

**Филатов Александр Юрьевич (alexander.filatov@gmail.com),
Смирнова Евгения Олеговна (smirnovevgen-91@mail.ru)
Иркутский государственный университет**

ПРОГНОЗИРОВАНИЕ ОСНОВНЫХ ХАРАКТЕРИСТИК РЫНКА ЭЛЕКТРОЭНЕРГИИ «НА СУТКИ ВПЕРЕД» ЭКОНОМЕТРИЧЕСКИМИ МЕТОДАМИ

Реформирование электроэнергетики [1] и введение с 2006 года Нового оптового рынка электроэнергии и мощности (НОРЭМ) ставит электроэнергетические компании в новые конкурентные условия, поэтому задача прогнозирования цен и объемов продаж на рынке «на сутки вперед» (РСВ) становится все более актуальной. В частности, экстраполированные значения цен и объемов РСВ необходимы организациям, регулирующим энергосистему, для обоснования своей финансовой стратегии на рынке и определения доли электроэнергии, вырабатываемой на собственных станциях [3].

Решение задачи прогнозирования проходило по трем направлениям.

1. Прогнозирование суточных цен осуществлялось на основе статистики (рис.1) по скорректированным (в частности, на индекс цен) ценам РСВ [3] (y) за 4 года (2008 – 2011). [8]. В силу длительного интервала времени важен учет изменения цен в стране, поэтому исходный ряд был индексирован. Скорректированная на инфляцию динамика изменения цен на электроэнергию представлена на рис.1:

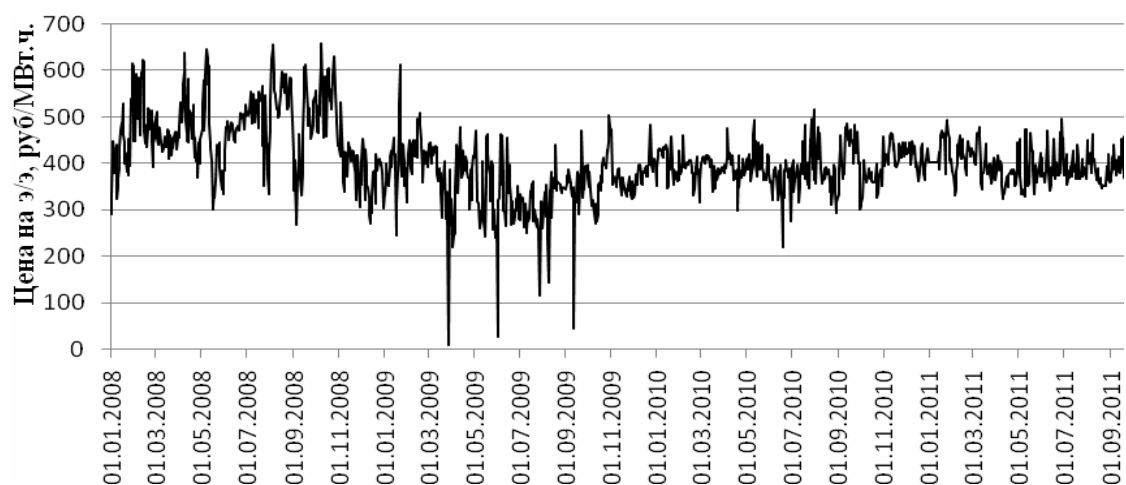


Рис.1. Динамика индексированных цен на электроэнергию

Разрешение	Временной ряд	Длина ряда	Среднее значение	Станд. отклон.	Минимум	Максимум
Суточные	Цена РСВ СЦЗ (руб./МВт.ч)	1388	404,79	72,85	10,00	657,65
	Ночная температура (°С)	1388	-2,80	14,22	-40,00	23,00
	Дневная температура (°С)	1388	6,41	15,03	-31,00	40,00
	Длина светового дня (мин.)	1388	749,84	182,11	461,00	1009,00
	Цена на газ (руб./м ³)	1388	159,37	64,67	70,11	379,33
	Курс доллара	1388	28,98	3,03	23,13	36,43
	Курс евро	1388	40,37	3,06	34,08	46,84
Почас.	Нефть (тыс.руб./баррель)	1388	2,01	0,43	0,17	3,14
	Цена РСВ СЦЗ (руб./МВт.ч)	8758	392,96	49,02	170,62	607,62

