

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА И
ГОСУДАРСТВЕННОЙ СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(РАНХиГС)

Касьянова К.А., Каукин А.С., Миллер Е.М.

МОДЕЛИРОВАНИЕ СПОТОВЫХ ЦЕН НА ЭЛЕКТРОЭНЕРГИЮ НА ОПТОВОМ
РЫНКЕ В РОССИИ
(препринт)

Москва 2020

Аннотация

В ходе проводимых реформ в электроэнергетике был сформирован оптовый рынок электроэнергии и мощности. Целью данного исследования является разработка модели ценообразования на рынке на сутки вперед (далее РСВ). Для моделирования цен на электричество традиционно используется два подхода: первый – структурные модели, например, модели равновесия или оптимизационные модели, второй – модели временных рядов, причем популярной является модель, рассматривающая цену на электричество как диффузионно-скачкообразный процесс, что мотивировано наличием тяжелых хвостов у эмпирического распределения цен на этом рынке. Недостатками первого подхода является необходимость введения определенных предпосылок об агентах, участвующих на рынке, и упрощений, позволяющих получить численный ответ без использования сложных оптимизационных алгоритмов. Другой проблемой является невозможность учесть в модели случайную компоненту. Последняя решается с помощью второго подхода, когда особенности эмпирического распределения учитываются внутри модели временных рядов, например, добавлением в модель ненаблюдаемых компонент, отвечающих за частоту скачков и тенденцию возвращения процесса к среднему, либо моделированием изменяющейся во времени дисперсии с помощью моделей авто-регрессионной условной гетероскедастичности (модели типа GARCH) или с помощью моделей с марковскими переключениями, позволяющих уловить моменты переходов цены из одного «режима» в другой. Однако модели временных рядов заточены под задачу прогнозирования и часто ограничивается использованием одного-двух наиболее значимых факторов. Но в отличие от цен на акции, цены на электроэнергию связаны со множеством наблюдаемых факторов (потребление, цены на топливо, и др.), которые могут быть включены в модель.

В рамках данного исследования была разработана модель, позволяющая учесть недостатки двух классов моделей: цены на оптовом рынке электроэнергии рассматриваются как диффузионно-скачкообразный процесс возвращающий среднее, но при этом учитываются также фундаментальные факторы спроса и предложения (температура, уровень деловой активности и другие), факторы, влияющие на сезонную и трендовую компоненту в ценах, а также особенности российского рынка.

Структура работы представлена четырьмя разделами. В первом разделе представлен обзор мирового опыта моделирования рынка электроэнергии, акцент сделан на рассмотрении моделей, описывающих динамику цен. Во втором разделе представлен анализ современного состояния и основные проблемы развития рынка электроэнергии России. В третьем разделе представлен эмпирический анализ предложенной модели, ее калибровка. В четвертом разделе описаны полученные результаты и их интерпретация.

Ключевые слова: энергетика, цены на электроэнергию, рынок на сутки вперед, диффузионно-скачкообразный процесс.

Abstract

The aim of the study is to simulate the dynamics of electricity prices in the market for the day ahead. Within the framework of this study, a model was developed that allows one to take into account the disadvantages of two classes of models: prices on the wholesale electricity market are considered as a diffusion-step process that returns the average, but this also takes into account the fundamental factors of supply and demand (temperature, level of business activity, etc.), factors influencing the seasonal and trend component in prices, as well as the specifics of the Russian market. The structure of the work is presented in four sections. The first section provides an overview of the experience in modeling the electricity market, with an emphasis on considering models that describe price dynamics. The second section presents an analysis of the current state and the main problems of the development of the Russian electricity market. The third section presents an empirical analysis of the proposed model and its calibration. The fourth section describes the results obtained and their interpretation.

Key words: energy, electricity prices, day-ahead market, diffusion-jump process.

СОДЕРЖАНИЕ

| | |
|--|----|
| ВВЕДЕНИЕ | 4 |
| 1 ОБЗОР МИРОВОГО ОПЫТА МОДЕЛИРОВАНИЯ РЫНКОВ ЭЛЕКТРОЭНЕРГИИ | 5 |
| 2 ОСОБЕННОСТИ РОССИЙСКОГО РЫНКА ЭЛЕКТРОЭНЕРГИИ | 23 |
| 3 ЭМПИРИЧЕСКИЙ АНАЛИЗ И КАЛИБРОВКА МОДЕЛИ ЦЕН НА ЭЛЕКТРИЧЕСТВО НА РЫНКЕ НА СУТКИ ВПЕРЕД | 31 |
| 4 РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ И ИХ ИНТЕРПРЕТАЦИЯ | 64 |
| ЗАКЛЮЧЕНИЕ..... | 87 |
| СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ | 89 |

ВВЕДЕНИЕ

Характерной чертой российской энергетической системы является ее сложная географическая структура, обусловленная как неоднородным расположением месторождений первичных источников энергии, так и исторически сложившейся концентрацией населения, производства и потребления конечной продукции системы (топлива, тепла и электричества). Помимо этого, преобладает технологическая неоднородность в элементах системы: из-за разнообразия перерабатываемых первичных источников энергии (нефть, природный газ, уголь, ВИЭ); из-за длительной истории существования системы; и из-за, по большей части, нескоординированного развития отдельных элементов системы с другими.

Целью работы является разработка модели, учитывающей недостатки двух классов моделей: моделирование цены на оптовом рынке электроэнергии как диффузионно-скаккообразного процесса с учётом фундаментальных факторов спроса и предложения (температура, уровень деловой активности и других), особенностей российского рынка.

Объектом исследования является электроэнергетическая отрасль Российской Федерации.

В работе были применены различные статистические и аналитические методы сбора и интеграции данных, включая автоматизированные самообучающиеся методы. Анализ собранной статистики осуществлялся с использованием современных методов эконометрического моделирования.

Результаты данной работы могут быть использованы в интересах:

- для научно-методологического обеспечения прикладных исследований и моделирования рынков электроэнергии, в частности совершенствования методов статистического анализа случайных процессов;
- для разработки конкретных направлений мер и механизмов государственной политики в сфере регулирования рынка электроэнергии.

1 Обзор мирового опыта моделирования рынков электроэнергии

Рынок электроэнергии в отличие от других энергетических рынков обладает рядом особенностей. Во-первых, возникает проблема невозможности хранения производимого товара, что может учитываться при моделировании цены фьючерсного контракта на электричество, поскольку невозможно открыть короткую позицию.

Во-вторых, любой производитель электроэнергии решает оптимизационную задачу по выбору состава работающего оборудования (unit commitment problem), заключающуюся в том, что спрос на электричество должен быть удовлетворен при минимальных затратах; при этом учитывается тот факт, что для многих производителей при очень низком спросе или в его отсутствии в определенные дни выгоднее продавать электричество по нулевой или близкой к нулю цене, так как издержки на остановку генератора и его перезапуск впоследствии могут быть намного выше, чем краткосрочное снижение выручки.

В-третьих, может наблюдаться проблема с ограничениями ЛЭП, в случае повышенного спроса в конкретном узле может наблюдаться перенапряжение сети, в таком случае локальные узловые цены могут сильно отличаться от общеустановленных по системе; такая проблема обычно решается единым системным оператором.

В-четвертых, если цены на электричество определяются на конкурентном рынке, в России на рынке на сутки вперед (РСВ), то между принятием решений агентами рынка и системным оператором существует лаг, также отсутствует непрерывность торговли, а решения по равновесным ценам на каждый час будущих суток принимаются на основании одного и того же информационного множества; в России, например, эти цены могут быть скорректированы на балансирующем рынке (БР), однако, системный оператор поощряет тех, кто правильно оценивает объемы потребления на все следующие сутки и штрафует тех, кто отклоняется.

В-пятых, для цен на электричество в отличие от других производственных товаров характерна невозможность перераспределить волатильность цен по производственной цепочке.

В-шестых, электроэнергетическая отрасль довольно инертна, принятые решения могут достаточно долго внедряться повсеместно, инвестиционная деятельность направленная по большей части на модернизацию устаревшего оборудования, может усложняться социальными причинами, и т.д.; все это при долгосрочном накоплении проблем может вылиться в энергетический кризис, причинами которому могут стать изменения налогообложения, рыночные манипуляции, устаревшая инфраструктура,

провалы рынка, излишняя зарегулированность, перебои с поставками топлива, резкое изменение климата, доставка электричества дешевле стоимости производства.

Для моделирования цен на электричество традиционно используется два подхода: первый – структурные модели, например, модели равновесия [1] или оптимизационные модели [2], второй – модели временных рядов, причем популярной является модель, рассматривающая цену на электричество как диффузионно-скачкообразный процесс [3], [4], что мотивировано наличием тяжелых хвостов у эмпирического распределения цен на этом рынке. Более подробная классификация представлена на рисунке 1.



Примечание: Источник – составлено авторами.

Рисунок 1 – Классификация подходов к моделированию различных компонент энергосистем

Очевидно, выбор метода зависит от целей моделирования. Например, временные ряды традиционно используются в решении задачи прогнозирования. Структурные модели, учитывающие особенности динамических систем, нужны для выявления неэффективностей рыночного механизма всей энергетической системы. Однако разбиение на классы моделей оценки отдельных компонент энергетической системы довольно условно, поскольку многие методы пересекаются между собой, а авторы работ, посвящённых этому рынку, используют достижения и преимущества как одного класса моделей, так и другого.

Рассмотрим основные подклассы, начиная со структурных моделей. Во-первых, к этому классу можно отнести оптимизационные задачи. Например, в статье [5] структура

энергосети основывается на решении задачи оптимизации индивидуальных целей заинтересованных сторон – участников рынка. Такие модели называются агентскими моделями, или моделями поведения агентов, в которых каждый игрок моделируется как отдельный агент, нацеленный на конкретную цель, например, уменьшение своих расходов на покупку энергии или увеличение своих доходов от продажи энергии. Текущее состояние системы, в рамках которых работают агенты, является неопределенным из-за стохастичности эксплуатационных параметров и параметров окружающей среды, а также возможности технических неполадок и невозможности производства или перепроизводства энергии возобновляемыми источниками энергии, связанных с погодными условиями. Неопределенность эксплуатационных параметров и параметров окружающей среды энергосети в рамках данной модели определяются количественно в терминах интервалов прогнозирования (РІ) с помощью нейронных сетей, обученной с использованием недоминируемого генетического алгоритма сортировки (NSGA-II) [6]. Недоминируемое решение – это решение, которое обеспечивает подходящий компромисс между всеми целями, не умаляя ни одну из них, по концепции схожее с доминированием по Парето, состоящее в сравнении каждого решения x с любым другим решением в выборке, пока в нем не находится решение, которое доминирует среди них. С учетом всех неопределенностей каждый агент ищет оптимальное целенаправленное планирование действий с помощью робастного алгоритма оптимизации. Показано, что такая структура ведет к увеличению производительности системы, оцениваемой с точки зрения типичных показателей надежности (адекватности) для энергетических систем, таких как ожидаемые потери в нагрузке или передаваемой энергии.

Во-вторых, к структурным моделям относятся модели равновесия. Так, например, в статье [7] сравниваются два подхода к анализу конкурентных рынков электричества с помощью игрового подхода, равновесий Нэша, и описанного выше агентского моделирования. Для последнего используется предпосылка о том, что поставщики электроэнергии подают свои заявки на рынок, чтобы максимизировать свои прибыль. С учетом этого, с помощью методов машинного обучения, а именно обучения с подкреплением, оценивается модель поведения агента. Механизм клиринга рынка основан на схеме локальных предельных цен. В итоге, авторами показано, что эволюция в рамках модели поведения агентов напрямую связано с существованием уникального равновесия Нэша или множественных равновесий в системе.

В-третьих, также можно выделить класс динамических систем. Например, в работе [8] отмечено, что в национальной системе электроснабжения любой страны часто существует большой потенциал для повышения экономической эффективности

использования электроэнергии. Однако если экономические стимулы для улучшения использования системы слишком слабы, скорее всего, этот потенциал не будет использован. Анализ этой системы выполняется с помощью имитационной модели использования электроэнергии в промышленных энергетических системах. Оптимизация проводится на основе методов линейного программирования.

Второй подход к моделированию компонент энергосистемы заключается в применению различных статистических методов анализа временных рядов с целью определения и оценки различных особенностей энергосистем.

На большинстве конкурентных рынков электроэнергии временные ряды моделирующие цены на электричество характеризуются следующими свойствами:

- высокая частотность данных, поскольку системные операторы публикуют информацию о текущих ценах за каждый час,
- отсутствие стационарности: среднее и дисперсия непостоянны,
- множественная сезонность, цены характеризуются похожей динамикой в течение дня (обычно день делится на пиковые и непиковые часы), недели (цены стабильно ниже в выходные дни из-за снижения спроса), года (зимой может наблюдаться рост спроса на электричество за счет отопительного сезона, также иногда отмечается резкий всплеск в январе),
- календарный эффект (например, выходные и праздничные дни),
- высокая волатильность,
- высокая вероятность кратковременных скачков цен (в основном в периоды высокого спроса).

В дополнение к тенденции к возвращению среднего значения и сильного влияния сезонности на годовом, недельном и дневном уровне, спотовые цены на электроэнергию демонстрируют нечастые, но большие скачки. Спотовая цена может увеличиться в десять раз за один час. Скачки спотовых цен являются следствием колебаний экстремальной нагрузки, вызванных суровыми погодными условиями, часто в сочетании с перебоями в работе генераторов или отказами передачи. Эти пики, как правило, довольно недолгосрочны, и как только влияние погодных явлений или перебои в работе прекращаются, цены возвращаются к нормальному уровню.

Во-первых, классическим подходом к анализу временных рядов являются авторегрессионные модели (AR). Однако, из-за статистических особенностей эмпирического распределения рядов, связанных с энергетическим рынком, они редко используются в чистом виде.

Альтернативные модели в этом классе – авторегрессионные модели с регрессорами (ARX), и пороговые авторегрессионные модели (TAR). В качестве примера, в статье [9] спотовые цены на электроэнергию моделируются с помощью TAR-моделей. Предварительно ряды цен и спроса анализируются с применением вейвлет-анализа, позволяющего выявить, что является рыночным сигналом, а что шумом. Особенность ряда цен, а именно, возвращение к среднему, учитывается при очистке данных от высоких частот. Используя отфильтрованные данные, различные модели временных рядов рассматриваются в качестве возможных кандидатов для объяснения процесса ценообразования. В результате было показано, что в классе авторегрессионных моделей модели из класса TAR дают прогнозы, которые лучше всего отражают средние и дисперсионные компоненты фактических данных.

Наиболее популярными моделями в классе авторегрессионных моделей, применяемых к оцениванию и прогнозированию компонент энергетического рынка, являются обобщенные авторегрессионные модели условной гетероскедастичности (GARCH). В эконометрике ARCH модели применяются для анализа временных рядов, в которых дисперсия ошибки текущего члена описывается как функция от фактических значений размеров дисперсий ошибок предыдущих периодов времени и других факторов. В GARCH моделях дисперсии ошибок моделируются как авторегрессия со скользящим средним (ARMA). Чаще всего модели ARCH обычно используются при моделировании финансовых временных рядов, которые демонстрируют изменяющуюся во времени волатильность и кластеризацию волатильности, то есть периоды колебаний, чередующиеся с периодами относительного спокойствия. По этой же причине они популярны и при моделировании динамики компонент энергосистемы. Модели типа ARCH иногда считаются принадлежащими к семейству моделей стохастической волатильности, хотя это строго неверно, поскольку в момент времени t волатильность полностью предопределена (детерминирована) с учетом предыдущих значений.

В качестве примера использования GARCH- моделей для анализа энергетического рынка рассмотрим статью [10], в которой динамика цен на электроэнергию на нерегулируемых рынках анализируется с помощью специального класса моделей, который одновременно учитывает несколько факторов: сезонность, возврат к среднему, поведение GARCH и скачки, зависящие от времени. Восемь различных вложенных моделей были оценены для сравнения относительной важности каждого фактора, включенного в модель. В итоге на примере восьми различных рынков, даже после поправки на сезонность, показано, что цены на электроэнергию действительно имеют

тенденцию возвращения к среднему значению, с кластеризацией волатильности по времени (GARCH-компоненты) и скачками с интенсивностью, зависящей от времени.

Во-вторых, при анализе структуры временных рядов, наряду с применением вейвлет-анализа и различных фильтров для выделения низкочастотных компонент рядов, используются модели экспоненциального сглаживания. Например, в [11] для краткосрочного прогнозирования нагрузки энергосистемы были разработаны одномерные методы, которые моделируют внутринедельные и внутридневные сезонные циклы в данных о внутридневной нагрузке. В работе сравниваются три метода, ARMA с двойной сезонностью, адаптированный алгоритм экспоненциального сглаживания Холта – Уинтерса учитывающий двойную сезонность и недавно предложенный метод экспоненциального сглаживания. В данных о нагрузке за несколько лет, помимо внутридневных и внутринедельных циклов, также можно выделить внутригодичный сезонный цикл. Авторы разработали расширенную модель двойных сезонных методов, чтобы учесть в том числе и внутригодичный сезонный цикл. Показано, что для прогнозирования на день вперед «тройные» сезонные методы превосходят «двойные» сезонные, а также одномерный подход нейронной сети. Также отмечено, что дальнейшее повышение точности прогнозов достигается за счет использования комбинирования прогнозов из двойных и тройных сезонных методов.

Важным для анализа высокочастотных временных рядов являются модели тренд-сезонного разложения, поскольку, в отличие от цен на акции, ряды по электричеству могут обладать ярко выраженной сезонностью на трех уровнях (день, неделя, год), а также иметь трендовую компоненту. Далее подробно будут рассмотрены два подхода, используемых для моделирования трендовой и сезонной компоненты в рядах по электричеству. Стоит отметить основной недостаток таких моделей. В них "случайная" компонента моделируется как стационарный ряд, что на самом деле не так, что подтверждается в теоретических работах, моделирующих ряды по электричеству с помощью моделей стохастической волатильности, а также эмпирическими наблюдениями.

В-третьих, при анализе компонент энергосистем используются модели, учитывающие стохастическую структуру временных рядов, такие как модели стохастической волатильности, моделирование временных рядов как диффузионно-скачкообразного процесса и моделирование рядов с марковскими переключениями. Так при анализе высокочастотных данных по рынку электричества, например, на РСВ, можно отметить, что эмпирическое распределение цен на электричество имеет тяжелые хвосты.

Часто в литературе процесс, описывающий цену на электричество, задается, как диффузионно-скачкообразный процесс (jump-diffusion process) или модель Мертона [3].

В модели [12] так же учитывается нейтральность к риску, поэтому процесс, описывающий изменение цены на электричество – это риск-нейтральный диффузионно-скачкообразный процесс, представленный уравнением (1):

$$\frac{dS_t}{S_t} = (r - \lambda \cdot \bar{k})dt + \sigma dW_t + kdq_t \quad (1)$$

где

S_t – цена на электричество,

r – изменение цены,

λ – параметр Пуассоновского процесса q_t ,

\bar{k} – среднее случайного скачка k , имеющая логнормальное распределение,

представленной уравнением (2), где $\bar{k} = E(k) = e^{\gamma + \frac{\delta^2}{2}} - 1$,

t – момент времени,

σ – параметр, отвечающий за волатильность диффузионной компоненты (при $\lambda = 0$ получаем классическую модель Блэка-Шоулза),

W_t – Винеровский процесс,

k – величина случайного скачка.

$$\ln(1 + k) \sim N(\gamma, \delta^2) \quad (2)$$

Далее детали оценивания этой модели будут рассмотрены подробнее.

Такие модели очень популярны, поскольку оценка скачкообразной компоненты позволяет выявить выбросы, наблюдаемые в данных с учетом всех остальных факторов модели. С другой стороны, в классической модели Мертона предполагается постоянство параметра волатильности.

Однако, стоит отметить, что хотя многие исследователи рассматривают цены на рынках электричества как процессы со стохастическими скачками, в работе [13] показано, что такая спецификация может привести к потенциальным проблемам с указанием истинной величины среднего скачка в этом процессе. В качестве альтернативы, в этой статье предлагается модель переключения режимов, которая моделирует скачки цен без предположения о структуре процесса, порождающего данные.

В-четвертых, из-за сложной нелинейной структуры компонент энергосистемы популярны для анализа также и классические методы машинного обучения, например, метод опорных векторов, и модели нейронных сетей, в частности, рекуррентные нейросети (RNN) и нейросети с прямой связью (Feedforward NN). Например, в [14] метод опорных векторов применяется для задачи краткосрочного прогнозирования нагрузки энергосистемы. Результаты показывают, что метод опорных векторов выигрывает по сравнению с авторегрессионной моделью в смысле среднеквадратичных ошибок между фактическими и прогнозными данными.

В работе [15] модели нейронных сетей (ANN) применяются для прогнозирования цены на электроэнергию на сутки вперед. Поскольку эффективность прогноза по ANN зависит от выбора соответствующих наборов входных параметров, основное внимание уделяется отбору и подготовке данных о фундаментальных факторах, которые оказывают заметное влияние на цены на электроэнергию. Выбор регрессоров проводится с помощью различных кластерных алгоритмов, а также путем сравнения результатов предварительно выбранных конфигураций модели в сочетании с различными настройками входных параметров. После определения оптимальных входных параметров, влияющих на цены на электроэнергию на сутки вперед, модель ANN применяется для прогнозирования в выборке и вне выборки. Результаты показывают, что данная методология приводит к хорошим прогнозам цен на электроэнергию в том смысле, что ошибки прогноза столь же низки или даже ниже, чем по другим моделям цен на электроэнергию.

Недостатками структурных моделей является необходимость ведения определенных предпосылок об агентах, участвующих на рынке, и упрощений, позволяющих получить численный ответ без использования сложных оптимизационных алгоритмов. Другой проблемой является невозможность учесть в модели случайную компоненту. Последняя решается с помощью второго подхода, когда особенности эмпирического распределения учитываются внутри модели временных рядов, например, добавлением в модель ненаблюдаемых компонент, отвечающих за частоту скачков и тенденцию возвращения процесса к среднему [12], либо моделированием изменяющейся во времени дисперсии с помощью моделей автогрессионной условной гетероскедастичности (модели типа GARCH) [10] или с помощью моделей с марковскими переключениями, позволяющих уловить моменты переходов цены из одного “режима” в другой [16]. Однако модели временных рядов заточены под задачу прогнозирования и часто ограничивается использованием одного-двух наиболее значимых факторов.

В отличие от процесса оценки цен на акции, который трудно прогнозируем, цены на электроэнергию связаны с множеством наблюдаемых факторов: совокупное

потребление, цены на топливо, простои предприятий и т.д. Поэтому логичным является использование этих факторов в регрессии, а в рамках данной работы разрабатывается модель для российской энергосистемы, позволяющая учесть недостатки двух классов моделей: цены на оптовом рынке электроэнергии рассматриваются как диффузионно-скачкообразный процесс возвращающий среднее, но при этом учитываются также фундаментальные факторы спроса и предложения (температура, уровень деловой активности и др.), факторы, влияющие на сезонную и трендовую компоненту в ценах, а также особенности российского рынка.

Рассмотрим подробнее, один из примеров моделей цен на электричество, учитывающей также помимо лагированных значений цен, фундаментальные факторы спроса и предложения, а именно, температуру, дамми на выходные и понедельник, учитывающие недельную динамику деловой активности, а также эффект рычага, отражающий то, что большим значениям логарифма цены соответствуют большие значения дисперсии, является модель стохастической волатильности SVDEJX [12], которую мы рассмотрим подробнее далее.

Модель стохастической волатильности [12]. Для прогнозирования цен на электричество используется модель с экзогенными переменными (температура) и дамми-переменными на выходные и понедельник (далее SVDEJX, stochastic volatility model with a double exponential distribution of jumps, a leverage effect and exogenous variables).

В отличие от классической диффузионно-скачкообразной модели Мертона в этой модели учитывается изменение волатильности во времени, то есть модель различает периоды более высокой и низкой волатильности.

Особенностью данной модели является наличие ненаблюдаемых компонент:

— скачков вверх и вниз, пример представлен в уравнение (3):

$$q_t = \begin{cases} 1, & p_D \\ 0, & p_0 \\ -1, & p_U \end{cases} \quad (3)$$

где величины скачков вверх и вниз распределены экспоненциально, с разными параметрами $\xi_t^U \sim \text{Exp}(\eta_U) \text{ i.i.d.}, \xi_t^D \sim \text{Exp}(\eta_D) \text{ i.i.d.}$, ненаблюдаемое значение прироста логарифма цены,

— $h_{t-1} = y_t - y_{t-1}$, где $y_t = \ln(S_t)$ – логарифм цены, необходимое для оценки эффекта возвращения к среднему и эффекта рычага, когда большим значениям логарифма цены соответствуют большие значения дисперсии логарифма цен.

Оцениваемая SVDEJX модель задается следующей системой уравнений (смотри уравнения (4)-(6)):

$$y_t = y_{t-1} + \mu + \psi X_t + d_{Sat} D_{Sat,t} + d_{Sun} D_{Sun,t} + d_{Mon} D_{Mon,t} + \sqrt{\exp(h_{t-1})} \epsilon_t^{(1)} + J_t; \quad (4)$$

$$h_{t-1} = h_{t-2} + \kappa_h (\theta_h - h_{t-2}) + \sigma_h (\rho \epsilon_{t-1}^{(1)} + \sqrt{1 - \rho^2} \epsilon_{t-1}^{(2)}); \quad (5)$$

$$J_t = -\xi_t^D \cdot \mathbb{I}(q_t = -1) + 0 \cdot \mathbb{I}(q_t = 0) + \xi_t^U \cdot \mathbb{I}(q_t = 1), \quad (6)$$

где

y_t – зависимая переменная в момент времени t , в данном случае цена электричества,

y_{t-1} – лаг зависимой переменной, в момент времени $t-1$,

$\mu, \psi, d_{Sat}, d_{Sun}, d_{Mon}$ – константы,

X_t – логарифм почасовой температуры в момент времени t ,

$D_{Sat,t}, D_{Sun,t}, D_{Mon,t}$ – дамми переменные, учитывающие недельную сезонность,

h_{t-1} – разность между логарифмом цены в момент времени t и логарифмом цены в момент времени $t-1$,

h_{t-2} – разность между логарифмом цены в момент времени t и логарифмом цены в момент времени $t-2$,

$\kappa_h, \theta_h, \sigma_h$ – параметры распределения,

J_t – скачок,

$\epsilon_t^{(1)}, \epsilon_t^{(2)} \sim N(0,1) i.i.d.$,

$\xi_t^D \sim \exp(\eta_D) i.i.d.$,

$\xi_t^U \sim \exp(\eta_U) i.i.d.$

Для q_t значения переменной ненаблюдаемы, но можно оценить вероятность скачка вверх и вниз. Для оценки эффекта рычага вводится параметр $\rho < 0$, причем если $\rho > 0$ – это обратный параметр "рычага", то есть большим значениям логарифма цены соответствуют большие значения дисперсии.

Стоит отметить, что само понятие скачка – это общепринятый термин, который часто используется в литературе, однако точного его определения не существует. При

анализе временных рядов наличие скачков легко определить при визуальном анализе графиков из-за их выраженного отличия их положения по отношению к другим наблюдениям. Однако в то же время определить границу, классифициирующую наблюдения, зачастую бывает сложно. Простейшими методами являются квантильный метод или метод 3-сигма, однако такой подход имеет множество недостатков. Эмпирическое распределение диктует правила классификации наблюдений на скачки и не-скачки, однако, если величина скачка сильно варьируется, то при анализе подвыборок (окон) временного ряда граница будет сильно меняться в зависимости от того, сколько и каких скачков попали в это окно.

В рамках моделирования цен на электроэнергию термин скачок будет означать короткий период времени, во время которого цена на электроэнергию резко увеличивается (или уменьшается), а затем возвращается к ее предыдущему уровню последовательностью нисходящих (или восходящих) прыжков.

Обычно при моделировании величины скачков используется нормальное распределение. Однако такой выбор предполагает симметричность распределения скачков вверх и вниз, что не всегда верно. Например, скачки вверх могут быть единовременными и достаточно большими по величине, в то время как снижение к среднему уровню может происходить несколькими скачками меньшего размера.

В модели [12] для моделирования распределения отрицательных или положительных скачков используется экспоненциальное распределениями с различными значениями параметров («double exponential distribution»).

Байесовская оценка вектора параметров модели и ненаблюдаемых переменных $\theta \in \Theta$ основан на построении апостериорного распределения с плотностью, представленного уравнением (7):

$$p(\theta|y) = \frac{(p(y|\theta)p(\theta))}{\int_{\Theta} p(y|\theta)p(\theta)d\theta}, \quad (7)$$

где

θ – ненаблюдаемая переменная,

$y = (y_1, \dots, y_n)$ – вектор наблюдений,

$p(y|\theta)$ – функция правдоподобия,

$p(\theta)$ – плотность априорного распределения,

$\int_{\Theta} p(y|\theta)p(\theta)d\theta$ – нормирующая константа, или "безусловная" вероятность увидеть набор наблюдений y .

В байесовской модели константа опускается и оценивание происходит с помощью численного интегрирования по $p(y, \theta) = p(y|\theta)p(\theta)$. Значение цены на РСВ по модели M определяется по плотности прогнозного распределения (уравнение (8)), где $k > 0$:

$$p(y_n + k | y, M) = \int_{\theta} p(y_{n+1}|\theta, y, M)p(\theta|y, M)d\theta, \quad (8)$$

Такой подход позволяет учесть неопределенность относительно будущего значения цены и неопределенность относительно параметров и ненаблюдаемых переменных θ , которые зависят от предыдущих наблюдений y и предположения модели M .

Вектор неизвестных величин выглядит следующим образом (θ, h, q, ξ) , где:

- θ — вектор неизвестных параметров,
 - $h = (h_0, \dots, h_{n-1})$ — вектор ненаблюдаемых значений разности логарифма цен необходимая для оценивания стохастической волатильности,
 - $q = (q_1, \dots, q_n)$ — вектор ненаблюдаемых значений скачков вверх и вниз,
 - $\xi = (\xi_1^D, \xi_1^U, \dots, \xi_n^D, \xi_n^U)$ — вектор ненаблюдаемых значений величины скачков вверх и вниз.
- Помимо моделей диффузинно-скачкообразных компонент при анализе временных рядов часто используются модели, в которых в явном виде задается трендовая и сезонная составляющие. Одной из таких моделей является обобщенная аддитивная модель Prophet.

В Prophet-модели лежит хорошо известная модель ненаблюдаемых компонент [17], однако в отличие от классической модели Prophet включает в себя помимо трендовой и сезонной компоненты также компоненту «праздников», учитывающую аномально низкие (или высокие) значения временного ряда во время определенных дат, в отличие от «будних» дней. Модель можно описать следующим уравнением (9):

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, \quad (9)$$

где

$g(t)$ — функция тренда, которая моделирует апериодические изменения в значении временного ряда,

$s(t)$ — представляет периодические изменения, причем модель позволяет учесть множественную сезонность, например для цен на электричество, как еженедельную, так и годовую сезонность),

$h(t)$ – учитывает влияние праздников, которые происходят с нерегулярным графиком в течение одного или нескольких дней,

ϵ_t – ошибка, в которой отражаются любые специфические изменения, которые не учитываются моделью.

Рассмотрим более подробно каждую из компонент этого разложения.

Функция тренда (T_t) моделирует апериодические изменения в значении временного ряда, задается как кусочно-линейная функция, представлена уравнение (10):

$$T_t = (k + a(t)^T \delta) t + (m + a(t)^{T\gamma}), \quad (10)$$

где

k – темп роста,

$a(t)$ – вектор точек поворота,

δ – скорость подстройки,

m – параметр смещения,

γ – корректировочная константа.

