

МЕТОДЫ РЕГРЕССИОННОГО АНАЛИЗА В ИССЛЕДОВАНИЯХ ТЕПЛОПОТРЕБЛЕНИЯ В РОССИИ

В. А. Стенников, Т. В. Добровольская

Институт систем энергетики им. Л. А. Мелентьева Сибирского отделения
Российской академии наук, Иркутск, Россия

Суровые климатические условия России обуславливают высокую социальную и экономическую значимость теплоснабжения потребителей. В соответствии с этим качественный мониторинг состояния и планирование развития теплоснабжения страны и ее регионов являются стратегическим вектором научных исследований и процесса выработки и принятия решений по организации эффективного энергоснабжения потребителей. В статье представлены методы регрессионного анализа применительно к исследованиям потребления тепловой энергии населением, промышленными и социальными объектами. Авторами представлены результаты исследования зависимости изменения теплопотребления от ряда социальных, технических и экономических параметров. По результатам проведенных исследований описана ретроспективная динамика изменения теплопотребления с помощью одно- и многопараметрических линейных и нелинейных регрессионных уравнений. Полученные регрессионные модели апробированы при прогнозировании уровня теплопотребления на среднесрочную (краткосрочную) перспективу.

Ключевые слова: теплопотребление, регрессия, ретроспектива, анализ, прогнозирование.

METHODS OF REGRESSIVE ANALYSIS IN RESEARCHING THERMAL-CONSUMPTION IN RUSSIA

Valery A. Stennikov, Tatyana V. Dobrovolskaya

Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences,
Irkutsk, Russia

Severe climate in Russia stipulates high social and economic significance of heat supply for customers. Due to the fact, high-quality monitoring of its condition and plans for developing heat supply in the country and its regions form the strategic line of investigations and decision-making dealing with organization of efficient heat supply of customers. The article proposes methods of regressive analysis applied to research in the field of thermal power consumption by the population, industrial and social entities. The authors demonstrate findings of the research concerning the dependence of changes in thermal consumption on a number of social, technical and economic parameters. Retrospective dynamics of thermal consumption changes was described with the help of one- and multi-parametric linear and non-linear regressive equations. The obtained regressive models were tested in forecasting the level of thermal consumption for the medium and short-term period.

Keywords: thermal consumption, regression, retrospective, analysis, forecast.

Изменение экономической ситуации в стране и в теплоснабжении в частности, рост стоимости топлива и тепловой энергии, возрастающие требования потребителей к комфорту во многом определяют целесообразность более внимательного отношения к рациональному и

эффективному использованию энергетических и топливных ресурсов. В связи с этим возникает потребность в детальном и всестороннем анализе ситуации с целью выявления механизмов взаимодействия различных индикативных показателей и формирующихся тенденций развития.

На основе статистических данных за период с 1990 по 2015 г. авторами выполнен сводный баланс производства и потребления тепловой энергии в целом по стране. За рассматриваемый период наблюдается значительное сокращение величины общего теплопотребления более чем на 39,5%. При этом сокращение потребления промышленными предприятиями за рассматриваемый период составило около 52,0%, что связано главным образом с сокращением объемов промышленных производств, с территориальным перераспределением в результате деления СССР, а также с внедрением менее энергоемких технологий в различных отраслях добычи, обработки и переработки ископаемых ресурсов. Потребление тепловой энергии населением сохранилось примерно на прежнем уровне, сокращение за рассматриваемый период составило 2,9%, при этом площадь жилищного фонда в целом по стране увеличилась на 38,5%. Применение при строительстве современных материалов с низкой теплопроводностью, а также разработка и внедрение энергосберегающих мероприятий привели к сокращению удельного теплопотребления населением и сферой социально-бытовых услуг [1].

Анализ методических подходов к мониторингу и прогнозированию энергопотребления [2–4] показал, что, как правило, они ориентированы на использование в расчетах энергопотребления удельных показателей. Однако усредненные показатели потребления тепловой энергии в жилых зданиях определяются расчетным путем и не отображают реальной картины, так как не учитывают климатические особенности территорий, фактическое состояние ограждающих конструкций зданий, количественные и качественные характеристики проведенных энергосберегающих мероприятий. В промышленности основой прогноза служат показатели энергоемкости различных отраслей, принятые на основе анализа отечественного и зарубежного опыта. В этом случае довольно сложно произвести аналогии в связи со значитель-

ными различиями в экономическом развитии и технологическом потенциале регионов и стран (в случае ориентации на зарубежные показатели).

Многие из этих погрешностей могут быть устранены в результате предметного изучения ретроспективных данных и применения их для интерполяции прогнозных оценок на соответствующем отрезке времени. Достаточно продуктивным может быть применение подхода, основанного на ретроспективном анализе, который широко используется в мировой практике.

Выполненные исследования по применению различных методических подходов для описания динамики изменения теплопотребления на ретроспективном отрезке времени показали обоснованность использования многопараметрических регрессионных моделей теплопотребления. Эти исследования основывались на сравнительном анализе различных модификаций линейной и нелинейной регрессионных моделей теплопотребления. В качестве объекта исследования рассматривался верхний иерархический территориальный уровень с характеристическим показателем, соответствующим суммарному потреблению тепловой энергии в стране.

