|  |  |
| --- | --- |
| lu135925on3bu_tmp_3360867a00ce4d37 | **Министерство науки и высшего образования Российской Федерации Федеральное государственное бюджетное образовательное учреждение высшего образования** **«Московский государственный технический университет имени Н.Э. Баумана** **(национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ Информатика и системы управления и искусственный интеллект

КАФЕДРА                  Системы обработки информации и управления

**Лабораторная работа №3**

**По курсу**

**«Анализ временных рядов»**

**«Адаптивные модели экспоненциального сглаживания**

**нестационарных временных рядов»**

Подготовил:

Студент группы

**ИУ5-34Б Журавлев Н.В**

12.10.2024

Проверил:

**Лабунец Л.В.**

*2024 г*.

**1 Цели работы**

1.1 Изучение моделей экспоненциального сглаживания НВР и методики выбора параметров этих моделей.

1.2 Приобретение навыков моделирования и прогнозирования НВР с помощью процедур адаптивного экспоненциального сглаживания в пакете«STATISTICA».

**2 Задачи работы**

2.1 Освоение методики выбора параметров экспоненциальной скользящей средней (Exponential Moving Average - EMA) в пакете«STATISTICA». Анализ и компенсация запаздывания EMA.

2.2 Изучение адаптивной модели Хольта - Брауна линейного темпа изменения нестационарного процесса и формирование оценки тренда НВР.

2.3 Изучение адаптивной модели Тейла – Вейджа моделирования аддитивной сезонной компоненты нестационарного процесса.

2.4 Изучение адаптивной модели Уинтерса моделирования мультипликативной сезонной компоненты нестационарного процесса.

2.5 Освоение методики прогнозирования сезонных компонент НВР с помощью адаптивных моделей Тейла – Вейджа и Уинтерса.

**3 Теоретическая часть**

На практике исследователь нередко сталкивается с проблемой переобучения модели данных. Такого рода нежелательный эффект возникает, когда модель обладает чрезмерно большим количеством параметром. В этом случае часть параметров настраивается по объективно существующей закономерности, скрытой в данных, а оставшаяся часть избыточных параметров настраивается по ошибкам измерений или вычислений. В результате незначительное изменение исходных данных приводит к катастрофически большому изменению параметров модели. Иными словами, модель становится некорректной. Решение указанной выше проблемы основано на применении методологии теории регуляризации Тихонова – Филипса.

**3.1 Простая экспоненциальная скользящая средняя**

Показательным примером практической реализации теории регуляризации является семейство моделей экспоненциального сглаживания данных. В частности проанализируем смысл целевой функции EMA [11]

. (4.1)



Первое слагаемое целевой функции представляет собой квадратичную меру близости исходного ряда и оценки его тренда. Второе слагаемое – это штраф за сложность (гладкость) модели тренда. Таким образом, EMA удовлетворяет фундаментальному принципу регуляризации, а именно в классе моделей гарантирующих заданную точность выбрать наиболее простую.



Положительные константы и рационально трактовать как нормированные веса штрафов за точность аппроксимации данных и сложность модели соответственно. Константу принято называть параметром сглаживания EMA. Ясно, что этот параметр может принимать значения из интервала . При EMA игнорирует ошибку аппроксимации данных и формирует максимально гладкую оценку тренда. При EMA игнорирует гладкость модели и формирует максимально точное описание данных, т.е. модель краткосрочного прогноза [15].



Решением задачи квадратичного программирования (1) является линейное разностное уравнение первого порядка

. (2)



Это уравнение описывает рекурсивный ЦФ (БИХ- фильтр) с импульсной характеристикой для всех и ноль в противном случае. В качестве начального значения экспоненциальной скользящей средней обычно выбираю величину НВР усредненную на некотором начальном участке ряда



.



Для значений уравнение EMA (4.2) удобно представить в прогнозной форме



,



где - локально постоянный («наивный») прогноз НВР на один шаг времени; - ошибка прогноза.



Временной интервал запаздывания *L* EMA относительно исходного ВР удобно рассчитывать с помощью методики Эйлерса (Ehlers). Уравнению (2) скользящей средней соответствует следующее алгебраическое уравнение в терминах запаздывания отсчетов ряда

.



Решение этого уравнения дает оценку запаздывания . На практике предпочитают контролировать «эффективный» период сглаживания EMA



; (3)



Типичные значения параметра и соответствующие ему величины запаздывания и эффективного периода сглаживания EMA сведены в таблицу 4.1.



Таблица 4.1. Параметры EMA.

|  |  |  |  |
| --- | --- | --- | --- |
| Тип модели |  | *L* | *K* |
| Тренд | 0,05 | 19 | 39 |
| Краткосрочный прогноз | 0,5 | 1 | 3 |

**3.2 Модель Хольта локально линейного тренда НВР**

Локально постоянный прогноз НВР , реализованный в модели простой экспоненциальной средней, очевидно является грубым. Достаточно адекватное описание динамики данных в среднем возможно с помощью локально линейного прогноза НВР , где и - оценки среднего значения процесса и его темпа изменения за один шаг времени для предшествующего отсчета ряда. Очевидно, что эти параметры необходимо пересчитывать по мере поступления новых нестационарных данных. Простую и, вместе с тем, эффективную процедуру адаптации параметров локально линейного прогноза к динамике НВР предложил Хольт (Holt). Процедура основана на модели экспоненциального сглаживания и имеет вид



;



.



