

Методика построения аналитической модели потребления природного газа

Ольга Башарина^{1,2}, Алексей Еделев^{1,3}, Татьяна Фер²

¹ Институт динамики систем и теории управления им. В.М. Матросова СО РАН, ул. Лермонтова, д. 134, 664033, Иркутск, Россия

² Уральский государственный экономический университет, ул. 8 марта, д. 62, 620144, Екатеринбург, Россия

³ Институт систем энергетики им. Л.А. Мелентьева СО РАН, ул. Лермонтова, д. 130, 664033, Иркутск, Россия

Аннотация

Представлена методика построения математической модели на основе анализа данных временных рядов потребления природного газа на территории Евросоюза. Методика включает три основных этапа: предварительный анализ данных; построение тренд-сезонной модели; оценка качества и точности модели. Использование методики продемонстрировано на примере и показывает хорошее качество и точность полученной аналитической модели.

Ключевые слова

Временной ряд, тренд-сезонная модель, гармонический анализ

1. Введение

Необходимость в разработке новых моделей систем газо-, нефте- и нефтепродуктоснабжения для исследований живучести основана, в частности, на однопериодности разработанных в ИСЭМ СО РАН моделей этих больших трубопроводных систем [1, 2, 3]. Следствиями этого недостатка являются отсутствие динамики и упрощенное моделирование экстремальных условий функционирования систем энергетики [3].

Новые модели систем газо-, нефте- и нефтепродуктоснабжения позволят более реалистично и точно моделировать функционирование этих систем энергетики в экстремальных условиях. При том серьезной проблемой является отсутствие в открытом доступе исходных данных, необходимых для построения многопериодных моделей больших трубопроводных систем, в частности для получения посуточных графиков потребления энергоресурсов. В [1] показан подход к решению задачи построения схемы газоснабжения Европейского союза на основе открытых данных, являющихся результатом проекта SciGRID_gas [4] Ольденбургского института сетевых энергетических систем.

В рамках проекта SciGRID_gas были получены и опубликованы массивы данных SciGRID_gas CONS [5] ежесуточного потребления природного газа за 10 лет для всех регионов стран Евросоюза на уровнях NUTS-1 (федеральный уровень), NUTS-2 (административные округа) и NUTS-3 (районы и города без районов).

Цель нашего исследования – разработать методику построения модели потребления газа для территориальных субъектов Евросоюза уровня NUTS-3.

6th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2024), July 01–05, 2024, Irkutsk, Russia

EMAIL: basharinaolga@mail.com (A. 1); flower@isem.irk.ru (A. 2); tatiana.fer@yandex.ru (A. 3)

ORCID: 0000-0002-7151-782X (A. 1); 0000-0003-3034-4907 (A. 2)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

ICCS-DE 2024 Workshop Proceedings

DOI: 10.47350/ICCS-DE.2024.10

2. Материалы и методология

Исходными данными послужили показатели ежедневного потребления природного газа в 1364 городах и районах стран Евросоюза за период 2010–2019 гг. с детализацией по промышленному, жилому и коммерческому секторам. В рамках нашего исследования интерес представляет суммарное потребление газа, таким образом исходные данные по каждому территориальному субъекту представлены в виде одномерного временного ряда, содержащего 3652 уровня.

Методика построения аналитической модели включает три основных этапа:

1. Предварительный анализ данных.
2. Построение тренд-сезонной модели.
3. Оценка адекватности и точности модели.

Предварительный анализ ряда проводится на основе графического представления анализируемых данных ежедневного потребления природного газа, а также с использованием статистических критериев. На данном этапе выполняется начальное описание ряда, предварительная обработка данных, включая устранение аномальных наблюдений, определение стационарности ряда.

Визуальный анализ графика временного ряда позволяет выдвинуть предположение, о том что ряд содержит трендовую $T(t)$, сезонную $S(t)$ и случайную $E(t)$ составляющие. Поскольку амплитуды колебаний уровней ряда примерно одинаковые, для моделирования ряда выбрана аддитивная модель:

$$Y(t) = T(t) + S(t) + E(t). \quad (1)$$

Проверка наличия тенденции и определение периодичности сезонных колебаний во временном ряде выполняется на основе анализа автокорреляционной функции уровней ряда и коррелограммы [6]. Такой анализ позволяет найти величину лага, при котором автокорреляция уровней ряда наиболее высокая.