Анализ динамики изменения теплопотребления за рассматриваемый ретроспективный отрезок времени позволяет сделать вывод о существовании некой функциональной зависимости изменения теплопотребления от момента времени. При описании ретроспективы теплопотребления с помощью линейного уравнения функциональная зависимость имеет монотонно убывающий характер. Графическая интерпретация динамики теплопотребления в зависимости от времени за рассматриваемый ретроспективный период, полученная по линейному регрессионному уравнению, представлена на рис. 1.

Самым распространенным методом оценки параметров уравнения линейной многопараметрической регрессии является метод наименьших квадратов (МНК). Суть метода состоит в минимизации суммы

квадратов отклонений наблюдаемых значений зависимой переменной y_i от ее значений, получаемых по уравнению регрессии \hat{y}_i . На основе выдвинутого требования стандартная ошибка ε_i должна быть минимальна, что формально может быть записано в следующем виде:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (\varepsilon_i)^2 \rightarrow \min. \quad (1)$$

В соответствии с МНК были получены значения коэффициентов регрессионного уравнения зависимости теплопотребления от времени. Однопараметрическое линейное уравнение регрессии для описания теплопотребления на ретроспективном отрезке времени может быть представлено в следующем виде:

$$Q_i^{\text{mod}(1)} = -31,5 \cdot T_i + 2190,2 + \varepsilon_i^{(1)}, \quad (2)$$

где $Q_i^{\text{mod}(1)}$ – значение теплопотребления в i -й год, определенное с помощью линейного однопараметрического регрессионного уравнения;

T_i – порядковый номер года;

$\varepsilon_i^{(1)}$ – отклонение ретроспективных значений теплопотребления от значений теп-

лопотребления, полученных по уравнению регрессии.

Проверка общего качества уравнения регрессии осуществлялась на основе оценки коэффициента детерминации R^2 . Значение данного коэффициента состоит в доле общего разброса значений зависимой переменной, объясненной уравнением регрессии. Определение коэффициента детерминации производится по следующей формуле:

$$R^2 = 1 - \frac{\sum \varepsilon_i^2}{\sum (y_i - \bar{y})^2}. \quad (3)$$

Для данного линейного регрессионного уравнения теплопотребления коэффициент детерминации составит 0,84. Однако его использование для прогнозирования теплопотребления невозможно, так как при любом изменении социальной или экономической обстановки в стране теплопотребление будет сокращаться, а уравнение не включает в себя социальные и экономические показатели.

Аналогичная ситуация будет иметь место в случае описания изменения теплопотребления с помощью полинома второго порядка (рис. 1).

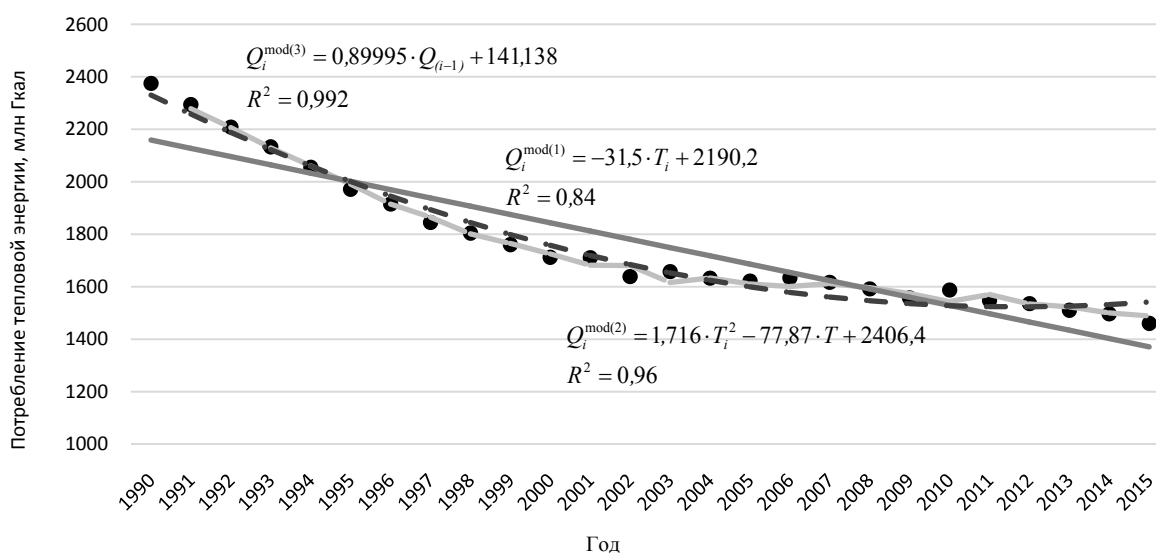


Рис. 1. Динамика изменения потребления тепловой энергии в России в период с 1990 по 2015 г. и значения теплопотребления, полученные с помощью регрессионных моделей:

- наблюдаемые данные; — авторегрессионное уравнение; — линейное однопараметрическое регрессионное уравнение; — полиномиальное уравнение второго порядка

Полиномиальное уравнение второго порядка, описывающее изменение теплопотребления на рассматриваемом отрезке времени, будет иметь следующий вид:

$$Q_i^{\text{mod}(2)} = 1,61 \cdot T_i^2 - 74,9 \cdot T + 2390,6 + \varepsilon_i^{(2)}, \quad (4)$$

где $Q_i^{\text{mod}(2)}$ – значение теплопотребления в i -й год, определенное с помощью полиномиального уравнения второго порядка;

T_i – порядковый номер года;

$\varepsilon_i^{(2)}$ – отклонение выборочных значений объясняющей переменной от значений, полученных с помощью полиномиального уравнения второго порядка.

