

РАЗРАБОТКА МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ РЕЖИМОВ ЭЛЕКТРОПОТРЕБЛЕНИЯ ПРОМЫШЛЕННЫХ ПОТРЕБИТЕЛЕЙ НА ПРИМЕРЕ ПРЕДПРИЯТИЙ ТРУБОПРОВОДНОГО ТРАНСПОРТА НЕФТИ

Н. В. ТОКОЧАКОВА, А. С. ФИКОВ

*Учреждение образования «Гомельский государственный
технический университет имени П. О. Сухого»,
Республика Беларусь*

Введение

В Республиканской программе энергосбережения [1] ставились задачи разработки и ввода в действие системы прогрессивных норм расхода ТЭР, скоординированной с проведением энергетических обследований, и продолжения работы по корректировке действующих и разработке новых нормативно-технических документов в части повышения эффективности использования ТЭР. Поэтому разработка новых методов, позволяющих выявлять и оценивать потенциал энергосбережения в технологических процессах промышленных потребителей (ПП), прогнозировать и нормировать как общие, так и удельные расходы электрической энергии (ЭЭ) с учетом изменения объемов выпускаемой продукции, прогнозировать и нормировать целевой показатель по энергосбережению в сопоставимых условиях, является актуальной.

Постановка задачи

Ранее [2], для решения комплекса задач управления энергоэффективностью ПП со сложной взаимосвязью между энергетикой и технологией, было предложено использовать регрессионные модели зависимости режимов электропотребления от влияющих факторов. С использованием таких моделей предусматривается оценка экономии ЭЭ от внедрения мероприятий в технологической системе ПП, прогнозирование и оценка текущего состояния показателей энергетической эффективности: общих и удельных расходов ЭЭ, целевого показателя по энергосбережению. Для решения указанных задач необходима разработка *адекватных* моделей режимов электропотребления промышленных потребителей.

Предлагаемый способ решения

Важным этапом моделирования является определение вида модели и набора независимых переменных, включенных в нее. На практике указанные задачи решаются по-разному. В [3] предлагается прекратить ввод независимых переменных, когда остаточная дисперсия относительно возмущающей переменной начнет увеличиваться либо когда приведенный коэффициент детерминации модели с вводом новых переменных увеличивается незначительно. Наряду с этим в [4] показано, что отбор независимых переменных, произведенный при условиях устранения мультиколлинеарности и сохранения только существенных и линейно независимых коэффициентов регрессии, не всегда приводит к достаточно точным и надежным моделям. При формальной оценке надежности модели различными статистическими критериями ускользают те реальные причинно-следственные связи между параметрами, которые определяются особенностью исследуемого процесса. Как правило, в прикладных исследованиях для выявления количественной меры влияния важна оценка именно причинных связей. Анализ структуры связей производится на основе теоретических или профессиональных

предположений о наличии и направлении связи. Таким образом, набор независимых переменных, введенных в модель, должен быть обоснованным с использованием теоретических зависимостей и соответствовать цели создания модели [5]. Для целей нормирования, прогнозирования, энергетического обследования модель целесообразно представить в линейном аддитивном виде как наиболее простую в построении, использовании и интерпретации. Схема построения модели в наиболее общем виде представлена на рис. 1.

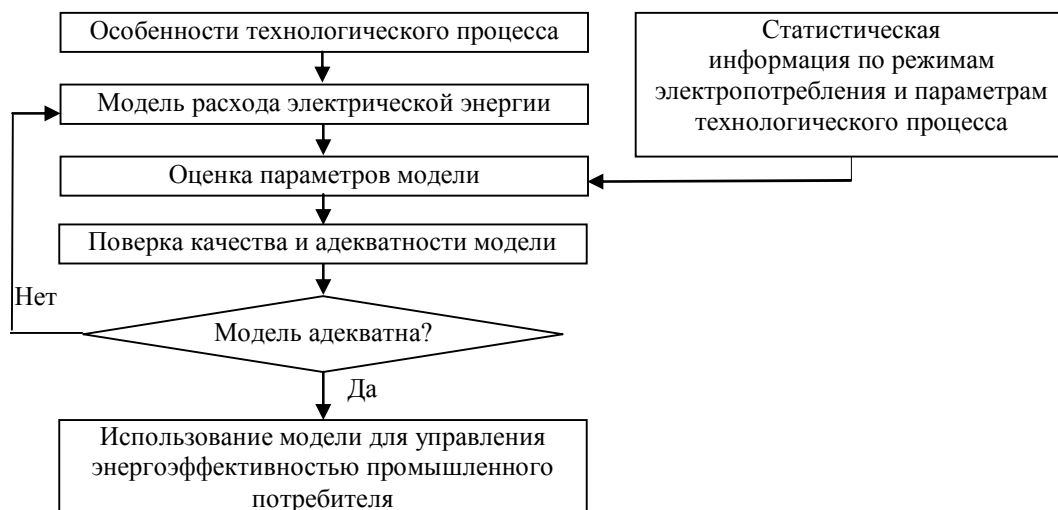


Рис. 1. Схема построения модели режимов электропотребления промышленного потребителя

Для предприятий трубопроводного транспорта нефти на основе анализа физики протекания технологического процесса было предложено построение модели режимов электропотребления участков нефтепровода в виде линейного аддитивного уравнения регрессии, где в качестве независимых переменных включены грузооборот нефти P , ее вязкость ν , эквивалентный диаметр нефтепровода d_3 [2]:

$$W = \beta_p \cdot P + \beta_\nu \cdot \nu + \beta_d \cdot d_3 + \beta_0 + \varepsilon, \text{ кВт} \cdot \text{ч/сут}, \quad (1)$$

где β_p – коэффициент регрессии, кВт·ч/тыс. т·км; P – грузооборот нефти, тыс. т·км/сут; β_ν – коэффициент регрессии, кВт·ч·с/(м²·сут); ν – вязкость нефти, м²/с; β_d – коэффициент регрессии, кВт·ч/(м·сут); d_3 – эквивалентный диаметр нефтепровода, м; β_0 – свободный член уравнения регрессии, кВт·ч/сут; ε – стохастическая составляющая зависимой переменной, возникшая в результате воздействия неконтролируемых или неучтенных факторов, а также измерений независимых переменных, неизбежно сопровождающихся некоторыми случайными ошибками, кВт·ч/сут.

