации каждый из объектов выступает в виде набора признаков (факторов). И, на наш взгляд, необходимо, чтобы была проведена в первую очередь локальная классификация по каждому фактору в отдельности. На этом этапе важно добиваться наиболее строгого отнесения объектов к тому или иному классу с тем, чтобы в дальнейшем (при анализе объекта по совокупности признаков) можно было наиболее обоснованно говорить о значимости, потерях, экономической целесообразности включения объекта в ту или иную классификационную группу.

Введем условные обозначения:

n — число наблюдений (объектов);

i — номер наблюдения ($i = \overline{1,\ n}$);

m — число факторов, по которым проводится классификация наблюдений;

j — номер фактора ($j = \overline{1,\ m}$);

$(a_{ij})_{n \times m}$ — матрица значений фактора j по наблюдению i.

Задача заключается в том, чтобы распределить все n наблюдений (объектов) по каждому из m факторов на некоторое заранее заданное число групп. Обозначим число групп через k.

Поскольку группы должны быть составлены таким образом, чтобы выделились наиболее характерные сочетания объектов по каждому фактору, то естественно потребовать такого разбиения объектов на k групп, при котором достигается максимальная величина межгрупповой дисперсии (D_j) каждого факторного признака. Таким

образом, задача заключается в поиске

$$D_j = \max \frac{\sum_{r=1}^{k} (\bar{x}_{rj} - \bar{x}_j)^2 n_{rj}}{\sum_{r=1}^{k} n_{rj}} \quad (j = \overline{1, m}). \tag{1}$$

В этой задаче неизвестно n_{rj} — число наблюдений, относимых к классу r. Оно определяет и величину $\bar{x}_{rj}$ — среднее значение фактора по группе r. Всего необходимо решить m задач по числу рассматриваемых факторов. (Ниже излагается алгоритм решения по отдельному фактору, поэтому индекс j в дальнейшем опускается.)

Задачу оптимального разбиения n наблюдений на k групп с приведенной выше нелинейной целевой функцией и целочисленным ограничением удобнее, на наш взгляд, решать поэтапно, используя известный в динамическом программировании принцип оптимальности [1]. Сокращенно всю процедуру классификации, основанную на сочетании дискретного анализа и динамического программирования, будем обозначать ДАДП. Перед началом расчета следует перенумеровать все n наблюдений в порядке возрастания факторного признака. Процесс перенумерации наблюдений носит не только формальный характер, так как в противном случае необходим, по существу, полный ненаправленный перебор вариантов разбиения n наблюдений на k групп.

После упорядочения наблюдений все объекты, близкие по значению признака и относимые к одному и тому же классу, выступают компактной группой, а процесс оптимизации является управляемым и сводится к назначению в каждый класс такого числа примыкающих наблюдений, чтобы получить

$$\max \sum_{r=1}^{k} f_r(n_r), \tag{2}$$

где

$$f_r(n_r) = [(\bar{x}_r - \bar{x})^2 n_r] \Big/ \sum_{r=1}^{k} n_r; \tag{3}$$

$$\bar{x}_r = \left(\sum_{i=\alpha}^{\beta} a_i \right) \Big/ n_r; \tag{4}$$

$$\alpha = 1 + \sum_{\lambda=1}^{r-1} n_\lambda; \quad \beta = \sum_{\lambda=1}^{r} n_\lambda; \tag{5}$$

$$\bar{x} = \left(\sum_{i=1}^{n} a_i \right) \Big/ n \tag{6}$$

при условии

$$\sum_{r=1}^{k} n_r = n. \tag{7}$$

Поэтапное решение этой задачи можно начинать с последней стадии распределения, когда предполагается, что на предыдущих $(k-1)$ этапах произошло закрепление $(n-L)$ наблюдений, в результате чего образованы группы с номерами k, $k-1$, $k-2$, . . ., 2. Оставшиеся L наблюдений, по условию, будут иметь наибольшие значения факторного признака. Оптимальный подход по отношению к оставшимся L наблюдениям сводится к тому, что все они должны быть включены в один и тот же последний класс, которому и присваивается первый номер. Таким образом, каков бы ни был результат распределения наблюдений по предыдущим $(k-1)$ классам, заключительный этап (выделение наблюдений в первый класс) дает приращение межгрупповой дисперсии в размере

$$F_1(L) = f_1(L) = \left[\left(\frac{\sum_{i=n-L+1}^{n} a_i}{L} - \bar{x} \right)^2 L \right] \Big/ n. \tag{8}$$

Так как значение L заранее неизвестно, то расчет делается для всего интервала изменения $0 \leqslant L \leqslant n$. В результате получаем функцию приращения межгрупповой дисперсии, вычисленную до любого исхода предыдущих $(k-1)$ этапов.

Предположим теперь, что сначала выделено $(k-2)$ класса, после чего также осталось L нераспределенных наблюдений. Их следует наилучшим образом, т. е. так, чтобы получить максимальное приращение межгрупповой дисперсии, распределить между заключительными двумя классами с номерами 1 и 2. Управление оставшимися наблюдениями будет оптимальным, если

$$F_2(L) = \max_{0 \leqslant n_2 \leqslant L} \{f_2(n_2) + F_1(L - n_2)\}. \tag{9}$$

Здесь

$$f_2(n_2) = \left[\left(\frac{\sum_{i=\alpha}^{\beta} a_i}{n_2} - \bar{x}\right)^2 n_2\right]\Big/n, \qquad (10)$$

$$\alpha = n - L + 1; \qquad \beta = n - L + n_2,$$

а F_1 определена на предыдущем шаге для любого аргумента в пределах от 0 до n. Задавая L в пределах от 0 до n, получаем функцию оптимального поведения $F_2(L)$ для двухэтапного процесса (разбиение L наблюдений на два класса).

Нетрудно заметить, что и для любого числа классов r разбиение наблюдений должно удовлетворять функциональным соотношениям для $r = 2, 3, \ldots, k$:

$$F_r(L) = \max_{1 \leqslant n_r \leqslant L} \{f_r(n_r) + F_{r-1}(L - n_r)\}, \qquad (11)$$

где

$$f_r(n_r) = \left[\left(\frac{\sum_{i=\alpha}^{\beta} a_i}{n_r} - \bar{x}\right)^2 n_r\right]\Big/n. \qquad (12)$$

При выделении наблюдений в класс k нет необходимости табулировать все значения функции $F_k(L)$, так как значение L известно заранее и равно n. Тогда из соотношения

$$F_k(n) = \max_{0 \leqslant n_k \leqslant n} \{f_k(n_k) + F_{k-1}(n - n_k)\} \qquad (13)$$

определяется оптимальное число наблюдений n_k, относимых к классу k. Здесь

$$l_k(n_k) = \left[\left(\frac{\sum_{i=1}^{n_k} a_i}{n_k} - \bar{x}\right)^2 n_k\right]\Big/n. \qquad (14)$$

После выделения n_k наблюдений в группу k весь процесс можно повторить, потребовав наилучшего распределения оставшихся n' наблюдений ($n' = n - n_k$) на k' групп ($k' = k - 1$). Именно такой путь оказался более удобным

в вычислительном отношении при программировании алгоритма классификации для ЭВМ М-220.

Для иллюстрации можно привести пример решения, когда исходная совокупность наблюдений $\{a_i\}$ ($i = \overline{1, 9}$) разбивается на 4 класса с наибольшей межгрупповой дисперсией. Пусть $\{a_i\} = (1, 2, 2, 3, 6, 7, 7, 8, 9)$, тогда $\bar{x} = 5$. Результаты расчета функций $F_r(L)$ представлены в табл. 1 (расчеты выполнены без операции деления на $n = 9$).

Таблица 1

Значения функций оптимального поведения $F_r(L)$ для условного примера

L	$F_1(L)$	$F_2(L)$	$F_3(L)$	$F_4(L)$
0	0	0	0	—
1	16,0	16,0	16,0	—
2	24,5	25,0	25,0	
3	27,0	28,5	29,0	—
4	30,25	32,5	33,0	—
5	28,8	32,8	33,5	—
6	16,67	32,8	36,83	—
7	7,0	41,3	45,33	—
8	2,0	50,1	54,17	—
9	0	64,2	68,83	70,17

Напомним порядок расчета функций $F_r(L)$. Так, для одношагового процесса может оставаться нераспределенным любое число наблюдений ($L = \overline{0, 9}$). Если $L = 1$, то остается одно наблюдение $a_9 = 9$ и $F_1(1) = (a_9/1 - \bar{x})^2 \cdot 1 = 16$. При $L = 2$ остается два наблюдения: $a_8 = 8$, $a_9 = 9$, и тогда $F_1(2) = [(a_8 + a_9)/2 - \bar{x}]^2 \cdot 2 = 24{,}5$. Если $L = 3$, то остается три наблюдения: $a_7 = 7$, $a_8 = 8$, $a_9 = 9$, и для них $F_1(3) = 27$ и т. д. Принцип перебора наблюдений сохраняется и при расчете $F_2(L)$ и $F_3(L)$ по формулам (11) и (12). Например, $F_2(3) = 28{,}5$ находим следующим образом:

$$F_2(3) = \max_{0 \leqslant n_2 \leqslant 3} \{f_2(n_2) + F_1(L - n_2)\};$$

тогда $L = 3$;

при $n_2 = 0$ $F_2(3) = f_2(0) + F_1(3 - 0) = 27{,}0$;

при $n_2 = 1$ $F_2(3) = f_2(1) + F_1(3-1) =$

$$= \left(\frac{\sum_{i=n-L+1}^{n-L+n_2} a_i}{n_2} - \bar{x} \right)^2 n_2 + 24{,}5 = \left(\frac{\sum_{i=9-3+1}^{9-3+1} a_i}{1} - 5 \right)^2 \cdot 1 +$$

$$+ 24{,}5 = (7-5)^2 + 24{,}5 = 28{,}5;$$

при $n_2 = 2$ $F_2(3) = f_2(2) + F_1(3-2) = \left(\frac{\sum_{i=7}^{8} a_i}{2} - 5 \right)^2 \cdot 2 +$

$$+ 16 = 28{,}5;$$

при $n_2 = 3$ $F_2(3) = f_2(3) + F_1(3-3) =$

$$= \left(\frac{\sum_{i=7}^{9} a_i}{3} - 5 \right)^2 \cdot 3 + 0 = 27{,}0.$$

Следовательно, $F_2(3)$ будет равно 28,5 — максимальному из полученных значений.

Оптимальное разбиение на классы определяется следующим образом. Согласно формуле (13), в 4-й класс выделяется наблюдение a_1, так как при $n_4 = 1$ достигается максимальное значение $F_4(9)$ из всего ряда значений.

Действительно, при $n_4 = 0$ $f_4(0) + F_3(9-0) = F_3(9) = 68{,}83$; при $n_4 = 1$ $f_4(1) + F_3(9-1) = (a_1/1 - \bar{x})^2 \times 1 + 54{,}17 = 70{,}17$ и так далее, $n_4 = 2, 3, \ldots, 9$. В итоге получаем последовательность чисел: 68,83; 70,17; 69,83; 70,17; 69,5; 57,2; 42,5; 32,18; 0.

С остальными наблюдениями процесс решения повторяется, в результате в 3-й класс попадают наблюдения $\{a_2, a_3, a_4\}$, во 2-й — $\{a_5, a_6, a_7\}$ и в 1-й — $\{a_8, a_9\}$.

Интересно отметить, что в данном примере полученное оптимальное решение не является единственным, ибо при $n_4 = 3$ достигается тот же максимум межгрупповой дисперсии: $D = \frac{70{,}17}{9} = 7{,}8$. Таким образом можно построить и другое разбиение, а именно выделить в 4-й класс наблюдения $\{a_1, a_2, a_3\}$, в 3-й — $\{a_4\}$, во 2-й — $\{a_5, a_6, a_7\}$ и в 1-й — $\{a_8, a_9\}$. Насколько часто такая ситуация может сложиться в реальных экономических задачах, неизвестно. Во всяком случае анализ классификаций, дающих одинаковую величину межгрупповой

дисперсии наблюдаемого признака, может оказаться полезным при построении классификационной системы по совокупности факторов.

Весь процесс счета на ЭВМ М-220 занимает для данного примера 1—2 мин, включая трансляцию программы на машинный язык. Для реальных примеров время счета возрастает. Так, при разбиении 140 наблюдений на два класса время счета $\approx$ 3 мин; на три класса $\approx$ 4 мин; на четыре класса $\approx$ 5—6 мин; общее время решения пропорционально числу признаков, по которым проводится классификация наблюдений. Сопоставление группировки объектов по факторам осуществляется автоматически, по программе классификации. На печать выдаются номера наблюдений, попавших в один класс, номер класса по каждому из факторов, число наблюдений, попавших в один класс, и средние значения факторов по классу.

Группировка факторов по «венгерскому» алгоритму. Исследование степени взаимного соответствия группировок объектов по каждому фактору кладется в основу второго этапа — группировки факторов. Наиболее наглядно взаимное соответствие классификационных факторов выступает, если построить таблицу максимального числа попарных совпадений (несовпадений) номеров групп между признаком r и l $(r, l = \overline{1, m})$.

Используя эти данные, можно поставить и решить задачу, аналогичную известной задаче «о назначениях» [2], в данном случае попытаться определить, насколько согласованным оказался набор факторов, выбранных для группировки объектов. Так как нас интересуют связи каждого фактора с другими, то формируем матрицу коэффициентов совпадений (C_{ij}) таким образом, чтобы элементы на главной диагонали были равны нулю, а остальные значения — числу соответствующих попарных совпадений (несовпадений). Далее решаем задачу о назначениях, которая, как известно, сводится к выбору таких величин $\{x_{ij}\}$, чтобы максимизировать

$$\sum_{i=1}^{m} \sum_{j=1}^{m} C_{ij} x_{ij} \qquad (15)$$

при ограничениях:

$$\sum_{j=1}^{m} x_{ij} = 1 \quad (i = \overline{1, m}); \qquad (16)$$

$$\sum_{i=1}^{m} x_{ij} = 1 \quad (j = \overline{1, m}); \tag{17}$$

$$x_{ij} \geqslant 0 \quad (i,\ j = \overline{1,\ m}).$$

Решая задачу «венгерским» алгоритмом, получаем таблицу $\{x_{ij}\}$, откуда и следуют связи между факторами.

Экспериментальная проверка алгоритмов. Она осуществлялась на объектах разной экономической природы: мартеновских цехах и целлюлозных продуктах. Различно со статистической точки зрения и факторное описание этих объектов. Факторы мартеновского производства относятся к количественным и имеют сложный характер распределения. Описание целлюлозных продуктов по своему характеру ближе к качественным признакам. Приведем результаты группировки этих продуктов. Рассматривалось 18 видов целлюлозы: сульфатная беленая марок АС — 0, АС — I, АС — II, БС; сульфитная небеленая марок Ж — 1, Ж — 2, Ж — 3, Ж — 4, Ж — 5; сульфитная беленая марок А — О, А — I, Б — I, Б — II, С — I, С — II, С — III, С — IV. Согласно ГОСТам на эти продукты, их потребительские свойства имеют 8 общих факторов. Это разрывная длина, м (x_1); влажность при поставке, % (x_2); излом, число двойных перегибов (x_3); pH водной вытяжки (x_4); белизна, % (x_5); содержание смол и жиров по дихлорэтану (x_6); сорность, число соринок на 1 м² размером 0,1—1 мм² (x_7) и 9—2 мм² (x_8).

В табл. 2 приведены результаты классификации целлюлоз на два класса по каждому фактору.

Анализ классификации продуктов по всем факторам показывает, что имеют одинаковое распределение продукты: 1) 1 и 2, 2) 3 и 4, 3) 5, 6 и 7, 4) 8 и 9, 5) 10 и 18, 6) 11, 12, 13 и 17, 7) 15 и 16. Продукт 14 представляет самостоятельную группу. Таким образом, в результате двухклассовой группировки получилось 8 групп продуктов. Попытка уменьшить количество групп путем дальнейшей классификации факторов на основе их средних значений ($\overline{x}$) не привела к образованию новых групп продуктов. Это позволяет сделать вывод об устойчивости полученных групп.

Классификация на несовпадающее число классов по каждому фактору. Поскольку алгоритмом ДАДП клас-

Таблица 2

Группировка целлюлоз (двухклассовое разбиение по каждому фактору)

Номер продукта	Факторы							
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	2	2	2	2	2	1	1	2
2	2	2	2	2	2	1	1	2
3	2	2	1	2	2	1	1	2
4	2	2	1	2	2	1	1	2
5	2	2	2	2	1	2	2	1
6	2	2	2	2	1	2	2	1
7	2	2	2	2	1	2	2	1
8	1	2	1	2	1	2	2	1
9	1	2	1	2	1	2	2	1
10	2	2	2	2	2	2	1	1
11	1	2	1	2	2	2	1	1
12	1	2	1	2	2	2	1	1
13	1	2	1	2	2	2	1	1
14	1	2	1	1	2	2	1	1
15	1	1	1	2	2	2	1	1
16	1	1	1	2	2	2	1	1
17	1	2	1	2	2	2	1	1
18	2	2	2	2	2	2	1	1

сификация проводится по отдельному фактору, представляет интерес исследовать возможности для детального анализа структуры изучаемой совокупности путем комбинирования классификаций, различающихся числом классов по каждому фактору. Такой анализ позволил выявить нецелесообразность разбиения целлюлоз по фактору x_2 на 4 и более классов, по факторам x_4, x_5 — на 5 и более классов, по x_6, x_7, x_8 — на 6 классов. Нецелесообразность разбиения по указанным факторам на большее число классов проявляется в пропуске одного из номеров классов. Например, по фактору x_2 при разбиении на 4 класса получаем: 3, 3, 3, 3, 4, 4, 2, 3, т. е. нет представителя 1-го класса. Анализ исходных данных по этому фактору показал, что имеется всего лишь три градации признака: 20; 21,2 и 15. Аналогичное положение по факторам x_5, x_6, x_7, и x_8.

Из классификаций 18 продуктов по 8 факторам на два и на три класса составлена комбинированная классификация. Для этого первоначально анализировался диапазон изменений значений факторов и характер данных —

непрерывность, повторяемость. В данном примере колеблемость значений факторов была следующей:

x_1 . . .	4400—8700	x_5 . . .	0—92
x_2 . . .	15—25	x_6 . . .	0—1,5
x_3 . . .	80—2800	x_7 . . .	0—850
x_4 . . .	0—7	x_8 . . .	0—225

По факторам x_2, x_4, x_6 различные марки целлюлоз характеризуются одной и той же величиной, т. е. четко прослеживаются группы целлюлоз (например, для четырех марок сульфатной целлюлозы $x_4 = 7{,}0$; для пяти марок сульфитной небеленой — 5,5; для семи марок сульфитной беленой — 6,5). Фактор x_5, по существу, носит характер качественного (его значение либо 0, либо 81—92). С учетом этого исследовалась следующая комбинация классификации факторов:

x_1 — на три класса	x_5 — на два класса
x_2 — на два »	x_6 — на два »
x_3 — на три	x_7 — на три »
x_4 — на три »	x_8 — на два »

Сводная группировка продуктов по всем факторам представлена в табл.3. Анализ ее показал, что в результате смешанной классификации получилось 13 групп продуктов.

В целом же можно заключить, что смешанная классификация приводит к определенному усреднению унифицированных группировок, поэтому количество групп продуктов в ней меньше, чем в максимальной унифицированной классификации, но больше, чем в минимальной. Так, в нашем примере двухклассовое разбиение привело к образованию 8 групп продуктов, трехклассовое — 15 и смешанное — 13 групп продуктов. Окончательно вопрос о построении рациональной иерархии группировок может быть решен при соединении процедуры ДАДП с механизмом расчетов по критерию МПИ [1]. Эта работа находится в стадии исследований.

Группировка факторов. Задача «о назначениях» решалась для анализа взаимосвязи и группировки технико-

[1] См. ст. «О критерии минимума потерь информации» в имитационном агрегировании» в настоящем сборнике.

Таблица 3

Сводная группировка продуктов

Факторы	Продукты																		Номер класса
	1	2	3	5	6	7	8	9	10	11	12	13	14	15	16	17	18		
x_1	3	3	3	2	3	2	1	1	2	1	1	1	1	1	1	2	2		1
x_2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	2	2		2
x_3	3	3	3	3	3	2	2	1	3	3	3	3	1	3	3	3	3		3
x_4	3	3	3	2	2	2	2	2	3	3	3	3	1	3	3	3	3		4
x_5	2	2	2	1	1	1	1	1	2	2	2	2	2	2	2	2	2		5
x_6	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2		6
x_7	2	2	2	3	3	3	3	3	1	1	1	1	1	1	1	1	1		7
x_8	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1		8
Номер класса	1	2	2	3	4	5	6	9	8	9	9	9	10	11	11	12	13		

экономических показателей производства некоторых видов продукции Братского лесопромышленного комплекса. Исследование включает предварительную классификацию наблюдений на заданное число классов по каждому фактору в отдельности таким образом, чтобы межгрупповая дисперсия была максимальной. Полученные результаты служат исходной базой для разбиения факторов на непересекающиеся между собой классы. Количество классов при этом заранее не задается, а получается в результате решения задачи. Для исследования взята информация за ряд лет, единицами наблюдения приняты продуктотоды [2].

Результаты группировки наблюдений на два класса по 26 факторам приведены в табл. 4. Данные этой таблицы используются для расчета числа совпадений (несовпадений) номеров групп для факторов i и j ($i, j = \overline{1,26}$). По каждой паре факторов выбирается наибольшее из двух чисел: совпадений или несовпадений номеров групп и используется в качестве коэффициента целевой функции при решении задачи «о назначениях». Но так как связь каждого фактора с самим собой нас, естественно, не интересует, то полагаем все коэффициенты C_{ij} равными нулю. Матрица коэффициентов C_{ij} приводится в табл. 5. В результате решения задачи о прикреплении фактора к фактору на максимум общего числа совпадений была получена следующая система назначений (стрелками указано направление прикрепления):

Группа	
1	$x_1 \to x_6 \to x_3 \to x_2 \to x_1 \to x_1$
2	$x_5 \to x_{10} \to x_5$
3	$x_7 \to x_{15} \to x_7$
8	$x_8 \to x_9 \to x_8$
5	$x_{11} \to x_{18} \to x_{11}$
6	$x_{12} \to x_{13} \to x_{12}$
7	$x_{14} \to x_{19} \to x_{16} \to x_{20} \to x_{14}$
8	$x_{17} \to x_{22} \to x_{21} \to x_{25} \to x_{26} \to x_{17}$
9	$x_{23} \to x_{24} \to x_{23}$

В одну группу вошли факторы, характеризующие затраты на производство единицы продукции (группа 1). Группы факторов 3, 4 и 6 отражают различные свойства состава и качества щепы.

[2] В упомянутой выше статье этого сборника.

