

Прогнозирование рыночной конъюнктуры на оптовом рынке электроэнергии

Введение

С либерализацией электроэнергетики модель торговли электрической энергией в России прошла трансформацию от государственного регулирования к конкурентному рынку и рыночному ценообразованию. В рамках новой рыночной парадигмы точное прогнозирование цены и спроса на электроэнергию, даже «на сутки вперед», является важным критерием, определяющим конкурентные преимущества производителей или потребителей электроэнергии. К настоящему моменту данная тематика является широко проработанной в научной и прикладной литературе. Предложено множество различных регрессионных моделей [8], моделей на основе искусственного интеллекта [7], а также теории игр [4] и фундаментального анализа [2]. Не смотря на высокую проработанность вопроса, остается ряд открытых задач, решение которых направленно, главным образом, на повышение качества моделирования рыночной конъюнктуры, включая важное направление - выбор оптимального набора факторов для прогнозной модели, оказывающих влияние на формирование цен на электрическую энергию. Целью настоящего исследования является всесторонний анализ факторов ценообразования на электроэнергию для определения оптимального с точки зрения точности прогноза набора экзогенных переменных для прогнозных моделей спроса, равновесной цены на рынке на сутки вперед и балансирующем рынке электроэнергии России.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-36-00389.

1 Подход к прогнозированию

В работе представлен прогноз трех основных параметров, в значительной степени характеризующих функционирование оптового рынка электроэнергии и его динамику: объема спроса, спотовой цены рынка на сутки в перед (PCB) и цены балансирующего рынка (БР).

Прогнозирование сводится к трем основным этапам:

- определение оптимального набора факторов (регрессоров), максимально точно описывающих историческую динамику рассматриваемых индикаторов;
- оптимизация конфигурации модели по каждому из трех прогнозных параметров;
- почасовое прогнозирование параметров на сутки, неделю и месяц вперед и оценка точности прогноза.

Выбор оптимального набора факторов производится с применением подхода на основе последовательного набора факторов [6]. Настоящий подход широко используется в машинном обучении и анализе данных с целью включения в модель только значащих факторов, а также оптимизации времени вычисления. Данный подход заключается в последовательном включении в модель объясняющих факторов, при этом на каждом шаге отбирается тот параметр, который вносит наибольший вклад в точность прогнозирования. Если включение дополнительного фактора не приводит к снижению ошибки прогнозирования, то отбор останавливается.

Моделирование исторической динамики осуществляется на основе искусственных нейронных сетей, а именно многослойного перцептрона, подробно описанного в [5]. Сеть состоит из входного, выходного и одного скрытого слоев. Число нейронов во входном слое определяется числом рассматриваемых факторов, число нейронов в скрытом слое на первом этапе составляет фиксированное число (10), а на втором этапе, этапе оптимизации конфигурации прогнозных моделей, определяется при помощи дополнительной кросс-валидации. Обучение модели производится с применением k-fold кросс-валидации [3], в рамках которой

¹ Стажер-исследователь отдела мониторинга и прогнозирования цен ИПЦиРЕМ ВШЭ

² Заместитель директора ИПЦиРЕМ ВШЭ

выборку произвольно разбивают на 10 равных участков и оценивают точность «подгонки» теоретической и эмпирической кривой на каждом участке выборки. По итогам k-fold кросс-валидации оценивается усредненная точность «подгонки».

По итогам отбора факторов и определения конфигурации моделей производится прогнозирование спроса, цены РСВ и БР на сутки, неделю и месяц вперед. При соотнесении модельных и фактических данных делаются выводы о точности предлагаемых моделей. Точность прогноза оценивается на основе средней абсолютной ошибки и аналогичного показателя в процентах.

2 Исходные данные и факторы прогноза

В качестве исходных данных были приняты временные ряды почасовой динамики спроса, цены РСВ и цены БР первой ценовой зоны оптового рынка электроэнергии и мощности [1]. Данные охватывают период с 8 июля 2010 года по 30 сентября 2012 года.

