

Лабораторная работа

Целью работы является рассмотрение наиболее методов краткосрочного прогнозирования и осуществление прогнозирования (на примере объемов производства электроэнергии в РФ).

Задачами лабораторной работы является:

1. Изучение методов прогнозирования.
2. Изучение оценка качества прогнозной модели.
3. Практическое изучение методов прогнозирования.

Теоретическая часть

Методы прогнозирования и оценки качества прогнозной модели

Прогнозирование – разработка прогноза, т.е. специальное научное исследование перспектив (прошлых тенденций) развития каких-либо явлений (технических, социально-экономических).

Временной ряд – это совокупность значений какого-либо показателя за несколько последовательных моментов времени.

Важность прогнозирования для народного хозяйства отражается в сферах его применения как средства:

- а) предпланового анализа систем, процессов, объектов с целью установления достоверных показателей функционирования;
- б) расчета показателей, не принятых к утверждению как плановых, например, показателей развития социально-экономических систем;
- в) исследования многомерных и многофакторных зависимостей;
- г) моделирования сложных, а также трудно формализуемых и комплексных показателей;
- д) решения вероятностных задач получения интервальных оценок прогнозируемых показателей.

Виды прогнозирования во времени:

- 1) краткосрочное – 1 год;
- 2) среднесрочное – 2-5 лет;
- 3) долгосрочное – 5-10 лет;

4) сверхдолгосрочное – более 10 лет.

Период, на который делается прогноз, называется шагом прогнозирования. Различают прогнозирование с постоянным и переменным шагом прогнозирования.

При прогнозировании используются три группы методов:

- экстраполяция (интерполяция);
- моделирование;
- опрос экспертов.

Экстраполяция изучает явление и переносит тенденции этого изученного явления на другую часть этого явления или переносит прошлые тенденции явления на будущие периоды. В научном плане примером является математическая статистика.

Интерполяция ищет промежуточные значения величины (параметра) по некоторым известным ее значениям.

Каждый прогноз имеет свои цели и задачи. Прогнозы могут быть классифицированы следующим образом:

- по виду объекта;
- по масштабу объекта;
- по затратам ресурсов;
- по времени;
- по числу факторов;
- по применяемым методам и пр.

Под **методами прогнозирования** следует понимать совокупность приемов и способов мышления, позволяющих на основе анализа ретроспективных данных, экзогенных (внешних) и эндогенных (внутренних) связей объекта прогнозирования, а также их измерений в рамках рассматриваемого явления или процесса вывести суждения определенной достоверности относительно его (объекта) будущего развития.

Классификация методов по данному признаку отображена на рисунке 1.. Как свидетельствует схема, представленная на рисунке 1, по степени

формализации, методы прогнозирования можно разделить на *интуитивные* и *формализованные*. Интуитивные методы прогнозирования используются в тех случаях, когда невозможно учесть влияние многих факторов из-за значительной сложности объекта прогнозирования.

Классификационная схема методов прогнозирования

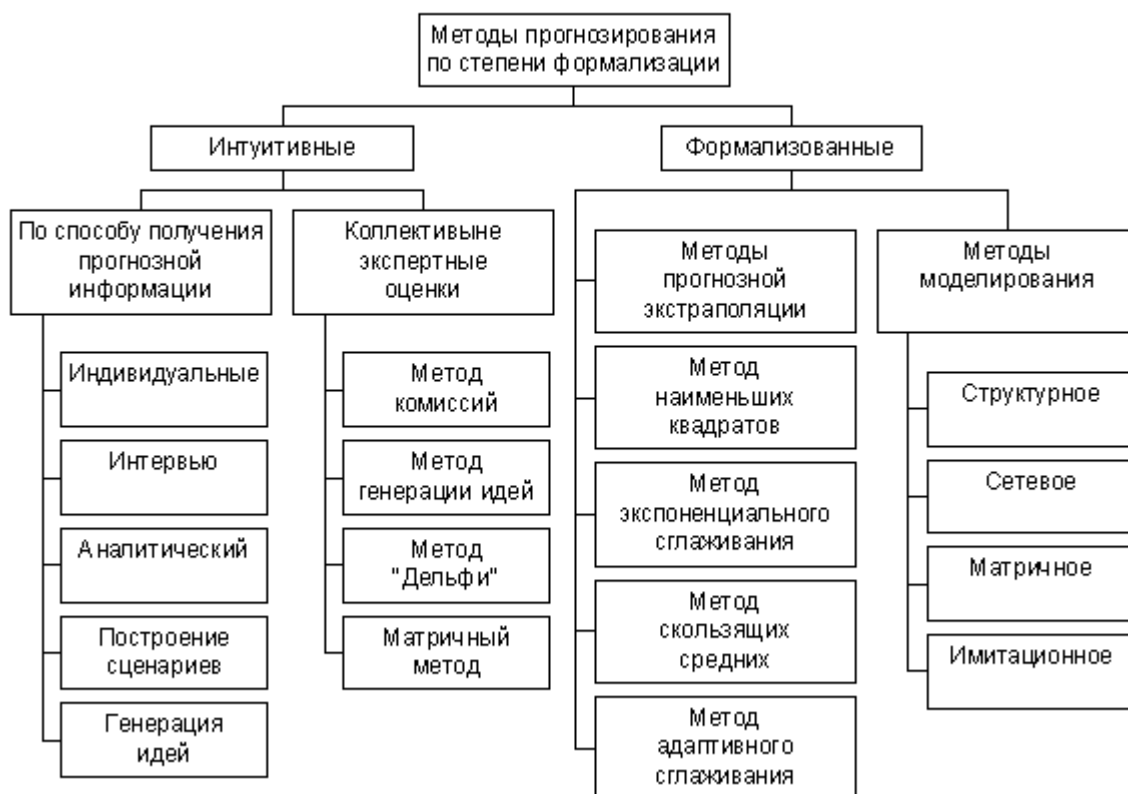


Рисунок 1 – Классификация методов прогнозирования

В лабораторной работе представлен один из способов построения прогноза выручки предприятия с сезонным характером производства электроэнергии. Важно заметить, что понятие <сезон> в прогнозировании применим к любым систематическим колебаниям, например, если речь идёт об изучении объемов выработки электроэнергии в течение года под термином <сезон> понимается один месяц. Кроме того, цикл колебаний может существенно отличаться (как в большую, так и в меньшую сторону) от величины один год. И если удаётся выявить величину цикла этих колебаний, как в данной работе, то такой временной ряд исходных данных можно

использовать для прогнозирования с использованием аддитивных и мультипликативных моделей.

Общая идея таких моделей состоит в том, что прогнозы вычисляются не только по предыдущим наблюдениям (как в простом экспоненциальном сглаживании, модели Хольта и Брауна, методе наименьших квадратов), но и с некоторыми задержками, что позволяет независимо оценить тренд и сезонную составляющую. Gardner (1985) обсудил различные модели в терминах сезонности (отсутствует, аддитивная сезонность, мультипликативная) и тренда (отсутствует, линейный тренд, экспоненциальный, демпфированный).

Качество прогнозирования, прежде всего, характеризуется ошибкой прогноза, (математическое определение качества прогноза): предполагается, что чем меньше ошибка прогноза, тем выше его качество.

Все существующие методики оценки качества прогнозирования можно условно разделить на три группы показателей: абсолютные; сравнительные; качественные.