Таблица 4

Таблица кофакторной группировки наблюдений

Номер наблюдения	Факторы																									
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}
1	2	2	2	2	1	2	2	2	1	2	1	1	2	1	1	1	2	2	2	1	2	2	1	1	1	1
2	2	2	2	2	1	2	2	2	1	2	1	1	2	1	1	1	2	2	2	1	2	2	1	1	1	1
3	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
4	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
5	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
6	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
7	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
8	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
9	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
10	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
11	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
12	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
13	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
14	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
15	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
16	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
17	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2
18	2	2	2	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2	2	1	2	2	1	1	2	2

Таблица 5

Матрица коэффициентов целевой функции

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	0	18	18	18	15	18	10	11	12	14	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
2	18	0	18	18	15	18	10	11	12	14	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
3	18	18	0	18	15	18	10	11	12	14	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
4	18	18	18	0	15	18	10	11	12	14	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
5	15	15	15	15	0	15	9	10	12	17	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
6	18	18	18	18	15	0	10	11	12	14	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
7	10	10	10	10	9	10	0	11	12	14	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
8	11	11	11	11	10	11	11	0	13	14	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
9	12	12	12	12	9	12	12	13	0	10	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
10	14	14	14	14	17	14	14	14	10	0	11	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
11	11	11	11	11	10	11	11	11	13	14	0	16	12	11	14	11	10	11	10	10	15	13	18	18	15	13
12	16	16	16	16	16	16	16	16	16	16	16	0	12	11	14	11	10	11	10	10	15	13	18	18	15	13
13	12	12	12	12	12	12	12	12	12	12	12	12	0	11	14	11	10	11	10	10	15	13	18	18	15	13
14	11	11	11	11	11	11	11	11	11	11	11	11	11	0	14	11	10	11	10	10	15	13	18	18	15	13
15	14	14	14	14	14	14	14	14	14	14	14	14	14	14	0	11	10	11	10	10	15	13	18	18	15	13
16	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	0	10	11	10	10	15	13	18	18	15	13
17	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	0	11	10	10	15	13	18	18	15	13
18	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	10	0	10	10	15	13	18	18	15	13
19	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	0	10	15	13	18	18	15	13
20	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	0	15	13	18	18	15	13
21	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	0	13	18	18	15	13
22	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	0	18	18	15	13
23	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	0	18	15	13
24	18	18	18	18	15	18	10	11	12	14	11	16	12	11	14	11	10	11	10	10	15	13	18	0	15	13
25	15	15	15	15	14	15	11	10	9	13	12	13	13	10	11	10	11	10	13	9	14	16	15	15	0	16
26	13	13	13	13	13	13	9	10	11	11	10	11	11	10	13	10	15	10	11	11	16	18	13	13	16	0

По характеру группировки наблюдений подобными оказались цеховые расходы и содержание в сырье нормальной щепы (группа 2); процент мелкой щепы в сырье и сульфидность белого щелока (группа 5); содержание в сырье сучков и характеристики белого и черного щелоков: каустизация белого щелока, общая щелочность белого щелока и плотность слабого черного щелока (группа 7); величина активного белого щелока, коэффициенты использования варочного оборудования и сушильных агрегатов, время работы сушильных агрегатов и выработка вареной целлюлозы (группа 8); выработка сушильных агрегатов и скорость машин (группа 9). Например, от подготовки древесины в виде щепы и от ее состава действительно зависят сульфидность щелока и состав белого и черного щелоков, хотя это и не является единственным условием, определяющим их расход.

Полученные группы факторов могут быть приняты для отбора существенных из них при агрегировании продукции целлюлозно-бумажной промышленности по технико-экономическим показателям, а изложенный выше метод можно использовать и в других отраслях промышленного производства.

ЛИТЕРАТУРА

1. **Беллман Р.** Динамическое программирование. Пер. с англ. М. ИЛ, 1960. 400 с.
2. **Юдин Д. Б., Гольштейн Е. Г.** Задачи и методы линейного программирования. М., «Сов. радио», 1964. 736 с.

статистической модели, которая как раз и предназначена для прогнозирования результирующих показателей деятельности производственного объекта. Таким образом, величины $\overline{N}_k^j$ и $\underline{N}_k^j$ можно оценить с помощью статистических зависимостей (8), (9), с учетом многообразных, не обязательно вытекающих из модели соображений.

Но такое задание ограничений возможно лишь для некоторого предприятия данной группировки объектов. Все остальные предприятия этой группировки должны копировать задание ограничений в модели базового объекта. Индивидуальные особенности конкретных предприятий могут быть учтены с помощью статистических зависимостей. Таким образом, не принимаются во внимание дополнительные, неформальные соображения по поводу развития каждого данного предприятия.

В качестве базового рассмотрим идеализированный усредненный объект данной группировки предприятий I. Значения независимых факторов для него определяются как средние арифметические индивидуальных значений факторов отдельных объектов группировки I:

$$\hat{x}_i^I = \sum_{j \in S^I} \hat{x}_i^j / l_i.$$

Исходя из значений $\hat{x}_i^I$, определяем границы ограничений $\underline{M}_i^I$, $\overline{M}_i^I$, $\underline{N}_k^I$, $\overline{N}_k^I$, а по ним — индивидуальные для каждого отдельного объекта ограничения $\underline{M}_i^j$, $\overline{M}_i^j$, $\underline{N}_k^j$, $\overline{N}_k^j$ через коэффициенты отклонения индивидуальных значений независимых факторов от их усредненных значений каждой группировки I: $\lambda_i^j = \hat{x}_i^j / \hat{x}_i^I$,
где $\hat{x}_i^j$ — фактическое значение i-го независимого фактора для j-го предприятия ($j \in S^I$; $I = 1, \ldots, \tau$). Он же используется для определения $\hat{x}_i^I$. Очевидно, $\sum_{j \in S^I} \lambda_i^j = 1$.

Значения $\overline{M}_i^j$, $\underline{M}_i^j$, $\overline{N}_k^j$, $\underline{N}_k^j$ рассчитываются так:

$$\underline{M}_i^j = \underline{M}_i^I / \lambda_i^j, \quad \overline{M}_i^j = \overline{M}_i^I / \lambda_i^j; \tag{15}$$

$$\overline{N}_k^j = \overline{N}_k^I - \sum_{i=1}^{n} (1 - \lambda_i^j) \overset{\circ}{x}_i^I a_{ik}^I,$$

$$\underline{N}_k^j = \underline{N}_k^I - \sum_{i=1}^{n} (1 - \lambda_i^j) \overset{\circ}{x}_i^I a_{ik}^I. \tag{16}$$

В формуле (16)

$$\overset{\circ}{x}_i^I = \begin{cases} \overline{M}_i^I \text{ при } (1 - \lambda_i^j) < 0, \\ \underline{M}_i^I \text{ при } (1 - \lambda_i^j) > 0; \end{cases}$$

$$\overset{\circ}{x}_i^I = \begin{cases} \underline{M}_i^I \text{ при } (1 - \lambda_i^j) < 0, \\ \overline{M}_i^I \text{ при } (1 - \lambda_i^j) > 0. \end{cases}$$

Если ограничения в задаче А для всех рассматриваемых в ней объектов заданы посредством перерасчета их значений для усредненного объекта каждой группировки I с помощью соотношений (15)—(16), то можно показать, что решение этой задачи после разагрегирования будет решением задачи АА. Доказательство этого факта приведено в работе [5].

Изложенный выше способ формирования агрегированной системы моделей и получения на его основе решения для расширенной совокупности моделируемых объектов был экспериментально проверен на примере лесозаготовительных предприятий Красноярского края (Всесоюзное производственное объединение «Красноярсклеспром», в состав которого входит около 60 леспромхозов). При статистическом моделировании использовалась отчетная информация за 1970—1974 гг.

Статистическое моделирование показателей лесозаготовительной отрасли уже имеет свою историю. Это объясняется рядом обстоятельств, в частности сравнительно однородным потребительским назначением продукции, высокой степенью монопродуктивности и известной однотипностью технологических процессов и принципов организации производства. Довольно большое количество предприятий в отрасли и в отдельных территориальных группировках, четко выраженная дифференциация последних создают благоприятные условия для применения математико-статистических моделей в анализе производственно-хозяйственной деятельности лесозаготовительных предприятий (леспромхозов).

Не вызывает возражений и гипотеза о том, что в течение относительно небольшого промежутка времени процесс функционирования лесозаготовительного предприятия не претерпевает сколько-нибудь существенных изме-

нений. За этой гипотезой стоит возможность распространения полученных в результате ретроспективного анализа закономерностей на ближайшую перспективу и т. п. [2].

Накопленный опыт и методические приемы были использованы при конструировании экономико-статистической модели лесозаготовительного предприятия. При этом учитывались следующие ограничения, исходя из которых формировались определенные структурные особенности модели.

1. *Конструкция модели верхнего уровня.* Здесь ориентиром служили стандартные постановки моделей оптимального отраслевого планирования. Поэтому в соответствии с назначением модели предприятия необходимо было среди моделируемых выходных показателей предусмотреть те, которые одновременно выступают параметрами технологических способов (вариантов) модели и коэффициентов ее целевой функции. Такими показателями стали: численность промышленно-производственного персонала (чел.), среднегодовая стоимость основных промышленно-производственных фондов (тыс. руб.), затраты на заготовку древесины (тыс. руб.), объем заготавливаемой древесины (тыс. $м^3$).

Кроме того, в качестве локального ограничения выступал натуральный показатель производительности труда — комплексная выработка в кубометрах вывезенной древесины на одного занятого в лесозаготовках.

2. Включение *экономико-статистической модели предприятия в разработанную схему координации моделей разных уровней.* Это нашло свое отражение в форме модели — линейные уравнения регрессии были сведены к общей задаче линейного программирования с функционалом на минимум затрат.

3. *Адекватность описания моделируемых показателей.* Данное требование необходимо рассматривать по крайней мере в двух аспектах. Во-первых, насколько всесторонне структура экономико-статистической модели отражает особенности формирования того или иного моделируемого показателя; во-вторых, в какой степени колебания включенных в модель независимых признаков совпадают с колебаниями моделируемого показателя, другими словами, насколько значимы факторы, составляющие структуру экономико-статистической модели. Может случиться

так, что в состав значимых факторов войдут далеко не все представители выделенных групп.

Проблема отбора факторов для моделей предприятий лесозаготовительной промышленности в последнее время получила известное развитие. В частности, в работе [1] на базе исследования более чем 100 факторов и нескольких моделируемых показателей для леспромхозов, расположенных на территории европейской части СССР, Западной и Восточной Сибири, Дальнего Востока и т. д., был получен достаточно устойчивый круг значимых факторов как в привязке к району, так и в общем случае. Поэтому при построении экономико-статистической модели отбор факторов как самостоятельный вопрос не ставился.

Рекомендуемый набор факторов взят за основу и дополнен рядом других. В конечном счете в состав независимых переменных вошли: фондовооруженность труда, производительность (на среднесписочные трактор, погрузчик и автомобиль), удельный вес рабочих, занятых на лесозаготовках, удельный вес деловой древесины, породный состав насаждений, запас древесины на 1 га, средний объем хлыста, среднее расстояние вывозки, удельный вес вывоза леса к железной дороге. При построении экономико-статистической модели предприятия приходилось учитывать возможности имеющегося информационного обеспечения, что накладывало свой отпечаток на состав как независимых переменных, так и моделируемых показателей.

4. *Необходимость проведения расчетов на перспективу.* Как правило, для прогнозирования используются динамические экономико-статистические модели. Их структура в общем случае должна характеризовать влияние двух групп факторов: статических (различие природно-географических условий, технологии производства, объема и структуры промышленно-производственных фондов и т. д.) — пространственное сечение информации — и динамических (предыстория развития каждого объекта, влияние научно-технического прогресса и т. д.) — временное сечение информации. Будем считать, что состав значимых факторов первой группы нам известен. Факторы же, отражающие динамику, привносились в модель двояким образом.

Во-первых, отправным моментом при построении экономико-статистической модели лесозаготовительного пред-

приятия послужил вывод о том, что значение и уровень важнейших экономических показателей его работы определяется не только природными, организационно-техническими и экономическими условиями, но и его положением на траектории развития [2]. Рядом исследований установлено, что леспромхозы имеют отчетливо выраженную траекторию развития во времени, на которой выделяются три этапа [2]: постепенное нарастание объемов лесозаготовок за счет использования избытков исходного запаса спелых лесов, стабилизация на каком-то высшем уровне и последующее сокращение в меру исчерпания избытка спелых лесов в данном районе и продвижения лесной промышленности в новые районы.

Это положение предопределяет необходимость разбиения всей исследуемой совокупности предприятий на три соответствующие этапам группы и описания каждой из них своей особой моделью или набором внутриэтапных моделей (с определенным для каждого конкретного случая принципом внутриэтапного дробления). В таких моделях динамика в косвенном виде уже учтена.

Во-вторых, в конструируемую экономико-статистическую модель были введены специальные факторы, несущие временну́ю (динамическую) нагрузку — лаговые переменные (показатели-функции с лагом в один год). Ими стали следующие показатели предшествующего года: численность промышленно-производственного персонала, стоимость основных промышленно-производственных фондов, себестоимость лесозаготовок, объем заготовок древесины, производительность труда.

Модель, построенная на таких принципах, может рассматриваться как один из возможных вариантов динамической экономико-статистической модели лесозаготовительного предприятия.

В соответствии с гипотезой о трехэтапном развитии лесозаготовительного предприятия, а также с учетом таких признаков, как мощность, уровень фондовооруженности и пункт примыкания, исходный массив информации был разбит на 6 однородных групп. При этом 5 групп наблюдений соответствовали второму и 1 группа — третьему этапу развития леспромхоза. Это объясняется тем, что рассматриваемый период (1970—1974 гг.), как и непосредственно предшествовавший ему (1965—1968 гг.), не был отмечен вводом в действие новых лесозаготовительных предприятий на территории Красноярского края.

При классификации наблюдений выделены 6 групп леспромхозов, представленных в таблице. Точность прогноза (прогнозируемая величина результирующих показателей на 1974 г. сравнивалась с ее фактическим значением в этом году) по моделям всех типов оказалась приблизительно одинаковой (со средней ошибкой прогноза в 4%).

Ставилась задача при установленных ограничениях на независимые факторы, производственную мощность каждого из предприятий, задание по заготовке древесины, лимиты основных производственных фондов и численность трудовых ресурсов для всех выделенных леспромхозов территориальной группировки минимизировать полные затраты на заготовку древесины.

С учетом операций агрегирования и разагрегирования такая задача была сведена к распределению плановых заданий по заготовке древесины, лимитов ресурсов многоцелевого назначения только между шестью группировками лесозаготовительных предприятий. Это позволило существенно сократить размерность решаемой задачи (вместо 506 переменных и 1415 ограничений имели 168 переменных и 615 ограничений).

При таком подходе решение задачи включало следующие этапы:

а) формирование усредненного объекта для каждой статистически однородной группы предприятий;

б) определение ограничений на результирующие показатели усредненного объекта;

Классификация леспромхозов Красноярского края

Номер группировки	Этап развития	Объем заготовки древесины, тыс. м³	Фондовооруженность, руб.	Пункт примыкания, %
1	II	200—360	2,4—4,5	40—80
2	II	90—250	3,6—6,3	0—15
3	III	310—200	3,5—5,0	0—60
4	II	100—200	3,5—5,7	50—80
5	II	400—670	3,7—6,6	0—20
6	II	250—400	3,5—6,5	0—20

в) построение агрегированных моделей путем умножения ограничений для усредненного предприятия на число объектов агрегируемой группы и задание ограничений, распределяемых между группами;

г) координацию и нахождение согласованного решения для агрегированных моделей;

д) получение индивидуальных значений независимых признаков и результирующих показателей с помощью коэффициентов отклонения индивидуальных значений независимых факторов от их усредненных значений по каждой группировке (разагрегирование).

Наиболее трудоемким этапом оказался процесс координации моделей. Ускорению агрегирования и разагрегирования получаемых планов способствует их автоматизированное выполнение. В нашем случае для достижения согласованного решения, обеспечивающего 2%-ную точность, потребовалось 8 итераций. При этом затраты на заготовку древесины по сравнению с исходным планом сократились на 13,1%. Получены разагрегированные планы, удовлетворительные с точки зрения принятых гипотез и содержательной трактовки тенденций развития предприятий лесозаготовительной отрасли.

Включение экономико-статистических моделей в процесс согласования решений в многоуровневой системе налагает определенные требования на форму этих моделей. О некоторых требованиях уже говорилось выше. Однако проблема не исчерпывается только выполнением формальных условий.

Статистическая модель по своему существу является информационной. Ее цель — генерирование информации, необходимой для определенных целей исследователя, как правило, для прогноза результирующих показателей. Обмен информацией между моделями разных уровней, один из которых (верхний) представлен нормативной, а другой (нижний) — экономико-статистической моделью, предполагает устранение принципиальных различий в характере и содержательном назначении вырабатываемых показателей.

Нормативная модель — это формализация идеальных представлений о развитии объекта. Поэтому у верхнего уровня (центра) есть заложенные априори представления о желаемом поведении подсистем. Они могут совпадать с действительностью, которую отображают статистические

модели, а могут и не совпадать. Необходимо, следовательно, добиться некоторого соответствия статистических моделей нижнего уровня представлениям центра. Почему не наоборот? Мы исходим из примата нормативных представлений над статистическими, из примата интересов более высокого уровня над интересами нижнего уровня.

Достижение соответствия статистических моделей нижнего уровня представлениям центра выдвигает новую проблему, которую необходимо решать при построении экономико-статистических моделей. Речь идет о том, чтобы структурные особенности этих моделей согласовывались с требованиями модели верхнего уровня.

Прежде всего уясним требования нормативной модели (верхнего уровня). Все результирующие факторы-ограничения делятся на качественно однородные группы, например продукты и ресурсы. Первое требование в том и состоит, чтобы ограничения каждой группы совершенно определенным образом воздействовали на значение целевой функции модели нижнего уровня. Так, например, если в качестве целевой функции выступает уравнение полной себестоимости, то увеличение используемых ресурсов при прочих равных условиях должно по крайней мере не повышать затрат.

Второе требование сводится к тому, чтобы все независимые факторы также определенным образом влияли на величину результирующих показателей. Например, увеличение среднего расстояния вывозки при прочих равных условиях должно уменьшать комплексную выработку на одного рабочего. Не всегда, однако, удается заранее предвидеть влияние каждого фактора на результат. В этом случае можно довериться тому, как объясняет то или иное взаимодействие факторов статистическая модель.

Требования нормативной модели отрасли к статистическим моделям нижнего уровня можно отразить в виде системы знаков, с которыми те или иные факторы должны входить в уравнения результирующих показателей.

Тогда построение экономико-статистической модели в соответствии с требованиями нормативной превращается в итеративный процесс: формируется последовательность систем регрессионных уравнений, отвечающих априорно заданной структуре и обеспечивающих удовлетворительный коэффициент детерминации. Такой процесс не всегда заканчивается построением экономико-статистической мо-

дели в соответствии с представлениями центра. Он может показать и невозможность формирования такой модели. В последнем случае необходимо выполнить корректировку постановки задачи верхнего уровня в соответствии с выявленными особенностями поведения объектов нижнего уровня и возобновить процесс построения экономико-статистической модели.

ЛИТЕРАТУРА

1. **Кулешов В. В., Лукацкая М. Л.** Исследование и отбор факторов для статистической модели лесозаготовительного предприятия.— В кн.: Распознавание образов и регрессионный анализ. Новосибирск, 1972, с. 74—94.
2. **Кулешов В. В., Лукацкая М. Л., Талышева Л. П.** Исследование некоторых особенностей динамики функционирования лесозаготовительных предприятий.— В кн.: Распознавание образов и регрессионный анализ в экономических исследованиях. Новосибирск, 1972, с. 179—198.
3. **Базилевич А. А.** Постановка задачи линейного программирования с применением данных корреляционного анализа.— «Экономика и математические методы», 1967, № 1, с. 83—87.
4. **Поспелов Г. С., Ириков В. А.** Программно-целевое планирование и управление. М., «Сов. радио», 1976 (гл. 3, § 5, 6).
5. **Третьяков А. А., Шишацкий Н. Г.** Элементы агрегации на базе экономико-статистических моделей производственных объектов.— В кн.: Проблемы экономико-статистического моделирования отдельных объектов и отраслевого звена в целом. Новосибирск, 1976, с. 96—103.

Н. Е. АЧЕЛАШВИЛИ

НЕКОТОРЫЕ ВОПРОСЫ ПОСТРОЕНИЯ СИСТЕМЫ ЭКОНОМИЧЕСКИХ ОЦЕНОК ПРИРОДНЫХ УСЛОВИЙ В УГОЛЬНОЙ ПРОМЫШЛЕННОСТИ [1]

Исходными данными в настоящем исследовании служили показатели работы угольных шахт Кузнецкого бассейна за период с 1963 по 1972 г. В качестве основы экономических оценок использовались четыре показателя: производительность труда рабочего по добыче, себестоимость и фондоемкость продукции, годовой объем добычи угля. Полный набор факторов производства и их измерителей, включенных в анализ, представлен в работе [1, с. 95—123].

В процессе исследования применялись различные модификации двух типов экономико-статистических моделей — непрерывных и дискретно-непрерывных.

Анализ непрерывных моделей показал, что в силу неоднородности исходной совокупности они не дают удовлетворительных результатов как с позиции формальных статистических критериев, так и применительно к существу поставленной задачи [1, с. 164—182].

Прежде всего следует отметить, что построение уравнений регрессий по всей совокупности не всегда возможно вследствие низкой корреляционной связи между экономическими показателями и производственными факторами. Так, например, среди исходных данных за 1972 г. достаточную корреляционную связь для построения моделей имели лишь объем добычи и себестоимость (множествен-

[1] Общую постановку задачи построения экономических оценок природных условий и методические подходы к ее решению см. в работе [1].

ные коэффициенты детерминации составили соответственно 0,75 и 0,50), а по фондоемкости и производительности труда вообще нельзя было получить представительного описания. Такая слабая корреляционная связь объясняется тем, что в исходной общей совокупности в силу ее неоднородности разнонаправленное, а часто и противоположное влияние факторов искажает реальную картину. В частности, разнонаправленное влияние имеет изменение характера падения пластов: при пологом падении с увеличением угла падения уровень экономических показателей снижается, а при крутом — растет.

Кроме того, при построении системы экономических оценок на основе общей непрерывной модели неприродные факторы закрепляются на среднем уровне. При этом следует быть уверенным в том, что те сочетания факторов, для которых определяется значение зависимой переменной, будут совместимы между собой. В данном случае надо установить, совместимы ли сочетания средних значений факторов технического уровня с конкретными природными условиями. Однако вариация технических уровней вследствие дифференциации природных условий для шахт Кузнецкого бассейна весьма существенна. Поэтому усреднение факторов технического уровня по всей совокупности лишено смысла: определенному типу природных условий соответствует определенный технико-технологический уровень. Следовательно, выбрать типичное среднее предприятие в данном случае не представляется возможным и система экономических оценок не будет отражать реальной ценности тех или иных комплексов природных условий.

Во избежание указанного недостатка использовалась классификация шахт по природным условиям, и для каждого выделенного класса брались соответствующие средние значения факторов технического уровня. Далее, подставляя в уравнения индивидуальные значения природных условий и усредненные характеристики технического уровня (по соответствующим классам природных условий) и экономико-организационных факторов (по всей совокупности), мы получали условную величину экономического показателя, которая и служила основой расчета экономических оценок природных условий.