Табл.1. Параметры временных рядов

В качестве наиболее значимых факторов, влияющих на уровень цен помимо тренда (t), в модели были оставлены дамми для дней недели ($z^{(1)}-z^{(6)}$) и праздников ($z^{(7)}$), дневная температура ($x^{(1)}$) и длина светового дня ($x^{(2)}$) [9], скорректированные цены на газ ($x^{(3)}$) и нефть ($x^{(4)}$) [10], курс евро ($x^{(5)}$) [12]. Характеристики временных рядов приведены в табл.1.

Для прогнозирования суточных узловых цен необходимо использование подхода, позволяющего производить анализ временных рядов [3]. С этой целью была разработана методика прогнозирования цен на электроэнергию на оптовых энергетических рынках. Методика включает в себя следующие шаги (рис.2):

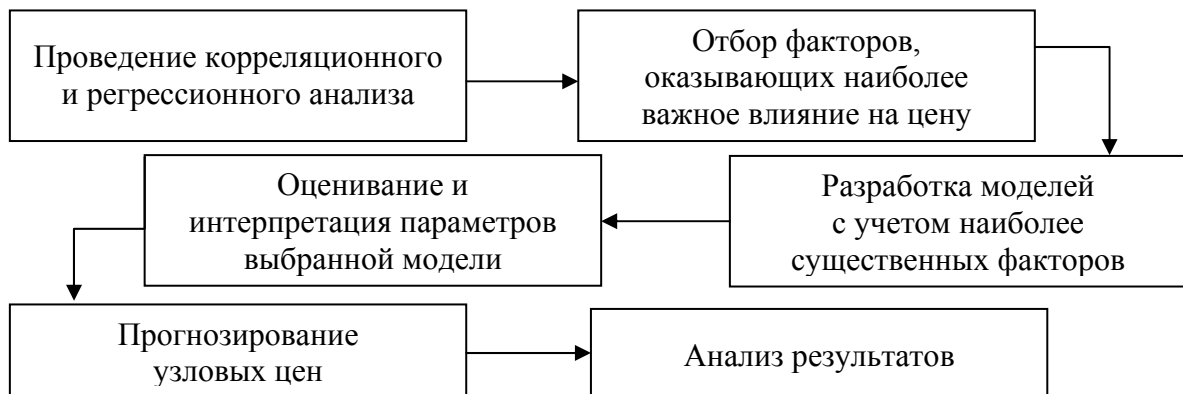


Рис.2. Методика прогнозирования цен на электроэнергию

Для исходных данных был найден оптимальный состав регрессоров. Полученная модель множественной регрессии имеет следующий вид:

$$\hat{y} = 616,14 + 0,038t^{**} + 26,97z^{(1)**} + 24,36z^{(2)**} + 18,49z^{(3)*} + 23,18z^{(4)**} + 23,53z^{(5)**} + 13,45z^{(6)**} - 30,93z^{(7)*} - 0,297x^{(1)} - 0,059x^{(2)**} + 0,024x^{(3)**} + 0,058x^{(4)*} - 7,31x^{(5)**}, \quad \hat{R}^2 = 0,357, \quad \hat{\sigma} = 65,8.$$

Здесь одной и двумя звездочками обозначены регрессоры, значимые соответственно при уровне значимости 5% и 0,1%. Поведение остатков описывается авторегрессией первого порядка AR(1):

$$\varepsilon(t) = 0,713\varepsilon(t-1) + \delta(t).$$

Коэффициент детерминации при этом увеличивается до $\hat{R}^2 = 0,685$, а среднеквадратическая ошибка прогноза будет равна $\hat{\sigma} = 40,46$.