Коэффициент детерминации, описывающий качество уравнения, в данном случае составляет 0,96. Использование полиномиального уравнения второго порядка для мониторинга изменения теплопотребления на рассматриваемом отрезке времени показывает хорошую сходимость с наблюдаемыми данными. Однако при прогнозировании теплопотребления его использование не обеспечивает адекватное отражение изменения теплопотребления в зависимости от состояния социальной и экономической ситуации.

Следующим этапом в исследовании точностного описания изменения теплопотребления является выявление авторегрессии. Очевидно, что при прогнозировании рассматриваемых показателей следует учитывать наличие устойчивой связи исследуемого признака с его значениями в прошлом и будущем. Взаимосвязь между последовательными элементами динамического ряда представляет частный случай корреляционной зависимости или автокорреляции. Для определения порядка уравнения авторегрессии были определены коэффициенты корреляции для различных периодов запаздывания. Для уравнений первого, второго и третьего порядка имеет место наиболее высокое значение коэффициента автокорреляции (0,991–0,996). Уравнение авторегрессии первого порядка будет иметь следующий вид:

$$Q_i^{\text{mod}(3)} = 0,89995 \cdot Q_{(i-1)} + 141,138 + \varepsilon_i^{(3)}, \quad (5)$$

где $Q_{(i-1)}$ – значение теплопотребления, наблюдаемое в предыдущий момент времени (год);

$\varepsilon_i^{(3)}$ – отклонение выборочных значений объясняющей переменной от значений, полученных с помощью авторегрессионного уравнения.

Коэффициент детерминации данного линейного регрессионного уравнения составит 0,992, что означает высокую сходимость модельных данных с наблюдаемыми (см. рис. 1). Однако использование данного уравнения для прогнозирования теплопотребления на краткосрочную и среднесрочную перспективу нецелесообразно ввиду отсутствия связи с изменяющейся экономической и социальной обстановкой (в уравнении отсутствуют показатели, отражающие состояние экономической и социальной сферы).

Более точное описание динамики изменения теплопотребления может быть представлено с помощью многопараметрической линейной регрессионной модели, которая достаточно хорошо отражает зависимость теплопотребления от социальных и экономических параметров и в полной мере соответствует действительным изменениям теплопотребления. Важным этапом формирования такой модели, как известно, является выбор параметров для многопараметрической регрессионной модели. В связи с характерной для регрессионных моделей особенностью, обусловленной необходимостью иметь в наличии перспективные оценки регрессоров уравнения (объясняющих переменных), выбор параметров в качестве объясняющих переменных в регрессионном уравнении значительно усложняется. Наиболее авторитетными прогнозами социальных и экономических показателей можно считать прогнозы социально-экономического развития страны, разработанные Министерством экономического развития Российской Федерации, которые принимались за основу в проводимых исследованиях.

В связи с этим первым из возможных регрессоров уравнения может быть принята численность населения. Население является одним из основных потребителей тепловой энергии (около 28–31% от общего теплопотребления в стране), направляемой на цели отопления и горячего водоснабжения. В последнее время намечается тенденция увеличения доли населения в общей структуре теплопотребления, что, вероятно, еще в большей степени усилит влияние данного регрессора на модель.

В качестве экономических показателей, оказывающих непосредственное влияние на уровень теплопотребления, выбраны инвестиции в электроэнергетику и валовой внутренний продукт (ВВП). Инвестиции в электроэнергетику определяют объемы планируемой модернизации и финансирования строительства новых объектов электроэнергетики, в том числе и тепловых электростанций. ВВП выбран как основной макроэкономический показатель, отражающий общее состояние экономики страны.

В качестве показателей, представляющих общее состояние энергетики, для рассмотрения были приняты установленная электрическая мощность электростанций и потребление электроэнергии. Выбор потребления электроэнергии в качестве возможного регрессора в уравнении для определения теплопотребления объясняется изменением уровня электропотребления в результате изменения теплопотребления в жилищно-коммунальном секторе. В случае снижения или прекращения подачи тепловой энергии в результате аварийной ситуации на источнике тепловой энергии или тепловых сетях потребление электрической энергии в жилых домах может возрасти за счет включения электрообогревателей и нагревателей воды. Подобные меры могут быть предприняты в результате непредвиденного похолодания. Однако такая доля изменения электропотребления довольно мала, и в связи с рассмотрением уровней электро- и теплопотребления в годовом разрезе сезонные или разовые из-

менения энергопотребления из-за их малой величины нивелируются.

С учетом вышеизложенного для исследования выбраны следующие показатели: установленная электрическая мощность электростанций (N , млн кВт), потребление электрической энергии (W , млрд кВт·ч), потребление тепловой энергии (Q , млн Гкал), численность населения (P , млн чел.), инвестиции в электроэнергетику (с учетом централизованного теплоснабжения) (I , млрд руб.), внутренний валовой продукт (ВВП) (GDP , млрд руб.), временной отрезок (T , год). Рассматриваемые показатели представлены в виде временных рядов с 1990 по 2015 г.