Функция S_t представляет периодические изменения, причем модель позволяет учесть множественную сезонность, смотри уравнение (11):

$$S_t = \sum_{\{n=1\}}^{\{N\}} \left(a_{\{n\}} \cos \left(\frac{\{2 \pi n t\}}{\{P\}} \right) + b_{\{n\}} \sin \left(\frac{\{2 \pi n t\}}{\{P\}} \right) \right) \quad (11)$$

где

$a_n, b_n, n = 1, \dots, N$ – коэффициенты Фурье,

P – рассматриваемый ряд цен,

\cos, \sin – функции косинуса и синуса,

π – математическая константа, которая выражает отношение длины окружности к её диаметру (примерно 3,14),

t – длина временного ряда (рассматриваемый период).

$J_t = d_{\{event\}} * D_{\{event\}}$, $d_{\{event\}}$ – коэффициент при дамми праздника, учитывает влияние праздников, которые происходят с нерегулярным графиком в течение одного или нескольких дней.

Недостатком данной модели является параметрическое предположение о том, что остатки нормально распределены, что, вообще говоря, может быть неверно.

Эта спецификация аналогична обобщенной аддитивной модели (GAM), классу регрессионных моделей с потенциально нелинейными параметрами сглаживания, применяемыми к регрессорам. Здесь в качестве регрессора используется время, но, возможно использование нескольких линейных и нелинейных функций времени в качестве структурных компонент ряда. В этой модели сезонность – аддитивная компонента. Мультипликативная сезонность, может быть учтена с помощью логарифмического преобразования.

Преимущество аддитивной GAM модели в том, что он легко разлагается и допускает включение новых компонент по мере необходимости, например, когда определен дополнительный источник сезонности. GAM легко оцениваются, например, с помощью метода обратной подгонки, либо с помощью L-BFGS алгоритма. Эта модель также обеспечивает ряд практических преимуществ:

- в модели можно легко учесть многоуровневую сезонность, что позволяет делать различные предположения о трендовой составляющей ряда; в случае с ценами на электричество будут учитываться недельная сезонность, внутригодовая динамика, можно рассмотреть модели с учетом месячной сезонности и без,
- в отличие от моделей класса arima, коэффициенты модели могут меняться во времени, что важно, поскольку было показано, что сезонная динамика рядов цен не постоянна,
- не требуется интерполяция пропущенных значений, появившихся вследствие удаления выбросов,
- подгонка выполняется очень быстро, в отличие от классических байесовских моделей,
- полученная модель имеет легко интерпретируемые параметры.

Если функции сезонности и праздничных дней для каждого наблюдения объединить в матрицу X , и индикаторы точки изменения направления тренда $a(t)$ в матрице A , то Prophet модель в может быть выражена на языке кода для оценивания байесовских моделей Stan:

```
model {/ Priors
  k ~ normal(0, 5);
  m ~ normal(0, 5);
  epsilon ~ normal(0, 0.5);
  delta ~ double_exponential(0, tau);
  beta ~ normal(0, sigma);
  // Logistic likelihood
```

```

y ~ normal(C ./ (1 + exp(-(k + A * delta) .* (t - (m + A * gamma)))) + X * beta, epsilon);
// Linear likelihood
y ~ normal((k + A * delta) .* t + (m + A * gamma) + X * beta, sigma);}

```

Для оценивания параметров модели используем L-BFGS алгоритм, чтобы найти максимум апостериорной плотности, также реализованный в Stan.

Для сравнения рассмотрим также альтернативную обобщенную аддитивную модель тренд-сезонного разложения с учетом регрессоров (STR). Одним из преимуществ данной модели является то что для каждой из компонент разложения стоятся доверительные интервалы.

Кратко опишем эту модель и опишем ее особенности. Декомпозиция временного ряда производится на основе регрессии (STR). В некотором смысле используемая для оценивания процедура похожа на ridge-регрессию, а робастная STR процедура связана с LASSO регрессией. Разложение исходного ряда представлено уравнением (12):

$$Y_{\{t\}} = T_{\{t\}} + S_{\{t\}} + R_{\{t\}} \quad (12)$$

где

$T_{\{t\}}$ – трендовая компонента сглаживается так, что $\Delta_{\{t\}}^{(2)T} = T_{\{t+1\}} - 2T_{\{t\}} + T_{\{t-1\}} \sim N(0, \sigma_{\{T\}}^{(2)})$,

$S_{\{t\}}$: $\sum_{\{s\}} S_{\{s,t\}} = 0$ – аддитивная сезонная компонента, не изменяющаяся во времени,

$S = [S_{\{s,t\}}]$ – матрица $k \times n$, в которой определены наряду с наблюдаемой сезонно компонентой отражены также и «ненаблюдаемые» на данный момент сезоны (k – количество сезонов, а n – длина временного ряда),

$R_{\{t\}} \sim N(0, \sigma_{\{R\}}^{(2)})$ – «остаток».

Также эта модель может быть записана в форме модели пространства состояний (аналогично Harvey, 1985) или в виде многомерной модели ARIMA. Тогда, T_t рассматривается как ARIMA (0,2,0), а S_t как многомерный ARIMA (0,2,0) процесс с одновременно коррелированными ошибками (но без автокорреляции).

Модель можно записать в кратком виде уравнения (13):

$$y_{\{e x t\}} = X \beta + \varepsilon, \quad (13)$$

где

X – матрица наблюдений,

$\beta = [s', l']'$ – вектор неизвестных коэффициентов,

ε – ошибка.

Оценивание STR модели с помощью ML, где минус логарифм правдоподобия имеет вид уравнения (14):

$$-\log(L) = \left| \frac{\{y - Q s - \ell\}}{\{\sigma_{\{R\}}\}} \right|_{\{L_{\{2\}}\}}^{\{2\}} + \left| \left\{ \frac{\mathbb{E}_{\{s\}} \left\{ -\frac{\{1\}}{\{2\}} D \right\}}{\{\sigma_{\{s\}}\}} \right\} \right|_{\{L_{\{2\}}\}}^{\{2\}} + \left| \left\{ \frac{D_{\{\ell\}} \ell}{\{\sigma_{\{L\}}\}} \right\} \right|_{\{L_{\{2\}}\}}^{\{2\}}, \quad (14)$$

где второе слагаемое отвечает за сглаживание сезонных компонент (D_s - матрица, вычисляющая вторые производные по времени для компонент сезонности), третья - за сглаживание трендовой компоненты (D_ℓ - матрица, вычисляющая вторые производные по времени для трендовой компоненты).

Тогда, учитывая, что оценки по линейной модели и ML-оценки STR модели совпадают, имеем уравнения (15) и (16):

$$\hat{\beta} = (X'X)^{\{-1\}} X' y_{\{e x t\}} = (X'X)^{\{-1\}} [Q \ I]' y \quad (15)$$

$$\Sigma_{\{\beta\}} = X'X^{\{-1\}} X' \Sigma X (X'X)^{\{-1\}} = \sigma^{\{2\}} (X'X)^{\{-1\}} [Q \ I]' [Q \ I] (X'X)^{\{-1\}} \quad (16)$$

Трендовая компонента \hat{T} и ее доверительные интервалы могут быть получены непосредственно из $\hat{\beta}$ и $\Sigma_{\{\hat{\beta}\}}$. Для получения сезонной компоненты \hat{S} , $\hat{\beta}$ необходимо преобразовать с помощью матрицы R , возвращающей последнюю сезонную компоненту, удаленную ранее в целях оценивания. Используя матрицу R , рассчитывается новая ковариационная матрица и доверительные интервалы.

STR модель позволяет учесть регрессоры трех видов, представленных в уравнении (17):

$$Y_{\{t\}} = T_{\{t\}} + \sum_{\{i=1\}} S_{\{i \ t\}} + \sum_{\{i=1\}} P_{\{i \ t\}} + \sum_{\{i=1\}} \tilde{P}_{\{i \ t\}} \tilde{\{P\}}_{\{i \ t\}} + \sum_{\{i=1\}} \tilde{\tilde{P}}_{\{i \ t\}} + R_{\{t\}}, \quad (17)$$

где

$T_{\{t\}}$ – трендовая компонента,

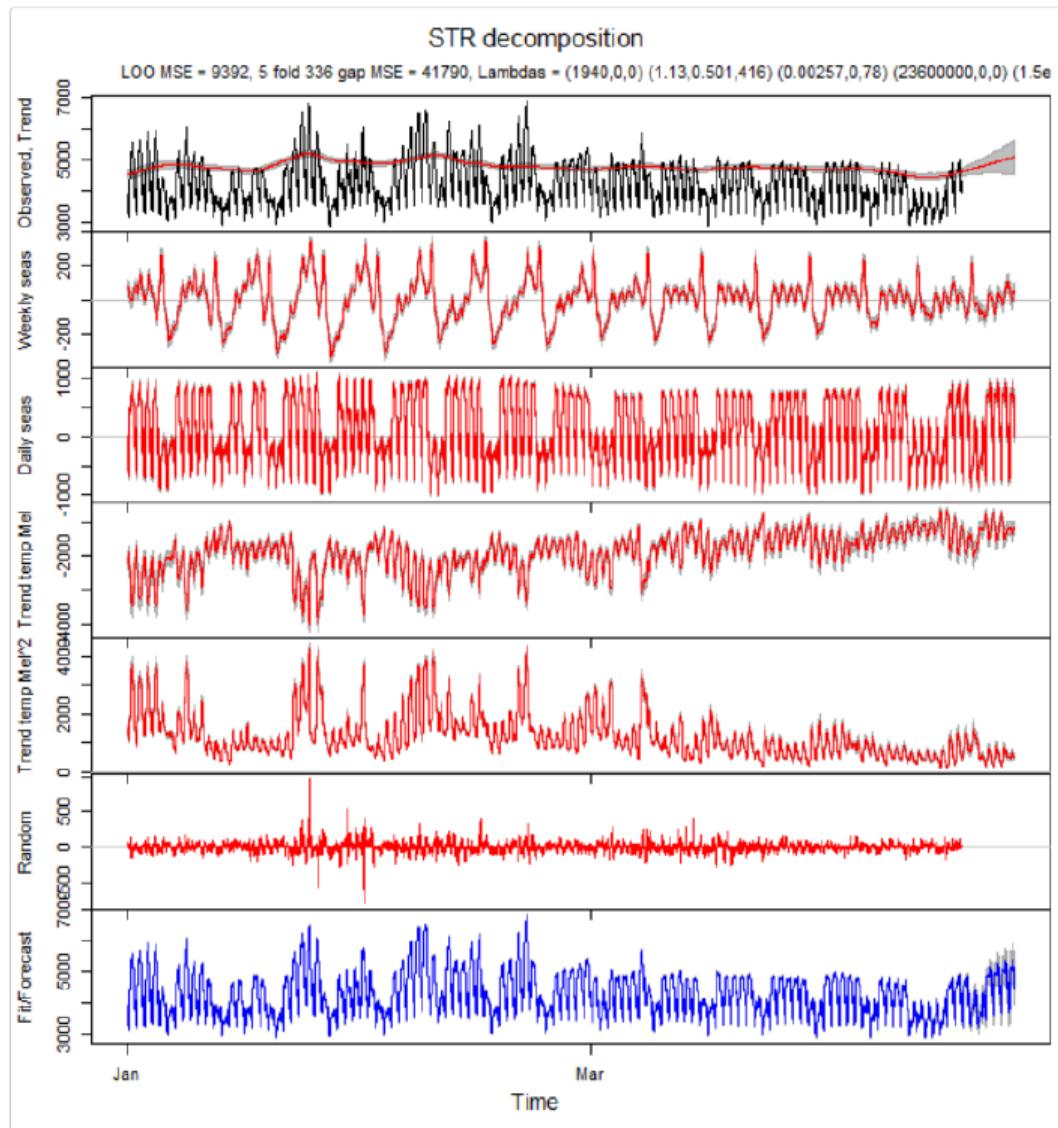
$S_{\{i\} t}$ – сезонная компонента,

$P_{\{i\} t}$ - аддитивные компоненты, линейно зависящие от предикторов p_i , с постоянными коэффициентами,

$\tilde{P}_{\{i\} t}$ - аддитивные компоненты, линейно зависящие от предикторов $\tilde{p}_{\{i\}}$ с изменяющимися во времени (несезонными) коэффициентами,

$\tilde{P}_{\{i\} t}$ - аддитивные компоненты, линейно зависящие от предикторов $\tilde{p}_{\{i\}}$; с изменяющимися во времени сезонными коэффициентами.

В качестве примера в статье [18] рассматривается разложение ряда потребления электричества с учетом дополнительных регрессоров (температуры и квадрата температуры), представленного на рисунке 2.



Примечание: Источник - [18].

Рисунок 2 – Пример разложения ряда потребления электричества с регрессорами
(температура и квадрат температуры)

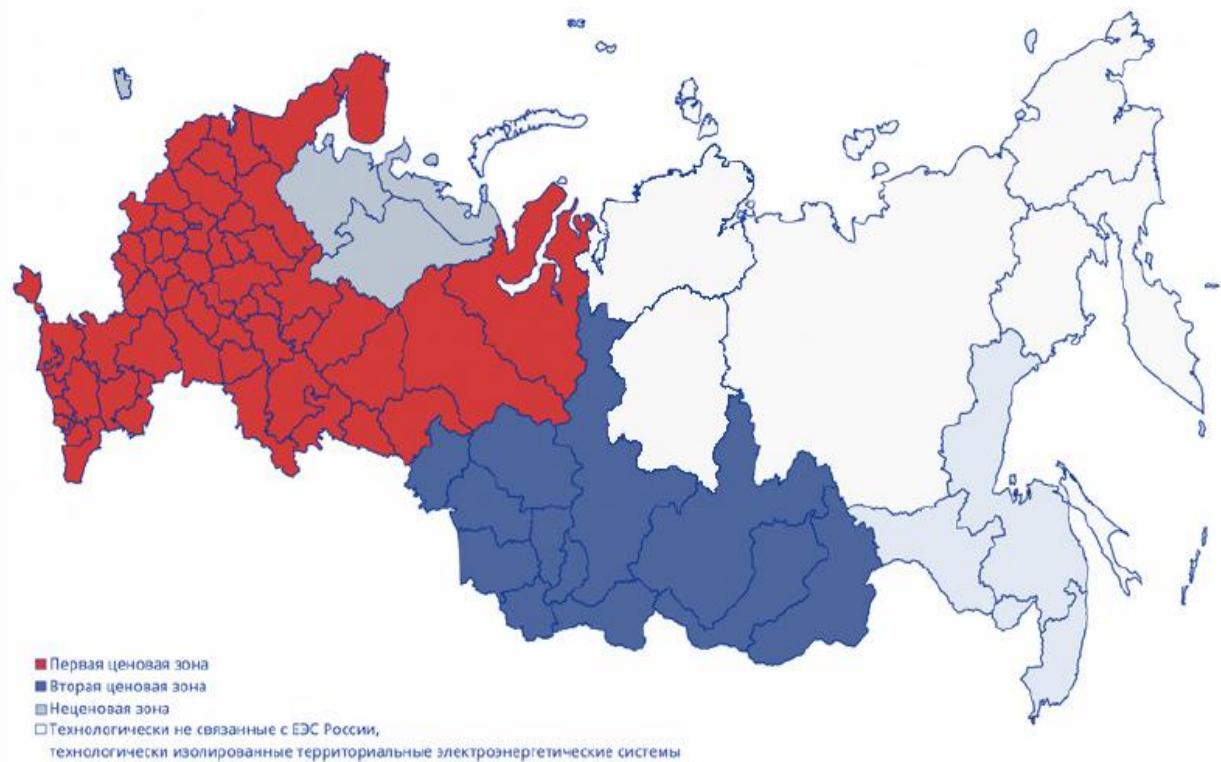
Можно заметить, что доверительные интервалы получаемые по STR очень узкие, особенно для трендовой компоненты. Однако интересно, что доверительный интервал прогноза тренда включает в себя неопределенность в оценке трендовой компоненты.

При работе с российскими данными такая модель оказалось очень чувствительной к настройке гиперпараметров. Доверительные интервалы для компонент разложения очень узкие, хотя оценки компонент сильно различаются для разных наборов гиперпараметров. Поэтому для дальнейшего анализа в качестве базовой модели для моделирования трендовой и сезонной компоненты ряда цен на электричество была выбрана именно Prophet-модель.

Таким образом, разделе были подробно описаны подходы к моделированию спотовых цен на рынке электроэнергии, их преимущества и недостатки, тем самым обозначена мотивация для создания модели, учитываящей недостатки двух классов моделей: моделирование цены на оптовом рынке электроэнергии как диффузионно-скачкообразного процесса с учётом фундаментальных факторов спроса и предложения (температура, уровень деловой активности и др.), особенностей российского рынка.

2 Особенности российского рынка электроэнергии

Рынок электроэнергии – сложная структура, состоящая из множества взаимосвязанных компонент. Энергосистема Российской Федерации состоит из ЕЭС России (семь объединенных энергосистем (ОЭС) – ОЭС Центра, Средней Волги, Урала, Северо-Запада, Юга и Сибири) и территориально изолированных энергосистем (Чукотский автономный округ, Камчатский край, Сахалинская и Магаданская область, Норильско-Таймырский и Николаевский энергорайоны, энергосистемы северной части Республики Саха (Якутия)). На рисунке 3 представлены ценовые зоны энергетического рынка России.



Примечание: Источник – АТС.

Рисунок 3 – Ценовые зоны энергорынка России

В ходе проводимых реформ в электроэнергетике был сформирован оптовый рынок электроэнергии и мощности. Так, например, в качестве объектов моделирования могут быть выбраны подструктуры энергосистемы [19]:

— рынок мощности, отдельная структура, являющаяся одной из важнейших компонент энергосистемы. На этом рынке товаром является гарантия того, что электричество будет поставлено потребителю в определенный период, иными словами, мощность представляет собой товар, покупка которого дает участнику оптового рынка право требовать от продавца мощности поддержания в готовности генерирующего

оборудования для выработки электроэнергии установленного качества в объеме, необходимом для удовлетворения потребности указанного участника в электроэнергии,

— рынок на сутки вперед (PCB) – рынок электроэнергии, на котором объемы спроса и предложения определяются на основе заявок покупателей и производителей, то есть электричество бронируется заранее на день вперед,

— балансирующий рынок (БР) – рынок на котором корректируются объемы производства электроэнергии в течение дня; покупки в течение дня совершаются на меньшие суммы, но цены на этом рынке более высокие, поскольку со штрафует участников при сильном отклонении от заявленных объемов потребления и производства, тем самым стимулируя формирования более точных заявок участниками рынка, тем самым позволяя проводить точное планирование управления всей энергосистемой

— рынок регулирования частоты, на котором избыточная мощность может использоваться для контроля частоты.

Также можно моделировать отдельные компоненты энергосистемы:

— цены на электричество на PCB,

— спрос на электричество,

— предложение электричества,

— факторы спроса на электричество и факторы предложения электричества,

— нагрузка на ЛЭП, пропускная способность энергосистемы,

— механизмы регулирования торговли электроэнергией, в том числе с помощью спотовых контрактов и производных финансовых инструментов; в России ОРЭМ достаточно молодой рынок, его либерализация и расширение постепенно продолжается, поэтому вполне возможно, что в будущем расширится список финансовых инструментов, позволяющих нивелировать риски, связанные с высокой волатильностью цен на электричество.

Рынок электроэнергии и мощности в России представляет собой двухуровневую систему: оптовый и розничный рынки, подробнее смотри рисунок 4. Электроэнергия и мощность продаются как два отдельных товара. Участниками оптового рынка являются генерирующие компании, выступающие в качестве продавцов электроэнергии и мощности, а также покупатели – крупные промышленные предприятия, сбытовые компании и гарантирующие поставщики.

| | | Тип рынка | |
|----------------|---|---|--------------------------|
| | | Оптовый рынок | Розничный рынок |
| Электроэнергия | Рынок на сутки вперед 74.7 % 767.89 млн Мвт·ч | Балансирующий рынок 5.7 % 59.04 млн Мвт·ч | Рынок регулируемых цен |
| | Рынок регулируемых договоров 15.4 % 158.2 млн Мвт·ч | Рынок свободных договоров 4.2 % 42.93 млн Мвт·ч | Рынок нерегулируемых цен |
| | | Оптовый рынок | Розничный рынок |
| Мощность | Конкурентный отбор мощности | ДПМ и реализация в статусе вынужденного генератора | Рынок регулируемых цен |
| | Рынок регулируемых договоров | Рынок свободных договоров | Рынок нерегулируемых цен |

Примечание: Структура объемов реализации электрической энергии в секторах оптового рынка электроэнергии за 2016 г. Потребление электроэнергии по субъектам Российской Федерации в 2016 г -- 1077.94 млн. МВт час. Источник – АТС.

Рисунок 4 – Структура российского энергорынка

Рынок на сутки вперед (PCB) предусматривает конкурентный отбор ценовых заявок поставщиков и покупателей за сутки до реальной поставки электроэнергии с определением цен и объемов поставки на каждый час следующих суток, что обеспечивает работу рыночных механизмов энергосистемы [20].

Целью данного исследования является разработка модели ценообразования на рынке на сутки вперед (далее PCB). Однако при определении фундаментальных факторов, влияющих на цену электричества на PCB, необходимо понимать механизмы влияния остальных компонент энергосистемы.

Причем краткосрочное прогнозирование спроса (до 1 недели) обычно основано на метеорологических прогнозах и на исторических данные о потреблении. Сезонные прогнозы также производятся с учетом справочных метеорологических данных. Долгосрочные сценарии изменения спроса на электричество создаются на основе разработки наиболее вероятных сценариев изменения динамики экономической активности в различных секторах и использования структурных моделей потребления.

Спрос на электричество не является стационарным, его годовое среднее постепенно растет [21]. В долгосрочной перспективе может ожидаться рост потребления электричества, поскольку постепенно происходит электрификация транспорта и других индустрий в связи с проведением мероприятий по снижению выбросов углекислого газа в атмосферу.

Рассмотрим подробнее факторы, влияющие на спрос и предложение на электроэнергию [22]:

- погодные условия: например, температура (причем в зимой довольно сильно растет спрос на электроэнергию из-за продолжительного отопительного сезона, а летом – из-за охладительного), скорость ветра и облачность (по большей части влияют на предложение электроэнергии, в наибольшей степени влияют на ВИЭ) и т.д.,
- неожиданные погодные условия и погодные бедствия: могут влиять на спрос, так как в крайнем случае при плохих погодных условиях может дойти до остановки производства, а следовательно падения спроса на электричество, с другой стороны при природных катализмах могут участиться аварии на ЛЭП и генераторах, что может привести к краткосрочному снижению предложения,
- дисбалансы спроса и предложения, возникающие из-за воздействия делового цикла,
- политические события и внутренняя политика, в частности нормативные акты, регулирующие деятельности энергосистемы или косвенно влияющие на нее, например, касательно энергетики и регулирования производства ресурсов, на которых работают электрогенераторы (уголь, мазут, природный газ и др.),
- интенсивность бизнеса и повседневна деловая активность.

Так на российском рынке тариф для конечного потребителя на электроэнергию и мощность формируется на основе пяти составляющих [23]:

- цена электроэнергии,
- цена мощности,
- цена передачи по сети с дифференциацией по уровню напряжения,
- инфраструктурные платежи: плата за услуги со ЕЭС, АТС, ЦФР,
- сбытовая надбавка.

Основной объем (около 72%) производимой электроэнергии продается на рынке на сутки вперед (РСВ).

При моделировании российской энергосистемы необходимо учитывать ее особенности, а именно:

- высокая степень изношенности основных фондов,
- перекрестное субсидирование,
- проблема неплатежей,
- вынужденная генерация,
- высокие потери тепла.

Все эти факторы являются сдерживающими при оценке динамики развития и модернизации энергосистемы. Рассмотрим подробнее особенности российского рынка. Во-первых, это высокая степень изношенности основных фондов. Несмотря на проведение инвестиционных программ, направленных на модернизацию энергосистемы, этот процесс дорогостоящий и длительный. Согласно Генеральной схеме размещения объектов электроэнергетики до 2035 года, около 46% установленных мощностей в России были введены до 1980 года [24].

На фоне спада промышленного производства в 1990-х годах, наблюдалось снижение потребления электроэнергии, что усугублялось нерегулярными и неполными платежами за предоставленную электроэнергию, что привело к дефициту собственных средств, выделяемых на обновление основных фондов электроэнергетики. В результате много лет сектор электроэнергетики не получал достаточного объема инвестиций, что на текущий момент привело к критическому износу оборудования [25].

Дефицит инвестиций в обновление основных производственных фондов, а также в их реконструкцию и модернизацию может привести к техническим ограничениям и снизить надежность энергоснабжения потребителей.

Во-вторых, это проблема перекрестного субсидирования [26]. Перекрестное субсидирование – это ценовая дискриминация, при которой для одних покупателей (потребителей) цена устанавливается ниже предельных издержек за счет других покупателей (потребителей), для которых цена устанавливается выше предельных издержек (при этом средние цены соответствуют средним издержкам). Данный механизм вносит значительные искажения в конечную цену электрической и тепловой энергии и в тарифы на услуги по ее передаче.

Наиболее ярким примером тарифной политики последних лет является перекрестное субсидирование между категориями потребителей, приводящее к перераспределению платежного бремени, часть которого переносится с населения на промышленность. Перекрестное субсидирование между категориями потребителей в сфере электроэнергетики возникло в России в начале 1990-х годов, когда в период экономического спада государством были приняты меры социальной поддержки населения.

В конце 1990-х годов Правительство России поставило цели по ликвидации перекрестного субсидирования. К 2007 году экономическая справедливость цен для различных категорий потребителей была практически достигнута, однако после кризиса 2008 года динамика ухудшилась: расхождение тарифов для населения и промышленности снова увеличилось.

Перекрестное субсидирование населения в существенной мере негативно влияет на отрасль. Среди его основных недостатков можно выделить:

- стимулирование крупных потребителей к уходу из «большой» энергетики и к постройке собственных генерирующих объектов, что приводит к росту тарифной нагрузки для остальных потребителей,
- отсутствие адресности при распределении субсидий,
- непрозрачность распределения нагрузки по перекрестному субсидированию между группами потребителей.

В-третьих, это проблема неплатежей. По данным Ассоциации «НП Совет рынка» [23], на конец октября 2017 года на оптовом рынке задолженность составила 65,2 млрд руб., а на розничном – 243 млрд руб. Если на оптовом рынке электроэнергии есть ряд законодательных мер, которые позволяют воздействовать на неплательщика, то на розничном рынке такой механизм практически отсутствует.

Фактически отключить электроэнергию населению и приравненной к нему группе потребителей невозможно из-за особого статуса предоставляемой услуги и ее высокой социальной значимости.

В-четвертых, это вынужденная генерация [27]. Сокращение доли выработки тепловой энергии на ТЭЦ приводит к тому, что наиболее экономичные и надежные теплоисточники в массовом порядке становятся неэффективными на рынке электроэнергии, где они сначала получают статус генераторов, работающих в вынужденном режиме, а впоследствии выводятся из эксплуатации. Мощности, работающие в режиме вынужденной генерации, оплачиваются по существенно более высокой цене, чем рыночная, что приводит к росту расходов потребителей.

В-пятых, это высокие потери тепла [28]. Проблема связана в том числе с первой проблемой изношенности основных фондов. Объемы реконструкции и реновации тепловых сетей недостаточны для поддержания их в требуемом техническом состоянии, что приводит к значительным потерям энергии при передаче. Износ тепловых сетей приводит к высоким теплопотерям: в России они достигают 20-30%, что практически в четыре раза выше, чем в европейских странах. В 2016 году потери в системе теплоснабжения, по данным Росстата, составили 132,3 млн Гкал.

По данным Министерства энергетики, для того, чтобы российская система теплоснабжения стала надежной и приблизилась по эффективности к европейским стандартам, до 2025 года нужно вложить в отрасль 2,5 трлн руб. Из них около 40% требуется потратить на обновление генерирующих объектов, 30% – на обновление теплосетей, а оставшуюся треть средств – на оптимизацию операционных расходов.

Ввиду всех особенностей российской энергосистемы встает вопрос необходимости системного анализа всех оставляющих компонент энергосистемы с целью создания практических рекомендаций, помогающих преодолеть недостатки инертной отрасли. В частности, часто ставится задача прогнозирования спроса, связанная с особенностью рынка электроэнергии, а именно невозможностью хранения и проблемой обязательства энергоустановок.

Для оценки влияние отдельного события, например, изменений в регламенте НП «Совет рынка» необходимо иметь декомпозицию ряда. Поскольку некоторые события могут иметь краткосрочный эффект и вызвать скачок на рынке, который нивелируется за счет эффекта рычага, например, крупная авария. Другие же последствия могут иметь долгосрочные последствия и повлиять на уровень наклона трендовой компоненты, например, при наступлении экономического кризиса долгосрочное снижение темпов роста производства влечет за собой также долгосрочное снижение темпов роста объемов спроса на электричество, что напрямую влияет на равновесные цены на этом рынке.

Также в качестве особенности российской энергосистемы стоит отметить использование инвестиционной программы договоров на поставку мощности ДПМ [29]. Механизм строительства и модернизации электрогенерирующих мощностей по договорам о поставке мощности (ДПМ) стал частью реформы РАО ЕЭС в конце 2000-х годов. Он использовался для привлечения в отрасль негосударственных инвестиций на рыночных, взаимовыгодных для банков и генерирующих компаний условиях. Однако эта система имеет ряд недостатков, например, перекладывание ценового бремени на покупателя (в том числе проблема перекрёстного субсидирования). Также результатом применения системы ДПМ стал тот факт, что многие производители готовы были заплатить штрафы за установку генерирующих мощностей иностранного производства, что не способствует развитию отечественного рынка электрогенераторов.

Совокупная установленная мощность российских генерирующих компаний в последние годы демонстрировала устойчивый рост, в связи с проведением государственных реформ в энергетической отрасли в 2001-2011 годах и последующей реализации инвестиционных программ по модернизации и установке новых генерирующих мощностей.

Отметим также крупнейших игроков в отрасли российской электроэнергетики, в конкурентном сегменте генерации и регулируемом сегменте передачи и распределения. Во-первых, в сегменте генерации это государственные компании, такие как ПАО «РусГидро», ООО «Газпромэнергохолдинг», ПАО «Интер РАО», АО «Концерн Росэнергоатом», частные российские, такие как АО «ЕвроСибЭнерго», ПАО «Т Плюс»,

частные иностранные, а именно, ПАО «Юнипро», ПАО «Фортум», ПАО «Энел Россия» [30].

Во-вторых, в сегменте передачи и распределения это группа российских сетевых компаний ПАО «Россети» (Федеральная сетевая компания, Межрегиональные распределительные сетевые компании, Региональные распределительные сети) и прочие частные компании АО «БЭСК», ПАО «СУЭНКО».

Таким образом, в разделе были кратко описаны особенности функционирования оптового рынка электроэнергии и мощности (ОРЭМ) и формирования равновесных цен на рынке на сутки вперед (PCB). Также были описаны ценовые факторы, которые потенциально могут быть включены в модель, объяснен выбор первоначального набора внешних регрессоров на основе изученной литературы по структурным моделям для рынка электроэнергии и доступных данных.

3 Эмпирический анализ и калибровка модели цен на электричество на рынке на сутки вперед

В рамках данной работы мы сосредоточимся на задаче моделирования цены на электричество на РСВ и декомпозиции цены на трендовую и сезонную составляющие.

Прогнозы цен содержат важную информацию для производителей и потребителей при планировании стратегий торгов с целью максимизации прибыли и минимизации издержек. В условиях постепенной либерализации рынка электроэнергетики в России задача моделирования цены на электричество, а также построения качественного прогноза становится все более важной.

