

Прогнозирование является одной из основных целей и задач большого числа специалистов, занимающихся анализом данных. Современные методы статистического прогнозирования позволяют с высокой точностью прогнозировать практически все возможные показатели. Прогнозирование является весьма полезным инструментом поддержки принятия решений для участников оптового рынка электрической энергии и мощности (ОРЭМ). Задача прогнозирования состоит в том, чтобы по значениям наблюдений, собранных в данный момент времени, определить значения в следующий момент. Не существует универсальных методик прогнозирования на все случаи жизни. Выбор метода прогнозирования и его эффективность зависят от многих условий, и в частности от требуемой длины или времени прогнозирования. По времени упреждения экономические прогнозы делятся на: • оперативные (с периодом упреждения до одного месяца); • краткосрочные (период упреждения - от одного, нескольких месяцев до года); • среднесрочные (период упреждения более 1 года, но не превышает 5 лет); • долгосрочные (с периодом упреждения более 5 лет). Наибольший практический интерес, безусловно, представляют краткосрочные и оперативные прогнозы. Статистическое описание развития экономических процессов во времени осуществляется с помощью временных рядов. Временным рядом (рядом динамики, динамическим рядом) называется последовательность значений показателя (признака), упорядоченная в хронологическом порядке, т.е. в порядке возрастания временного параметра. Отдельные наблюдения временного ряда называются уровнями этого ряда. На стадии проведения графического анализа можно исследовать компонентный состав временных рядов, а также сделать первые шаги к выбору модели для описания их динамики и последующего прогнозирования. В теории временных рядов разработаны различные методы исследования: корреляционный и спектральный анализы, методы экспоненциального сглаживания и фильтрации, модели авторегрессии и проинтегрированного скользящего среднего (АРПСС - ARIMA), которые позволяют отслеживать качество прогноза в зависимости от срока прогноза. В анализе временных рядов, как и в большинстве статистических методов, предполагается, что исходные данные содержат детерминированную и случайную составляющие. В общем случае детерминированная составляющая может быть представлена в виде комбинации следующих компонент: • тренд представляет собой общую систематическую линейную или нелинейную компоненту, которая может изменяться во времени; • цикл - более или менее регулярные колебания относительно тренда; • сезонная составляющая - это периодически повторяющаяся компонента. При различных сочетаниях этих компонент в изучаемом процессе или явлении зависимость уровней ряда от времени может принимать различные формы. Во-первых, большинство временных рядов экономических показателей имеют тенденцию, характеризующую долговременное совокупное воздействие множества

факторов на динамику изучаемого показателя. Очевидно, что эти компоненты, взятые в отдельности, могут оказывать разнонаправленное влияние на исследуемый показатель. Однако в совокупности они формируют его возрастающую или убывающую тенденцию. Во-вторых, изучаемый показатель может быть подвержен циклическим колебаниям. Эти колебания могут носить сезонный характер, поскольку деятельность ряда отраслей экономики зависит от времени года. Временной ряд может быть представлен различными математическими моделями. Аддитивная модель записывается в виде:
$$Y_t = T_t + S_t + I_t$$
 где T_t - тренд, S_t - соответственно циклическая, сезонная и случайная остаточная составляющие. Выбор модели зависит от конкретной совокупности явлений определяющих данный временной ряд, и их взаимосвязей. Целью проведенного исследования является прогнозирование оперативных показателей ОРЭМ - динамики продаж электроэнергии на различных секторах: Регулируемые договоры (РД), Рынок на сутки вперед (РСВ). Статистическая база исследования сформирована из данных по оперативным показателям продажи электроэнергии по различным договорам (секторам ОРЭМ) ежемесячно за период 2009-2013г., в процентных показателях. Исследования проводились на реальных данных, предоставленных ОАО «Генерирующая компания» [1]. Компьютерный эксперимент выполнялся в программной среде Statistica, которая является комплексным аналитическим инструментом, предназначенным для выбора моделей АРПСС и построения прогнозов в любых областях. Предварительный визуальный анализ в среде Statistica дает основания полагать, что наблюдаемый временной ряд порождается моделью АРПСС. Подбор конкретной модели из этого класса предусматривает три этапа: - идентификация модели; - оценивание модели; - диагностика модели. До того, как начать оценивание параметров модели, необходимо решить, какой тип модели будет подбираться к данным, и какое количество параметров присутствует в модели, иными словами, нужно идентифицировать модель. В АРПСС модели имеется 4 типа параметров: p - параметр авторегрессии, P - сезонный параметр авторегрессии, q - параметр скользящего среднего, Q - сезонный параметр скользящего среднего. По полученным графикам автокорреляционной функции (АКФ, рис.1.) и частной автокорреляционной функции (ЧАКФ, рис.2), можно определить какие параметры, и в каком количестве должны присутствовать в модели. Опираясь на свойства АКФ и ЧАКФ за годы практической работы с авторегрессионными моделями, исследователями были сформулированы правила выбора наилучшей модели [2]. Рис. 1 - Автокорреляционная функция для АРПСС модели. Полученные данные (рис.1, 2) соответствуют модели АРПСС (1,0), то есть модели авторегрессии первого порядка, так как АКФ - экспоненциально затухает, а ЧАКФ - имеет выброс (пик) на лаге 1. Аналогичные функции получены и для исследуемого показателя РСВ. Рис. 2 - Частная автокорреляционная функция для