Абсолютные показатели оценки качества прогноза позволяют количественно определить величину ошибки прогноза в единицах измерения прогнозируемого объекта или в процентах. Это среднеквадратическая ошибка (σ_t), абсолютная ошибка (Δ_{np}), средняя абсолютная ошибка ($\overline{\Delta_{np}}$), относительная ошибка (ε_{np}) и средняя относительная ошибка ($\overline{\varepsilon_{np}}$).

Абсолютная ошибка прогноза может быть определена как разность между фактическим значением (x_i) и прогнозом (x_i^*), значит,

$$\Delta_{np} = x_t - x_t^*;$$

среднее абсолютное значение ошибки

$$\overline{\Delta_{np}} = \frac{\sum_{i=1}^n |x_i - x_i^*|}{n}.$$

Среднее абсолютное значение всегда неотрицательно. Среднеквадратическая ошибка прогноза рассчитывается по формуле:

$$\sigma_i = \sqrt{\frac{\sum_{i=1}^n (x_i - x_i^*)^2}{n}}, \text{ где } n - \text{период упреждения.}$$

Недостатком рассматриваемых показателей является то, что значение этих характеристик существенно зависит от масштаба измерения уровней исследуемых явлений. Поэтому абсолютная ошибка прогноза может быть выражена в процентах относительно фактических изменений показателя следующим образом:

$$\Delta_{np} = \frac{x_i - x_i^*}{x_i} \cdot 100\%$$

а средняя относительная ошибка рассчитывается как

$$\overline{\Delta_{np}} = \frac{1}{n} \cdot \sum_{t=1}^n \frac{|x_t - x_t^*|}{x_t} \cdot 100\%$$

Данный показатель используется, как правило, при сравнении точности прогнозов разнородных объектов прогнозирования.

Сравнительные показатели оценки качества прогноза основаны на сравнении ошибки рассматриваемого прогноза с эталонными прогнозами определенного вида. Один из типов таких показателей может быть представлен следующим образом:

$$k = \sqrt{\frac{\sum_{t=1}^n (p_t - x_t)^2}{\sum_{t=1}^n (p_t^* - x_t)^2}},$$

где p_t^* - прогнозируемое значение величины эталонного прогноза. В качестве эталонного прогноза может быть выбрана простая экстраполяция, постоянный темп прироста и т.д.

Качественные показатели позволяют повести некоторый анализ видов ошибок прогнозов, разложить их на какие-либо составляющие. Особенно такой анализ важен для циклически меняющихся переменных, когда

необходимо прогнозировать не только общее направление развития, но и поворотные точки цикла, в которых меняются коэффициенты адаптации прогнозной модели.

Все рассмотренные выше показатели точности прогноза используются при проверке точности прогноза, полученного в виде точечных оценок.

Выбор показателей точности зависит от задач, которые ставит перед собой исследователь при анализе точности прогноза.

Важным критерием правильности применения прогнозной модели является проверка на адекватность. Адекватными моделями считаются такие, для которых остаточная компонента имеет свойства независимости, случайности и нормальности распределения.

В лабораторной работе использовались абсолютные показатели оценки качества прогноза.

Математическое обеспечение задачи прогнозирования

В работе использовались аддитивная модель, мультипликативная модель, моделирование с помощью фиктивных переменных для прогноза временного ряда. Остановимся поподробнее на каждом из них.

В большинстве случаев фактический уровень временного ряда можно представить как сумму трендовой, циклической и случайной компонент. Модель, в которой временной ряд представлен как сумма перечисленных компонент, называется **аддитивной моделью временного ряда** (общий вид аддитивной модели: $Y=T+S+E$).

Модель, в которой временной ряд представлен как произведение перечисленных компонент, называется **мультипликативной моделью временного ряда** (общий вид мультипликативной модели: $Y=T*S*E$).

Выбор одной из 2 моделей осуществляется на основе анализа структуры сезонных колебаний. Если амплитуда колебаний приблизительно постоянна, строят аддитивную модель временного ряда, в которой значения сезонной компоненты предполагаются постоянными для различных циклов.

Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель временного ряда, которая ставит уровни ряда в зависимости от значений сезонной компоненты.

Процесс построения модели включает в себя следующие шаги:

1. Выравнивание исходного ряда методом скользящей средней
2. Расчет значений сезонной компоненты S
3. Устранение сезонной компоненты из исходных уровней ряда и получение выровненных данных $(T+E)$ в аддитивной или $(T*S)$ в мультипликативной модели.
4. Аналитическое выравнивание уровней $(T+E)$ или $(T*S)$ и расчет значений T с использованием полученного уравнения тренда.
5. Расчет полученных по модели значений $(T+S)$ и $(T*S)$.
6. Расчет абсолютных и относительных ошибок.

Существует еще один метод моделирования временного ряда содержащего сезонные колебания построение модели регрессии с включением фактора времени и фиктивных переменных. Количество фиктивных переменных в такой модели должно быть на единицу меньше числа моментов (периодов) времени внутри одного цикла колебаний. Каждая фиктивная переменная отражает сезонную (циклическую) компоненту временного ряда, для какого либо одного периода. Она равна единице для данного периода и нулю для всех остальных периодов.

Пусть имеется временной ряд, содержащий циклические колебания периодичностью k . Модель регрессии с фиктивными переменными для этого ряда будет иметь вид:

$$y_t = a + b * t + c_1 * x_1 + c_j * x_j + \dots + c_{k-1} * c_{k-1} + \varepsilon_t$$

a, b, c_j - параметры, характеризующие среднее абсолютное изменение уровней ряда под воздействием тенденции.

t - время,

ε_t - случайная компонента,

y_t - моделируемый показатель в точке t .

$$\text{где } x_j = \begin{cases} 1 & \text{для каждого } j \text{ внутри каждого цикла,} \\ 0 & \text{во всех остальных случаях.} \end{cases}$$

В сущности, модель регрессии с фиктивными переменными есть аналог аддитивной модели временного ряда, поскольку фактический уровень временного ряда - это сумма трендовой, сезонной и случайной компонент.

Практическая часть

Полное исследование проблем, связанных с прогнозированием конкретного показателя, сводится к выполнению следующих этапов:

1. постановка задачи исследования и сбор исходной информации;
2. предпрогнозный анализ данных;
3. выбор моделей прогнозирования;
4. численное оценивание параметров моделей;
5. проверка моделей на точность и адекватность;
6. выбор лучшей модели или построение обобщенной модели;
7. получение точечного интервального прогнозов;
8. анализ и интерпретация полученных прогнозных значений.

Задача исследования и исходные данные

Задачей данного исследования является прогнозирование объемов выработки электроэнергии в Российской Федерации. Исходная информация представляет из себя временной ряд размерностью равной 48 – приведена статистика производства электроэнергии в млн.кВт*час (данные в соответствии с федеральной службы государственной статистики) с 2005 по 2008 год. Для расчетов будет использоваться MS Excel.