При построении прогнозов цен на электричество необходимо верно оценить доверительные интервалы прогноза, например, для участников рынка эти оценки позволяют верно оценить риски при принятии стратегических решений. Традиционные методы временных рядов не могут точно моделировать нелинейную структуру цены на электроэнергию, также без учета скачкообразной компоненты доверительные интервалы могут быть заужены.

Именно поэтому нас интересует SVDEJX модель [12], подробнее описанная ранее (смотри уравнения (4)-(6)).

Данная модель хорошо показала себя в прогнозировании по сравнению с альтернативными моделями. Однако недостатком данной модели является то, что в качестве регрессоров, объясняющих динамику в этих рядах, используется только температура и дамми-переменные на недельную сезонность.

Помимо этих факторов в качестве регрессоров могут быть также включены и другие показатели, прямо или косвенно отражающие влияние на цены на электричество:

- среднемесячная цена на электроэнергию,
- расход электроэнергии на душу населения,
- отопительные дни,
- «охладительные» дни,
- индекс промышленного производства,
- гидроэлектрический индекс производительности,
- спотовая цена на brent,
- импорт сырой нефти на душу населения,
- производство электроэнергии с помощью ВИЭ на душу населения,
- баланс импорта-экспорта электроэнергии на душу населения.

В связи с этим можно оценить дополненную SVDEJX модель, представленную уравнениями (18)-(20):

$$\begin{aligned}
p_t = & p_{t-1} + \mu + \beta_1 y_{t-1} + \beta_2 y_{t-7} + \beta_3 y_{t-30} + \beta_4 y_{t-365} + \gamma_1 p_t^{coal} + \\
& + \gamma_2 p_t^{gas} + \psi_1 T_{t-1} + \psi_2 T_{t-30} + d_{Sat} D_{Sat,t} + d_{Sun} D_{Sun,t} + \\
& + d_{Mon} D_{(Mon,t)} + \sqrt{\exp(h_{t-1})} \epsilon_t^{(1)} + J_t;
\end{aligned} \tag{18}$$

$$\begin{aligned}
h_{t-1} = & h_{t-2} + \kappa_h (\theta_h - h_{t-2}) + \\
& + \sigma_h
\end{aligned} \tag{19}$$

$$\begin{aligned}
J_t = & -\xi_t^D \cdot I(q_t = -1) + 0 \cdot I(q_t = 0) + \\
& + \xi_t^U \cdot I(q_t = 1),
\end{aligned} \tag{20}$$

где

$p_{t-k} = \ln(S_{t-k})$ – k -ые лаги логарифма средней цены за день,

$y_t = \ln(Y_t)$ – уровень деловой активности,

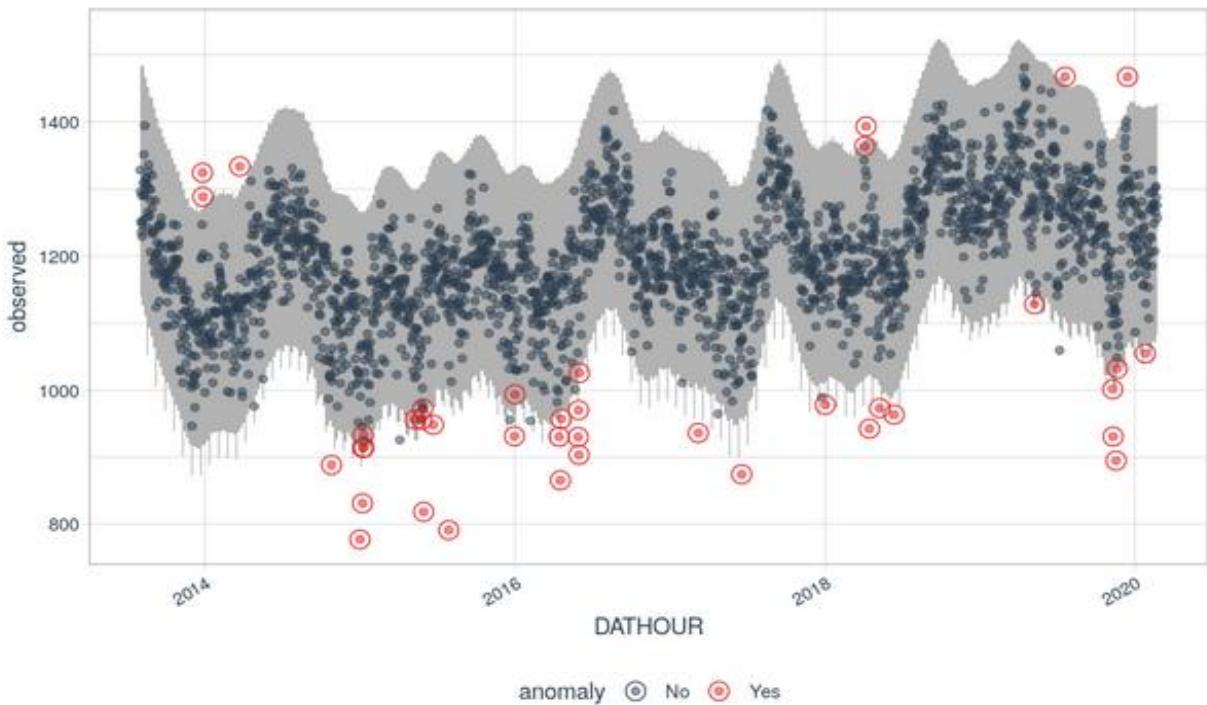
p_t^R – логарифм цены или прогноза цен на ресурсы, где $R = \{coal, gas\}$,

T_t – логарифм модуля средней дневной температуры (если $\psi > 0$, то большим отклонениям от температуры нуля соответствуют большие цены); отметим, что погодные условия сильнее всего влияют на цены, причем при более точном прогнозировании погодных условий можно уменьшить ошибку прогноза цены на электричество,

$D_{Sat,t}, D_{Sun,t}, D_{Mon,t}$ – дамми переменные, учитывающие недельную сезонность, ненаблюдаемые переменные: $h_{t-1} = p_t - p_{t-1} = \ln(S_t/S_{t-1})$ и q_t – скачки вверх и вниз.

Далее подробно опишем факторы, которые потенциально можно использовать в качестве регрессоров, объясняющих динамику цен на электричество с учетом особенностей российского рынка. В первую очередь описывается показатель, источник данных. Далее рассматриваются его статистические характеристики. Также описывается предполагаемый канал влияния на цены. Отметим, что влияние различных факторов может идти исключительно на трендовую, сезонную или случайную компоненту, что будет отражено в последующем анализе.

Объясняемая переменная в модели – это индекс равновесных цен на покупку электроэнергии, руб./МВт.ч. Данные по этому показателю доступны для первой и второй ценовой зоны за период с 01-08-2013 по текущие сутки с ежедневной периодичностью. Показатель, публикуется с усреднением по ЦЗ (ценовым зонам). Доступны индексы равновесных цен как потребителей, так и производителей. Смотри рисунок 5.



Примечание – Источник: АТС.

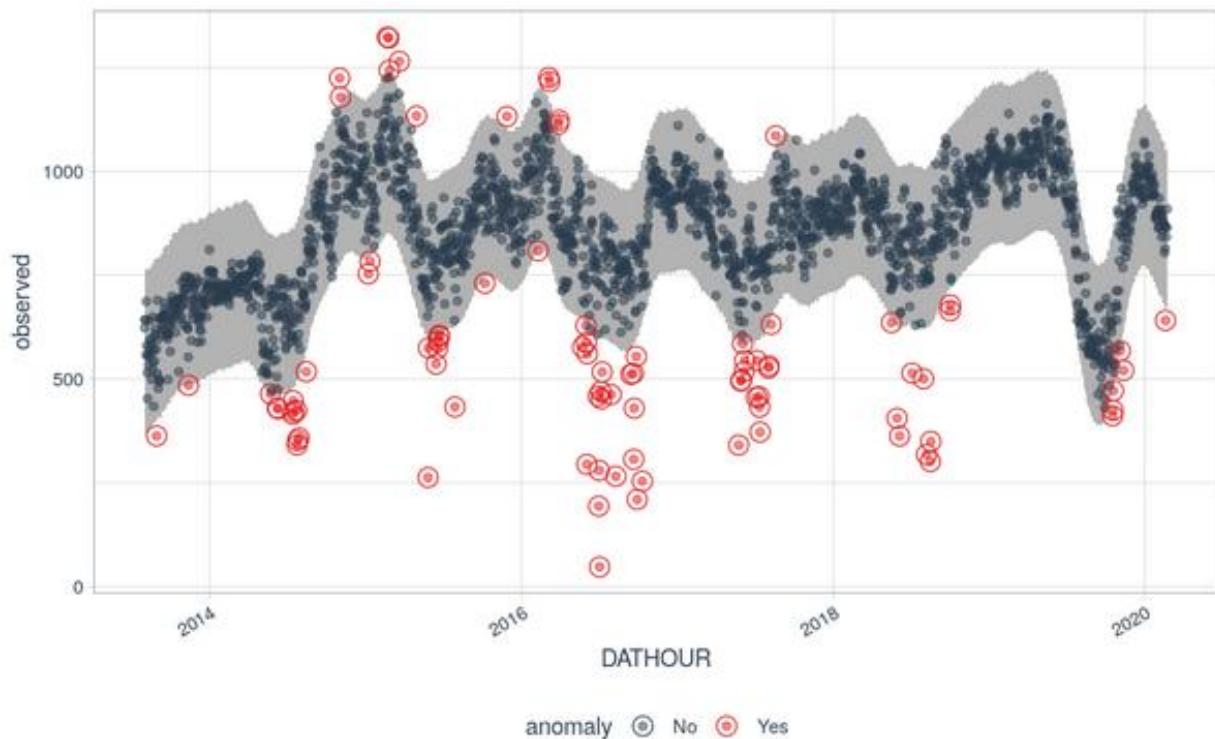
Рисунок 5 – Цены на PCB для 1 ЦЗ. Красным цветом отмечены потенциальные «аномальные» значения – скачки

Ряды цен для обеих ЦЗ p_t^1 и p_t^2 характеризуются наличием тренда, сезонности и множественных выбросов. В p_t^1 с помощью метода GESD было выявлено 10 выбросов, во втором – 89. Корреляция между ценами – 0.125, цены во второй ЦЗ более волатильны – $\sigma(p_t^1) = 95.7$, $\sigma(p_t^2) = 157.9$. С помощью теста Канова-Хансена (CH) на стабильность сезонных колебаний было определено, что для обоих рядов годовые сезонные колебания нестабильны, следовательно, необходимо либо рассматривать ряды на более коротких временных отрезках, либо использовать модели, учитывающие эти изменения, например, модели с коэффициентами изменяющимися во времени (TVC).

Поскольку объясняемая переменная – индекс равновесных цен потребителей для каждой из ЦЗ, ожидается, что факторы, влияющие на цены в разных ЦЗ, а также выделяемые трендовые и сезонные компоненты должны различаться.

Далее рассмотрим группу показателей, отражающих индекс деловой активности, которые по большей части отражают влияние на цены через их влияние на спрос на электричество.

Во-первых, в качестве показателя отражающего общую динамику спроса можно использовать ВВП, в ценах 2016 года, млн рублей. Данные в постоянных ценах позволяют учесть влияние только изменения в объемах производства, без учета инфляции. Данные доступны с 2014 по 2019 гг. с ежеквартирной периодичностью. Пропуски из-за различия в периодичности данных заполнены с помощью линейной интерполяции, смотри рисунок 6.



Примечание – Источник: АТС.

Рисунок 6 – Цены на РСВ для 2 ЦЗ. Красным цветом отмечены потенциальные «аномальные» значения – скачки

Ряд, публикуемый Росстатом сезонно сглаженный. Также в текущих ценах доступны данные по квартальному ВВП с сезонной корректировкой, который менее привлекателен в качестве регрессора, поскольку хуже отражает спад после 2014 из-за инфляции, а также ряда без сезонной корректировки, динамика которого за вычетом тренда сопоставима с циклическими регрессорами (СН тест не отвергает гипотезу о стабильности сезонных колебаний на 5% уровне).

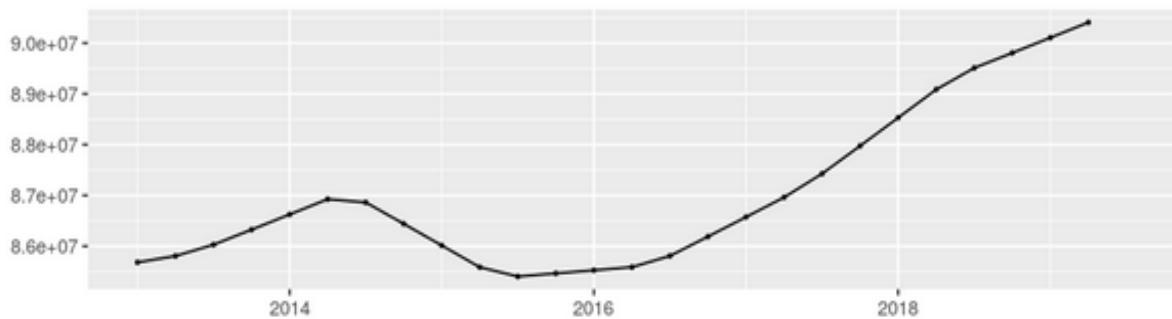
Предполагается, что большим значениям ВВП соответствуют большие цены, поскольку для производства большего объема требуется больше электроэнергии и спрос на нее растет, и наоборот. Предполагается, что влияние этого показателя идет только на трендовую составляющую цен на электричество.

Во-вторых, аналогичным показателем является ВРП, в текущих ценах, млн рублей. Данные по ВРП позволяют учесть производство в рамках отдельной ЦЗ, однако данные доступны с 1998 по 2017 гг. с ежегодной периодичностью с прогнозом на 2018, 2019 гг. Пропуски заполнены с помощью линейной интерполяции.

Также, как и в случае с ВВП, большим значениям ВРП соответствуют большие цены, поскольку для производства большего объема требуется больше электроэнергии и спрос на нее растет, и наоборот. Предполагается, что влияние этого показателя идет

только на трендовую составляющую цен на электричество. В качестве регрессоров для каждой из ЦЗ используется соответствующий ей ВРП.

Корреляция между ВРП 1 ЦЗ и 2 ЦЗ – 0.9997. Корреляция между приростами ВРП за год 1 ЦЗ и 2 ЦЗ – 0.925. Корреляция между ВРП 1 ЦЗ и предсказаниям по ARMA(0,0) с дрифтом – 0.999, между ВРП 2 ЦЗ и предсказаниям по ARMA(0,0) с дрифтом – 0.9991, следовательно, нет смысла использовать этот показатель в регрессии, поскольку он не обладает достаточной вариацией, способной улучшить оценку трендовой составляющей цен на электричество по сравнению с регрессором t , учитывающим линейный тренд, смотри рисунки 7-9.



Примечание – Источник: Росстат.

Рисунок 7 – ВВП, в ценах 2016 года, млн рублей

В-третьих, индекс промышленного производства (в %) – относительный показатель, характеризующий изменение масштабов производства в сравниваемых периодах. Индекс производства используется при анализе динамики физического объема продукции. Сезонно сглажен по STL, поскольку нам наиболее интересно его влияние на трендовую составляющую ряда цен. Данные доступны с ежемесячной периодичностью.

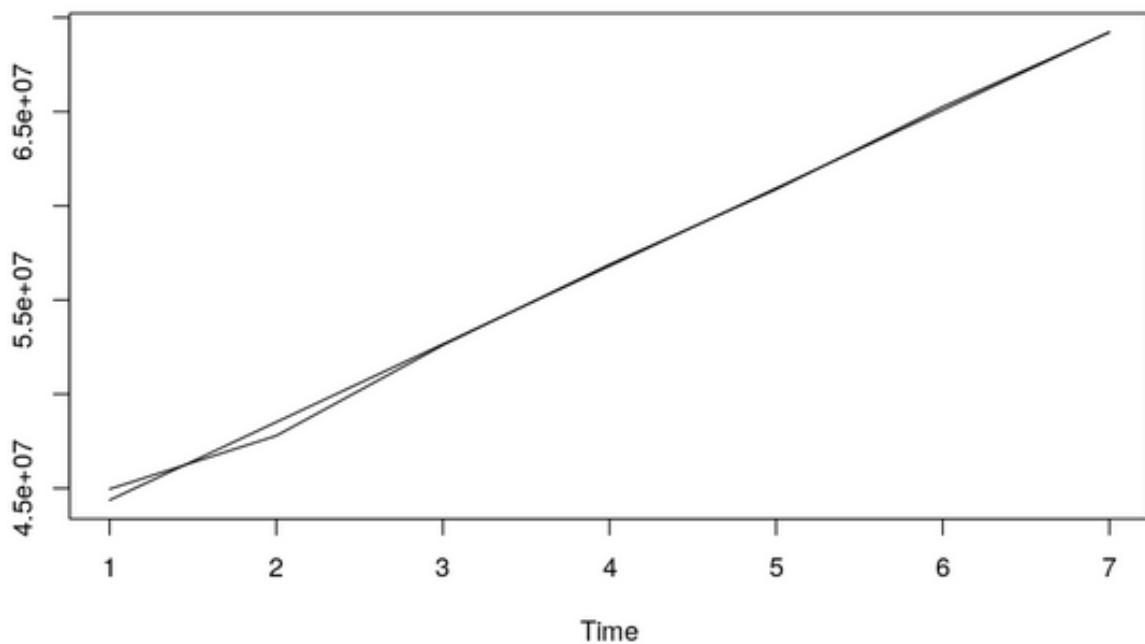
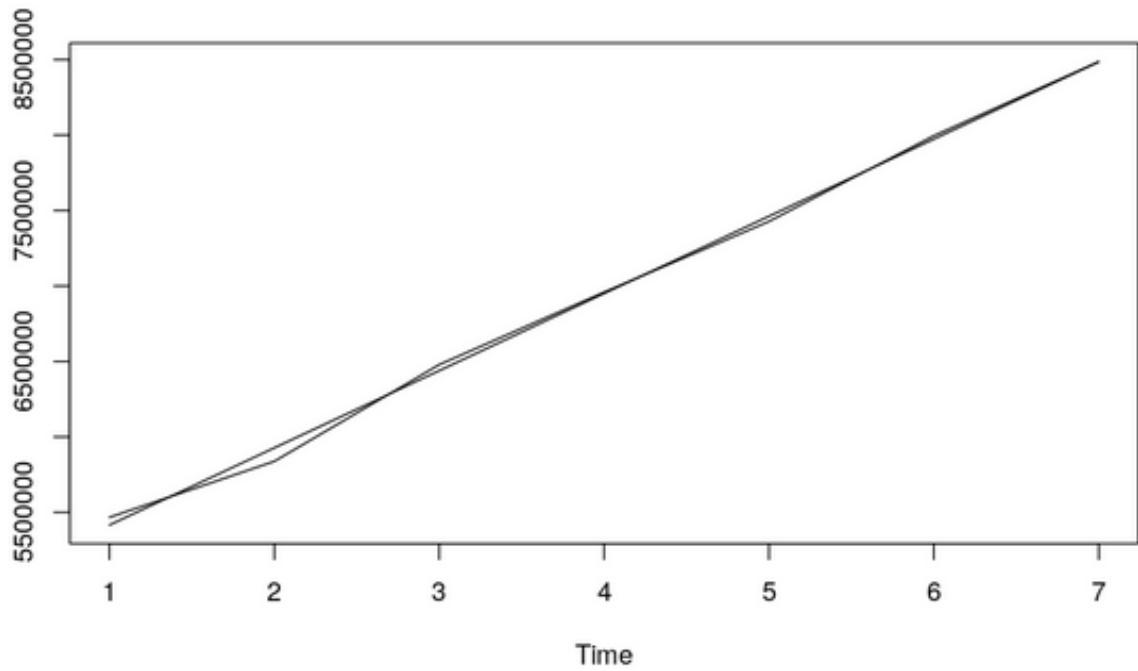


Рисунок 8 – ВРП 1 ЦЗ и оценки ВРП 2 ЦЗ по ARMA(0,0) с дрифтом



Примечание – Источник: Росстат и собственные расчеты.

Рисунок 9 – ВРП 2 ЦЗ и оценки ВРП 2 ЦЗ по ARMA(0,0) с дрифтом

Данный показатель сильно коррелирует с ВВП (в базовых ценах 2016 года) – 0.75, и сильнее коррелирует с ценами на электроэнергию в обеих ЦЗ, в 1 ЦЗ – 0.26, 2 ЦЗ – 0.24 (в то время как корреляция ВВП (2016) с ценами 1 ЦЗ – 0.17, 2 ЦЗ – 0.12). Вероятно, причиной этому является то, что данные по ИПП доступны с месячной периодичностью, в

отличие от квартального ВВП, поэтому предпочтительней использовать ИПП в качестве регрессора, отражающего долгосрочный тренд деловой активности.

Индекс промышленного производства – агрегированный индекс производства по видам деятельности «Добыча полезных ископаемых», «Обрабатывающие производства», «Производство и распределение электроэнергии, газа и воды». Рост обрабатывающего производства приводит к росту спроса на электроэнергию.

В-четвертых, индекс производства по виду экономической деятельности добыча топливно-энергетических полезных ископаемых (СА). Выделение трендовой составляющей было осуществлено с помощью пакета Demetra с использованием процедуры X12-ARIMA. Источник: Горшкова Т.Г., Каукин А.С., Миллер Е.М., Турунцева М.Ю.

Данный показатель отрицательно коррелирует с ВВП (в базовых ценах 2016 года) – -0.48, также отрицательно коррелирует с ценами на электроэнергию в обеих ЦЗ, в 1 ЦЗ – -0.27, 2 ЦЗ – -0.17 (в то время как корреляция ВВП (в базовых ценах 2016 года) с ценами 1 ЦЗ – 0.17, 2 ЦЗ – 0.12).

Рост ИПП добычи топливно-энергетических полезных ископаемых (ТЭПИ) с одной стороны может приводить к росту спроса на электроэнергию и цен, с другой стороны при больших объемах добычи ТЭПИ при прочих равных, цена на тепло-энергетические ресурсы может снижаться, предложение может вырасти, цены снизиться.

В-пятых, индекс производства по виду экономической деятельности производство и распределение электроэнергии, газа и воды (Е). Выделение трендовой составляющей было осуществлено с помощью пакета Demetra с использованием процедуры X12-ARIMA. Источник: Горшкова Т.Г., Каукин А.С., Миллер Е.М., Турунцева М.Ю.

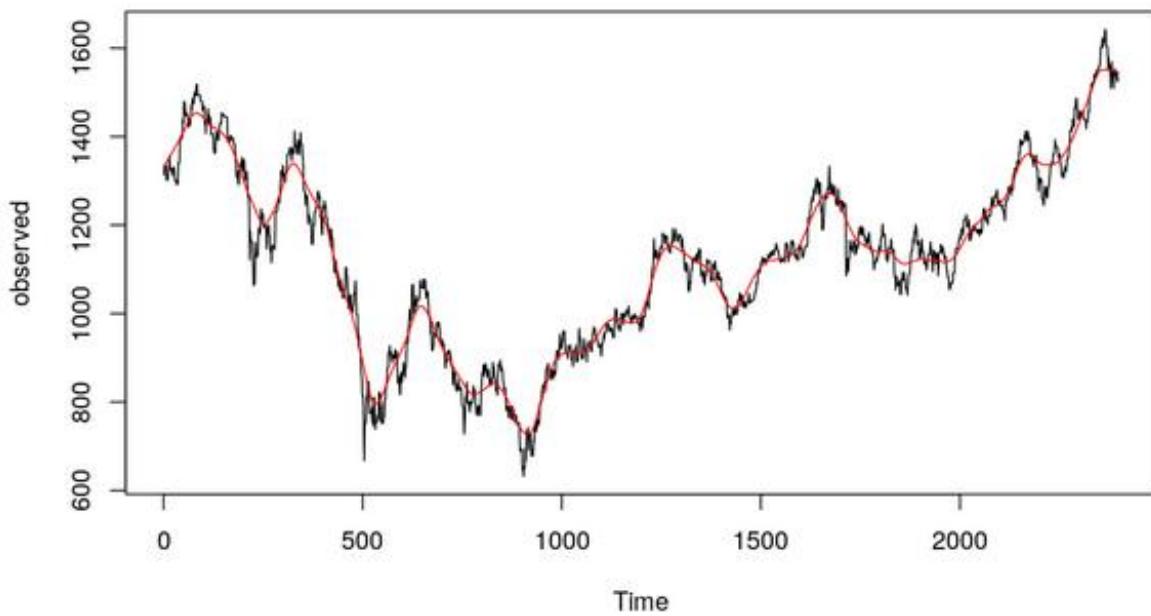
Данный показатель отрицательно коррелирует с ВВП (в базовых ценах 2016 года) – 0.62, также положительно коррелирует с ценами на электроэнергию в 1 ЦЗ – 0.15, отрицательно в 2 ЦЗ – -0.26 (в то время как корреляция ВВП (2016) с ценами 1 ЦЗ – 0.17, 2 ЦЗ – 0.12).

Рост ИПП производства электроэнергии, газа и воды с одной стороны может приводить к росту спроса на электроэнергию и цен за счет роста производства газа и воды, с другой стороны при росте ИПП за счет производства электроэнергии или производства газа при прочих равных предложение может вырасти (напрямую или за счет снижения цен на газ), цены снизиться.

В-шестых, индекс РТС (RTSI), в текущих ценах, млн рублей. RTSI -- основной индикатор фондового рынка России, рассчитанный по рыночной капитализации крупнейших и динамично развивающихся компаний, виды экономической деятельности

которых относятся к основным секторам экономики. Компании, включенные в список для расчета RTSI включают в себя имеющие производство, как в 1, так и 2 ЦЗ. Данные получены на момент открытия за каждый день. Пропуски за нерабочие дни заполнены последними известными значениями. Источник данных Финам.

Финансовые индексы отличаются сильной волатильностью, которая может мешать использованию этого показателя в качестве регрессора для моделирования цен на электричество, поскольку, зная, что зависимость между этими показателями не может быть прямой, оба показателя отличаются частыми случайными скачками вверх и вниз, который могут коррелировать случайным образом. Во избежание смещения оценок влияния RTSI на трендовую составляющую цен на электричество, в качестве регрессора будет использоваться тренд ряда RTSI, полученный с помощью STL разложения. Корреляция цен в 1 ЦЗ и тренда RTSI – - 0.235, цен во 2 ЦЗ и тренда RTSI – - -0.29. Более подробный анализ влияния метода декомпозиции может быть проведен, если RTSI, будет значительно влиять на цены на электричество. В основе STL процедуры LOESS регрессия, близкая по логике оценивания к непараметрическим статистическим методам, поэтому трендовая составляющая нелинейна и может уловить дневные изменения трендовой компоненты, смотри рисунок 10.



Примечание – Источник: АТС и собственные расчеты.

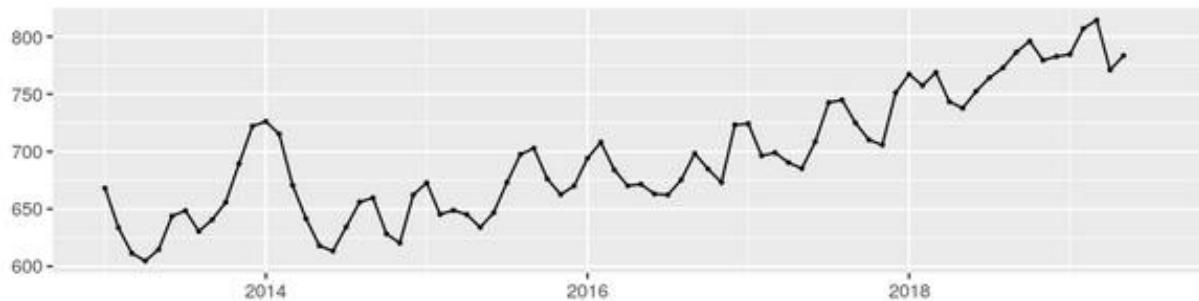
Рисунок 10 – Цены на PCB для 2 ЦЗ. Красным цветом выделена сглаженная по STL компонента.

Поскольку индекс RTSI – композитный финансовый индекс, нельзя говорить о прямой взаимосвязи между ним и ценами. Однако, трендовая компонента этого индекса может отражать общие тенденции деловой активности крупнейших компаний –

резидентов 1 и 2 ЦЗ, а следовательно, отражать состояние российской экономики. Крупные компании могут использовать заемный капитал для наращивания объемов производства или инвестирования в основной капитал, что так или иначе положительно влияет на спрос на электроэнергию. Однако, даже если этот показатель действительно может отразиться в ценах на электричество, то имеет смысл учитывать только влияние лагированной переменной.

Следующей группой показателей, влияющих на цены на электричество, являются цены на ресурсы. По большей части они должны отражать влияние на цены через предложение электроэнергии.

Во-первых, это цены на уголь (бурый), руб/т. Данные с ежемесячной периодичностью, доступны также в разбивке по федеральным округам, что позволяет получить усредненные цены на уголь для каждой из ЦЗ. Динамика цен на уголь представлена на рисунке 11.



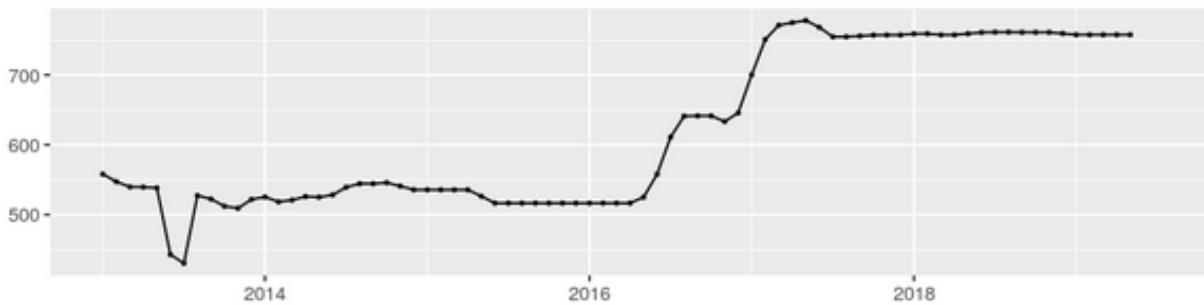
Примечание – Источник: Росстат.

Рисунок 11 – Средние цены на уголь по России

Отметим, что средний уровень цен в 1 ЦЗ определялся, как среднее значение по Центральному, Приволжскому и Уральскому ФО, по 2 ЦЗ – значения цен для Сибирского ФО. Средний уровень цен в 1 ЦЗ не отличается волатильностью, цены находятся примерно на одном уровне в с 2013 по конец 2016, наблюдается резкий рост в 2017 г, цены находятся на одном уровне с начала 2018 года. Эта динамика лучше аппроксимируется с помощью дамми переменных для соответствующих периодов роста и стабильности цен на уголь, смотри рисунки 12-13.

