

Факторный анализ потребления электроэнергии на тягу поездов

А.С. Маниковский

Забайкальский институт железнодорожного транспорта, Чита

Аннотация: Рыночные отношения между потребителем и поставщиком электроэнергии влечет к увеличению обязательств для участников рынка. В условиях развития конкуренции на розничном рынке для энергосбытовой компании становится актуальной проблема эффективного распределения электроэнергии, приобретаемой на оптовом рынке. Прогноз электропотребления становится важным показателем для дальнейшего планирования спроса потребителей, уменьшении затрат на генерацию и транспортировку электроэнергии. В случае составления неточного прогноза предприятие вынуждено закупать или продавать электроэнергию по заранее невыгодной цене.

Применение многофакторных математических моделей в задаче прогнозирования позволяет получить прогноз электропотребления высокой точности. Автором предложено сложную систему потребления электроэнергии представить в виде уравнения множественной линейной регрессии, которое учитывает взаимосвязь между факторами и дает возможность выделить влияние того или иного фактора на выходную величину.

Ключевые слова: электропотребление, тяга поездов, прогнозирование электропотребления, множественная регрессия, коэффициент корреляции, парная корреляция, математическая модель, анализ остатков, коэффициент детерминации, критерий Фишера.

Введение

Прогнозирование электропотребления (ЭП) представляет исходные данные для планирования режимов работы, позволяющие оценить качественные параметры и рассчитать действительные режимы работы энергосистемы [1].

Согласно правилам функционирования розничных рынков (постановление Правительства РФ от 27.12.2010 № 1172 (ред. от 1.10.2020), потребители обязаны точно планировать объемы электропотребления. При отклонении в потреблении электроэнергии (ЭЭ) более 5 % от запланированного объема, предприятие несет дополнительные издержки [2].

В связи с этим, перед предприятием стоит задача составления достоверной заявки по потреблению ЭЭ. Это требует совершенствования методов и алгоритмов прогнозирования потребления ЭЭ с применением современных технологий обработки информации.

Моделирование системы электропотребления

Для повышения точности прогнозов необходимо ввести в модель прогнозирования возможность учета нескольких факторов, оказывающих влияние на величину ЭП. Таким образом, систему электропотребления необходимо представить математической моделью, в которой можно выделить значимые закономерности изучаемого процесса, который протекает в реальной системе [3].

Для определения влияющих факторов, необходимо систему электропотребления представить уравнением множественной линейной регрессии. Множественная регрессия используется в тех случаях, когда из множества факторов нельзя выделить один доминирующий, а необходимо учитывать влияние нескольких.

Процесс электропотребления можно представить уравнением (1) [4]

$$y=b_0+b_1x_1+b_2x_2+b_3x_3+...+b_nx_n, \quad (1)$$

где y – потребление ЭЭ;

x_1-x_n – влияющие факторы;

b_0-b_n – коэффициенты множественной регрессии.

Среди всех возможных факторов следует выделить наиболее значимые. Для оценки степени влияния каждого фактора на величину электропотребления необходимо рассчитать все возможные комбинации парных коэффициентов корреляции между потреблением ЭЭ y и факторами x .

Для расчета парных коэффициентов корреляции использовалась почасовая информация по показателям работы участка и потреблению ЭЭ тяговой подстанцией Белогорск Забайкальской железной дороги за пятидневный период 2019 г.

В результате анализа работ специалистов по вопросам планирования потребления ЭЭ выделены следующие показатели, оказывающие влияние на величину электропотребления [5, 6]:

- количество поездов, одновременно находящихся на участке;
- средняя масса состава;
- грузооборот участка;
- средняя участковая скорость;
- температура окружающей среды.

Расчет степени взаимосвязи между каждым фактором в отдельности и электропотреблением проведен в пакете анализа данных *Statistica*, с использованием встроенного инструмента «Матрица парных корреляций». Результаты расчета приведены в таблице 1.

Таблица № 1

Результаты расчета парных коэффициентов корреляции.

Переменная	Потребление ЭЭ, кВт·ч (y)	Средняя масса состава $m_{ср}$, т (x_1)	Количество поездов, одновременно находящихся на участке N (x_2)	Средняя участковая скорость $V_{уч}$, км/ч (x_3)	Грузооборот участка $\Sigma p \cdot L$, тыс. ткм брутто (x_4)	Температура окружающей среды (x_5)
Потребление ЭЭ, кВт·ч (y)		0,974	0,976	0,332	0,883	0,026
Средняя масса состава $m_{ср}$, т (x_1)			0,998	0,284	0,874	0,022
Количество поездов, одновременно находящихся на участке N (x_2)				0,289	0,701	0,028
Средняя участковая скорость $V_{уч}$, км/ч (x_3)					0,338	0,033
Грузооборот участка $\Sigma p \cdot L$, тыс. ткм брутто (x_4)						0,049
Температура окружающей среды (x_5)						