Таблица 1 – Исходные данные

	2005	2006	2007	2008
январь	91395,7	99611,9	94970,7	103326
февраль	86617,1	90059,3	89921,2	94475,3
март	89423	91218,5	90871	93516,2
апрель	77481,9	80568	81573,1	83794,2
май	70033	74381,1	77447,2	80032,7
июнь	64712,6	68150,9	70837	73708,6
июль	66854	70567,3	72363,3	76225,5
август	68959,8	73188,4	74952,8	78022,3
сентябрь	70986,5	74016,3	77068,7	81778,9
октябрь	82235,4	86496,2	87984,6	89080,1
ноябрь	87633,1	90915,2	94944,8	89271,1
декабрь	96738,7	96607,5	102368,6	97192,2

Для наглядности посмотрим диаграмму рассеяния объема выработки электроэнергии в млн. кВт*час (рис. 2):



Рисунок 2 - Графическое представление временного ряда

Из рисунка 2 видно, что временной ряд содержит сезонные колебания. Из рисунка видно, что весной и летом объем выработки заметно падает, а осенью и зимой он растет. Это может связано с тем, что в весенне-летний период потребность в электроэнергии снижается.

Исходя из этого выбираем следующие модели:

- 1) Построение аддитивной модели $y = T + S + E$;
- 2) Построение мультипликативной модели $y = T * S * E$;
- 3) Построение модели регрессии с включением фиктивных переменных.

Оценка параметров моделей

Аддитивная модель

Для построения модели используем данные без последнего уровня ряда.

1 шаг. Проведем выравнивание исходных уровней ряда методом скользящей средней. Результаты представлены в таблице 2:

Таблица 2 – Выравнивание исходных уровней ряда

t	y	итого за год	скользящая средняя	центр. скользящая средняя	оценка сезонной компоненты S
(1)	(2)	(3)	(4)	(5)	(6)
1	91395,7				
2	86617,1	953070,8	79422,5667		
3	89423	961287	80107,25	79764,90833	9658,091667
4	77481,9	964729,2	80394,1	80250,675	-2768,775
5	70033	966524,7	80543,725	80468,9125	-10435,9125
6	64712,6	969610,8	80800,9	80672,3125	-15959,7125
7	66854	973958,9	81163,2417	80982,07083	-14128,07083
8	68959,8	977397,2	81449,7667	81306,50417	-12346,70417
9	70986,5	981110,5	81759,2083	81604,4875	-10617,9875
10	82235,4	985339,1	82111,5917	81935,4	300
11	87633,1	988368,9	82364,075	82237,83333	5395,266667
12	96738,7	992629,7	82719,1417	82541,60833	14197,09167
13	99611,9	995911,8	82992,65	82855,89583	16756,00417
14	90059,3	995780,6	82981,7167	82987,18333	7072,116667
15	91218,5	991139,4	82594,95	82788,33333	8430,166667
16	80568	991001,3	82583,4417	82589,19583	-2021,195833
17	74381,1	990653,8	82554,4833	82568,9625	-8187,8625
18	68150,9	991658,9	82638,2417	82596,3625	-14445,4625
19	70567,3	994725	82893,75	82765,99583	-12198,69583

20	73188,4	997411,1	83117,5917	83005,67083	-9817,270833
21	74016,3	999207,1	83267,2583	83192,425	-9176,125
22	86496,2	1000971,5	83414,2917	83340,775	3155,425
23	90915,2	1004023,9	83668,6583	83541,475	7373,725
24	96607,5	1005512,3	83792,6917	83730,675	12876,825
25	94970,7	1009541,9	84128,4917	83960,59167	11010,10833
26	89921,2	1015303	84608,5833	84368,5375	5552,6625
27	90871	1023658,3	85304,8583	84956,72083	5914,279167
28	81573,1	1028212,4	85684,3667	85494,6125	-3921,5125
29	77447,2	1030857,6	85904,8	85794,58333	-8347,383333
30	70837	1033078,7	86089,8917	85997,34583	-15160,34583
31	72363,3	1035664,2	86305,35	86197,62083	-13834,32083
32	74952,8	1038535,8	86544,65	86425	-11472,2
33	77068,7	1042398	86866,5	86705,575	-9636,875
34	87984,6	1045467,5	87122,2917	86994,39583	990,2041667
35	94944,8	1050177,7	87514,8083	87318,55	7626,25
36	102368	1051273,2	87606,1	87560,45417	14808,14583
37	103326	1045599,5	87133,2917	87369,69583	15956,30417
38	94475,3	1040423,1	86701,925	86917,60833	7557,691667
39	93516,2			43350,9625	50165,2375
40	83794,2				83794,2
41	80032,7				80032,7
42	73708,6				73708,6
43	76225,5				76225,5
44	78022,3				78022,3
45	81778,9				81778,9
46	89080,1				89080,1
47	89271,1				89271,1
48	97192,2				97192,2

В 3 столбце просуммировали уровни ряда последовательно за каждый месяц со сдвигом на 1 момент времени и определили условные суммарные объемы выработки за год.

В 4 столбце, разделив полученные суммы на 12, нашли скользящие средние. Полученные таким образом выровненные значения уже не содержат сезонной компоненты.

В 5 столбце привели эти значения в соответствие с фактическими моментами времени, для чего нашли средние значения из двух

последовательных скользящих средних – центрированные скользящие средние.

2 шаг. Найдем оценки сезонной компоненты как разность между фактическими уровнями ряда и центрированными скользящими средними. Используем эти оценки для расчета значений сезонной компоненты S . Для этого найдем средние за каждый месяц по всем годам оценки сезонной компоненты S_i . Теперь для вычисления S_i посчитаем сумм всех полученных ср. оценок сезонной компоненты (ср. ОСК).

Определим корректирующий коэффициент: $k = -9948,68/12 = -829,0571$

Рассчитаем скорректированные значения сезонной компоненты как разность между её средней оценкой и корректирующим коэффициентом k : $S_i = \text{Ср. ОСК} - k$.

Таблица 3-Расчет скорректированного значения сезонной компоненты

год/ месяц	январь	февраль	март	апрель	май	июнь	июль	август	сентябрь	октябрь	ноябрь	декабрь
2005			9658,09	-2768,77	-10435,9	-15959,7	-14128,1	-12346,7	-10617,99	300	5395,27	14197,1
2006	16756	7072,12	8430,17	-2021,2	-8187,86	-14445,5	-12198,7	-9817,27	-9176,125	3155,43	7373,72	12876,8
2007	11010,1	5552,66	5914,28	-3921,51	-8347,38	-15160,3	-13834,3	-11472,2	-9636,875	990,204	7626,25	14808,1
2008	15956,3	7557,69										
итого	43722,4	20182,5	24002,5	-8711,48	-26971,2	-45565,5	-40161,1	-33636,2	-29430,99	4445,63	20395,2	41882,1
ср. оценка сез. комп. (Ssr)	14574,1	6727,49	8000,85	-2903,83	-8990,39	-15188,5	-13387	-11212,1	-9810,329	1481,88	6798,41	13960,7
скорректир. сез. комп. S	15403,2	7556,55	8829,9	-2074,77	-8161,33	-14359,4	-12558	-10383	-8981,272	2310,93	7627,47	14789,7
корректир. коэф. k	-829,057											

В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период взаимопогашаются. В аддитивной модели это выражается в том, что сумма значений сезонной компоненты по всем месяцам должна быть равна 0. В нашем случае так и есть.