Однако этот подход также не дал удовлетворительных результатов, так как велика ошибка аппроксимации,

вызванная тем, что общая зависимость описывает «средний» объект исследуемой совокупности. Проведенный анализ позволил сделать вывод, что использование непрерывных моделей для построения системы оценок природных условий для шахт Кузнецкого бассейна неэффективно. Решить поставленную задачу для условий Кузнецкого бассейна, как показал и предшествующий экономико-статистический анализ [1, с. 164—182], помогут дискретно-непрерывные модели.

При построении системы экономических оценок природных условий на основе дискретно-непрерывных моделей использовались два способа.

1. Проводилась укрупненная классификация шахт по характеру падения пластов и строились внутриклассовые регрессионные модели по «полному» набору факторов производства. Необходимость построения уравнений от всех факторов производства связана с тем, что данная классификация имеет ограниченный характер. Во-первых, используется не полный комплекс природных условий, а только одна его характеристика, правда, наиболее существенная. Во-вторых, данное разбиение не элиминирует в полной мере влияния самого группировочного фактора.

2. С использованием результатов многомерной классификации по «полному» комплексу природных условий строились внутриклассовые регрессионные модели, характеризующие зависимость экономических показателей от природных факторов. Модели были двух видов: для расчета укрупненной экономической оценки на уровне определенных типов природных условий и для получения экономической оценки природных условий на уровне отдельных объектов внутри классов.

В первом случае использовались внутриклассовые регрессионные уравнения, характеризующие зависимость экономического показателя от факторов технического и экономико-организационного уровней. При этом оказалось, что в результате многомерной классификации элиминируется влияние природных условий и, следовательно, внутри полученных классов вариация экономических показателей обусловлена только вариацией факторов технического и экономико-организационного уровней.

Во втором случае внутри каждого класса строились уравнения, описывающие связь между отклонениями эко-

номических показателей и отклонениями технических и экономико-организационных факторов от соответствующих внутриклассовых средних:

$$\Delta y_i = y_i - \overline{y};$$

$$\Delta x_{ij} = x_{ij} - \overline{x_j};$$

$$\Delta y = f(\Delta x. \ . \ . \ \Delta x_j. \ . \ . \ \Delta x_m),$$

где i — номер объекта, j — номер фактора.

ГРУППОВЫЕ ЭКОНОМИЧЕСКИЕ ОЦЕНКИ ОТДЕЛЬНЫХ ТИПОВ ПРИРОДНЫХ УСЛОВИЙ

Поскольку система экономических оценок природных условий охватывает значительный промежуток времени, естественно ожидать изменений технических уровней объектов, а в зависимости от этого и абсолютных величин оценок. Поэтому для обеспечения сопоставимости оценок во времени необходимо учитывать особенности технического развития объектов, находящихся в различных природных условиях.

В качестве примера проанализируем динамику технического уровня трех наиболее типичных классов шахт Кузнецкого бассейна. В основе классификации лежит одна из наиболее важных характеристик природных условий — характер падения пластов.

Первый класс характеризуется следующим комплексом природных условий: пологим залеганием пластов (на пласты с углом падения 0—15° приходится 70—90% добычи угля, остальное—на пласты с углом падения 16—35°), мощностью пластов в пределах 1,8—3,6 м, газоносностью 3—10 м³/т, глубиной разработки 180—220 м, водообильностью 200—350 м³/ч.

Этот комплекс природных условий является наиболее благоприятным, так как для него хорошо отработана технологическая схема, имеется возможность широко и эффективно использовать высокопроизводительные средства механизации во всех основных звеньях производственного процесса.

Анализ данных, приведенных в табл. 1, показывает, что на шахтах рассматриваемого класса уровень механизации подготовительных работ за 1963—1972 гг. возрос

Таблица 1

Динамика характеристик шахт первого и второго классов

Год	Характеристика технического уровня								Нагрузка на очистной забой, т/сут	Среднемесячное подвигание, м	Производительность труда 1 рабочего, т	Себестоимость, руб./т	Фондоемкость, руб./т	Объем добычи, тыс. т
	общая фондовооруж. (на 1 рабочего), руб.	техн. фондовооруж. (на 1 рабочего), руб.	уд. вес механизации подготовит. работ, %			уд. вес механизации очистных работ, %								
			всего	в том числе комбайновой	комплексной	комбайновой	взрывной и отбойным молотком	удельный вес конвейеризации, %						
Первый класс														
1963	6,3	1,8	43,8	—	—	42,9	16,1	35,8	382	31,5	57,6	7,84	11,2	879,4
1966	7,7	2,2	41,6	10,0	10,0	44,3	10,0	14,3	411,9	36,8	56,4	8,88	14,2	930,7
1968	8,0	2,6	48,0	18,7	27,9	19,4	24,6	24,7	424,3	35,6	51,5	10,1	14,2	1104,2
1970	10,2	3,6	57,9	26,7	50,1	15,5	27,9	19,0		41,5	63,8	9,2	13,0	1244,5
1971	11,4	4,2	73,1	36,4	55,7	3,2	19,6	37,6	583,8	44,3	74,8	8,3	12,9	1347,8
1972	12,5	4,7	75,9	47,3	64,7	2,7	12,2	43,5	609,2	46,0	76,4	8,7	13,9	1324,4
Второй класс														
1963	7,8	2,1	15,5	—	—	—	82,8	3,9	271,1	38,1	55,0	8,20	13,4	928,9
1966	7,8	1,8	23,1	10,0	—	—	89,1	9,7	243,2	37,0	49,7	9,05	14,6	994,9
1968	8,7	1,9	23,5	0,8	—	—	86,3	11,2	269,1	39,4	50,1	9,94	15,3	1055,9
1970	10,2	2,6	26,6	5,7	—	—	92,6	17,7		43,3	52,3	10,26	15,9	936,4
1971	10,4	2,5	30,5	2,9	—	—	89,0	15,4	226,2	46,7	56,2	9,77	16,2	1010,2
1972	13,2	3,5	37,9	3,3	—	—	77,4	10,5	241,0	23,0	85,7	10,75	20,0	982,5

почти в 1,8 раза (в частности, комбайновой механизации — в 4 раза), а уровень механизации основного звена производственного процесса — очистной добычи угля — самый высокий в бассейне. Удельный вес добычи угля механизированным способом, в частности с помощью наиболее прогрессивной комплексной механизации, увеличился с 1966 г. в 6,5 раза. Уровень конвейеризации был выше, чем в остальных классах, за все годы рассматриваемого периода, что в значительной степени объясняется характером природных условий.

В результате сравнительно высокого уровня механизации основных звеньев производственного процесса увеличилась общая и техническая фондовооруженность. Причем по сравнению с другими классами она росла более высокими темпами: за рассматриваемый период техническая фондовооруженность выросла более чем в 2,5 раза, тогда как, например, на шахтах второго класса она практически не менялась до 1972 г.[2]

В угольной промышленности экономический эффект от внедрения средств механизации реализуется через такие важные технические показатели, как нагрузка на очистной забой и среднемесячное подвигание очистной линии забоев.

За рассматриваемый период нагрузка на очистной забой возросла почти в 2 раза, а среднемесячное подвигание — в 1,5 раза. Тем не менее достигнутый уровень и динамику этих показателей нельзя считать удовлетворительными, так как их изменение не сопровождалось соответствующим улучшением экономических показателей, а в ряде случаев последние даже ухудшались. Это обусловлено, во-первых, тем, что недостаточно полно и эффективно использовались возможности высокопроизводительной и в то же время дорогостоящей техники [3]. Кроме того,

[2] Следует отметить, что рост общей фондовооруженности является результатом увеличения как технической фондовооруженности, так и пассивной части основных фондов в связи со «старением» шахты, ухудшением природных условий, ростом удельного веса выработок, закрепленных металлической и железобетонной крепью, и т. д.

[3] Одной из причин неудовлетворительных темпов роста нагрузки на очистной забой являются длительные сроки освоения средств механизации, что в определенной мере было неизбежно, так как технические средства внедрялись на процессах, выполнявшихся до этого либо вручную, либо примитивными средствами.

имело место существенное ухудшение природных условий.

Второй класс характеризуется следующим комплексом природных условий: крутым падением пластов (удельный вес добычи угля с крутопадающих пластов составил 83—93% за рассматриваемый период), мощными (1,8—3,5 м) и сверхмощными (свыше 3,5 м) пластами (удельный вес добычи угля с этих пластов был выше 80%), газообильностью пород по 10—20 м³/т, глубиной разработки 160—220 м, водообильностью 200—250 м³/ч и т. д.

Этот комплекс природных условий относится к числу более сложных и менее эффективных для эксплуатации, чем условия первого класса, что, прежде всего, сказывается на возможности механизации основных звеньев производственного процесса. Уровень механизации подготовительных, очистных работ и подземного транспорта значительно ниже, чем в первом классе. Но, несмотря на сравнительно невысокий уровень механизации, определенный эффект при эксплуатации шахт данного типа природных условий достигнут за счет внедрения и усовершенствования специальных систем разработки (щитовой и комбинированной с гибким перекрытием), учитывающих особенности и сложность этого типа природных условий.

В рассматриваемом периоде имели также место определенные изменения технического уровня, связанные с применением прогрессивного способа крепления (изменилась структура крепления: возрос удельный вес металлической и железобетонной крепей), с улучшением техники безопасности, усовершенствованием технологической схемы и т. д., что привело к росту общей фондовооруженности при неизменной технической фондовооруженности. На шахтах данного класса нагрузка на очистной забой практически оставалась на одном и том же уровне. В результате такие экономические показатели, как себестоимость и фондоемкость, росли, а производительность оставалась на одном и том же уровне до 1972 г.

Третий класс шахт, менее многочисленный по количеству объектов, характеризуется следующими природными условиями: разработкой пластов с наклонным падением (средневзвешенный угол 50°), преобладающей мощностью пластов 1,8—3,5 м, небольшой газообильностью 5,6 м³/т), сравнительно небольшой глубиной разработки 185 м), водообильностью порядка 330 м³/ч и т. д.

Этот комплекс природных условий наименее благоприятен для эксплуатации. Применительно к нему нет удовлетворительных технических и технологических решений, что обусловливает низкий технический уровень шахт. Применяемые в этих условиях системы разработки — длинными столбами и реже комбинированная с гибким перекрытием — не дают должного экономического эффекта. Использование высокопроизводительной техники здесь, как правило, неэффективно, а в ряде случаев и невозможно. Поэтому удельный вес комбайновой выемки составил лишь 10%, остальное приходилось на выемку взрывную и отбойными молотками. В результате имели место низкая нагрузка на очистной забой (156 т/сут) и малое среднемесячное подвигание очистной линии забоев (32,2 м).

Естественно, перечисленные условия производства (природные и технико-технологические) не могли не отразиться на экономических показателях: уровень экономических показателей на шахтах данного класса — один из самых низких в бассейне: себестоимость — 12,78 руб/т, производительность — 47,3 т на 1 раб., фондоемкость — 19,80 руб/т, средний объем добычи — 734,9 тыс. т в год.

Итак, рассмотренные классы шахт с различными типами природных условий существенно отличаются друг от друга по условиям эксплуатации. Каждый тип природных условий обладает определенной спецификой, которая накладывает отпечаток на все стороны производственного процесса. В частности, обнаружены следующие зависимости: каждому типу природных условий соответствует определенный тип технологических схем, учитывающий особенности и сложность производственного процесса в данных природных условиях; от типа природных условий зависят возможность и эффективность применения высокопроизводительных средств механизации; изменение технических уровней во времени носит различный характер для разных классов природных условий (в одних имело место принципиальное изменение технических уровней — переход от ручной выемки угля к механизированной, в других — лишь некоторые изменения технических уровней, не связанные с основными звеньями производственного процесса); уровень и характер изменения экономических показателей также различен для каждого из выделенных типов природных условий.

Специфика добычи угля в определенных условиях нашла свое отражение в системах их экономических оценок. Так, по трем из принятых за основу показателям эти оценки для шахт первого и второго классов были следующие:

	Первый класс	Второй класс
Себестоимость . .	8 руб. 90 коп.	10 руб. 72 коп.
Фондоемкость . .	13 руб. 96 коп.	18 руб. 80 коп.
Объем добычи . .	1323,6 тыс. т	988,3 тыс. т

СИСТЕМА ИНДИВИДУАЛЬНЫХ ЭКОНОМИЧЕСКИХ ОЦЕНОК ПРИРОДНЫХ УСЛОВИЙ

Процедуру разработки и анализа индивидуальных оценок природных условий рассмотрим на материалах первых двух наиболее многочисленных по составу классов.

В результате многомерной классификации по природным условиям значительно уменьшилась внутриклассовая относительная вариация природных факторов, а также характеристик технического уровня, в наибольшей степени зависящих от типа природных условий. В частности, уменьшение внутриклассовой вариации наблюдается для основных характеристик производственного процесса — системы разработки и способа механизации очистной добычи угля. Что касается уровня механизации подготовительных работ и подземного транспорта, то в первом классе относительная вариация этой характеристики существенно снизилась, а во втором — осталась по-прежнему значительной.

Несмотря на близость природных условий внутри классов, технический уровень производства во многом определяет уровень экономических показателей, а следовательно, и их вариацию. В связи с этим внутриклассовая вариация экономических показателей по сравнению с общей исходной совокупностью в первом классе уменьшилась, а во втором — оставалась примерно на том же уровне.

Для анализа связей переменных были рассчитаны матрицы парных коэффициентов корреляции для всей исходной совокупности и для двух рассматриваемых клас-

сов. Их сравнительный анализ показал, что характер внутриклассовых связей переменных претерпевает существенные изменения по сравнению с исходной совокупностью [4]. Изменяются направление и степень влияния природных условий на технический уровень производства, характер связи факторов технического уровня между собой, влияние природных и технических факторов на уровень экономических показателей.

Вместе с тем существенно различается характер связей при сравнении двух рассматриваемых классов. Так, если в первом классе влияние природных условий охватывает практически все характеристики технического уровня, то во втором степень тесноты связи факторов природных условий и технического уровня значительно меньше. При этом особенно существенно различается внутриклассовый характер влияния мощности и угла падения пласта. Так, связь этих факторов с общей и технической фондовооруженностью, нагрузкой на очистной забой, среднемесячным подвиганием очистной линии забоев имеет противоположное направление, а влияние, например, угла падения пласта на уровень комбайновой механизации добычи угля во втором классе вообще отсутствует, а в первом оно довольно существенно.

Связь природных факторов и экономических показателей также различается по классам (табл. 2). Направление влияния угла падения в данных классах противоположное: для шахт, разрабатывающих крутопадающие пласты (второй класс), увеличение угла падения ведет к улучшению экономических показателей, а для шахт, разрабатывающих пологие пласты (первый класс),— к ухудшению.

Статистический анализ системы индивидуальных экономических оценок включал в себя сопоставление: оценок индивидуальных объектов, полученных двумя различными способами (табл. 3);

оценок по разным показателям;

оценок с фактическими значениями экономических показателей.

---

[4] Здесь следует отметить, что в силу неоднородности исходной совокупности общие коэффициенты парной корреляции нередко не поддаются содержательной интерпретации.

Таблица 2

Характеристика относительной вариации производственных факторов и экономических показателей

Факторы и экономические показатели	Общий массив		Первый класс		Второй класс	
	средние значения	коэффициент вариации	средние значения	коэффициент вариации	средние значения	коэффициент вариации
Природные условия						
Газоносность	21,3	2,54	8,4	1,34	21,7	1,35
Максимальная глубина разработки, м	227,3	0,40	222,9	0,24	229,7	0,31
Водообильность, м³/ч	329,0	0,98	308,3	0,75	227,1	0,54
Удельный вес добычи, %						
С пластов с углом пад., град						
0—15	42,1	1,06	85,9	0,31	—	—
15—35	16,9	1,55	—	—	—	—
свыше 35	40,8	1,08	—	—	90,0	0,16
С пластов мощностью, м:						
1,2—1,8	16,4	1,17	13,0	0,94	—	—
1,8—3,5	41,1	0,75	70,4	0,34	211,	0,21
свыше 3,5	35,8	1,03	—	—	723,	0,22
Технический уровень						
Нагрузка на очистной забой, т/сут	408,2	0,60	609,2	0,43	241,0	0,30
Удельный вес механизации подготовительных работ (проходка), %	55,9	0,50	75,9	0,23	37,9	0,61
В том числе комбайнами	21,1	1,36	47,3	0,62	—	—
Уровень конвейеризации, %	27,0	0,90	43,5	0,52	10,5	1,55
Удельный вес механизации очистных работ, %						
В том числе добычи комплексами	0,5	1,33	64,7	0,59	—	—
взрывной выемки	44,3	1,01	—	—	77,4	0,46
Среднемесячное подвигание очистной линии забоев, м	44,1	0,50	46,0	0,44	52,1	0,50
Экономические показатели						
Производительность труда, т на 1 раб.	74,1	1,22	7,6	0,35	85,7	1,79
Себестоимость, руб/т	9,89	0,28	8,66	0,23	10,75	0,27
Объем добычи, тыс. т	1173,0	0,54	1324,4	0,37	982,5	0,47

Таблица 3

Система экономических оценок природных условий конкретных объектов, полученных способами 1 и 2

Класс	Номер объекта	Производительность труда			Себестоимость продукции			Фондоемкость продукции			Объем добычи		
		фактическое значение	абс.	отн.	фактическое значение	абс.	отн.	фактическое значение	абс.	отн.	фактическое значение	абс.	отн.
						Способ 1							
	1	39,4	24,9	1	14,72	12,76	1,3	27,0	21,1	1,5	1389,1	1164,1	1,1
Первый	2	48,4	49,5	2	10,93	10,36	1	19,5	18,4	1,3	742,8	1046,8	1
	3	51,8	75,4	3	10,76	10,07	1	23,4	17.7	1,3	1077,6	1081,9	1,03
	4	69,1	34,0	1,4	8,35	10,40	1	19,2	19,3	1,6	1275,3	1289,8	1,2
	5	59,5	38,0	1,5	9,40	10,01	1	14,6	17,6	1	1060,4	1070,5	1,02
	6	64,4	72,0	1,05	10,5	8,20	1,23	13,4	7,5	1	1054,8	1725,0	1
Второй	7	99,8	79,6	1,16	6,28	6,68	1	8,90	17,4	2,32	1472,2	1723,8	1
	8	60,6	68,5	1	9,81	10,15	1,52	19,6	17,4	2,32	1746,0	1715,5	1
	9	114,1	88,0	1,3	5,81	8,34	1,25	7,3	11,6	1,55	1881,9	1735,0	1
	10	78,7	77,7	1,13	7,50	8,70	1,30	11,8	11,9	1,58	1332,4	1734,4	1
						Способ 2							
	1	—	—	—	14,72	12,17	1,43	27,0	20,7	1,41	1389,1	846,0	1,0
Первый	2	—	—	—	10,93	8,51	1	19,5	17,8	1,21	742,8	1063,7	1,26
	3	—	—	—	10,76	10,58	1,24	23,4	22,5	1,53	1077,6	1060,9	1,25
	4	—	—	—	8,35	9,27	1,09	19,2	19,1	1,31	1275,3	1041,0	1,23
	5	—	—	—	9,40	11,23	1,32	14,6	14,7	1	1060,4	936,8	1,11
	6	64,4	88,1	1,32	10,50	9,98	1,62		13.2	1,02	1054,8	1340,3	1,18
Второй	7	99,8	93,7	1,43	6,28	6,15	1		12,9	1	1472,2	1352,6	1,18
	8	60,6	68,3	1,03	9,81	9,52	1,54		17,0	1,32	1746,0	1349,6	1,18
	9	114,1	94,7	1,44	5,81	8,83	1,43		13,7	1,06	1881,9	1355,0	1,18
	10	78,7	66,4	1	7,50	8,48	1,38		14,0	1,09	1332,4	1134,9	1,0

Примечание. В таблице приведены экономические оценки только некоторых объектов рассматриваемых классов.

Полученные двумя различными способами оценки в определенной степени близки: как правило, в качестве базисной (объекта с худшими природными условиями) выступает одна и та же шахта, и, кроме того, относительные величины оценок отклоняются друг от друга незначительно (см. табл. 3).

Соотношения оценок по различным экономическим показателям не идентичны. Это обусловлено тем объективным фактом, что каждый экономический показатель отражает определенную сторону производственного процесса, а следовательно, и влияние природных условий по-разному отражается на уровне различных экономических показателей.

Сравнительный анализ вариации фактических значений экономических показателей и системы оценок показал, что вариация первых более значительна внутри полученных классов. Некоторые различия в природных условиях отдельных шахт нашли свое отражение в вариации экономических оценок внутри классов.

При сопоставлении индивидуальных оценок природных условий и фактических значений экономических показателей возможны следующие ситуации:

1) система оценок лучше фактических экономических показателей по всем показателям;

2) система оценок хуже фактических экономических показателей по всем показателям;

3) система оценок по всем показателям близка к уровню их фактических значений;

4) различные комбинации соотношений: по одним показателям система оценок лучше, по другим — хуже, по третьим — близка к их фактическим значениям и т. д.

Первая ситуация означает, что возможна более эффективная эксплуатация шахт (№ 3, 6 и др.). Во втором случае технический и экономико-организационный уровни выше средних по классу, в результате чего достигнуты лучшие результаты (№ 5, 7, 10 и др.). В третьем случае близость оценок и фактических экономических показателей обусловлена наличием «среднего» технического уровня на данной шахте (№ 2, 8, 10 и др.). И наконец, четвертая ситуация означает, что с позиции одних показателей эксплуатация данного комплекса природных условий эффективна, а с позиции других — нет (№ 1 и др.).

ЛИТЕРАТУРА

1. Распознавание образов в экономико-статистическом моделировании. Новосибирск, 1974. 228 с.
2. Беккер А. В., Ягольницер М. А. Алгоритм выделения однородных совокупностей объектов.— В кн.: Распознавание образов и регрессионный анализ в экономических исследованиях. Новосибирск, 1972, с. 95—105.
3. Айзерман М. А., Браверман Э. М., Розоноэр Л. И. Метод потенциальных функций в теории обучения машин. М., «Наука», 1970. 384 с.

Г. Т. БЕРСЕНЕВА, Г. П. БОНДАРЕНКО,
Н. В. ОВЧИННИКОВА

УЧЕТ ВЛИЯНИЯ ПРИРОДНЫХ ФАКТОРОВ НА ЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ ПРЕДПРИЯТИЙ

Для удовлетворения растущих потребностей народного хозяйства существенно увеличивается использование природных ресурсов, многие из которых являются невоспроизводимыми. Отсюда возникает необходимость в повышении отдачи эксплуатируемых и рациональном освоении новых источников. В этой связи вопросам экономической оценки природных ресурсов в настоящее время уделяется все большее внимание, разрабатываются научно обоснованные методы такой оценки. Ее основное содержание сводится к учету влияния закономерных территориальных различий в природных свойствах ресурсов и их источников на производительность общественного труда [1]. Это значит, что использование одноименных ресурсов в силу объективно существующих различий в их качестве обеспечивает получение равного экономического эффекта только при дифференциации затрат. Вопросы экономической оценки природных ресурсов разрабатываются в ряде работ [1, 2, 5, 6 и др.]. Однако современное состояние проблемы требует решения широкого круга конкретных методологических и практических вопросов. К ним относятся:

определение частных (поэлементных) оценок отдельных природных факторов;

количественное выражение зависимости экономической эффективности производства от отдельных (элементарных) природных факторов с использованием массового материала и математико-статистических методов;

методика перехода от частных оценок к интегральным и способы их выражения в сопоставимом виде;

решение вопроса определения коэффициентов соизмерения (значимости) отдельных частных показателей или отдельно оцениваемых элементов комплекса природных условий. Введение этих коэффициентов связано с ранжированием компонентов и показателей, с приданием каждому члену ранжированного ряда определенного веса;

выявление закономерностей и формулировка правил перехода от количественных и качественных натуральных показателей в характеристике свойств природных комплексов к оценкам этих свойств, выраженным в показателях шкалы оценок.