Параметры сглаживания удовлетворяют неравенству , поскольку процесс имеет низкочастотный спектр по сравнению со спектром процесса , т.е. обладает б*о*льшими характерными периодами сглаживания *K* (см. формулу (4.3)).



С целью экономии вычислительных затрат модель Хольта линейного темпа изменения НВР рационально формулировать в терминах ошибки линейного прогноза процесса на один шаг времени [15]

. (4)



Алгоритм адаптации параметров модели в этом случае принимает вид

;



. (5)



В качестве начальных значений среднего уровня и темпа изменения процесса обычно выбираю параметры модели линейной регрессии НВР для некоторого начального участка ряда



;



.



Определенные сложности, связанные с применением модели Хольта, обусловлены необходимостью выбора двух параметров сглаживания и . В ряде практических случаев количество параметров модели удается сократить. В частности, Браун (Brown) предложил применять коэффициент дисконтирования (старения) информации и установил его взаимосвязь с параметрами и для некоторых НВР



; .



Эти зависимости позволяют оперировать в формуле (4.5) единственным параметром сглаживания с учетом приближенного равенства . В итоге, алгоритм адаптации параметров локально линейного тренда Хольта – Брауна принимает следующий вид



***Шаг* 0:** Выбрать параметры сглаживания и начальные значения ; ; ; .



***Шаг* 1:** Вычислить ошибку прогноза НВР на один шаг времени

.



***Шаг* 2:** Корректировать среднее значение НВР за один шаг времени

.



***Шаг* 3:** Корректировать темп изменения НВР за один шаг времени



или для упрощенной модели

.



***Шаг* 4:** Вычислить прогноз НВР на один шаг времени

.



***Шаг* 5:** Цикл по времени *n* = *n* + 1. Идти к ***Шагу* 1**.

В теории адаптивного экспоненциального сглаживания НВР известны более сложные локально квадратичные модели тренда, однако их анализ выходит за рамки данных практических занятий.

**3.3 Сезонные модели адаптивного экспоненциального сглаживания**

Достаточно часто НВР, помимо трендовой компоненты, содержит сезонную составляющую. В этом случае прогноз ряда на один шаг времени рационально описывать следующими моделями [15]

;



соответственно для аддитивного и мультипликативного сезонных эффектов. Здесь - локально линейный прогноз процесса в среднем; и - оценки сезонной (квазипериодической) компоненты в момент времени с характерным периодом *T*. Ясно, что значения указанных выше структурных временных рядов необходимо пересчитывать по мере поступления новых нестационарных данных. Адаптация процессов , и , к динамике НВР основана на модели экспоненциального сглаживания. Процедуру обновления параметров локально линейного тренда и аддитивного сезонного цикла предложили Тейл и Вейдж (Theil and Wage - TV):



;



;



.



Аналогичная процедура для мультипликативного характера поведения данных была предложенаУинтерсом (Winters - W):

;



;



.



Спектры процессов , и , локализованы в интервалах низких, средних и высоких частот соответственно. Поэтому параметры сглаживания удовлетворяют неравенству (см. формулу (3)).



Для экономии вычислительных затрат указанные выше алгоритмы моделирования и прогнозирования тренда и квазипериодических составляющих НВР рационально формулировать в терминах ошибки прогноза ряда (4.4) на один шаг времени.

***Шаг* 0:** Выбрать параметры сглаживания и начальные значения



или .



***Шаг* 1:** Вычислить прогноз рынка на один шаг времени



или

.



***Шаг* 2:** Вычислить ошибку прогноза рынка на один шаг времени



или

.



***Шаг* 3:** Корректировать среднее значение рынка за один шаг времени



или

.



***Шаг* 4:** Корректировать темп изменения рынка за один шаг времени



или

.



***Шаг* 5:** Корректировать сезонный цикл за один шаг времени



или

.



***Шаг* 6:** Цикл по времени *n* = *n* + 1. Идти к ***Шагу* 1**.

Наряду с рассмотренными выше локально линейными моделями трендов на практике широко применяют локально квадратичные, экспоненциальные и гиперболические модели в сочетании с квазипериодическими компонентами аддитивного и мультипликативного вида [15, с.59]. Однако, из анализ и практическое освоение выходи за рамки данных практических занятий.

**4 Практическая часть**

Построим график начальный данных.

