

Экономико-математические модели

А.Ю. Домников, доц., канд. экон. наук,
К.Б. Кожов, доц., канд. техн. наук,
ГОУ УГТУ-УПИ, Екатеринбург,
ОАО ВНИПИЭНЕРГОПРОМ

РАЗРАБОТКА ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ ПОКАЗАТЕЛЕЙ ЭНЕРГЕТИКИ

Одним из основных направлений повышения эффективности развития энергетики является разработка методов и моделей прогнозирования показателей энергетики. Это связано с тем, что данные показатели лежат в основе выработки стратегий развития энергетики и в конечном итоге определяют траектории ее развития. Следовательно, разработанные имитационные эконометрические модели прогнозирования показателей энергетики, позволят рассмотреть основные сценарии возможного перспективного развития ТЭК (в зависимости от экзогенно заданных социально-экономических условий) и заблаговременно наметить и реализовать мероприятия по выходу из сложного экономического состояния изучаемого объекта.

Решение данной задачи осложняется тем, что разработка прогнозов показателей энергетики обусловлена экономической нестабильностью и малыми размерностями обучающих выборок.

Отмеченные обстоятельства требуют соответствующего развития средств и процедур прогнозирования показателей энергетики, предназначенных для обоснования принимаемых решений, преодоления кризисных явлений в отраслях ТЭК, промышленном комплексе и создания условий для устойчивого социально-экономического роста.

Вследствие возросшей роли энергетических факторов в функционировании народнохозяйственного комплекса и обеспечении жизнедеятельности населения одним из важнейших вопросов становится разработка методического аппарата прогнозирования основных показателей энергетики (ПЭ) для решения задач развития региональных энергосистем и определения оптимальных сценариев выхода из кризиса в социально-экономической и промышленной сферах. Многообразие форм взаимодействия ТЭКа с социально-экономическими системами, необходимость исследования множества альтернативных вариантов их развития, а также высокая степень инерционности, присущая народнохозяйственным процессам, создают известные предпосылки для формирования имитационной системы (ИС) управления развитием ТЭК с учетом требований надежности и безопасности.

Таким образом, из-за отмеченного малого объема выборки по отдельно взятому административно - территориальному образованию возникает необходимость в значительном сокращении пространства переменных. Для преодоления указанных препятствий рассмотрено два методических подхода. Первый из них основан на описании зависимостей между переменными с помощью регрессионных уравнений [4, 5]. В этом случае отбор переменных производится с помощью методов кластерного анализа [6, 7, 8]. Эти методы позволяют разбить все пространство показателей на кластеры, включающие показатели, обладающие наиболее тесными связями. Разработанные алгоритмы кластерного анализа базируются на использовании корреляционных матриц. Таким образом, появляется возможность в каждом кластере выбрать один представительный (наиболее информативный) показатель, а остальные показатели исключить из рассмотрения. Этим самым, с одной стороны, достигается желаемое уменьшение числа независимых переменных в регрессионных уравнениях, а с другой - преодолевается мультиколлинеарность пространства переменных, что, как известно, улучшает прогностические возможности моделей.

Другой методический подход основывается на применении специальных методов многомерного статистического анализа, к которым относится факторный анализ [9]. Процедуры факторного анализа также базируются на использовании матриц коэффициентов корреляции между переменными. Прибегая к регулярным методам многомерного анализа, предусматривающим выделение главных факторов или главных компонент, как правило, удается заменить исходную совокупность переменных Y небольшим числом ортогональных факторов f :

$$\begin{aligned} f_1 &= \sum_{i=1}^n \omega_{1i} Y_i \\ &\dots\dots\dots \\ f_k &= \sum_{i=1}^n \omega_{ki} Y_i \end{aligned} \quad (2)$$

После этого становится возможным построение уравнений в факторно-регрессионной форме:

$$\begin{aligned} Y_1 &= \sum_k a_{1k} f_k + b_{11} X_1 + \dots + b_{1m} X_m \\ &\dots\dots\dots \\ Y_n &= \sum_k a_{nk} f_k + b_{n1} X_1 + \dots + b_{nm} X_m \end{aligned} \quad (3)$$

Отмеченное сочетание факторного и регрессионного анализа облегчает построение искомых зависимостей при использовании малых выборок. Основные преимущества такого подхода перед традиционным регрессионным анализом состоят в следующем: 1) значительно повышается корректность регрессионных зависимостей благодаря ортогональности выделяемых факторов; 2) обеспечивается необходимая степень агрегирования с минимальными потерями информации; 3) сужаются доверительные интервалы неопределенности прогнозируемых характеристик (за счет увеличения числа наблюдений, приходящихся на одну переменную).