Некоторые из этих вопросов рассматриваются в данной статье на конкретных примерах. При оценке различные природные условия сравниваются с точки зрения их влияния на результативные экономические показатели производственного процесса.

Следует заметить, что влияние природных условий на выходные экономические показатели не является абсолютным, а носит относительный характер, так как в значительной степени зависит от способов использования самих природных факторов. С развитием технического прогресса совершенствуются старые и появляются принципиально новые технико-технологические способы, а в связи с этим изменяется и характер влияния природных условий на уровень экономических показателей.

Экономическую оценку влияния природных условий на результаты производственного процесса можно проводить, во-первых, для отдельных предприятий, групп предприятий, месторождений и т. д.; во-вторых, для отдельных компонент природных условий или же их комплекса в целом.

Нами рассматриваются вопросы построения системы экономических оценок влияния природных условий на экономические показатели отдельных предприятий. Имеется в виду система сравнительных величин (абсолютных или относительных), позволяющих сопоставлять между собой природные условия отдельных предприятий с точки зрения экономической эффективности их эксплуатации.

Рассматриваемая задача носит статистический характер. Действительно, предприятие, бассейн и т. д. можно представить как статистическую совокупность в определенной степени однородных объектов, экономические по

казатели производства которых варьируют в пространстве (с переходом от одного объекта к другому) и во времени (одно предприятие за ряд лет). Варьирование экономических показателей обусловлено различием в уровне соответствующих факторов (природных и неприродных). Количественное влияние отдельных факторов или их совокупности (в данном случае комплекса природных условий) можно учесть статистически.

При оценке природных факторов экономико-статистическими методами возникает ряд трудностей. Во-первых, природные условия — комплексное понятие, представляющее собой совокупность характеристик, для оценки которых необходимо иметь систему измерителей. Во-вторых, совокупность характеристик природных условий включает в себя не только количественные, но и качественные признаки. В-третьих, все эти признаки неравноценны между собой с точки зрения влияния их на уровень экономических показателей, т. е. имеют различный вес. В-четвертых, природные условия, оказывая влияние на результаты производственного процесса, одновременно выступают условиями, определяющими протекание самого процесса [3]. Эти особенности обусловливают специфичность учета влияния природных условий на экономические показатели и задачи соизмерения природных условий отдельных предприятий.

Процесс решения поставленной задачи сводится к следующему: на основе качественного и количественного статистического анализа выделяется набор существенных факторов (как природных, так и неприродных) и определяются закономерности и сила их воздействия на уровень экономических показателей. С использованием полученных результатов находятся объективно возможные значения экономических показателей путем элиминирования различий в неприродных факторах (в том числе и технико-технологической структуре), которые и кладутся в основу системы экономических оценок.

Элиминирование влияния отдельных факторов или их комплекса на уровень экономических показателей возможно двумя способами. Первый заключается в следующем. Определяем статистическую зависимость экономических показателей производства от всей совокупности факторов производства, среди которых необходимо выделить и оценить влияние комплекса природных условий.

Фиксируя значения одних факторов на среднем уровне и варьируя значениями других, получаем чистое уравнение регрессии, характеризующее зависимость выходного показателя только от комплекса характеристик природных условий. Подставляя индивидуальные значения характеристик природных условий, получаем расчетные значения экономических показателей, которые и будут основой построения системы экономических оценок природных условий.

Уравнение регрессии строится для всей совокупности факторов, а не для одного лишь комплекса природных условий, потому что фактические показатели отражают взаимодействие и влияние природных и неприродных факторов. Если уравнение построить только для природных условий, это не даст нужных результатов, так как коэффициенты регрессии будут искажены неявным влиянием неприродных факторов и, следовательно, расчет по ним значений экономических показателей утратит свою достоверность.

Элиминирование влияния отдельных факторов или их комплекса возможно и методом классификации. С его помощью осуществляется подбор однородных с точки зрения отдельных факторов или их комплекса совокупностей предприятий. В каждой из них влияние тех факторов, по которым была проведена классификация, устраняется. Различия экономических показателей внутри классов будут обусловлены вариацией оставшихся факторов, не учтенных при классификации.

В работе рассматривается первый подход к решению задачи и используются данные о природных и технико-производственных условиях добычи железной руды открытым способом по горно-обогатительным комбинатам (ГОКам) и рудоуправлениям (РУ) страны. В качестве моделируемого показателя принята себестоимость добычи руды. Для нее можно принять гипотезу об обобщенно-нормальном распределении с поправкой на эксцесс и асимметрию по критерию Пирсона. Аналогичные результаты дает проверка этой гипотезы по дифференциальному критерию (доверительный уровень $\beta = 0{,}01$).

Построению регрессионных уравнений предшествовал анализ взаимосвязей факторов по матрице парных коэффициентов корреляции. Сначала выявлялись взаимосвязанные переменные, так как совмест-

Таблица 1

Коэффициенты уравнений множественной регрессии по себестоимости добычи сырой руды

Номер варианта	Свободный член регрессии (a_0)	Объем производства	Текущий коэффициент вскрыши		Содержание в сырой руде		Коэффициент использования		Годовая производительность буровых станков	Коэффициент использования		Годовая производительность экскаваторов по карьеру	Состав основных фондов, %
			т/т	м³/т	% загрязненности	влаги, %	парка буровых станков	календарного времени		парка экскаваторов	календарного времени		
Стандартизованный масштаб													
I		−0,214	0,697	0,853		−0,336	−0,322	−0,018	−0,012	0,150	−0,089	0,017	−0,155
II		−0,276				−0,293	−0,330	−0,053	0,049	0,099	−0,115	0,062	−0,020
III		−0,238		0,834		−0,302	−0,319			0,135			−0,185
Натуральный масштаб													
I	2,52	0,00002	0,3856	0,889		−0,065	−1,712	−0,207	−0,000040	0,960	−0,659	0,0004	−0,095
II	2,65	0,00002				−0,056	−0,018	−0,600	0,0002	0,6325	−0,839	0,0016	−0,0012
III	2,26	0,00002		0,8908		−0,058	−1,698			0,862			−0,011

ное использование двух или более таких переменных искажает результаты дальнейшего исследования.

К числу взаимосвязанных мы относили переменные с высоким значением парного коэффициента корреляции (более 0,8). В таких случаях в модель включался один из пары дублирующих переменных. Из анализа нашей матрицы выявлена сильная связь рудничной себестоимости 1 т сырой руды с производительностью на одного работающего ($r = -0,8459$), удельным весом пустой породы ($r = 0,8051$) и с текущим коэффициентом вскрыши ($r = 0,8271$); производительности на одного работающего по добыче сырой руды с удельным весом пустой породы ($r = -0,8087$) и текущим коэффициентом вскрыши ($r = -0,7964$); удельного веса пустой породы с текущим коэффициентом вскрыши ($r = 0,9033$).

Просчитывалось три варианта уравнений множественной регрессии (табл. 1). Варианты I и II включали: объем производства, текущий коэффициент вскрыши (м³/т), содержание в сырой руде влаги, коэффициент использования календарного времени отдельно буровыми станками и экскаваторами, годовую производительность буровых станков (тыс/м³), коэффициент использования экскаваторного парка, годовую производительность экскаватора (на 1 м³ емкости ковша), состав основных фондов по карьерам.

Анализ полученных стандартизованных коэффициентов полных уравнений регрессии (I и II варианты) позволил сделать вывод о том, что коэффициенты использования календарного времени буровыми станками и экскаваторами, а также показатели годовой производительности этих механизмов оказывают слабое влияние на себестоимость 1 т сырой руды и могут быть исключены из модели. В III варианте уравнения регрессии число факторов уменьшено почти в 2 раза. Из сравнения оценок вариантов уравнений множественной регрессии (табл. 2) видно, что коэффициенты детерминации (R^2) различаются незначительно.

Проверка значимости полных и конечных уравнений множественной регрессии по критерию Фишера дает положительные результаты: $F_{факт} \gg F_{табл}$. Коэффициенты детерминации ($R^2 = 0,86-0,88$) как полных, так и конечных уравнений свидетельствуют о высокой степени зависимости уровня затрат на добычу 1 т сырой руды от включенного в модели набора признаков. Себестоимость добычи руды зависит в основном от коэффициента вскрыши, содержания в руде влаги и от факторов, характеризующих использование основных производственных фондов.

Таблица 2

Оценки уравнений множественной регрессии по себестоимости добычи сырой руды

Номер варианта	Число факторов в модели	R^2	$\sigma^2_{ост}$	$F_{факт}$	$F_{табл}$ при $P_{ош}=5\%$
I	10	0,88	0,12	25,3	1,75 $F_{факт} > F_{табл}$
II	10	0,87	0,14	21,7	1,75 $F_{факт} > F_{табл}$
III	6	0,86	0,12	47,9	1,71 $F_{факт} > F_{табл}$

Таблица 3

Система суммируемых поправок для себестоимости

Фактор	Отклонения	
1. Объем производства, тыс. т.	—918,4	918,4
2. Текущий коэффициент, $м^3/т$	—0,0211	—0,0211
3. Содержание в руде влаги, %	—0,3223	0,3223
4. Коэффициент использования парка буровых станков	—0,011	0,011
5. Состав основных фондов, %	—1,6624	+1,6624
Поправки, %	1	—1

На основе полученных моделей рассчитаны оценки, представляющие собой систему суммируемых поправок (табл. 3), варьирующих под влиянием отклонений условий производства от базисных [4]. В качестве базисных взяты средние условия по исследуемой совокупности объектов. Система поправок оценивает относительное влияние каждого из факторов без учета конкретного производственного объекта. Например, увеличение содержания в руде влаги на 0,32% соответствует снижению себестоимости добычи на 1%. Зная условия производства конкретного предприятия и рассчитав их отклонения от средних, можно по системе поправок найти общую величину коэффициента, характеризующего суммарное влияние факторов. Последнее рассчитывается и непосредственно по уравнению регрессии [4].

Для некоторых ГОКов и РУ коэффициенты, характеризующие суммарное влияние факторов, составляют: Оленегорский — 0,90, Ковдорский — 0,85, Михайловский — 1,50, Тейское — 1,20, Азербайджанский — 0,95, Днепровский — 1,00, Гороблагодатское — 1,00. Отсюда себестоимость, например, Тейского рудоуправления по природным производственным условиям должна быть выше среднего по исследуемой совокупности значения на 20%.

ЛИТЕРАТУРА

1. **Минц А. А.** Экономическая оценка естественных ресурсов, М., 1972. 303 с.
2. **Лопатина Е. Б., Минц А. А., Мухина Л. И., Назаревский О. Р., Преображенский В. С.** Состояние и задачи разработки теории и методики оценки природных условий и ресурсов.— «Изв. АН СССР. Серия геогр.», 1970, № 4, с. 45—54.
3. **Розин Б. Б., Ачелашвили Н. Е.** Некоторые вопросы статистического моделирования экономических показателей горнодобывающей промышленности.— В кн.: Распознавание образов в экономико-статистическом моделировании. Новосибирск, 1974, с. 164—182.
4. **Розин Б. Б., Овчинникова Н. В., Лукацкая М. Л.** Построение системы коэффициентов-соизмерителей для предприятий лесозаготовительной промышленности.— В кн.: Статистическое моделирование и прогнозирование технико-экономических показателей предприятий. Новосибирск, «Наука», 1974, с. 29—76.
5. **Хачатуров Т. С., Лойтер М. Н.** Экономическая оценка природных ресурсов при проектировании и строительстве.— «Вестник АН СССР», 1973, № 3, с. 57—69.
6. **Фейтельман Н.** К проблеме экономической оценки минеральных ресурсов.— «Экономические науки», 1973, № 5, с. 50—56.

Л. А. СЕРГЕЕВА, Г. Л. ЧУДНОВСКИЙ,
М. А. ЯГОЛЬНИЦЕР

ИССЛЕДОВАНИЕ ДИНАМИКИ НЕФТЕДОБЫЧИ

При проектировании основного периода разработки нефтяных месторождений большое значение имеют методы и подходы, позволяющие планировать динамику отбора нефти и жидкости по отдельным залежам. В этом направлении в последние годы выполнено немало работ, появились обобщающие исследования по сравнению различных методик проектирования [3, 4, 7].

В то же время анализ проектных документов и фактических показателей по многим нефтяным месторождениям страны [4] показывает, что состояние проектных работ в данной отрасли остается неудовлетворительным. В частности, отмечается несовершенство методик расчета отдельных показателей, использование принципиально отличных друг от друга подходов к коренным вопросам рациональной разработки залежи (оптимальная плотность сетки скважин, обоснование коэффициентов текущей и конечной нефтеотдачи и т. д.). Все это создает большие трудности при оценке технологических и экономических показателей разработки, определенных различными методами. Нередко динамика отбора нефти и жидкости в реальных условиях значительно отличается от проектной. Поэтому предприятия при перспективном планировании добычи нефти по залежам нередко используют опыт разработки подобных залежей, находящихся в более поздней стадии эксплуатации.

В последние годы во Всесоюзном нефтегазовом научно-исследовательском институте была разработана методика по определению технологических экономических показателей разработки и коэффициента нефтеотдачи с учетом неоднородности пласта при водонапорном режиме (гидродинамический расчет) [3]. Однако при решении задач долгосрочного планирования развития и размещения нефтедобывающей промышленности, когда определяется стратегия отбора запасов и подготовки сырьевой базы на длительную перспективу, использование этой методики не представляется целесообразным. Это объясняется несоответствием степени достоверности исходной информации о геолого-промысловых особенностях открытых и тем более предполагаемых (неоткрытых) месторождений в предплановый период и сложностью моделирования гидродинамических процессов, протекающих в залежах.

Кроме того, для используемых при перспективном планировании многоуровневых схем [6] с их многовариантностью разработки нефтяных месторождений использование упомянутой методики и основанной на ней программы для ЭВМ приводит либо к предварительному формированию технологических показателей по всевозможным вариациям управляющих параметров, либо к текущему формированию вариантов. В обоих случаях использование ресурсов ЭВМ (время работы процессора и обращения к долговременной памяти, а также размеры последней) будет неэффективным по отношению к результатам исходных задач. Отметим еще, что сама программа гидродинамического расчета основана на некоторых упрощающих предположениях и, следовательно, полученные по ней результаты не могут быть абсолютно адекватными реальному процессу.

Очевидна необходимость использования иных подходов к расчету технологических показателей нефтедобычи. И наиболее обнадеживающими представляются здесь экономические методы анализа.

Цель данного исследования как раз и состоит в построении экономико-статистических моделей влияния комплекса природных, управляющих и временны́х факторов на динамику нефтедобычи по месторождениям. Возможны два методических подхода к исследованию динамики нефтедобычи: с помощью имитационных моделей программы гидродинамического расчета и построенных на фактическом материале экономико-статистических моделей.

В соответствии с конкретными целями рассматриваемых подходов используются различные элементы и пара-

метры их описания. Для более узкой задачи имитационного моделирования в качестве исследуемой совокупности используются проектируемые нефтяные месторождения. Моделируются уровни нефтедобычи, коэффициенты изменения нефтедобычи по годам и динамика нефтедобычи с момента пуска объекта в эксплуатацию.

Для задачи экономико-статистического моделирования по фактическим данным исследуемая совокупность формируется из действующих нефтяных месторождений СССР. Исходные данные анализируются за 15 — 20-летний период.

Моделируются абсолютные уровни добычи по годам, изменение нефтедобычи по годам, коэффициент нефтеотдачи.

Специфика разработки нефтяных месторождений обусловлена геологическими условиями залегания и физическими свойствами нефти [1, 5, 9]. Основное влияние на технологические и экономические показатели разработки нефтяных залежей оказывают геолого-физические характеристики — горно-геологические параметры пласта и свойства флюидов. К первым относятся: размер залежи (площадь в га), глубина залегания продуктивного пласта, тип коллектора (терригенные, карбонатные пласты), геологический возраст пласта, количество пластов в залежи (одно- и многопластовые), степень расчлененности (структура) продуктивных горизонтов (пластов), коэффициент песчанистости ($K_п$), неоднородность, температура, проницаемость пласта, пластовое давление (начальное, текущее, забойное), мощности (эффективная, нефтенасыщенная, общая) залежи.

Вторую группу характеристик образуют физико-химические свойства нефти, которые определяют режим пласта: вязкость пластовой нефти и воды ($\mu_н$, $\mu_в$), их соотношение ($\mu_н/\mu_в$) — относительную вязкость пластовой нефти (μ_0), подвижность нефти в пластовых условиях $\left(k_н/\mu_н — \frac{\text{проницаемость}}{\text{вязкость нефти}}\right)$, процент сернистости, содержание парафина в нефти и т. д. Изучение материалов разработки нефтяных залежей и результатов предшествующих исследований дает основание считать определяющими геолого-физическими характеристиками нефтяных залежей вязкость пластовой нефти и воды, проницаемость и показатель неоднородности пласта. Эти показатели на момент ввода элемента в эксплуатацию и могут быть использованы для решения задачи имитационного моделирования.

В задаче многофакторного экономико-статистического исследования фактической динамики нефтедобычи круг горно-геологических факторов должен быть значительно расширен с учетом всех вышеизложенных характеристик.

Совокупность мероприятий, с помощью которых можно воздействовать на процесс эксплуатации залежи при данных геологических условиях, принято называть системой разработки залежи. Управляющими факторами (варьирующими независимыми переменными) при экономико-статистическом моделировании могут выступать плотность сетки скважин, фонд действующих скважин (эксплуатационных — $n_э$ и нагнетательных — $n_н$), интенсивность воздействия ($N = n_э/n_н$), режим залежи (упруговодонапорный, водонапорный), вид воздействия на залежь (заводнение, тепловые, физико-химические методы воздействия на пласт и др.), система размещения скважин (рядная, площадная, избирательная).

Плотность сетки эксплуатационных скважин в значительной степени зависит от особенностей геологического строения нефтяных пластов, степени их литологической изменчивости, проницаемости пород, физико-химических свойств нефти, дебитов скважин, системы воздействия на залежь. Вместе с тем экономика разработки залежи является определяющей для данного параметра [2, 10].

Залежи маловязких нефтей, приуроченные к однородным, хорошо проницаемым пластам, могут разрабатываться ограниченным числом скважин в зависимости от планируемого уровня добычи нефти и срока разработки залежи. С применением редкой сетки скважин могут разрабатываться также залежи, приуроченные к трещиноватым карбонатным толщам. Нефтяные залежи, отличающиеся малыми дебитами, связанные со слабопроницаемыми глинистыми коллекторами или с высоковязкой нефтью, в противоположность рассмотренным выше должны разрабатываться с применением плотных сеток скважин. В этом случае активные методы воздействия на залежи, например тепловые, не исключают необходимости бурения скважин по достаточно плотной сетке.

Применение заводнения с начала разработки позволило увеличить расстояние между скважинами, хотя опти-

мальная плотность сетки скважин продолжает оставаться одной из основных проблем разработки нефтяных месторождений, по которой среди ученых нет единого мнения [7]. В силу определяющего влияния на динамику нефтедобычи данный фактор используется для решения первой задачи в качестве управляющего. Кроме перечисленных варьирующих независимых переменных (плотность сетки скважин, вязкость нефти, проницаемость и неоднородность) в задаче учитывается еще и время с момента ввода объекта в эксплуатацию.

При втором подходе необходим более широкий набор управляющих переменных. В этом случае предлагается учет и «возраста» месторождения — темпа ввода нефтяной залежи в разработку, т. е. очередности ввода в эксплуатацию скважин внутри месторождения.

Коротко охарактеризуем возможные способы построения имитационных моделей.

I-А. Построение набора дискретно-непрерывных моделей [8] уровней нефтедобычи для каждого года исследуемого периода. Модели строятся для каждого года и каждого вида управления. Исходная информация — результаты, полученные по программе гидродинамического расчета в пределах диапазонов варьирования значений горно-геологических условий исследуемого нефтедобывающего района. Выборка должна быть репрезентативной для фактического распределения горно-геологических условий региона. Поэтому необходимо иметь распределение объектов по горно-геологическим условиям (по трем вышеуказанным факторам).

I-В. Построение классификационных моделей прогнозирования кривых нефтедобычи методами распознавания образов [8]. Здесь возможны два подхода.

1. Способом многомерной классификации элементов для каждого управления по горно-геологическим условиям (вход способа I-А) формируется обучающая выборка. Для каждого полученного класса выделяется наиболее типичный представитель (элемент), и классу приписывается кривая этого элемента (типичное сочетание геологических условий).

2. Для каждого нового объекта (он может формироваться методом Монте-Карло) в диапазоне вариации независимых переменных определяется вероятность его принадлежности к выделенным классам. Здесь объекту приписывается соответствующая (или усредненная) кривая (с учетом вероятности). Отметим, что этот подход позволяет прогнозировать не только динамику добычи нефти по разрабатываемым и открытым (вовлекаемым в разработку) месторождениям, но и основные геологопромысловые характеристики и состав неоткрытых месторождений. Последнее позволяет использовать их добычные возможности наравне с открытыми объектами при решении задач перспективного планирования.

Основная возникающая при этом проблема — формирование представительной выборки для условий исследуемого нефтедобывающего района по горно-геологическим факторам (как и при способе I-А).

Для статистического анализа при моделировании фактической работы объектов используются также два способа: II-А — исследование одномерных динамических рядов коэффициентов нефтедобычи (авторегрессия, тренды и т. д.) и II-В — построение дискретно-непрерывных многофакторных моделей [8] зависимости коэффициента изменения нефтедобычи от его прошлых значений, «возраста» объекта (если выполнять для определенного года, то будет чисто пространственная выборка), горно-геологических факторов, характеристик управления с лагами и т. д.

Следует отметить, что статистические методы анализа динамики еще довольно слабо разработаны, а известные из литературы подходы принципиально отличаются от предложенных в данной работе. Достоинство рассматриваемого подхода, на наш взгляд, заключается в том, что при сравнительно небольших материальных затратах (ресурсы машинного времени и т. д.) он позволяет просмотреть достаточно большое число вариантов и выбрать наиболее приемлемый исходя из поставленной цели исследования. Кроме того, эти модели могут быть использованы в сочетании с другими, например в оптимизационных отраслевых системах.