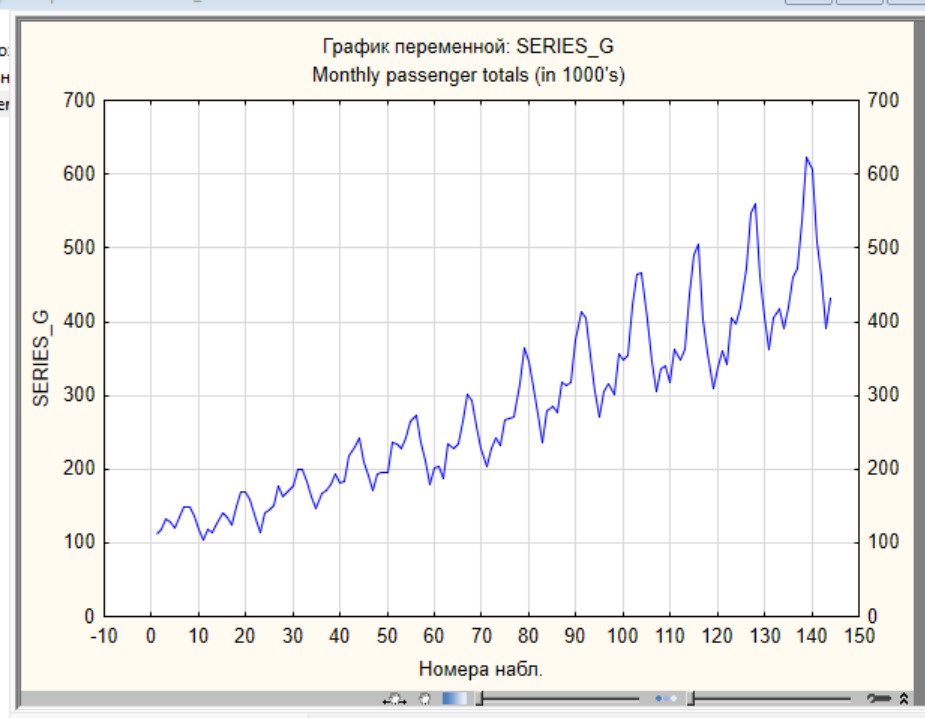


Рисунок 1 – Начальные данные

Построим линейную модель на участке от 0 до 24, что нужно указать в поле range.

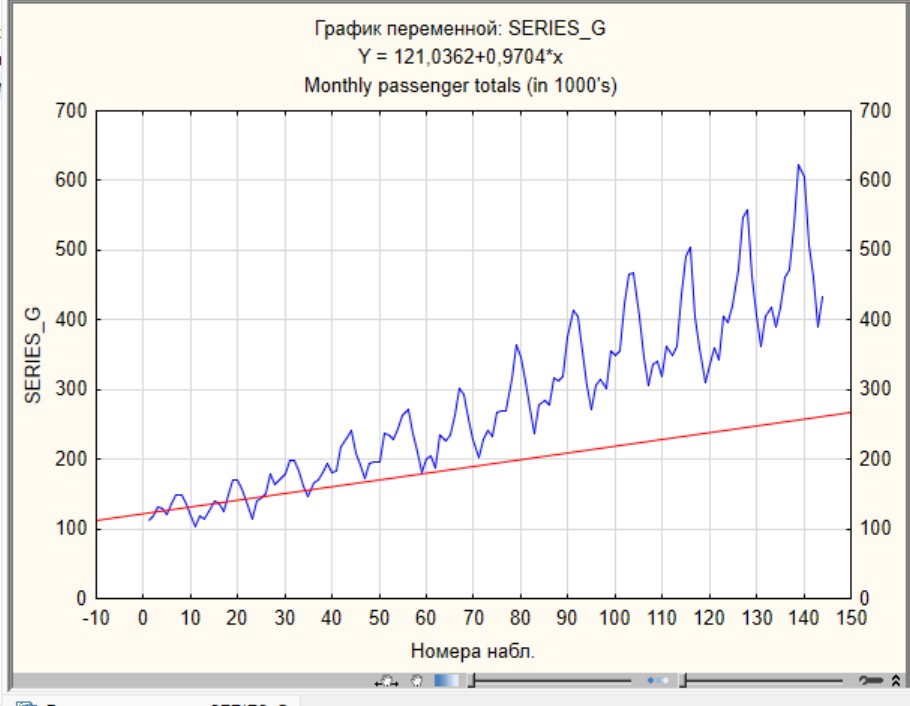


Рисунок 2 – Линейная модель на графики начальных значений

Выбираем исходные данные и выбираем кнопку Exponential smoothing & forecasting и переходим на вкладку дополнительно. На панели выбираем коэффициент альфа равный 0,3, начальное значение 121.