В развитие данной модификации были построены лаговые модели, учитывающие запаздывание во влиянии цены газа на цену электроэнергии. В частности, применив преобразование Койка, добавив тренд, цикличность и внешние факторы, получили

$$\hat{y}(t) = 418,64 - 0,043t^{**} + 28,66z^{(1)**} + 25,69z^{(2)**} + 20,78z^{(3)*} + 24,53z^{(4)**} + 23,86z^{(5)**} + 14,60z^{(6)*} - 21,14z^{(7)*} + l(t),$$

$$l(t) = 0,795l(t-1)^{**} + 0,012x^{(3)**} + m(t),$$

$$m(t) = 78,40 - 0,176x^{(1)*} - 0,0035x^{(2)} - 1,89x^{(5)**}.$$

Также в работе был осуществлен анализ устойчивости, показавший, что при добавлении в модель дополнительных незначимых регрессоров и смене выборки не происходит существенного изменения коэффициентов.

2. Вторым направлением в работе являлось построение прогноза объемов продаж электроэнергии на РСВ. В качестве факторов, влияющих на объем продаж электро-

энергии, помимо вышеперечисленных, была включена цена РСВ. Динамика объемов продаж представлена на рис.3.

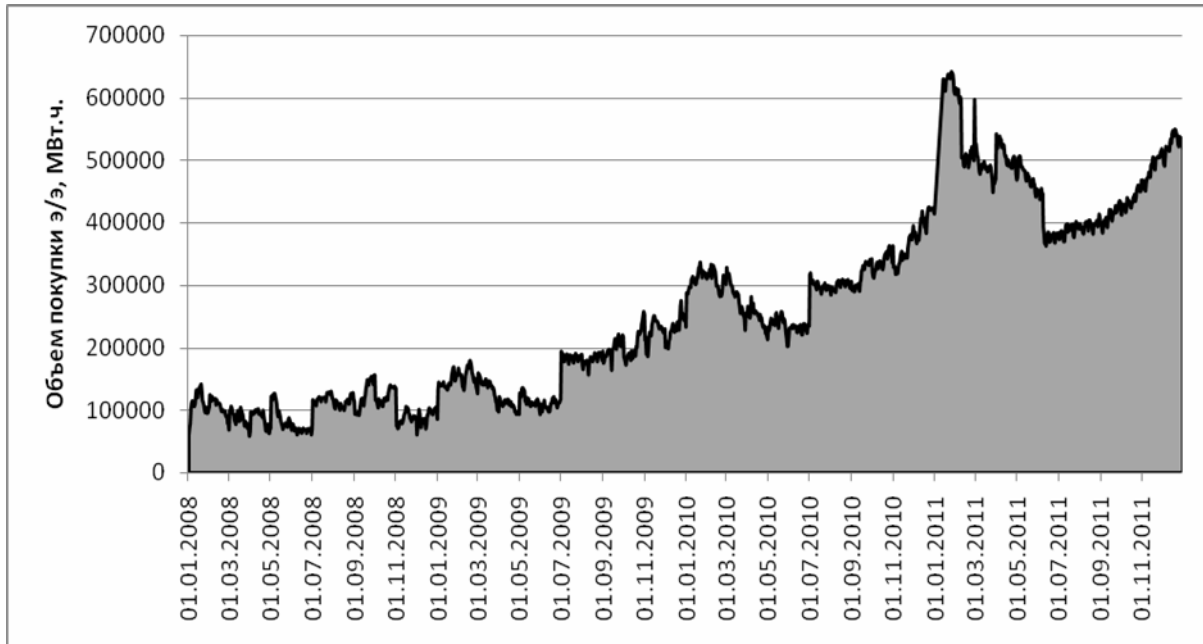


Рис.3. Динамика объемов покупки электроэнергии

Из графика видны резкие полугодовые повышения объемов, что может быть связано с окончанием действия срочных договоров и соответствующим переходом участников на рынок РСВ. Данное явление было учтено вводом дамми-переменных для каждого участка резкого повышения. Модель множественной регрессии приняла вид:

$$\hat{y} = 327,70 + 4,20z^{(1)} + 4,08z^{(2)} + 4,36z^{(3)*} + 4,32z^{(4)*} + 5,17z^{(5)*} - 15,42z^{(7)**} - 374,9g^{(1)**} - 345,1g^{(2)**} - 320,0g^{(3)**} - 242,0g^{(4)**} - 176,2g^{(5)**} - 106,5g^{(6)**} + 62,1g^{(7)**} - 0,91x^{(1)**} - 0,08x^{(2)**} + 8,19x^{(3)**} + 2,62x^{(4)**} + 0,11x^{(6)**}, \hat{R}^2 = 0,97.$$