ЛИТЕРАТУРА

1. Аммкс Дж., Басс Д., Уайтинг Р. Физика нефтяного пласта. М., Гостоптехиздат, 1962. 572 с.
2. Иванова М. М. Некоторые итоги разработки залежей при вытеснении нефти водой.— «Нефтяное хозяйство», 1975, № 5, с. 32—38.

3. **Борисов Ю. П., Воинов В. В., Рябинина З. К.** Особенности проектирования разработки нефтяных месторождений с учетом их неоднородности. М., «Недра», 1976. 286 с.
4. **Мирчинк М. Ф., Желтов Ю. В.** О мероприятиях по улучшению проектных работ в добыче нефти.— «Нефтяное хозяйство», 1975, № 8, с. 26—30.
5. **Пирсон С. Д.** Учение о нефтяном пласте. М., Гостоптехиздат, 1961. 570 с.
6. **Планирование** отраслевых систем (Модели и методы оптимизации). М., «Экономика», 1974. 319 с.
7. **Совершенствование** систем разработки нефтяных месторождений Западной Сибири. Свердловск, Средне-Уральское кн. изд-во, 1975. 173 с.
8. **Статистические** модели в оптимальном отраслевом планировании. М., «Статистика», 1975. 189 с.
9. **Физические** свойства горных пород Западно-Сибирской нефтегазоносной провинции. М., «Недра», 1975. 182 с.
10. **Щелкачев В. Н.** Влияние на нефтеотдачу плотности сетки скважин и их размещения.— «Нефтяное хозяйство», 1974, № 6, с. 26—30.

Г. Т. БЕРСЕНЕВА, Т. Р. САЗОНОВА,
Л. А. СЕРГЕЕВА, М. А. ЯГОЛЬНИЦЕР

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ДИНАМИКИ НЕФТЕДОБЫЧИ

(на примере Западной Сибири)

В статье рассматривается один из способов моделирования динамики нефтедобычи, а именно построение имитационной модели программы гидродинамического расчета (способ I-А [1]). Задача исследования состояла в построении дискретно-непрерывных моделей [3] уровней нефтедобычи для каждого года исследуемого периода.

Развитие математических методов и появление современных быстродействующих ЭВМ способствовали возникновению нового направления в исследовании природы экономических и социальных явлений — имитационного моделирования. «Имитационная система — это совокупность модели, имитирующей изучаемое явление и закодированной в ЭВМ, системы внешнего математического обеспечения и системы внутреннего обеспечения» — такова наиболее удачная формулировка имитационной системы, предложенная Н. Н. Моисеевым в предисловии к работе [4, с. 6]. Имитационная система представляет собой аналог сложного реального процесса и позволяет проводить эксперимент не на самом процессе, а на его математической модели.

В работе [5] для исследования зоны неопределенности оптимального развития сложной системы предложен количественный метод, который заключается в имитации на ЭВМ множества возможных условий ее развития. Это происходит за счет изменения параметров модели (коэффициенты функционала, вектор ограничений, коэффициенты технологической матрицы). Различные сочета-

ния значений параметров имитируют колебания внешних условий, а полученные для каждого сочетания решения образуют зону неопределенности оптимального развития. Исследуя зону неопределенности, можно дать экономическую оценку выявленных вариантов для различных условий развития системы.

В. Е. Лихтенштейн [6] разработал способ имитации, известный под названием эволюционно-симультативного метода планирования. Суть его в том, что для задачи стохастического программирования формулируется такая модель, которая позволяет использовать симуляции для поиска решения.

Основной, как нам кажется, принципиальный недостаток теории оптимальных решений связан с тем, что она пока мало ориентируется на имитационные средства. Ее процедуры поиска решения основаны на построении модели взаимодействия различных факторов, влияющих на качество решения. Такая модель задается легко обозримыми соотношениями, которые можно «охватить» и «выбрать» нужные направления поиска.

В имитационных моделях нет налицо таких зависимостей, они скрыты от глаз наблюдателя, и можно анализировать только отдельные случайные проявления последствий тех или иных решений. Ситуация здесь во многом напоминает ту, с которой имеет дело математическая статистика, изучающая закономерности по их отдельным случайным проявлениям. В этом случае причинно-следственные отношения не могут быть отражены с помощью математических зависимостей, для их исследования целесообразны средства имитационного моделирования.

Имитация дает возможность изучать сложнейшие взаимодействия в рассматриваемой системе, и не случайно она взята на вооружение специалистами в области системного анализа. Уже сам процесс построения имитационной модели обеспечивает исследователю информацию о моделируемом явлении. «Знания, полученные во время разработки имитационной модели, часто становятся источником изменений в имитируемой системе. Влияние этих изменений можно проверить с помощью имитации еще до их практического внедрения» [4, с. 19].

В настоящей работе имитация рассматривается именно как опыт построения модели сложной программы гидродинамического расчета [2]. Проектирование систем разработки нефтяных месторождений — сложный процесс, требующий учета большого количества факторов. Поэтому в строгой постановке решение этой задачи невозможно, да и вряд ли целесообразно ввиду стохастической природы многих геологопромысловых характеристик. Следует отметить, что до сих пор нет даже унифицированных формулировок при характеристике различных показателей и понятий. С этих позиций программа гидродинамического расчета выступает в качестве имитатора реальных процессов, протекающих в залежи.

Однако любая подобная программа очень сложна, требует больших затрат машинного времени и вследствие этого непригодна для оперативных расчетов в текущем и перспективном планировании. Отсюда понятно наше стремление разработать более упрощенную модель динамики нефтедобычи, исходя из результатов анализа гидродинамического расчета. Именно в этом смысле понимается имитация сложной гидродинамической программы.

Объектом исследования выступали 30 проектируемых нефтяных месторождений Западной Сибири. Из их геолого-физических характеристик вначале были взяты для рассмотрения следующие наиболее существенные факторы (вариант 1): вязкость пластовой нефти и воды, проницаемость и показатель неоднородности пласта. Для случая водонапорного режима разработки залежи и трехрядной системы размещения скважин (условия, необходимые для гидродинамического расчета по методике [2]) плотность сетки скважин (управляющий параметр) принималась равной 64 га. Такие крупные сетки скважин наиболее характерны для условий Западной Сибири.

Строился набор дискретно-непрерывных моделей (ДНМ) зависимости годовой добычи нефти ($q_н$) от вышеперечисленных факторов за 15-летний период эксплуатации исследуемой группы месторождений. Отметим, что для плотности сетки в 64 га влияние показателя неоднородности пласта на объем годовой добычи (по результатам анализа парных корреляций) незначительно. Поэтому при построении дискретно-непрерывных моделей данный фактор не учитывался. Характеристики ДНМ за каждый год рассматриваемого периода приведены в табл. 1.

Из таблицы видно, что адекватность моделей с годами падает. Наилучшие прогнозные результаты по данным

Таблица 1

Набор дискретно-непрерывных моделей нефтедобычи (вариант 1)

Год	R^2	$\overline{S}^2$	F	a_0	x_1	x_2
1-й	0,71	41,08	33,17	15,462	—0,640 —3,904	0,727 68,513
2-й	0,82	8,84	60,75	9,977	—0,587 —2,096	0,846 46,657
3-й	0,79	4,62	51,68	7,474	—0,513 —1,242	0,864 32,277
4-й	0,79	3,29	51,32	5,966	—0,460 —0,937	0,882 27,730
5-й	0,77	2,24	44,99	4,972	—0,390 —0,623	0,886 21,170
6-й	0,68	1,84	28,57	4,530	—0,377 —0,462	0,831 15,715
7-й	0,67	1,26	27,20	3,920	—0,332 —0,331	0,832 12,819
8-й	0,56	1,25	16,87	3,488	—0,262 —0,225	0,765 10,128
9-й	0,44	0,93	10,86	3,364	—0,254 —0,168	0,683 6,990
10-й	0,30	1,07	5,83	2,986	—0,130 —0,083	0,566 5,533
11-й	0,28	0,76	5,32	2,603	—0,012 —0,006	0,534 4,340
12-й	0,23	0,69	3,95	2,380	0,065 0,031	0,456 3,393
13-й	0,21	0,59	3,56	2,174	0,159 0,070	0,391 2,673
14-й	0,21	0,51	3,53	1,987	0,246 0,101	0,328 2,078
15-й	0,21	0,45	3,51	1,844	0,319 0,123	0,255 1,510

Таблица 2

Оценка качества имитационной модели (вариант)

Номер объекта	Относительная ошибка расчета по годам														
	1-й	2-й	3-й	4-й	5-й	6-й	7-й	8-й	9-й	10-й	11-й	12-й	13-й	14-й	15-й
1	0,02	0,06	0,18	0,26	0,15	0,01	0,08	0,10	0,07	0,04	0,02	0,03	0,06	0,09	0,16
2	0,20	0,20	0,14	0,15	0,18	0,20	0,19	0,17	0,13	0,11	0,10	0,07	0,03	0,03	0,04
3	0,33	0,23	0,19	0,20	0,20	0,16	0,08	0,01	0,02	0,09	0,08	0,05	0,03	0,06	0,08
4	0,23	0,08	0,08	0,12	0,15	0,20	0,14	0,07	0,06	0,05	0,09	0,10	0,10	0,10	0,09
5	0,42	0,11	0,01	0,01	0,02	0,04	0,08	0,14	0,20	0,32	0,23	0,26	0,28	0,29	0,30
6	0,09	0,20	0,12	0,08	0,02	0,21	0,17	0,14	0,04	0,00	0,03	0,06	0,07	0,08	0,09
7	0,06	0,02	0,07	0,01	0,04	0,15	0,21	0,27	0,13	0,15	0,16	0,17	0,16	0,13	0,09
8	0,37	0,15	0,05	0,02	0,03	0,01	0,00	0,01	0,01	0,06	0,10	0,13	0,15	0,17	0,20
9	0,18	0,09	0,13	0,13	0,16	0,07	0,07	0,10	0,08	0,12	0,15	0,21	0,22	0,19	0,17
10	0,05	0,18	0,18	0,11	0,12	0,17	0,22	0,16	0,27	0,27	0,28	0,27	0,25	0,23	0,20
11	0,09	0,24	0,26	0,20	0,22	0,30	0,26	0,29	0,32	0,32	0,33	0,31	0,29	0,27	0,26
12	0,09	0,03	0,00	0,05	0,12	0,14	0,21	0,23	0,22	0,27	0,29	0,28	0,23	0,21	0,20
13	0,12	0,13	0,25	0,28	0,28	0,26	0,26	0,29	0,30	0,40	0,35	0,39	0,41	0,42	0,43
14	0,14	0,30	0,35	0,48	0,30	0,28	0,32	0,42	0,41	0,44	0,47	0,48	0,48	0,48	0,48
15	0,45	0,43	0,35	0,26	0,44	0,48	0,50	0,50	0,47	0,43	0,39	0,33	0,27	0,19	0,12

Таблица 3

Набор дискретно-непрерывных моделей нефтедобычи (вариант 2)

Год	R^2	$\bar{S}^2$	F	c_0	Коэффициенты регрессии для факторов (в стандартизованном и натуральном масштабе соответственно)						
					x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
1-й	0,8867	21,880	17,384	165,182	−0,6701 −2,6046	0,6835 64,396	−0,1702 −96,788	−0,0729 −13,351	0,2093 96,757	0,3363 0,2856	−0,2602 −25,120	−0,3002 −179,885	−0,0095 −2,420
2-й	0,9205	5,606	25,739	60,389	−0,6700 −1,5627	0,6802 38,741	−0,3027 −104,087	−0,2893 −32,046	0,0863 24,121	0,2701 0,1387	−0,1774 −10,354	−0,0762 −27,619	−0,0326 −5,035
3-й	0,9389	2,0139	34,126	31,166	−0,6197 −0,9878	0,6393 24,880	−0,3212 −75,474	−0,3780 −28,615	0,0762 14,553	0,2704 0,0949	−0,1823 −7,273	0,0157 3,895	−0,0757 −7,995
4-й	0,9463	1,186	39,123	22,149	−0,5709 −0,7449	0,6743 21,480	−0,2793 −53,712	−0,3772 −23,371	0,1149 17,959	0,2680 0,0770	−0,1855 −6,055	−0,0029 −0,5852	−0,0362 −3,1312
5-й	0,9246	1,142	27,231	20,189	−0,5201 −0,5621	0,5950 15,700	−0,3132 −49,894	−0,4275 −21,942	0,0899 11,644	0,2555 0,0608	−0,2431 −6,573	0,0432 7,248	−0,0945 −6,771
6-й	0,8792	1,2317	16,177	17,408	−0,5011 −0,4444	0,4616 9,995	−0,4541 −59,361	−0,5076 −21,381	0,0374 3,976	0,2217 0,0433	−0,2498 −5,544	0,0973 13,401	−0,0954 −5,610
7-й	0,8388	1,174	11,560	10,864	−0,4747 −0,3557	0,4145 7,584	−0,4830 −53,359	−0,5008 −17,824	0,0037 0,3355	0,1653 0,0273	−0,2180 −4,089	0,1910 22,236	−0,1638 −8,135
8-й	0,8075	1,081	9,322	1,8756	−0,4544 −0,2991	0,3149 5,061	−0,4706 −45,660	−0,5500 −17,190	−0,0287 −2,262	0,1519 0,022	−0,1645 −2,710	0,3111 31,806	−0,2094 −9,136
9-й	0,8226	0,6477	10,304	−0,79653	−0,4537 −0,2408	0,2083 2,699	−0,5377 −42,060	−0,6205 −15,638	−0,0736 −4,681	0,1405 0,0164	−0,1233 −1,637	0,3895 32,098	−0,2260 −7,950
10-й	0,8194	0,5581	10,085	−5,1853	−0,3888 −0,1898	0,0867 1,0334	−0,5017 −36,106	−0,6404 −14,849	−0,0337 −1,974	0,1333 0,0143	−0,0739 −0,9029	0,4681 35,498	−0,2768 −8,957
11-й	0,8218	0,3780	10,247	−6,135	−0,3036 −0,1228	0,0441 0,4356	−0,5181 −30,889	−0,6808 −13,076	−0,0677 −3,279	0,1631 0,0145	−0,0560 −0,5668	0,5083 31,932	−0,2670 −7,158
12-й	0,8224	0,3066	10,292	−9,372	−0,2658 −0,0970	−0,0596 −0,5312	−0,5050 −27,162	−0,7078 −12,267	−0,0115 −0,5009	0,1479 0,0119	0,0147 0,1343	0,5548 31,447	−0,2312 −5,592
13-й	0,8171	0,2590	9,929	−10,080	−0,1844 −0,0610	−0,1224 −0,988	−0,4873 −23,742	−0,7192 −11,289	−57,053 −0,2260	0,1283 0,0093	0,0372 0,3079	0,5837 29,965	−0,2222 −4,868
14-й	0,8055	0,2306	9,204	−10,632	−0,1026 −0,0310	−0,1789 −1,321	−0,4691 −20,909	−0,7158 −10,278	0,0016 0,0579	0,1050 0,0070	0,0568 0,4299	0,6060 28,461	−0,2070 −4,148
15-й	0,7986	0,2038	8,8122	−10,870	−0,0316 −0,0088	−0,2438 −1,663	−0,4558 −18,771	−0,7161 −9,500	0,0155 0,5182	0,0907 0,0056	0,0735 0,5137	0,6147 26,675	−0,1801 −3,335

моделям получаются в первые 5—7 лет. Влияние факторов, включенных в ДНМ, также уменьшается.

Затем для случайно выбранных 15 объектов рассматриваемой совокупности сопоставлялись результаты, полученные непосредственно по программе гидродинамического расчета [2] и по системе ДНМ.

Относительная ошибка рассчитывалась по формуле

$$(q_{\text{н}}^{0} - q_{\text{н}}^{\text{р}})/q_{\text{н}}^{0},$$

где $q_{\text{н}}^{0}$ — объем годовой добычи нефти по программе гидродинамического расчета; $q_{\text{н}}^{\text{р}}$ — то же, по ДНМ.

Результаты сопоставления даны в табл. 2. Расхождения расчетов по ДНМ и по методике [2] составляют в среднем 18%, т. е. сравнительно большие. На наш взгляд, это можно объяснить тем, что из 12 варьирующих факторов, включенных в гидродинамический расчет, в ДНМ входили только два — относительная вязкость пластовой нефти (x_1) и проницаемость пласта (x_2).

Построенный имитатор отличается простотой (включает два фактора) и может быть легко номографирован. Но даже такой простейший имитатор дает неплохие результаты. Точность может оказаться вполне достаточной, если учесть погрешность исходной информации, ограничения программы гидродинамического расчета, прогрессирующую технологию добычи нефти и т. п. При необходимости точность имитатора можно повысить, увеличив число переменных и введя условия нелинейности.

Для проверки этого утверждения был выполнен 2-й вариант расчетов по ДНМ, в котором число геолого-физических характеристик рассматриваемых месторождений увеличено с 2 до 9. В модели вводились следующие факторы:

x_1 — относительная вязкость пластовой нефти, б/р;
x_2 — проницаемость, дарси;
x_3 — коэффициент остаточной нефтенасыщенности;
x_4 — коэффициент насыщенности связанной водой;
x_5 — пористость, б/р;
x_6 — перепад давления, атм;
x_7 — показатель неоднородности по проницаемости, б/р;
x_8 — удельный вес нефти на поверхности, г/см³;
x_9 — удельный вес воды на поверхности, г/см³.

Таблица 4

Оценка качества имитационной модели (вариант 2)

Номер объекта	Относительная ошибка расчета по годам														
	1-й	2-й	3-й	4-й	5-й	6-й	7-й	8-й	9-й	10-й	11-й	12-й	13-й	14-й	15-й
1	0,0559	0,0598	0,0368	0,0145	0,0599	0,0990	0,0995	0,1703	0,1516	0,1873	0,1994	0,1839	0,1753	0,1476	0,1077
2	0,0473	0,0454	0,0262	0,0123	0,0801	0,1075	0,1325	0,1397	0,1192	0,1051	0,1125	0,0714	0,0458	0,0463	0,0629
3	0,1541	0,0778	0,0604	0,0453	0,0653	0,0282	0,0020	0,0475	0,0802	0,1299	0,1192	0,0984	0,0746	0,0967	0,1160
4	0,0080	0,0295	0,0530	0,0615	0,0768	0,1155	0,1380	0,1791	0,0384	0,1005	0,0664	0,0592	0,0549	0,0487	0,0407
5	0,0247	0,0121	0,0455	0,0479	0,0225	0,0024	0,0450	0,0385	0,0116	0,1097	0,0171	0,0313	0,0469	0,0363	0,0664
6	0,1369	0,1342	0,0975	0,0665	0,0418	0,1340	0,1171	0,1298	0,0481	0,0558	0,0091	0,0093	0,0110	0,0134	0,0209
7	0,0748	0,0896	0,0762	0,0075	0,0346	0,1916	0,2242	0,2442	0,0851	0,0393	0,0614	0,0519	0,0322	0,0071	0,0529
8	0,2272	0,2655	0,1776	0,1329	0,0798	0,1270	0,1162	0,1545	0,1490	0,1179	0,0756	0,0603	0,0386	0,0154	0,0138
9	0,1395	0,0637	0,0735	0,0809	0,1202	0,0567	0,0709	0,0717	0,0630	0,0430	0,0315	0,0659	0,0640	0,0295	0,0015
10	0,1435	0,0247	0,0007	0,0811	0,0498	0,0197	0,0518	0,0982	0,1058	0,1217	0,1314	0,1087	0,0941	0,0723	0,0436
11	0,0751	0,0520	0,0975	0,0013	0,0091	0,0567	0,0033	0,0065	0,0112	0,0025	0,0055	0,0078	0,0566	0,0880	0,1070
12	0,0496	0,0313	0,0219	0,0577	0,0868	0,0777	0,1065	0,1176	0,1108	0,1097	0,1452	0,1220	0,0693	0,0478	0,0361
13	0,2563	0,0415	0,0609	0,0732	0,0456	0,0707	0,0609	0,0262	0,0734	0,1707	0,0904	0,1240	0,1520	0,1758	0,1911
14	0,1454	0,1581	0,1401	0,0974	0,0303	0,0313	0,0424	0,1116	0,1093	0,0779	0,1323	0,1227	0,1199	0,1178	0,1106
15	0,2313	0,1994	0,0562	0,0677	0,1425	0,2087	0,2536	0,2379	0,2138	0,1432	0,0908	0,0349	0,0396	0,1271	0,1958

Набор ДНМ для данного сочетания факторов приведен в табл. 3. Анализ таблицы показывает, что адекватность модели в течение рассматриваемого периода в отличие от результатов предыдущего варианта расчетов остается на одном уровне. На наш взгляд, это вызвано тем, что с течением времени влияние одних факторов на моделируемый показатель уменьшается, а других увеличивается (например, влияние x_1 уменьшается, а x_2 растет).

Оценка качества имитационной модели проводилась по той же формуле и приведена в табл. 4. Ошибка варьирует на уровне 9%. Точность имитатора во втором варианте (ср. табл. 2 и 4) возросла вдвое, что и подтверждает выдвинутую ранее гипотезу. Дальнейшее повышение точности имитатора может быть обеспечено путем введения нелинейности в рассмотренную модель.

Для окончательной же оценки качества имитатора следует сравнить результаты дискретно-непрерывного моделирования и гидродинамического расчета с фактической динамикой. Из-за отсутствия таких данных по исследуемой совокупности оценку качества имитатора в полном объеме на данном этапе провести не удалось. Однако приведенные результаты расчетов показали принципиальную возможность и конструктивность применения имитационного моделирования в исследовании динамики нефтедобычи.

ЛИТЕРАТУРА

1. **Сергеева Л. А., Чудновский Г. Л., Ягольницер М. А.** Исследование динамики нефтедобычи (ст. в настоящем сборнике).
2. **Борисов Ю. П., Воинов В. В., Рябинина З. К.** Особенности проектирования разработки нефтяных месторождений с учетом их неоднородности. М., «Недра», 1976. 286 с.
3. **Статистические** модели в оптимальном отраслевом планировании. М., «Статистика», 1975. 189 с.
4. **Нейлор Т., Ботон Дж., Бердик Д.** Машинные имитационные эксперименты с моделями экономических систем. М., «Мир», 1975. 500 с.
5. **Макаров А. А., Макарова А. С., Зейлигер А. Н.** Исследование зоны неопределенности оптимального развития сложных экономических систем.— «Экономика и математические методы», 1970, т. VI, вып. 6, с. 849—863.
6. **Лихтенштейн В. Е.** Дискретность и случайность в экономико-математических задачах. М., «Наука», 1973. 375 с.

И. Д. МАНДЕЛЬ, А. И. НИКИТИН

КЛАССИФИКАЦИЯ МНОГОМЕРНЫХ НАБЛЮДЕНИЙ ПРИ АНАЛИЗЕ ОПРЕДЕЛЕНИЙ ПОНЯТИЯ АСУ

В обеспечении идейного единства многочисленных разработок автоматизированных систем управления в нашей стране и в создании общей теории АСУ существует ряд проблем. Среди них не последнее место занимает вопрос об упорядочении понятийной базы и терминологического аппарата теории. Расплывчатость многих понятий (и многозначное толкование их определений), используемых в литературе по АСУ, а также в практике внедрения систем, может привести к ряду нежелательных последствий. В частности, различные определения самого понятия «автоматизированная система управления» у многих авторов представляются одной из причин несовместимости систем и методологической путаницы.