Первоначальный анализ системы рассматриваемых показателей заключался в выявлении тесноты связи между показателями в рамках представленной выборки. Кроме того, корреляционный анализ позволил определить необходимость включения показателей в многопараметрическую регрессионную модель. Математической мерой корреляции двух величин служит коэффициент корреляции. В данном случае показатели представлены в виде некоторых выборок абсолютных величин, поэтому они были интерпретированы с помощью линейного коэффициента корреляции (коэффициента корреляции Пирсона) [1].

Кроме того, при проведении корреляционного анализа выборки данных должны выполняться следующие условия:

- применение корреляционного анализа возможно только при наличии достаточного количества наблюдений для исследования, в 5–6 раз превышающего количество параметров;
- необходимо, чтобы совокупность значений всех факторных и результативных признаков подчинялась многомерно нормальному распределению (определено графически);
- исходная совокупность значений должна быть качественно однородна (рассматриваемые показатели представлены в

виде временных рядов, измерения выполнены через равный промежуток времени).

Учитывалось и то, что сам по себе факт корреляционной зависимости не дает оснований утверждать, что одна из переменных предшествует или является причиной изменения, либо то, что переменные вообще имеют причинно-следственную связь, а не наблюдается действие третьего фактора. В результате были получены все

необходимые основания, чтобы продолжить корреляционный анализ системы показателей путем определения парного коэффициента корреляции. Полученная корреляционная матрица представлена в виде таблицы. Элементы этой матрицы являются коэффициентами корреляции между всеми переменными данной совокупности. Матрица коэффициентов корреляции является симметричной.

Матрица коэффициентов корреляции рассматриваемых показателей в рамках выборки с 1990 по 2015 г.

Переменные	<i>T</i>	<i>N</i>	<i>W</i>	<i>Q</i>	<i>P</i>	<i>I</i>	<i>GDP</i>
<i>T</i>	1						
<i>N</i>	0,8559	1					
<i>W</i>	0,4079	0,6007	1				
<i>Q</i>	-0,928	-0,6522	-0,0667	1			
<i>P</i>	-0,858	-0,5382	-0,3786	0,8358	1		
<i>I</i>	0,8818	0,8946	0,4088	-0,7454	-0,606	1	
<i>GDP</i>	0,987	0,8698	0,4524	-0,8952	-0,8542	0,8905	1

Принято говорить о существовании достаточно сильной зависимости между показателями при значении коэффициента корреляции больше 0,75. Если значение коэффициента корреляции находится в пределах 0,4–0,7, то имеет место умеренная зависимость между рассматриваемыми показателями, если коэффициент корреляции меньше 0,4, то между показателями наблюдается слабая зависимость. В связи с тем что объектом наших исследований является теплопотребление, можно сделать выводы о существовании зависимости теплопотребления от численности населения, ВВП и времени (значения временного периода), так как парные коэффициенты корреляции имеют значения, превышающие 0,8 (таблица).

Наряду с вышеизложенными данными исследования были проведены расчеты коэффициентов автокорреляции рассматриваемых показателей. Для уменьшения автокорреляции, как известно, применяют различные методы, цель которых – исключение основной тенденции развития в первоначальных данных – линейных трендов. Введение в уравнение множественной регрессии показателя времени *T* в качестве независимой переменной позво-

лит учесть отмеченную особенность [4]. Линейные однопараметрические регрессионные модели относительно выбранных показателей были представлены в работе [1]. Рассмотрим более подробно зависимость теплопотребления от остальных показателей, используя линейную многопараметрическую регрессию.

Теоретическое линейное уравнение регрессии в общем виде может быть записано в виде выражения:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon, \quad (6)$$

где $X = (X_1, X_2, \dots, X_m)$ – вектор независимых (объясняющих) переменных;

β – вектор параметров (подлежащих определению);

ε – случайная ошибка (отклонения);

Y – зависимая (объясняемая) переменная;

m – количество объясняющих переменных;

n – количество наблюдений.

Для того чтобы однозначно решить задачу поиска параметров $\beta_0, \beta_1, \beta_2, \dots, \beta_m$ (т. е. найти некоторый вектор β), должно выполняться неравенство $n \geq m + 1$. При условии невыполнения этого неравенства существует бесконечно большое число раз-

личных векторов параметров, при которых линейная зависимость между Y и X будет абсолютно точно соответствовать имеющимся наблюдениям. Следовательно, если число наблюдений больше минимально необходимого, т. е. когда $n > m + 1$, то уже нельзя подобрать линейную форму, в точности удовлетворяющую всем наблюдениям, и возникает необходимость оптимизации, т. е. оценивания параметров $\beta_0, \beta_1, \beta_2, \dots, \beta_m$, при которых формула дает наилучшее приближение для имеющихся наблюдений. При оценивании множественной линейной регрессии для обеспечения статистической надежности требуется, чтобы число наблюдений по крайней мере в 3 раза превосходило число оцениваемых параметров [1]. В нашем случае $n = 26$, $m = 6$, в связи с чем дальнейшие расчеты были направлены на поиск необходимых коэффициентов регрессии. Уравнение (6) в случае рассматриваемой линейной многопараметрической регрессии теплопотребления принимает следующий вид:

$$y_i = b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{3i} + b_4x_{4i} + b_5x_{5i} + b_6x_{6i} + e_i, \quad (7)$$

где y_i – значения объясняемой переменной (теплопотребления), $i = 1, 2, \dots, n$;

$x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}, x_{6i}$ – значения объясняющих переменных, $i = 1, 2, \dots, n$;

$b_0, b_1, b_2, b_3, b_4, b_5, b_6$ – коэффициенты рассматриваемой линейной многопараметрической регрессии;

e_i – отклонение выборочных значений объясняющей переменной от значений, полученных по уравнению регрессии.