3 шаг. Устраним влияние сезонной компоненты, вычитая её значение из каждого уровня ряда, тогда получим величины $T+E=Y-S$. Эти значения рассчитываются за каждый момент времени содержат только тенденцию и случайную компоненту:

Таблица 4 – Устранение влияния сезонной компоненты

у	S	тренд + ошибка прогноза (Т+Е)	t	тренд Т	S+T
(1)	(2)	(3)	(4)	(5)	(6)
91395,7	15403,2	75992,50405	1	78895,2986	94298,49
86617,1	7556,547	79060,55266	2	79088,2092	86644,76
89423	8829,903	80593,09711	3	79281,1198	88111,02
77481,9	-2074,771	79556,67072	4	79474,0305	77399,26
70033	-8161,32	78194,32905	5	79666,9411	71505,61
64712,6	-14359,4	79072,04988	6	79859,8517	65500,4
66854	-12557,9	79411,97211	7	80052,7623	67494,79
68959,8	-10383	79342,80127	8	80245,6729	69862,67
70986,5	-8981,27	79967,77211	9	80438,5835	71457,31
82235,4	2310,933	79924,46655	10	80631,4941	82942,43
87633,1	7627,471	80005,62905	11	80824,4047	88451,88
96738,7	14789,74	81948,95544	12	81017,3153	95807,06
99611,9	15403,2	84208,70405	13	81210,2259	96613,42
90059,3	7556,547	82502,75266	14	81403,1365	88959,68
91218,5	8829,903	82388,59711	15	81596,0471	90425,95
80568	-2074,77	82642,77072	16	81788,9578	79714,19
74381,1	-8161,32	82542,42905	17	81981,8684	73820,54
68150,9	-14359,4	82510,34988	18	82174,779	67815,33
70567,3	-12557,9	83125,27211	19	82367,6896	69809,72
73188,4	-10383	83571,40127	20	82560,6002	72177,6
74016,3	-8981,27	82997,57211	21	82753,5108	73772,24
86496,2	2310,933	84185,26655	22	82946,4214	85257,35
90915,2	7627,471	83287,72905	23	83139,332	90766,8
96607,5	14789,74	81817,75544	24	83332,2426	98121,99
94970,7	15403,2	79567,50405	25	83525,1532	98928,35
89921,2	7556,547	82364,65266	26	83718,0638	91274,61
90871	8829,903	82041,09711	27	83910,9744	92740,88
81573,1	-2074,77	83647,87072	28	84103,885	82029,11
77447,2	-8161,32	85608,52905	29	84296,7957	76135,47
70837	-14359,4	85196,44988	30	84489,7063	70130,26
72363,3	-12557,9	84921,27211	31	84682,6169	72124,64
74952,8	-10383	85335,80127	32	84875,5275	74492,53
77068,7	-8981,27	86049,97211	33	85068,4381	76087,17
87984,6	2310,933	85673,66655	34	85261,3487	87572,28
94944,8	7627,471	87317,32905	35	85454,2593	93081,73

102368,6	14789,74	87578,85544	36	85647,1699	100436,9
103326	15403,2	87922,80405	37	85840,0805	101243,3
94475,3	7556,547	86918,75266	38	86032,9911	93589,54
93516,2	8829,903	84686,29711	39	86225,9017	95055,8
83794,2	-2074,77	85868,97072	40	86418,8123	84344,04
80032,7	-8161,32	88194,02905	41	86611,7229	78450,39
73708,6	-14359,4	88068,04988	42	86804,6336	72445,18
76225,5	-12557,9	88783,47211	43	86997,5442	74439,57
78022,3	-10383	88405,30127	44	87190,4548	76807,45
81778,9	-8981,27	90760,17211	45	87383,3654	78402,09
89080,1	2310,933	86769,16655	46	87576,276	89887,21
89271,1	7627,471	81643,62905	47	87769,1866	95396,66
97192,2	14789,74	82402,45544	48	87962,0972	102751,8

4 Шаг. Определим компоненту Т:

Таблица 5- Определение компоненты Т

тренд, Т:	$T=78702,39-192,91*t$
$T=a+bt$	
$a=(T+E)_{cp.} - b*tc_{p.}$	78702
$b=\{[(T+E)t]_{cp.} - (T+E)_{cp.}*tc_{p.}\} / \{t^2_{cp.} - tc_{p.}^2\}$	192,91

5 Шаг. Найдем значения уровней ряда, полученные по аддитивной модели. Для этого прибавим к уровням Т значения сезонной компоненты для соответствующих месяцев. Результат представлен в 6 столбце таблицы 4. Также для наглядности посмотрим график с расчетными и фактическими данными (Рисунок 3).

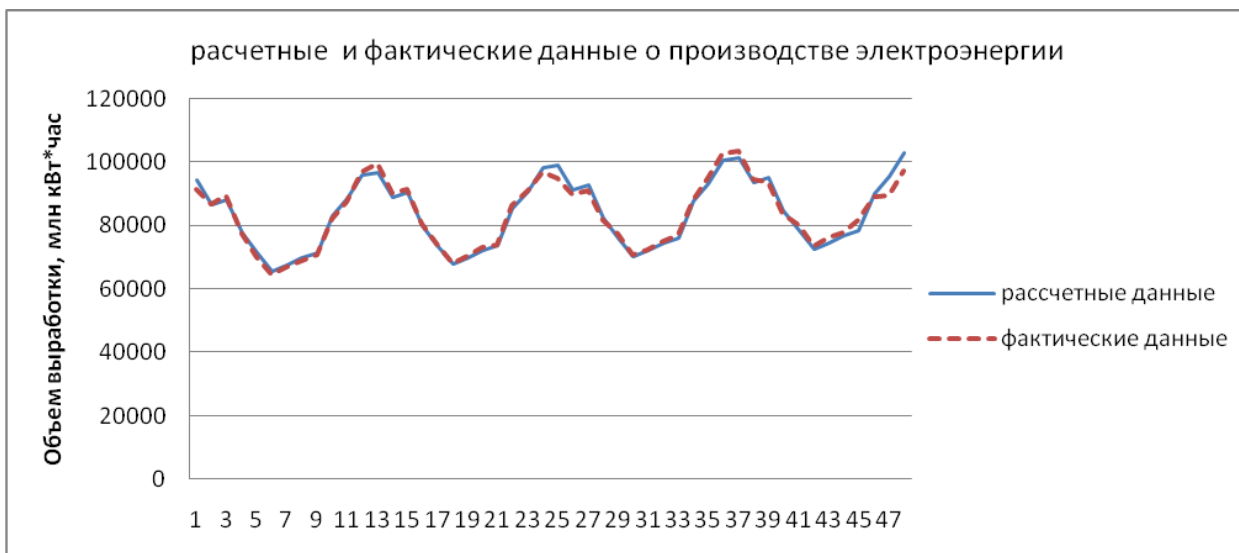


Рисунок 3 –Расчетные и фактические значения временного ряда

По графику можно сделать вывод, что расчетные значения незначительно отличаются от фактических.

6 Шаг. Расчет ошибки (или случайной компоненты) производится по формуле: $E=Y-(T+S)$ - это абсолютная ошибка. Численные значения абсолютных ошибок полученных по этой формуле приведены в таблице 6.