Выбор линейного аддитивного вида уравнения регрессии основывался на характере рассеивания (не противоречащем линейному) суточных значений электропотребления от каждой из независимых переменных в отдельности. Для увеличения точности построение модели осуществляется на выделенном интервале грузооборота. Совокупность таких моделей, охватывающих весь рабочий диапазон грузооборота нефти, представляет собой дискретно-непрерывную модель. Следует отметить, что, как правило, достаточная точность моделирования может быть достигнута построением одной линейной модели на всем рабочем интервале грузооборота [6].

Традиционно оценка коэффициентов уравнения регрессии производится методом наименьших квадратов, основными желательными предпосылками которого являются [7], [8]:

- 1) равенство нулю математического ожидания стохастической составляющей ($\mu_\varepsilon = 0$);
- 2) гомоседакционность или постоянство дисперсии стохастической составляющей ($D(\varepsilon_i) = D(\varepsilon_j)$);
- 3) отсутствие автокорреляции стохастической составляющей ($\text{cov}(\varepsilon_i, \varepsilon_j) = 0; \forall i \neq j$);
- 4) независимость стохастической составляющей от объясняющих параметров ($\text{cov}(\varepsilon_i, X_j^k) = 0; X_j^k$ – j -я реализация k -го параметра);
- 5) отсутствие между независимыми переменными сильной (по Чеддоку) линейной зависимости (мультиколлинеарности);
- 6) распределение ε по нормальному закону.

Как видно из рис. 1, оценка параметров модели основывается на статистической информации по режимам электропотребления и параметрам технологического процесса. Для качественного выполнения указанного этапа моделирования важным является обоснование глубины выборки ретроспективной статистической информации. В [4] показано, что при известной дискретности проведения экспериментов и 5-процентной вероятности попадания случайной величины W в крайние значения рабочего диапазона общая продолжительность необходимого времени наблюдения может быть определена из выражения:

$$T = 79,6 \cdot \Delta t, \text{ ч}, \quad (2)$$

где Δt – дискретность проведения экспериментов, ч.

В случае использования суточной статистики ($\Delta t = 24$ ч) для моделирования режимов электропотребления время наблюдения составит $T = 79,6 \cdot 24 = 1910,4$ ч, а количество проведенных опытов должно составить 80 единиц. Однако найденное время наблюдения T не охватывает годовой интервал времени, и такой технологический параметр, как вязкость нефти не включает все возможные значения рабочего диапазона в виду сезонного характера ее изменения. Всякая же экстраполяция модели режимов электропотребления за пределы обследованного интервала значений технологических параметров может привести к большим погрешностям. Энергетическая эффективность технологической системы ПП постоянно изменяется. Выбор в качестве основы моделирования режимов электропотребления суточной статистики, охватывающей временной интервал более одного года, приведет к увеличению влияния на результаты прогноза расхода ЭЭ состояний технологической системы имевших место в прошлом, но не планируемых в будущем. Таким образом, моделирование режимов электропотребления для решения задач нормирования и прогнозирования расхода ЭЭ предприятий трубопроводного транспорта нефти должно основываться на суточной статистике, охватывающей годовой интервал времени. Что с учетом реального времени прокачки нефти по трубопроводам составляет порядка 360 сут. Следует отметить, что при моделировании режимов электропотребления ПП вследствие особенностей системы учета электрических и технологических параметров возможна ситуация, когда дискретность съема статистических данных составляет один месяц. Необходимое время наблюдения согласно выражению (2) составит порядка 6,5 лет. Использование столь глубокой ретроспективной информации по режимам электропотребления нецелесообразно в силу вероятных значительных изменений технологического процесса за рассматриваемый период. В этом случае минимальный объем выборки должен составлять не менее утроенного количества оцениваемых параметров уравнения регрессии [7].

Качество построенной модели характеризуется как максимальной относительной погрешностью δ^{\max} , так и среднеквадратичным отклонением относительной погрешности σ [2]:

$$\delta^{\max} = \max \left\{ \delta_i = \left| 100 \frac{\varepsilon_i}{W_i} \right| \right\}_{i=1,2,\dots,n}, \quad \%, \quad (3)$$

$$\sigma = \sqrt{\frac{1}{n-k-1} \cdot \sum_{i=1}^n (\delta_i - \bar{\delta})^2}, \%, \quad (4)$$

где $\bar{\delta}$ – среднеарифметическое значение относительной погрешности, %.