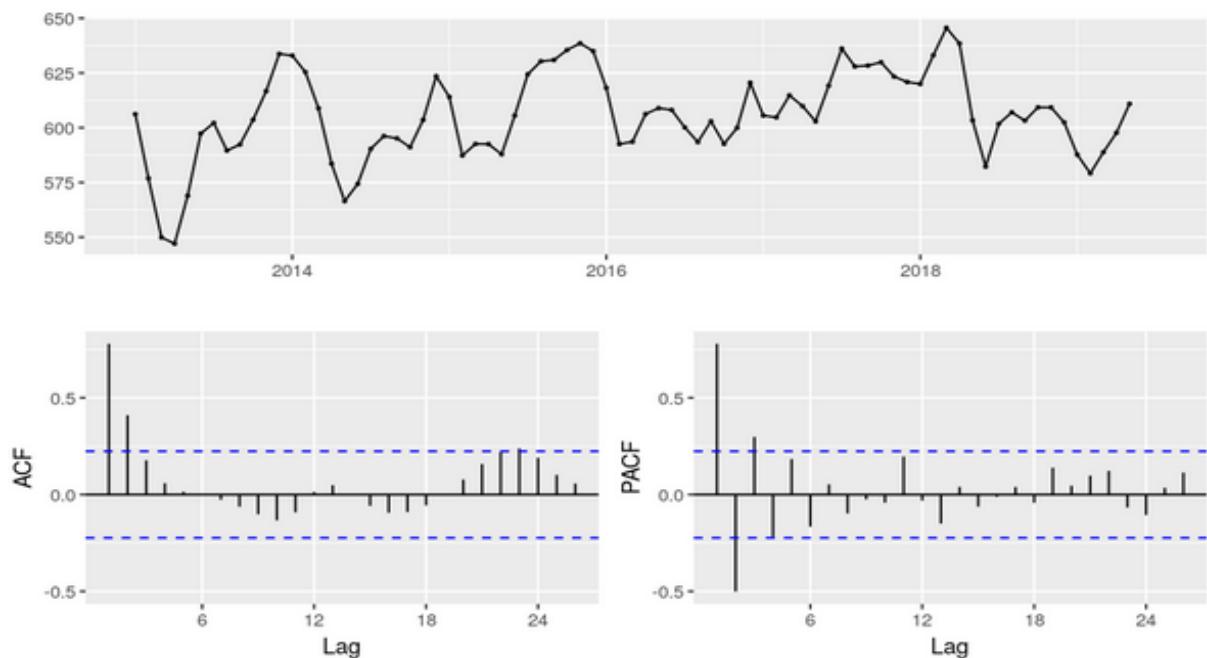
Уголь является одним из основных ресурсов для производства электроэнергии в РФ. Рост цен на ресурсы должен соответствовать росту цен на электроэнергию. Также предполагается, что цены на уголь имеют большее влияние на цены на электричество во второй ценовой зоне. В модели цен на электричество 1 ЦЗ будет использована средняя цена по РФ, если фактор окажется значимым, то также можно будет учесть скачки цен в

рамках 1 ЦЗ с помощью дамми переменных. В модели цен на электричество 2 ЦЗ будет использована средняя цена по Сибирскому ФО.



Примечание – Источник: Росстат.

Рисунок 12 – Средние цены на уголь по 1 ЦЗ

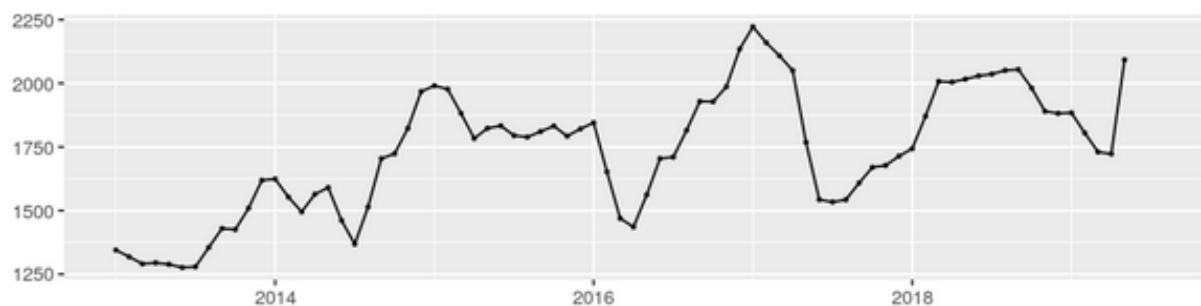


Примечание – Источник: Росстат и собственные расчеты.

Рисунок 13 – Средние цены на уголь по 2 ЦЗ, их автокорреляционная функция и частная автокорреляционная функция

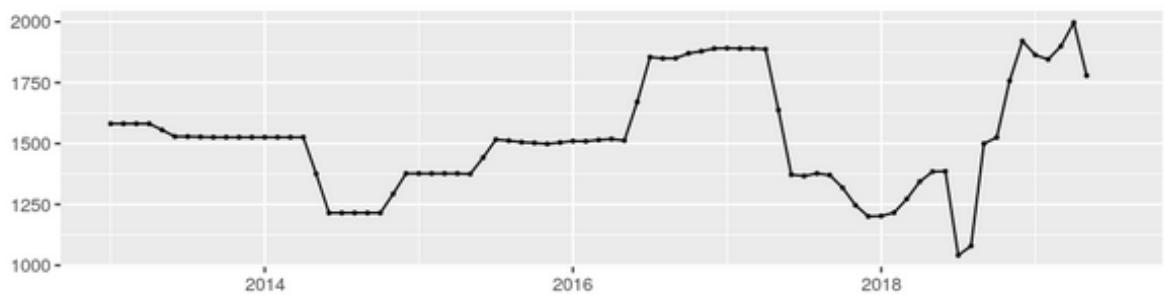
Во-вторых, это цены на газ горючий природный (естественный), руб./тыс. куб. м. Данные с ежемесячной периодичностью, доступны также в разбивке по федеральным округам, что позволяет получить усредненные цены на газ для каждой из ЦЗ. Динамика цен на газ по всей России представлена на рисунке 14, в разбивке по ценовым зонам на рисунках 15-16. Отметим, что средний уровень цен в 1 ЦЗ определялся, как среднее значение по Центральному, Приволжскому, Северо-Западному, Южному, Северо-Кавказскому и Уральскому ФО, по 2 ЦЗ – значения цен для Сибирского ФО. Средний уровень цен во 2 ЦЗ не отличается волатильностью, цены находятся примерно на одном

уровне с 2013 по конец 2014, далее меняется скачками один или два раза в год по 2017 год. Эта динамика лучше аппроксимируется с помощью дамми переменных для соответствующих периодов роста и стабильности цен на уголь.



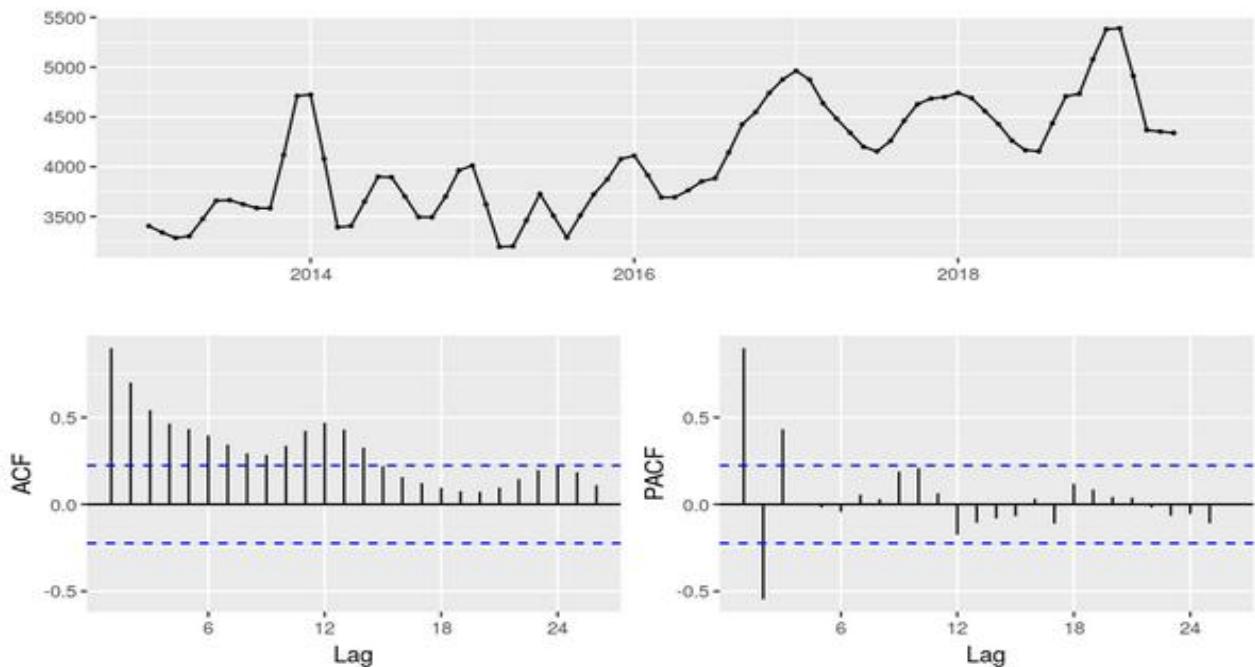
Примечание – Источник: Росстат.

Рисунок 14 – Средние цены на газ по России



Примечание – Источник: Росстат.

Рисунок 15 – Средние цены на газ по 2 ЦЗ



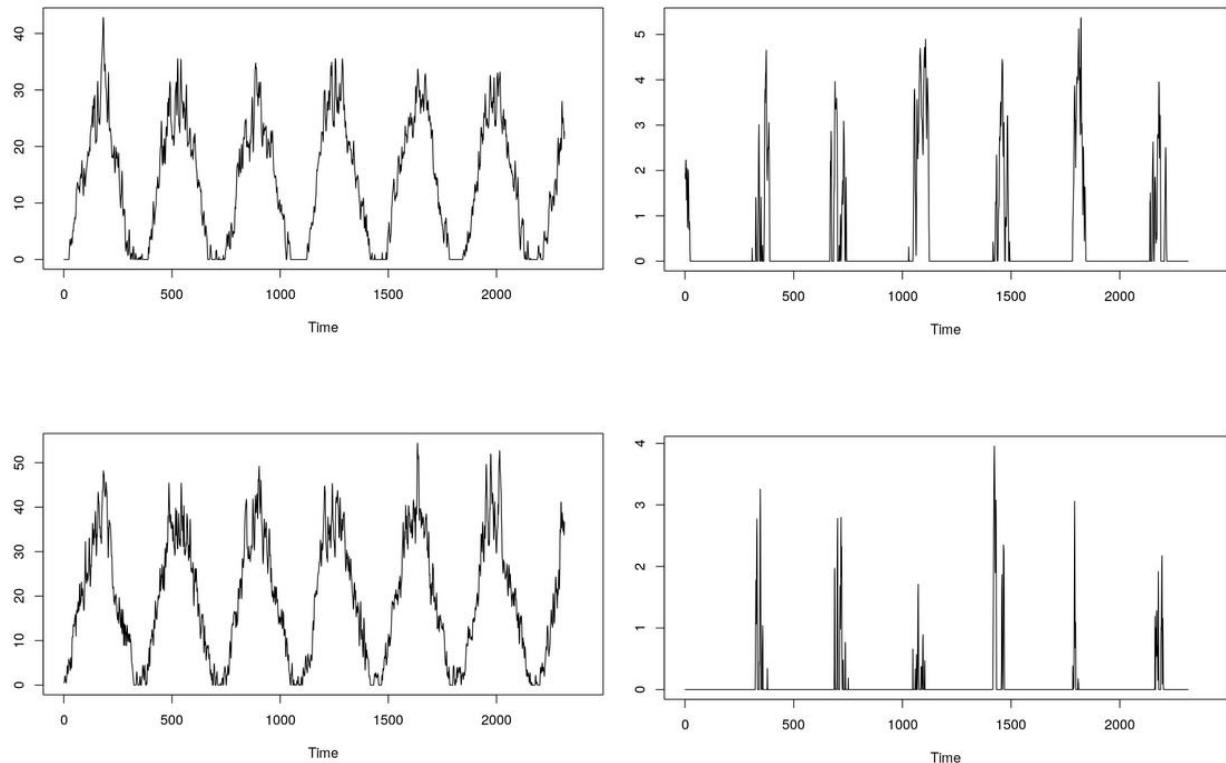
Примечание – Источник: Росстат и собственные расчеты.

Рисунок 16 – Средние цены на газ по 1 ЦЗ, их автокорреляционная функция и частная автокорреляционная функция

Газ является одним из основных ресурсов для производства электроэнергии в РФ. Рост цен на ресурсы должен соответствовать росту цен на электроэнергию. Также предполагается, что цены на газ имеют большее влияние на цены на электричество в первой ценовой зоне. В модели цен на электричество 1 ЦЗ будет использована средняя цена по ФО входящих в состав 1 ЦЗ. В модели цен на электричество 2 ЦЗ будет использована средняя цена по РФ, если фактор окажется значимым, то также можно будет учесть скачки цен в Сибирском ФО с помощью дамми переменных.

Далее, важным фактором является температура. Однако ее влияние на цены на электричество нелинейно, поскольку по большей части идет за счет роста спроса на электроэнергию из-за отопительного или охладительного периодов.

Так, рассчитываются два показателя: градусо-сутки отопительного периода (Heating Degree-Days – HDD) и охладительного периода (Cooling Degree Days – CDD). В классическом определении разница между наружной температурой воздуха и базовой температуры – 18.3°C умножается на коэффициент затрат тепла и холода, соответственно. Данные по температуре – усредненные за день показатели по всей территории 1 ЦЗ и 2 ЦЗ, смотри рисунок 17.



Примечание – Источник: Climate Data Store и собственные расчеты.

Рисунок 17 – Регрессоры, связанные с температурой: справа – HDD, слева – CDD, первый ряд – 1 ЦЗ, второй ряд – 2 ЦЗ

Если разность наружной температуры и базовой отрицательна, CDD приравнивается к 0; если разность базовой и наружной температуры отрицательна, HDD приравнивается к 0. Для задачи регрессии точное значение коэффициента не важно, поскольку роль играет только факт пропорциональности потребности в отоплении (охлаждении) и разности температур. Корреляция HDD_t^1 и p_t^1 – -0.26, CDD_t^1 и p_t^1 – 0.24, HDD_t^2 и p_t^2 – 0.43, CDD_t^2 и p_t^2 – -0.33.

Данные показатели используются для определения потребности в тепле от отопительных систем и в холода от систем кондиционирования воздуха, соответственно, большим отклонениям от базовой цены, т.е. большим значениям HDD и CDD, соответствуют большие объемы спроса на электроэнергию, а соответственно и большие цены.

Последняя группа показателей – дамми переменные на понедельник, субботу и воскресенье, а также дамми на государственные праздники. Если суббота, воскресенье или понедельник совпадают с государственными праздниками, то значение соответствующих им дамми переменных приравнивается к нулю. Дамми переменные должны отражать недельные сезонные колебания.

Предполагается, что в выходные дни спрос на электричество ниже, в понедельник – выше, причем снижение спроса на электричество во время государственных праздников должно быть ниже, чем в воскресенье.

Расширенная SVDEJX модель может учесть влияние дополнительных регрессоров на различные составляющие компоненты цен на электричество. При работе с байесовскими моделями очень важно правильно подобрать априорные распределения коэффициентов, важно выбрать правильно специфицировать модель, выбрать оптимальный алгоритм численного интегрирования и его параметры.

Для начала анализа цен на РСВ в первой и второй ценовых зонах будет проводиться на основе более простой Prophet-модели [31]. Prophet-модель описывает практический подход к прогнозированию «в масштабе», то есть сочетает в себе «ручную» настройку параметров модели в сочетании с автоматической калибровкой параметров внутри самой модели. В результате авторами предлагается «модульная» регрессионная модель с интерпретируемыми параметрами.

Напомним, что целью данного исследования является создание модели цены на оптовом рынке электроэнергии как диффузионно-скачкообразного процесса с трендовой и сезонной компонентой с учётом фундаментальных факторов спроса и предложения

(температура, уровень деловой активности и др.), особенностей российского рынка. Основная гипотеза заключается в том, что с помощью построенной модели можно получить трендсезонное разложение цены на электричество с более точными оценками коэффициентов при интересующих нас переменных.

При визуальном анализе рядов по электричеству были определены желаемые характеристики модели. Модель должна учитывать следующие особенности:

- должна включать в себя компоненту трендовой и годовой сезонности. Коэффициент, оценки которого нас интересуют больше всего – коэффициент при линейном тренде для простейшей модели, учитывающей тренд, либо набор коэффициентов, относящийся к моделированию трендовой компоненты (например, как по обобщенной аддитивной модели типа Prophet),
- поведение волатильности предполагается стохастичным: волатильность цен должна изменяться согласно определённому случайному процессу,
- большое число выбросов подразумевает то, что в финальной модели должна быть скачкообразная компонента,
- интересен тот факт, что коэффициент корреляции Пирсона между рядами по ценам на электричество для первой и второй ЦЗ имеет близкое к 0 значение ($\hat{\rho}(p_1^m, p_2^m) = 0.07$). Для того, чтобы понять за счет влияния каких факторов существует такая разница между рядами, в финальную модели должны быть экономические факторы, оценки коэффициентов при которых нас интересуют.

То есть в результате по полученной модели можно будет определять темпы роста цен на электроэнергию; причины различия в динамике цен в ценовых зонах; влияние структуры генерирующих мощностей на волатильность цен; оценивать влияние резких изменений экономических факторов на цены. Учитывая вышесказанное, было решено использовать следующую стратегию эмпирических оценок для построения модели, обладающей желаемыми характеристиками. Здесь и далее будем называть обобщенную аддитивную модель типа Prophet Моделью (1), модель стохастической волатильности с экспоненциально распределенными скачками, эффектом рычага и экзогенными переменными, предложенную в [12] Моделью (2), модель, создаваемую нами в рамках данного исследования Моделью (3).

На первом шаге происходит выбор основных детерминантов и первичный анализ с помощью Модели (1):

- решение о добавлении трендовой составляющей,
- решение о добавлении годовой сезонности,
- решение о добавлении месячной сезонности,

- решение о добавлении недельной сезонности,
- решение о добавлении дамми на государственные праздники,
- решение о добавлении регрессора градусо-сутки отопительного периода, для 1 ЦЗ и 2 ЦЗ,
- решение о добавлении регрессора градусо-сутки охладительного периода, для 1 ЦЗ и 2 ЦЗ,
- решение о добавлении регрессора цен на газ в 1 ЦЗ для 1 ЦЗ,
- решение о добавлении регрессора цены на уголь во 2 ЦЗ для 2 ЦЗ.

На втором шаге происходит выбор функциональной формы (упрощенной Модель (1) + (2))

- с помощью репараметризации, можно отделить трендовую и сезонную компоненту (использовав аналоги компонент Prophet-модели) от «стохастической», которая уже будет моделироваться по модели Kostrzewski, смотри уравнение (21):

$$\begin{aligned} r_{\{t\}} &= y_{\{t\}} - T_{\{t\}} - S_{\{t\}} \\ r_{\{t\}} &= r_{\{t-1\}} + \mu + \psi X_{\{t\}} + \sqrt{\{\exp(h_{\{t-1\}})\} \epsilon_{\{t\}}^{\{(1)\}}} + J_{\{t\}} \end{aligned} \quad (21)$$

где

$y_{\{t\}}$ – исходный ряд,

$T_{\{t\}}$ – трендовая составляющая,

$S_{\{t\}}$ – сезонная составляющая,

$r_{\{t\}}$ – случайная составляющая,

$X_{\{t\}}$ – изменение случайных факторов, например, температуры,

$h_{\{t-1\}}$ – ненаблюдаемые переменные,

J_t – скачок.

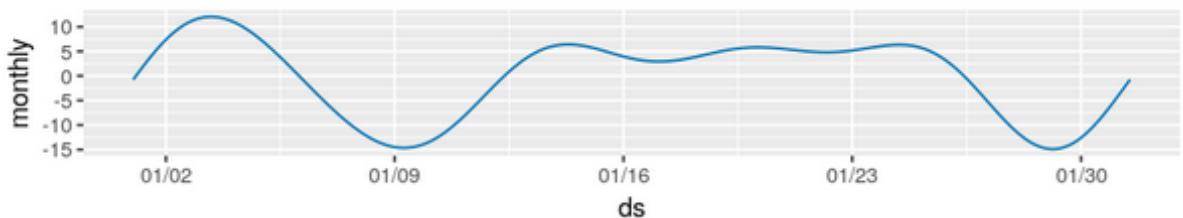
На третьем шаге происходит оценивание Модели (3):

- сравнить OLS с байесовскими методами,
- сравнить классическую линейную модель с моделью стохастической волатильности,
- сравнить модель с неинформативными априорными распределениями и модель с информативными априорными распределениями,
- добавление в модель дополнительных регрессоров (ИПП, цены на ресурсы).

На четвертом шаге происходит дальнейшее пошаговое усложнение Модели (3):

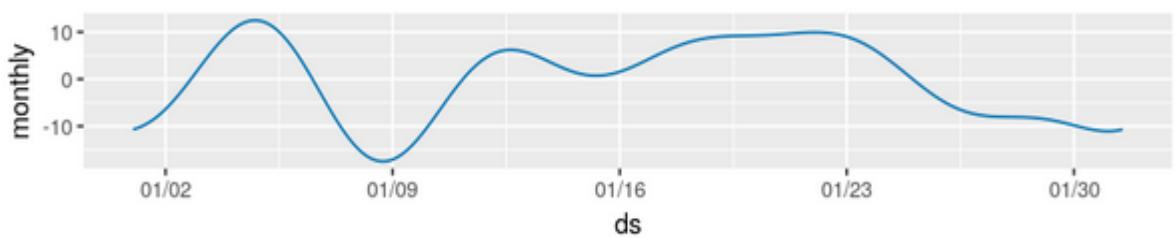
- добавление кусочно-линейного тренда как в Модели (1),
- добавление сезонной функции как в Модели (1),
- добавление диффузионной компоненты как в Модели (2),
- добавление скачкообразной компоненты как в Модели (2).

Перейдем к Шагу 1 стратегии эмпирических оценок. Для начала сравним результаты моделирования по модели, включающей и не включающей месячную сезонность, которая для ценовой зоны 1 представлена на рисунке 18, а для ценовой зоны 2 – на рисунке 19.



Примечание – Источник: Собственные расчеты.

Рисунок 18 – Месячная компонента цен в 1 ЦЗ

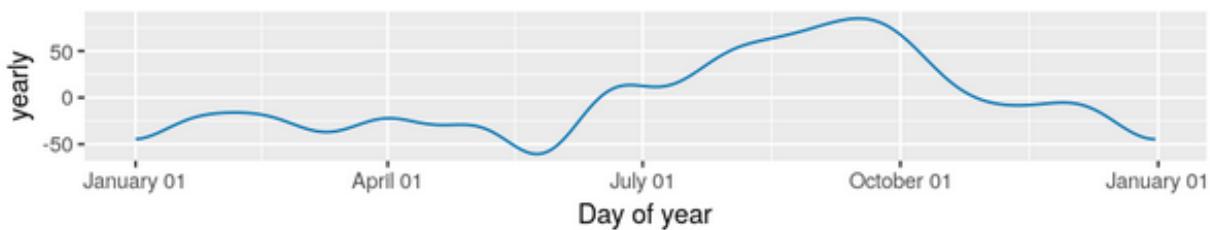


Примечание – Источник: Собственные расчеты.

Рисунок 19 – Месячная компонента цен во 2 ЦЗ

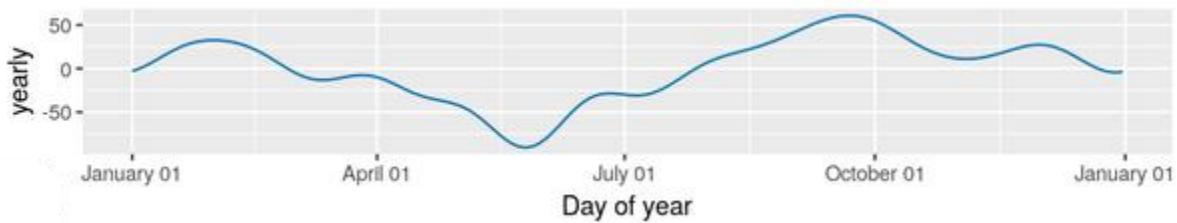
Как можно заметить на рисунках 18, 19 динамика месячной сезонности не отличается в первой и второй ценовых зонах. Причиной этому является то, что в обеих моделях порядок разложения Фурье одинаков и равен 5. При этом для обоих рядов вероятно, что такие внутримесячные колебания, были получены случайно, поскольку нет никаких теоретических оснований предполагать, что цены в течение месяца имеют такую динамику. К тому же среднее значение внутримесячных колебаний для обеих ЦЗ близко к нулю: -0.15 для 1 ЦЗ и 0.07 для второй.

При анализе рисунков 20-23 можно заметить, что учет температуры, а именно регрессора HDD 1 ЦЗ и 2 ЦЗ, соответственно, позволяет более точно оценить внутригодовую динамику цен, поскольку большая часть изменений в течение года связана с продолжительным отопительным периодом.



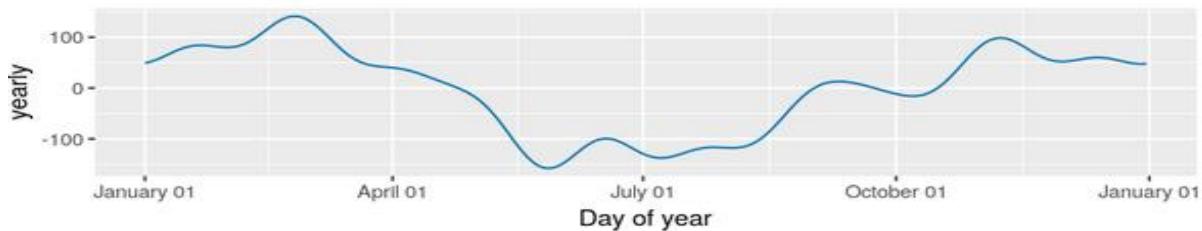
Примечание – Источник: Собственные расчеты.

Рисунок 20 – Годовая сезонная компонента цен в 1 ЦЗ без учета температуры



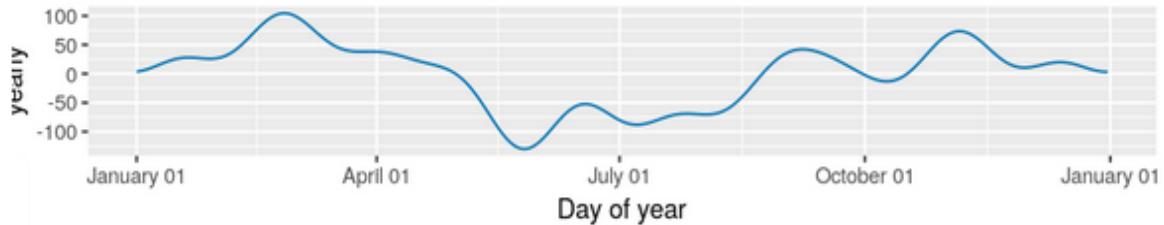
Примечание – Источник: Собственные расчеты.

Рисунок 21 – Годовая сезонная компонента цен в 1 ЦЗ с учетом температуры



Примечание – Источник: Собственные расчеты.

Рисунок 22 – Годовая сезонная компонента цен во 2 ЦЗ без учета температуры



Примечание – Источник: Собственные расчеты.

Рисунок 23 – Годовая сезонная компонента цен во 2 ЦЗ с учетом температуры

Отметим, что после учета влияния температуры общая внутригодовая динамика в ценах стала значительно более схожей, хотя никаких предположений относительно годовой сезонности в разных ЦЗ не было учтено в модели. Далее годовая динамика и влияние температуры на цены на электричество будут рассмотрены более подробно.

Также отметим, что учет цен на ресурсы, который предположительно должен влиять на трендовую компоненту никак на нее не повлиял, как для первой, так и для второй ценовой зоны. Это связано с тем, что в основе этой модели лежит задача

«сглаживания», а не регрессии, поэтому для того, чтобы повлиять на «оптимальную» надо либо изменить спецификацию модели, либо более тщательно откалибровать гиперпараметры оптимизационного алгоритма.

Помимо цен на ресурсы, предполагалось, что большую роль в оценивании трендовой составляющей цен на электричество сыграет индекс деловой активности. Однако все предложенные ранее индексы имели стабильный рост начиная с 2015 года. Из-за этого трендовая компонента оценивалась неадекватно, также завышенная потенциальный рост цен. Если в более сложной модели необходимо будет учесть эти факторы, надо будет сгладить влияние этого резкого роста, например, использовать стандартизованный временной ряд, либо учесть затухающее влияние этого показателя с помощью изменяющихся во времени коэффициентов, либо с помощью дамми-переменных учитывать влияние только в период кризиса, когда отмечался резкий спад производства.

В итоге финальная Prophet модель включает в себя следующие компоненты:

- трендовая составляющая,
- внутригодовая сезонность,
- недельная сезонность,
- государственные праздники,
- градусо-сутки отопительного периода, для 1 ЦЗ и 2 ЦЗ,
- цены на газ в 1 ЦЗ для 1 ЦЗ,
- цены на уголь во 2 ЦЗ для 2 ЦЗ.

На следующем Шаге этапе перейдем к моделированию цен по упрощенной модели стохастической волатильности с целью дальнейшего пошагового усложнения этой модели. Необходимо выбрать функциональную форму упрощенной модели, поэтому на данном этапе начнем с сравнения модели классической линейной регрессии (Модель 1А), байесовскую линейную регрессию (Модель 1Б) и байесовскую модель стохастической волатильности (Модель 1В).

Первая модель основана на Модели (2).

Модель 3А Классическая линейная регрессия представлена уравнением (22):

$$\begin{aligned}
 p_{\{t\}} = & p_{\{t-1\}} + \mu + c_{\{HDD\}} * HDD_{\{t\}} + c_{\{CDD\}} * CDD_{\{t\}} + c_{\{IPP\}} * IPP_t \\
 & + c_{\{gas\}} * p_{\{gas,t\}} + c_{\{coal\}} * p_{\{coal,t\}} + d_{\{Sat\}} D_{\{Sat,t\}} + d_{\{Sun\}} D_{\{Sun,t\}} \\
 & + d_{\{Mon\}} D_{\{Mon,t\}} + \varepsilon_t,
 \end{aligned} \tag{22}$$

где

$p_{\{t\}} = \ln(S_{\{t\}})$ - логарифм средней цены в ЦЗ за день,

$HDD_{\{t\}}$ - Heating-Degree-Days среднее по ЦЗ,

$CDD_{\{t\}}$ - Cooling-Degree-Days среднее по ЦЗ,

IPP - сглаженная трендовая составляющая ИПП,

$p_{\{gas\}}$ - среднемесячная цена на газ (по 1 ЦЗ для 1 ЦЗ, по РФ для 2 ЦЗ),

$p_{\{coal\}}$ - среднемесячная цена на уголь (по РФ для 1 ЦЗ, по 2 ЦЗ для 2 ЦЗ),

$D_{\{Sat,t\}}, D_{\{Sun,t\}}, D_{\{Mon,t\}}$ - учитывают недельную сезонность,

$\varepsilon_t \sim N(0, \sigma)$ *i. i. d.*

| <i>Dependent variable:</i> | | |
|----------------------------|--|--|
| | $\Delta \log(\text{consumer_price_1})$ | $\Delta \log(\text{consumer_price_2})$ |
| | (1) | (2) |
| hdd_1 | 0.00004 (0.0001) | |
| cdd_1 | 0.001 (0.001) | |
| hdd_2 | | 0.0001 (0.0002) |
| cdd_2 | | 0.002 (0.006) |
| ipp_trend | 0.606 (1.899) | -0.287 (3.912) |
| coal_rf | -0.00002 (0.00004) | 0.00003 (0.0001) |
| gaz_rf | 0.00000 (0.00000) | 0.00000 (0.00001) |
| mon_dummy | 0.067*** (0.003) | 0.013* (0.007) |
| sat_dummy | -0.019*** (0.003) | 0.019*** (0.007) |
| sun_dummy | -0.052*** (0.003) | -0.040*** (0.007) |
| Constant | -0.608 (1.910) | 0.266 (3.938) |
| Observations | 1,820 | 1,505 |
| R ² | 0.343 | 0.039 |
| Adjusted R ² | 0.340 | 0.034 |
| Residual Std. Error | 0.042 (df = 1811) | 0.084 (df = 1496) |
| F Statistic | 118.172*** (df = 8; 1811) | 7.551*** (df = 8; 1496) |

Note:

*p<0.1; **p<0.05; ***p<0.01

Примечание – Источник: Собственные расчеты.