На основе имеющейся статистики все параметры (факторы) можно разделить на следующие группы:

- **Отраслевые.** К настоящей группе относятся факторы, отражающие специфику функционирования рынка, а именно объем спроса и цены РСВ и БР. В связи с маргинальным принципом ценообразования на рынке РСВ, спрос на электроэнергию рассматривается как один из ценообразующих факторов, используемых при прогнозировании спотовой цены на электроэнергию. При прогнозировании цены БР помимо величины спроса на электроэнергию, также учитывается значение цены РСВ, так как в силу «природы» балансирующего рынка электроэнергии (а именно балансирование объемов спроса и предложения, которые в течение операционных суток могут отклоняться от соответствующих значений, указанных в заявках игроков на РСВ) цена БР колеблется вокруг значения цены РСВ.
- **Сезонные.** Рынку электроэнергии присуща высокая степень сезонности. В настоящую категорию отнесены следующие факторы: час; день недели; фиктивная переменная, отражающая рабочий или выходной день; фиктивная переменная, отражающая один из четырех сезонов года. На рисунке 1 представлены диаграммы, характеризующие сезонность спроса на электроэнергию. Наиболее ярко выражена часовая сезонность спроса на рынке: пики потребления электроэнергии наблюдаются в утренние и вечерние часы, в ночное время можно отметить падение уровня электропотребления. Наименьшие сезонные колебания спроса на электроэнергию наблюдаются в разрезе дней недели, но тем не менее можно выделить небольшие отклонения средних значений.
- **Исторические.** В данную группу факторов отнесены исторические данные по динамике индексов цен: значение параметра час назад; значение параметра в аналогичный час сутки назад; значение параметра в аналогичный час неделю назад. На рисунке 2 представлен график автокорреляции, на котором показано, что текущее значение цены сильно взаимосвязано со значением цены с лагом 1, 24 и 168 (час, сутки и неделя).

3 Результаты отбора факторов прогнозирования

Последовательный отбор факторов позволил определить те входные параметры, подача на вход нейронной сети которых позволяет получить наиболее «аккуратную» теоретическую кривую, описывающую ретроспективную динамику анализируемых показателей. На рисунке 3 представлены результаты отбора.

Отбор факторов при моделировании спроса прошел в 5 шагов. На первых четырех шагах были отобраны два сезонных фактора (час и фиктивная переменная «рабочий/выходной день») и два исторических фактора (значение спроса в предшествующий час и в аналогичный час в предшествующий день).

При моделировании исторической динамики цены РСВ был учтен фактор спроса, так как, как уже было отмечено выше, в силу маргинального ценообразования объем спроса определяет равновесную цену в точке пересечения с величиной предложения. Результаты расчетов показали, что также как при моделировании спроса для моделирования динамики цены РСВ четырех принятых факторов (два сезонных и два исторических) достаточно.