Подобного рода проявления неустоявшейся терминологии неоднократно отмечались советскими специалистами по АСУ. «В настоящее время назрела необходимость обеспечения единого методического подхода к созданию АСУ различных типов... В качестве первоочередной задачи можно выделить формирование единой концепции с целью определения понятия АСУ и основных принципов ее создания»,— писали Н. П. Лапшин и А. А. Модин в 1971 г. [3, с. 3]. По нашему мнению, эта проблема не потеряла своей актуальности и сейчас. «От точного определения понятия АСУ отраслью зависит как состав проекта разработок, так и направление проектных решений. Такое определение должно включать цели системы, критерии ее оценки, условия работы, ее составные части, определение ее границ. Поэтому вовсе не обязательно и, по-видимо-

му, невозможно раскрывать смысл этого понятия в одном предложении»,— замечает Ю. И. Черняк [4, с. 36]. Подобные мнения, а также попытки решения вопроса приведением различных точек зрения к одной «наилучшей» можно найти у многих других авторов [5, с. 115—128; 6, с. 15 и др.].

Конечно, множество различных определений одного понятия само по себе естественно для развивающейся науки и говорит о споре, в котором отыскивается истина. Однако, как замечает А. А. Модин, «нельзя забывать, что... разнообразие подхода к проектированию систем управления (а на этом, безусловно, сказывается мнение специалистов о сущности АСУ.— *И. М., А. Н.*) дорого обходится народному хозяйству, поскольку это, в конечном счете, затягивает процесс разработки и сам процесс внедрения системы» [8, с. 25].

Подобные соображения делают задачу выработки однозначного, «правильного» и, что особенно важно, «всех устраивающего» определения АСУ достаточно актуальной. Весьма существенно, что в методической литературе за последние годы несколько раз давались определения, которые по характеру изданий должны были обеспечить единое толкование содержания данного понятия. Мы имеем в виду определения, содержащиеся во «Временных межотраслевых руководящих материалах по составу и содержанию проектов АСУП...» [17], в «Общеотраслевых РММ по созданию АСУП» [29] и «Руководящих указаниях по разработке и внедрению ОАСУ» [29].

Определения появлялись и в некоторых других изданиях (рекомендации ЦНИИТУ — 1965 г., «вариант ЦЕМИ» определения ВРММ — 1971 г. [3] и др.).

Исследование данного вопроса показало, что ни одно из вышеупомянутых определений не прижилось в литературе достаточно прочно; количество «собственных точек зрения» продолжало расти чуть ли не пропорционально выпуску новых книг и статей по АСУ. Наконец, в 1974 г. был выпущен ГОСТ «АСУ. Основные положения. Термины и определения» (что, кстати, говорит о достаточно высокой значимости проблемы). Так как ГОСТ является обязательным документом, можно было ожидать «закрытия вопроса». Однако этого не произошло. Хотя период 1975—1976 гг. не входил в схему нашего исследования, есть основания считать, что и ГОСТовское определение пока «не

принято» всеми исследователями (см., например, [7, с. 41], где дается определение АСУ, явно не совпадающее с ГОСТовским).

Одной из причин плохой «приживаемости» централизованных определений является то обстоятельство, что в них, по нашему мнению, слабо отражен богатый опыт советских и зарубежных специалистов в разработке АСУ. Это приводит к неполному охвату всех точек зрения и, значит, к недостаточной объективности даваемых определений.

Заметим, что объединение различных точек зрения в «единую», «унифицированную» может и не дать «истинного» представления об объекте — нередко именно одно, уникальное воззрение оказывается единственно верным. Однако применительно к конструкциям, «правдивость» которых не поддается экспериментальной проверке, тезис о соответствии обобщенной точки зрения «настоящему положению вещей» можно ослабить. Достаточно предположить, что унифицированное определение есть просто усредненное обобщение коллективного опыта, «истинность» которого, вообще говоря, фактом унификации еще не доказывается.

Из литературы известны примеры проведения подобного анализа. Так, А. И. Уемовым сделана унификация определений понятия «модель» (по 37 определениям советской и зарубежной литературы) [9].

Основная идея метода унификации, предлагаемого в [10], заключается в следующем. Определения делятся на условно единичные, недробимые признаки (как это было обосновано А. И. Уемовым). Все множество признаков разбивается на классы по типу самих признаков (разбиение носит содержательный, а не формальный характер), затем проводится унификация признаков внутри классов, в процессе которой рассматриваются различные возможные отношения между ними (4 вида). Результатом унификации в классах будут наиболее общие, не противоречащие другим признаки. Итоговое унифицированное определение — логическая функция (обычно конъюнкция) от полученных единичных признаков.

Этот метод был неоднократно опробован, в частности, проведена унификация определений понятия «автоматизированная система управления» (61 определение, 148 признаков), взятых из работ советских авторов с 1963 по

1972 г. Унифицированное определение формулируется так: АСУ — система, в которой людьми на базе научных (в том числе экономико-математических) методов, ЭВМ и других технических средств осуществляется сбор и переработка информации для принятия решений с целью наилучшего управления объектом.

По ГОСТовскому определению, приведенному в [1], АСУ — человеко-машинная система, обеспечивающая автоматизированный сбор и обработку информации, необходимую для оптимизации управления в различных сферах человеческой деятельности.

Как видно, оба определения весьма близки друг другу по содержанию, однако ГОСТовское описывает понятие большего объема. В нем допускается существование АСУ без ЭВМ и о методах управления говорится лишь косвенно; очевидно, имеется в виду, что оптимизация управления сама по себе предполагает применение новых научных методов. Авторы далеки от мысли оспаривать какие-либо точки зрения на данный предмет, тем более официальные определения. Целью работы была попытка дать сравнительно объективную картину дифференциации мнений в этой области.

С этих позиций представляет интерес, по нашему мнению, не только получение унифицированного определения (хотя в конечном итоге именно оно может быть наиболее полезным), но и изучение структуры всего множества определений, которое дало бы возможность проанализировать характерные типы суждений и выявить наличие «противостоящих групп». Для решения этой задачи наиболее адекватны методы классификации многомерных наблюдений, или, в более широком смысле, методы распознавания образов [2].

Было рассмотрено 61 определение АСУ, на которых проводилась унификация. В качестве контрольных в исходное множество включено еще 2 определения — само унифицированное и ГОСТОвское, т. е. классификации подвергалось множество из 63 объектов. Каждое определение записывалось в виде нуль-единичной строки, где единица означала наличие некоторого признака, нуль — его отсутствие. Заметим, что если при унификации определений рассматривались несколько логических отношений между признаками, среди которых существенную роль играли включение (по содержанию) и взаимное отрицание,

то в модели классификации принималось, что на множестве признаков задано одно отношение — тождества (совпадение — несовпадение). Это основывалось на гипотезе, что более тонкие свойства объектов, в том числе интенсиональную связь признаков в определении, можно будет учесть при анализе полученной классификации (что впоследствии подтвердилось).

Конечно, непосредственный учет возможных типов отношений в исходном описании данных более предпочтителен, но он связан с рядом методологических трудностей или с более высокой степенью субъективизма (если, например, описывать отношения включения мерами принадлежности). Дихотомических нетождественных признаков, таким образом, насчитывалось 56, т. е. исходная матрица имела размерность 63×56.

Для решения задачи классификации в пространстве качественных признаков можно применять методы, использующие понятие меры связи между объектами. Подобных мер связи насчитывается довольно большое количество (коэффициенты Юла, Рао, Роджерса — Танимото и др.— см., например, [11]). При выборе меры связи учитывались соображения относительно природы изучаемых объектов: число признаков в разных определениях различно, а представление каждого определения вектором в 56-мерном пространстве логически слабо обосновано. При столь большом числе признаков отсутствие какого-либо признака в данном определении, видимо, не может обладать той же информативностью, что и его наличие. Поэтому многие меры, учитывающие в той или иной степени совпадение или несовпадение нулевых признаков, были отброшены. Определения с точки зрения связи между ними описывались только единичными признаками (их число колебалось от 1 до 7) и по смыслу задачи учитывались только совпадения этих единичных признаков.

Предъявленным требованиям удовлетворяла мера связи В. Л. Устюжанинова [12], имеющая вид

$$S_{ij} = \frac{2m_{ij}}{t_i \log_2 (1 + t_j/t_i) + t_j \log_2 (1 + t_j/t_i)}, \qquad (1)$$

где S_{ij} — информационная мера связи, $0 \leqslant S_{ij} \leqslant 1$;

m_{ij} — число попарных совпадений единичных признаков i-го и j-го объектов;

t_i, t_j — соответственно число единичных признаков i-го и j-го объектов.

В качестве алгоритма таксономии был выбран эвристический метод диагонализации матрицы связи, предложенный А. В. Беккером и М. Л. Лукацкой [13]. Его основная идея заключается в постепенном (пошаговом) подформировании объектов друг к другу по максимальным значениям мер связи данного объекта к каждому из уже рассмотренных. К достоинствам алгоритма относятся его инвариантность относительно выбора первого элемента, с которого начинается работа, а также сравнительно высокая скорость работы. Существенным недостатком, отличающим метод от «автоматической классификации», представляется отсутствие критерия выделения классов в диагонализированной матрице: это делается исследователем, причем вопрос иногда решается неоднозначно.

Была составлена программа на языке ФОРТРАН[1]. Реализация проводилась на ЭВМ ЕС-1020, время работы программы (с полной выдачей результатов) — около 20 мин для нашей матрицы 63×63.

Фрагмент диагонализированной матрицы связей приведен в табл. 1. По нему можно видеть, что визуальное выделение классов в матрице, предусматриваемое в алгоритме диагонализации, не всегда является однозначным: классы 1 и 2 можно принять за один, а в некоторых случаях возникает желание разбить выделенный класс на более мелкие. Для проведения более объективной классификации логично применить функционалы качества разбиения, ряд которых описан, например, в [15]. К сожалению, алгоритмическое совмещение идей упорядочения структуры матрицы связей (с разбиением, допустим, сначала на «плохие» с точки зрения функционала классы) и изменения классификации по критерию качества еще не разработано (или нам об этом неизвестно). Методы таксономии, использующие функционалы качества, как правило, комбинаторны, т. е. обладают недостатком (с вычислительной точки зрения), который отсутствует у алгоритмов диагонализации, а они, в свою очередь, не дают четких границ классов, позволяющих применять критерии качества.

[1] К сожалению, авторы поздно узнали о выходе сборника с работой [14], в которой Л. В. Говорская описывает программу по данному алгоритму на α-языке.

Таблица 1

Фрагмент диагонализированной матрицы связей (классы 1 и 2)

Объект	Класс 1									Класс 2									
	1	58	31	42	33	2	37	48	40	56	62	53	54	13	50	45			
1	1	0,82	0,82	0,67	0,67	0,67	0,82	0,67	0,58	0,33	0,33	0,0	0,67	0,41	0,29	0,23			
58		1	1,00	0,82	0,82	0,82	0,50	0,41	0,73	0,41	0,41	0,0	0,41	0,50	0,36	0,29			
31			1	0,82	0,82	0,82	0,50	0,41	0,73	0,41	0,41	0,0	0,41	0,50	0,36	0,29			
42				1	0,67	0,67	0,41	0,33	0,58	0,33	0,33	0,0	0,33	0,41	0,29	0,23			
33					1	0,67	0,41	0,33	0,58	0,67	0,67	0,33	0,33	0,41	0,29	0,23			
2						1	0,41	0,33	0,58	0,67	0,33	0,0	0,33	0,41	0,58	0,45			
37							1	0,82	0,36	0,0	0,0	0,0	0,41	0,0	0,0	0,0			
48								1	0,29	0,0	0,0	0,0	0,33	0,0	0,0	0,0			
40									1	0,29	0,58	0,29	0,29	0,36	0,50	0,38			
56										1	0,67	0,33	0,33	0,41	0,58	0,45			
62											1	0,67	0,33	0,41	0,58	0,45			
53												1	0,0	0,0	0,29	0,45			
54													1	0,82	0,29	0,23			
13														1	0,36	0,29			
50															1	0,77			
45																1			

Таким образом, расчеты, связанные с вычислением функционала, проводились вручную. Выбранный функционал обобщает некоторые другие результаты в этой области и предложен В. Л. Куперштохом, Б. Г. Миркиным и В. А. Трофимовым (см. [16]), а также статью этих авторов в «Автоматике и телемеханике» — 1976, № 3).

Показателем качества разбиения $R = \{R_1, \ldots, R_m\}$ является число

$$f(a, R) = \sum_{s=1}^{m} \sum_{i,j \in R_s} (a_{ij} - a), \qquad (2)$$

где $f(a, R)$ — показатель качества при пороге существенности a;

a_{ij} — мера связи i-го и j-го объектов;

R_s — s-й класс исходного множества;

m — количество классов в разбиении.

Относительно функционала (2) в [16] доказан ряд предложений, утверждающих, что классификация, максимизующая (2), является оптимальной в смысле четырех обычных требований, которые предъявляются в таксономии объектов в пространстве их признаков (см. также [15, с. 87] и др.). Кроме того, в [16] дан алгоритм получения оптимального разбиения (опять-таки комбинаторный) и способ определения наилучшего значения порога существенности. Подобные соображения, как отмечалось, не ложились сразу в основу классификации определений, и нами была проделана следующая работа: в диагонализированной матрице выделялись классы в нескольких вариантах, по каждому из которых считался функционал (2); порог существенности принимался 0,2 и 0,3 (отношения функционалов не менялись).

Максимальному значению (2) соответствует классификация, основные характеристики которой приведены в табл. 2, где Q_s — меры средней концентрации в класах; $Q_s = \sum_{i,j \in R_s} a_{ij}/m_s^2$; m_s — число элементов в s-м классе; $\rho_{ks} = \sum_{i \in R_s, j \in R_k} a_{ij}/m_s m_k$ — меры средней связи между k-м и s-м классами (см., например, [15, с. 83]).

Полученное разбиение было проверено на оптимальность в смысле предложений, доказанных в [16]. Видно, что $\rho_{1,9} > Q_9$; $\rho_{8,9} > Q_9$; $\rho_{6,11} > Q_{11}$, т. е., строго говоря,

Таблица 2

Характеристики классификации в диагонализированной матрице связей

Номер класса	1	2	3	4	5	6	7	8	9	10	11	12
m_s	9	7	5	6	2	4	2	6	3	7	2	2
Q_s	0,598	0,415	0,550	0,513	0,500	0,437	0,820	0,483	0,330	0,624	0,330	0,410
1	1											
2	0,286	1										
3	0,053	0,046	1									
4	0,000	0,100	0,273	1								
5	0,002	0,080	0,072	0,227	1							
6	0,004	0,169	0,084	0,066	0,062	1						
7	0,307	0,346	0,000	0,048	0,000	0,171	1					
8	0,372	0,038	0,000	0,104	0,427	0,074	0,088	1				
9	0,344	0,000	0,074	0,051	0,000	0,023	0,000	0,350	1			
10	0,134	0,130	0,005	0,001	0,000	0,008	0,179	0,000	0,013	1		
11	0,052	0,121	0,144	0,000	0,000	0,335	0,000	0,000	0,000	0,018	1	
12	0,041	0,060	0,180	0,000	0,010	0,000	0,000	0,000	0,000	0,002	0,01	1

разбиение не является компактным. Однако все классы образуют сгущения в среднем [16, с. 135], т. е. меры внутриклассовой концентрации выше среднего расстояния элементов класса от всех прочих элементов множества. Таким образом, разбиение можно считать близким к оптимальному, так сказать, не по максимуму функционала (2), а по выполнению условий оптимальности.

Рассмотрим содержание полученных классов определений, давая лишь краткие характеристики объектов в классах с приведением типичных определений.

К л а с с 1 включает в себя совокупность очень близких друг другу определений, варьирующих преимущественно три признака: АСУ — совокупность методов и средств (1), предусматривающая оптимизацию управления (2) на основе автоматизированной обработки данных и экономико-математических методов (3). Приведем официальное определение: «АСУП — совокупность методов и технических средств, обеспечивающих наиболее эффективное выполнение функций управления предприятием на основе использования экономико-математических методов и средств обработки данных, включающих в себя ВТ, средства сбора, фиксации, передачи и хранения информации» [17, с. 4—5].

В к л а с с 2 попали ГОСТовское и унифицированное определения; первое — под номером 50 в табл. 1, второе — под номером 62; мера связи между ними достаточно высока (0,58). В классе варьируют восемь признаков, из которых наиболее плотно расположены пять: система (1); человеко-машинная система (2); осуществляется сбор и переработка информации для принятия решений (3); направлена на оптимизацию управления (4); управление осуществляется на базе научных (и экономико-математических) методов, ЭВМ и других технических средств (5). Остальные три признака даны по одному в различных определениях. Приведем с некоторыми сокращениями одно определение из этого класса: «АСУП представляет собой систему, базирующуюся на ЭВМ и других технических средствах, предназначенных для комплексной автоматизированной... переработки всей информации, используемой для управления объектом путем выработки оптимальных управленческих решений ... и реализации принятых решений через обратную автоматизированную связь» [18, с. 48—49].

Устойчивая разница типов признаков 1(1) и 2(1) с 2(2) у определений классов 1 и 2 характеризует субстанциональное толкование понятия «система» в классе 1 (автоматизированная система управления — комплекс методов и средств), а в классе 2 даются определения через род и видовое отличие (АСУ — система...). Эти типы признаков повторяются почти во всех классах полученного разбиения.

Заметим, что в целом класс 2 менее компактен, чем класс 1; его определениям присущ некоторый акцент на технические средства управления. Существенным отличием является также указание на «промежуточную цель системы» — ее помощь при принятии решений, тогда как общие цели в определениях обоих классов совпадают — оптимизация управления.

В к л а с с е 3 пять признаков из шести образуют весьма компактную группу: «система» (1); «используются методы экономической кибернетики» (2); «средства автоматизации» (3); «предназначена для переработки информации» — «промежуточная цель» (4); «направлена на наиболее эффективное функционирование объекта» (5). Шестой признак (в двух определениях) — «система управления». Типичное определение класса формулируется так: «Под АСУ следует понимать систему, использующую методы экономической кибернетики с применением автоматизированных устройств для получения, передачи и обработки информации. Цель ее — обеспечение наиболее экономически эффективного функционирования объекта» [19, с. 6].

Существенное отличие определений этого класса от двух предыдущих состоит в ином понимании целей системы: не оптимизация управления, а оптимизация функционирования объекта. Эта разница имеет, видимо, немалое значение.

К л а с с 4 близок к классу 3. Компактной группой расположены признаки 3(1), 3(3), 3(4) — во всех определениях. Добавляется признак «определенные функции выполняются человеком» (4). Приведем характерное определение: «АСУ — система, в которой для получения, обработки информации и управления используются различные устройства, однако определенные функции выполняются человеком или группой людей» [20, с. 28]. Специфичная акцентировка определений этого типа — выявление от-

личий АСУ от «простых» систем управления, с одной стороны, и от автоматических — с другой; в них ничего не говорится о методах управления и целях системы.

В классе 5 два определения утверждают, что АСУ — человеко-машинная система. «АСУ — понятие несколько условное. По существу, это человеко-машинная система, в которой ведущее место принадлежит человеку» [21, с. 72].

Класс 6 — признаки расположены весьма некомпактно; сравнительно высокая мера внутриклассовой концентрации (0,437) объясняется «цепными» связями определений по различным от пары к паре признакам. Единственный общий для всех признак— «наличие технических средств в системе». Под АСУ понимается то «комплекс задач», то «организационно-технологический комплекс», то «комплекс процессов управления». Приведем более или менее представительное определение: «АСУ представляет собой организационно-технологический комплекс, обеспечивающий выполнение функций управления на основе использования экономико-математических методов и современных средств обработки данных (ЭВМ, устройств накопления, регистрации, отображения и др.)» [22, с. 425].

Заметим, что цель системы видится в выполнении (а не улучшении) функций управления.

В классе 7 два определения близки к некоторым признакам класса 6 и почти совпадают между собой: «Под АСУП понимается единый организационно-технологический комплекс, направленный на выполнение функций управления промышленным производством» [23, с. 12].

Для класса 8, как и 6, характерна весьма сильная разбросанность признаков; один общий для всех определений признак — «комплекс методов и средств» (1); в три определения входит признак «для решения задач управления» (2); в три другие —«решает задачи планирования и учета» (3); остальные признаки (их всего 8) единичны. Приведем определение, имеющее максимальную связь с другими объектами класса: ««Под АСУ следует понимать совокупность взаимоувязанных между собой технических средств, правил, инструкций, математических методов и форм организации сбора, передачи и переработки информации, позволяющих осуществлять с помощью частично остающегося технического персонала рациональное управление контролируемыми процессами по установленным для этого алгоритмам и регламентам» [24, с. 21].

В этом классе есть один интересный признак — «под АСУ следует понимать и сам объект управления», а также мнение о том, что в условиях АСУ необходимо минизировать затраты на управление.

Класс 9 мал, и для него типично следующее определение: «АСУП — комплекс методов и устройств, служащих информационно-технической основой общего управления и реализующих функции механизированной обработки информации» [25, с. 72].

В классе 10 варьируют в основном пять признаков: «технологический комплекс» (1), «осуществляет обработку планово-экономической информации» (2), «предназначена для своевременного получения данных» (3), «обеспечивает эффективное управление« (4). Пятый признак, из которого только и состоят практически три определения класса, формулируется так: «АСУ — совокупность всех устройств системы управления без участия человека» (5). Его можно условно считать тождественным признаку 10(1). Приведем определение, почти идентичное всем определениям класса (кроме упомянутых трех): «АСУ управления и планирования — некоторый технологический комплекс, наиболее рационально осуществляющий обработку планово-экономической информации для своевременного получения управляющих данных, позволяют организовать эффективное экономическое и административное управление» [26, с. 19].

Таким образом, определения класса 10, как и 9, носят явно выраженный «технократический» характер.

В классе 11 два определения описываются четырьмя довольно близкими попарно признаками: «ОАСУ — целенаправленная неразрывная совокупность методов (правил), органов и средств управления, обеспечивающая выполнение поставленных перед промышленностью отрасли задач, базирующаяся на использовании математических методов и средств ВТ» [27, с. 528]. Кроме выражения методов через методы, здесь характерно понимание цели АСУ как выполнения поставленных задач.

В класс 12 попали два близких определения «кибернетического» характера. Одно из них: «АСУ — это человеко-машинная структура, создаваемая в организации применительно к схеме замкнутого контура регулиро-

вания и предназначенная для выработки оптимальных или близких к ним решений» [28, с. 17].

Восемь определений ни в какие классы не попали.

Была сделана попытка связать полученное разбиение со временем публикации соответствующих определений. Однако какой-либо закономерности не обнаружено, т. е. близкие определения давались в разное время, хотя определений со ссылками на более ранний источник обнаружено всего два.