Рисунок 24 – Модель 3А Классическая линейная регрессия

Можно проверить, надо ли брать $p_{\{t-1\}}$ с коэффициентом 1 с помощью теста Дики-Фуллера. Оценим следующую регрессию, включающую в себя все регрессоры, что и в Модели 3А, но добавим лаг логарифма температуры и лаги разности логарифма температуры до 6 порядка.

Модель 3Б Аналог теста Дики Фуллера представлена уравнением (23):

$$\begin{aligned}
\Delta p_{\{t\}} = & \psi p_{\{t-1\}} + \beta_1 \Delta p_{\{t-1\}} + \beta_2 \Delta p_{\{t-2\}} + \beta_3 \Delta p_{\{t-3\}} + \beta_4 \Delta p_{\{t-4\}} + \beta_5 \Delta p_{\{t-5\}} + \\
& + \beta_6 \Delta p_{\{t-6\}} + c_{\{HDD\}} * HDD_{\{t\}} + c_{\{CDD\}} * CDD_{\{t\}} + c_{\{IPP\}} * IPP_t + \\
& + c_{\{gas\}} * p_{\{gas,t\}} + c_{\{coal\}} * p_{\{coal,t\}} + d_{\{Sat\}} D_{\{Sat,t\}} + d_{\{Sun\}} D_{\{Sun,t\}} \\
& + d_{\{Mon\}} D_{\{Mon,t\}} + \varepsilon_t,
\end{aligned} \tag{23}$$

где

$p_{\{t\}} = \ln(S_{\{t\}})$ - логарифм средней цены в ЦЗ за день,

$HDD_{\{t\}}$ - Heating-Degree-Days среднее по ЦЗ,

$CDD_{\{t\}}$ - Cooling-Degree-Days среднее по ЦЗ,

IPP - сглаженная трендовая составляющая ИПП,

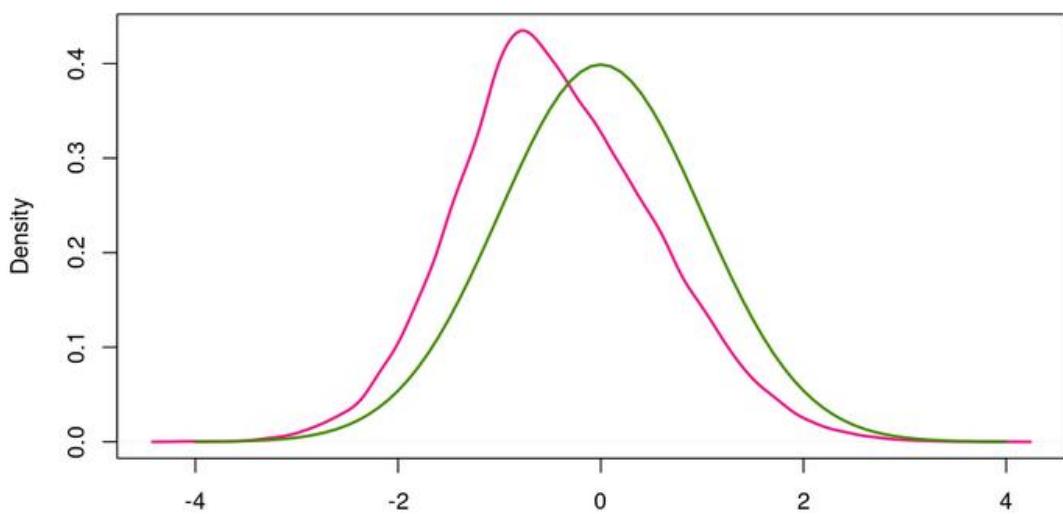
$p_{\{gas\}}$ - среднемесячная цена на газ (по 1 ЦЗ для 1 ЦЗ, по РФ для 2 ЦЗ),

$p_{\{coal\}}$ - среднемесячная цена на уголь (по РФ для 1 ЦЗ, по 2 ЦЗ для 2 ЦЗ),

$D_{\{Sat,t\}}, D_{\{Sun,t\}}, D_{\{Mon,t\}}$ - учитывают недельную сезонность,

$\varepsilon_t \sim N(0, \sigma) \text{ i. i. d.}$

Для тестирования наличия единичного корня будем использовать аналог t-теста для коэффициента ψ . Критическое значение теста Дики-Фуллера для такой модели было сгенерировано по методу Монте-Карло. Для левостороннего теста на 5% уровне значимости критическое значение теста равно -2.53. Плотность распределения статистики Дики-Фуллера и стандартного нормального распределения представлены на рисунке 25., результаты оценки для Модели 3Б аналога теста Дики Фуллера представлены на рисунке 26.



Примечание – Источник: Собственные расчеты.

Рисунок 25 – Плотность распределения статистики Дики-Фуллера (розовый) и стандартного нормального распределения (зеленый)

| Dependent variable: | | |
|-------------------------|----------------------------|---------------------------|
| | Δ log(consumer_price_1) | Δ log(consumer_price_2) |
| | (1) | (2) |
| lag(price, 1) | -0.117*** (0.021) | -0.184*** (0.024) |
| lag(Δ price, 1) | -0.347*** (0.028) | -0.248*** (0.032) |
| lag(Δ price, 2) | -0.209*** (0.028) | -0.150*** (0.032) |
| lag(Δ price, 3) | -0.158*** (0.027) | -0.139*** (0.031) |
| lag(Δ price, 4) | -0.137*** (0.026) | -0.151*** (0.030) |
| lag(Δ price, 5) | -0.176*** (0.026) | -0.054* (0.029) |
| lag(Δ price, 6) | -0.153*** (0.024) | -0.063** (0.026) |
| bdd_1 | -0.0001 (0.0001) | |
| cdd_1 | 0.002* (0.001) | |
| bdd_2 | | 0.001*** (0.0002) |
| cdd_2 | | -0.001 (0.005) |
| ipp_trend | 0.778*** (0.141) | 1.224*** (0.170) |
| coal_rf | 0.00005* (0.00003) | -0.00003 (0.0001) |
| gaz_rf | 0.00000 (0.00000) | -0.00000 (0.00001) |
| mon_dummy | 0.040*** (0.003) | 0.007 (0.006) |
| sat_dummy | -0.012*** (0.003) | 0.015** (0.006) |
| sun_dummy | -0.048*** (0.003) | -0.030*** (0.006) |
| Observations | 1,641 | 1,381 |
| R ² | 0.492 | 0.223 |
| Adjusted R ² | 0.488 | 0.214 |
| Residual Std. Error | 0.037 (df = 1626) | 0.072 (df = 1366) |
| F Statistic | 105.110*** (df = 15; 1626) | 26.125*** (df = 15; 1366) |

Note: *p<0.1; **p<0.05; ***p<0.01

Примечание – Источник: Собственные расчеты.

Рисунок 26 – Оценки для Модель 3Б аналога теста Дики Фуллера

Для обеих моделей гипотеза о наличии единичного корня отвергается. Рассмотрим в таком случае модель, где оценивается коэффициент при $p_{\{t-1\}}$.

Модель 3В AR(1) с регрессорами представлена уравнением (24):

$$\begin{aligned}
p_{\{t\}} = & \psi p_{\{t-1\}} + \mu + c_{\{HDD\}} * HDD_{\{t\}} + c_{\{CDD\}} * CDD_{\{t\}} + c_{\{IPP\}} * IPP_t \\
& + c_{\{gas\}} * p_{\{gas,t\}} + c_{\{coal\}} * p_{\{coal,t\}} + d_{\{Sat\}} D_{\{Sat,t\}} + d_{\{Sun\}} D_{\{Sun,t\}} \\
& + d_{\{Mon\}} D_{\{Mon,t\}} + \varepsilon_t,
\end{aligned} \tag{24}$$

где

$p_{\{t\}} = \ln(S_{\{t\}})$ - логарифм средней цены в ЦЗ за день,

$HDD_{\{t\}}$ - Heating-Degree-Days среднее по ЦЗ,

$CDD_{\{t\}}$ - Cooling-Degree-Days среднее по ЦЗ,

IPP - сглаженная трендовая составляющая ИПП,

$p_{\{gas\}}$ - среднемесячная цена на газ (по 1 ЦЗ для 1 ЦЗ, по РФ для 2 ЦЗ),

$p_{\{coal\}}$ - среднемесячная цена на уголь (по РФ для 1 ЦЗ, по 2 ЦЗ для 2 ЦЗ),

$D_{\{Sat,t\}}, D_{\{Sun,t\}}, D_{\{Mon,t\}}$ - учитывают недельную сезонность,

$\varepsilon_t \sim N(0, \sigma) \text{ i.i.d.}$

Согласно полученным оценкам можно предположить, что модель, предложенная [12], хоть и показа себя хорошо при решении задачи прогнозирования, может быть не лучшим выбором, если стоит задача получения интерпретируемых коэффициентов.

Использование байесовского подхода предпочтительно для решения поставленных в исследовании задач, поскольку он довольно легко позволяет учитывать в модели сложные компоненты, а также получать доверительные интервалы для компонент любого вида. Перейдем к анализу байесовских моделей.

Модель 1А отличается от Модели 1Б только подходом к оцениванию. Первая оценивается с помощью МНК, вторая с помощью байесовского подхода. Байесовский подход подразумевает наличие априорных распределений. Вышеперечисленные априорные распределения для коэффициентов модели 1А являются неинформативными. Как будет видно ниже по полученным апостериорным распределениям, дисперсия апостериорного распределения меньше дисперсии априорного минимум в 10, а для некоторых коэффициентов в более, чем в 1000 раз. Для первичного анализа и сравнения с классическими методами достаточно использования неинформативных распределений. Результаты оценивания модели 3В представлены на рисунке 27.

| <i>Dependent variable:</i> | | |
|----------------------------|---------------------------|---------------------------|
| | log(price_1) | log(price_2) |
| | (1) | (2) |
| lag(log(price_1), 1) | 0.759*** (0.015) | |
| hdd_1 | -0.0004*** (0.0001) | |
| cdd_1 | 0.003** (0.001) | |
| lag(log(price_2), 1) | | 0.689*** (0.018) |
| hdd_2 | | 0.002*** (0.0002) |
| cdd_2 | | -0.003 (0.005) |
| mon_dummy | 0.056*** (0.003) | 0.008 (0.006) |
| sat_dummy | -0.014*** (0.003) | 0.019*** (0.006) |
| sun_dummy | -0.050*** (0.003) | -0.034*** (0.006) |
| Constant | 1.713*** (0.108) | 2.077*** (0.121) |
| Observations | 1,820 | 1,505 |
| R ² | 0.666 | 0.671 |
| Adjusted R ² | 0.665 | 0.670 |
| Residual Std. Error | 0.040 (df = 1813) | 0.077 (df = 1498) |
| F Statistic | 601.612*** (df = 6; 1813) | 510.190*** (df = 6; 1498) |

Note:

*p<0.1; **p<0.05; ***p<0.01

Примечание – Источник: Собственные расчеты.

Рисунок 27 – Модель 3В AR(1) с регрессорами

Отметим, чем байесовский подход более привлекателен в рамках данного исследования. Часто при моделировании с помощью классических методов анализа регрессии игнорируются следующие фундаментальные проблемы.

Во-первых, за исключением гауссовой линейной модели с постоянной дисперсией и некоторых других моделей, основанных на выводе из нее, вычисленные р-значения имеют неизвестную точность для конкретного набора данных и модели. Причем при анализе последовательных экспериментов или адаптивных моделей, часто бывает так, что р-значения не может быть вычислено, и можно установить только общее α .

Например, одной из часто используемых моделей является бинарная логистическая модель. Для нее логарифм функции правдоподобия не является квадратичным, поэтому подавляющее большинство доверительных интервалов и р-значений, вычисленных для таких моделей, не очень точны. В отличие от классического подхода байесовские логистические модели, обеспечивают точный вывод.

Во-вторых, при частотном подходе главный способ анализа значимости коэффициентов является сравнение вероятности ошибки I рода с конкретным значением, причем чаще всего это 0,05, независимо от того растет размер выборки или нет. В частотном подходе не простого способа моделирования таких поправок.

Модель 4А Байесовская линейная регрессия представлена уравнением (25):

$$\begin{aligned}
 p_{\{t\}} = & p_{\{t-1\}} + \mu + c_{\{HDD\}} * HDD_{\{t\}} + c_{\{CDD\}} * CDD_{\{t\}} + c_{\{IPP\}} * IPP_t \\
 & + c_{\{gas\}} * p_{\{gas,t\}} + c_{\{coal\}} * p_{\{coal,t\}} + d_{\{Sat\}} D_{\{Sat,t\}} + d_{\{Sun\}} D_{\{Sun,t\}} \\
 & + d_{\{Mon\}} D_{\{Mon,t\}} + \varepsilon_t,
 \end{aligned} \tag{25}$$

где

$p_{\{t\}}$ = $\ln(S_{\{t\}})$ - логарифм средней цены в ЦЗ за день,

$HDD_{\{t\}}$ - Heating-Degree-Days среднее по ЦЗ,

$CDD_{\{t\}}$ - Cooling-Degree-Days среднее по ЦЗ,

IPP - сглаженная трендовая составляющая ИПП,

$p_{\{gas\}}$ - среднемесячная цена на газ (по 1 ЦЗ для 1 ЦЗ, по РФ для 2 ЦЗ),

$p_{\{coal\}}$ - среднемесячная цена на уголь (по РФ для 1 ЦЗ, по 2 ЦЗ для 2 ЦЗ),

$D_{\{Sat,t\}}, D_{\{Sun,t\}}, D_{\{Mon,t\}}$ - учитывают недельную сезонность,

$\varepsilon_t \sim N(0, \sigma)$ i. i. d..

Априорные распределения параметров представлены уравнениями (26)-(32):

$$\mu \sim N(0, 10) \tag{26}$$

$$c_{\{gas\}} \sim N(0, 10); c_{\{coal\}} \sim N(0, 10); \quad (27)$$

$$d_{\{Sat\}} \sim N(0, 10); d_{\{Sun\}} \sim N(0, 10); d_{\{Mon\}} \sim N(0, 10); \quad (28)$$

$$c_{\{CDD\}} \sim N(0, 10); c_{\{HDD\}} \sim N(0, 10); \quad (29)$$

$$c_{\{IPP\}} \sim N(0, 10); \quad (30)$$

$$c_{\{gas\}} \sim N(0, 10); c_{\{coal\}} \sim N(0, 10); \quad (31)$$

$$\sigma \sim Cauchy(0, 5) * 1_{\{(0, \infty)\}}. \quad (32)$$

Результаты оценки модели 4А для ценовой зоны 1 представлены на рисунке 28, для ценовой зоны 2 на рисунке 29.

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|------------|----------|-------|-----------------|---------|----------|----------|----------|
| μ | 25.73476 | 2 | 0.37008 | 0.95040 | -0.92836 | 0.33784 | 1.72426 |
| d_{Sat} | 1.07718 | 36 | -0.02438 | 0.01476 | -0.05105 | -0.02425 | 0.00032 |
| d_{Sun} | 1.23235 | 8 | -0.04559 | 0.01354 | -0.07139 | -0.04563 | -0.02130 |
| d_{Mon} | 1.26225 | 13 | 0.06938 | 0.02552 | 0.03559 | 0.06777 | 0.12472 |
| c_{CDD} | 1.17545 | 8 | 0.00380 | 0.00482 | -0.00546 | 0.00349 | 0.01371 |
| c_{HDD} | 2.42558 | 2 | -0.00004 | 0.00100 | -0.00198 | 0.00013 | 0.00151 |
| c_{IPP} | 21.47695 | 2 | -0.52968 | 0.73721 | -1.51883 | -0.49669 | 0.33932 |
| c_{gas} | 2.14065 | 3 | -0.00003 | 0.00002 | -0.00007 | -0.00003 | 0.00002 |
| c_{coal} | 7.78041 | 2 | 0.00042 | 0.00097 | -0.00120 | 0.00059 | 0.00166 |
| σ | 1.28569 | 9 | 0.17583 | 0.00384 | 0.16859 | 0.17590 | 0.18303 |

Примечание – Источник: Собственные расчеты.

Рисунок 28 – Результаты оценивания Модели 4А для 1 ЦЗ

Байесовские модели классически оцениваются с помощью метода Монте Карло Марковских цепей. В Stan используется алгоритм Метрополиса-Гастингса [32] или Монте Карло сэмплер по Гамильтону классический [33] и с No-U-Turn модификацией [34]. Для получения апостериорных оценок требуются относительно длительные вычисления (относительно классических методов), однако в результате можно проводить диагностику моделей, что является одним из преимуществ байесовских методов.

Для начала, до получения численных оценок необходимо определить число итераций для построения оценок. При выборе слишком большого числа итераций время на оценку моделей может расти слишком быстро, однако бывает и что число итераций алгоритма может быть недостаточным, и полученные оценки будут несостоительными. Если при оценивании по МСМС возникают проблемы еще на предварительном этапе, то обычно предлагается перезапустить алгоритм повторно с удвоенным числом итераций

[35]. Увеличение числа итераций в таком случае значительно повысит точность оценок, но кардинально не изменит результаты. В нашем анализе используется 5000 итераций, не считая итераций, относящихся к «фазе прожига».

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|------------|---------|-------|----------|---------|----------|----------|---------|
| μ | 4.23957 | 2 | 0.33493 | 0.77255 | -1.00406 | 0.47021 | 1.27865 |
| d_{Sat} | 1.05332 | 81 | 0.01454 | 0.01522 | -0.01534 | 0.01466 | 0.04546 |
| d_{Sun} | 1.03727 | 95 | -0.02096 | 0.01299 | -0.04690 | -0.02093 | 0.00487 |
| d_{Mon} | 1.03758 | 69 | 0.00220 | 0.01475 | -0.02491 | 0.00215 | 0.03250 |
| c_{CDD} | 1.04280 | 94 | -0.00934 | 0.01279 | -0.03287 | -0.00994 | 0.01599 |
| c_{HDD} | 1.24906 | 7 | -0.00014 | 0.00057 | -0.00115 | -0.00018 | 0.00103 |
| c_{IPP} | 8.32425 | 2 | -0.63380 | 0.87501 | -1.40739 | -1.07842 | 1.03292 |
| c_{gas} | 1.00205 | 57 | -0.00003 | 0.00002 | -0.00008 | -0.00003 | 0.00001 |
| c_{coal} | 1.90607 | 3 | 0.00062 | 0.00072 | -0.00027 | 0.00033 | 0.00183 |
| σ | 1.04175 | 69 | 0.21803 | 0.00413 | 0.20968 | 0.21802 | 0.22638 |

Примечание – Источник: Собственные расчеты.

Рисунок 29 – Результаты оценивания Модели 4А для 2 ЦЗ

Обычно считается, что параметры дисперсии и стандартного отклонения требуют большего числа итераций для достижения сходимости, чем параметры, влияющие на среднее. По полученным оценкам для моделей стохастической можно заметить, что n_eff для показателей дисперсии низкий по сравнению с n_eff для обычных регрессоров.

Показатель Rhat – это статистика потенциального уменьшения разброса, известной как статистика Гельмана-Рубина [36]. Грубо говоря, эта статистика представляет собой отношение дисперсии апостериорного распределения, полученного по объединенной по всем цепям выборке, к дисперсии полученным по каждой из цепей по отдельности. Чем дальше значение статистики от 1, тем хуже. В нашем анализе использовалось 4 цепи для каждой из байесовских моделей.

Техническая трудность, связанная с методами МCMC, заключается в том, что сгенерированные величины обычно автокоррелированы (или антикоррелированы) внутри цепи. Это увеличивает неопределенность оценки апостериорных интересующих величин, таких как средние, дисперсии или квантили [37].

Stan оценивает эффективный размер выборки (n_eff) для каждого параметра, который играет такую же роль в центральной предельной теореме для Монте-Карло Марковских цепей (MCMC CLT), какую играет число независимых случайных величин в

стандартной центральной предельной теореме (CLT). Stan для расчета n_{eff} использует статистику R_{hat} , чтобы учесть информацию между цепями при вычислении ESS.

Грубо говоря, эффективный размер выборки (ESS) интересующей величины отражает, сколько независимых розыгрышей содержит такой же объем информации, как и зависимая выборка, полученная с помощью алгоритма МCMC. Понятно, что чем выше ESS, тем лучше, поскольку выборочное среднее апостериорного распределения имеет дисперсию обратно пропорциональную ESS, смотри уравнение (33):

$$\bar{X}_{\{n\}} \approx \sim N\left(\mu, \frac{\{\sigma^2\}}{\{n\}}\right) \quad (33)$$

В математической теории случайных процессов центральная предельная теорема для цепей Маркова имеет вывод, несколько схожий с заключением классической центральной предельной теоремы (ЦПТ) из теории вероятностей, но величины дисперсии предельного распределения определяются несколько сложнее, представленная уравнениеми (34)-(37):

$$\mu = E(g(X_{\{1\}})) \quad (34)$$

$$\sigma^2 = \text{Var}(g(X_{\{1\}})) + 2 \sum_{\{k=1\}}^{\{\infty\}} \text{Cov}(g(X_{\{1\}}), g(X_{\{1+k\}})) \quad (35)$$

$$\hat{\mu}_{\{n\}} = \frac{\{1\}}{\{n\}} \sum_{\{k=1\}}^{\{n\}} g(X_{\{k\}}) \quad (36)$$

$$\hat{\mu}_{\{n\}} \approx N\left(\mu, \frac{\{\sigma^2\}}{\{n\}}\right), n \rightarrow \infty \quad (37)$$

Степень, на которую автокорреляция в цепях увеличивает неопределенность в оценках, может быть измерена с помощью эффективного размера выборки (ESS или n_{eff}). Для независимых выборок центральная предельная теорема неопределенность оценок на основе объема выборки N . Учитывая зависимость выборки, количество независимых выборок заменяется эффективным размером выборки n_{eff} , которая представляет собой количество независимых выборок с той же мощностью оценивания, что и N автокоррелированных наблюдений.

Соответственно, обычно n_{eff} для классического алгоритма МCMC будет меньше N . Часто для достижения желаемой точности оценивания достаточно относительно

небольшого ESS, но для оценки самой ESS и диагностики сходимости требуется более высокие показатели ESS. Рекомендуется, чтобы ESS превышал количество цепей минимум в 100 раз. Например, в нашем случае, при использовании четырех цепей необходим эффективный размер выборки не менее 400.

Эффективный размер выборки n_{eff} при оценивании в Stan может быть больше, чем N , если полученная цепь Маркова является антитетической и имеет отрицательную автокорреляцию для нечетных лагов. Алгоритм выборки без разворота (NUTS), используемый в Stan, может привести к $n_{eff} > N$ для параметров, которые имеют близкие к гауссовским апостериорным распределениям и мало зависят от других параметров.

Если при диагностике с какими-то из вышеперечисленных статистик наблюдаются проблемы, то рекомендуется увеличить число итераций после периода прожига, рассмотреть график рангов, чтобы увидеть, различия между цепями, рассмотреть график локальной и квантильной эффективности. Также можно посмотреть, как меняется ESS при увеличении количества итераций. Если R-hat меньше 1,01 и ESS линейно растет с количеством итераций и в конечном итоге превышает рекомендуемый предел, цепи смешиваются достаточно, но MCMC имеет высокую автокорреляцию, требующую большого количества итераций. Обычно нужно беспокоиться только в том, случае, когда это число меньше 1/100 или 1/1000 от вашего количества итераций.

В данном случае для всех коэффициентов показатель n_{eff} очень маленький, что еще раз подтверждает то, что для данной модели спецификация предложенная [12] не является подходящей для имеющихся данных.

Перед тем как делать какие-то выводы о выборе функциональной формы рассмотрим также Модель стохастической волатильности для аналогичной модели.

Модель 4Б Байесовская модель стохастической волатильности с регрессорами, представлена уравнением (38):

$$\begin{aligned}
 p_{\{t\}} = & p_{\{t-1\}} + \mu + c_{\{HDD\}} * HDD_{\{t\}} + c_{\{CDD\}} * CDD_{\{t\}} + c_{\{IPP\}} * \ln(IPP_t) \\
 & + c_{\{gas\}} * \ln(p_{\{gas,t\}}) + c_{\{coal\}} * \ln(p_{\{gas,t\}}) + d_{\{sat\}} D_{\{sat,t\}} + d_{\{sun\}} D_{\{sun,t\}} \\
 & + d_{\{mon\}} D_{\{mon,t\}} + \epsilon_t^{\{(1)\}} * \exp(h_t/2), \\
 h_t = & \kappa + \phi(h_{\{t-1\}} - \kappa) + \epsilon_{\{t-1\}}^{\{(2)\}} \sigma
 \end{aligned} \tag{38}$$

где

$p_{\{t\}} = \ln(S_{\{t\}})$ - логарифм средней цены в ЦЗ за день,

$HDD_{\{t\}}$ - Heating-Degree-Days среднее по ЦЗ,

$CDD_{\{t\}}$ - Cooling-Degree-Days среднее по ЦЗ,

IPP - сглаженная трендовая составляющая ИПП,

$p_{\{gas\}}$ - среднемесячная цена на газ (по 1 ЦЗ для 1 ЦЗ, по РФ для 2 ЦЗ),

$p_{\{coal\}}$ - среднемесячная цена на уголь (по РФ для 1 ЦЗ, по 2 ЦЗ для 2 ЦЗ),

$D_{\{Sat,t\}}, D_{\{Sun,t\}}, D_{\{Mon,t\}}$ - учитывают недельную сезонность,

$\epsilon_t^{\{(1)\}} \sim N(0, \sigma) i.i.d, \epsilon_t^{\{(2)\}} \sim N(0, \sigma) i.i.d.$

Априорные распределения параметров, представлены уравнениями (39)-(47):

$$\mu \sim N(0, 10) \quad (39)$$

$$\psi \sim N(0, 10); \quad (40)$$

$$d_{\{Sat\}} \sim N(0, 10); d_{\{Sun\}} \sim N(0, 10); d_{\{Mon\}} \sim N(0, 10); \quad (41)$$

$$c_{\{CDD\}} \sim N(0, 10); c_{\{HDD\}} \sim N(0, 10); \quad (42)$$

$$c_{\{IPP\}} \sim Laplace(0, 1); \quad (43)$$

$$c_{\{gas\}} \sim Laplace(0, 1); c_{\{coal\}} \sim Laplace(0, 1); \quad (44)$$

$$\sigma \sim Cauchy(0, 5) * 1_{\{(0, \infty)\}}; \quad (45)$$

$$\phi \sim U(-1, 1); \quad (46)$$

$$\kappa \sim Cauchy(0, 10); \quad (47)$$

Сравним модели 4А и 4Б. Первая модель представлена для сравнения с моделью стохастической волатильности, а также с результатами, полученными по Prophet-модели (в которой так же, как и в модели 1 никак не учитывались особенности дисперсии ошибки).

В таблицах ниже представлены результаты оценивания моделей 4А и 4Б для 1 и 2 ЦЗ. Стоить отметить, что для Модели 2 регрессоры $IPP, p_{\{gas\}}, p_{\{coal\}}$ взяты с логарифмом, поэтому при сравнении моделей имеет смысл сравнивать не полученные значения оценок, а знаки (вспоминая также полученные знаки в модели Prophet).

В качестве априорного распределения параметров в Модели 1Б для всех коэффициентов взят $N(0, 10)$. Однако в Модели 4Б для регрессоров $IPP, p_{\{gas\}}, p_{\{coal\}}$, влияние которых на цены не подтвердилось классическими регрессионными моделями (к тому же был получен неправильный знак, противоречащий логике и выдвинутым гипотезам), в качестве априорного было выбрано двойное экспоненциальное распределение со средним в 0, поскольку оно является ``стягивающим'' к 0, что является неким аналогом байесовской регуляризации [38].

В Модели 4А используется классическая версия модели стохастической волатильности, более сложные компоненты (диффузионная с эффектом рычага и скачки, как в Kostrzewski) будут добавлены на следующих этапах работы. Диагностику сходимости МСМС по распределению для каждого из коэффициентов можно проводить по статистике Rhat. Если Rhat сильно отклоняется от 1, значит цепи не очень хорошо смешались, что обычно отражает проблемы в самой модели, а не в процедуре оценивания [39].

Заметим, что в Модели 4А как для 1 ЦЗ, так и для 2 для многих коэффициентов Rhat намного больше 1, что отражает тот факт, что модели с постоянной волатильностью не подходят для анализа этих рядов.

В большинстве оцененных моделей (а также в моделях, не представленных в отчете) единственными значимыми коэффициентами были коэффициенты при дамми-переменных. Причем знаки полученных оценок согласуются, как между моделями 1 и 2, так и со знаками, полученными в Prophet модели: для 1 ЦЗ при прочих равных цены на электричество в субботу и воскресенье в среднем ниже, а в понедельник выше, для 2 ЦЗ -- в среднем ниже только в воскресенье. Результаты оценивания представлены на рисунках 30 и 31 для ценовой зоны 1 и ценовой зоны 2, соответственно.

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|------------|--------|-------|----------------|--------|---------|---------|---------|
| μ | 1.0034 | 300 | -0.0904 | 0.1108 | -0.3312 | -0.0848 | 0.1124 |
| d_{Sat} | 1.0074 | 292 | -0.0182 | 0.0028 | -0.0237 | -0.0182 | -0.0129 |
| d_{Sun} | 1.0067 | 550 | -0.0444 | 0.0025 | -0.0493 | -0.0444 | -0.0396 |
| d_{Mon} | 1.0154 | 434 | 0.0658 | 0.0027 | 0.0608 | 0.0658 | 0.0712 |
| c_{CDD} | 1.0047 | 952 | 0.0005 | 0.0010 | -0.0014 | 0.0005 | 0.0024 |
| c_{HDD} | 1.0055 | 1345 | 0.0001 | 0.0001 | -0.0001 | 0.0001 | 0.0003 |
| c_{IPP} | 1.0058 | 511 | -0.0694 | 0.9457 | -2.1004 | -0.0308 | 1.7872 |
| c_{gas} | 1.0101 | 417 | 0.0043 | 0.0105 | -0.0163 | 0.0043 | 0.0245 |
| c_{coal} | 1.0040 | 235 | 0.0086 | 0.0224 | -0.0316 | 0.0070 | 0.0551 |
| σ | 1.1361 | 45 | 0.6114 | 0.0638 | 0.4940 | 0.6095 | 0.7498 |
| κ | 1.0030 | 279 | -6.4263 | 0.0889 | -6.6007 | -6.4250 | -6.2496 |
| ϕ | 1.1294 | 50 | 0.8144 | 0.0389 | 0.7294 | 0.8159 | 0.8851 |

Примечание – Источник: Собственные расчеты.