Проверка адекватности построенной модели основывается на выполнимости предпосылок метода наименьших квадратов. В этих условиях оценки коэффициентов регрессии являются несмещенными, состоятельными и эффективными. Выполнимость последней предпосылки используется для проверки статистических гипотез. В [7] показано, что оценки коэффициентов регрессии являются линейными комбинациями ε . Поскольку линейная комбинация нормально распределенных случайных величин является нормально распределенной случайной величиной, то распределение оценок коэффициентов регрессии также является нормальным. Следует отметить, что при возрастании объема выборки (от 30 и более) форма распределения выборочной статистики критерия приближается к нормальной, даже если распределение исследуемых величин не является таковым [9], что позволяет использовать критерии Стьюдента и Фишера для проверки значимости коэффициентов регрессии при отклонениях ε от нормального закона распределения.

При проведении регрессионного анализа для ограниченной по объёму совокупности параметры уравнения регрессии, коэффициенты детерминации могут быть сильно искажены действием случайных факторов. В этих условиях необходима дополнительная проверка модели на адекватность. При этом производится проверка значимости (существенности) каждого коэффициента регрессии. Значимость коэффициентов линейной регрессии определяют с помощью t -критерия Стьюдента [7], [8], например для коэффициента β_p :

$$t_{\beta_p} = |\beta_p| \sqrt{(n-2) \frac{\sum (P - \bar{P})^2}{\sum (W - \bar{W})^2}}, \quad (5)$$

где \bar{P} – среднеарифметическое значение грузооборота нефти, тыс. т · км/сут.

Вычисленные по формуле (5) значения сравнивают с критическими, которые определяют по таблице Стьюдента с учетом принятого уровня надежности $1 - \alpha$ и объема выборки. Уровень надежности $1 - \alpha$ обычно принимают равным 0,95. Параметр признаётся значимым (существенным) при условии, если $t_{\text{расч}} > t_{\text{табл}}$. В таком случае практически невероятно, что найденное значение коэффициента регрессии обусловлено только случайными совпадениями.

Для проверки значимости уравнения регрессии вычисляется статистический критерий Фишера (F -критерий) [7], [8]:

$$F = \frac{\sum (\bar{W} - \bar{W})^2}{\sum (W - \bar{W})^2} \frac{n-k-1}{k}. \quad (6)$$

Рассчитанное значение F -критерия сравнивается с табличным при k , $n - k - 1$ степенями свободы и заданном уровне надежности $1 - \alpha$. При превышении рассчитанным значением табличного практически невероятно, что все найденные значения коэффициентов регрессии обусловлены только случайными совпадениями:

$$F > F_{1-\alpha}(k, n-k-1). \quad (7)$$

Для предприятий трубопроводного транспорта нефти алгоритм построения модели режимов электропотребления может быть представлен в следующем виде [2], [6]:

1) формируются суточные временные ряды расхода ЭЭ и факторов, включенных в модель на временном интервале $t = 1, 2, \dots, n''$:

$$\{W_t\}_{t=1,2,\dots,n''}, \{P_t\}_{t=1,2,\dots,n''}, \{v_t\}_{t=1,2,\dots,n''}, \{d_{эt}\}_{t=1,2,\dots,n''};$$

2) сформированные временные ряды очищаются от календарных эффектов (исключаются сутки, соответствующие первому и последнему числам месяца, из-за несовпадения начала отсчета объемов транспортируемой нефти и расхода ЭЭ в первый день месяца):

$$\{W_t\}_{t=1,2,\dots,n'}, \{P_t\}_{t=1,2,\dots,n'}, \{v_t\}_{t=1,2,\dots,n'}, \{d_{эt}\}_{t=1,2,\dots,n'};$$

$$n' = n'' - NK,$$

где NK – количество суток, исключенных из временных рядов;

3) в зависимости от цели построения модели выбирается период усреднения τ и по выражениям (8) рассчитываются сглаженные временные ряды (9):

$$\bar{W}_t = \frac{1}{\tau} \sum_{i=t}^{t+\tau-1} W_i, \quad \bar{P}_t = \frac{1}{\tau} \sum_{i=t}^{t+\tau-1} P_i, \quad \bar{v}_t = \frac{1}{\tau} \sum_{i=t}^{t+\tau-1} v_i, \quad \bar{d}_{эt} = \frac{1}{\tau} \sum_{i=t}^{t+\tau-1} d_{эi}; \quad (8)$$

$$\{\bar{W}_t\}_{t=1,2,\dots,n}, \{\bar{P}_t\}_{t=1,2,\dots,n}, \{\bar{v}_t\}_{t=1,2,\dots,n}, \{\bar{d}_{эt}\}_{t=1,2,\dots,n}; \quad (9)$$

$$n = n' - (\tau - 1); \quad t = 1, 2, \dots, n.$$

Взаимосвязь между независимыми и зависимой переменными, выраженная через оценки коэффициентов регрессии, вследствие снижения вариации параметров $W, P, v, d_{э}$, при увеличении временного цикла прогнозирования (месяц, квартал, год) объективно изменяется. Выбор значения τ непосредственно зависит от величины временного цикла прогнозирования. На практике это означает, что для прогнозирования среднеквартально расхода ЭЭ желательно искать взаимосвязь между среднеквартальным потреблением ЭЭ и среднеквартальными значениями независимых переменных, что обеспечивается выбором $\tau = 90$ сут. Для задачи нормирования в настоящее время принято разрабатывать среднеквартальные удельные расходы на транспортировку нефти и, соответственно, $\tau = 90$ сут; при оценке экономии ЭЭ от проведения энергосберегающих мероприятий τ принимается равным 30 сут, поскольку в Республике Беларусь сложилась система ежемесячной статистической отчетности по экономии ТЭР; при оценке ЦП τ увеличивается от 30 сут (за январь) до 360 сут (за декабрь), поскольку целевой показатель рассчитывается ежемесячно с нарастающим итогом;