Таким образом, с одной стороны, налицо весьма сильная дифференциация мнений по данному вопросу. С другой стороны, при анализе полученных результатов оказывается, что классификация в немалой степени объясняется разными сочетаниями в определениях сравнительно небольшого числа признаков. Действительно, из 56 исходных признаков 28 являются единичными, а 10 повторяются лишь дважды. Можно считать, что чем больше часто встречающихся признаков объясняют связи в отдельном классе, тем значимее этот класс с точки зрения представления им «единого мнения». Далее, чем больше размер класса, т. е. чем больше определений данного типа, тем этот класс «информативнее», «представительнее» (по крайней мере, в нашем случае).

Исходя из этих простых соображений, мы «для узких целей» подсчитали общую информативность каждого класса по формуле

$$w_s = \mu_s \sum_{q \in L_s} F_q^s p_q, \qquad (3)$$

где $\mu_s = m_s/M$ — удельный вес s-го класса в классификации (M — общее число объектов, $M = 63$, m_s — число объектов в классе);

$p_q = P_q/N$ — удельный вес q-го признака ($q = \overline{1,56}$) в общем числе реализаций единичных признаков (P_q — количество реализаций признака q, $P_q = \overline{1,18}$); N — общее число реализаций всех признаков, $N = 155$; L_s — множество признаков, описывающих объекты класса s; F_q^s — количество реализаций признака на объектах класса s.

Большему значению функционала w_s соответствует интуитивное представление о большей значимости в общей картине данного класса, хотя его приведенная форма,

возможно, не самая лучшая (например, здесь не учитывается связь между классами).

Расчеты показали, что наибольшее значение (3) принимает в классах 1 и 2, причем $w_1 > w_2$. Различие между R_1 и R_2 в значительной степени объясняется четким «развалом» по классам двух очень частых признаков: «комплекс методов и средств» 1(1) и «система» 2(1). В то же время большая доля связи между классами вызвана совпадением признаков «оптимизация управления» и (в меньшей степени) «применение экономико-математических методов управления». Таким образом, если заменить понятие «комплекс методов и средств» в классе 1 понятием «система» или «человеко-машинная система», то два класса, по-видимому, сольются в один большой, в котором и будут сконцентрированы почти все часто встречающиеся признаки. Так как в классе 2 находятся ГОСТовское и унифицированное определения, можно сделать вывод, что оба они описывают весьма распространенные точки зрения и представляют тем самым удовлетворительное обобщение опыта советских специалистов по АСУ.

Заметим, что «индивидуальный индекс представительности» определения $\left(\sum_{q \in a_i} F_q^i p_q/m_i, \text{ где } F_q^i \in \{0,1\}\right)$ означает наличие или отсутствие признака q в a_i-определении, m_i — общее число признаков в определении), показывающий среднее количество повторений, приходящихся на один признак, несколько выше для унифицированного определения, чем для ГОСТовского (11 и 9,33); средняя связь этого определения с объектами классов 1 и 2 (0,411) также выше, чем у ГОСТовского (0,369). Это понятно, если учесть, что в унифицированном определении больше признаков (4, а не 3), причем один из них — «применение научных (в том числе экономико-математических) методов управления» — встречается часто (9 раз).

* * *

По нашему мнению, в изучаемую выборку попали практически все типы возможных определений АСУ или, по крайней мере, почти все возможные признаки. Представляет интерес дальнейшее изучение проблемы с точки зрения набора новых (после ГОСТовских) определений. Весьма интересно также было бы проследить связь между

определениями АСУ (особенно даваемыми «производственниками», а таких немало) и реально внедряемыми системами. Заметим по этому поводу, что существенный признак многих определений (в том числе унифицированного и ГОСТовского) — «оптимизация управления» — ориентирован фактически на системы завтрашнего дня, указывая на конечную цель системы, которая в строгом смысле слова, как известно, нигде не достигнута.

Впрочем, подробный содержательный разбор полученных результатов является самостоятельной темой и, возможно, будет проведен в дальнейшем. Пока существенным выводом, как нам представляется, может быть, кроме самой по себе таксономии, подтверждение достаточно высокой значимости ГОСТовского и унифицированного определений и тем самым выявление признаков, наиболее полно характеризующих объект.

Авторы дают себе отчет в том, что при анализе множества объектов, описанных номинальными признаками, можно пользоваться и более тонкими методами, чем диагонализация матрицы связей (см., например, [16] и др.). Однако при работе с методами классификации многомерных наблюдений и смежных разделов нередко окончательным аргументом в пользу какого-то данного алгоритма является «хорошее качество» полученных результатов, так как единой теории, вообще говоря, еще нет. В нашем случае классы в целом весьма логично разбивают множество и поддаются доступной интерпретации. Было бы интересно углубить анализ изучением наиболее информативных признаков, а также выявлением непосредственной структуры матрицы «объект-признак».

ЛИТЕРАТУРА

1. **ГОСТ 19675—74.** АСУ. Основные положения. Термины и определения. М., Госстандарт, 1974.
2. **Загоруйко Н. Г.** Методы распознавания и их применение. М., «Сов. радио», 1972. 206 с.
3. **АСУП.** Кишинев, «Картя молдовеняскэ», 1971. 92 с.
4. **Черняк Ю. И.** Автоматизированные системы отраслевого управления.— В кн.: Вопросы проектирования ОАСУ. М., «Наука», 1970, с. 33—48.

5. **Евдокимов В. В.** К вопросу классификации АСУП.— В кн.: Труды ЛИЭИ им. П. Тольятти, вып. 72. Л., 1970, с. 117—127.
6. **Васманов В. В.** Автоматизированные системы оперативного управления. М., «Энергия», 1970. 264 с.
7. **Проектирование** АСУ на автомобильном транспорте. М., «Транспорт», 1975. 93 с.
8. **Модин А. А.** О разработке проектов АСУ промышленным предприятием.— В кн.: Применение экономико-математических методов и ЭВМ в управлении промышленным предприятием. М., «Машиностроение», 1968, с. 24—46.
9. **Уемов А. И.** Логические основы метода моделирования. М., «Мысль», 1971. 311 с.
10. **Мандель И. Д.** К вопросу об унификации определений.— «Философские науки», 1975, № 6, с. 134—138.
11. **Гладких Б. А.** Некоторые проблемы классификации.— В кн.: Распознавание образов в экономико-математическом моделировании. Новосибирск, ИЭиОПП СО АН СССР, 1974, с. 5—31.
12. **Устюжанинов В. Л.** Информационные меры и их использование в социологическом анализе.— В кн.: Измерение и моделирование в социологии. Новосибирск, «Наука», 1969, с. 62—86.
13. **Беккер А. В., Лукацкая М. Л.** Об анализе структуры матрицы коэффициентов связи.— В кн.: Вопросы экономико-статистического моделирования и прогнозирования в промышленности. Новосибирск, ИЭиОПП СО АН СССР, 1970, с. 273—285.
14. **Говорская Л. В.** Группировка элементов методом диагонализации матрицы расстояний.— В кн.: Алгоритмы статистической обработки информации. Новосибирск, ИЭиОПП СО АН СССР, 1974, с. 55—72.
15. **Айвазян С. А., Бежаева З. И., Староверов С. В.** Классификация многомерных наблюдений. М., «Статистика», 1974. 240 с.
16. **Куперштох В. Л., Миркин Б. Г., Трофимов В. Л.** К обоснованию одного критерия классификации.— В кн.: Методы моделирования и обработка информации. Новосибирск, «Наука», 1976, с. 133—141.
17. **Временные** межотраслевые руководящие методические материалы по составу и содержанию проектов АСУП с дискретным характером производства, последовательности их разработки и внедрения. М., Госкомитет по науке и технике, 1968.
18. **Твердохлеб Н. Г.** МОЭИ промышленных предприятий. М., «Статистика», 1971. 496 с.
19. **АСУП.** Одесса, «Маяк», 1971. 75 с.
20. **Справочник** проектировщика систем автоматизации управления производством. М., «Машиностроение», 1971. 426 с.

21. **АСУ.** Теория и методология. Т. 1. М., «Мысль», 1972. 455 с.
22. **Основы** экономики и управления производством. М., «Экономика», 1972. 447 с.
23. **Материалы** 1-й Всесоюзной конференции по экономической кибернетике. М., ЦЭМИ АН СССР, 1966.
24. **Лоскутов В. И.** Проблемы современной техники управления. М., «Знание», 1972. 64 с.
25. **Автоматизированная** система управления предприятием. М., Дом научно-технической пропаганды им. Дзержинского, 1969. 58 с.
26. **Истомин Л. И., Пирмухаммедов А. Н.** Методические аспекты разработки АСУ и планирования. Ташкент, ИНТИ и пропаганды, 1970. 24 с.
27. **Проблемы** научной организации управления социалистической промышленностью. Материалы 2-й Всесоюзной научно-технической конференции, сб. 1, ч. 2. ВИНИТИ, 1972, с. 437—901.
28. **Авдеев Ю. А.** Методологические вопросы разработки и применения АСУ при выполнении сложных проектов.— В кн.: Моделирование процессов управления. Вып. 2. Новосибирск, Институт гидродинамики СО АН СССР, 1969, с. 7—37.
29. **Автоматизированные** системы управления. Применение ВТ и АСУ на предприятиях и в отраслях промышленности. М., «Экономика», 1972. 399 с.

Г. П. ГВОЗДЕВА, Е. Е. ГОРЯЧЕНКО,
И. Б. МУЧНИК, Е. Э. СИНИЦЫН

ИСПОЛЬЗОВАНИЕ ТИПОЛОГИЧЕСКОГО АНАЛИЗА ПРИ ПЛАНИРОВАНИИ ВЫБОРОЧНЫХ ОБСЛЕДОВАНИЙ

В настоящей работе показано, как типология сложных социально-экономических объектов может быть использована для построения выборки при проведении экономико-статистических исследований.

МЕТОДИЧЕСКИЕ ЗАМЕЧАНИЯ

Выборки, которые строятся на основе того или иного расслоения рассматриваемой совокупности объектов, называют стратифицированными. В процессе формирования таких выборок предлагается выделять четыре задачи: 1) расслоение (стратификацию) совокупности; 2) размещение выборки и отбор объектов из слоев (формирование выборки); 3) оценивание существенных для исследования параметров, значения которых для объектов совокупности известны заранее (тестирование выборки); 4) оценивание параметров, полученных в результате выборочного обследования (использование выборки). При этом первая из названных задач чаще всего рассматривается как уже решенная, т. е. используется готовое расслоение, либо решается на основе информации о каком-то одном параметре, существенно связанном с изучаемым явлением.

На практике обычно имеется много параметров, информация о которых известна до намеченного обследования для всех объектов совокупности. Однако эта априорная информация, как правило, лишь косвенно связана с задачами конкретного исследования. Многомерный характер

описаний объектов и косвенный характер параметров, лежащих в основе этих описаний, требует разделения всей совокупности объектов на однородные группы. В таких случаях естественно воспользоваться алгоритмами автоматической классификации [1]. При этом оценка однородности окончательно формируемых групп должна подвергаться тщательному анализу.

Обычные алгоритмы автоматической классификации предоставляют для такого анализа крайне скудную информацию. Фактически это лишь сведения о том, какой из групп принадлежит определенный объект совокупности. В случае неудачного результата такое оценивание не дает указаний, как менять само расслоение. Лингвистический же метод типологического анализа [2—4], представляющий собой естественное сочетание человеко-машинных процедур, дает возможность активно вмешиваться в процесс формирования однородных групп объектов (типов) и содержательно анализировать получающиеся результаты. Поэтому он дает уверенность в том, что формируемые группы объектов действительно качественно различны.

Лингвистический подход основан на сочетании методов экстремальной группировки параметров и автоматической классификации объектов. В процессе анализа решаются две основные задачи. Одна из них состоит в разбиении множества параметров на группы «близких» или «сильно коррелированных» между собой и в формировании обобщенных факторов, наиболее тесно связанных со всеми параметрами группы (факторный анализ пространства свойств). Другая задача заключается в классификации объектов по каждому из выделенных факторов (частная классификация) и по всем факторам одновременно (типология объектов).

При использовании лингвистического подхода для построения выборки открываются две дополнительные возможности. Первая из них состоит в том, что каждый раз, когда возникает необходимость строить новую выборку, предварительно определяется, по какому фактору она будет строиться. Классификация объектов по выбранному фактору служит основой построения выборки. Использование частной классификации объектов по отдельному фактору рассматривалось в ряде работ [5, 6].

Вторая возможность связана с использованием в качестве основы построения выборки не частной классификации, а типологии, учитывающей классификации по всем факторам. Одновременный учет нескольких факторов важен при организации комплексных исследований. В этом случае выборка строится таким образом, чтобы выборочное распределение объектов по типам хорошо отражало соответствующее распределение всей совокупности.

Распределение объектов по типам можно рассматривать как важную структурную характеристику изучаемой совокупности объектов при условии, что сама типология содержательно обоснована и проверена на существенность. Такое «повышенное» требование к качеству используемого расслоения особенно важно в тех случаях, когда вся совокупность объектов относительно небольшая, а формируемая выборка должна содержать буквально несколько единиц наблюдения.

Именно для такого случая предлагается описываемая ниже методика построения выборки. Методические положения излагаются на примере решения конкретной задачи построения двух подвыборок поселений для пробного и основного обследований бюджетов времени сельского населения Новосибирской области. При этом раскрываются как общий характер предлагаемой методики, так и методические проблемы, вытекающие из особенностей конкретной задачи.

ПОСТРОЕНИЕ ВЫБОРКИ

В рамках крупного комплексного исследования постоянно возникает потребность в проведении более частных, но в то же время углубленных исследований по отдельным проблемам. Такие исследования, проводимые по существенно расширенным программам, невозможно базировать на большой выборке, используемой для основного исследования, так как ее реализация очень трудоемка. Для углубленного изучения отдельных проблем имеет смысл построение специальных подвыборок, репрезентативных в рамках этих проблем.

Так, если основная выборка, используемая для комплексного изучения сибирской деревни, включает в себя 157 сельских поселений Новосибирской области, в кото-

рых проживает около 10% сельского населения [1], то для исследования бюджетов времени потребовалось построение двух подвыборок небольшого объема. При этом условия проведения бюджетных обследований налагали достаточно жесткие ограничения на выборочную совокупность: во-первых, ее максимально возможный объем не должен был превышать 800 чел. в основном и 300 чел. в пробном обследовании, и, во-вторых, территориально подвыборка должна была охватывать поселения, принадлежащие не более чем трем сельсоветам в основном и одному сельсовету в пробном обследовании.

Была выдвинута гипотеза, что затраты времени сельского населения на основные виды деятельности существенно дифференцируются по поселениям, различающимся достигнутым уровнем социально-экономического и демографического развития. Так, уровень развития социальной инфраструктуры (урбанизация поселения) в значительной степени определяет затраты и структуру свободного времени. Чем более развита сфера обслуживания, тем меньше затраты времени в быту, богаче возможности проведения свободного времени. Уровень развития экономической структуры производства существенно влияет на режим труда, размеры личного подсобного хозяйства, затраты труда на его ведение, характер жилого фонда, его благоустройство, а следовательно, и на организацию быта. Фактор естественного прироста и возрастного состава населения определяет соотношение детей и взрослых в составе населения, а вместе с тем и занятость женщин, структуру распределения их суточного фонда времени. Кроме того, в зависимости от соотношения трудоспособного населения и пенсионеров существенно меняется совокупный баланс времени населения, распределение ролей в быту между отдельными членами семьи. Следовательно, достигнутый уровень социально-экономического и демографического развития поселения в значительной степени определяет распределение и использование времени населением, а типология поселений по указанным факторам может успешно использоваться для построения выборки бюджетного обследования.

[1] Подробно методика построения выборки для основного обследования изложена в работе [7].

На основе 28 индикаторов социально-экономического и демографического развития поселений была сформирована трехфакторная социально-демографическая типология [2]. В табл. 1 приведена структура выделенных факторов (состав входящих в них параметров и факторные нагрузки). По факторам урбанизации и естественного прироста [3] выделялись два класса поселений, а по фактору экономической структуры — три. Типология, использованная для построения выборки бюджетного обследования, основывалась на трех указанных факторах и была получена пересечением частных классификаций по каждому из факторов. Всего было образовано 12 типов. Распределение поселений по типам приведено в табл. 2.

Таким образом, задача состояла в том, чтобы отобрать такие поселения, входящие в состав не более чем трех сельсоветов для основного и одного сельсовета для пробного обследования, распределение которых по типам максимально соответствует распределению всей совокупности. На языке теории выборки это означает, что мы предлагаем использовать принцип пропорционального отбора с соответствующими ограничениями. Однако практическое использование этого принципа нуждается в уточнении. Дело в том, что не существует единственного критерия, позволяющего количественно оценить качество такой выборки, и, кроме того, отсутствует априорная возможность судить, какой из существующих критериев лучше. Было решено использовать 4 наиболее естественных для данного случая и существенно различных критерия, отбор осуществлять независимо по каждому из них, а полученные выборки сравнивать по результатам их тестирования. Таким путем не только сформирована окончательная выборка, но и уточнена процедура ее нахождения.

[2] Обоснование системы используемых индикаторов, методика построения типологии поселений, содержательное описание полученных типов подробно изложены в работе [8].

[3] Авторы сознательно отказались от использования при построении типологии других демографических факторов, помимо естественного прироста. Это связано с тем, что как показатели трудового потенциала, так и показатели механического движения населения самостоятельных факторов не образуют, а объединяются с социальными факторами, причем с достаточно высокими нагрузками.

Таблица 1

Структура факторов социального и демографического развития сельских поселений

Фактор	Наименование параметра	Коэффициент корреляции параметра с фактором
Урбанизация (I)	Отраслевая сложность непроизводственной сферы поселения	0,886
	Уровень библиотечного обслуживания	0,875
	Административное значение поселения	0,821
	Уровень медицинского обслуживания	0,821
	» торгового »	0,815
	Людность	0,800
	Уровень бытового обслуживания	0,796
	» почтового »	0,774
	» школьного »	0,760
	» развития общественного питания	0,750
	Место поселения во внутрихозяйственном расселении	0,705
	Число рабочих мест в торговле	0,700
	Доля трудоспособного населения, занятого в непроизводственной сфере	0,644
	Отраслевая сложность сферы материального производства	0,645
	Обеспеченность дошкольными учреждениями	0,636
	Уровень клубного обслуживания	0,622
	Расстояние до центральной усадьбы	—0,593
Экономическая структура (II)	Социально-отраслевой тип предприятия	0,818
	Доля трудоспособного населения, занятого в сельском хозяйстве	—0,771
	Средний доход от личного подсобного хозяйства на семью в год	—0,738
	Доля трудоспособного населения, занятого в промышленности, строительстве и транспорте	0,702
	Доля семей, проживающих в домах, принадлежащих государству или колхозам	0,531
Естественный прирост (III)	Коэффициент плодовитости женщин	0,992
	Коэффициент рождаемости населения	0,904
	Естественный прирост населения	0,855
	Доля детей и подростков в возрасте до 15 лет во всем населении	0,739
	Средний размер семьи	0,666
	Доля лиц пенсионного возраста во всем населении	—0,634

Таблица 2

Распределение поселений по типам

Описание типов	Число поселений	Доля поселений в типе, %	Доля населения в типе, %
Неурбанизированные с менее развитой экономической структурой и низким естественным приростом	31	19,8	8,8
Неурбанизированные с менее развитой структурой и высоким естественным приростом	17	10,8	3,9
Неурбанизированные со среднеразвитой экономической структурой и низким естественным приростом	29	18,5	10,8
Неурбанизированные со среднеразвитой экономической структурой и высоким естественным приростом	24	15,3	8,1
Неурбанизированные с высокоразвитой экономической структурой и низким естественным приростом	12	7,6	6,3
Неурбанизированные с высокоразвитой экономической структурой и высоким естественным приростом	2	2,3	0,8
Урбанизированные с менее развитой экономической структурой и низким естественным приростом	5	3,2	4,6
Урбанизированные с менее развитой экономической структурой и высоким естественным приростом	2	1,3	1,2
Урбанизированные со среднеразвитой экономической структурой и низким естественным приростом	13	8,3	18,5
Урбанизированные со среднеразвитой экономической структурой и высоким естественным приростом	6	3,8	4,2
Урбанизированные с высокоразвитой экономической структурой и низким естественным приростом	12	7,6	27,2
Урбанизированные с высокоразвитой экономической структурой и высоким естественным приростом	4	2,5	5,6

ОСНОВНЫЕ ПРОЦЕДУРЫ ОТБОРА

Перейдем к детальному изложению реализованных процедур отбора поселений для изучения бюджетов времени сельского населения. Эти процедуры типичны в решении задач подобного рода.

Обозначим выбранные критерии оценки соответствия выборочного распределения всей рассматриваемой совокупности через $I_l (l = 1, \ldots, 4)$. Пусть $P_{\alpha j}$ — доля поселений, принадлежащих классу α по j-му фактору ($\alpha = \overline{1, K}$; $j = \overline{1, F}$), во всей совокупности объектов, а P_τ — доля поселений, отнесенных к τ-му типу, образованному пересечением классов по всем выделенным факторам ($\tau = \overline{1, T}$). $\hat{P}_{\alpha j}$ и $\hat{P}_\tau$ — соответствующие доли в выборочной совокупности. Тогда при разбиении совокупности на K_j классов по каждому из факторов и общем числе полученных типов T

$$I_1 = \sum_{\tau=1}^{T} (P_\tau - \hat{P}_\tau)^2;$$

$$I_2 = \sum_{\tau=1}^{T} |P_\tau - \hat{P}_\tau|;$$

$$I_3 = \sum_{j=1}^{F} \sum_{\alpha=1}^{K} (P_{\alpha j} - \hat{P}_{\alpha j})^2;$$

$$I_4 = \sum_{j=1}^{F} \sum_{\alpha=1}^{K} |P_{\alpha j} - \hat{P}_{\alpha j}|.$$

Каждый из предложенных критериев дает правило приписывания количественной оценки соответствия распределения данной выборки распределению всей совокупности. Иными словами, значение критерия является функцией от выборки. Если конкретную выборку из множества S допустимых выборок заданного объема обозначить через s, то численное значение критерия на конкретной выборке будет равно $I_l(s)$. Выборка s^*, минимизирующая значение критерия I_l, является наилучшей (в смысле критерия I_l):

$$I_l(s^*) = \min_{s \in S} I_l(s),$$

где S — множество допустимых выборок.

Отметим две особенности использованных критериев.

1. Первые два оценивают разницу между распределениями по типам непосредственно, а вторые два — косвенно, через оценку отклонений между распределениями по классам частных классификаций, из которых составлена типология. В этом смысле первые два критерия более точно соответствуют поставленной задаче. Вместе с тем в тех случаях, когда вся рассматриваемая совокупность, а соответственно и выборка имеют малый объем, возникает эффект безразличия. Он состоит в том, что, с одной стороны, доли многих типов во всей совокупности одинаковы, а с другой — необходимый объем выборки не позволяет представить все типы. В результате оказывается несколько вариантов выборок с одинаковым значением критерия.

В этих условиях даже оптимальные выборки характеризуются большими отклонениями, а эффект безразличия подчеркивает искажения: в то время как доли одних классов некоторых классификаций получают в выборке относительно малые отклонения, на доли других приходится основная часть суммарных, причем больших отклонений. В таких случаях более предпочтительными могут оказаться вторые два критерия, специально контролирующие равномерность представления в выборке классов частных классификаций.