Рисунок 30 – Результаты оценивания Модели 4Б для 1 ЦЗ

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|------------|--------|-------|----------------|--------|----------------|----------------|----------------|
| mu | 1.0636 | 83 | -0.0884 | 0.2499 | -0.5529 | -0.0910 | 0.4050 |
| d_{Sat} | 1.0692 | 50 | 0.0112 | 0.0042 | 0.0029 | 0.0113 | 0.0189 |
| d_{Sun} | 1.0311 | 120 | -0.0225 | 0.0032 | -0.0289 | -0.0225 | -0.0162 |
| d_{Mon} | 1.0507 | 100 | 0.0117 | 0.0042 | 0.0034 | 0.0117 | 0.0199 |
| c_{CDD} | 1.0331 | 144 | -0.0005 | 0.0047 | -0.0094 | -0.0005 | 0.0085 |
| c_{HDD} | 1.0121 | 561 | 0.0002 | 0.0001 | -0.0000 | 0.0002 | 0.0004 |
| c_{IPP} | 1.0271 | 257 | -0.1319 | 1.0087 | -2.3814 | -0.0730 | 1.7860 |
| c_{gas} | 1.0051 | 382 | -0.0021 | 0.0084 | -0.0176 | -0.0023 | 0.0153 |
| c_{coal} | 1.0648 | 89 | 0.0159 | 0.0412 | -0.0655 | 0.0161 | 0.0940 |
| σ | 1.0867 | 45 | 0.8549 | 0.0604 | 0.7365 | 0.8539 | 0.9707 |
| κ | 1.0324 | 195 | -5.3745 | 0.1387 | -5.6462 | -5.3797 | -5.0895 |
| ϕ | 1.0521 | 65 | 0.8455 | 0.0233 | 0.7980 | 0.8457 | 0.8899 |

Примечание – Источник: Собственные расчеты.

Рисунок 31 – Результаты оценивания Модели 4Б для 2 ЦЗ

Важный результат заключается в том, что знак при переменной HDD в модели стохастической волатильности положительный для обеих ЦЗ. Согласно результатам полученным по Prophet модели, а так же по Модели 4А, в 1 ЦЗ, более высоким значениям HDD соответствуют более низкие цен, что противоречит логике и вызывает сомнения в правильности спецификации данной модели. Но при использовании модели стохастической волатильности мы получаем положительные оценки коэффициентов для обоих регрессоров. Вероятно, при добавлении в Prophet модель диффузионной компоненты можно будет улучшить оценки годовой компоненты.

По модели 4Б оценки коэффициентов при IPP, как для 1, так и для 2 модели имеют очень широкие доверительные интервалы, в сравнении с остальными коэффициентами. Причиной этого может быть тот факт, что данный регрессор не подходит для данного анализа и необходимо использовать другой индекс деловой активности.

Также отметим, что по модели 4Б (в отличие от 4А модели) для 1 ЦЗ знаки при коэффициентах цен на ресурсы положительные, а для 1В ЦЗ -- знак при ценах на уголь. Однако для всех регрессоров, кроме дамми дней недели, 95-% доверительные интервалы содержат 0. Вероятно, что на это могли повлиять ``стягивающие'' к нулю априорные распределения.

Также для модели 4А показатель R_{hat} для некоторых показателей сильно превышают 1, хотя для модели 4Б, напротив, близки к 1. На основе этого анализа было решено, что упрощенная Модель (2) не подходит для анализа временных рядов по электричеству для РФ. Следовательно модель стохастической волатильности лучше подходит для имеющихся данных.

Также для модели 4Б, несмотря на то, что в ней параметр логарифма волатильности оценивался для каждого наблюдения, т.е. число оцениваемых параметром в раз больше, значения n_{eff} для многих коэффициентов в несколько раз превышают n_{eff} для модели 4А, что опять же является подтверждением того, что модель стохастической волатильности лучше подходит для имеющихся данных.

Однако, даже учитывая, что n_{eff} для модели 4Б выше, чем для модели 4А, этот показатель достаточно мал, учитывая, что используется 4 цепи Маркова для порядка 1500-2000 наблюдений. Следовательно, n_{eff} для хорошей модели должен быть не менее 200-400 для каждого коэффициента.

Значительные проблемы с для показателя n_{eff} , вероятно, вызваны неправильной спецификацией модели и могут быть решены с помощью правильной параметризации. В частности, можно отделить трендовую и сезонную компоненту (использовав аналогичные Prophet-модели компоненты) от «стохастических» остатков, которые уже будут моделироваться по модели Kostrzewski, представленная уравнением (48):

$$\begin{aligned} r_{\{t\}} &= y_{\{t\}} - T_{\{t\}} - S_{\{t\}} \\ r_{\{t\}} &= r_{\{t-1\}} + \mu + \psi X_{\{t\}} + \sqrt{\{\exp(h_{\{t-1\}})\}} \epsilon_{\{t\}}^{\{(1)\}} + J_t \end{aligned} \quad (48)$$

где

$y_{\{t\}}$ – исходный ряд,

$T_{\{t\}}$ – трендовая составляющая,

$S_{\{t\}}$ – сезонная составляющая,

$r_{\{t\}}$ – случайная составляющая,

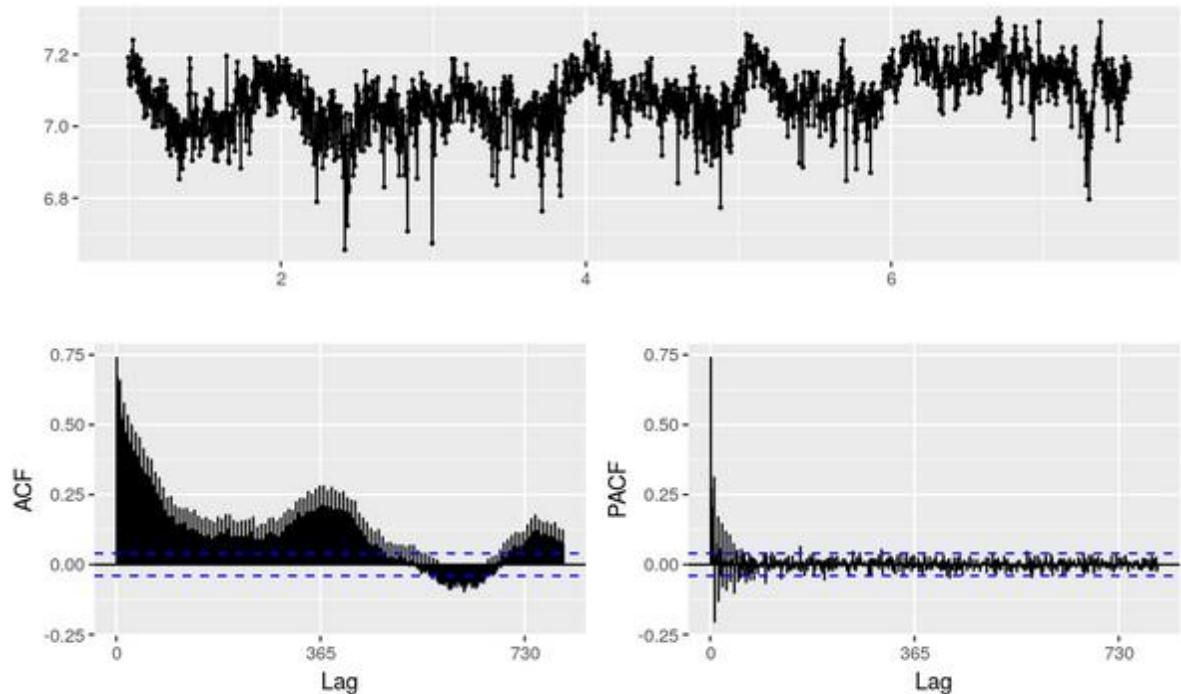
$X_{\{t\}}$ – изменение случайных факторов, например, температуры,

$h_{\{t-1\}}$ – ненаблюдаемые переменные,

J_t – скачок.

4 Результаты моделирования и их интерпретация

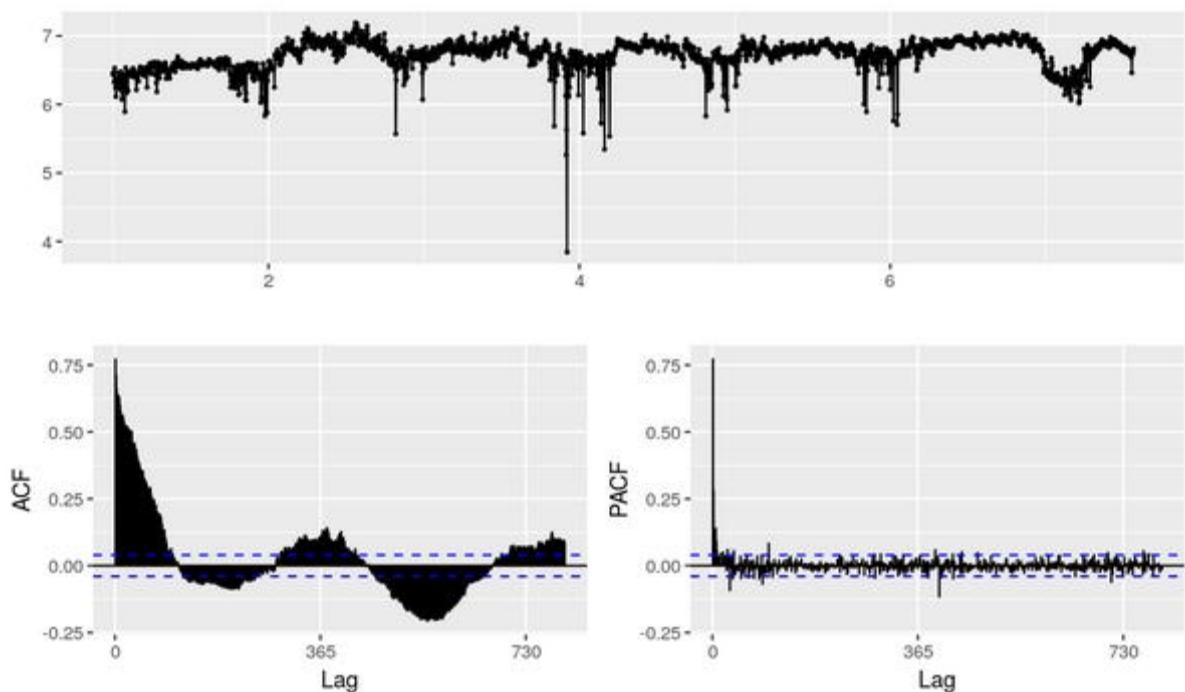
Для того, чтобы детально проанализировать результаты моделирования по Prophet модели цен на электричество для 1 и 2 ЦЗ, рассмотрим каждую из оцененных компонент ряда цен по отдельности. На рисунках 32-33 представлен логарифм цен на электроэнергию на PCB в первой и второй ценовых зонах, их автокорреляционные функции и функции частной автокорреляции.



Примечание – Источник: АТС и собственные расчеты.

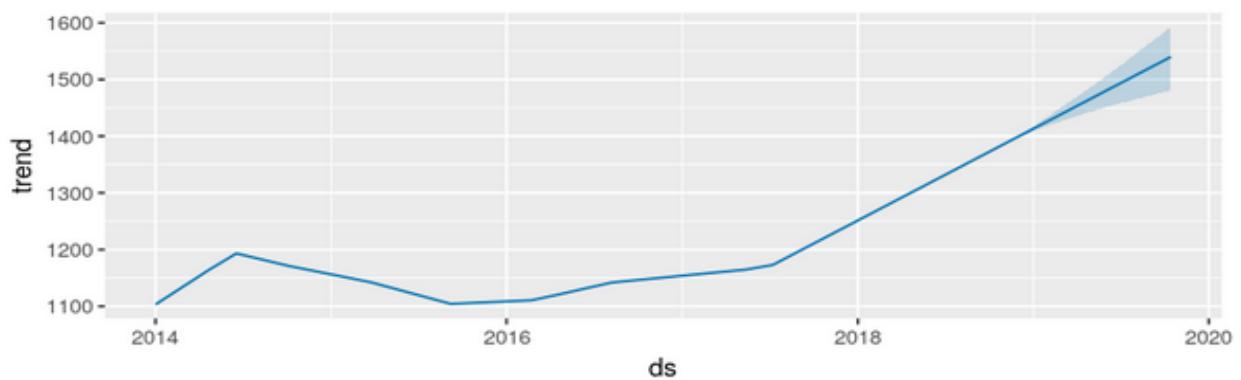
Рисунок 32 – Логарифм цен на электроэнергию на PCB в 1 ЦЗ, их автокорреляционные функции и частные автокорреляционные функции

Во-первых, рассмотрим оцененную трендовую компоненту ряда для цен на электроэнергию в 1 и 2 ЦЗ, рисунки 34 и 35, соответственно. В 1 ЦЗ трендовая составляющая цен выросла на 8,1 % от 1103,2 до 1193,1 за период с 2014-01-01 по 2014-06-08. Далее обратно снижалась до уровня 1104,2 в 2015-09-03 и постепенно росла. За период с 2018-01-01 по 2019-01-01 трендовая составляющая ряда цен выросла на 12,9%. При этом годовая инфляция за 2018 год составила 4%. Если данная модель верно оценивает трендовую компоненту, то рост цен в 1 ЦЗ превышает на 8,9 п.п. Стоит отметить, что данная модель откалибрована не идеально, могут иметь место влияние и других неучтенных факторов, и т.д. В любом случае такой рост указывает на возможные проблемы, то есть требует более внимательного анализа с помощью более сложных моделей.



Примечание – Источник: АТС и собственные расчеты.

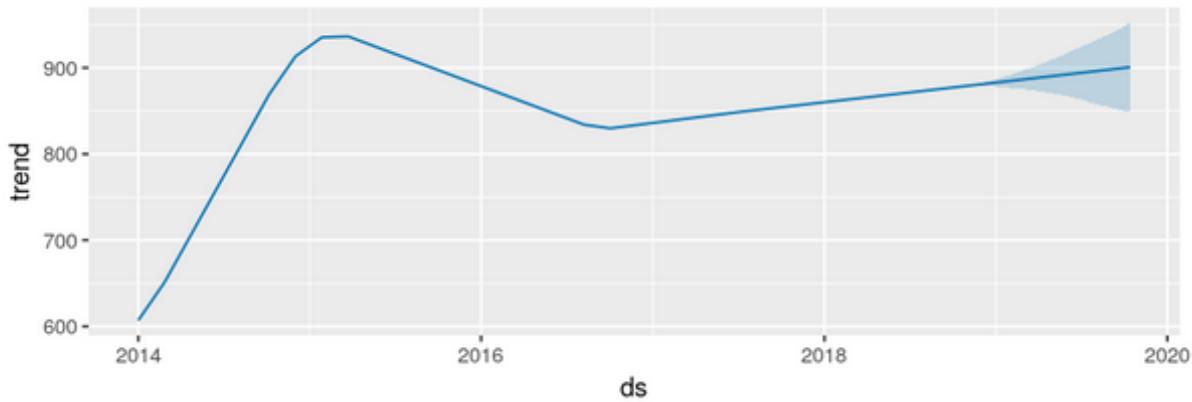
Рисунок 33 – Логарифм цен на электроэнергию на РСВ во 2 ЦЗ, их автокорреляционные функции и частные автокорреляционные функции



Примечание – Источник: Собственные расчеты.

Рисунок 34 – Трендовая компонента цен в 1 ЦЗ

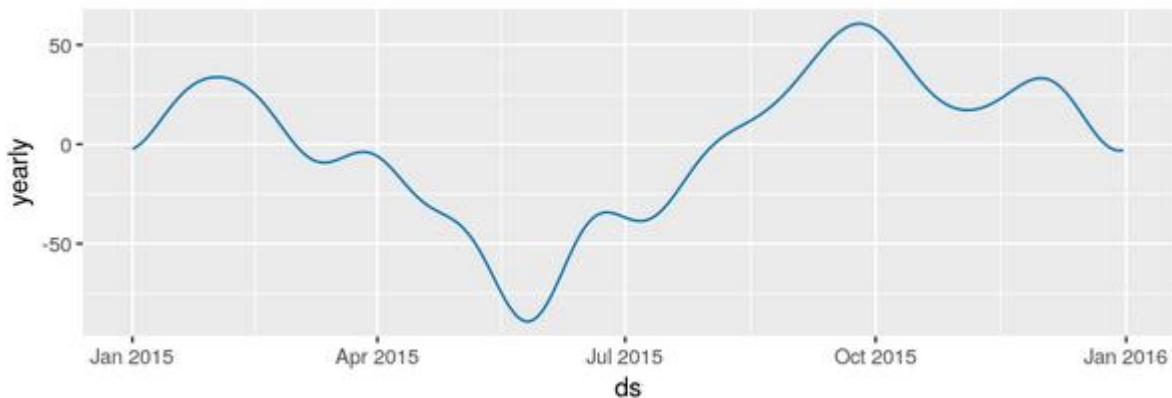
Во второй ЦЗ наблюдался резкий рост трендовой составляющей с 2014-01-01 по 2015-03-24, когда трендовая составляющая цен достигла своего максимума – 936,4. Частично это может быть связано с тем, что визуально можно заметить, что динамика цен до 2015 года значительно отличается от остальных периодов. Скорее всего Prophet модель увидела структурные изменения в динамике этого ряда, что и отразилось в достаточно резком росте трендовой компоненты, однако, если не учитывать эту неоднородность ряда напрямую, при более тщательном анализе цен во 2 ЦЗ рекомендуется рассматривать ряд, начиная с 2015-01-01.



Примечание – Источник: Собственные расчеты.

Рисунок 35 – Трендовая компонента цен в 2 ЦЗ

Во-вторых, рассмотрим подробнее внутригодовые сезонные колебания, параллельно с оцененной компонентой влияния температуры, представленных на рисунках 36, 37 для ценовой зоны 1, соответственно (для второй ценовой зоны смотри рисунки 39, 40, соответственно).

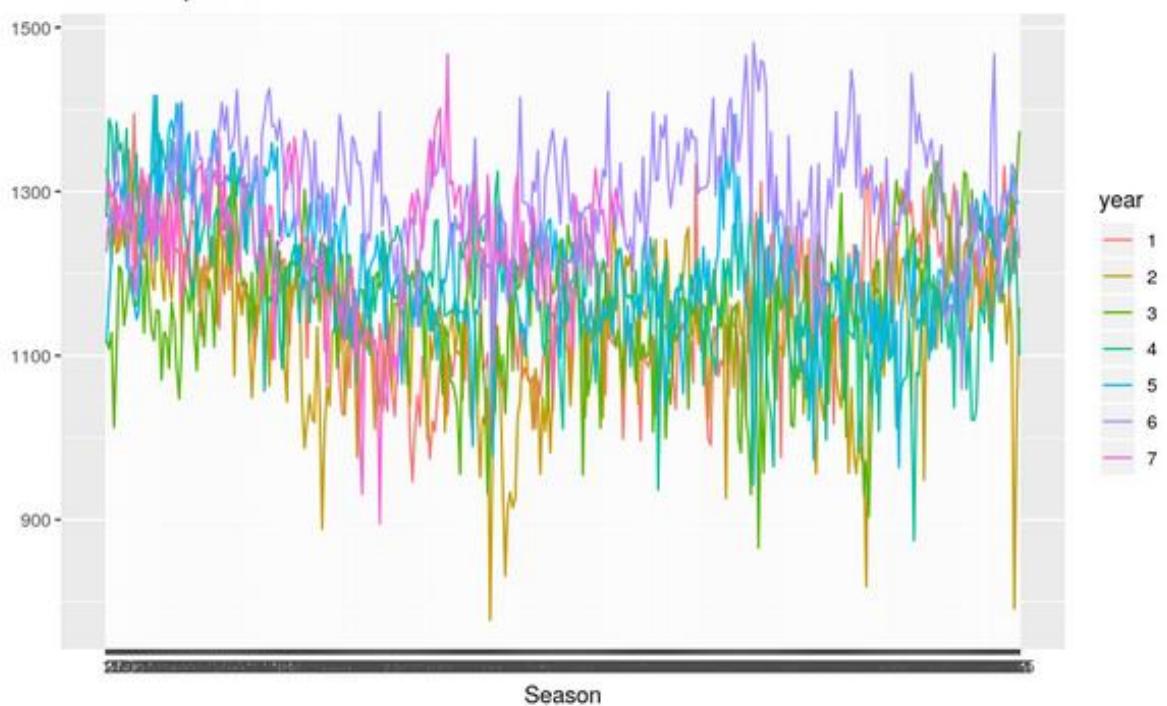


Примечание – Источник: Собственные расчеты.

Рисунок 36 – Оцененная компонента годовой сезонности цен в 1 ЦЗ

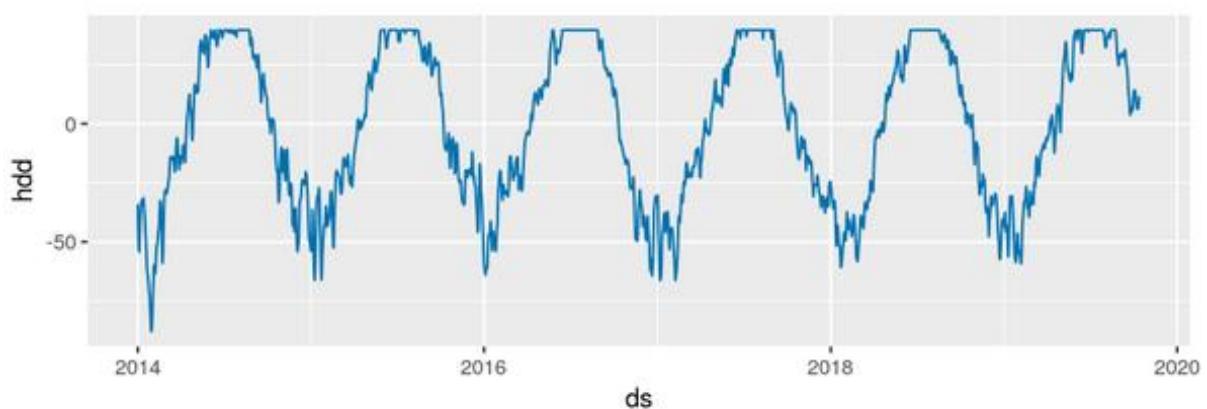
В обеих ЦЗ годовая динамика, не зависящая от температуры, в обеих ЦЗ довольно похожа: в 1 ЦЗ в цены на электричество в период с марта по сентябрь ниже, чем в остальные месяцы года, во 2 ЦЗ – в период с мая по сентябрь.

Для 2 ЦЗ, как и предполагалось ранее, более высоким значениям HDD в период отопительного сезона соответствуют более высокие цены. В 1 ЦЗ, более высоким значениям HDD соответствуют более низкие цен, что противоречит логике и вызывает сомнения в правильности спецификации данной модели, смотри рисунок 38 для первой ценовой зоны и рисунок 41 для второй ценовой зоны.



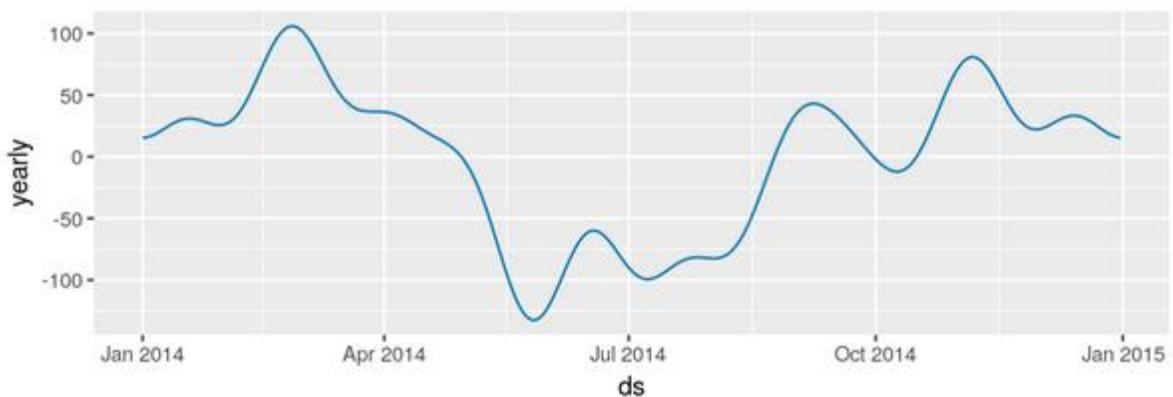
Примечание – Источник: Собственные расчеты.

Рисунок 37 – Сезонные графики годовой динамики для цен в 1 ЦЗ



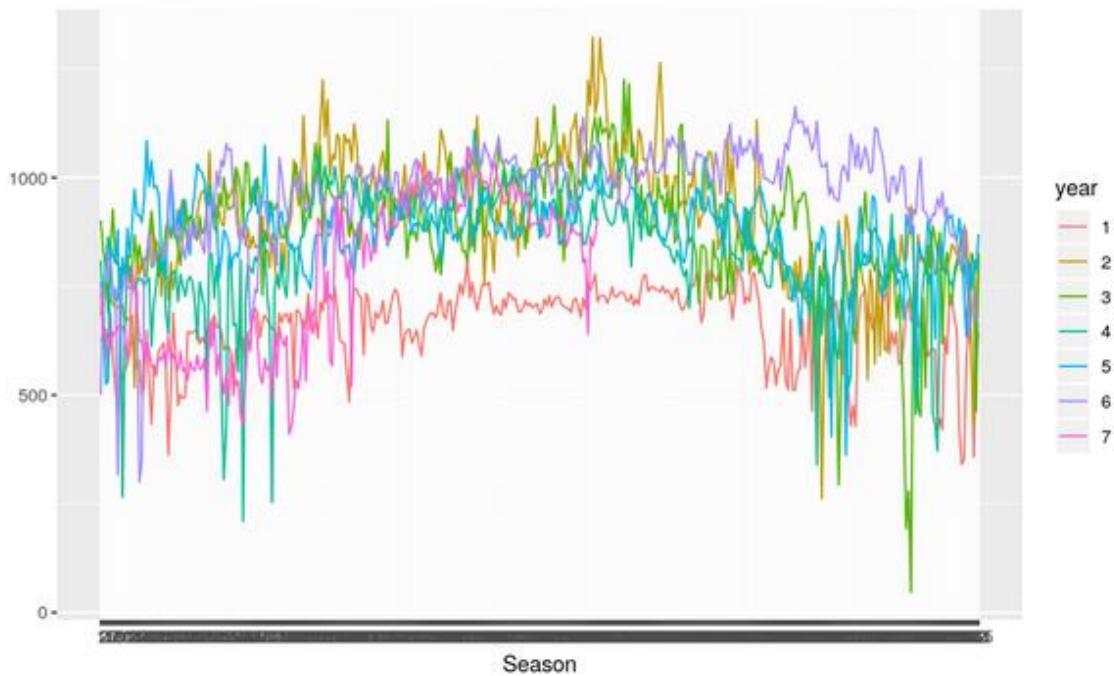
Примечание – Источник: Собственные расчеты.

Рисунок 38 – Влияние компоненты HDD на цены в 1 ЦЗ



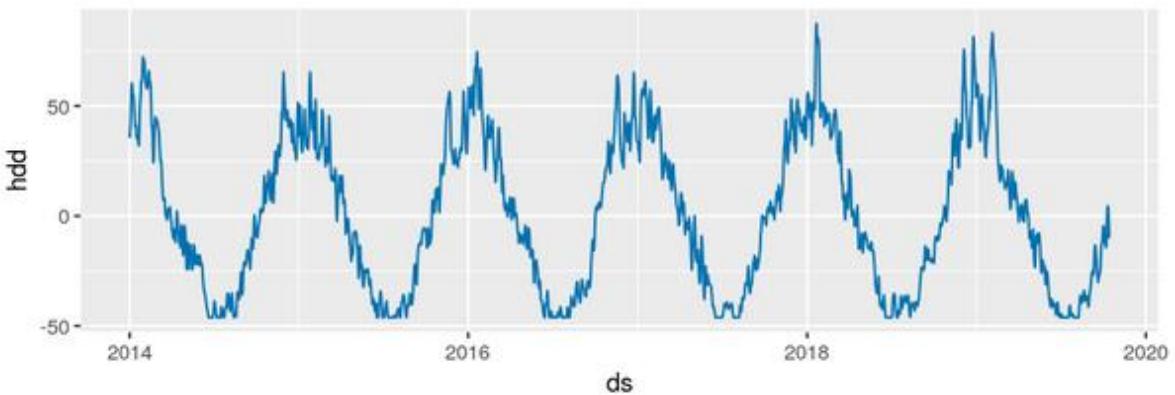
Примечание – Источник: Собственные расчеты.

Рисунок 39 – Оцененная компонента годовой сезонности цен во 2 ЦЗ



Примечание – Источник: Собственные расчеты.

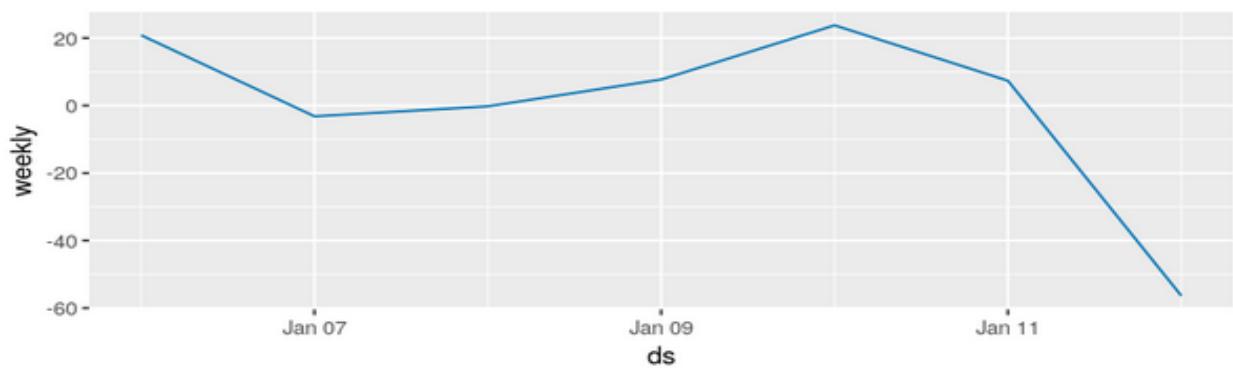
Рисунок 40 – Сезонные графики годовой динамики для цен во 2 ЦЗ



Примечание – Источник: Собственные расчеты.

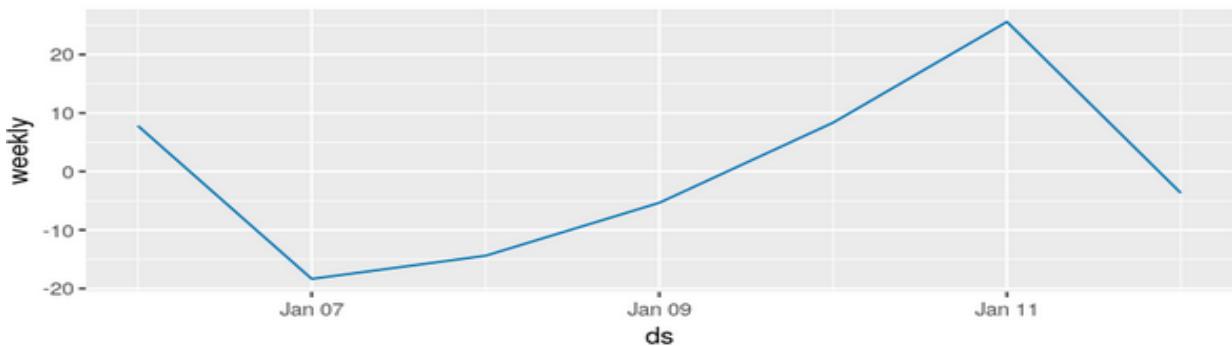
Рисунок 41 – Влияние компоненты HDD на цены во 2 ЦЗ

Далее перейдем к анализу недельной динамики и влияния праздничных дней, на рисунках 42 и 43 представлены оцененные компоненты недельной сезонности цен, а на рисунках 44 и 45 праздничных дней, для первой и второй ценовой зоны, соответственно. В обеих ЦЗ цены на электричество в понедельник выше, чем в остальные будние дни. В 1 ЦЗ цены в субботу ниже, чем в пятницу, во 2 ЦЗ – выше. В 1 ЦЗ, как и предполагалось, в воскресные и праздничные дни цен на электричество ниже, чем в остальные дни. Во 2 ЦЗ, как и предполагалось, в воскресные дни цены на электричество ниже, чем в остальные дни. Однако в праздничные дни цены выше, чем в остальные дни, что противоречит предположениям.