Основные этапы формирования эконометрических моделей для прогнозирования показателей энергетики приведены на рис. 1.

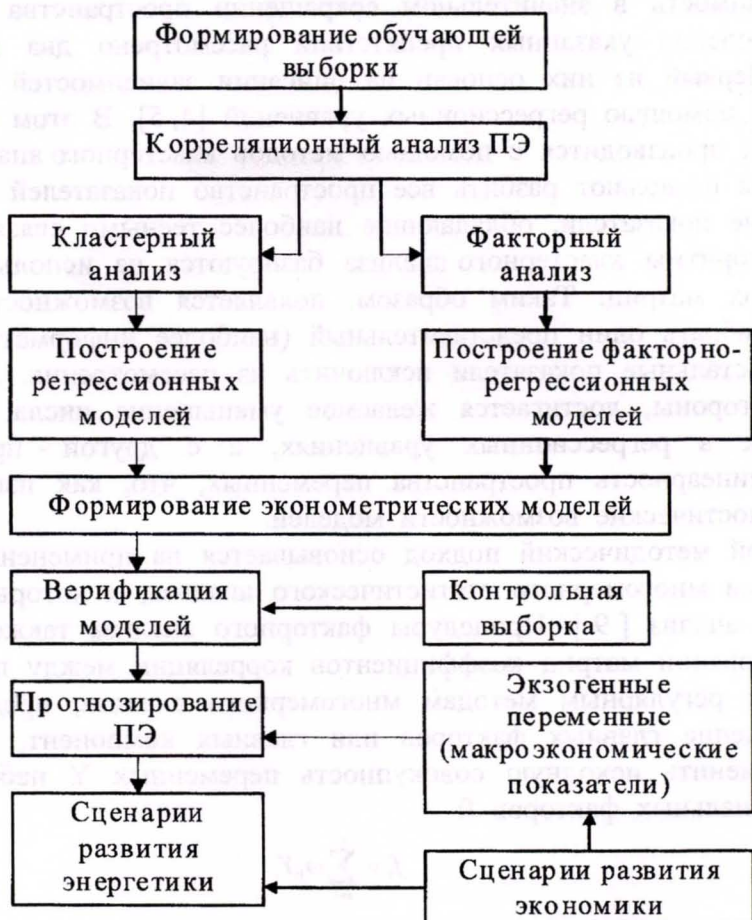


Рис. 1. Схема прогнозирования показателей энергетики на основе эконометрических моделей

Предложенная методика иллюстрируется на примере прогнозирования показателей энергетики Свердловской области. В качестве обучающей выборки для формирования эконометрической модели использовались динамические ряды по 15 народнохозяйственным показателям за промежуток времени с 1990 по 2000 год.

Обучающая выборка для формирования эконометрической модели была сформирована по следующим показателям:

1. Эндогенные:

- 1.1. Коэффициент естественного прироста населения
- 1.2. Доля экономически активного населения
- 1.3. Уровень безработицы
- 1.4. Плотность выбросов в атмосферу
- 1.5. Доля сброса загрязненных вод в общем стоке

- 1.6. Уровень износа основных фондов
- 1.7. Доля производства предметов потребления в общем объеме производства
- 1.8. Душевое потребление ТЭР
- 1.9. Душевое потребление электроэнергии
- 1.10. Душевое потребление тепловой энергии
- 1.11. Душевое потребление моторного топлива
- 1.12. Доля собственных энергоисточников
- 1.13. Доля научных кадров в численности трудящихся
2. Экзогенные:
 - 2.1. Индекс объема производства
 - 2.2. Численность промышленно - производственного персонала

В состав моделируемых показателей вошли показатели, отражающие различные аспекты, - производственные, социально - демографические, экологические, энергетические и др. По условиям моделирования все рассматриваемые показатели были подразделены на две упомянутые группы - эндогенную Y и экзогенную X . В состав экзогенной группы включены 13 показателей, в том числе пять - по топливно-энергетическому комплексу [10].