В соответствии с МНК вычисляются значения коэффициентов уравнения линейной многопараметрической регрессии для теплопотребления. Коэффициенты рассматриваемой регрессии, обозначенные через вектор B , имеют следующие значения:

$$B = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix} = \begin{bmatrix} 3093,673 \\ -50,312 \\ 10,467 \\ 0,775 \\ -24,46 \\ 0,199 \\ -0,00476 \end{bmatrix}.$$

Подставив полученные значения коэффициентов регрессии в уравнение (7), получаем уравнение многопараметрической регрессии для описания изменения теплопотребления:

$$Q_i^{\text{mod}(4)} = 3093,67 - 50,31 \cdot T_i + 10,47 \cdot N_i + 0,775 \cdot W_i - 24,46 \cdot P_i + 0,199 \cdot I_i - 0,048 \cdot GDP_i + \varepsilon_i^{(4)}. \quad (8)$$

Наилучшие результаты при использовании метода наименьших квадратов получаются, когда выполняется ряд предпосылок относительно случайного отклонения. В рамках статистического анализа стандартная ошибка регрессии определяется по формуле

$$S^2 = \frac{\sum e_i^2}{n - m - 1}. \quad (9)$$

В соответствии с полученным уравнением регрессии стандартная ошибка (9) составляет $S = \sqrt{S^2} = 21,137$. Дальнейшие расчеты были связаны с определением стандартных ошибок коэффициентов регрессии. В общем виде формулу для их расчета можно записать в следующем виде:

$$S_{b_j}^2 = S^2 z'_{jj} = \frac{\sum e_i^2}{n - m - 1} z'_{jj}, \quad j = 1, 2, \dots, m, \quad (10)$$

где z'_{jj} – j -й диагональный элемент матрицы $Z^{-1} = (X^T X)^{-1}$.

Расчеты стандартных ошибок коэффициентов регрессии (10) для полученной многопараметрической линейной регрессионной модели теплопотребления выполнялись для оценки статистической значимости коэффициентов регрессии. По результатам расчетов были получены следующие значения стандартных ошибок:

$$S_{b_0}^2 = 1214,1; S_{b_1}^2 = 5,1; S_{b_2}^2 = 3,27; S_{b_3}^2 = 0,113;$$

$$S_{b_4}^2 = 11,346; S_{b_5}^2 = 0,045; S_{b_6}^2 = 0,0023.$$

Для определения интервальных оценок коэффициентов регрессии β_j была построена t -статистика, имеющая распределение Стьюдента с числом степеней свободы $\nu = n - m - 1$:

$$t = \frac{b_j}{S_{b_j}}. \quad (11)$$

Полученная t -статистика (11) для коэффициентов рассматриваемой линейной многопараметрической регрессии будет иметь следующие значения:

$$t_{b_0} = 2,54; t_{b_1} = -9,858; t_{b_2} = 3,196; t_{b_3} = 6,85; \\ t_{b_4} = -2,15; t_{b_5} = 4,38; t_{b_6} = -2,05.$$

В дальнейшем эти значения сравнивались с их табличными уровнями t -статистики согласно распределению Стьюдента. Принимаем заведомо высокий уровень значимости ($\alpha = 0,05$), тогда при $|t| > t_{\alpha/2, m-n-1}$ коэффициент уравнения регрессии b_j считается статистически значимым. В соответствии с распределением Стьюдента получены следующие соотношения величин:

$$\left\{ \begin{array}{l} |t_{b_0}| = 2,54 > t_{0,05;18} = 1,734; \\ |t_{b_1}| = 9,858 > t_{0,05;18} = 1,734; \\ |t_{b_2}| = 3,126 > t_{0,05;18} = 1,734; \\ |t_{b_3}| = 6,85 > t_{0,05;18} = 1,734; \\ |t_{b_4}| = 2,15 > t_{0,05;18} = 1,734; \\ |t_{b_5}| = 4,38 > t_{0,05;18} = 1,734; \\ |t_{b_6}| = 2,05 > t_{0,005;18} = 1,734. \end{array} \right.$$

Выполненные расчеты позволяют сделать выводы о статистической значимости коэффициентов рассматриваемой линейной многопараметрической регрессии.

Проверка общего качества уравнения регрессии осуществлялась путем вычисления коэффициента детерминации. В соответствии с (3) для рассматриваемого уравнения множественной регрессии коэффициент детерминации равен $R^2 = 0,994$.

При выполнении проверки статистической значимости коэффициента детерминации была выдвинута гипотеза об одновременном равенстве нулю всех коэффициентов регрессии, т. е. ни одна из объясняющих переменных, включенных в регрессию, не оказывает влияния на зависимую переменную (в нашем случае – теплопотребление). Проверка данной гипотезы осуществлялась путем дисперсионного анализа сравнения объясненной и оста-

точной дисперсий. При этом рассматривались две гипотезы:

H_0 : объясненная дисперсия равна остаточной дисперсии;

H_1 : объясненная дисперсия больше остаточной дисперсии.