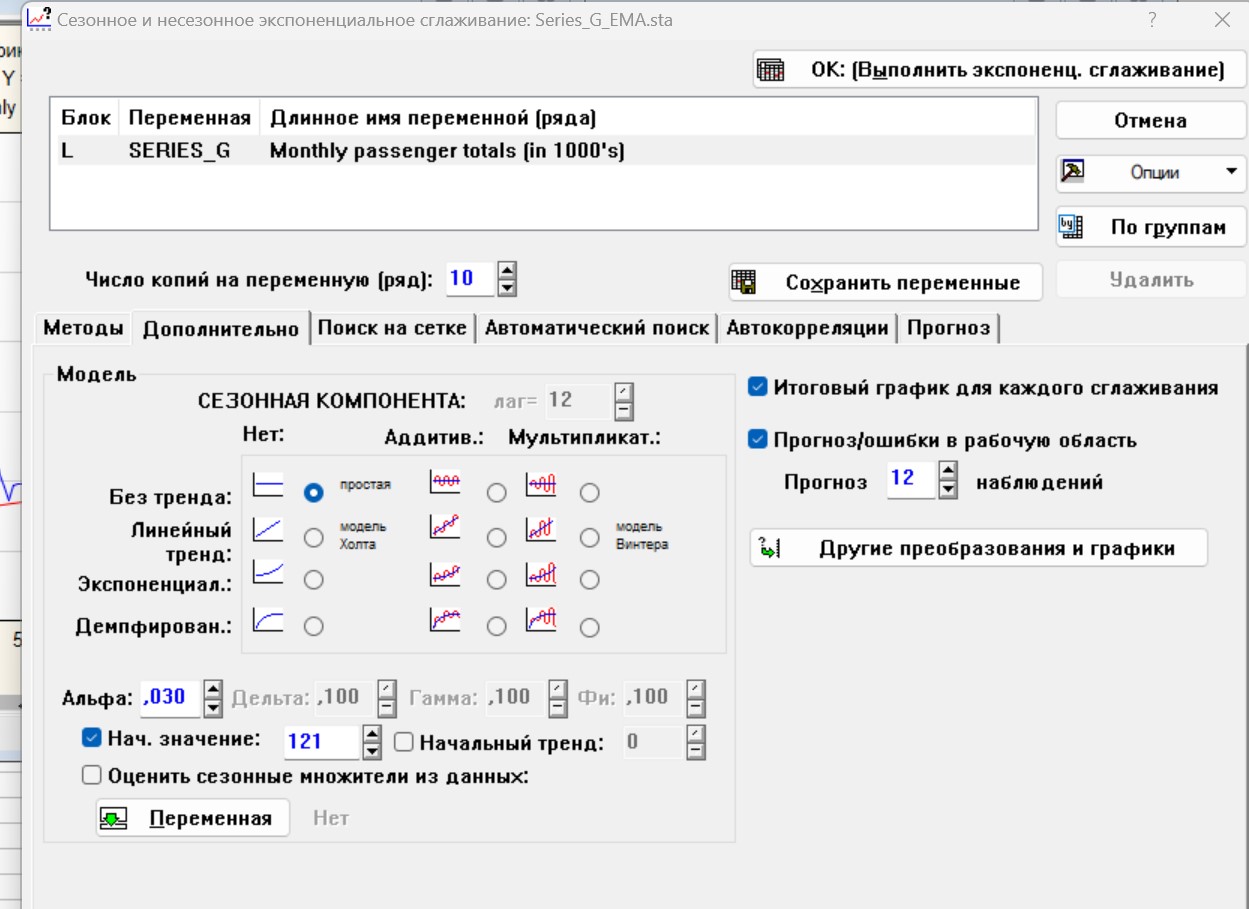


Рисунок 3 – Панель с заданными параметрами

После нажатия получаем модель, которая представляет собой реакцию фильтра, смещённого вправо на 32 и наивный прогноз.

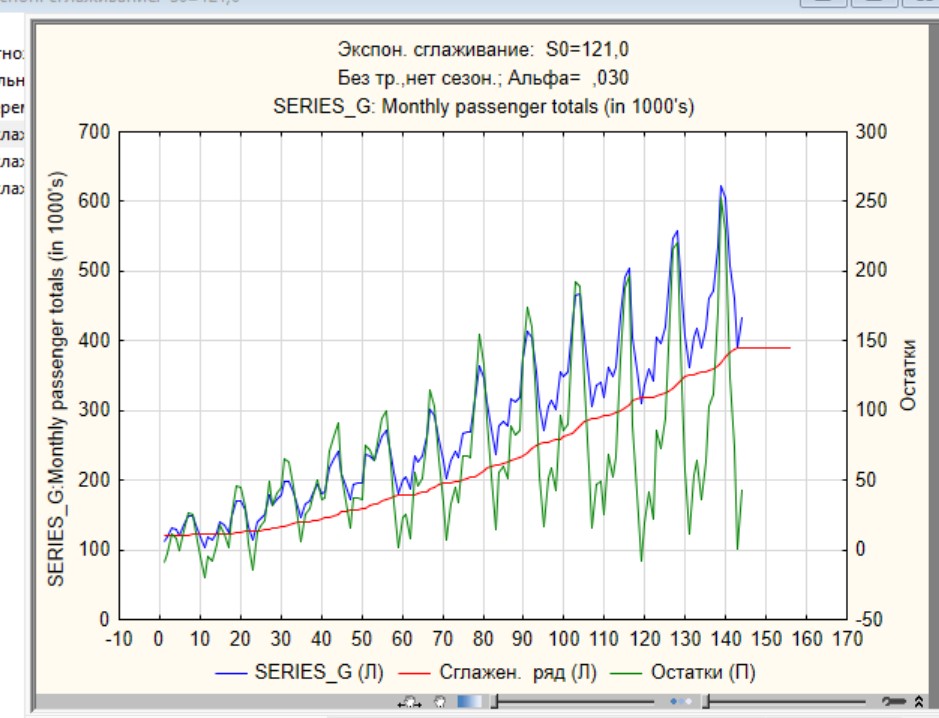


Рисунок 4 – Модель, которая представляет собой реакцию фильтра, смещённого вправо на 32

Удалим график остатков. Выбрать закладку shift, выбрать сдвиг назад и установить на 32.