Здесь $g^{(i)}$ – дамми-переменные для полугодовых участков и $x^{(6)}$ – цена РСВ.

Для долгосрочного прогноза значения средних за полугодие уровней продаж аппроксимировались логистической функцией.

3. На третьем этапе прогнозировались уже не среднесуточные, а почасовые цены РСВ. Рассмотрим временной ряд цен РСВ с часовым разрешением за 2010 год, длина которого составила 8758 значений [8]. Для данного временного ряда применим линейную модель множественной регрессии. Предварительная обработка данных заключается в устранении аутлаеров и сглаживании пиковых цен, которая проводилась уже двумя способами: скользящим средним по 3 и 5 точкам. Также необходимо ввести дамми-переменные для часов, чтобы выявить суточную циклическую составляющую. Тогда, общий вид модели будет представлен формулой:

$$\hat{Y}_i = c + at + \sum_{i=1}^5 \theta_i x_i + \sum_{i=1}^{23} d_i z_i + \sum_{i=1}^7 \tilde{d}_i \tilde{z}_i,$$

где c – константа, t – тренд, x – внешние факторы, z – дамми-переменные для часов, \tilde{z} – дамми-переменные для дней недели и праздников.

Численное значение каждого коэффициента, а также соответствующий уровень значимости каждой переменной представлены в табл.2. Следует заметить, что практически все переменные за малым исключением очень сильно значимы (уровень значи-

мости $\alpha=0,01$). Коэффициенты детерминации этих моделей практически одинаковы и равны примерно 30%. Зато применение моделей авторегрессии первого порядка значительно улучшает основные характеристики:

$$MA(3): \varepsilon(t)=0,98\varepsilon(t-1)+\delta(t), R^2 = 0,96, \sigma = 8,91$$

$$MA(5): \varepsilon(t)=0,92\varepsilon(t-1)+\delta(t), R^2 = 0,98, \sigma = 6,52$$

Обозначение	Название	МА (3)		МА (5)	
		Коэффициент	Значимость	Коэффициент	Значимость
c	const	229,38	0,01	233,08	0,01
t	тренд	0,01	0,01	0,01	0,01
x_1	днев.темпер.	-1,17	0,01	-1,16	0,01
x_2	свет. день	0,07	0,01	0,07	0,01
x_3	цена топлива	0,46	0,01	0,46	0,01
x_4	динамика \$	-6,39	0,01	-6,40	0,01
x_5	динамика €	4,10	0,01	4,06	0,01
z_1	1 час	-0,24	0,05	1,31	0,05
z_2	2	3,22	0,05	6,83	0,01
z_3	3	13,6	0,01	15,56	0,01
z_4	4	28,18	0,01	26,26	0,01
z_5	5	41,64	0,01	37,11	0,01
z_6-z_{18}	6-18	54,74	0,01	52,79	0,01
z_{19}	19	44,54	0,01	39,43	0,01
z_{20}	20	29,10	0,01	29,23	0,01
z_{21}	21	14,72	0,01	17,49	0,01
z_{22}	22	6,16	0,01	8,15	0,01
z_{23}	23	2,10	0,05	2,49	0,05
\tilde{z}_1	понедельник	18,21	0,01	17,93	0,01
$\tilde{z}_2 - \tilde{z}_3$	вторник-среда	14,61	0,01	14,39	0,01
\tilde{z}_4	четверг	13,52	0,01	13,36	0,01
\tilde{z}_5	пятница	15,44	0,01	15,29	0,01
\tilde{z}_6	суббота	10,36	0,01	9,97	0,01
\tilde{z}_7	праздник	5,69	0,01	5,72	0,01
R^2		0,31		0,30	

Табл.2. Параметры модели, построенной на почасовых данных

Таким образом, получены очень высокие коэффициенты детерминации и очень низкие среднеквадратические ошибки для прогнозных моделей, которые ранее получить не удавалось.