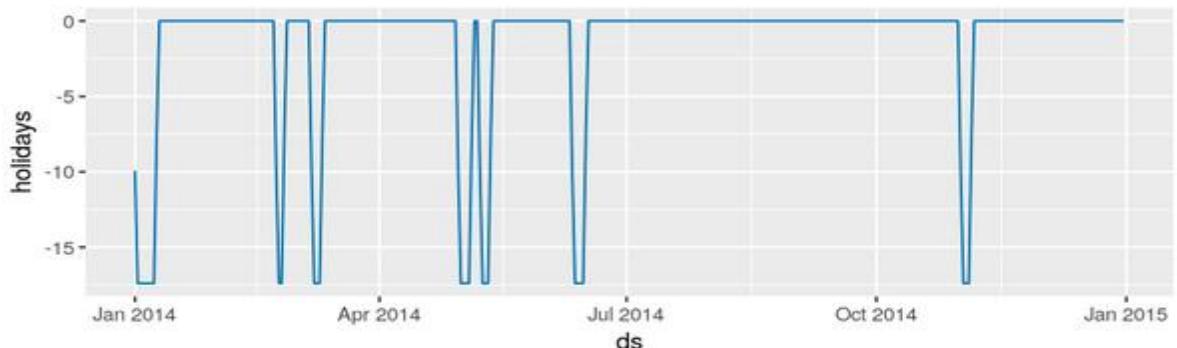
Примечание – Источник: Собственные расчеты.

Рисунок 42 – Оцененная компонента недельной сезонности цен в 1 ЦЗ



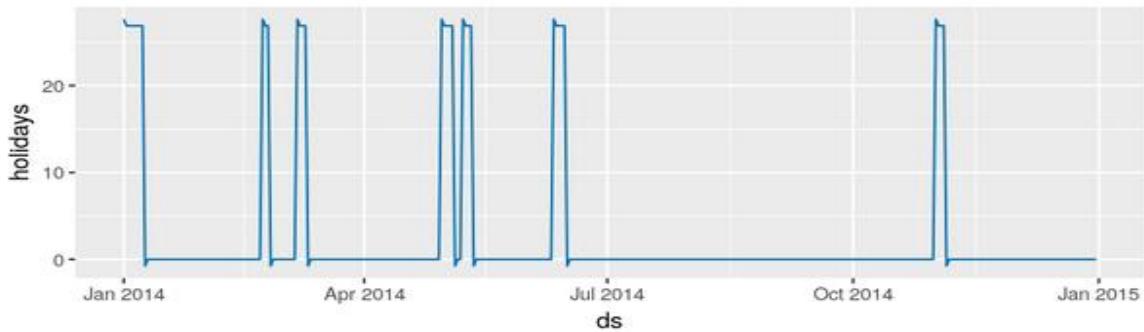
Примечание – Источник: Собственные расчеты.

Рисунок 43 – Оцененная компонента недельной сезонности цен во 2 ЦЗ



Примечание – Источник: Собственные расчеты.

Рисунок 44 – Оцененная компонента влияния праздничных дней на цены в 1 ЦЗ

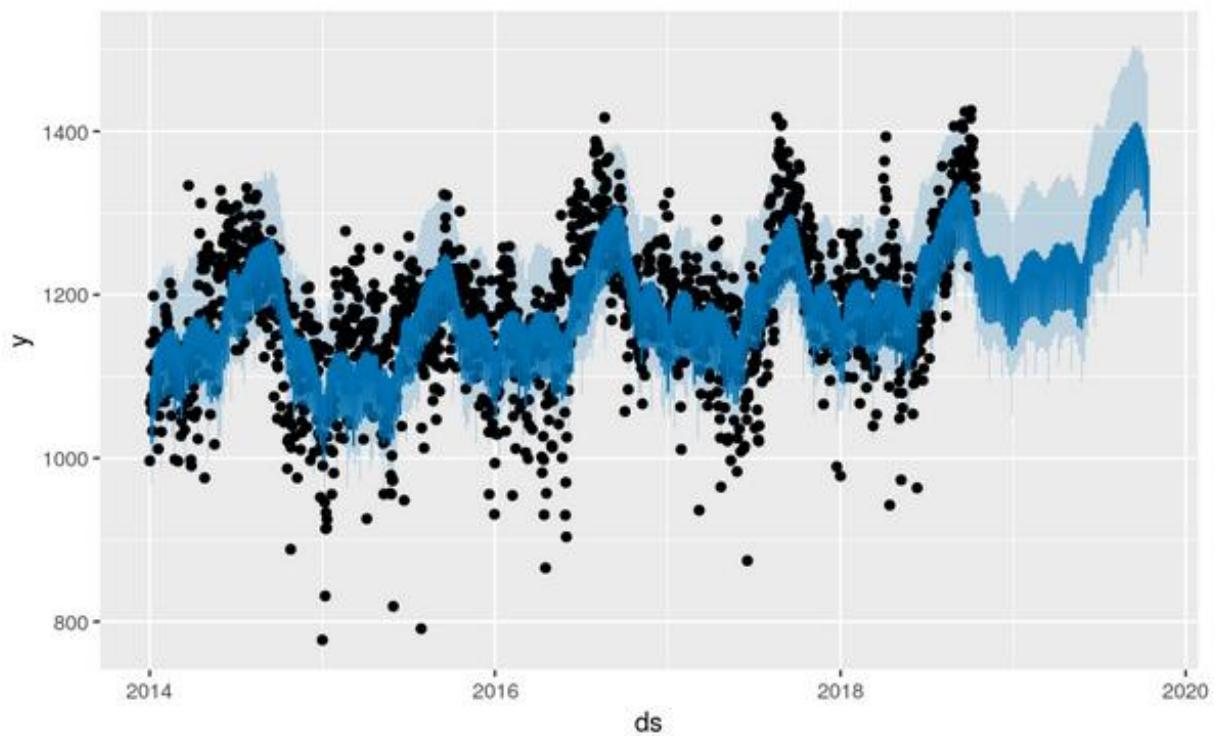


Примечание – Источник: Собственные расчеты.

Рисунок 45 – Оцененная компонента влияния праздничных дней на цены во 2 ЦЗ.

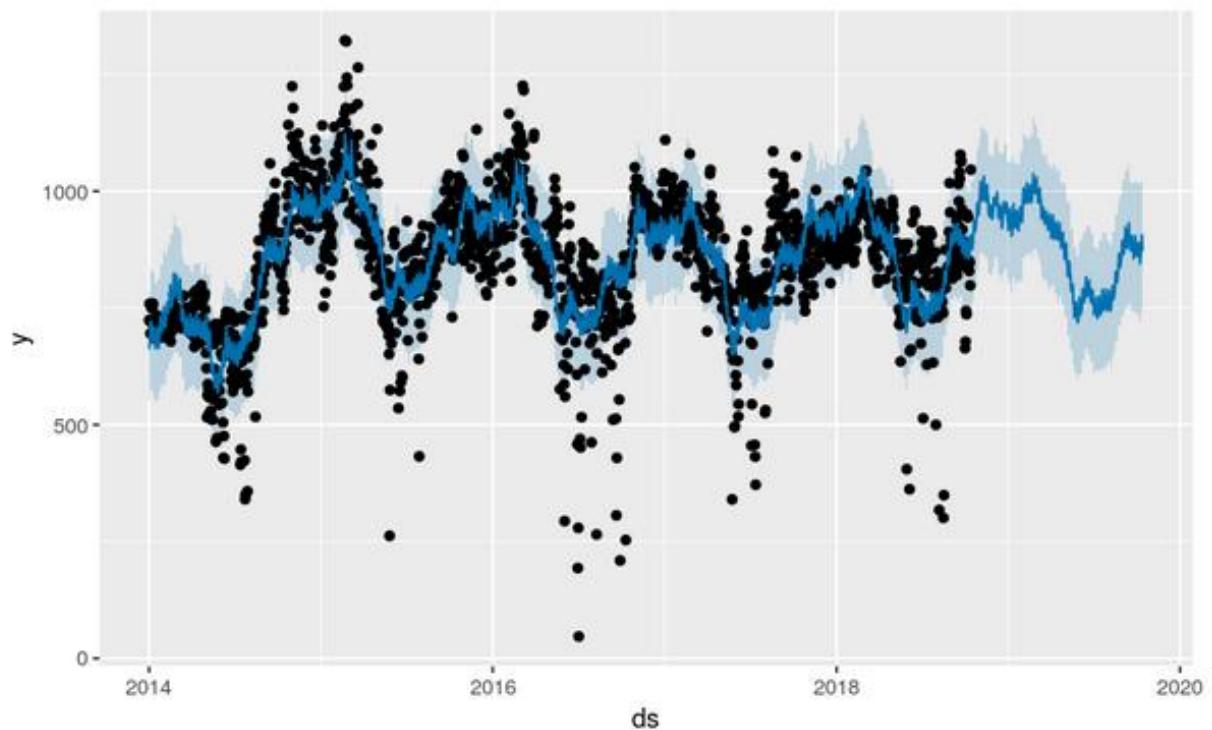
Частично это может быть объяснено тем, что в первой ценовой зоне доля предприятий легкой промышленности выше, чем во второй, в во второй ценовой зоне выше доля предприятий непрерывного технологического цикла, работающих в таких отраслях, как металлургия, газо-нефтяная отрасль. Предприятиям, занимающимся тяжелой промышленностью зачастую невыгодно останавливать производство, поэтому высокоенергозатратные производства функционируют даже в выходные дни.

В качестве проверки и сравнения модели Prophet с альтернативными моделями можно использовать среднеквадратическую ошибку как внутри, так и вневыборочного прогноза. Предполагается, что модель, идеально учитывая все компоненты и внешние факторы, должна иметь преимущества во вневыборочном прогнозе. Результаты прогноза представлены на рисунках 46 и 47 для первой и второй зоны, соответственно.



Примечание – Источник: Собственные расчеты.

Рисунок 46 – Прогноз на 2019 год для 1 ЦЗ.



Примечание – Источник: Собственные расчеты.

Рисунок 47 – Прогноз на 2019 год для 2 ЦЗ.

Отметим, что задача прогнозирования в рамках данной работы не основная, главная задача получить интерпретируемое тренд-сезонное разложение, поэтому при построении вне-выборочного прогноза мы можем пользоваться известными наблюдаемыми значениями внешних факторов, температуры и цен. В классической задаче прогнозирования это невозможно, приходится использовать прогнозные значения этих параметров, что приводит к более широким доверительным интервалам прогноза.

Перейдем к Шагу 3 стратегии эмпирических оценок. Теперь вместо упрощенной модели (2), представленной уравнением (49):

$$\begin{aligned}
 p_{\{t\}} = & p_{\{t-1\}} + \mu_{\{win\}} \cdot D_{\{win,t\}} + \\
 & \mu_{\{spr\}} \cdot D_{\{spr,t\}} + \mu_{\{sum\}} \cdot D_{\{sum,t\}} + \mu_{\{aut\}} \cdot D_{\{aut,t\}} + c_{\{HDD\}} \cdot HDD_{\{t\}} + \\
 & + c_{\{CDD\}} \cdot CDD_{\{t\}} + d_{\{Sat\}} D_{\{Sat,t\}} + d_{\{Sun\}} D_{\{Sun,t\}} + d_{\{Mon\}} D_{\{Mon,t\}} + \varepsilon_t,
 \end{aligned} \quad (49)$$

где

$p_{\{t\}}$ – спотовые цены на рынке на сутки вперед на электричество,

$\mu_{\{win\}}, \mu_{\{spr\}}, \mu_{\{sum\}}, \mu_{\{aut\}}$ – коэффициенты при дамми-переменных на сезон,

$D_{\{win,t\}}, D_{\{spr,t\}}, D_{\{sum,t\}}, D_{\{aut,t\}}$ – дамми на зимние месяцы, весенние, летние и осенние,

$c_{\{HDD\}}, c_{\{CDD\}}$ – коэффициенты перед градусо-сутками отопительных дней и охладительных дней,

$HDD_{\{t\}}$ – градусо-сутки отопительных дней,

$CDD_{\{t\}}$ – градусо-сутки охладительных дней,

$d_{\{Sat\}}, d_{\{Sun\}}, d_{\{Mon\}}$ – коэффициенты при дамми-переменных на день недели,

$D_{\{Sat,t\}}, D_{\{Sun,t\}}, D_{\{Mon,t\}}$ – дамми-переменные на день недели,

ε_t – ошибка.

Будет использоваться упрощенная Модель (2) с репараметризацией или Модель (3а), записанная уравнением (50):

$$\begin{aligned}
r_{\{t\}} = & p_t - (\beta \cdot t + \mu_{\{win\}} \cdot D_{\{win,t\}} + \\
& \mu_{\{spr\}} \cdot D_{\{spr,t\}} + \mu_{\{sum\}} \cdot D_{\{sum,t\}} + \mu_{\{aut\}} \cdot D_{\{aut,t\}} + \\
& + c_{\{HDD\}} \cdot HDD_{\{t\}} + c_{\{CDD\}} \cdot CDD_{\{t\}} + \\
& + d_{\{Sat\}} D_{\{Sat,t\}} + d_{\{Sun\}} D_{\{Sun,t\}} + d_{\{Mon\}} D_{\{Mon,t\}}) \\
r_{\{t\}} = & r_{\{t-1\}} + \varepsilon_t
\end{aligned} \tag{50}$$

где

$r_{\{t\}}$ – случайная составляющая,

$p_{\{t\}}$ – спотовые цены на рынке на сутки вперед на электричество,

$\mu_{\{win\}}, \mu_{\{spr\}}, \mu_{\{sum\}}, \mu_{\{aut\}}$ – коэффициенты при дамми-переменных на сезон,

$D_{\{win,t\}}, D_{\{spr,t\}}, D_{\{sum,t\}}, D_{\{aut,t\}}$ – дамми на зимние месяцы, весенние, летние и осенние,

$c_{\{HDD\}}, c_{\{CDD\}}$ – коэффициенты перед градусо-сутками отопительных дней и охладительных дней,

$HDD_{\{t\}}$ – градусо-сутки отопительных дней,

$CDD_{\{t\}}$ – градусо-сутки охладительных дней,

$d_{\{Sat\}}, d_{\{Sun\}}, d_{\{Mon\}}$ – коэффициенты при дамми-переменных на день недели,

$D_{\{Sat,t\}}, D_{\{Sun,t\}}, D_{\{Mon,t\}}$ – дамми-переменные на день недели.

Которая будет сравниваться с моделью стохастической волатильности (Модель (36)), записанной уравнением (51):

$$\begin{aligned}
r_{\{t\}} = & p_t - (\beta \cdot t + \mu_{\{win\}} \cdot D_{\{win,t\}} + \\
& \mu_{\{spr\}} \cdot D_{\{spr,t\}} + \mu_{\{sum\}} \cdot D_{\{sum,t\}} + \mu_{\{aut\}} \cdot D_{\{aut,t\}} + \\
& + c_{\{HDD\}} \cdot HDD_{\{t\}} + c_{\{CDD\}} \cdot CDD_{\{t\}} + \\
& + d_{\{Sat\}} D_{\{Sat,t\}} + d_{\{Sun\}} D_{\{Sun,t\}} + d_{\{Mon\}} D_{\{Mon,t\}}) \\
r_{\{t\}} = & r_{\{t-1\}} + \varepsilon_t^{\{(1)\}} \cdot \exp\left(\frac{h_t}{2}\right), \\
h_t = & \kappa + \phi(h_{\{t-1\}} - \kappa) + \varepsilon_{\{t-1\}}^{\{(2)\}} \sigma
\end{aligned} \tag{51}$$

где

$r_{\{t\}}$ – случайная составляющая,

$p_{\{t\}}$ – спотовые цены на рынке на сутки вперед на электричество,

$\mu_{\{win\}}, \mu_{\{spr\}}, \mu_{\{sum\}}, \mu_{\{aut\}}$ – коэффициенты при дамми-переменных на сезон,

$D_{\{win,t\}}$, $D_{\{spr,t\}}$, $D_{\{sum,t\}}$, $D_{\{aut,t\}}$ – дамми на зимние месяцы, весенние, летние и осенние,

$c_{\{HDD\}}$, $c_{\{CDD\}}$ – коэффициенты перед градусо-сутками отопительных дней и охладительных дней,

$HDD_{\{t\}}$ – градусо-сутки отопительных дней,

$CDD_{\{t\}}$ – градусо-сутки охладительных дней,

$d_{\{Sat\}}$, $d_{\{Sun\}}$, $d_{\{Mon\}}$ – коэффициенты при дамми-переменных на день недели,

$D_{\{Sat,t\}}$, $D_{\{Sun,t\}}$, $D_{\{Mon,t\}}$ – дамми-переменные на день недели,

$h_{\{t-1\}}$ – ненаблюдаемые переменные.

Априорные распределения параметров, представлены уравнением (52):

$$\begin{aligned} \beta, \mu_{\{win\}}, \mu_{\{spr\}}, \mu_{\{sum\}}, \mu_{\{aut\}}, d_{\{Sat\}}, d_{\{Sun\}}, d_{\{Mon\}}, c_{\{CDD\}}, c_{\{HDD\}} &\sim N(0, 10); \\ \sigma &\sim \text{Cauchy}(0, 5) \cdot 1_{\{(0, \infty)\}}, \phi \sim U(-1, 1); \kappa \sim \text{Cauchy}(0, 10) \end{aligned} \quad (52)$$

Такая репараметризация позволит не только решить проблему с состоятельностью оценок полученных по МСМС, как можно заметить во всех приведенных ниже таблицах для моделей с репараметризацией показатели R_{hat} близки к 1, а n_{eff} для некоторых коэффициентов даже выше чем число наблюдений в изначальной выборке, что говорит, что апостериорные распределения для соответствующих коэффициентов близки по распределению к гауссовским, но и позволить получить интерпретируемые коэффициенты в модели. Результаты оценивания Модели (3а) для первой ценовой зоны представлены на рисунке 48, а для второй ценовой зоны на рисунке 49. Результаты оценивания Модели (3б) для первой ценовой зоны представлены на рисунке 50, а для второй ценовой зоны на рисунке 51.

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|-------------|--------|-------|----------------|--------|---------|---------|---------|
| β | 0.9997 | 14572 | 0.0000 | 0.0011 | -0.0021 | 0.0000 | 0.0021 |
| d_{sat} | 0.9998 | 7581 | -0.0121 | 0.0027 | -0.0175 | -0.0122 | -0.0067 |
| d_{sun} | 0.9999 | 6570 | -0.0635 | 0.0031 | -0.0696 | -0.0635 | -0.0572 |
| d_{mon} | 0.9999 | 7543 | 0.0154 | 0.0027 | 0.0100 | 0.0154 | 0.0207 |
| d_{hol} | 1.0000 | 8093 | -0.0365 | 0.0058 | -0.0479 | -0.0365 | -0.0252 |
| μ_{win} | 1.0013 | 2969 | 7.2031 | 0.0536 | 7.0985 | 7.2029 | 7.3079 |
| μ_{spr} | 1.0012 | 2978 | 7.1682 | 0.0524 | 7.0654 | 7.1676 | 7.2707 |
| μ_{sum} | 1.0012 | 2964 | 7.2000 | 0.0491 | 7.1048 | 7.1997 | 7.2964 |
| μ_{aut} | 1.0013 | 2939 | 7.2076 | 0.0524 | 7.1050 | 7.2068 | 7.3109 |
| c_{cdd} | 0.9998 | 8138 | -0.0048 | 0.0051 | -0.0147 | -0.0048 | 0.0051 |
| c_{hdd} | 0.9997 | 14519 | -0.0037 | 0.0012 | -0.0061 | -0.0037 | -0.0013 |
| σ | 1.0003 | 7720 | 0.0484 | 0.0008 | 0.0469 | 0.0484 | 0.0499 |

Примечание – Источник: Собственные расчеты.

Рисунок 48 – Результаты оценивания Модели (3а) для 1 ЦЗ

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|-------------|--------|-------|---------------|--------|---------|--------|--------|
| β | 1.0007 | 12713 | 0.0002 | 0.0036 | -0.0069 | 0.0002 | 0.0072 |
| d_{sat} | 1.0003 | 7746 | 0.0275 | 0.0089 | 0.0098 | 0.0276 | 0.0450 |
| d_{sun} | 1.0001 | 6762 | 0.0079 | 0.0101 | -0.0117 | 0.0079 | 0.0283 |
| d_{mon} | 0.9999 | 7255 | 0.0253 | 0.0086 | 0.0083 | 0.0254 | 0.0424 |
| d_{hol} | 1.0001 | 10981 | 0.0310 | 0.0192 | -0.0063 | 0.0310 | 0.0685 |
| μ_{win} | 1.0000 | 3689 | 6.4020 | 0.1734 | 6.0677 | 6.4014 | 6.7397 |
| μ_{spr} | 1.0000 | 3736 | 6.3844 | 0.1696 | 6.0570 | 6.3852 | 6.7172 |
| μ_{sum} | 0.9999 | 3744 | 6.4220 | 0.1577 | 6.1159 | 6.4231 | 6.7262 |
| μ_{aut} | 0.9999 | 3729 | 6.3931 | 0.1688 | 6.0658 | 6.3941 | 6.7203 |
| c_{cdd} | 0.9998 | 11096 | 0.0098 | 0.0203 | -0.0299 | 0.0096 | 0.0501 |
| c_{hdd} | 0.9998 | 13933 | 0.0059 | 0.0025 | 0.0011 | 0.0059 | 0.0108 |
| σ | 0.9998 | 11536 | 0.1587 | 0.0026 | 0.1538 | 0.1586 | 0.1638 |

Примечание – Источник: Собственные расчеты.

Рисунок 49 – Результаты оценивания Модели (3а) для 2 ЦЗ

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|-------------|--------|-------|----------------|--------|---------|---------|---------|
| β | 1.0006 | 10590 | 0.0006 | 0.0008 | -0.0010 | 0.0006 | 0.0022 |
| d_{sat} | 1.0000 | 11134 | -0.0125 | 0.0021 | -0.0167 | -0.0125 | -0.0083 |
| d_{sun} | 0.9999 | 11197 | -0.0608 | 0.0024 | -0.0656 | -0.0608 | -0.0561 |
| d_{mon} | 1.0003 | 10953 | 0.0117 | 0.0020 | 0.0077 | 0.0117 | 0.0157 |
| d_{hol} | 1.0000 | 10821 | -0.0289 | 0.0062 | -0.0410 | -0.0288 | -0.0168 |
| μ_{win} | 0.9997 | 9824 | 7.2274 | 0.0443 | 7.1382 | 7.2276 | 7.3162 |
| μ_{spr} | 0.9997 | 9339 | 7.2034 | 0.0452 | 7.1118 | 7.2034 | 7.2953 |
| μ_{sum} | 0.9997 | 9134 | 7.2030 | 0.0399 | 7.1233 | 7.2026 | 7.2846 |
| μ_{aut} | 0.9998 | 9555 | 7.2158 | 0.0425 | 7.1291 | 7.2159 | 7.3015 |
| c_{cdd} | 0.9998 | 11007 | -0.0059 | 0.0043 | -0.0144 | -0.0059 | 0.0024 |
| c_{hdd} | 1.0000 | 9860 | -0.0039 | 0.0009 | -0.0057 | -0.0039 | -0.0022 |
| σ | 1.0449 | 133 | 0.4637 | 0.0541 | 0.3620 | 0.4615 | 0.5792 |
| κ | 1.0017 | 3558 | -6.4850 | 0.0812 | -6.6401 | -6.4863 | -6.3230 |
| ϕ | 1.0345 | 176 | 0.8483 | 0.0317 | 0.7784 | 0.8509 | 0.9037 |

Примечание – Источник: Собственные расчеты.

Рисунок 50 – Результаты оценивания Модели (3б) для 1 ЦЗ

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|-------------|--------|-------|----------------|--------|---------|---------|---------|
| β | 1.0001 | 10223 | 0.0004 | 0.0008 | -0.0013 | 0.0004 | 0.0020 |
| d_{sat} | 1.0001 | 11230 | -0.0126 | 0.0022 | -0.0169 | -0.0126 | -0.0084 |
| d_{sun} | 0.9999 | 11787 | -0.0604 | 0.0025 | -0.0653 | -0.0604 | -0.0554 |
| d_{mon} | 1.0007 | 12053 | 0.0118 | 0.0021 | 0.0078 | 0.0118 | 0.0159 |
| d_{hol} | 1.0001 | 9696 | -0.0309 | 0.0062 | -0.0434 | -0.0308 | -0.0189 |
| μ_{win} | 1.0000 | 8927 | 7.2088 | 0.0419 | 7.1266 | 7.2083 | 7.2931 |
| μ_{spr} | 1.0000 | 8767 | 7.1851 | 0.0432 | 7.0995 | 7.1853 | 7.2721 |
| μ_{sum} | 1.0002 | 8040 | 7.1871 | 0.0378 | 7.1113 | 7.1868 | 7.2644 |
| μ_{aut} | 1.0002 | 8437 | 7.1988 | 0.0403 | 7.1173 | 7.1984 | 7.2804 |
| c_{cdd} | 1.0002 | 10372 | 0.0021 | 0.0018 | 0.0001 | 0.0016 | 0.0067 |
| c_{hdd} | 0.9998 | 12218 | 0.0002 | 0.0002 | 0.0000 | 0.0002 | 0.0008 |
| σ | 1.0424 | 117 | 0.4435 | 0.0484 | 0.3497 | 0.4438 | 0.5463 |
| κ | 1.0008 | 8015 | -6.4663 | 0.0819 | -6.6263 | -6.4666 | -6.3083 |
| ϕ | 1.0286 | 172 | 0.8564 | 0.0286 | 0.7948 | 0.8581 | 0.9075 |

Примечание – Источник: Собственные расчеты.

Рисунок 51 – Результаты оценивания Модели (3б) для 2 ЦЗ

Сравнивая Модели (3а) и (3б) можно отметить, что для второй ЦЗ дисперсия для всех дамми переменных доверительные стали уже. Также для второй ЦЗ коэффициент при дамми переменной воскресенье стал значимым и положительным, коэффициент при HDD стал незначимым. Для первой ценовой зоны он значимый и отрицательный для обеих моделей.

Одним из преимуществ байесовского подхода является возможность использовать априорную информацию о коэффициентах. Сравним Модель (3б) с аналогичной спецификацией, представленной уравнением(53), но с более узкими априорными распределениями и только положительной частью для коэффициентов градусо-суток:

$$\begin{aligned}
 r_{\{t\}} = & p_t - (\beta \cdot t + \mu_{\{win\}} \cdot D_{\{win,t\}} + \\
 & \mu_{\{spr\}} \cdot D_{\{spr,t\}} + \mu_{\{sum\}} \cdot D_{\{sum,t\}} + \mu_{\{aut\}} \cdot D_{\{aut,t\}} + \\
 & + c_{\{HDD\}} \cdot HDD_{\{t\}} + c_{\{CDD\}} \cdot CDD_{\{t\}} + \\
 & + d_{\{Sat\}} D_{\{Sat,t\}} + d_{\{Sun\}} D_{\{Sun,t\}} + d_{\{Mon\}} D_{\{Mon,t\}}) \\
 r_{\{t\}} = & r_{\{t-1\}} + \epsilon_t^{\{(1)\}} \cdot \exp\left(\frac{h_t}{2}\right), \\
 h_t = & \kappa + \phi(h_{\{t-1\}} - \kappa) + \epsilon_{\{t-1\}}^{\{(2)\}} \sigma
 \end{aligned} \tag{53}$$

Априорные распределения параметров Модели (3б), представлены уравнением (54):

$$\begin{aligned}
 \beta, \mu_{\{win\}}, \mu_{\{spr\}}, \mu_{\{sum\}}, \mu_{\{aut\}}, d_{\{Sat\}}, d_{\{Sun\}}, d_{\{Mon\}}, c_{\{CDD\}}, c_{\{HDD\}} & \sim N(0, 10); \\
 \sigma & \sim Cauchy(0, 5) \cdot 1_{\{(0, \infty)\}}, \phi \sim U(-1, 1); \kappa \sim Cauchy(0, 10)
 \end{aligned} \tag{54}$$

Априорные распределения параметров Модели (3в), представлены уравнением (55):

$$\begin{aligned}
 \beta, d_{\{Sat\}}, d_{\{Sun\}}, d_{\{Mon\}} & \sim N(0, 1); \\
 \mu_{\{win\}}, \mu_{\{spr\}}, \mu_{\{sum\}}, \mu_{\{aut\}}, & \sim N(7, 1); \\
 c_{\{CDD\}}, c_{\{HDD\}} & \sim N(0, 1) \cdot 1_{\{(0, \infty)\}} \\
 \sigma & \sim Cauchy(0, 5) \cdot 1_{\{(0, \infty)\}}, \phi \sim U(-1, 1); \kappa \sim Cauchy(0, 10)
 \end{aligned} \tag{55}$$

При использовании априорной информации о знаках коэффициентов при градусо-сутках можно отметить, что знаки и значимость остальных коэффициентов не изменилась. Однако мы не можем проверить значимость коэффициентов при градусо-сутках. Поскольку по построению значения левее 0 не рассматриваются для этого коэффициента. На рисунках 52 и 53 представлены результаты оценивания модели 3в для первой ценовой зоны и второй ценовой зоны, соответственно.

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|-------------|--------|-------|----------------|--------|---------|---------|---------|
| β | 1.0001 | 10223 | 0.0004 | 0.0008 | -0.0013 | 0.0004 | 0.0020 |
| d_{sat} | 1.0001 | 11230 | -0.0126 | 0.0022 | -0.0169 | -0.0126 | -0.0084 |
| d_{sun} | 0.9999 | 11787 | -0.0604 | 0.0025 | -0.0653 | -0.0604 | -0.0554 |
| d_{mon} | 1.0007 | 12053 | 0.0118 | 0.0021 | 0.0078 | 0.0118 | 0.0159 |
| d_{hol} | 1.0001 | 9696 | -0.0309 | 0.0062 | -0.0434 | -0.0308 | -0.0189 |
| μ_{win} | 1.0000 | 8927 | 7.2088 | 0.0419 | 7.1266 | 7.2083 | 7.2931 |
| μ_{spr} | 1.0000 | 8767 | 7.1851 | 0.0432 | 7.0995 | 7.1853 | 7.2721 |
| μ_{sum} | 1.0002 | 8040 | 7.1871 | 0.0378 | 7.1113 | 7.1868 | 7.2644 |
| μ_{aut} | 1.0002 | 8437 | 7.1988 | 0.0403 | 7.1173 | 7.1984 | 7.2804 |
| c_{cdd} | 1.0002 | 10372 | <i>0.0021</i> | 0.0018 | 0.0001 | 0.0016 | 0.0067 |
| c_{hdd} | 0.9998 | 12218 | <i>0.0002</i> | 0.0002 | 0.0000 | 0.0002 | 0.0008 |
| σ | 1.0424 | 117 | 0.4435 | 0.0484 | 0.3497 | 0.4438 | 0.5463 |
| κ | 1.0008 | 8015 | -6.4663 | 0.0819 | -6.6263 | -6.4666 | -6.3083 |
| ϕ | 1.0286 | 172 | 0.8564 | 0.0286 | 0.7948 | 0.8581 | 0.9075 |

Примечание – Источник: Собственные расчеты.