2. Критерии, которые оценивают квадраты отклонений между соответствующими долями (I_1 и I_3), стремятся правильно воспроизвести в выборке прежде всего большие доли (за счет искажения малых). В то же время критерии, оценивающие отклонения по модулю (I_2 и I_4), воспринимают отклонения в больших и малых долях более равномерно.

Отмеченные особенности еще не дают правила выделения критерия, который следует использовать в конкретных условиях. Поэтому, чтобы сделать процедуру построения выборки более строгой, необходимо с большой тщательностью организовать процесс тестирования выборок, сформированных с помощью различных критериев. При этом результат тестирования следует представлять в виде количественных оценок, позволяющих определять наилучшую выборку.

В случае, когда информация для тестовых параметров известна по всем объектам рассматриваемой совокупности, процедура тестирования сводится к отбору лучших

(в смысле введенных критериев) выборок и фиксации той из них, которая получает лучшую оценку по тестовым параметрам.

На практике, однако, информация о тестовых параметрах зачастую бывает известна не для всех объектов совокупности, а лишь для основной их части. Так, при планировании выборки для исследования бюджетов времени в качестве тестовых параметров нами были использованы 10 показателей, содержательно связанных с использованием времени. При этом имевшиеся в нашем распоряжении значения этих параметров были известны лишь для 133 из 157 поселений. Поэтому необходимо было при тестировании и выделении лучшей выборки учитывать неполноту имеющейся информации.

В качестве тестовых параметров для изучения бюджетов времени были выбраны следующие:

средняя продолжительность рабочего дня в течение года (часов на одного работника);

средняя продолжительность рабочего дня зимой (часов на одного работника);

средняя продолжительность рабочего дня летом (часов на одного работника);

обеспеченность вещами длительного пользования, характерными для современного быта (единиц на 100 семей);

обеспеченность вещами длительного пользования, характерными для современного досуга (единиц на 100 семей);

индекс потребления бытовых услуг (суммарное число используемых видов общественного обслуживания в расчете на 100 семей);

индекс общественного обслуживания досуга (число используемых видов учреждений, организующих досуг сельских жителей, в расчете на 100 семей);

доля работников, имеющих, по их словам, свободное время, %;

индекс разнообразия использования свободного времени (в расчете на 100 работников);

средняя частота просмотра телевизионных передач (раз в месяц).

Для оценки качества выборки использовались все четыре критерия (I_1, I_2, I_3 и I_4). Если по какому-либо из критериев наилучшая выборка содержала поселения, для которых информация по тестовым параметрам отсутствует, предлагалась следующая процедура. По каждому из критериев выделялась не одна, а десять наилучших выборок, которые упорядочивались по значениям используемого критерия. Просматривая эти выборки, применительно к каждому критерию находили лучший из тех вариантов, для всех элементов которого определены значения тестовых параметров. В результате были получены искомые выборки для тестирования.

Использованные для такого анализа 40 выборок (по 10 для каждого критерия) пробного обследования представлены в табл. 3. Аналогичная работа была проведена и для построения выборки, предназначенной для основного обследования [4].

Перейдем к описанию этапа тестирования.

Допустим, что в распоряжении исследователя имеются M тестовых параметров, содержательно связанных с изучаемым явлением. Объем совокупности объектов, для которых известны значения тестовых параметров, обозначим через $\widetilde{N}(\widetilde{N} < N$, где N — объем всей совокупности). Пусть x — некоторый из тестовых параметров; $\bar{x}(\widetilde{N})$ — среднее арифметическое значение этого параметра, вычисленное на $\widetilde{N}$ поселениях, для которых известно его значение; $\bar{x}(\widetilde{N}) = 1/\widetilde{N} \sum_{i=1}^{\widetilde{N}} x_i$, где x_i — значения параметра, вычисленные на i-м поселении.

В качестве выборочной оценки среднего служит значение $\hat{\bar{x}}$. Выборочная средняя по типу (при отыскании лучшего варианта выборки используется обобщенная типология по нескольким факторам) рассчитывается по формуле

$$\hat{\bar{x}} = \sum_{\tau=1}^{T} P_\tau(\widetilde{N})\, \hat{\bar{x}}_\tau,$$

где $P_\tau(\widetilde{N})$ — доля τ-го типа в совокупности из $\widetilde{N}$ поселений, $P_\tau(\widetilde{N}) = \widetilde{N}_\tau / \widetilde{N}$ ($\widetilde{N}_\tau$ — число поселений, принадлежащих типу τ); $\hat{\bar{x}}_\tau$ — выборочная средняя по типу τ.

[4] Для основного обследования необходимо было взять 10 лучших выборок из всех возможных комбинаций сельсоветов по 3 (более 5000) по каждому из критериев.

Таблица 3

Варианты выборки для пробного обследования

Вариант выборки	I_1		I_2		I_3		I_4	
	значение критерия	номер сельсовета	значение критерия	номер сельсовета	значение критерия	номер сельсовета	значение критерия	номер сельсовета
1	6,4	**6**	8,3	1	9,1	**34**	10,8	**34**
2	6,6	**34**	8,4	6	10,3	10	12,6	**6**
3	6,6	1	8,6	**34**	11,5	6	12,6	20
4	7,0	20	9,4	20	11,5	20	12,8	10
5	7,5	10	9,7	10	11,5	5	13,3	1
6	7,6	33	9,9	26	12,2	33	14,7	5
7	7,9	**13**	10,0	**13**	12,9	1	15,3	33
8	8,5	26	10,9	33	13,2	**27**	15,9	**27**
9	8,9	2	12,1	25	13,2	**11**	19,0	**11**
10	8,9	**19**	12,2	**3**	13,2	12	19,0	12

Примечания: 1. Значения критериев даны в среднем (%) отклонении на класс. 2. Выделенные полужирным шрифтом выборки в соответствии с указанной процедурой были взяты для тестирования.

Выборочное среднее по классу (при отыскании лучшего варианта используются частные классификации по отдельным факторам) рассчитывается по формуле

$$\hat{\bar{x}} = 1/F \sum_{j=1}^{F} \sum_{\alpha=1}^{k} P_{\alpha j}(\tilde{N})\, \hat{\bar{x}}_{\alpha j},$$

где $P_{\alpha j}(\tilde{N})$ — доля класса α по j-му фактору в совокупности из $\tilde{N}$ поселений, $P_{\alpha j}(\tilde{N}) = \tilde{N}_{\alpha j}/\tilde{N}$ ($\tilde{N}_{\alpha j}$ — число поселений, принадлежащих классу α по j-му фактору);

$\hat{\bar{x}}_{\alpha j}$ — выборочная средняя по классу αj.

Качество такой оценки можно характеризовать двумя величинами:

$$\delta(x) = \frac{|\hat{\bar{x}} - \bar{x}(\tilde{N})|}{\bar{x}(\tilde{N})},$$

$$n_э(x) = \frac{\tilde{N}\sigma^2_{\tilde{N}}(x)}{\sigma^2_{\tilde{N}}(x) + \tilde{N}\delta^2(x) \cdot [\bar{x}(\tilde{N})]^2},$$

где $\sigma^2_{\tilde{N}}(x)$ — дисперсия параметра x, вычисленная на тестовом множестве поселений.

Величина $\delta(x)$ определяет относительное отклонение выборочной оценки $\hat{\bar{x}}$ от среднего значения на тестовом множестве объектов. Значение $n_э(x)$ характеризует объем случайной выборки, которая в среднем дает ошибку в оценке $\bar{x}(\tilde{N})$, равную $|\hat{\bar{x}} - \bar{x}(\tilde{N})|$ (такую выборку естественно называть эквивалентной случайной выборкой). Если объем тестируемой выборки s обозначить через Q_s, то вместо величины $n_э(x)$ удобнее пользоваться относительной величиной $\eta(x) = n_э(x)/Q_s$, которая показывает, во сколько раз эквивалентная случайная выборка должна превышать рассматриваемую стратифицированную выборку для оценки параметра x с той же степенью точности.

В случае нескольких тестовых параметров наиболее важными характеристиками выборки следует признать среднее значение введенных величин

$$\bar{\delta} = 1/M \sum_{m=1}^{M} \delta_m(x),$$

$$\bar{\eta} = 1/M \sum_{m=1}^{M} \eta_m(x),$$

где M — число тестовых параметров.

Эти характеристики для тестируемых выборок приведены в табл. 4. Сопоставление характеристик различных вариантов выборки показывает, что для пробного обследования следует признать лучшим сельсовет № 34 (такой вариант выборки дает в среднем 5%-ную ошибку на тестовом множестве параметров и почти в 8 раз эффективнее случайной выборки). Из троек сельсоветов лучшим оказался набор 25, 32, 34 (ошибка около 6%, объем выборки в 3,4 раза меньше эквивалентной случайной выборки).

В сформированных выборках, несмотря на их крайнюю малочисленность, распределение поселений по типам и классам хорошо соответствует распределению всех 157 поселений (табл. 5). В обеих выборках представлены как урбанизированные, так и неурбанизированные поселения с разным уровнем развития экономической структуры и различными показателями естественного прироста населения. В выборках представлены поселения, принад-

Таблица 4

Характеристики вариантов выборки для пробного и основного обследований

Вариант	Номера сельсоветов	δ	η
Пробное обследование			
1	34	5,0	7,7
2	6	8,6	3,7
3	19	12,6	3,2
4	3	14,8	3,1
5	11	16,1	2,0
Основное обследование			
1	25,32, 34	6,3	3,4
2	6,32, 34	6,7	2,0

лежащие и колхозам, и совхозам (а в выборке для основного обследования и пригородные поселения Новосибирска), с развитой системой транспортных коммуникаций и сравнительно удаленные от нее и т. д. Хотя при построении выборок не учитывались динамические характеристики поселений, но по фактору I (урбанизации) оказались представлены все типы динамики: девять поселений повысили свой статус, шесть остались стабильными, а у одного наблюдалось снижение достигнутого ранее уровня [5].

Таким образом, выборка поселений достаточно малого объема, построенная при условии жесткого ограничения на территориальную связанность, в результате использования строгих критериев оценки и тестирования хорошо соответствует распределению типов поселений с различным уровнем социально-демографического развития во всей совокупности поселений. Она обеспечивает также хорошее оценивание тестовых параметров, содержательно связанных с изучаемым явлением.

Таблица 5

Распределение всей совокупности поселений и построенных выборок по классам факторов, %

Выборка	Доля класса по фактору						
	I		II			III	
	1 кл.	2 кл.	1 кл.	2 кл.	3 кл.	1 кл.	2 кл.
Вся совокупность . .	0,74	0,26	0,19	0,46	0,35	0,65	0,35
Выборка из одного сельсовета	0,80	0,20	0,40	0,40	0,20	0,60	0,40
Выборка из трех сельсоветов	0,73	0,27	0,18	0,46	0,36	0,64	0,36

[5] Типология динамических схем развития сельских поселений построена и подробно описана в работе [7].

ПРОВЕРКА КАЧЕСТВА ВЫБОРКИ ПО ИТОГАМ ИССЛЕДОВАНИЯ

Об эффективности построенной выборки поселений для изучения бюджетов времени можно судить по результатам пробного обследования, проведенного летом 1975 г., в котором было получено 315 бюджетов времени. Анализ показывает, что бюджеты времени населения, проживающего в поселениях разного типа, существенно различаются между собой. Рассмотрим разницу в структуре затрат времени на примере двух типов поселений: неурбанизированных с менее развитой экономической структурой и урбанизированных со среднеразвитой экономической структурой.

В зависимости от типа поселения наблюдается дифференциация объема и особенно структуры общей трудовой нагрузки работников общественного производства (табл. 6). Так, общая трудовая нагрузка в неурбанизированных поселениях с менее развитой экономической структурой производства в среднем на 2 ч больше, чем в урбанизированных со среднеразвитой структурой производства. При этом затраты рабочего времени в общественном производстве в первом типе поселений на 6,2 ч

Таблица 6

Средняя недельная трудовая нагрузка на одного работника общественного производства, ч

Показатель	Неурбанизированные поселения с менее развитой экономической структурой			Урбанизированные поселения со среднеразвитой экономической структурой		
	Всего	Мужчины	Женщины	Всего	Мужчины	Женщины
Общая трудовая нагрузка	83,8	78,8	91,1	82,0	78,3	88,0
В том числе						
Рабочее время в общественном производстве	54,8	62,4	45,0	48,6	54,6	41,4
Время, связанное с работой в общественном производстве	4,4	3,9	5,0	3,1	2,7	3,7
Время труда в личном подсобном хозяйстве	11,4	5,6	19,2	15,5	15,9	15,1
Время труда в домашнем хозяйстве	13,2	6,9	21,9	14,8	5,1	27,8

в неделю больше, чем во втором. В целом структура затрат времени на труд в большей степени дифференцирована для работников-мужчин. В урбанизированных поселениях мужчины работают в общественном производстве примерно на 8 ч в неделю меньше, а в личном подсобном хозяйстве — на 10 ч больше по сравнению с мужчинами неурбанизированных поселений. Для женщин неурбанизированных поселений характерен более продолжительный труд в общественном производстве (на 3,6 ч) и в личном подсобном хозяйстве (на 4,1 ч), чем в урбанизированных со среднеразвитой структурой производства. При этом если в урбанизированных поселениях и мужчины, и женщины затрачивают примерно одинаковое время на ведение личного подсобного хозяйства, то в неурбанизированных затраты времени на труд в личном подсобном хозяйстве у женщин в 3,5 раза больше, чем у мужчин, что свидетельствует о различном характере разделения труда в семьях, проживающих в разных условиях. В результате у мужчин в поселениях обоих типов общая трудовая нагрузка одинакова, когда бóльшие затраты рабочего времени в общественном производстве компенсируются меньшими затратами труда в личном подсобном хозяйстве. У женщин же этой компенсации не происходит, и общая трудовая нагрузка у них в неурбанизированных поселениях на 3 ч больше.

Отраслевая структура занятости населения, дифференцированная по типам поселений, определяет пропорции распределения совокупного фонда рабочего времени между сферой материального производства и непроизводственной сферой. Так, совокупный фонд рабочего времени в урбанизированных поселениях со среднеразвитой экономической структурой в среднем за летний день составляет 694 ч (в расчете на 100 работников общественного производства), из которых 160 ч приходится на отрасли непроизводственной сферы. В то же время в неурбанизированных поселениях с менее развитой экономической структурой совокупные затраты рабочего времени составляют 762 ч, а в непроизводственной сфере из них расходуется лишь 43 ч.

Типы поселений определяют не только затраты рабочего времени, но и величину и структуру использования свободного времени населения (табл. 7). У жителей урбанизированных поселений доля времени, затрачиваемого на

Таблица 7

Структура свободного времени сельского работника в среднем за летний день, %

Затраты свободного времени	Неурбанизированные поселения с менее развитой экономической структурой			Урбанизированные поселения со среднеразвитой экономической структурой		
	Всего	Мужчины	Женщины	Всего	Мужчины	Женщины
Воспитание детей	5,1	1,5	12,7	4,5	1,3	10,1
Учеба, общественная и творческая деятельность	0,6	0,1	1,5	6,5	6,5	6,5
Просмотр кинофильмов, телепередач, чтение, слушание радио	13,8	14,0	13,4	23,6	22,3	26,5
Отдых, развлечения	65,8	63,0	72,0	52,4	53,5	51,3
В том числе						
Бездеятельный отдых	16,5	14,0	20,4	8,4	7,8	10,4
Прием и посещение гостей	35,3	32,8	43,9	27,8	29,8	22,9
Физкультура и спорт	14,7	21,3	0,4	13,0	16,4	5,6

учебу, общественную деятельность и любительские занятия, больше примерно в 10 раз; досуг более насыщен потреблением информации (в 1,7 раза), особенно за счет увеличения доли времени, расходуемого на чтение. В то же время в неурбанизированных поселениях с менее развитой экономической структурой значительно больше времени расходуется на бездеятельный отдых (его доля в структуре свободного времени в 2 раза выше, чем в урбанизированных поселениях), а также на прием и посещение гостей. Это связано как с существенно лучшими возможностями использования свободного времени, которые предоставляет урбанизированный поселок, так и с более высоким уровнем развития потребностей его жителей.

Результаты основного обследования бюджетов времени (как зимнего, так и летнего) подтверждают резкую дифференциацию затрат времени сельского населения, проживающих в населенных пунктах с разным уровнем социального и демографического развития. К сожалению,

Таблица 8

Основные характеристики бюджетов времени по типам поселений, % к типу неурбанизированных поселений с менее развитой экономической структурой

Показатель	Неурбанизированные с экономической структурой		Урбанизированные с экономической структурой	
	среднеразвитой	высокоразвитой	среднеразвитой	высокоразвитой
Трудовая нагрузка . .	99,2	96,2	95,6	90,9
Рабочее время в общественном производстве	96,0	76,1	88,5	83,7
Свободное время . . .	116,9	127,9	123,0	110,7

Примечание. Приведены данные бюджетов времени работников сельского хозяйства в среднем за зимний день.

объем настоящей работы не позволяет привести подробный анализ использования времени по всем типам поселений. Приведем лишь данные о дифференциации основных характеристик бюджетов времени по типам поселений (табл. 8). Она свидетельствует об обоснованности использования социально-демографической типологии поселений для построения выборки при исследовании бюджетов времени.

Предлагаемая методика формирования выборки с использованием типологии социально-экономических объектов, изложенная на примере сельских поселений, носит достаточно общий характер. Задача построения многомерной стратифицированной выборки на основе априорной информации об изучаемом объекте является типичной при подготовке целого ряда экономико-статистических исследований. Поэтому подобная схема может быть рекомендована для планирования различных выборочных обследований. Тот факт, что данная схема наиболее эффективна при малых объемах выборки с дополнительными ограничениями (например, территориальная связанность отобранных объектов, сравнительно часто встречающаяся в экономических задачах), делает ее особенно полезной при планировании углубленных исследований небольшого числа объектов.

ЛИТЕРАТУРА

1. **Воронов Ю. П.** Распознавание образов и выборка в социологических исследованиях.— В кн.: Социология и математика. Новосибирск, 1970, с. 69—77.
2. **Браверман Э. М.** Методы экстремальной группировки и задача выделения существенных факторов.— «Автоматика и телемеханика», 1970, № 1, с. 123—133.
3. **Дорофеюк А. А.** Алгоритмы автоматической классификации (обзор).— «Автоматика и телемеханика», 1971, № 12, с. 78—133.
4. **Браверман Э. М., Киселева Н. Е., Мучник И. Б., Новиков С. Г.** Лингвистический подход к обработке больших массивов информации.— «Автоматика и телемеханика», 1974, № 11, с. 73—88.
5. **Горяченко Е. Е., Мучник И. Б., Новиков С. Г.** Использование машинных методов обработки данных для планирования специализированных выборочных обследований деревни.— В кн.: Проблемы системного изучения деревни. Под ред. Т. И. Заславской, Р. В. Рывкиной. Новосибирск, 1975, с. 276—286.
6. **Браверман Э. М., Литваков Б. М., Мучник И. Б., Новиков С. Г.** Метод стратифицированной выборки в организации сбора эмпирических данных.— «Автоматика и телемеханика», 1975, № 10, стр. 65—78.
7. **Методика** выборочного обследования миграции сельского населения. Новосибирск, 1969, с. 57—76 (ротапринт).
8. **Развитие** сельских поселений (лингвистический метод типологического анализа социальных объектов). Под ред. Т. И. Заславской, И. Б. Мучника. М., «Статистика», 1977. 295 с.

СОДЕРЖАНИЕ

Предисловие 5

Б. Б. Розин. Природа неоднородности статистических совокупностей и методы ее учета при моделировании 7

Б. Б. Розин, В. В. Кулешов, М. Л. Лукацкая. Проблемы использования экономико-математических моделей в технологии плановых расчетов 27

А. В. Беккер, А. Н. Беккер, Б. Б. Розин. Статистический подход к одной задаче управления 48

И. Г. Багиров. Статистическое моделирование управляемых экономических процессов с помощью стохастических уравнений диффузионного типа 56

И. Г. Багиров. Прогнозная оценка вероятностных характеристик потребности в дефицитном ресурсе (на примере систем газоснабжения городов СССР) . . . 81

Л. Я. Шкрабин. Об интерпретации векторной оптимизации в терминах многомерной классификации наблюдений 89

В. Г. Ионин. Подбор пунктов размещения производства на перспективу на основе статистического моделирования 100

В. Г. Ионин. Точность расчета экономических показателей вариантов развития предприятий в задаче перспективного отраслевого планирования 108

Б. Б. Розин, М. А. Ягольницер. Статистические имитационные модели как инструмент анализа неопределенности оптимальных планов 119

Н. М. Журавель, Н. П. Ионина. О критерии минимума потерь информации в имитационном агрегировании . . . 127

Н. М. Журавель, В. Г. Ионин, Н. П. Ионина. Использование процедур оптимизации при классификации объектов и факторов 147

П. Г. Шишацкий. Использование экономико-статистических моделей и агрегирование информации на их основе в системе отрасль — предприятие 162

Н. Е. Ачелашвили. Некоторые вопросы построения системы экономических оценок природных условий в угольной промышленности 177

Г. Т. Берсенева, Г. П. Бондаренко, Н. В. Овчинникова. Учет влияния природных факторов на экономические показатели предприятий 192

Л. А. Сергеева, Г. Л. Чудновский, М. А. Ягольницер. Исследование динамики нефтедобычи 200

Г. Т. Берсенева, Т. Р. Сазонова, Л. А. Сергеева, М. А. Ягольницер. Имитационное моделирование динамики нефтедобычи (на примере Западной Сибири) 207

И. Д. Мандель, А. И. Никитин. Классификация многомерных наблюдений при анализе определений понятия АСУ 217

Г. П. Гвоздева, Е. Е. Горяченко, И. Б. Мучник, Е. Э. Синицын. Использование типологического анализа при планировании выборочных обследований 235

ЭКОНОМИКО-СТАТИСТИЧЕСКИЕ МОДЕЛИ В ПРОГНОЗИРОВАНИИ И ПЛАНИРОВАНИИ ПРОМЫШЛЕННОГО ПРОИЗВОДСТВА

Серия: Статистические методы в экономических исследованиях

Ответственные редакторы
Мэри Львовна Лукацкая
Бенциан Борисович Розин

Редактор *И. Г. Зыкова*
Художник *В. В. Растегаев*
Технический редактор *Н. М. Бурлаченко*
Корректоры *В. В. Борисова, М. В. Ржевцева*

ИБ № 10109

Сдано в набор 9 июня 1977 г. Подписано в печать 21 декабря 1977 г. МН 02104. Формат 84×108¹/₃₂. Бумага машинномелованная. 8 печ. л., 13,4 усл. печ. л., 13,9 уч.-изд. л. Тираж 2100 экз. Заказ № 571. Цена 1 р. 60 к.

Издательство «Наука», Сибирское отделение. 630099, Новосибирск, 99, Советская, 18.

4-я типография издательства «Наука». 630077, Новосибирск, 77, Станиславского, 25.

1