Не остановимся на достигнутом и попытаемся еще уменьшить ошибку прогнозирования путем использования новых моделей. Построим модель для краткосрочного прогноза с помощью метода экстраполяции по максимуму подобия [5].

Приведем обоснование выбора данной модели. На сегодняшний день неадаптированность большинства регрессионных методов прогнозирования для решения задачи прогноза цен РСВ, вероятно, обуславливается тем, что многие из них учитывают при прогнозировании лишь последние значения временного ряда, непосредственно предшествующие прогнозируемому периоду. Опираясь в прогнозировании цен на электроэнергию лишь на предшествующие ближайшие периоды невозможно.

Поэтому предложен альтернативный метод, который основан на предположении того, что существует множество факторов, оказывающих влияние на значения процес-

са, однако определить степень влияния каждого фактора невозможно по причине объема, конфиденциальности, трудностей в измерении информации. Иными словами, если общее влияние всего множества факторов в какой-то период времени привело к тому, что процесс имел определенный профиль, то существует или когда-то случится такой период времени, когда суперпозиция влияния всего множества факторов приведет к тому, что процесс будет иметь профиль подобный исходному. Опишем это формально.

Пусть есть исходная последовательность $X(t) = [x_1, x_2, \dots, x_T]$ длины T . Для нее необходимо решить задачу экстраполяции в точках $X_{T+1}^P = [x_{T+1}, x_{T+2}, \dots, x_{T+P}]$ на основании максимума подобия. Заметим, что верхним индексом определяется длина вектора, а нижним его начало на оси времени. В качестве меры подобия двух векторов внутри одной последовательности используется коэффициент линейной корреляции:

$$L_{N,M,i} = \text{corr}(X_N^M; X_i^M), \forall N, i \in [1, T-1], \forall M \in [1, T-1]: M+N < T \cup M+i < T$$

Тогда функция подобия

$$\text{Likeness}(X_N^M) = L(i) = |\text{corr}(X_N^M, X_i^M)|, \forall i \in [1, T-1]$$

возвращает вектор значений модулей коэффициентов линейной корреляции со всеми векторами длины M , лежащими левее X_N^M на оси времени.

Вектор $X_{i_{\max}}^M$, доставляющий максимум функции подобия $L(i)$ назовем максимумом подобия для исходного вектора X_N^M .

Далее введем предположение о том, что если X_N^M и $X_{i_{\max}}^M$ имеют высокое подобие, то есть модуль линейного коэффициента корреляции близок к 1, то вектора X_N^{M+1} и $X_{i_{\max}}^{M+1}$ будут также иметь высокое подобие. Данное рассуждение назовем предположением о подобии. На основании предположения о подобии решается задача экстраполяции вектора $X(t) = [x_1, x_2, \dots, x_T]$ в точках $X_{T+1}^P = [x_{T+1}, x_{T+2}, \dots, x_{T+P}]$. Схема действия данной модели наглядно представлена на рис.4.