Для подтверждения или опровержения основной гипотезы была построена F -статистика в соответствии со следующим выражением:

$$F = \frac{\sum(\hat{y}_i - \bar{y})^2 / m}{\sum(y_i - \hat{y}_i)^2 / (n - m - 1)}. \quad (12)$$

Согласно (12) и полученным уравнениям регрессии вычисляется значение F -статистики, которое составило $F_{\text{набл}} = 589,53$. Принимаем уровень значимости $\alpha = 0,05$, тогда в соответствии с распределением Фишера $F_{\alpha; m; n-m-1} = 2,66$ и вычисленное значение $F_{\text{набл}} \gg F_{\alpha; m; n-m-1}$. Это означает, что объясненная дисперсия существенно больше остаточной дисперсии, а следовательно, уравнение регрессии достаточно качественно отражает динамику изменения зависимой переменной. На основе проведенных вычислений можно сделать вывод, что полученное уравнение регрессии объясняет 99,4% разброса зависимой переменной (теплопотребления).

Коэффициент детерминации может быть достаточно высоким, и при наличии совпадающих трендов у рассматриваемых переменных может наблюдаться высокая степень мультиколлинеарности, устранить которую можно путем исключения коррелированных переменных из полученной регрессионной модели. Исследования показали, что наилучшим образом изменение теплопотребления описывает регрессионная модель с двумя переменными (ВВП и времени), которая может быть представлена в следующем виде:

$$Q_i^{\text{mod}(5)} = a_0 + a_1 T_i + a_2 GDP_i + \varepsilon_i^{(5)}. \quad (13)$$

В соответствии с МНК были проанализированы значения коэффициентов регрессии данного уравнения, и в результате уравнение теплопотребления в зависимо-

сти от времени и ВВП приняло следующий вид:

$$Q_i^{\text{mod}(5)} = 1971,5436 - 59,887 \cdot T_i + 0,01297 \cdot GDP_i + \varepsilon_i^{(5)} \quad (14)$$

Для данной модели теплопотребления соответственно был выполнен статистический анализ. Статистическая ошибка уравнения регрессии, определенная по формуле (9), равна величине $S^1 = \sqrt{S^{(1)^2}} = 94,13$. Стандартные ошибки коэффициентов регрессии в соответствии с (10) для полученной многопараметрической линейной регрессионной модели теплопотребления имеют следующие значения:

$$S_{a_0}^2 = 137,9548; S_{a_1}^2 = 16,2414; S_{a_2}^2 = 0,00757.$$

Определенная с помощью формулы (10) t -статистика для коэффициентов рассматриваемой линейной многопараметрической регрессии имеет следующие значения:

$$t_{a_0} = 14,291; t_{a_1} = -3,687; t_{a_2} = 1,712.$$

Сравнительный анализ полученных значений с табличными значениями согласно распределению Стьюдента при выбранном уровне значимости ($\alpha = 0,05$) позволил сделать вывод о статистической значимости коэффициентов рассматриваемой линейной многопараметрической регрессии. Проверка общего качества уравнения регрессии состоит в вычислении коэффициента детерминации в соответствии с (3), значение которого для рассматриваемого уравнения множественной регрессии составило $R^2 = 0,877$.

В соответствии с полученным уравнением регрессии выполнена проверка статистической значимости коэффициента детерминации. Сравнение полученного значения F -статистики с табличным значением согласно распределению Фишера дает право утверждать, что уравнение регрессии достаточно качественно отражает динамику изменения зависимой переменной. На основе проведенных вычислений можно сделать вывод, что полученное уравнение регрессии объясняет 87,7% раз-

броса зависимой переменной (теплопотребления).

Изменение теплопотребления кроме как линейной зависимостью может быть описано с помощью экспоненциального уравнения. В качестве переменных (регрессоров), также как и в предыдущей модели, выберем показатели времени и ВВП. Однако для данного исследования будем использовать не абсолютные величины показателей теплопотребления и ВВП, а базисный темп роста данных показателей. В качестве базового значения выбраны значения показателей в конечном году n , поскольку логарифмирование будет производиться именно по значениям базисного роста теплопотребления. Общее уравнение экспоненциальной зависимости теплопотребления принимает следующий вид:

$$Q_i^{\text{mod}(6)} = e^{c_0 + c_1 \cdot T_i + c_2 \cdot GDP_i^{(6)}} \cdot \varepsilon_i^{(6)}, \quad (15)$$

где $Q_i^{\text{mod}(6)} = \frac{Q_i}{Q_n}$ - базисный рост теплопотребления, $i = 1, 2, \dots, n$;

$GDP_i^{(6)} = \frac{GDP_i}{GDP_n}$ - базисный рост ВВП, $i = 1, 2, \dots, n$;

c_0, c_1, c_2 - коэффициенты уравнения регрессии;

$\varepsilon_i^{(6)}$ - отклонение выборочных значений объясняющей переменной от значений, полученных с помощью экспоненциального уравнения.