К экзогенным переменным отнесены индекс объема производства (x_1) и относительная численность промышленно-производственного персонала (x_2).

Как отмечалось выше, при формировании моделей, опирающихся на экономико-статистическую информацию, существенное значение имеет объем выборочной совокупности. Поскольку ретроспективный период, в течение которого проявились кризисные явления в экономике, относительно невелик, исследуемая выборочная совокупность, включающая всего лишь десять наблюдений на 15 переменных, явно относится к классу малых выборок.

Дальнейшие исследования по построению эконометрических моделей выполнены по схеме регрессионного анализа, предусматривающей отбор переменных, наиболее существенно влияющих на прогнозируемые показатели (см. рис. 1).

В соответствии с этой схемой отправным этапом исследований явилось проведение корреляционного анализа, в результате которого получены полная (расширенная) матрица коэффициентов парной корреляции для всего пространства переменных R_p и входящие в него подматрицы корреляционных связей между индикаторами (эндогенными) переменными R_y , между экзогенными переменными R_x , а также между обеими указанными группами переменных R_{yx} :

$$R_p = \begin{bmatrix} R_y & R_{yx} \\ R_{xy} & R_x \end{bmatrix}. \quad (4)$$

Расчетные значения коэффициентов корреляции приводятся в табл. 1.

Матрица парных коэффициентов корреляции переменных
эконометрической модели

	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅	Y ₆	Y ₇	Y ₈	Y ₉	Y ₁₀	Y ₁₁	Y ₁₂	Y ₁₃	X ₁	X ₂
Y ₁	1.0	.52	-.81	.97	.99	.76	.99	.99	.99	.82	-.10	-.99	.62	.98	.92
Y ₂	.52	1.0	-.33	.56	.58	.89	.52	.58	.49	.28	.41	-.57	-.34	.43	.65
Y ₃	-.81	-.33	1.0	-.84	-.81	-.48	-.80	-.85	-.89	-.91	.59	.77	-.63	-.83	-.92
Y ₄	.97	.56	-.84	1.0	.95	.70	.95	.96	.96	.90	-.22	-.97	.52	.91	.93
Y ₅	.99	.58	-.81	.95	1.0	.82	.99	.99	.98	.78	-.06	-.99	.56	.98	.94
Y ₆	.76	.89	-.48	.7	.82	1.0	.77	.80	.72	.37	.41	-.78	.04	.73	.79
Y ₇	.99	.52	-.80	.95	.99	.77	1.0	.99	.98	.78	-.08	-.99	.62	.99	.91
Y ₈	.99	.58	-.85	.96	.99	.80	.99	1.0	.99	.81	-.12	-.98	.52	.98	.96
Y ₉	.99	.49	-.89	.96	.98	.72	.98	.99	1.0	.86	-.24	-.97	.65	.98	.96
Y ₁₀	.81	.28	-.91	.90	.78	.37	.78	.81	.86	1.0	-.60	-.81	.61	.77	.82
Y ₁₁	-.10	.41	.59	-.22	-.06	.41	-.08	-.12	-.24	-.60	1.0	.06	-.52	.14	-.23
Y ₁₂	-.99	-.57	.77	-.97	-.99	-.78	-.99	-.98	-.97	-.81	.06	1.0	-.55	-.96	-.91
Y ₁₃	.62	-.34	.63	.52	.56	.04	.62	.57	.65	.61	-.52	-.55	1.0	.70	.45
X ₁	.98	.43	-.83	.91	.98	.73	.99	.98	.98	.77	-.14	-.96	.70	1.0	.91
X ₂	.92	.65	-.92	.93	.94	.79	.91	.96	.96	.82	-.23	-.91	.45	.91	1.0

Выявление переменных для регрессионных моделей осуществлено на основе алгоритмов кластер - анализа, предусматривающих группировку по тесноте корреляционных связей, при этом относительно сильные корреляционные связи ($r > 0,7$) обнаружены у значительной части переменных ($> 45\%$).

Как известно, результаты классификации переменных могут зависеть от задаваемой величины порогового значения коэффициентов корреляции - R_{Π} . В выполненных расчетах при $R_{\Pi} = 0,9$ получено шесть кластеров переменных (табл. 2). Наиболее мощным оказался первый кластер, в который вошли восемь переменных; остальные кластеры включают по 1-2 переменных. Уменьшение величины порога до $R_{\Pi} = 0,7$ практически мало сказывается на результатах группировки.