Исходя из результатов расчетов, можно заключить, что факторы x_1 , x_2 , x_4 наиболее значимые. Однако, высокий коэффициент корреляции между факторами x_1 и x_2 (мультиколлинеарность) нужно учитывать при построении регрессионной модели, поскольку могут возникнуть большие ошибки при вычислении коэффициентов регрессии. Для устранения мультиколлинеарности был исключен один фактор, и мы отдали предпочтение тому, который при тесной связи с результатом имеет наименьшую связь с другими факторами [7].

Для построения линейной регрессии будут приняты два фактора: количество поездов, одновременно находящихся на участке (x_2) и грузооборот участка (x_4). Остальные факторы принимать во внимание не следует, т.к. связь с величиной y несущественная.

Уравнение (1) примет вид

$$y=b_0+b_1x_1+b_2x_2.$$

Неизвестные коэффициенты b определялись с использованием инструмента «Множественная регрессия» в пакете *Statistica*.

Подставив значения коэффициентов в уравнение множественной регрессии, получим

$$y=21424,29+0,75957109x_1+07763534x_2.$$

На основании данных о количестве поездов, одновременно находящихся на участке и грузооборота участка, было спрогнозировано количество потребляемой электроэнергии. На рисунке 1 представлено сравнение прогнозных значений электропотребления и исходного.

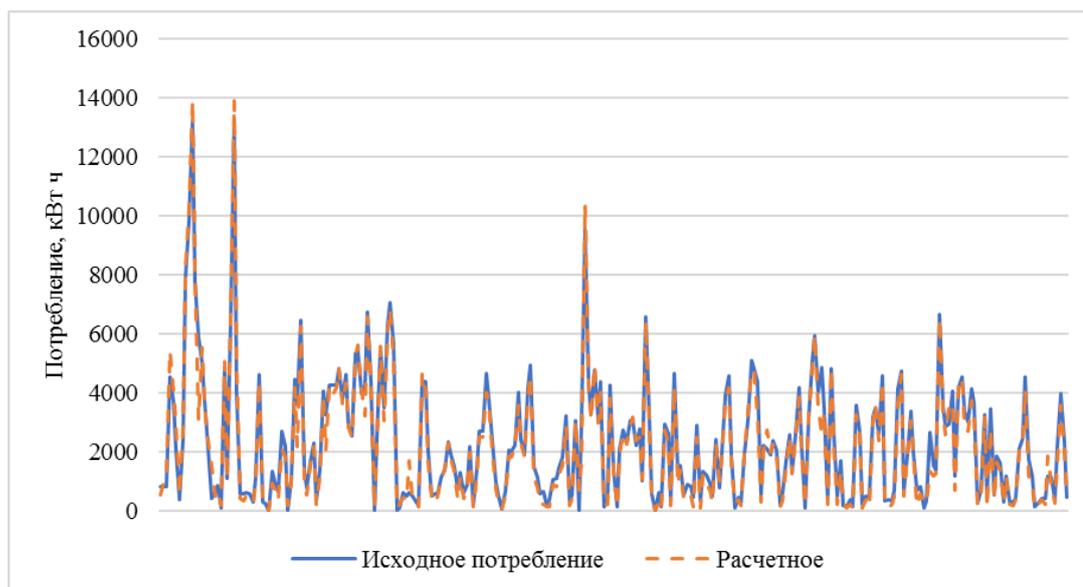


Рис. 1. – График исходного электропотребления и расчетного

Оценка точности построенной модели

Оценить точность построенного прогноза позволяет средняя относительная ошибка аппроксимации по формуле (2)

$$\delta = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \cdot 100\%, \quad (2)$$

где y_i – фактические значения ряда;

\hat{y}_i – спрогнозированные значения ряда;

n – количество элементов.

Полученная ошибка 3,7% свидетельствует о хорошем подборе модели к исходным данным.

Оценка качества построенной модели была проведена при помощи процедуры анализа остатков (рис.1, рис. 2).