В соответствии с МНК вычислены значения коэффициентов данного уравнения. Уравнение теплопотребления в зависимости от времени и ВВП примет следующий вид:

$$\ln Q_i^{\text{mod}(6)} = 0,2463 - 0,0315 \cdot T_i + 0,4681 \cdot GDP_i + \ln \varepsilon_i^{(6)} \quad (16)$$

Для данной модели теплопотребления также выполнен статистический анализ. Статистическая ошибка уравнения регрессии, определенная по формуле (9), составила $S^2 = \sqrt{S^{(2)^2}} = 0,043$. Стандартные ошибки коэффициентов регрессии в соответствии с (10) для полученной многопа-

раметрической линейной регрессионной модели теплопотребления имеют следующие значения:

$$S_{c_0}^2 = 0,063; S_{c_1}^2 = 0,0074; S_{c_2}^2 = 0,247.$$

Определенная с помощью формулы (11) t -статистика для коэффициентов рассматриваемой линейной многопараметрической регрессии будет иметь следующие значения:

$$t_{a_0} = 4,384; t_{a_1} = -4,243; t_{a_2} = 1,894.$$

Сравнение полученных значений с табличными данными согласно распределению Стьюдента при выбранном уровне значимости ($\alpha = 0,05$) позволяет сделать выводы о статистической значимости коэффициентов уравнения рассматриваемой линейной многопараметрической регрес-

сии. Проверка общего качества уравнения регрессии состоит в вычислении коэффициента детерминации в соответствии с (3), значение которого для рассматриваемого уравнения составляет $R^2 = 0,9097$. Однако данные вычисления являются верными для логарифма базисного роста теплопотребления. В дальнейшем путем приведения к экспоненциальной функции и умножения на значение теплопотребления в базисном году расчетный коэффициент детерминации составил 0,9071. На рис. 2 представлено графическое сравнение наблюдаемых значений теплопотребления и значений теплопотребления, полученных с помощью линейной и экспоненциальной регрессионных моделей.

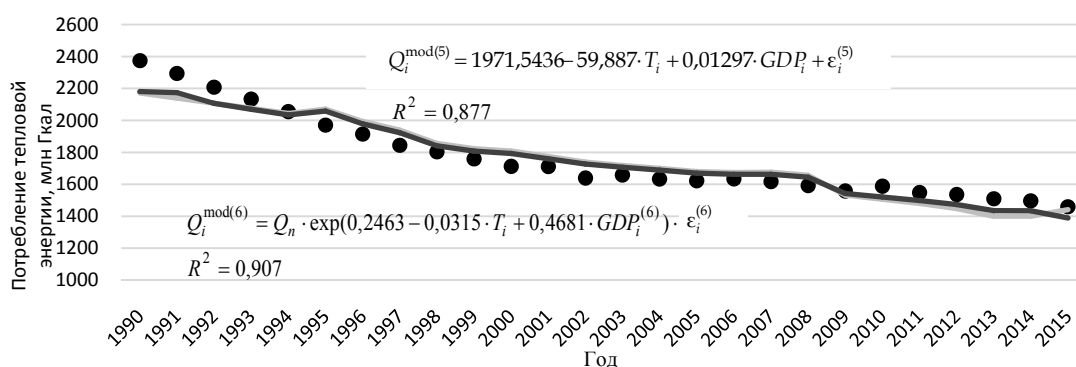


Рис. 2. Динамика наблюдаемых значений теплопотребления и значений теплопотребления, полученных с помощью регрессионной линейной и экспоненциальной моделей теплопотребления на ретроспективном отрезке времени:
 • наблюдаемые данные; — линейная многопараметрическая регрессионная модель;
 — экспоненциальная многопараметрическая модель

Как видно из рис. 2, рассматриваемые регрессионные модели достаточно хорошо описывают изменение теплопотребления на ретроспективном отрезке времени. Коэффициент детерминации экспоненциальной регрессионной модели теплопотребления имеет более высокое значение, что определяет высокую сходимость наблюдаемых данных и значений теплопотребления, полученных с помощью экспоненциальной модели.

Оценка прогнозных уровней теплопотребления с помощью полученных регрессионных моделей позволяет сделать срав-

нительный анализ полученных и имеющих в публикациях прогнозов. В связи с тем что в качестве исходных данных для прогнозирования теплопотребления было принято два варианта изменения ВВП (первый – консервативный, второй – целевой), получены интервальные оценки прогнозных уровней теплопотребления на перспективном отрезке времени. Длина прогнозируемого периода, на который распространяется действие регрессионной модели, напрямую зависит от рассматриваемого ретроспективного отрезка времени, на котором получена данная модель.

Принято, что прогнозируемый период не должен превышать 1/3 ретроспективного отрезка времени. В соответствии с данным обстоятельством прогноз теплопотребления

выполнен до 2025 г. На рис. 3 представлены результаты расчетов теплопотребления в соответствии с двумя вариантами экономического развития.



Рис. 3. Динамика наблюдаемых значений теплопотребления на ретроспективном отрезке времени с 1990 по 2015 г. и прогнозные значения теплопотребления, полученные с помощью регрессионных моделей:

- наблюдаемые данные; —•— авторегрессионная модель; — — — линейная многопараметрическая модель; — — — экспоненциальная многопараметрическая модель

Проведенные исследования показали, что в соответствии с прогнозами экономического развития при использовании линейной регрессионной модели сокращение теплопотребления к 2025 г. может составить 11,0% при первом варианте экономического развития или увеличиться на 11,1% при втором варианте по сравнению с уровнем 2015 г.