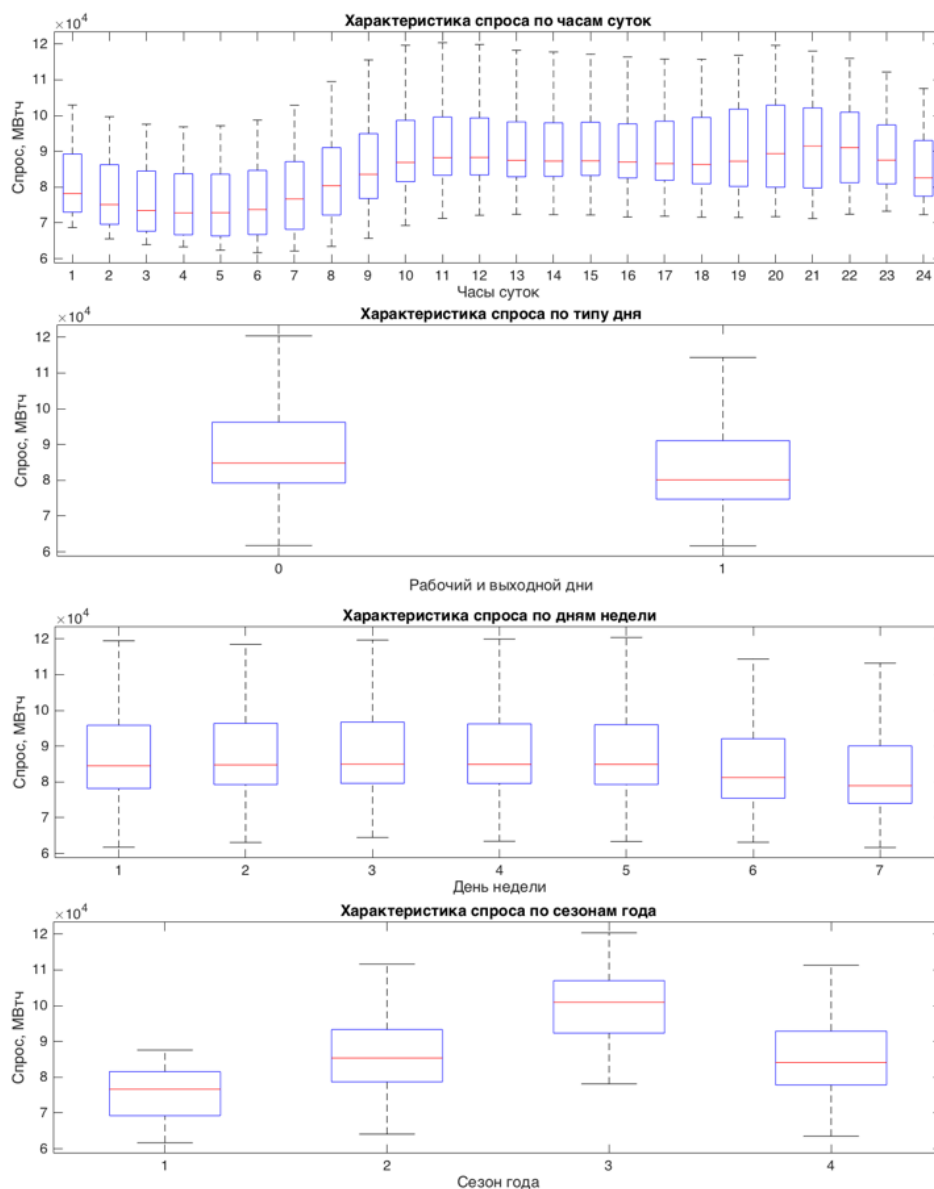


Рис. 1: Сезонность спроса в первой ценовой зоне ОРЭМ

При моделировании цены БР наряду с историческими и сезонными факторами, были использованы отраслевые параметры (цена РСВ и спрос). При этом максимально возможная точность «определилась» только при включении трех факторов: час дня, значение цены в предшествующем часе и в аналогичный час на предшествующей неделе. С точки зрения точности моделирования наилучшая «подгонка» эмпирической и теоретической кривой получилась для величины спроса - средняя абсолютная ошибка в процентах составила 0.56%. Моделирование цены РСВ показало большее отклонение - ошибка моделирования составила 2.56%. Наихудший результат показывает моделирование цены БР, ошибка которого составила 19.9%. Включение даже таких значимых системных отраслевых параметров как спрос и цена БР не позволило улучшить данный результат.

4 Результаты оптимизации конфигурации нейронных сетей

После завершения отбора факторов архитектура нейронных сетей (число нейронов в слоях перцептрона) практически predetermined. Число нейронов во входном слое соответствует числу прогнозных факторов, выходных нейронов - по одному в каждой модели, которые соответствуют моделируемому спросу, цене РСВ и БР.

Для окончательной конфигурации сетей необходимо определить оптимальное число нейронов в скрытом

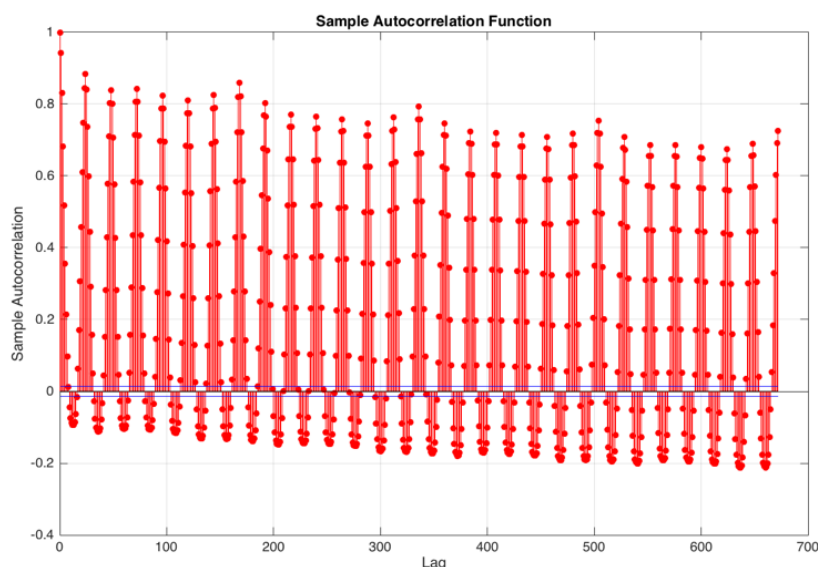


Рис. 2: График автокорреляция цены РСВ в первой ценовой зоне ОРЭМ.