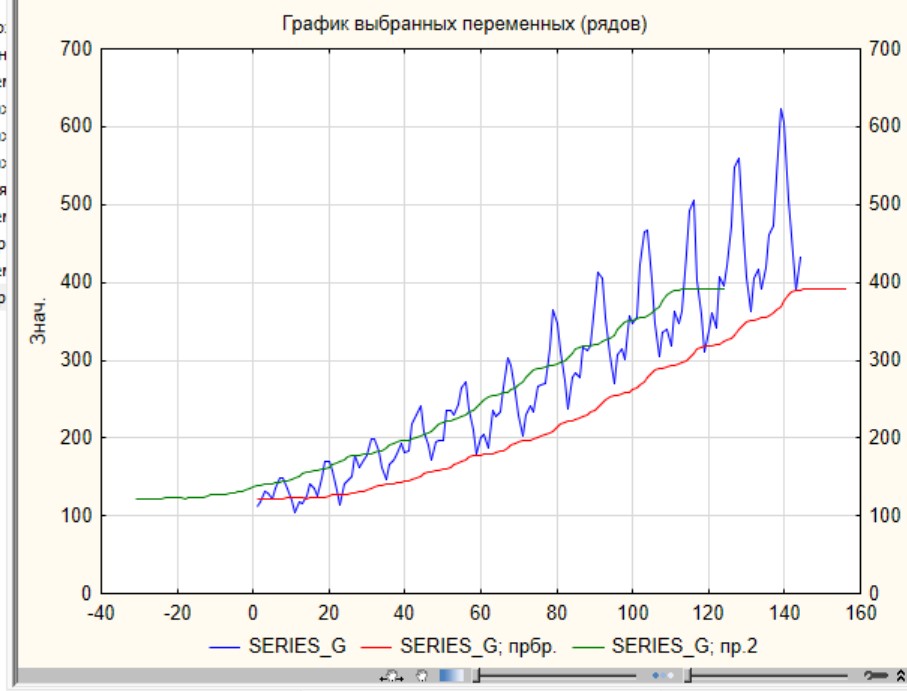


Рисунок 5 – Модель с сдвигом назад на 32

В основную таблицу скопируем значения, которые получили через кнопку save value.

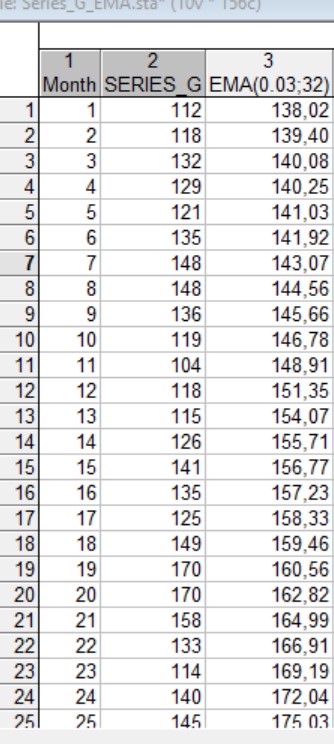


Рисунок 6 – Основная таблица с сохранёнными значениями

Далее построим модель Хольта для этого поменяем выбор на поле Holt. Параметр гамма установим 0,05.