4) методом наименьших квадратов на выбранном интервале грузооборота строится модель расхода ЭЭ от сглаженных значений грузооборота и вязкости нефти, эквивалентного диаметра нефтепровода:

$$\mathcal{E} = \beta_p \cdot P + \beta_v \cdot v + \beta_d \cdot d_{э} + \beta_0, \text{ кВт} \cdot \text{ч/сут}; \quad (10)$$

5) рассчитывается регрессионная статистика, позволяющая оценить качество модели; последнее характеризуется как максимальной относительной погрешностью δ^{\max} , так и среднеквадратичным отклонением относительной погрешности σ ;

6) с использованием F -критерия и t -критерия проверяется значимость коэффициентов регрессии (выражения (5), (6));

7) модель (10) для целей нормирования расхода ЭЭ на транспортировку нефти преобразуется в модель удельного расхода ЭЭ по следующему выражению:

$$W_{уд} = \beta_p + \frac{\beta_v v + \beta_d d_3 + \beta_0}{P}, \text{ кВт} \cdot \text{ч/тыс. т} \cdot \text{км.} \quad (11)$$

Ключевым моментом приведенного алгоритма является сглаживание временных рядов. Проанализируем влияние сглаживания скользящим средним на точность моделирования расхода ЭЭ. Фактическое электропотребление отличается от рассчитанного по модели на величину стохастической составляющей:

$$W_t = \mathcal{W}_t + \varepsilon_t, \text{ кВт} \cdot \text{ч/сут}, \quad (12)$$

где W_t – фактическое суточное электропотребление, кВт · ч/сут; \mathcal{W}_t – расчетное суточное электропотребление, кВт · ч/сут; ε_t – стохастическая составляющая, кВт · ч/сут.

Мерой неопределенности информации в поведении $\{W\}$ выступают максимальная относительная погрешность δ^{\max} и среднеквадратическое отклонение относительной погрешности σ на интервале времени $t = 1, 2, \dots, n$. Выражения (3), (4) показывают, что δ и σ связаны с ε прямопропорциональной зависимостью, следовательно, при снижении ε качество модели будет улучшаться.

С использованием выражения (12) выделяется стохастическая составляющая ε_t . Среднеарифметическое значение $\bar{\varepsilon}$, при достаточно большом периоде сглаживания τ , приближенно равно ее математическому ожиданию [6], [8]:

$$\bar{\varepsilon} = \sum_{i=1}^{\tau} \varepsilon_i \frac{m_i}{\tau} \approx \mu_\varepsilon = \sum_{i=1}^{\tau} \varepsilon_i \cdot p_i = 0, \quad (13)$$

где m_i – частота значения ε_i ; μ_ε – математическое ожидание величины ε_i ; p_i – вероятность значения ε_i .

Как следует из теоремы Бернулли [6], [8], точность этого равенства будет расти с увеличением выборки:

$$\lim_{\tau \rightarrow \infty} P\left(\left|\frac{m_i}{\tau} - p_i\right| < \xi\right) = 1, \quad (14)$$

где ξ – сколь угодно малая величина.

Как следует из ранее сделанных допущений, а также основываясь на оценке фактических данных, математическое ожидание стохастической составляющей близко к нулю ($\mu_\varepsilon = 0$) [6]. Данное важное свойство позволяет уменьшить ε методом скользящего среднего (*Moving Average*), и, следовательно, улучшить качество модели:

$$\bar{W}_t = \frac{1}{\tau} \sum_{i=t}^{t+\tau-1} W_i = \frac{1}{\tau} \sum_{i=t}^{t+\tau-1} \mathcal{W}_i + \frac{1}{\tau} \sum_{i=t}^{t+\tau-1} \varepsilon_i = \tilde{W}_t + \bar{\varepsilon}_t, \text{ кВт} \cdot \text{ч/сут}, \quad (15)$$

где \bar{W} – фактическое среднее электропотребление, кВт · ч/сут; \tilde{W} – расчетное среднее электропотребление, кВт · ч/сут; $\bar{\varepsilon}$ – среднее значение стохастической составляющей, кВт · ч/сут.

Из вышеизложенного следует, что с ростом периода усреднения τ стохастическая составляющая ε снижается и, как следствие, снижаются δ^{\max} и σ (рис. 2).

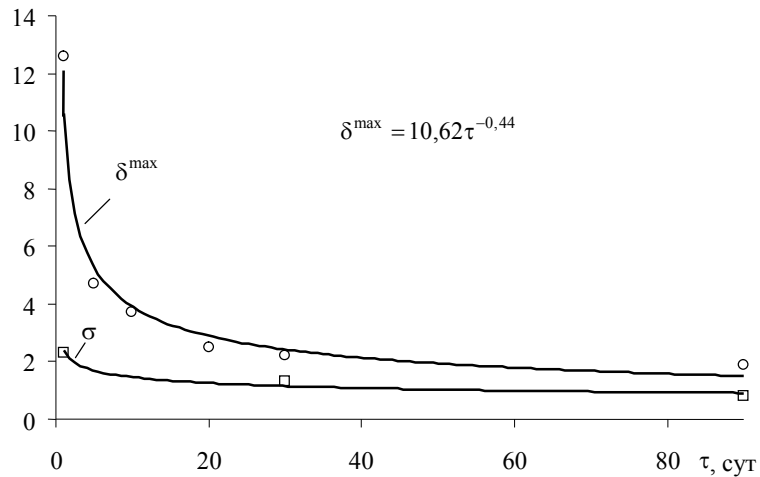


Рис. 2. Динамика δ^{\max} и σ в функции периода усреднения τ

Из представленного рисунка видно, что между мерой неопределенности в поведении W и периодом усреднения существует убывающая степенная зависимость. Для модели, построенной на несглаженных данных, δ^{\max} равно 12,6 %, что лежит за пределами точности, допускаемыми техническими расчетами. В пересчете на абсолютное значение расхода ЭЭ мера неопределенности может достигать 110 МВт · ч/сут. Следовательно, использование такой модели неприемлемо. Лучшее качество модели достигается при квартальном периоде усреднения ($\tau = 90$) $\delta^{\max} = 1,9$ %; σ снижается с 2,3 % до 0,9 %.