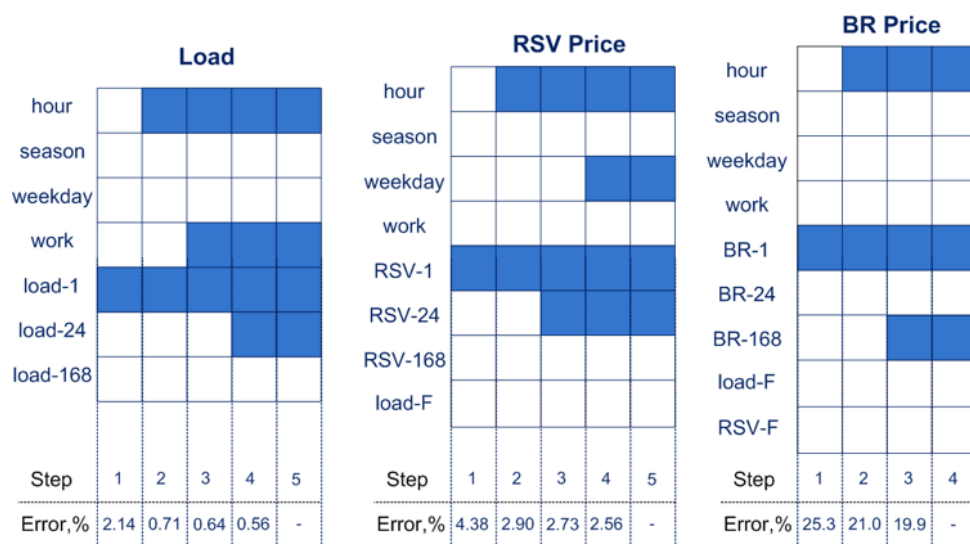


Рис. 3: Результаты отбора факторов с точки зрения ошибки моделирования

слое. Слишком большое число нейронов в скрытом слое (для рассматриваемых случаев это более 15) может привести к переобучению сети и сложному вычислительному процессу, а слишком малое число нейронов (например, один) - к высокой ошибке моделирования. Для поиска оптимального количества нейронов была использована 10-fold кросс-валидация [3], результаты которой представлены на рисунке 4. Для моделирования спроса в ходе исследования было выбрано 10 нейронов, для цены РСВ – 12, а для цены БР – 15 нейронов в скрытом слое сети.

5 Результаты прогноза спроса и цен на оптовом рынке электроэнергии

Сформированные в ходе исследования три модели на основе нейронных сетей были использованы для почасового прогнозирования спроса, цены рынка на сутки вперед и цены балансирующего рынка на сутки, неделю и месяц вперед в первой ценовой зоне оптового рынка электроэнергии России. Прогноз осуществлялся на ноябрь 2012 года. Точность прогнозирования оценивалась на основе средней абсолютной ошибки MAE и аналогичного показателя в процентах MAPE. Результаты представлены в таблице 1. В целом, мо-

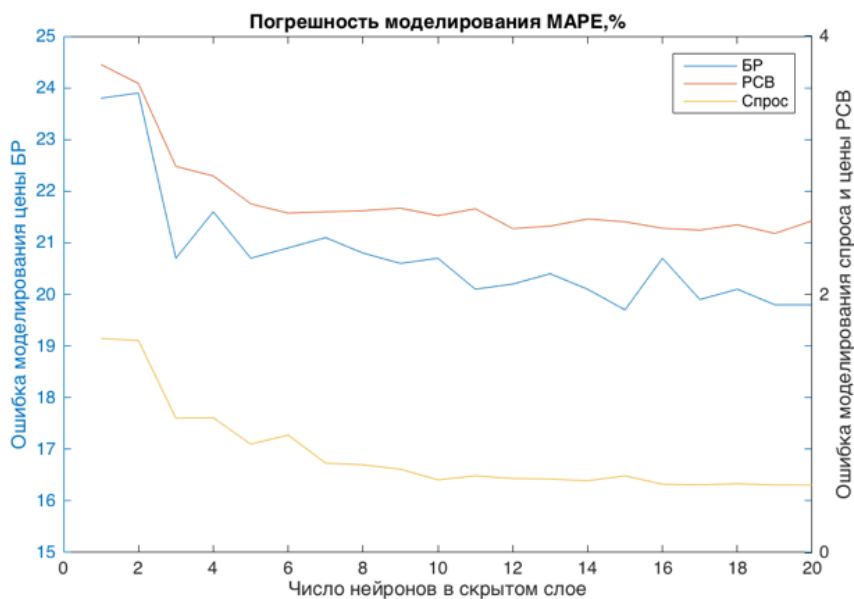


Рис. 4: Результаты 10-fold кросс-валидации. Выбор оптимального числа нейронов в скрытом слое.