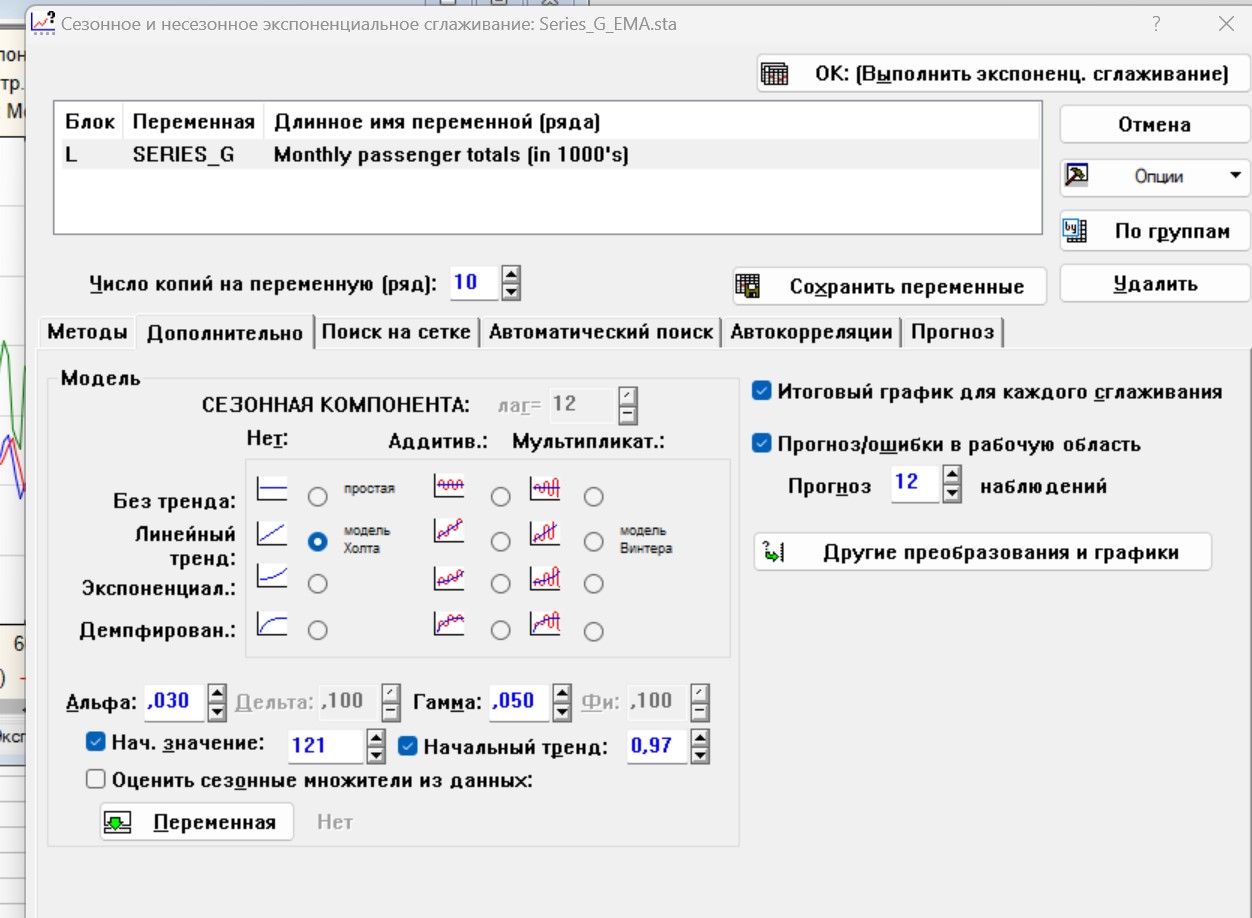


Рисунок 7 – Параметры для модели Хольта

Полученные результаты сохраняем в основную таблицу.

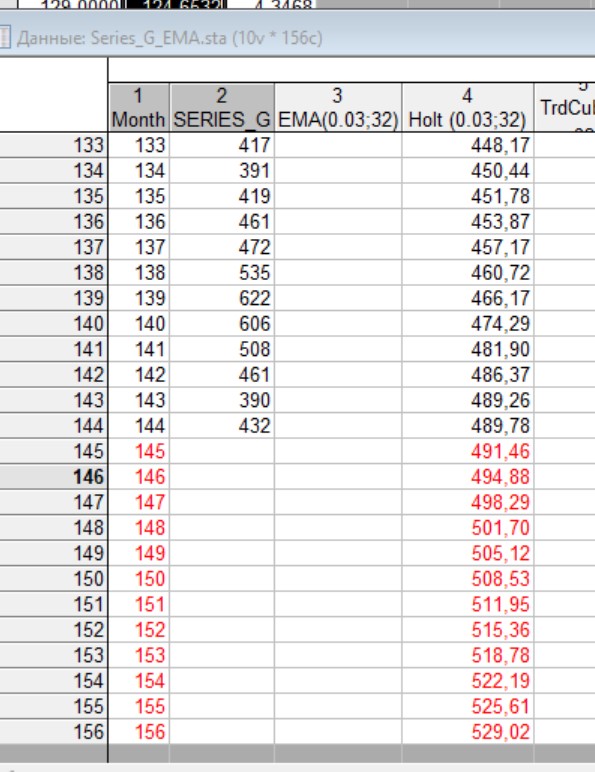


Рисунок 8 – Модель тренда с прогнозом

Посмотрим, как ведёт себя остаток. Выберем запись остатка и выберем кубическим полигоном.

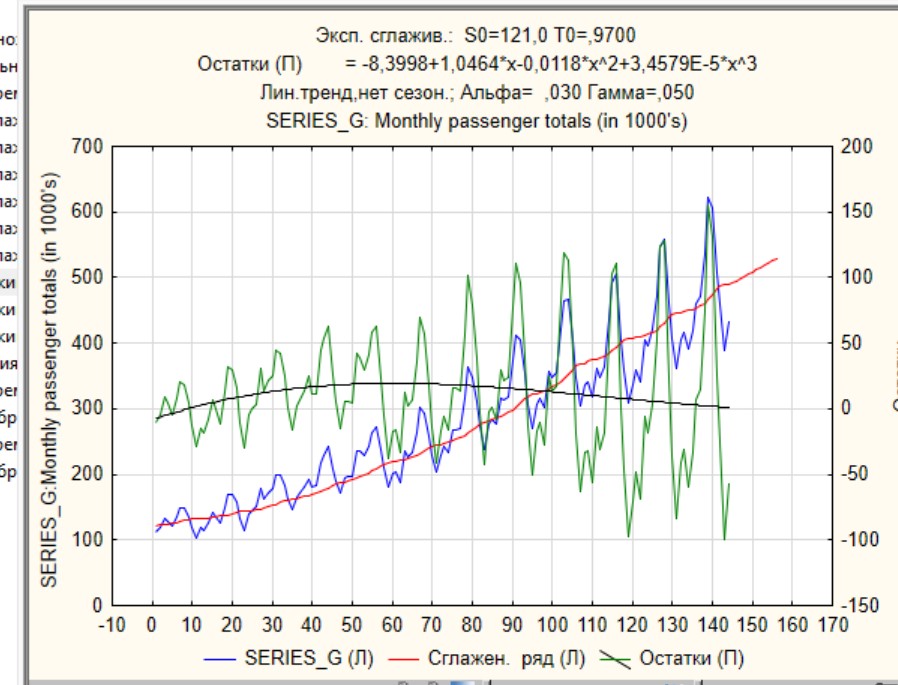


Рисунок 9 – График с аппроксимации остатка кубическим полиномом

Добавим полученный значения в основную таблицу.