Таблица 6- Расчет абсолютной и среднеквадратической ошибки

t	абсолютная ошибка прогноза($\Delta_{пр}$)	средняя квадратическая ошибка
(1)	(2)	(3)
1	-3,176073471	3,176073471
2	-0,031929698	2,246050056
3	1,467158626	3,498619294
4	0,106657505	3,843729066
5	-2,102740155	2,933423592
6	-1,217385466	3,449969302
7	-0,958491897	3,231043033
8	-1,309272379	4,047695724
9	-0,663240741	3,707308565
10	-0,85976058	3,715901196
11	-0,934322373	3,922328826
12	0,963048006	3,867996765
13	3,010160558	4,925149992

14	1,2209912	5,081954986
15	0,868847837	5,156319498
16	1,059742041	5,983498815
17	0,753633239	5,307405143
18	0,492393965	5,063797999
19	1,073560322	5,199266787
20	1,38109467	4,917689335
21	0,329739959	4,725923734
22	1,432253849	4,934396227
23	0,163225782	4,842352334
24	-1,567670391	4,859133951
25	-4,167231756	6,495512028
26	-1,505107991	6,446965769
27	-2,05772725	6,877635532
28	-0,559025373	7,331493876
29	1,693713135	6,305233915
30	0,997704059	6,476043435
31	0,329801486	6,083928552
32	0,614084861	6,273834547
33	1,27358321	6,233531077
34	0,468625032	6,088012315
35	1,962266235	6,268643491
36	1,886990281	6,628825803
37	2,015681954	6,678545582
38	0,937558853	7,002248467
39	-1,646350711	7,004064632
40	-0,656181001	5,744131454
41	1,977074502	5,57455554
42	1,714069091	5,626408959
45	4,129190695	7,252090065
46	-0,906049089	0,13358965
47	-6,861747578	1,000888752
48	-5,720255083	0,825647703

$$\overline{\Delta}_{np} = \frac{\sum_{i=1}^n |x_i - x_i^*|}{n} = 1,56, \quad \sigma_i = \sqrt{\frac{\sum_{i=1}^n (x_i^* - x_i)^2}{n}} = 2,0556$$

Оба значения ошибок принимают небольшие значения. Но при использовании данной модели на практике полученное значение

среднеквадратической ошибки (СКО) прогноза не является приемлемым. Для возможности применения данной модели для прогнозирования необходимо проверить ее по критерию Дарбина-Уотсона на адекватность модели.

Коэффициент Дарбина-Уотсона принимает значение $d=1,85$, это значение больше 1.33 (нижняя граница для критерия Дарбина-Уотсона для данной задачи), поэтому можно считать, что данная модель является адекватной, т.е. подходит для прогнозирования. Вычислим коэффициент детерминации. Он равен 96,99%.

Выводы по модели:

1. Модель можно считать адекватной (приемлемой для прогноза);
2. Значение средней абсолютной ошибки прогноза (1,56) и СКО прогноза (2,056) является приемлемыми;
3. Коэффициент детерминации равен 96,99%, т.е. объясняет 97% вариации значений объема выработки электроэнергии.

Прогнозирование по аддитивной модели.

Таблица 7- Вычисление прогнозных значений.

	t	S	T	Прогнозное значение
январь	49	15403,2	88155,01	103558
февраль	50	7556,55	88347,91	95904
март	51	8829,9	88540,83	97371

Мультипликативная модель

Для построения модели используем данные без последнего уровня ряда.

1 Шаг. Проведем выравнивание исходных уровней ряда методом скользящей средней, методика, применяемая на этом шаге, полностью совпадает с методикой аддитивной модели.

Таблица 8 – Расчет оценки сезонной компоненты

t	y	итого за год	скользящая средняя	центр. скольз. сред	оценка сезонной компоненты S
1	91395,7				
2	86617,1	953070,8	79422,56667		
3	89423	961287	80107,25	79764,91	1,121081963
4	77481,9	964729,2	80394,1	80250,68	0,965498421
5	70033	966524,7	80543,725	80468,91	0,870311252
6	64712,6	969610,8	80800,9	80672,31	0,802166171
7	66854	973958,9	81163,24167	80982,07	0,825540756
8	68959,8	977397,2	81449,76667	81306,5	0,848146169
9	70986,5	981110,5	81759,20833	81604,49	0,86988476
10	82235,4	985339,1	82111,59167	81935,4	1,003661421
11	87633,1	988368,9	82364,075	82237,83	1,065605652
12	96738,7	992629,7	82719,14167	82541,61	1,171999213
13	99611,9	995911,8	82992,65	82855,9	1,20223068
14	90059,3	995780,6	82981,71667	82987,18	1,085219384
15	91218,5	991139,4	82594,95	82788,33	1,101827955
16	80568	991001,3	82583,44167	82589,2	0,975527116
17	74381,1	990653,8	82554,48333	82568,96	0,900836074
18	68150,9	991658,9	82638,24167	82596,36	0,825107764
19	70567,3	994725	82893,75	82766	0,852612227
20	73188,4	997411,1	83117,59167	83005,67	0,881727709
21	74016,3	999207,1	83267,25833	83192,43	0,889699994
22	86496,2	1000971,5	83414,29167	83340,78	1,037861719
23	90915,2	1004023,9	83668,65833	83541,48	1,088264242
24	96607,5	1005512,3	83792,69167	83730,68	1,15378862
25	94970,7	1009541,9	84128,49167	83960,59	1,131134239

26	89921,2	1015303	84608,58333	84368,54	1,065814374
27	90871	1023658,3	85304,85833	84956,72	1,069615201
28	81573,1	1028212,4	85684,36667	85494,61	0,954131466
30	70837	1033078,7	86089,89167	85997,35	0,823711468
31	72363,3	1035664,2	86305,35	86197,62	0,839504609
32	74952,8	1038535,8	86544,65	86425	0,867258316
33	77068,7	1042398	86866,5	86705,58	0,888855186
34	87984,6	1045467,5	87122,29167	86994,4	1,01138239
35	94944,8	1050177,7	87514,80833	87318,55	1,087338257
36	102368,6	1051273,2	87606,1	87560,45	1,169119107
37	103326	1045599,5	87133,29167	87369,7	1,182629732
38	94475,3	1040423,1	86701,925	86917,61	1,086952366
39	93516,2				
40	83794,2				
41	80032,7				
42	73708,6				
43	76225,5				
44	78022,3				
45	81778,9				
46	89080,1				
47	89271,1				
48	97192,2				

2 Шаг. Найдем оценки сезонной компоненты (ОСК) как частное от деления фактических уровней ряда на центрированные скользящие средние. Используем эти оценки для расчета значений сезонной компоненты S . Находим значения сезонной компоненты: средние за каждый год оценки сезонной компоненты. Результат вычисления представлен на рисунке 4.

год/	месяц	январь	февраль	март	апрель	май	июнь	июль	август	сентябрь	октябрь	ноябрь	декабрь
	2005			1,121082	0,965498	0,870311	0,802166	0,825541	0,848146	0,869885	1,003661	1,065605652	1,171999
	2006	1,202231	1,085219	1,101828	0,975527	0,900836	0,825108	0,852612	0,881728	0,8897	1,037862	1,088264242	1,153789
	2007	1,131134	1,065814	1,069615	0,954131	0,902705	0,823711	0,839505	0,867258	0,888855	1,011382	1,087338257	1,169119
	2008	1,18263	1,086952										
итого		3,515995	3,237986	3,292525	2,895157	2,673852	2,450985	2,517658	2,597132	2,64844	3,052906	3,241208152	3,494907
ср.оценка сез.комп.(Ssr)		1,171998	1,079329	1,097508	0,965052	0,891284	0,816995	0,839219	0,865711	0,882813	1,017635	1,080402717	1,164969
скорректир.сез. комп. S		1,159586	1,067898	1,085885	0,954832	0,881845	0,808343	0,830332	0,856543	0,873464	1,006858	1,068960982	1,152632
корректир. коэф. k		0,98941											

Рисунок 4 – Оценка сезонной компоненты

3 Шаг. Разделим каждый уровень исходного ряда на соответствующие значения сезонной компоненты. Далее находим среднюю за каждый квартал оценку сезонной компоненты.