Рассмотрим пример построения модели расхода ЭЭ согласно приведенному алгоритму. На первом этапе формируются суточные временные ряды расхода ЭЭ и технологических факторов P, v, d_3 . Невозможность расчета эквивалентного диаметра участка нефтепровода за все сутки годового интервала времени из-за частичного отсутствия исходной статистической информации приводит к сокращению объема выборки до $n'' = 245$ сут. Далее из сформированных временных рядов исключаются календарные эффекты, что приводит к сокращению объема исходной статистической информации до $n' = 233$ сут. При этом достигнутое улучшение статистики можно оценить графически (рис 3).

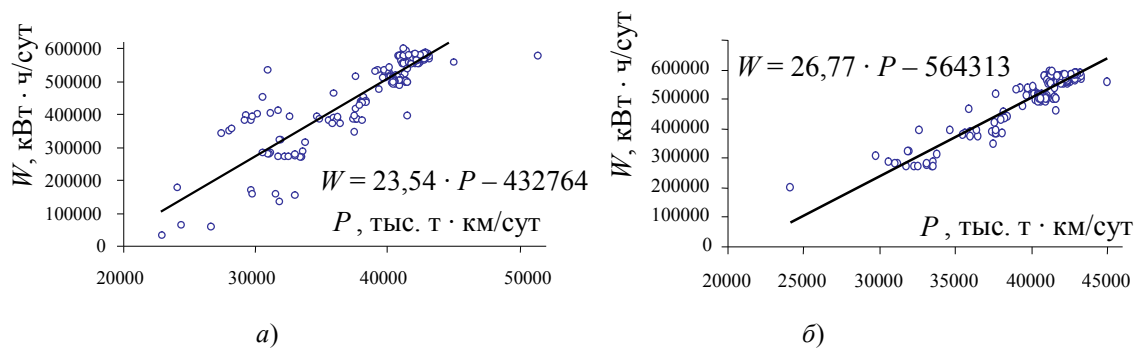


Рис. 3. Модель зависимости электропотребления от грузооборота нефти:
 а – исходная суточная статистика; б – статистика без календарных эффектов

Суточная статистическая информация длиной n' сглаживается с периодом усреднения $\tau = 30$ и строится модель зависимости электропотребления от технологических параметров транспортировки нефти (рис. 4). Полученная после сглаживания длина временного ряда составляет $n = 204$ сут, что превышает минимальный необходимый объем суточной статистической информации в 2,55 раза.

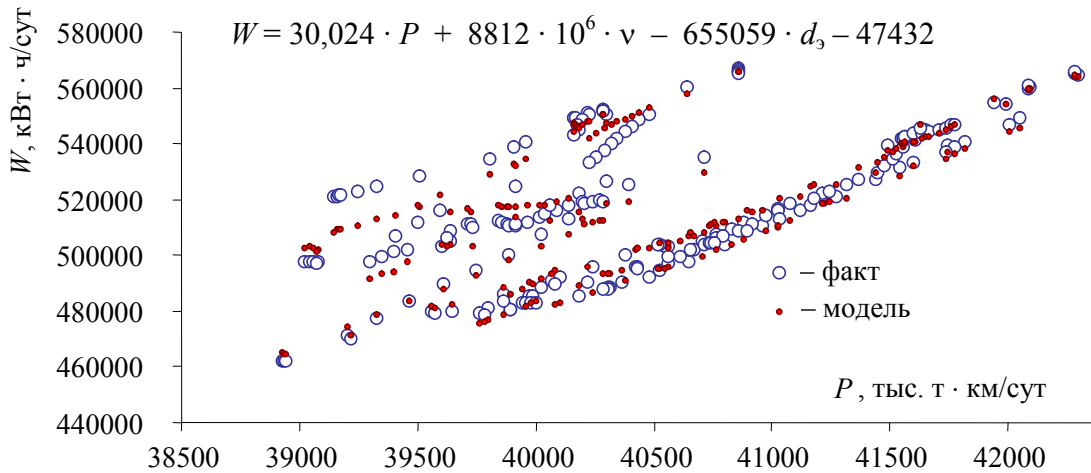


Рис. 4. Модель зависимости электропотребления от технологических параметров транспортировки нефти

Перед оценкой качества построенной модели произведем проверку выполнимости предпосылок метода наименьших квадратов. Одной из важнейших предпосылок является отсутствие мультиколлинеарности. В табл. 1 приведены результаты проверки взаимосвязей характеристик подсистем системы нефтепровода с помощью коэффициента парной корреляции Пирсона [7], [10]:

$$r_{xy} = \frac{\sum_{t=1}^n (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^n (x_t - \bar{x})^2 \sum_{t=1}^n (y_t - \bar{y})^2}}, \quad (16)$$

где x_t, y_t – значения случайных величин в t -й интервал времени; \bar{x}, \bar{y} – средние значения случайных величин; n – объем выборки.