Если наиболее высоким оказывается коэффициент автокорреляции первого порядка, то исследуемый ряд содержит только тенденцию. Если наиболее высоким оказывается коэффициент автокорреляции порядка p , то ряд содержит колебания с периодичностью в p моментов времени.

Для исследуемых временных рядов структура оказалась следующей: определено наличие линейного тренда и сезонности с периодом 365 дней.

Следующий этап представленной методики заключается в построении тренд-сезонной модели. При анализе графика временного ряда был сделан вывод о линейном характере тренда. Методом гармонического (спектрального) анализа, позволяющего представить периодические сезонные колебания в виде линейной комбинации синусов и косинусов, была определена гармоника под номером 1 с наиболее значимым спектром [6-7].

Таким образом, модель ежедневного потребления природного газа представляется суммой линейного тренда, сезонной компоненты с периодом 365, описываемой тригонометрическими функциями, и случайной составляющей:

$$Y(t) = a_0 + a_1 t + a_2 \cos \frac{2\pi t}{365} + a_3 \sin \frac{2\pi t}{365} + E(t), \quad (2)$$

где t – уровень ряда; a_0, a_1, a_2, a_3 – параметры уравнения регрессии, рассчитанные методом наименьших квадратов; $E(t)$ – случайная составляющая.

На заключительном этапе моделирования проводится оценка адекватности полученной модели по следующим характеристикам.

Коэффициент детерминации R^2 . Измеряет долю дисперсии, объясненную моделью, в общей дисперсии уровней временного ряда. Чем ближе коэффициент детерминации к единице, тем более хорошо полученная аналитическая модель соответствует исходному временному ряду.

F-критерий Фишера. Используется для оценки значимости модели «в целом». Расчетное значение F-критерия сравнивают с табличным при выбранном уровне значимости. Если выполняется неравенство $F_{\text{расч}} > F_{\text{табл}}$, то модель признается значимой, а количество наблюдений достаточным [6-8].

Точность модели оценивают с помощью ошибок аппроксимации. Средняя абсолютная ошибка (MAE, Mean Absolute Error) отражает среднюю абсолютную разницу между модельными (расчетными) и фактическими значениями временного ряда. MAE устойчива к выбросам и не усиливает большие ошибки. Средняя квадратическая ошибка (MSE, Mean Squared Error), напротив, выявляет грубые ошибки за счет того, что величина отклонения фактических уровней от расчетных возводится во вторую степень. На практике из величины MSE извлекают квадратный корень и получают RMSE (Root Mean Squared Error), называемую стандартной ошибкой. Средняя абсолютная процентная ошибка (MAPE, Mean Absolute Percentage Error) показывает среднее абсолютное отклонение расчетных значений от фактических в процентах и имеет простую интерпретацию.

3. Пример реализации методики

В качестве примера реализации представленной методики приведем процедуру построения модели потребления газа для территориального субъекта Греции CZ041.

На этапе предварительного анализа проведена проверка исходных данных, визуальный анализ графика временного ряда. Анализ автокорреляционной функции и коррелограммы временного ряда (рисунок 1) показал наличие в его составе тенденции и сезонной компоненты с периодом 365 дней.