4 Шаг. Определим компоненту Т в мультипликативной модели.

Таблица 9 – Определение тренда в мультипликативной модели

$Y=T*S*E$	
тренд, $T=a+bt$	$80264,609+205,863*t$
$a=(T*E)_{ср.} - b*t_{ср.}$	80264,609
$b=\{[(T*E)t]_{ср.} - (T*E)_{ср.}*t_{ср.}\} / \{t^2_{ср.} - t_{ср.}^2\}$	205,863

5 Шаг. Найдем значения уровней ряда, полученные по мультипликативной модели. Для этого умножим уровни Т на значения сезонной компоненты для соответствующих месяцев S.

Таблица 10- Получение расчетных значений объема выработки

t	Yt	S	тренд* ошибка прогноза (T*E)=Y/S	тренд, T	S*T	ошибка прогноза $E=Y/(T*S)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	91395,7	1,1595864	78817,49481	80470,4	93312,4	0,9794585
2	86617,1	1,0678983	81109,87367	80676,3	86154,1	1,00537381

3	89423	1,0858854	82350,30424	80882,2	87828,8	1,01815118
4	77481,9	0,9548321	81147,13866	81088,0	77425,4	1,00072858
5	70033	0,8818451	79416,43369	81293,9	71688,6	0,97690493
6	64712,6	0,8083429	80055,8722	81499,7	65879,7	0,98228323
7	66854	0,8303316	80514,8154	81705,6	67842,7	0,98542534
8	68959,8	0,8565426	80509,47722	81911,5	70160,7	0,98288357
9	70986,5	0,8734640	81270,0832	82117,3	71726,5	0,98968197
10	82235,4	1,0068581	81675,25758	82323,2	82887,8	0,99212886
11	87633,1	1,0689609	81979,69943	82529,1	88220,3	0,99334298
12	96738,7	1,1526316	83928,54611	82734,9	95362,9	1,01442662
13	99611,9	1,1595864	85902,95179	82940,8	96177,0	1,0357137
14	90059,3	1,0678983	84333,21418	83146,6	88792,2	1,01427031
15	91218,5	1,0858854	84003,79352	83352,5	90511,3	1,00781315
16	80568	0,9548321	84379,22493	83558,4	79784,2	1,00982324
17	74381,1	0,8818451	84347,11773	83764,2	73867,1	1,00695815
18	68150,9	0,8083429	84309,3886	83970,1	67876,6	1,00404016
19	70567,3	0,8303316	84986,88385	84176	69894	1,00963322
20	73188,4	0,8654263	85446,30092	84381,8	72276,6	1,01261456
21	74016,3	0,8734640	84738,80046	84587,7	73884,3	1,00178604
22	86496,2	1,0068581	85907,03535	84793,5	85375,1	1,01313129
23	90915,2	1,0689609	85050,06407	84999,4	90861,0	1,00059548
...
33	77068,7	0,8734640	88233,39171	87058,0	76042,1	1,0135003
34	87984,6	1,0068581	87385,29718	87263,94	87862,4	1,0013907
35	94944,8	1,0689609	88819,70587	87469,8	93501,8	1,0154328
36	102368,	1,1526316	88812,93386	87675,6	101057	1,0129713
37	103326	1,1595864	89105,90398	87881,5	101906,	1,0139321
38	94475,3	1,0678983	88468,43924	88087,3	94068,3	1,0043258
39	93516,2	1,0858854	86119,7625	88293,2	95876,3	0,9753833
40	83794,2	0,9548321	87758,03854	88499,1	84501,8	0,9916262
41	80032,7	0,8818451	90755,95238	88704,9	78224,0	1,0231213
42	73708,6	0,8083429	91184,81194	88910,8	71870,4	1,0255759
43	76225,5	0,8303316	91801,26936	89116,7	73996,4	1,0301242
44	78022,3	0,8565426	91089,80282	89322,5	76508,5	1,0197849
45	81778,9	0,8734640	93625,94306	89528,4	78199,8	1,0457677
46	89080,1	1,0068581	88473,33524	89734,2	90349,7	0,9859479
47	89271,1	1,0689609	83512,02852	89940,1	96142,5	0,9285289
48	97192,2	1,1526316	84321,99356	90146,0	103905,	0,9353934

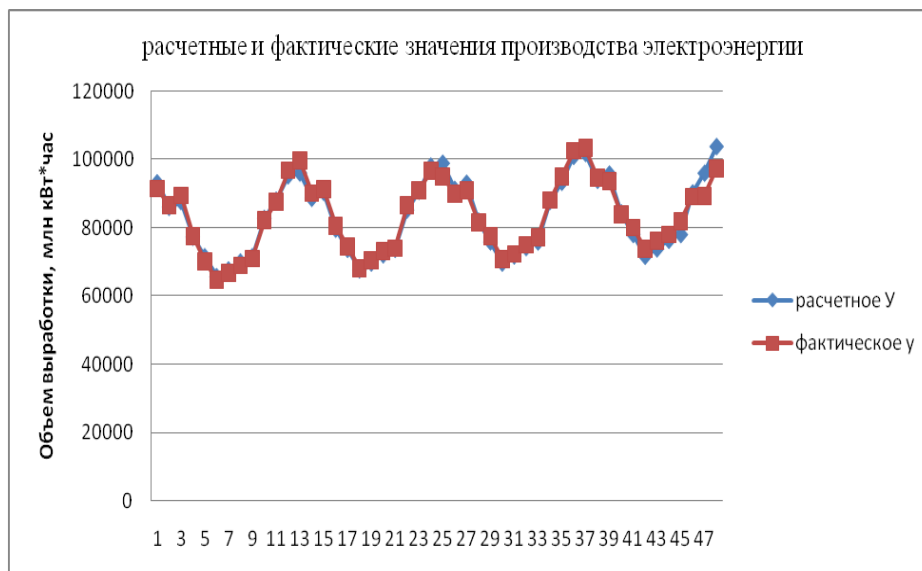


Рисунок 5 - Расчетные и фактические значения временного ряда

По графику можно сделать вывод, что значения расчетного производства электроэнергии и фактического практически идентичны.

6 Шаг. Расчет ошибки (или случайной компоненты) производится по формуле: $E = Y_t - (T \cdot S)$ Это абсолютная ошибка, численные значения абсолютных ошибок полученных по этой формуле приведены в таблице 10.

Для того чтобы сравнить мультипликативную модель и другие модели временного ряда, можно по аналогии с аддитивной моделью использовать сумму квадратов абсолютных ошибок. Абсолютные ошибки определяются по формуле: $E_0 = y_t - (T \cdot S)$.