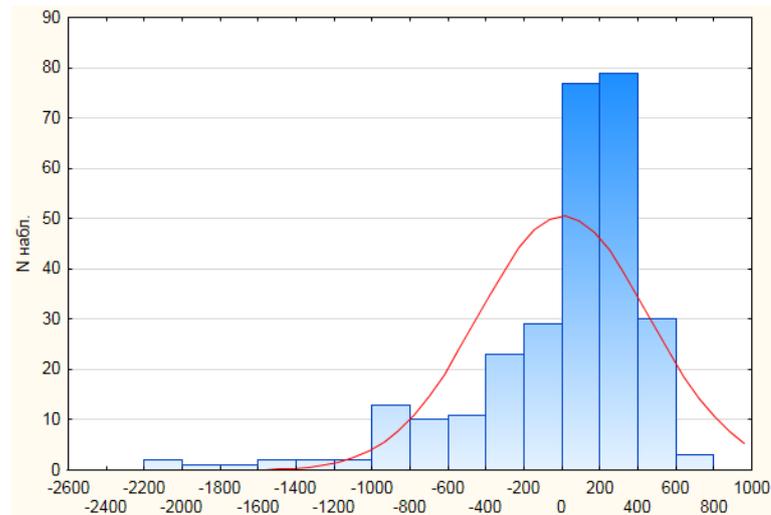


Рис. 2. – Гистограмма распределения остатков

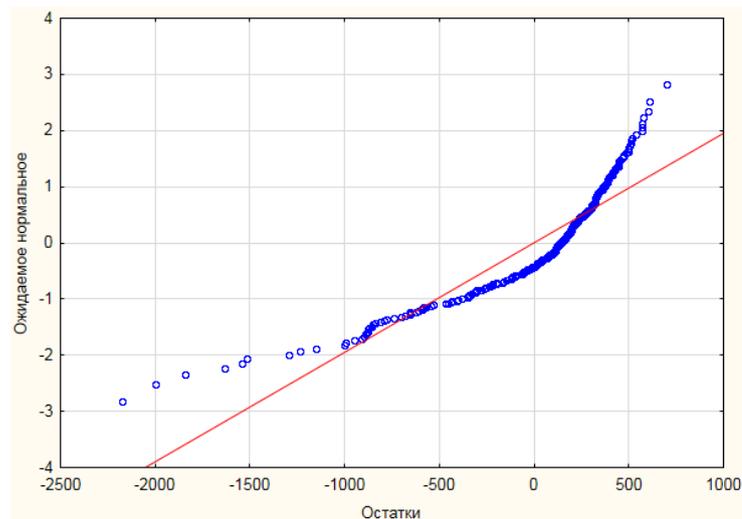


Рис. 3. – Нормальный вероятностный график остатков

Анализ гистограммы и графика позволяет заключить, что распределение остатков близко к нормальному закону распределения. Это говорит о хорошем качестве модели.

После анализа остатков был определен коэффициент детерминации, который позволяет проверить качество уравнения регрессии. Коэффициент детерминации показывает, какая доля дисперсии результативного признака

объясняется влиянием независимых переменных [8, 9], определяемы по формуле (3)

$$R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (3)$$

где e – отклонение фактического значения зависимой переменной от ее прогнозного значения;

y_i – фактическое значение переменной;

\bar{y} – среднее значение фактических значений.

Величина коэффициента детерминации составляет 0,93, т.е. в 93% случаев изменения факторов x приводят к изменению величины потребления электроэнергии y . Это говорит о том, что регрессионная модель соответствует данным.

Дать оценку значимости уравнения регрессии позволяет критерий Фишера [10].

Модель будет считаться адекватной если выполняется условие

$$F_{расч} > F_{табл.}$$

Расчетный коэффициент Фишера определяется по формуле (4)

$$F_{расч} = \frac{\sum (y_{iрасч} - y_{ср.расч})^2}{m} \cdot \frac{n-m-1}{\sum (y_i - y_{iрасч})^2}, \quad (4)$$

где $y_{iрасч}$ – расчетное значение электропотребления,

$y_{iрасч}$ – среднее значение электропотребления за исследуемый период,

y_i – фактическое значение электропотребления,

m – количество факторов, влияющих на электропотребление,

n – количество наблюдений.

Табличное значение коэффициента Фишера определяется при заданной вероятности 0,95, т.е. с вероятностью 95% убедиться, что полученная функция соответствует исходным данным.

Расчетное значение существенно превосходит табличное ($F_{расчл} = 399,1$; $F_{табл} = 3,027$), следовательно построенная модель является адекватной.

Вывод

Система электропотребления была представлена уравнением множественной линейной регрессии. Это позволяет определить факторы, влияющие на величину электропотребления. При помощи корреляционного анализа выявили, что количество поездов, одновременно находящихся на участке и грузооборот участка, оказывает существенное влияние на величину тягового электропотребления.

Определены коэффициенты уравнения множественной регрессии и построен прогноз на основании данных влияющих факторов.

Полученная средняя относительная ошибка аппроксимации (3,7%), говорит о хорошем подборе модели к исходным данным. Это же подтверждают анализ остатков, значение коэффициента детерминации и выполнение критерия Фишера. Построенная модель является адекватной и может быть использована для прогнозирования потребления электроэнергии. Располагая данными о воздействующих факторов (количество поездов, одновременно находящихся на участке и грузооборот участка) можно строить прогнозы, обладающие высокой точностью.