При использовании экспоненциальной модели для прогнозирования теплопотребления сокращение его уровня к 2025 г. может составить 8,9% в первом варианте развития экономики или увеличиться на 9,6% во втором варианте по сравнению с уровнем 2015 г. Интервал изменения теплопотребления при использовании экспоненциальной модели для прогнозирования имеет более узкие границы, что позволяет точнее определить возможный уровень теплопотребления. Расхождения между прогнозными значениями теплопотребления на уровне 2025 г., полученными с помощью линейной и нелинейной регрессионных моделей, составляет не более 2% (32 млн Гкал).

При использовании авторегрессионной модели для прогнозирования наблюдается сохранение тенденции сокращения объемов теплопотребления. В соответствии с данной моделью прогноз теплопотребления на период до 2025 г. показал сокращение уровня потребления тепловой энергии на 2,2% по сравнению с уровнем 2015 г. Кроме того, прогнозные значения теплопотребления, полученные с помощью авторегрессионной модели, располагаются на координатной плоскости примерно посередине по отношению к прогнозам, выполненным с помощью регрессионных моделей с учетом вариантности экономического развития. Данные обстоятельства позволяют сделать вывод о возможности использования полученных регрессионных моделей для мониторинга и среднесрочного прогнозирования теплопотребления.

Дальнейшие исследования планируется осуществлять в направлении расширения моделей объясняющими переменными (регрессорами), детализации описания структуры теплопотребления с помощью регрессионных моделей.

Список литературы

1. Добровольская Т. В., Стенников В. А. Мониторинг и прогнозирование теплопотребления с помощью регрессионных моделей // Моделирование и прогнозирование социально-эколого-экономического развития региона : материалы Всероссийской молодежной школы-конференции с международным участием. – Улан-Удэ : Изд-во БНЦ СО РАН, 2016. – С. 126–133.
2. Мастепанов А. М. Топливо-энергетический комплекс России на рубеже веков – состояние, проблемы и перспективы развития // Справочно-аналитический сборник. – 3-е изд., перераб. и доп. – Новосибирск : Наука, 2010.
3. Некрасов А. С. Анализ и прогнозы развития отраслей топливо-энергетического комплекса. Избранные труды. – М. : Издание ООО «ЛЕТО Индастриз», 2013.
4. Фёрстер Э., Рёнц Б. Методы корреляционного и регрессионного анализа / пер. с нем. и предисл. В. М. Ивановой. – М. : Финансы и статистика, 1983.

References

1. Dobrovolskaya T. V., Stennikov V. A. Monitoring i prognozirovanie teplopotrebleniya s pomoshch'yu regressionnykh modeley [Monitoring and Forecasting Thermal Consumption by Regressive Models]. *Modeling and Forecasting Social, Ecologic and Economic Development of the Region: materials of the All-Russian Youth School-Conference with Foreign Participation*. Ulan-Ude, Publishing house BNC CO RAN, 2016, pp. 126–133. (In Russ.).
2. Mastepanov A. M. Toplivno-energeticheskiy kompleks Rossii na rubezhe vekov – sostoyanie, problemy i perspektivy razvitiya [The Fuel and Power Complex of Russia at the Turn of Centuries – Condition, Challenges and Prospects of Development]. *Reference – Analytical Collection of Materials*, 3rd edition, amended and altered. Novosibirsk, Nauka, 2010. (In Russ.).
3. Nekrasov A. S. Analiz i prognozy razvitiya otrasley toplivno-energeticheskogo kompleksa. Izbrannyye trudy [Analysis and Forecasts of Development of the Fuel and Power Complex Industries. Selected Works]. Moscow, Published by 'LETO Industries', 2013. (In Russ.).
4. Ferster E., Renz B. Methods of Correlative and Regressive Analysis, translated from German, foreword by V.M. Ivanova. Moscow, Finance and Statistics, 1983. (In Russ.).

Сведения об авторах

Валерий Алексеевич Стенников

член-корр. РАН, доктор технических наук, профессор, врио директора ИСЭМ СО РАН.
Адрес: Федеральное государственное бюджетное учреждение науки Институт систем энергетики им. Л. А. Мелентьева Сибирского отделения Российской академии наук, 664033, г. Иркутск, ул. Лермонтова, д. 130.
E-mail: sva@isem.irk.ru

Татьяна Владимировна Добровольская

ведущий инженер ИСЭМ СО РАН.
Адрес: Федеральное государственное бюджетное учреждение науки Институт систем энергетики им. Л. А. Мелентьева Сибирского отделения Российской академии наук, 664033, г. Иркутск, ул. Лермонтова, д. 130.
E-mail: makarova@isem.irk.ru

Information about the authors

Valery A. Stennikov

Corresponding Member of the RAS, Doctor of Engineering Sciences, Professor, Acting Director of the ESI SB RAS.
Address: Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences, 130 Lermontova str., Irkutsk, 664033, Russian Federation.
E-mail: sva@isem.irk.ru

Tatyana V. Dobrovolskaya

Lead engineer of the ESI SB RAS.
Address: Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences, 130 Lermontova str., Irkutsk, 664033, Russian Federation.
E-mail: makarova@isem.irk.ru