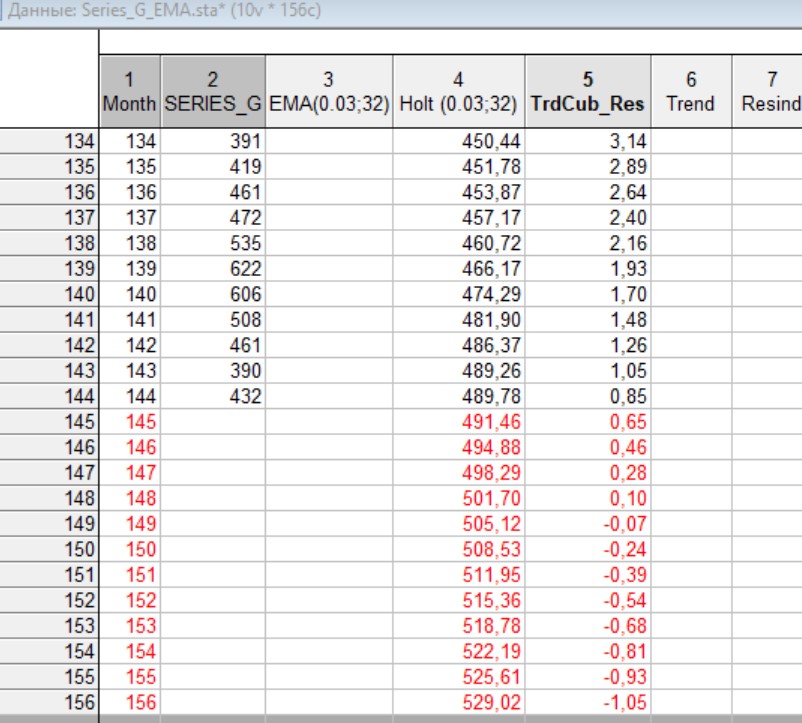


Рисунок 10 – Значения остатка в основной таблице

Посчитаем финальный тренд для этого в соседний столбец формулу v4+v5.

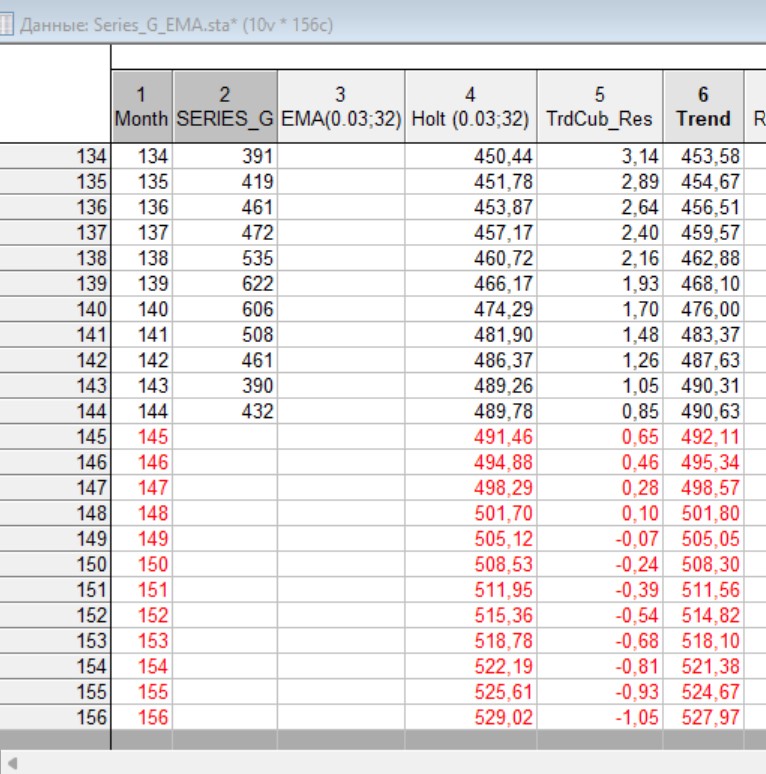


Рисунок 11 – Посчитанный финальный тренд

Посчитаем остаток для этого введём формулу v2-v6.

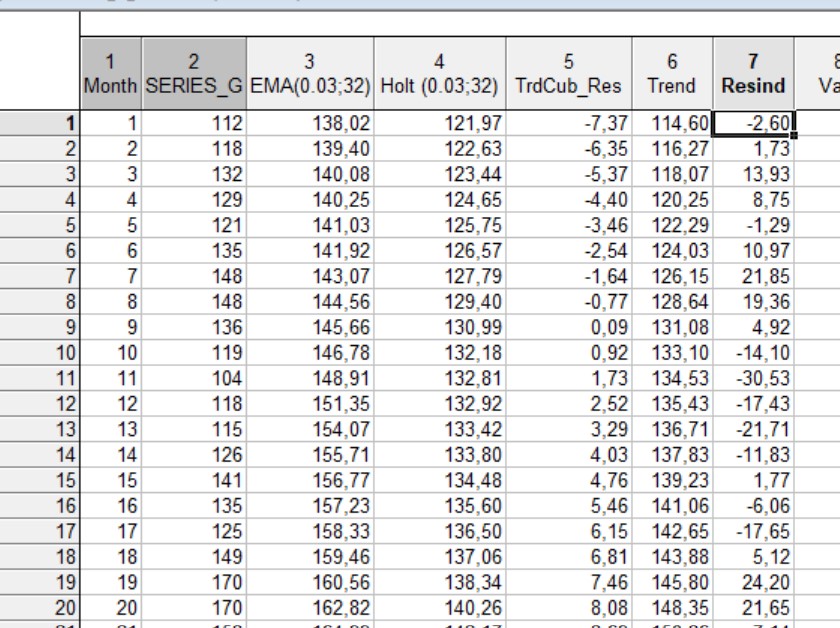


Рисунок 12 – Остаток

Убедимся что финальный остаток обладает 0 трендом, для этого выберем подгонка, затем остаток, затем кубический полином.