Литература

1. Староверов Б. А., Семенов И.В. Схемы взаимодействия поставщиков и получателей прогнозов электропотребления на основе использования нейросетевой информационной системы// Инженерный вестник Дона, 2018, №1. URL: ivdon.ru/ru/magazine/archive/n1y2018/4784.



2. Кретов Д. А., Рузанов Р.В. Прогнозирование электропотребления энергосбытовой компании с использованием искусственной нейронной сети // Инженерный вестник Дона, 2015, №2. URL: ivdon.ru/ru/magazine/archive/n2y2015/2880.

3. Шаюхов Т.Т. Математическое моделирование влияния внешних факторов на параметры электропотребления // Интернет-журнал «Науковедение» Том 9, 2017, №5.

4. Emmert-Streib, F.; Dehmer, M. Evaluation of Regression Models: Model Assessment, Model Selection and Generalization Error. *Mach. Learn. Knowl. Extr.* 2019, 1, pp.521-551.

5. Раевский Н. В., Литвинцев В.Г. Построение модели прогнозирования тягового электропотребления на основе интервальной регрессии // Современные технологии. Системный анализ. Моделирование / Иркутский гос. ун-т путей сообщения. Иркутск, 2010. № 4. С. 192 – 196.

6. Литвинцев В. Г. Влияние веса поезда на потери электрической энергии в контактной сети магистральных железных дорог переменного тока // Тр. четвертой междунар. науч. конф. творческой молодежи «Научно-техническое и экономическое сотрудничество стран АТР в XXI веке» / Дальневосточный гос. ун-т путей сообщения. Хабаровск, 2005. Т. I. С. 191 – 194.

7. Филатова Л.Ф. Множественная регрессия: Учебное пособие / Северск: Изд. СТИ НИЯУ МИФИ, 2013. – 50 с.

8. Джеймс Г., Уиттон Д., Хасты Т., Тибширани Р. Введение в статистическое обучение с примерами на языке R. Пер. с англ. С.Э. Мастицкого. – М.: ДМК Пресс, 2016. – 450 с.

9. Катаева Е.С. Макростатистический анализ и прогнозирование / Томск, 2016. – 56 с.

10. Кремер Н. Ш., Путко Б. А. Эконометрика: Учебник. 2-е изд. – М.: Юнити, 2008. – 328 с.

References

1. Staroverov B. A., Semenov I.V. Inzhenernyj vestnik Dona, 2018, №1. URL: ivdon.ru/ru/magazine/archive/n1y2018/4784.
2. Kretov D. A., Ruzanov R.V. Inzhenernyj vestnik Dona, 2015, №2. URL: ivdon.ru/ru/magazine/archive/n2y2015/2880.
3. Shayukhov T.T. Internet-zhurnal «Naukovedenie» Tom 9, 2017, №5.
4. Emmert-Streib, F.; Dehmer, M. Evaluation of Regression Models: Model Assessment, Model Selection and Generalization Error. *Mach. Learn. Knowl. Extr.* 2019, 1, 521-551.
5. Raevskiy N. V., Litvintsev V. G. Sovremennye tekhnologii. Sistemnyy analiz. Modelirovanie. Irkutskiy gos. un-t putey soobshcheniya. Irkutsk, 2010. № 4. pp. 192 – 196.
6. Litvintsev V. G. Tr. chetvertoy mezhdunar. nauch. konf. tvorcheskoy molodezhi «Nauchno-tekhnicheskoe i ekonomicheskoe sotrudnichestvo stran ATR v XXI veke». Dal'nevostochnyy gos. un-t putey soobshcheniya. Khabarovsk, 2005. T. I. pp. 191 – 194.
7. Filatova L.F. Mnozhestvennaya regressiya: Uchebnoe posobie. [Multiple Regression: A Tutorial]. Seversk: Izd. STI NIYaU MIFI, 2013. 50 p.
8. Dzheymys G., Uitton D., Khasti T., Tibshirani R. Vvedenie v statisticheskoe obuchenie s primerami na yazyke R. Per. s angl. S.E. Mastitskogo [An introduction to statistical learning with examples in R]. M.: DMK Press, 2016. 450 p.
9. Kataeva E.S. Makrostatisticheskyy analiz i prognozirovanie [Macrostatistical analysis and forecasting]. Tomsk, 2016. 56 p.
10. Kremer N. Sh., Putko B.A. Ekonometrika: Uchebnik. 2-e izd. [Econometrics: A Textbook. 2nd ed.] M.: Yuniti, 2008.