Таблица 11- Расчет абсолютной и среднеквадратической ошибки

t	Δ_{np}	σ_t
1	-2,097220768	3,176073471
2	0,534508826	2,246050056
3	1,782758976	3,498619294
4	0,072805153	3,843729066
5	-2,364105447	2,933423592
6	-1,80363071	3,449969302
7	-1,479021774	3,231043033
8	-1,741450135	4,047695724
9	-1,042559513	3,707308565
10	-0,79335797	3,715901196
11	-0,670163316	3,922328826
12	1,422145734	3,867996765
13	3,44822768	4,925149992
14	1,406953596	5,081954986
15	0,7752578	5,156319498

16	0,972768842	5,983498815
17	0,691007669	5,307405143
18	0,402390952	5,063797999
19	0,954130812	5,199266787
20	1,245742111	4,917689335
21	0,17828613	4,725923734
22	1,296110203	4,934396227
23	0,059513201	4,842352334
...
33	1,332055818	3,233531077
34	0,138880267	3,088012315
35	1,519828342	3,268643491
36	1,280525751	3,628825803
37	1,374072863	2,678545582
38	0,430721672	2,002248467
39	-2,523795353	2,004064632
40	-0,844450367	2,744131454
41	2,259882838	3,57455554
42	2,493809161	3,626408959
43	2,924327804	3,870260154
44	1,940108429	2,12795849
45	4,376477805	2,252090065
46	-1,425233671	1,13358965
47	-7,697239741	1,000888752
48	-6,906879587	1,825647703

$$\overline{\Delta}_{np} = \frac{\sum_{i=1}^n |x_i - x_i^*|}{n} = 1,71 \quad \sigma_i = \sqrt{\frac{\sum_{i=1}^n (x_i^* - x_i)^2}{n}} = 2,28$$

Оба значения ошибок принимают небольшие значения. Но при использовании данной модели на практике полученное значение среднеквадратической ошибки (СКО) прогноза не является приемлемым. Для возможности применения данной модели для прогнозирования необходимо проверить ее по критерию Дарбина-Уотсона на адекватность модели.

Коэффициент Дарбина-Уотсона принимает значение $d=1,89$, это значение больше 1.33 (нижняя граница для критерия Дарбина-Уотсона для данной задачи), поэтому можно считать, что данная модель является адекватной, т.е. подходит для прогнозирования. Вычислим коэффициент детерминации. Он равен 99,1%.

Выводы по модели:

1. Модель можно считать адекватной (приемлемой для прогноза);

2. Значение средней абсолютной ошибки прогноза (1,71) и СКО прогноза (2,28) является приемлемыми;
3. Коэффициент детерминации равен 99,1%, т.е. объясняет 99% вариации значений объема выработки электроэнергии.

Прогноз на первые 3 три месяца 2009года:

Таблица 12 – Прогноз объема выработки электроэнергии

t	месяцы	S	тренд	прогноз
49	январь	1,159586463	90351,87469	104770,8108
50	февраль	1,067898347	90557,73725	96706,45789
51	март	1,085885484	90763,59981	98558,87553

Применение фиктивных переменных для моделирования сезонных колебаний

Имеется временной ряд, который содержит циклические колебания с периодом 12. Тогда общий вид уравнения регрессии:

$$Y_t = a + bt + c_1 * x_1 + c_2 * x_2 + c_3 * x_3 + c_4 * x_4 + c_5 * x_5 + c_6 * x_6 + c_7 * x_7 + c_8 * x_8 + c_9 * x_9 + c_{10} * x_{10} + c_{11} * x_{11} + c_{12} * x_{12} + \varepsilon_t.$$

где $x_j = 1$ для каждого j внутри каждого цикла и нулю во всех остальных случаях.

- 1) Уравнение тренда для каждого квартала будет иметь следующий вид: $Y_t = a + bt + c_1 + \varepsilon_t$

- 2) $Y_t = a + bt + c_2 + \varepsilon_t$

- 3) $Y_t = a + bt + c_3 + \varepsilon_t$

- 4) $Y_t = a + bt + c_4 + \varepsilon_t$

- 5) $Y_t = a + bt + c_5 + \varepsilon_t$

- 6) $Y_t = a + bt + c_6 + \varepsilon_t$

- 7) $Y_t = a + bt + c_7 + \varepsilon_t$

- 8) $Y_t = a + bt + c_8 + \varepsilon_t$

- 9) $Y_t = a + bt + c_9 + \varepsilon_t$

- 10) $Y_t = a + bt + c_{10} + \varepsilon_t$

$$11) Y_t = a + bt + c11 + \varepsilon_t$$

$$12) Y_t = a + bt + \varepsilon_t$$

Т.о. фиктивная переменная позволяет дифференцировать величину свободного члена уравнения для каждого квартала. Матрица исходных данных представлена на рисунке 6:

t	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11
1	1	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0
10	0	0	0	0	0	0	0	0	0	1	0
11	0	0	0	0	0	0	0	0	0	0	1
12	0	0	0	0	0	0	0	0	0	0	0
13	1	0	0	0	0	0	0	0	0	0	0
14	0	1	0	0	0	0	0	0	0	0	0
15	0	0	1	0	0	0	0	0	0	0	0
16	0	0	0	1	0	0	0	0	0	0	0
17	0	0	0	0	1	0	0	0	0	0	0
18	0	0	0	0	0	1	0	0	0	0	0
19	0	0	0	0	0	0	1	0	0	0	0
20	0	0	0	0	0	0	0	1	0	0	0
21	0	0	0	0	0	0	0	0	1	0	0
22	0	0	0	0	0	0	0	0	0	1	0
23	0	0	0	0	0	0	0	0	0	0	1
24	0	0	0	0	0	0	0	0	0	0	0
25	1	0	0	0	0	0	0	0	0	0	0
26	0	1	0	0	0	0	0	0	0	0	0
27	0	0	1	0	0	0	0	0	0	0	0
28	0	0	0	1	0	0	0	0	0	0	0
29	0	0	0	0	1	0	0	0	0	0	0
30	0	0	0	0	0	1	0	0	0	0	0
31	0	0	0	0	0	0	1	0	0	0	0
32	0	0	0	0	0	0	0	1	0	0	0
33	0	0	0	0	0	0	0	0	1	0	0
34	0	0	0	0	0	0	0	0	0	1	0
35	0	0	0	0	0	0	0	0	0	0	1
36	0	0	0	0	0	0	0	0	0	0	0
37	1	0	0	0	0	0	0	0	0	0	0
38	0	1	0	0	0	0	0	0	0	0	0
39	0	0	1	0	0	0	0	0	0	0	0
40	0	0	0	1	0	0	0	0	0	0	0
41	0	0	0	0	1	0	0	0	0	0	0
42	0	0	0	0	0	1	0	0	0	0	0
43	0	0	0	0	0	0	1	0	0	0	0
44	0	0	0	0	0	0	0	1	0	0	0
45	0	0	0	0	0	0	0	0	1	0	0
46	0	0	0	0	0	0	0	0	0	1	0
47	0	0	0	0	0	0	0	0	0	0	1
48	0	0	0	0	0	0	0	0	0	0	0