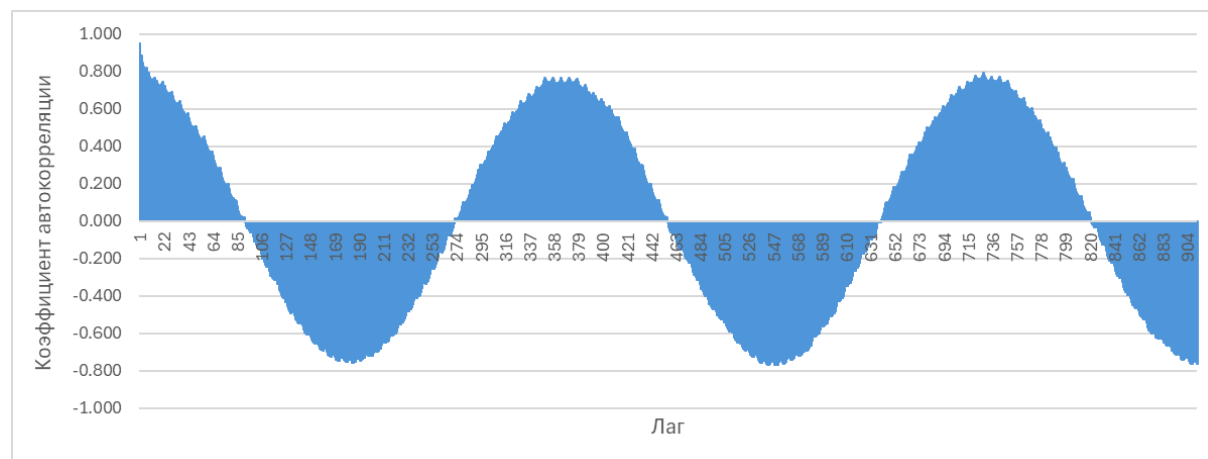


Рисунок 1: Коррелограмма временного ряда

В результате построения тренд-сезонной модели в соответствии с (1) получена оценка ежесуточного потребления газа:

$$\hat{Y}(t) = 0,487 - 0,00001t + 0,176 \cos \frac{2\pi t}{365} + 0,056 \sin \frac{2\pi t}{365}, \quad (2)$$

где t – номер дня.

На рисунке 2 представлен график сравнения фактических и расчетных значений объема потребления газа.

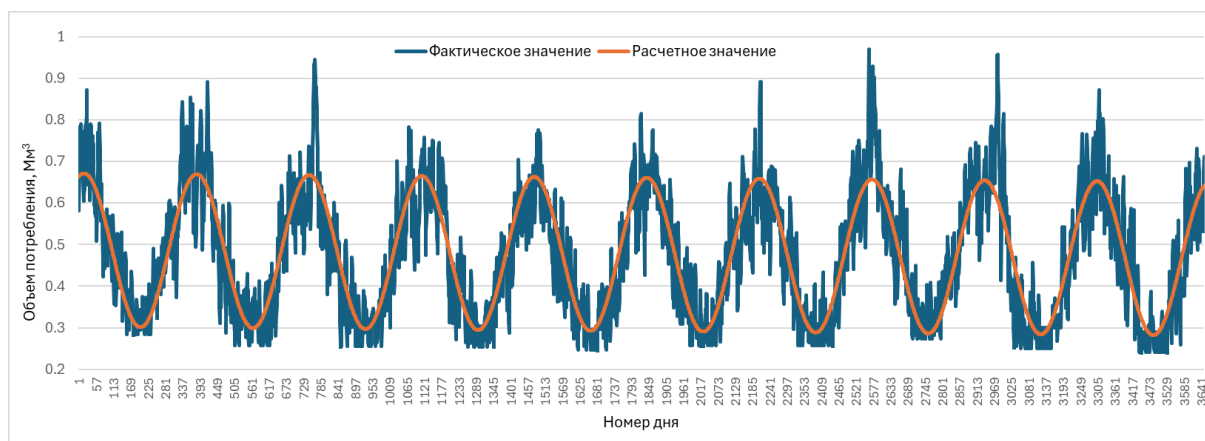


Рисунок 2: Ежесуточный объем потребления газа (фактическое и расчетное значения)

График остаточной компоненты построенной модели представлен на рисунке 3. Судя по графику, остаточная компонента носит случайный характер

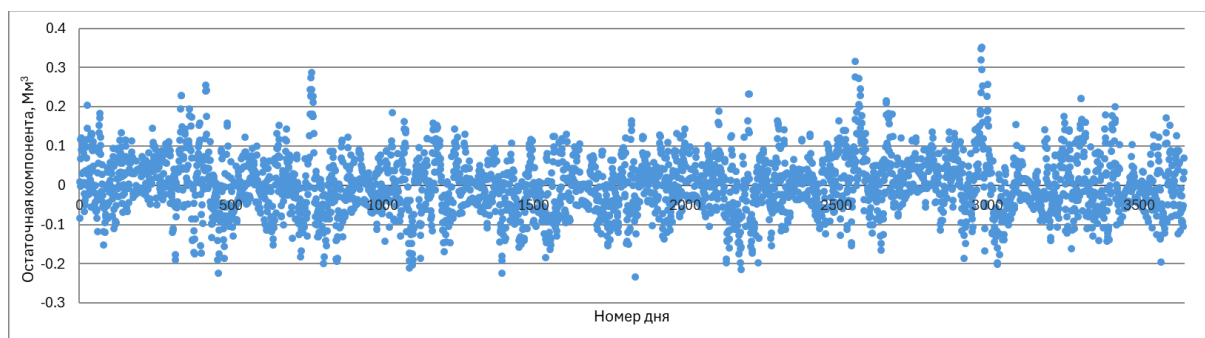


Рисунок 3: Остаточная компонента временного ряда

Коэффициент детерминации $R^2 = 75.8\%$. Построенная тренд-сезонная модель объясняет 75.8% вариации потребления природного газа, а остальные 24.2% приходятся на влияние внешних факторов.