Таблица 1

Коэффициенты Пирсона для характеристик подсистем участка нефтепровода РУП «Гомельтранснефть Дружба»

r_{Pv}	r_{Pd}	r_{vd}
-0,41	0,09	0,06

Полученные значения r_{xy} согласно известной шкале Чеддока характеризуют взаимосвязь между характеристиками подсистем P, v как обратную «умеренную» ($|R_{xy}| < 0,5$), между характеристиками P, d и v, d взаимосвязь не выявлена ($|R_{xy}| < 0,1$).

Проверка значимости коэффициентов корреляции Пирсона основывается на предположении о нормальности распределения величин x и y и осуществляется по величине t_s [7], распределенной по закону Стьюдента с $n - 2$ степенями свободы:

$$t_s = |r_{Pv}| \sqrt{\frac{n-2}{1-r_{Pv}^2}} = 0,41 \cdot \sqrt{\frac{204-2}{1-0,41^2}} = 6,38.$$

Поскольку найденное значение t_s больше критического $t_{кр}(\frac{\alpha}{2}, n-2) \approx 1,97$ при уровне значимости $\alpha = 0,05$, то найденное значение r_{Pv} статистически значимо.

На практике отклонения закона распределения величин x и y от нормального могут исказить оценку их взаимосвязи по критерию Пирсона [10], поэтому произведем

дополнительную проверку взаимосвязей по непараметрическим критериям Спирмена и Кендалла. В табл. 2 приведены результаты проверки взаимосвязей характеристик подсистем системы нефтепровода с помощью коэффициента ранговой корреляции Спирмена [7], [10]:

$$R_{xy} = 1 - \frac{6 \sum d^2}{n(n-1)(n+1)}, \quad (17)$$

где d – разность между рангами характеристик x и y ; n – объем выборки.

Таблица 2

Коэффициенты Спирмена характеристик подсистем участка нефтепровода РУП «Гомельтранснефть Дружба»

R_{Pv}	R_{Pd}	R_{vd}
-0,41	0,09	0,06

Полученные значения R_{xy} совпадают с параметрическими аналогами r_{xy} и согласно шкале Чеддока характеризуют взаимосвязь между характеристиками подсистем P, v как обратную «умеренную», между характеристиками P, d и v, d взаимосвязь не выявлена. При большом объеме выборки ($n > 20$; в нашем случае $n = 204$) проверка значимости коэффициентов ранговой корреляции Спирмена осуществляется по величине t_s [7], распределенной по закону Стьюдента с $n - 2$ степенями свободы:

$$t_s = |R_{Pv}| \sqrt{\frac{n-2}{1-R_{Pv}^2}} = 0,41 \cdot \sqrt{\frac{204-2}{1-0,41^2}} = 6,38.$$

Поскольку найденное значение t_s больше критического $t_{кр}(\frac{\alpha}{2}, n-2) \approx 1,97$ при уровне значимости $\alpha = 0,05$, то найденное значение R_{Pv} статистически значимо.

В табл. 3 приведены результаты проверки взаимосвязей характеристик подсистем системы нефтепровода с помощью коэффициента ранговой корреляции Кендалла, дающего более осторожную оценку взаимосвязи по сравнению с коэффициентами Спирмена и Пирсона [10]:

$$\tau_{xy} = \frac{2}{n(n-1)} S = \frac{2}{n(n-1)} \sum_{i,j=1}^n \text{sgn}[(r_{xi} - r_{xj})(r_{yi} - r_{yj})], \quad (18)$$

где n – объем выборки; r_x, r_y – ранги характеристик x и y .

Таблица 3

Коэффициенты Кендалла характеристик подсистем участка нефтепровода РУП «Гомельтранснефть Дружба»

τ_{Pv}	τ_{Pd}	τ_{vd}
-0,23	0,07	0,08

Полученные значения τ_{xy} согласно шкале Чеддока характеризуют взаимосвязь между характеристиками подсистем P, v как обратную «слабую» ($|\tau_{xy}| < 0,3$), между характеристиками P, d и v, d взаимосвязь не выявлена ($|\tau_{xy}| < 0,1$). При большом ($n > 10$)

объеме выборки проверка значимости коэффициентов ранговой корреляции Кендалла осуществляется по величине z , распределенной по нормальному закону [10]:

$$z = \frac{|S - \text{sgn}(S)|}{\sqrt{\frac{1}{18}n(n-1)(2n+5)}} = \frac{|-4712 - (-1)|}{\sqrt{\frac{1}{18}204(204-1)(2 \cdot 204 + 5)}} = 0,21,$$

где величина S при подсчете $\tau_{P,v}$ составила -4712 .

Поскольку найденное значение z меньше критического $z_{\text{кр}} = 1,96$ при уровне значимости $\alpha = 0,05$, то найденное значение $\tau_{P,v}$ статистически не значимо. Следовательно, гипотеза о равенстве нулю коэффициента $\tau_{P,v}$ верна. Окончательно по коэффициентам Пирсона, Кендалла и Спирмена можно сделать вывод о *линейной* независимости характеристик подсистем P,v и v,d и наличии умеренной обратной линейной зависимости между характеристиками P,v . Это дает основание разрабатывать модель режимов электропотребления участка нефтепровода в виде аддитивного уравнения регрессии.