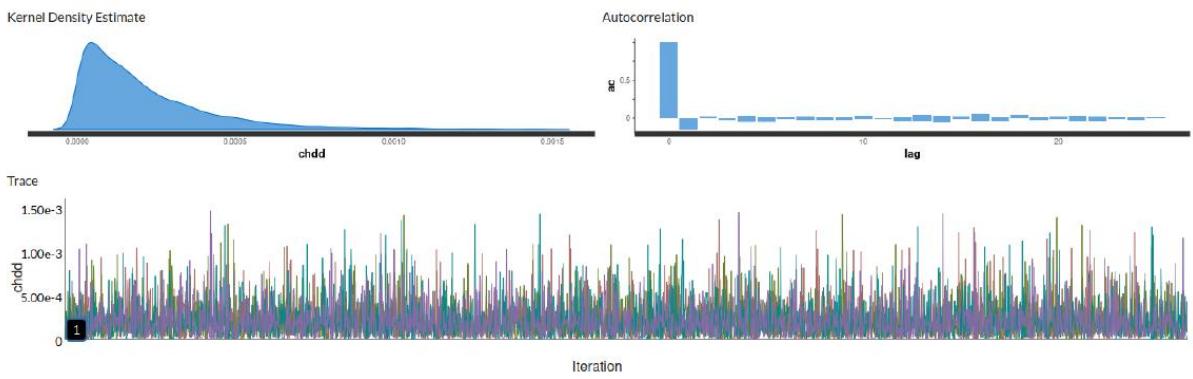
Рисунок 52 – Результаты оценивания Модели (3в) для 1 ЦЗ

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|-------------|--------|-------|----------------|--------|---------|---------|---------|
| β | 1.0000 | 11066 | -0.0009 | 0.0012 | -0.0032 | -0.0009 | 0.0014 |
| d_{sat} | 1.0003 | 10625 | 0.0141 | 0.0031 | 0.0079 | 0.0141 | 0.0201 |
| d_{sun} | 1.0001 | 9376 | -0.0125 | 0.0037 | -0.0200 | -0.0125 | -0.0053 |
| d_{mon} | 0.9997 | 11960 | 0.0061 | 0.0030 | 0.0003 | 0.0061 | 0.0120 |
| d_{hol} | 1.0001 | 8883 | 0.0063 | 0.0072 | -0.0082 | 0.0063 | 0.0199 |
| μ_{win} | 1.0006 | 3384 | 6.4479 | 0.1110 | 6.2516 | 6.4402 | 6.6988 |
| μ_{spr} | 1.0005 | 3427 | 6.4545 | 0.1120 | 6.2557 | 6.4468 | 6.7059 |
| μ_{sum} | 1.0006 | 3144 | 6.4553 | 0.1055 | 6.2728 | 6.4450 | 6.6954 |
| μ_{aut} | 1.0006 | 3327 | 6.4397 | 0.1095 | 6.2465 | 6.4313 | 6.6838 |
| c_{cdd} | 0.9998 | 9571 | <i>0.0072</i> | 0.0056 | 0.0003 | 0.0059 | 0.0209 |
| c_{hdd} | 1.0003 | 6242 | <i>0.0014</i> | 0.0008 | 0.0001 | 0.0014 | 0.0030 |
| σ | 1.0060 | 351 | 0.7987 | 0.0558 | 0.6931 | 0.7964 | 0.9112 |
| κ | 0.9997 | 10591 | -5.3833 | 0.1313 | -5.6391 | -5.3846 | -5.1168 |
| ϕ | 1.0040 | 514 | 0.8501 | 0.0228 | 0.8033 | 0.8511 | 0.8916 |

Примечание – Источник: Собственные расчеты.

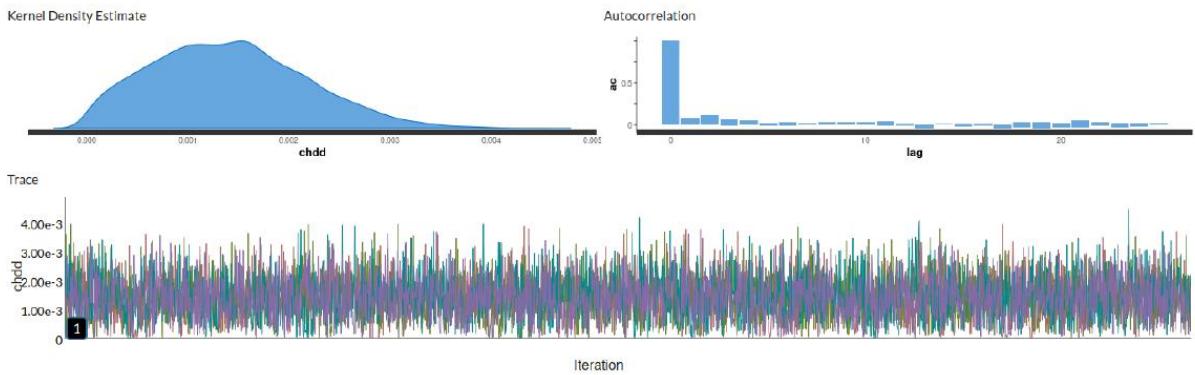
Рисунок 53 - Результаты оценивания Модели (3в) для 2 ЦЗ

В качестве аналога t-теста для таких коэффициентов можно применить визуальный анализ апостериорного распределения. Можно заметить, что для 1 ЦЗ коэффициент все еще достаточно близок к 0, а само распределение асимметричное и с большим коэффициентом эксцесса. В то время как апостериорное распределение для коэффициента во 2 ЦЗ стало более похожим на нормальное, более подробно смотри на рисунке 54 для первой ценовой зоны, на рисунке 55 для второй ценовой зоны.



Примечание – Источник: Собственные расчеты.

Рисунок 54 – Апостериорная плотность коэффициента при HDD для 1ЦЗ



Примечание – Источник: Собственные расчеты.

Рисунок 55 – Апостериорная плотность коэффициента при HDD для 2ЦЗ

Из результатов по приведенным выше моделям можно сказать, что значимость HDD по данной спецификации установить сложно. Теоретически эта зависимость должна быть ярко выражена, однако для того, чтобы можно было ее отследить, способ включения регрессора температуры в модель должен быть изменен.

Касательно использования информативных априорных распределений решение не было принято, поскольку теоретически это преимущество байесовского подхода может быть использовано при дальнейшем анализе, однако на данном этапе будет лишь мешать интерпретации и скрывать проблемные зоны моделей.

На последнем этапе анализа рассмотрим на влияние включения дополнительных регрессоров в модель. Добавим ИПП, цены на уголь и газ в Модель (3г), уравнение (56):

$$\begin{aligned}
 r_{\{t\}} = & p_t - (\beta \cdot t + \mu_{\{win\}} \cdot D_{\{win,t\}} + \\
 & \mu_{\{spr\}} \cdot D_{\{spr,t\}} + \mu_{\{sum\}} \cdot D_{\{sum,t\}} + \mu_{\{aut\}} \cdot D_{\{aut,t\}} + \\
 & + c_{\{HDD\}} \cdot HDD_{\{t\}} + c_{\{CDD\}} \cdot CDD_{\{t\}} + \\
 & + c_{\{IPP\}} \cdot \ln IPP_{\{t\}} + c_{\{gas\}} \cdot \ln Pgas_{\{t\}} + c_{\{coal\}} \cdot \ln Pcoal_{\{t\}} + \\
 & + d_{\{sat\}} D_{\{sat,t\}} + d_{\{sun\}} D_{\{sun,t\}} + d_{\{mon\}} D_{\{mon,t\}}) \\
 r_{\{t\}} = & r_{\{t-1\}} + \epsilon_t^{\{(1)\}} \cdot \exp\left(\frac{h_t}{2}\right), \\
 h_t = & \kappa + \phi(h_{\{t-1\}} - \kappa) + \epsilon_{\{t-1\}}^{\{(2)\}} \sigma
 \end{aligned} \tag{56}$$

Априорные распределения параметров представлены уравнением (57):

$$\beta, \mu_{\{win\}}, \mu_{\{spr\}}, \mu_{\{sum\}}, \mu_{\{aut\}}, d_{\{sat\}}, d_{\{sun\}}, d_{\{mon\}}, c_{\{CDD\}}, c_{\{HDD\}} \sim N(0, 10); \tag{57}$$

$$\sigma \sim \text{Cauchy}(0, 5) \cdot 1_{\{(0, \infty)\}}, \phi \sim U(-1, 1); \kappa \sim \text{Cauchy}(0, 10)$$

где

$p_{\{t\}} = \ln(P_{\{t\}})$ - логарифм средней за день цены в ЦЗ,

β - коэффициент при линейном тренде,

$c_{\{HDD\}}, c_{\{CDD\}}$ - коэффициент при средних за день градусо-сутках отопительного $HDD_{\{t\}}$ и охладительного периода $CDD_{\{t\}}$,

IPP - сглаженная трендовая составляющая ИПП,

$p_{\{gas\}}, p_{\{coal\}}$ - среднемесячная цена на газ и уголь (по РФ),

$d_{\{Sat\}}, d_{\{Sun\}}, d_{\{Mon\}}$ - коэффициент при дамми переменных на дни недели (суббота, воскресенье, понедельник),

$\mu_{\{win\}}, \mu_{\{spr\}}, \mu_{\{sum\}}, \mu_{\{aut\}}$ - коэффициент при дамми переменных на время года,

$\varepsilon_t, \varepsilon_t^{\{(1)\}}, \varepsilon_t^{\{(2)\}} \sim N(0, \sigma) i.i.d.$

Оценки полученные по данной модели имеют очень маленький n_{eff} , что указывает на проблемы в спецификации и использовании регрессоров. Несмотря на то, что теоретически модель стохастической волатильности лучше подходит для моделирования цен на электричество, на данном этапе необходимо скорректировать спецификацию модели и используемые в ней регрессоры перед тем, как добавлять в модель стохастические компоненты.

Итак, на данном этапе необходимо решить следующие проблемы. Во-первых, необходимо скорректировать, добавляемые в модель регрессоры. Регрессор градусо-сутки отопительного периода имеет ярко выраженную сезонность. На данном этапе сезонность учитывается с помощью дамми переменных, но линейная динамика в течение года не отражается в спецификации выбранных моделей. Для предварительного анализа этого было достаточно, но поскольку финальной целью данной работы является получение точных оценок коэффициентов тренд-сезонного разложения необходимо это будет учесть. Результаты оценивания Модели (3г) для первой ценовой зоны представлены на рисунке 56.

| Parameter | Rhat | n_eff | mean | sd | 2.5% | 50% | 97.5% |
|-------------|--------|-------|----------------|--------|---------|---------|---------|
| β | 1.0050 | 496 | 0.0006 | 0.0008 | -0.0011 | 0.0006 | 0.0022 |
| μ_{win} | 1.1634 | 44 | 3.6130 | 2.5804 | -1.1713 | 3.7836 | 8.3705 |
| μ_{spr} | 1.1634 | 44 | 3.5878 | 2.5807 | -1.2084 | 3.7564 | 8.3406 |
| μ_{sum} | 1.1634 | 44 | 3.5868 | 2.5804 | -1.2048 | 3.7556 | 8.3359 |
| μ_{aut} | 1.1635 | 44 | 3.6003 | 2.5800 | -1.1827 | 3.7720 | 8.3541 |
| d_{sat} | 1.1236 | 46 | -0.0122 | 0.0023 | -0.0164 | -0.0123 | -0.0078 |
| d_{sun} | 1.0837 | 48 | -0.0610 | 0.0025 | -0.0660 | -0.0610 | -0.0562 |
| d_{mon} | 1.0957 | 46 | 0.0115 | 0.0020 | 0.0076 | 0.0115 | 0.0153 |
| d_{hol} | 1.0262 | 132 | -0.0277 | 0.0057 | -0.0393 | -0.0276 | -0.0172 |
| c_{cdd} | 1.0398 | 99 | -0.0056 | 0.0045 | -0.0143 | -0.0055 | 0.0030 |
| c_{hdd} | 1.0155 | 236 | -0.0040 | 0.0009 | -0.0057 | -0.0040 | -0.0021 |
| c_{ipp} | 1.0334 | 108 | 0.0635 | 1.2755 | -2.6231 | 0.0695 | 2.7945 |
| c_{gas} | 1.0366 | 123 | 0.1946 | 0.2783 | -0.3194 | 0.1768 | 0.7715 |
| c_{coal} | 1.1202 | 48 | 0.3387 | 0.4158 | -0.3709 | 0.2912 | 1.2543 |
| σ | 1.4198 | 9 | 0.4860 | 0.0475 | 0.3921 | 0.4886 | 0.5698 |
| κ | 1.0406 | 81 | -6.4817 | 0.0804 | -6.6352 | -6.4837 | -6.3238 |
| ϕ | 1.1277 | 22 | 0.8401 | 0.0268 | 0.7901 | 0.8387 | 0.8916 |

Примечание – Источник: Собственные расчеты.

Рисунок 56 – Результаты оценивания Модели (3г) для 1 ЦЗ

При использовании рядов Фурье для моделирования годовой сезонности, как в модели Prophet вероятно, что этот регрессор будет сильно коррелирован с регрессором градусо-суток. По этой причине коэффициент при градусо-сутках может получаться значимым и отрицательным, поскольку оценивается по большей части именно влияние годовой компоненты сезонности. В результате проведенного анализа было решено использовать в качестве регрессора отклонение показателя градусо-суток от некоторого плавающего среднего значения. Эти отклонения должны учитываться в стохастической компоненте. Трендовая компонента градусо-суток должна быть трансформирована в дамми переменные начала – середины – конца отопительного периода и всего охладительного периода. Эти дамми можно учесть при моделировании сезонной компоненты.

В рамках приведенных в основном тексте работы гипотез индекс деловой активности (который в наших моделях представлен ИПП) должен отвечать за оценку

трендовой компоненты, однако доверительные интервалы для этого регрессора значительно шире в сравнении с другими коэффициентами. Имеет смысл вместо ИПП использовать кусочно-линейную функцию, как в Prophet-модели: $g(t) = (k + \overrightarrow{a}(t)\delta) + (m + \overrightarrow{a}(t)^{\{T\}}\gamma)$.

Проблемы с использованием регрессоров ИПП и цен могут быть вызваны теми же эффектами, что и для градусо-суток: вероятно, что за счет интерполяции и в целом плавного изменения этих показателей, эти регрессоры должны быть сильно коррелированы с линейным трендом.

Проблема с IPP заключается в том, что используется совокупный показатель по РФ по всем отраслям. Данных с разбивкой по ЦЗ (по регионам) нет. Однако, спрос на электричество для конкретного региона может быть лучше отражен через IPP по отраслям, наиболее развитым в конкретной области. Аналогично, данные по ценам на ресурсы берутся средние по РФ, поскольку данные, усредненные по ЦЗ вызывают сомнения, так как в дезагрегированных данных по регионам присутствуют пропуски и длительные периоды, когда цены оставались неизменными. Такие регрессоры можно добавить только в качестве регрессоров для гиперпараметров регрессии или для определения точек поворота в кусочно-линейной функции. Также можно будет попробовать использовать финансовые котировки для цен на нефть.

Как было отмечено ранее цены на электроэнергию на сутки вперед могут оперативно сигнализировать о состоянии рынка электроэнергии, поскольку эти цены используются в качестве справочных цен для решений, принимаемых в торговле энергией, участниками рынка. Краткосрочные стратегии покупки и продажи электроэнергии обычно основаны на спотовых ценах на день вперед. Таким образом, эффективные методы анализа, моделирования и прогнозирования цен на электроэнергию на сутки вперед крайне важны для правильного регулирования рынком электроэнергии.

Тарифы на электроэнергию отражают реальные затраты на электроэнергию. Если это действительно так, то в течение года у регуляторов рынка должны возникать экономически эффективные стимулы для введения мер поддержки энергосистемы, которые могут значительно снизят стоимость энергосистемы.

Построении многоуровневая структурная модели цен на электричество может применяться для разработки конкретных направлений, мер и механизмов государственной политики в сфере регулирования рынка электроэнергии:

— во-первых, в рамках предварительного анализа цен в 1 и 2 ЦЗ было отмечено, что в 1 ЦЗ наблюдается рост трендовой компоненты цен на электричество,

потенциально превышающим темпы роста инфляции, в то время как во 2 ЦЗ цены на электричество отличаются гораздо менее быстрыми темпами роста,

— во-вторых, при решении задачи выявления причин различия в динамике цен в ценовых зонах в первую очередь необходимо выявить статистические различия в этих рядах, что было подробно разобрано в пункте 3.2; помимо значительных различий в динамике трендовой составляющей рядов, внутригодовая динамика, зависящая от изменений температуры также сильно различается для этих рядов; различаются также и недельная динамика рядов; для детального изучения причин этих различий необходимо построить более сложные модели, учитывающие особенности российской электроэнергетики и каждой из ценовых зон в частности,

— в-третьих, при построении сложной структурной модели можно будет отследить влияние структуры генерирующих мощностей на волатильность цен,

— в-четвертых, уже в рамках Prophet модели можно оценить влияние резких изменений экономических факторов на цены.

Надо понимать, что в реформированной российской энергетики сильную роль играют как рыночные, так и административные механизмы. Поэтому при более детальном анализе цен на электричество необходимо тщательно изучить все нововведения, вводимые НП Советом рынка, корректирующие регламент. Поскольку, если, например, формально «рыночная» система никак не реагирует на резкие изменения экономических показателей, но сразу изменяется в результате введения новых норм и правил, можно оценить, насколько зарегулированной является эта система. Или, например, если цены слабо отзываются на изменения факторов спроса, то можно предположить, что причиной этому может стать тот факт, что совокупный спрос формируется только формально, поскольку подавляющее большинство фирм является ценополучателями, то есть никак не влияют на равновесную цену.

Например, одним из нововведений СО ЕЭС является отбор агрегаторов спроса, снижающих потребление электроэнергии своих клиентов в пиковые часы. Агрегаторы управления спросом – специализированные организации, координирующие способность группы розничных потребителей управлять своим электропотреблением, конвертирующие ее в услуги по управлению спросом на электрическую энергию и транслирующие потребителям часть выручки, полученной от реализации этой услуги. Агрегаторы по команде СО снижают нагрузку клиентов на два или четыре часа в периоды пиковых нагрузок на энергосистему, что в перспективе должно приводить к снижению цен на рынке на сутки вперед (РСВ). Однако, как показывает практика новый инструмент имеет незначительное влияние на стоимость электроэнергии.

В свете последних событий, связанных с обвалом цен на нефтяном и угольном рынках, со снижением совокупного спроса на электричество, регуляторам необходимо быстро реагировать на изменения рынка. Цены на РСВ один из немногих доступных оперативных показателей, транслирующих текущее состояние рынка. Последние три года энергопотребление в ЕЭС России стабильно росло более чем на 1%. Предположительно в результате карантинных мер, спрос на электроэнергию снизился более, чем на 10%. Естественно, после снятия ограничений спрос частично восстановится, но неизвестно, с какой скоростью. Модели применяемы для анализа цен на РСВ могут быть также применены к оперативному анализу совокупного ежедневного спроса и предложения на рынках на сутки вперед 1 и 2 ЦЗ.

Все долгосрочные инвестиционные проекты, такие как ДПМ, ДПМ-ВИЭ, инвестпрограммы сетевых компаний и т.д., планировались исходя из устойчивого роста спроса на электроэнергию и экономического роста в 1-3% в год.

Согласно схеме и программе развития электроэнергетики, утвержденной Минэнерго в 2019 году, в 2020 году спрос на электроэнергию должен составить 1092,2 млрд кВтч. Однако прогноз роста на 3,1% слишком оптимистичен. До развития пандемии можно было ожидать рост спроса в пределах 0,5%, что частично обусловлено ростом доли энергосберегающего потребления.

В новых условиях сохранение инвестпрограмм в первоначальном виде может привести к чрезмерной нагрузке на потребителей электроэнергии и росту цен выше инфляции.

В рамках предыдущего анализа уже было отмечено, что есть вероятность, что в 1 ЦЗ уже, трендовая составляющая цен на электроэнергию растет с темпами выше инфляции. Учитывая, что цены на РСВ – одна из пяти составляющих конечных цен на электричество, необходимо скорректировать текущие программы с целью сдерживания цен на электроэнергию ниже инфляции.

В частности, одной из инвестпрограмм, действие которой завершается в 2020 году является ДПМ-ВИЭ. Сейчас идут обсуждения принятия ДПМ-ВИЭ-2. Однако отметим, что при нынешнем снижении цен на полезные ископаемые, зеленая энергетика становится менее привлекательной альтернативой в смысле затрат.

В целом преимущества использования ДПМ-ВИЭ неочевидны и часто критикуются, поскольку их введение потенциально увеличило объемы платежей потребителей почти вдвое выше относительно заявленного на старте. Результаты проведенных отборов проектов ВИЭ демонстрируют цены, которые кратно превышают среднемировые показатели. Для поддержки развития ВИЭ существуют альтернативные

меры, например, налоговые льготы, субсидирование процентных ставок, использование «зеленых» сертификатов.

Также использование ДПМ приводит к тому, что участники конкурсов по отбору проектов ВИЭ закупают и устанавливают по ДПМ импортное оборудование, заплатив штраф за отказ от локализации, что негативно влияет на развитие отечественного рынка ВИЭ-генераторов. Учитывая существование «Стратегии развития энергетического машиностроения до 2030 года». Необходимо переопределить наиболее востребованные механизмы поддержки, которые позволяют обеспечить загрузку российских предприятий и развитие отечественной научно-технической базы.

На данный момент в рамках российской энергосистемы еще не достигнуты оптимальное сочетание использования различных источников генерации и способов доставки и потребления электричества и тепла для каждого региона. Поэтому необходимо точное понимание особенностей не только каждой ЦЗ, но и каждого региона, входящих в их состав.

ЗАКЛЮЧЕНИЕ

Практически всеми предприятиями в современной экономике потребляется энергия, как в первичном виде, так и в переработанном: в виде топлива, тепла и электричества. Целевые ориентиры развития энергетической системы определяются государством, в частности, энергетической стратегией, которая включает в себя перспективную структуру мощностей, политику в области обеспечения топливом, развития конкуренции, повышения привлекательности отрасли для инвесторов. Кроме того, одним из важных элементов государственного управления энергетической системы является ценовая политика.

На региональном уровне цены на электроэнергию могут выступать, как ключевые факторы оценки качества условий жизни населения и развития предпринимательской среды, поэтому в рамках данной работы нами были разработаны модели, описывающие экономику электроэнергетического рынка России с акцентом на возможность получение прогнозных значений цен в региональном разрезе. Оценка моделей осуществлялась в том числе с использованием данных, собранных в НИР 2019 года «Формирование банка данных и разработка аналитической оболочки для объектов энергетической системы России».

На рынке электрической энергии большая часть объемов реализуется по свободным ценам, сформированным на рынке на сутки вперед (далее – PCB). Формирование цен на PCB осуществляется на основании заявок участников рынка по принципу маржинального ценообразования. Для каждого узла расчетной модели под воздействием спроса и предложения определяется цена на электрическую энергию, которая распространяется на всех участников. Ценовые заявки формируются поставщиками исходя из топливных затрат, так на динамику цен на рынке PCB в значительной степени влияет изменение цен на топливо.

В рамках моделирования рынка электроэнергии России была разработана модель ценообразования на электроэнергию на рынке на сутки вперед. Цены на электричество отличаются значительной волатильностью, обладают двухуровневой структурой сезонности (не считая дневной), отличаются наличием частых выбросов — скачков. При построении прогнозов цен на электричество необходимо верно оценивать доверительные интервалы прогноза, так как для участников рынка это позволяет верно оценить риски при принятии стратегических решений.

Существует множество различных подходов к моделированию цен на электричество. Традиционные методы временных рядов не могут точно моделировать нелинейную структуру цены на электроэнергию, также без учета скачкообразной

компоненты доверительные интервалы могут быть заужены. В результате при анализе временных рядов цен на электроэнергию была выбрана гибкая модель, позволяющая учесть статистические особенности этих временных рядов: моделирование цены на оптовом рынке электроэнергии осуществлялось как диффузионно-скачкообразный процесс с учётом фундаментальных факторов со стороны спроса и предложения (температуры, уровня деловой активности и цен на используемое сырье при выработки электроэнергии).

Результатом моделирования является тренд-сезонное разложение цены на электричество с интерпретируемыми оценками составляющих компонент. Помимо этого с помощью построенной модели можно определить причины различия в динамике цен в Европейской и Сибирской ценовых зонах, оценить темпы роста цен на электроэнергию, что немаловажно, поскольку, во-первых, расходы на электроэнергию несут подавляющее большинство производителей, во-вторых, сбытовые компании, участвующие на РСВ, в результате конкурентного отбора заявок, приобретают электричество, которое впоследствии будет отпускаться населению. Также оценка случайной компоненты внутри модели позволяет определить влияние структуры генерирующих мощностей на волатильность цен и оценить воздействие резких изменений экономических факторов на цены на электричество, а скачкообразная компонента, позволяет учесть неоднородность волатильности ошибок. Данная модель также может быть применена на равновесных поузловых ценах, что позволит использовать полученные оценки трендовой и сезонной составляющих, для выявления особенностей ценообразования в разрезе конкретных регионов России и разработки мер государственной политики в сфере регулирования рынка электроэнергии.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Liu A.L. Repeated games in electricity spot and forward markets-An equilibrium modeling and computational framework // 2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton). 2010. pp. 66-71.
2. Krause T. Evaluating congestion management schemes in liberalized electricity markets using an agent-based simulator // 2006 IEEE Power Engineering Society General Meeting. 2006. pp. 8--pp.
3. Cartea A. Pricing in electricity markets: a mean reverting jump diffusion model with seasonality // Applied Mathematical Finance. 2005. pp. 313--335.
4. Lucia J. Electricity prices and power derivatives: Evidence from the nordic power exchange // Review of derivatives research. 2002. pp. 5--50.
5. Kuznetsova E. An integrated framework of agent-based modelling and robust optimization for microgrid energy management // Applied Energy. 2014. pp. 70-88.
6. Deb K., Agrawal S., Pratap A., and Meyari T., «A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II» Parallel Problem Solving from Nature PPSN VI, 2000. pp. 849-858.
7. Krause T. A comparison of Nash equilibria analysis and agent-based modelling for power markets // International Journal of Electrical Power & Energy Systems. 2006. pp. 599-607.
8. Andersson M. Cost-effective energy system measures studied by dynamic modelling // 2nd International Conference on Advances in Power System Control, Operation and Management. 1993. pp. 448-455.
9. Stevenson M. Filtering and forecasting spot electricity prices in the Australian electricity market. // University of Technology, Sydney. 2002.
10. Escribano A. Modelling electricity prices: International evidence // Oxford bulletin of economics and statistics. 2011. pp. 622-650.
11. Taylor J.W. Triple seasonal methods for short-term electricity demand forecasting. // European Journal of Operational Research. 2010. pp. 139-152.
12. Kostrzewski M.A. Probabilistic electricity price forecasting with Bayesian stochastic volatility models // Energy Economics. 2019. pp. 610-620.
13. Huisman R.A. Regime jumps in electricity prices // Energy economics. 2003. pp. 425-434.
14. Mohandes M. Support vector machines for short-term electrical load forecasting // International Journal of Energy Research. 2002. pp. 335-345.

15. Keles D. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks // *Applied energy*. 2016. pp. 218-230.
16. Huisman R. Regime jumps in electricity prices // *Energy economics*. 2003. pp. 425-434.
17. Harvey A. Estimation procedures for structural time series models // *Journal of Forecasting*. 1990. pp. 89-108.
18. Hyndman A. STR: A Seasonal-TrendDecomposition Procedure Basedon Regression, arXiv preprint arXiv:2009.05894, 2015.
19. BP. Statistical Review of World Energy 2019 // BP. 2019. URL: <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html> (дата обращения: 07.06.2020).
20. Федеральная антимонопольная служба Российской Федерации, «Доклад о состоянии конкуренции,» Федеральная антимонопольная служба Российской Федерации, Москва, Доклад 2017.
21. Аналитический центр при Правительстве Российской Федерации, «ТЭК России - 2017» Статистический сборник 2018.
22. АТС. Оптовый рынок [Электронный ресурс] [2020]. URL: <https://www.atsenergo.ru/optovyy-rynek>
23. Дробыш И.И., «Статистические модели прогнозирования цен на оптовых рынках электроэнергии: российский и зарубежный опыт» Вестник Южно-Уральского государственного университета., 2017.
24. Соловьева И. «Прогнозирование электропотребления с учетом факторов технологической и рыночной среды.,» Научный диалог 7, 2013.
25. Совет Рынка Н. Оптовый рынок электрической энергии и мощности [Электронный ресурс] [2020]. URL: <https://www.np-sr.ru/ru/market/wholesale/index.htm>
26. Минэнерго О.с.М.э.Р.Ф. Об утверждении Генеральной схемы размещения объектов электроэнергетики на период до 2035 года [Электронный ресурс] [2017]. URL: <https://minenergo.gov.ru/node/8504>
27. Седых К. «Сдерживающие факторы развития электроэнергетической отрасли россии,» Динамика взаимоотношений различных областей науки в современных условиях, 2019. pp. 82-84.
28. Трачук А. «Перекрестное субсидирование в электроэнергетике: проблемы и пути решения.,» 2017. pp. 5-102.
29. Лясковская Е.А. «К вопросу реформирования электроэнергетики РФ. ,» Экономика, управление и инвестиции (электронный научный журнал), 2014. pp. 3-24.

30. Рыбина И.А. «Основные аспекты стратегии развития электроэнергетической отрасли до 2035 года.» *Взаимодействие науки и бизнеса*, 2016. pp. 78-84.
31. Совет Рынка Н. // ДПМ (договор о предоставлении мощности): [сайт]. [2020]. URL: <https://www.np-sr.ru/ru/glossary/item/dpm-dogovor-o-predostavlenii-moshchnosti>
32. Ernst & Young E., «Обзор электро-энергетической отрасли России,» 2018.
33. Taylor S.J. Forecasting at scale // *The American Statistician*. 2018. pp. 37-45.
34. Yildirim I., «Bayesian Inference: Metropolis-Hastings Sampling» Department of Brain and Cognitive Sciences, 2012. pp. 1-6.
35. Betancourt M., «A Conceptual Introduction to Hamiltonian Monte Carlo» Preprint: arXiv:1701.02434v2 [stat.ME], 2018. pp. 1-60.
36. Monnahan C.C., Kristensen K., «No-U-turn sampling for fast Bayesian inference in ADMB and TMB: Introducing theadnuts and tmbstan R packages» *P L o S One*, 2018. pp. 1-10.
37. Adrian E. Raftery S.M.L., «The Number of Iterations, Convergence Diagnostics and Generic Metropolis Algorithms» University of Washington Seattle, 1970. pp. 1-15.
38. Dootika Vats C.K., «Revisiting the Gelman-Rubin Diagnostic» preprint: arXiv:1812.09384, 2018. pp. 1-28.
39. Begg C.B., «Systematic reviews of diagnostic accuracy studies require study by study examination: first for heterogeneity, and then for sources of heterogeneity» *Journal of clinical epidemiology* , 2005.
40. Anirban Bhattacharya D. «Dirichlet–Laplace Priors for Optimal Shrinkage» *Journal of the American Statistical Association*, 2015. pp. 1479-1490.
41. Roy V., «Convergence diagnostics for Markov chain Monte Carlo» *Annual Reviews*. Department of statistics, Iowa state University, 2019. pp. 1-25.
42. Burke P.J., Liao H. Is the price elasticity of demand for coal in China increasing? // *China Economic Review*. December 2015. Vol. 36. pp. 309-322.