Рисунок 6 – Матрица фиктивных переменных

Применив инструмент регрессии пакета анализа MS Excel, выделив, исходные данные, получаем модель (рис. 7):

$$Y = 98226,75t - 900,675x_1 - 79558,52x_2 - 6969,575x_3 - 17372x_4 - 22753,25x_5 - 28874,475x_6 - 26724,23x_7 - 24445,93x_8 - 11777,68x_9 - 7535,7x_{10} - 7535,7x_{11}$$

$$R^2 = 90,6.$$

	Коэффициенты	Стандартная ошибка t-статистик	P-значения	Нижние 95	Верхние 95	Нижние 95	Верхние 95,0%
Y-пересечение	98226,75	1813,71191	54,15786	4,26E-36	94548,37	101905,1	94548,37 101905,1
Переменная X 1	-900,675	2564,975982	-0,35114	0,727527	-6102,69	4301,337	-6102,69 4301,337
Переменная X 2	-7958,525	2564,975982	-3,10277	0,003719	-13160,5	-2756,51	-13160,5 -2756,51
Переменная X 3	-6969,575	2564,975982	-2,71721	0,010057	-12171,6	-1767,56	-12171,6 -1767,56
Переменная X 4	-17372,45	2564,975982	-6,77295	6,53E-08	-22574,5	-12170,4	-22574,5 -12170,4
Переменная X 5	-22753,25	2564,975982	-8,87075	1,38E-10	-27955,3	-17551,2	-27955,3 -17551,2
Переменная X 6	-28874,475	2564,975982	-11,2572	2,39E-13	-34076,5	-23672,5	-34076,5 -23672,5
Переменная X 7	-26724,225	2564,975982	-10,4189	2,05E-12	-31926,2	-21522,2	-31926,2 -21522,2
Переменная X 8	-24445,925	2564,975982	-9,53066	2,21E-11	-29647,9	-19243,9	-29647,9 -19243,9
Переменная X 9	-22264,15	2564,975982	-8,68006	2,36E-10	-27466,2	-17062,1	-27466,2 -17062,1
Переменная X 10	-11777,675	2564,975982	-4,59173	5,19E-05	-16979,7	-6575,66	-16979,7 -6575,66
Переменная X 11	-7535,7	2564,975982	-2,93792	0,005733	-12737,7	-2333,69	-12737,7 -2333,69

Рисунок 7 – Модель сезонных колебаний с фиктивными переменными

Найдем Y расчетное, абсолютную ошибку прогноза:

Таблица 13 – Вычисление расчетных значений объема выработки электроэнергии

t	y	урасч	Δпр
(1)	(2)	(3)	(4)
1	91395,7	97326,075	-6,48868054
2	86617,1	188494,98	-1,176186631
3	89423	287710,68	-2,217412467
4	77481,9	375534,55	-3,846739045
5	70033	468380,5	-5,687997087
6	64712,6	560486,03	-7,661157564
7	66854	660863,03	-8,885168053
8	68959,8	761368,08	-10,04075237
9	70986,5	861776,6	-11,1400069
10	82235,4	970489,83	-1,080136322
11	87633,1	1072958,6	-1,124375892
12	96738,7	1178721	-1,11845859
13	99611,9	1276047,1	-1,181018709
14	90059,3	1367216	-1,418128583
15	91218,5	1466431,7	-1,507603364
16	80568	1554255,6	-1,829122667
17	74381,1	1647101,5	-2,114408633

18	68150,9	1739207	-2,451994214
19	70567,3	1839584	-2,506850517
20	73188,4	1940089,1	-2,550814986
21	74016,3	2040497,6	-2,656821943
22	86496,2	2149210,8	-2,384745948
23	90915,2	2251679,6	-2,376681072
24	96607,5	2357442	-2,34022669
25	94970,7	2454768,1	-2,48476359
26	89921,2	2545937	-2,73129782
27	90871	2645152,7	-2,810887604
28	81573,1	2732976,6	-3,250340431
29	77447,2	2825822,5	-3,54870841
30	70837	2917928	-4,01921457
31	72363,3	3018305	-4,07104392
32	74952,8	3118810,1	-4,06103211
33	77068,7	3219218,6	-4,077076556
34	87984,6	3327931,8	-3,682402631
35	94944,8	3430400,6	-3,513047318
36	102368,6	3536163	-3,354343422
37	103326	3633489,1	-3,41652931
38	94475,3	3724658	-3,842467476
39	93516,2	3823873,7	-3,988995997
40	83794,2	3911697,6	-4,568219936
41	80032,7	4004543,5	-4,90363414
42	73708,6	4096649	-5,457898298
43	76225,5	4197026	-540,6065588
44	78022,3	4297531,1	-540,8080478
45	81778,9	4397939,6	-527,7841473
46	89080,1	4506652,8	-495,9101668
47	89271,1	4609121,6	-506,3061226

$$\Delta_{np} = \frac{\sum_{i=1}^n |x_i - x_i^*|}{n} = 2,71 \quad \sigma_i = \sqrt{\frac{\sum_{i=1}^n (x_i^* - x_i)^2}{n}} = 3,75$$

Оба значения ошибок принимают небольшие значения, но не являются приемлемыми. Для возможности применения данной модели для прогнозирования необходимо проверить ее по критерию Дарбина-Уотсона на адекватность модели. Коэффициент Дарбина-Уотсона принимает значение $d=1,743$. Данная модель является адекватной, т.е. подходит для прогнозирования. Вычислим коэффициент детерминации. Он равен 93,6%.

Выводы по модели:

1. Модель можно считать адекватной (приемлемой для прогноза);
2. Значение средней абсолютной ошибки прогноза (2,41) и СКО прогноза (3,75) не являются приемлемыми;
3. Коэффициент детерминации равен 93,6%, т.е. объясняет 94% вариации значений объема выработки электроэнергии.

Прогноз на 3 месяца осуществим по формуле $Y=98226,75t-900,675x_1 - 79558,52x_2 - 6969,575x_3 - 17372x_4-22753,25x_5-28874,475x_6-26724,23x_7-24445,93x_8-11777,68x_{10}-7535,7x_{11}$

Таблица 14 – Прогноз объема выработки электроэнергии

t	месяцы	прогноз
49	январь	101770,2
50	февраль	97816,12
51	март	99015,36

Выбор лучшей модели

Сведем все интересующие нас результаты в единую таблицу (табл.15).

Таблица 15 - Результаты вычислений по выбранным моделям

Прогноз	Аддитивная модель	Мультипликативная модель	Регрессия с включением фиктивных переменных
январь	103558	104770,8108	101770,2
февраль	95904	96706,45789	97816,12
март	97371	98558,87553	99015,36
средняя абсолютная ошибка	1,56	1,71	2,41
среднеквадратическая ошибка	2,06	2,28	3,75
коэффициент детерминации, %	96,99	99,1	93,6
статистика Дарбина-Уотсона	1,85	1,89	1,743

Полученные результаты позволяют сделать вывод, что лучшей является мультипликативная модель, у нее самое высокое значение коэффициента детерминации (98,1%), хотя ср. абсолютная и среднеквадратическая ошибка чуть больше, чем у аддитивной модели. Также значение статистики Дарбина-Уотсона для мультипликативной модели 1,89 наиболее близко к 2, а следовательно эта модель наиболее адекватно отражает данные и дает более точный прогноз для объема выработки электроэнергии.

Задание

1. Собрать данные не менее чем за 4 года. Данные необходимо выбрать по месяцам, то есть временной ряд будет состоять минимум из 48 наблюдений. Также необходимо взять данные таким образом, чтобы вы смогли сравнить прогнозные значения с фактическими данными. Для этого можно взять ретроспективу за 5 лет, и строить модель на 4 года, а прогнозные значения построить для части месяцев 5-го года.
2. Построить прогнозирование с помощью следующих моделей: аддитивная модель, мультипликативная модель, моделирование с помощью фиктивных переменных для прогноза временного ряда. Оценить качество прогнозной модели. На данном этапе можно использовать MS Excel (как в методических указаниях) или язык для статистической обработки данных R/ Python для анализа данных.
3. Выбрать наилучшую модель.