При проверке значимости модели в целом получено расчетное значение F-статистики $F_{\text{расч}} = 3813$, значительно превышающее табличное значение $F_{\text{табл}} = 2.6$ для уровня значимости 0.05. Модель признается значимой в целом с вероятностью 95% и пригодной для прогнозирования.

При оценке точности модели рассчитаны значения ошибок: $\text{MAE} = 0,057 \text{ Мм}^3$; $\text{RMSE} = 0.074 \text{ Мм}^3$; $\text{MAPE} = 12.4\%$. Точность модели можно считать удовлетворительной.

4. Заключение

Представленная в статье методика позволяет построить тренд-сезонную модель потребления природного газа. Результаты оценки адекватности и точности модели дают возможность использовать ее для анализа и прогнозирования потребления газа для территориальных субъектов Евросоюза уровня NUTS-3. Модель направлена на реализацию в распределенной вычислительной среде для ускорения расчетов по множеству потребителей. Описанная методика частично реализована в виде сервиса.

Дальнейшее направление исследования связано с проведением анализа зависимости потребления газа от природно-климатических показателей.

5. Благодарности

Работа выполнена в рамках гранта № 075-15-2024-533 Министерства науки и высшего образования РФ на выполнение крупного научного проекта по приоритетным направлениям научно-технологического развития (проект «Фундаментальные исследования Байкальской природной территории на основе системы взаимосвязанных базовых методов, моделей, нейронных сетей и цифровой платформы экологического мониторинга окружающей среды», рег. № 124052100088-3).

6. Список литературы

- [1] А. Еделев, Г. Данилов, О. Башарина, Разработка модели системы газоснабжения на основе открытых данных для исследования живучести, Материалы VI Международного семинара по информационным, вычислительным и управляющим системам для распределенных сред (ICCS-DE 2024, 1-5 июля 2024 г., г. Иркутск, Россия), Иркутск, Изд-во ИДСТУ СО РАН, 2024, С. 103-110.
- [2] А.В. Храмов, С.М. Еникеева Программное и информационное обеспечение решения задач живучести Единой системы газоснабжения СССР, Методы и модели исследования живучести систем энергетики, Новосибирск: Наука, Сиб. Отд, 1990, С. 86-91.
- [3] С.М. Клименко, С.М. Сендеров, В.А. Янченко, Исследование проблем повышения устойчивости и экологической безопасности магистральных нефте- и нефтепродуктопроводов, Новые информационные технологии управления развитием и функционированием трубопроводных систем энергетики, Иркутск: СЭИ СО РАН, 1993, С. 119-127.
- [4] A. Pluta, W. Medjroubi, J.C. Diettrich, J. Dasenbrock, H.P. Tetens, J.E. Sandoval, O. Lünsdorf, SciGRID_gas-Data Model of the European Gas Transport Network, In 2022 Open Source Modelling and Simulation of Energy Systems (OSMSES), 2022, April, pp. 1-7.
- [5] Jan Diettrich, Adam Pluta, Wided Medjroubi. SciGRID_gas IGGIELGN, Version 1.1.2. Zenodo, May 2021. doi: 10.5281/zenodo.4767098.
- [6] В. А. Алексеева, Анализ временных рядов : уч. пособие, Ульяновск : УлГТУ, 2020, 147 с.
- [7] Н. А. Дегтярева, Модели анализа и прогнозирования на основе временных рядов: монография, Челябинск: Изд-во ЗАО «Библиотека А.Миллера, 2018, 160 с.
- [8] Е. В. Радковская, Е. М. Кочкина, М. В. Дроботун, Т. В. Фер, Н. П. Попова, И. В. Иванов, Эконометрика, США, Open Science Publishing, 2019, 177 с., ISBN 978-0-244-47834-6.