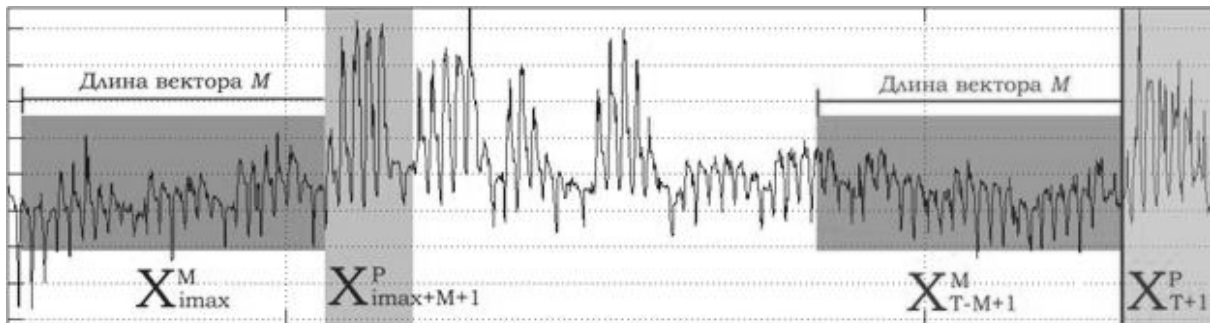


Рис 4. Положения векторов $X_{i_{\max}}^M$, $X_{i_{\max}+M+1}^P$, X_{T-M+1}^M и X_{T+1}^P на оси времени

Таким образом, экстраполированные значения исходной последовательности определяются согласно

$$X_{T+1}^P = A \times X_{i_{\max}+M+1}^P,$$

где A – матрица линейных коэффициентов.

Считаем, что для векторов X_{T-M+1}^M и $X_{i_{\max}}^M$ верно равенство

$$X_{T-M+1}^M = A \cdot X_{i_{\max}}^M + \varepsilon^M,$$

где вектор ε^M – вектор ошибок аппроксимации. Аппроксимация

$$X_{T-M+1}^M = A \cdot X_{i_{\max}}^M$$

позволяет определить матрицу A , найденную с помощью метода наименьших квадратов:

$$A = ((X_{i_{\max}}^M)^T \cdot X_{i_{\max}}^M)^{-1} \cdot (X_{i_{\max}}^M)^T \cdot X_{T-M+1}^M$$

Оценивание параметров проходило по следующей методике (рис.5). Временной ряд

был разделен на два непересекающихся отрезка: исследовательский и контрольный. Предварительная обработка данных проводилась аналогичным образом: устранение аутлаеров и сглаживание пиковых цен путем метода скользящего среднего по 3 и 5 точкам. Затем каждый из рядов (изначально не стационарный) приводился к стационарности первого порядка двумя алгоритмами: переход к ряду ошибок путем вычитания уравнения регрессии (устранение влияния тренда и внешних факторов), и переход к разностям соседних значений. Далее, проводилась экстраполяция исследовательского отрезка при различных значениях параметрах M ; базовый отрезок выступал в качестве исторических данных. Затем на основании полученного прогнозного ряда осуществляется переход к исходному ряду, в результате чего получен прогнозный ряд почасовых цен РСВ. В результате экстраполяции была получена зависимость среднеквадратической ошибки от параметра:

$$\hat{Y} = \hat{Y}_{стан} + \hat{\varepsilon}; \quad \sigma = \sqrt{\frac{1}{T} \sum_{i=1}^T (y_i - \bar{y})^2}$$