АРПСС модели После идентификации моделей переходим к непосредственной оценке ее параметров. В результате получили таблицу со значением $p(1)$ соответствующей модели авторегрессии первого порядка (рис.3). Во второй колонке этой таблицы - точечная оценка параметра $p(1)$, в третьей - асимптотическая стандартная ошибка, в четвертой - значения t - критерия, в пятой - уровни значимости, шестой и седьмой - соответственно верхние и нижние границы 95% доверительных интервалов для соответствующих параметров модели. Из таблицы видно, что согласно критерию Стьюдента данный параметр оценен статистически значимым. Таким образом, статистический критерий при $P 0,05$ говорит о том, что модель получилась согласованной с теорией.

а б Рис. 3 - Оценка параметров АРПСС модели для показателей: а - РД, б - РСВ Важнейшими характеристиками качества модели, выбранной для прогнозирования, являются показатели ее точности. Они описывают величины случайных ошибок, полученных при использовании модели. Таким образом, чтобы судить о качестве выбранной модели, необходимо проанализировать систему показателей, характеризующих как адекватность модели, так и ее точность. Для проверки адекватности модели воспользуемся визуальными методами, представленными в системе Statistica. Дополнительным тестом на проверку адекватности модели, является анализ случайной составляющей - остатков. Точки на графике расположения остатков на нормальной вероятностной бумаге лежат близко к прямой, что хорошо характеризует модель. Вследствие того, что приращения случайной составляющей должны быть независимы, автокорреляции и частные автокорреляции остатков, не должны выходить за допустимые интервалы, что также подтверждается для выбранной модели. На гистограмме распределении остатков они подчиняются нормальному закону распределения (рис.4).

Рис. 4 - Гистограмма остатков Результаты вышеперечисленных тестов подтверждают, что построенная динамическая модель является статистически значимой, адекватна фактическим данным и, следовательно, пригодна для прогнозирования неизвестных уровней ряда на перспективу. О точности прогноза можно судить по величине ошибки (погрешности) прогноза. Ошибка прогноза - величина, характеризующая расхождение между фактическим и прогнозным значением показателя. Очевидно, что все указанные характеристики могут быть вычислены после того, как период упреждения уже закончился, и имеются фактические данные о прогнозируемом показателе или при рассмотрении показателя на ретроспективном участке. Независимая проверка качества прогноза является одной из важных задач в прогнозировании. Мы должны быть уверены, что прогноз в будущем будет приближен к реальным (наблюдаемым) данным. Часто для этой цели используется алгоритм кросс - проверки. Алгоритм состоит в следующем - перед проведением анализа исходный временной ряд укорачивается на 7-10%, «хвост»

ряда сохраняется для дальнейшего анализа. После этого строится прогноз «укороченного» ряда и результат сравнивается с отложенными данными. Ошибки прогнозов, полученные ретроспективно (на втором участке) характеризуют точность применяемой модели. Таблица 1 - Результаты прогноза АРПСС модели Период прогноза 2013г. Значения РСВ,% Границы ДИ исходные прогноз-ные -90% +90% январь 76,10 74,81 58,53 91,09 февраль 76,44 74,06 51,03 97,08 март 78,68 73,33 45,13 101,5 апрель 82,14 74,10 41,54 106,7 май 78,69 72,02 35,62 108,4 июнь 73,28 72,47 32,58 112,3

Рис. 5 - Динамика уровней ряда РСВ фактических и прогнозных значений

Для оценки качества подгонки модели в целом (для всего ряда) рекомендуется использовать среднюю абсолютную относительную ошибку, которая составляет в данном случае 5%, поэтому прогноз можно считать хорошей точности. Данный показатель является относительным показателем точности прогноза и не отражает размерность изучаемых признаков, выражается в процентах и на практике используется для сравнения точности прогнозов полученных как по различным моделям, так и по различным объектам. При построении прогноза по изложенной методике предусмотрены возможности корректировки прогноза в зависимости от периода ретроспективных данных и времени предсказания. Алгоритм прогнозирования включает в себя следующие возможности: настройку прогнозных статистических моделей; отображение прогнозируемых данных в табличном и графическом виде. Одним из основных стратегических направлений на оптовом рынке электроэнергии и мощности является создание конкурентных преимуществ акционерного общества за счет использования информационных технологий, включая систему прогнозов и моделирования. Результаты данной работы и ранее выполненных [3-5] показывают, что разработку стратегии на оптовом рынке необходимо проводить на основе исторических данных результатов работы на всех секторах рынка и в первую очередь для таких оперативных показателей как: регулируемые договора (РД) и рынок на сутки вперед (РСВ).