В общем случае при выборе вида регрессионных моделей в качестве аргументов целесообразно использовать переменные из разных кластеров, что позволяет избежать мультиколлинеарности пространства переменных. С этой точки зрения наиболее информативными оказываются переменные, попадающие в центры кластеров. Однако в рассматриваемом примере данный принцип в полной мере не удалось реализовать. Из-за отмеченной малости выборочной совокупности пришлось ограничиться включением в регрессионные модели лишь двух переменных, одна из которых представляет соответственно экзогенные, а другая - эндогенные параметры, попадающие в общий кластер с зависимой переменной. В то же время внутри самого кластера отбор производился по тесноте корреляционных связей.

Таблица 2

Результаты кластер-анализа переменных эконометрической модели

Наименование	Обозначение	Номер кластера		
		R _П = 0,9	R _П = 0,8	R _П = 0,7
1. Коэффициент естественного прироста населения	y ₁	1	1	1
2. Доля экономически активного населения	y ₂	3*)	2*)	3*)
3. Уровень безработицы	y ₃	2*)	1	1
4. Плотность выбросов в атмосферу	y ₄	1	1	1
5. Доля сброса загрязненных вод в общем стоке	y ₅	1	1	1
6. Уровень износа основных фондов	y ₆	4*)	2	1
7. Доля производства предметов потребления в общем объеме производства	y ₇	1	1	1
8. Душевое потребление ТЭР	y ₈	1	1	1
9. Душевое потребление электроэнергии	y ₉	1*)	1*)	1*)
10. Душевое потребление теплоэнергии	y ₁₀	2	1	1
11. Душевое потребление моторного топлива	y ₁₁	5*)	3*)	2*)
12. Доля собственных энергоисточников	y ₁₂	1	1	1
13. Доля научных кадров в численности трудящихся	y ₁₃	6*)	4*)	4*)

Примечание: Переменные, соответствующие центрам кластеров обозначены *).

Результатом регрессионного анализа явилось формирование совокупности частных линейных моделей:

$$y_1 = -(30.98 \pm 2.70) + (35.35 \pm 6.03)y_7 - (3.18 \pm 3.32)x_1; \quad (5)$$

$$y_2 = (0.531 \pm 0.160) + (0.507 \pm 0.259)y_6 - (0.04 \pm 0.144)x_2; \quad (6)$$

$$y_3 = (2.713 \pm 1.916) - (2.526 \pm 2.021)y_{10} - (0.192 \pm 0.135)x_2; \quad (7)$$

$$y_4 = -(1.617 \pm 0.722) + (3.404 \pm 1.286)y_9 - (0.796 \pm 0.578)x_1; \quad (8)$$

$$y_5 = (0.484 \pm 0.072) + (0.501 \pm 0.118)y_8 + (0.013 \pm 0.046)x_1; \quad (9)$$

$$y_6 = -(0.852 \pm 1.049) + (2.149 \pm 1.133)y_5 - (0.313 \pm 0.277)x_1; \quad (10)$$

$$y_7 = -(1.082 \pm 0.541) + (1.939 \pm 0.687)y_5 + (0.048 \pm 0.143)x_1; \quad (11)$$

$$y_8 = (1.433 \pm 0.441) - (0.611 \pm 0.325)y_{12} + (0.171 \pm 0.116)x_1; \quad (12)$$

$$y_9 = (0.04 \pm 0.278) + (0.857 \pm 0.458)y_8 + (0.115 \pm 0.178)x_1; \quad (13)$$

$$y_{10} = (0.982 \pm 0.033) + (0.026 \pm 0.021)y_4 + (0.006 \pm 0.053)x_2; \quad (14)$$

$$y_{11} = (2.978 \pm 1.238) - (2.07 \pm 1.305)y_{10} + (0.089 \pm 0.087)x_2; \quad (15)$$

$$y_{12} = (3.087 \pm 0.634) - (2.2 \pm 0.806)y_5 + (0.106 \pm 0.168)x_1; \quad (16)$$

$$y_{13} = -(11.275 \pm 49.816) + (11.871 \pm 50.522)y_{10} + (0.765 \pm 1.046)x_1. \quad (17)$$