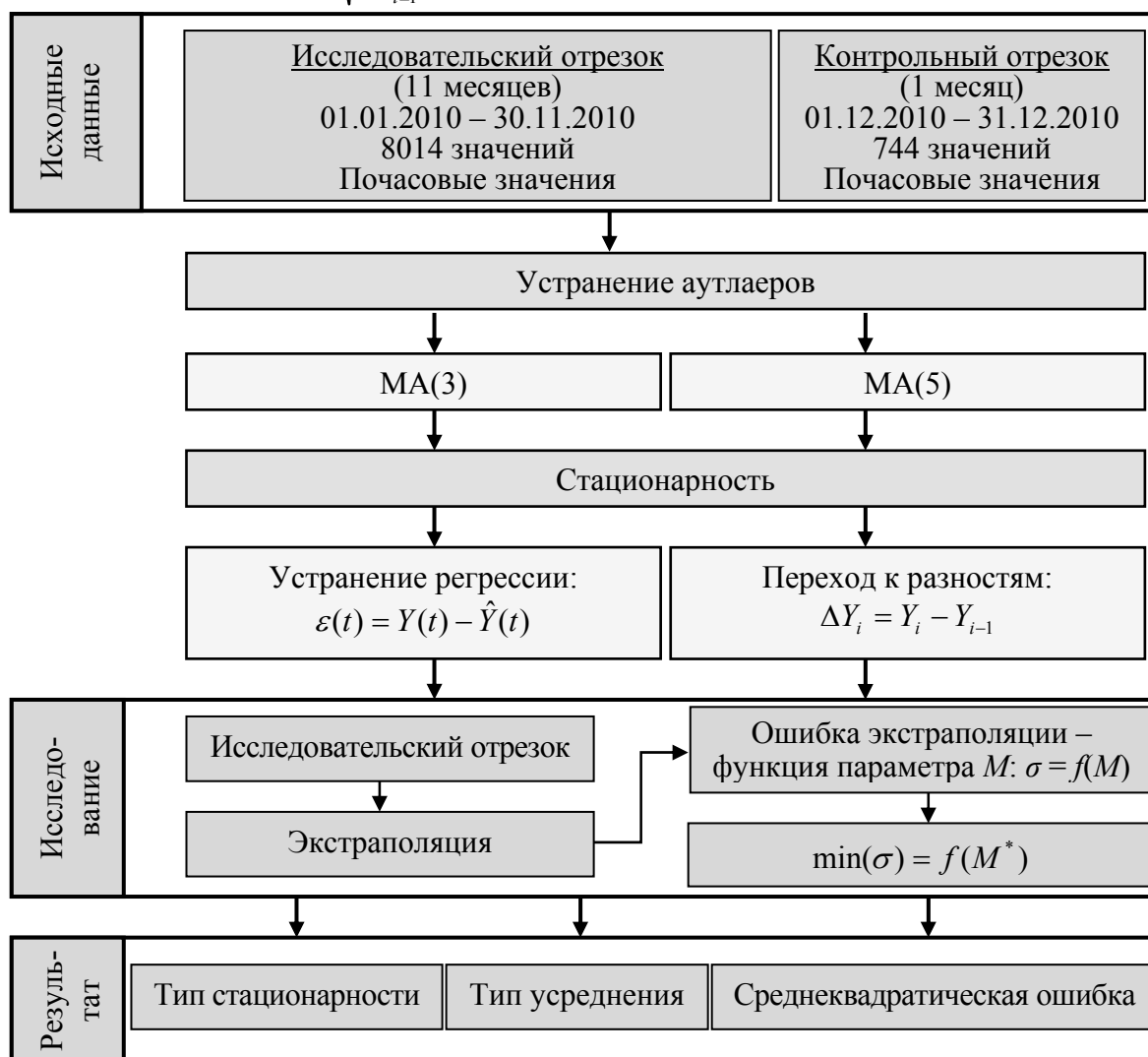


Рис.5. Алгоритм идентификации модели

На основе полученной эмпирической последовательности и построенной аналитической зависимости стандартного отклонения от длины ряда M , мы получали значение M^* , соответствующее минимуму ошибки для каждого исследуемого ряда и всех его модификаций. Результаты экстраполяции в зависимости от типа усреднения и типа стационарности представлены на рис.6. Также на каждом графике отмечено оптимальное значение M^* и соответствующая ему среднеквадратическая ошибка σ^* :