дели прогнозирования демонстрируют удовлетворительный результат.

Наибольшая точность прогнозирования достигается при определении величины спроса на электроэнергию - ошибка прогнозирования не превышает одного процента на любом рассматриваемом временном горизонте. Ошибка прогноза цены РСВ немногим выше ошибки прогноза спроса и не превышает 3%.

Наименьшая точность прогнозирования отмечена при прогнозировании цены БР - ошибка прогноза на неделю и месяц вперед оказалась на порядок выше, чем при прогнозировании спроса и цены РСВ на рынке электроэнергии. Вместе с тем, в ходе исследования было выявлено наличие аномально низкого значения цены БР в используемых фактических данных, что привело к значительному росту оценки MAPE (прогнозное значение в точке аномалии отличается от фактического более, чем на 1000%).

Исключение аномалии из временного ряда позволило снизить ошибку прогноза цены БР на неделю и месяц вперед до 4,4% и 6,5% соответственно. Таким образом, последние оценки также позволяют судить об адекватности прогнозирования цены БР.

Таблица 1: Ошибки прогноза показателей в первой ценовой зоне ОРЭМ на ноябрь 2012. Спрос в МВтч, цены в руб/МВтч.

Горизонт	Спрос		Цена РСВ		Цена БР	
	MAE	MAPE	MAE	MAPE	MAE	MAPE
День	538.4	0.60%	25.5	2.30%	50.8	5.56%
Неделя	738.5	0.83%	22.7	2.20%	45.9	47.49%
Месяц	718	0.78%	25.8	3.00%	48.1	17.26%

6 Заключение

Моделирование исторической динамики объема спроса на электрическую энергию, цены на рынке на сутки вперед и цены на балансирующем рынке электроэнергии в России выявило, что выбор набора исходных факторов в значительной степени определяет точность «подгонки» теоретической и фактической кривых.

Наиболее точными являются модели, включающие ограниченный набор исходных экзогенных факторов и не перегруженные избыточной информацией.

Высокая точность прогнозирования рассматриваемых параметров обеспечивается на основе сезонных и соответствующих исторических данных. При этом важно подчеркнуть, что факторы сезонности являются

четко детерминированными на перспективу.

Системные, или отраслевые, факторы, такие как величина спроса (при прогнозировании цены РСВ) и цена РСВ (при прогнозировании цены БР), в достаточной степени не оказывают положительного результата на аккуратность моделирования. Данный эффект вызван влиянием на данные факторы такого параметра, как «поведенческая стратегия» игроков рынка, который является в значительно меньшей степени предопределенным по сравнению с рассматриваемыми выше факторами.

Низкая ошибка прогнозирования достигается при прогнозировании параметров функционирования рынка электроэнергии, которые наименее подвержены влиянию «поведенческих стратегий» игроков рынка (такие как, спрос на электроэнергию). Наиболее высокие ошибки прогнозирования показывают ценовые параметры, ввиду наличия влияния на них не только детерминированных факторов.

Список литературы

- [1] Математическое бюро – Официальный сайт. Прогнозирование на ОРЭМ. - <http://www.mbureau.ru/>.
- [2] Bastian J, Zhu J, Banunaryanan V, Mukherji R. Forecasting energy prices in a competitive market. IEEE Comput Appl Power 1999(July):40–5.
- [3] Fushiki T. Estimation of prediction error by using K-fold cross-validation. Statistics and Computing — 2011, Volume 21, Issue 2, pp. 137-146
- [4] Ghosh D. Game Theoretic Bidding Strategies for Auctions in Green Electricity Markets"(2010). Honors Scholar Theses. Paper 149.
- [5] Mashor M. Y. Hybrid multilayered perceptron networks. International Journal of Systems Science — 2000, Volume 31, Issue 6, pp. 771-785.
- [6] Pudil P., Novovicova J., Kittler J. Floating search methods in feature selection. Pattern Recognition Letters. №15 (1994) pp. 1119-1125.
- [7] Voronin S., Partanen J. Forecasting electricity price and demand using a hybrid approach based on wavelet transform, ARIMA and neural networks. International Journal of Energy Research. Vol. 38. 2014. pp. 626-637.
- [8] Ziel F., Steinert R., Husmann S. Forecasting day ahead electricity spot prices: The impact of the EXAA to other European electricity markets. January 6, 2015.