Построенные регрессионные модели сами по себе обладают относительно высокой аппроксимирующей способностью, о чем в первом приближении можно судить по величине коэффициентов множественной корреляции R . Для большинства показателей значения этих коэффициентов лежат в пределах 0,85 — 0,99. Однако на данном этапе исследований по ряду показателей не удалось достичь необходимой статистической значимости моделируемых величин (по критерию Фишера), а также оценок параметров уравнений регрессии (по T -критерию). По данной причине первоначальная попытка включения в эконометрическую модель всех уравнений (5-17) привела к несовместности системы. Вследствие этого возникла необходимость отсева уравнений, которые в наибольшей мере ответственны за отмеченную несовместность. Таковыми оказались уравнения (5), (7) и (17), описывающие соответственно показатели естественного прироста населения (y_1), уровень безработицы (y_3) и долю научных кадров в численности трудящихся (y_{13}). По-видимому, прогнозирование данных показателей должно осуществляться с привлечением других инструментальных средств, выходящих за рамки рассматриваемой модификации эконометрической модели. Нельзя также полностью исключить того обстоятельства, что по некоторым из показателей негативный результат обусловлен недостаточной достоверностью исходной информации (в частности, это может касаться показателей y_3 и y_{13}).

Как и следовало ожидать, погрешности общей эконометрической модели получились несколько выше по сравнению с частными регрессионными моделями. Так, при верификации оценок индикаторов энергетического блока обнаружилось, что при использовании общей модели погрешности лежат в диапазоне 0,1 - 2,8 %, в то время как по частным моделям они не превышают 1,5 %. Аналогичная картина выявлена при моделировании остальных народнохозяйственных блоков. Однако, при всем этом, можно предположить, что результаты, получаемые по эконометрическим моделям, обладают более высокой прогностической ценностью. Последнее, по всей вероятности, обуславливается отмеченным учетом взаимодействия между показателями, а также сокращением числа независимых переменных, прогнозные оценки которых должны быть заданы экзогенно. Очевидно, такого рода оценкам свойственна неопределенность, возрастающая по мере расширения горизонта прогнозирования. В этих условиях эффективность эконометрического моделирования может быть существенно повышена при прогнозировании показателей исследуемых сценариев социально - экономического развития региональных народнохозяйственных систем.

Выводы

1. Сформирована имитационная эконометрическая модель для прогнозирования показателей энергетики с учетом влияния народнохозяйственных связей.
2. Для преодоления препятствий, обусловленных малым объемом статистических выборок, предложены два методических подхода, первый

из которых предусматривает использование регрессионных моделей в сочетании с кластерным анализом, а второй основывается на применении специальных методов многомерного статистического анализа.

3. Предложенная методика апробирована при проведении исследований по прогнозированию взаимоувязанной системы энергетических и экономических показателей (на примере Свердловской области).

Библиографический список

1. Н.И. Воропай, С.М. Клименко, Л.Д. Криворучий, И.И. Пяткова, С.М. Сендеров, Г.Б. Славин, М.Б.Чельцов. О сущности и основных проблемах энергетической безопасности России // Изв. РАН. Энергетика. 1996. № 3. С. 38-49.
2. Имитационное моделирование систем энергетики / Д.А. Арзамасцев, В.Р. Елохин, Л.Д. Криворучий, Л.И. Мардер, А.Л. Мызин. Иркутск, Свердловск: Изд. СЭИ, 1988. 156 с.
3. Джонсон Дж. Эконометрические методы. М.: Статистика, 1980. 448 с.
4. Четыркин Е.М. Статистические методы прогнозирования. М.: Статистика, 1975. 184 с.
5. Клейнен Дж. Статистические методы в имитационном моделировании. М.: Статистика, 1978. Вып. 1. 224 с.; Вып. 2. 336 с.
6. Айвазян С.А., Бежаева З.И., Староверов О.В. Классификация многомерных наблюдений. М.: Статистика, 1974. 240 с.
7. Классификация и кластер / Под ред. Дж. Вэн Райзина. М.: Мир, 1980. 389 с.
8. Богатырев Л.Л. Современные методы исследований электроэнергетических систем. Екатеринбург: изд. УПИ, 1991. 103 с.
9. Харман Г. Современный факторный анализ: Пер. с англ. М.: Статистика, 1972. 487 с.
10. Оценка угроз экономической безопасности Урала (препринт). Екатеринбург: Институт экономики РАН, 1995. 126 с.