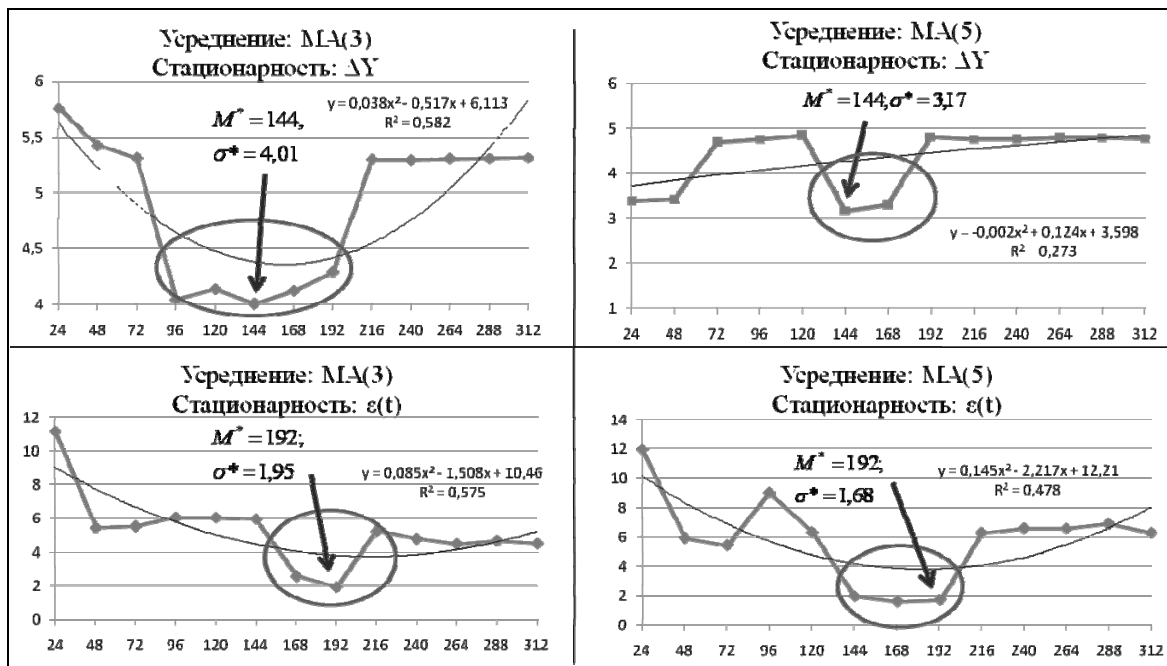


Рис.6. Зависимость точности экстраполяции от типа усреднения и стационарности

Из графиков можно сделать вывод, что для стационарных рядов разностей оптимальная длина ряда для прогноза $M^*=144$ часа или 6 суток, для ряда ошибок $M^*=192$ часа или 8 суток. Таким образом, среднее значение для обоих рядов $M^*=168$ часов или 7 суток подтверждает наличие недельных циклов.

Список литературы:

1. Постановление Правительства РФ от 31 августа 2006 г. №529 «О совершенствовании порядка функционирования оптового рынка электрической энергии (мощности)».
2. Айвазян С.А., Мхитарян В.С. Прикладная статистика. Основы эконометрики. – М.: ЮНИТИ, 2001. – 432 с.
3. Крупский А.В. Методика прогнозирования узловых цен на энергорынках // Труды всероссийской научно-технической конференции «Электроэнергетика глазами молодежи». – Екатеринбург: УрФУ. – 2010. – Т.1.
4. Таран А.С. Учет особенностей работы ТЭЦ при управлении режимами на оптовом рынке электроэнергии // Труды всероссийской научно-технической конференции «Электроэнергетика глазами молодежи». – Екатеринбург: УрФУ. – 2010. – Т.1.
5. Чучуева И.А. Сезонно-регрессионная модель прогнозирования в решении задачи прогнозирования цен РСВ // Информационно-аналитический журнал «Энерго рынок». – М. – 2009.
6. Васьковская Т.А. Анализ факторов, влияющих на ценообразование ОРЭМ. – [Электронный ресурс]. – Режим доступа: http://regeting.narod.ru/Analiz_faktorov.pdf
7. Новое в ценообразовании на рынке электроэнергии. – [Электронный ресурс]. – Режим доступа: <http://www.mosenergosbyt.ru>.
8. Открытое Акционерное Общество «Администратор торговой системы». – [Электронный ресурс]. – Режим доступа: <http://www.atsenergo.ru>
9. Расписание погоды. – [Электронный ресурс]. – Режим доступа: <http://rp5.ru/152525/ru>.
10. Рынок и аналитика «Финам». – [Электронный ресурс]. – Режим доступа: <http://finam.ru>
11. Ханин И. Алгоритмы, применяемые при прогнозировании цены и технический анализ. – [Электронный ресурс]. – Режим доступа: <http://regeting.narod.ru/Prognoz.ppt>.
12. Центральный Банк Российской Федерации. – [Электронный ресурс]. – Режим доступа: <http://www.cbr.ru>.