Произведем проверку выполнимости предпосылок относительно стохастической составляющей. С проблемой не выполнения требования гомоседаксичности в основном сталкиваются при моделировании показателей от заведомо неоднородных независимых переменных (например, доход предприятий) [7]. Следует отметить, что при моделировании зависимости электропотребления ПП от технологических параметров случаи невыполнения требования гомоседаксичности маловероятны. При моделировании удельных расходов ЭЭ ПП целесообразно разрабатывать модель общего расхода ЭЭ и от нее переходить к модели удельного, поскольку для большинства ПП дисперсия удельного расхода ЭЭ в области низкой загрузки технологического оборудования априори выше, чем в области высокой. Для общего расхода ЭЭ такой зависимости не наблюдается.

Независимость стохастической составляющей от технологических параметров может быть показана графически либо для каждого параметра в отдельности, либо от их линейной комбинации, в качестве которой может выступать модельное электропотребление \mathcal{W} . Поскольку на рис. 5 зависимости стохастической составляющей от величины \mathcal{W} не прослеживается, то требование гомоседаксичности выполняется.

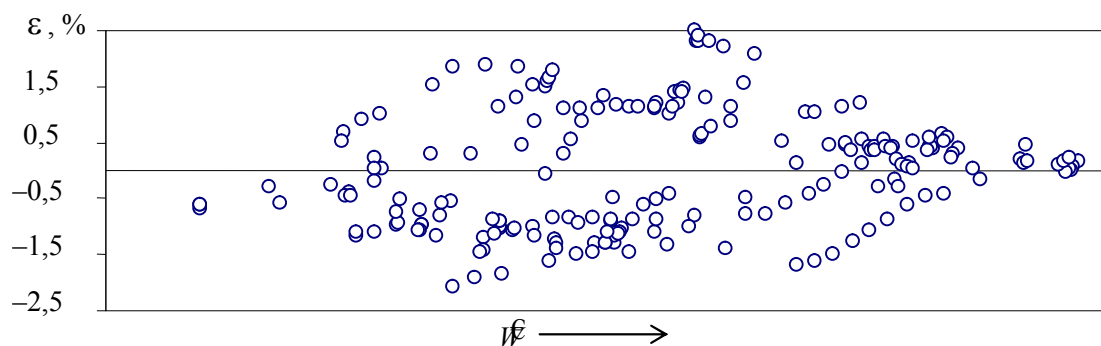


Рис. 5. График остатков в функции линейной комбинации технологических параметров

Из представленного графика остатков также видно, что дисперсия стохастической составляющей практически постоянна, а ее математическое ожидание близко к нулю ($\mu_\varepsilon \approx \bar{\varepsilon} = -0,01 \%$).

Произведем приближенную проверку на нормальность распределения стохастической составляющей. Известно [7], [8], что нормально распределенная величина подчиняется следующим правилам: практически все отклонения (99,7 %) лежат в диапазоне $\mu_\varepsilon \pm 3 \cdot \sigma_\varepsilon$;

примерно 2/3 всех отклонений (точнее, 68,3 %) должны лежать в диапазоне $\mu_\varepsilon \pm \sigma_\varepsilon$; примерно половина всех отклонений должна находиться в диапазоне $\mu_\varepsilon \pm 0,675 \cdot \sigma_\varepsilon$. Оценка среднеквадратического отклонения случайной величины ε определяется по выражению:

$$\sigma_\varepsilon = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2} . \quad (19)$$

Данные правила могут быть применены для упрощенной проверки гипотезы о нормальном распределении случайной величины. В табл. 4 представлена проверка на нормальность стохастической составляющей.

Таблица 4

Параметры закона распределения стохастической составляющей

n	μ_ε	σ_ε	Количество значений, лежащих в диапазоне $\mu_\varepsilon \pm 3 \cdot \sigma_\varepsilon$		Количество значений, лежащих в диапазоне $\mu_\varepsilon \pm \sigma_\varepsilon$		Количество значений, лежащих в диапазоне $\mu_\varepsilon \pm 0,675 \cdot \sigma_\varepsilon$	
			ед.	%	ед.	%	ед.	%
204	0	1,03	204	100	127	62,2	94	46,1

По данным табл. 4 можно сделать вывод о близости распределения стохастической составляющей к нормальному закону. На рис. 6 приведена гистограмма распределения стохастической составляющей, которая иллюстрирует схожесть характера исследуемого распределения с нормальным.

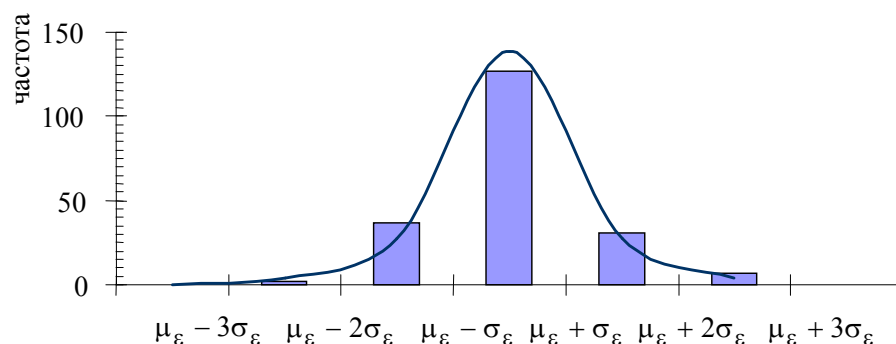


Рис. 6. Гистограмма стохастической составляющей

На практике при выявлении автокорреляции исследуют взаимосвязь между двумя соседними значениями стохастической составляющей $\text{cov}(\varepsilon_i, \varepsilon_{i-1}) = 0$. Произведем графическую проверку автокорреляции. Из рис. 7 видно, что взаимосвязи между величинами $\varepsilon_i, \varepsilon_{i-1}$ не прослеживается.

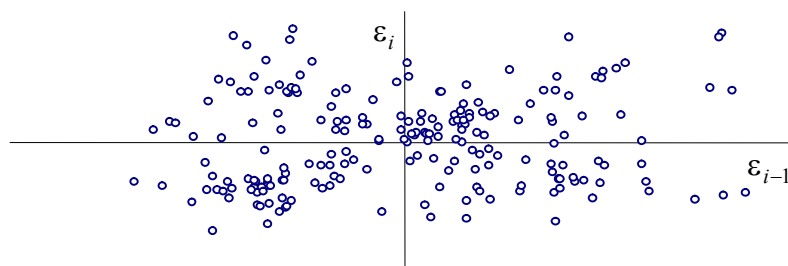


Рис. 7. Зависимость стохастической составляющей ε_i от ε_{i-1}

Последним этапом построения модели является проверка статистической значимости полученных коэффициентов регрессии. Расчетные значения t -критерия превышают табличное значение, при уровне надежности $\alpha = 0,95$ и числе степеней свободы $s = 202$ приблизительно равное 1,972: $t_p = 58,5$; $t_v = 51,1$; $t_d = 6,4$. Расчетное значение F -критерия превышает табличное ($F = 1434$), при уровне надежности $\alpha = 0,95$ и числами степеней свободы $s_1 = 3$, $s_2 = 200$ приблизительно равное 2,65. Таким образом, гипотезы о равенстве нулю как каждой оценки коэффициентов регрессии в отдельности, так и в совокупности отклоняется.

Выводы

1. Для предприятий трубопроводного транспорта нефти разработан алгоритм построения модели зависимости режимов электропотребления от воздействующих факторов, отличающийся использованием метода сглаживания временных рядов, что позволяет выявить тенденции режимов электропотребления в различных временных интервалах усреднения.

2. Произведена оценка влияния сглаживания скользящим средним на точность моделирования расхода электрической энергии. Установлено, что с ростом периода усреднения τ стохастическая составляющая ε снижается и, как следствие, уменьшаются максимальная относительная погрешность δ^{\max} и среднеквадратическое отклонение относительной погрешности σ . Так, при периоде сглаживания $\tau = 90$ сут δ^{\max} снижается с 12,6 % до 1,9 %; σ снижается с 2,3 % до 0,9 %.

3. С использованием корреляционного анализа доказано отсутствие сильной по Чеддоку линейной зависимости таких параметров технологической системы нефтепровода как грузооборот, вязкость нефти, эквивалентный диаметр нефтепровода, что позволяет в качестве вида модели зависимости режимов электропотребления участка нефтепровода от указанных параметров принять аддитивное линейное трехфакторное уравнение регрессии.

4. На основе представленного алгоритма произведена разработка модели зависимости режимов электропотребления участка нефтепровода от влияющих факторов, позволяющая оценивать среднемесячную экономию ЭЭ от воздействия на вязкость и эквивалентный диаметр нефтепровода, а также прогнозировать среднемесячный расход электрической энергии при среднеквадратической относительной погрешности 1,03 %.

Литература

1. Республиканская программа энергосбережения на 2001–2005 годы. – Минск : Ком. по энергоэффективности при Совете Министров Респ. Беларусь, 2001. – 27 с.
2. Токочакова, Н. В. Управление энергоэффективностью промышленных потребителей на основе моделирования режимов электропотребления / Н. В. Токочакова // Изв. высш. учеб. завед. и энерг. обь-ний СНГ – Энергетика. – 2006. – № 3. – С. 67–75.
3. Факторный, дискриминантный и кластерный анализ / Дж. Ким [и др.]; пер. А. М. Хотинского, С. Б. Королева ; под ред. И. С. Енюкова. – Москва : Финансы и статистика, 1989. – 215 с.
4. Олейников, В. К. Анализ и управление электропотреблением на металлургических предприятиях : учеб. пособие / В. К. Олейников, Г. В. Никифоров. – Магнитогорск : МГТУ им. Г. И. Носова, 1999. – 219 с.
5. Перегудов, Ф. И. Введение в системный анализ : учеб. пособие для вузов / Ф. И. Перегудов, Ф. П. Тарасенко. – Москва : Высш. шк., 1989. – 367 с.

6. Анищенко, В. А. Способ построения модели режимов электропотребления участка нефтепровода / В. А. Анищенко, Н. В. Токочакова, А. С. Фиков // Изв. высш. учеб. завед. и энерг. объ-ний СНГ – Энергетика. – 2006. – № 6. – С. 44–48.
7. Бородич, С. А. Вводный курс эконометрики : учеб. пособие / С. А. Бородич. – Минск : БГУ, 2000. – 354 с.
8. Андронов, А. М. Теория вероятностей и математическая статистика : учеб. для вузов / А. М. Андронов, Е. А. Копытов, Л. Я. Гринглаз. – Санкт-Петербург : Питер, 2004. – 461 с.
9. Элементарные понятия статистики // StatSoft Inc. [Электронный ресурс]. – 2001. – Режим доступа: <http://www.statsoft.ru/home/textbook/esc.html>. – Дата доступа: 10.07.2005.
10. Ван дер Варден, Б. Л. Математическая статистика / Б. Л. Ван дер Варден ; пер. Л. Н. Большева ; под ред. Н. В. Смирнова. – Москва : Изд-во иностр. лит., 1960. – 434 с.

Получено 15.01.2007 г