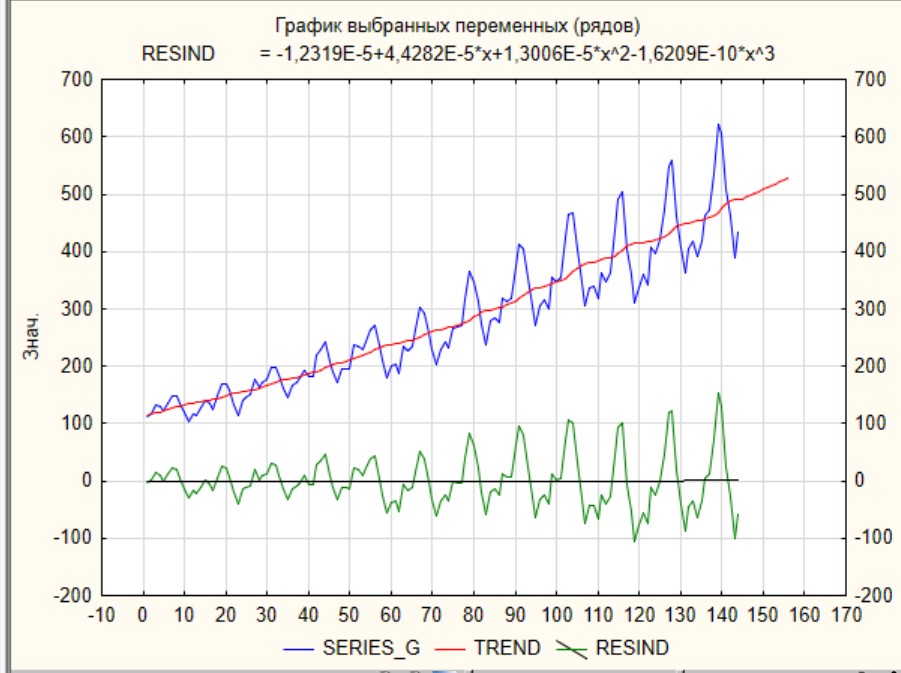


Рисунок 13 – График начальных данных, финального тренда и остатка

Прологарифмируем исходные данные. Для этого в панели 2 уровня выберем x=f(x), в котором выберем натуральный алгоритм. И на полученный график выберем 3 подгонки – линейную на всё диапазоне, линейную на диапазон от 0 до 24, экспоненциальную.

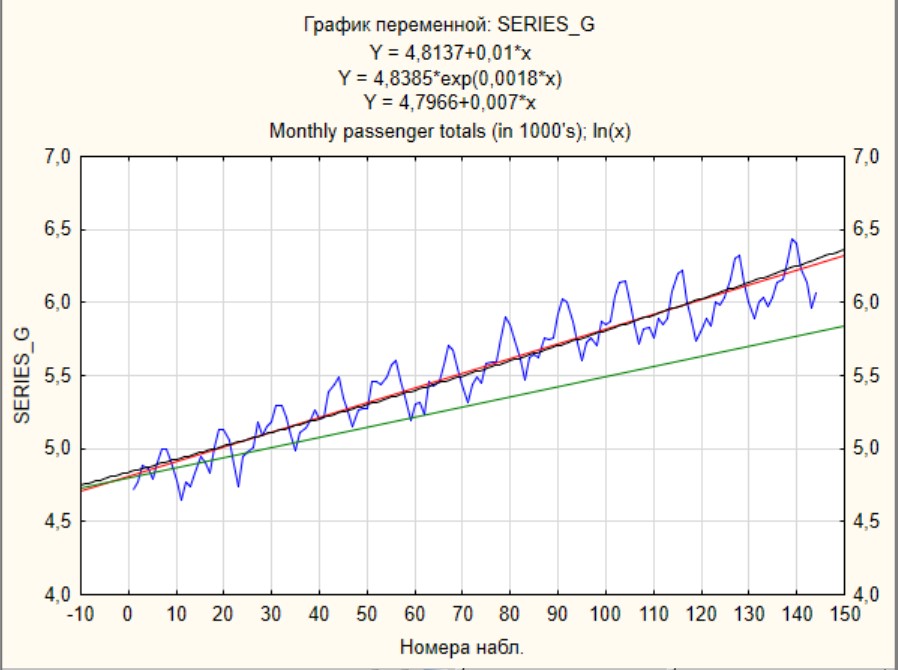


Рисунок 14 - Прологарифмированные данные

Далее перейдём на основную панель, где нажмём на кнопку Exponential smoothing & forecasting, в которой выберем модель Винтерса и установим параметры - указать окно прогноза 12, поля alpha -0,03, delta -0,07, gamma -0,05. Так же из предыдущего шага берёт параметы, которые заносим в поле нач. значение = 121 и начальный тренд 0,97.

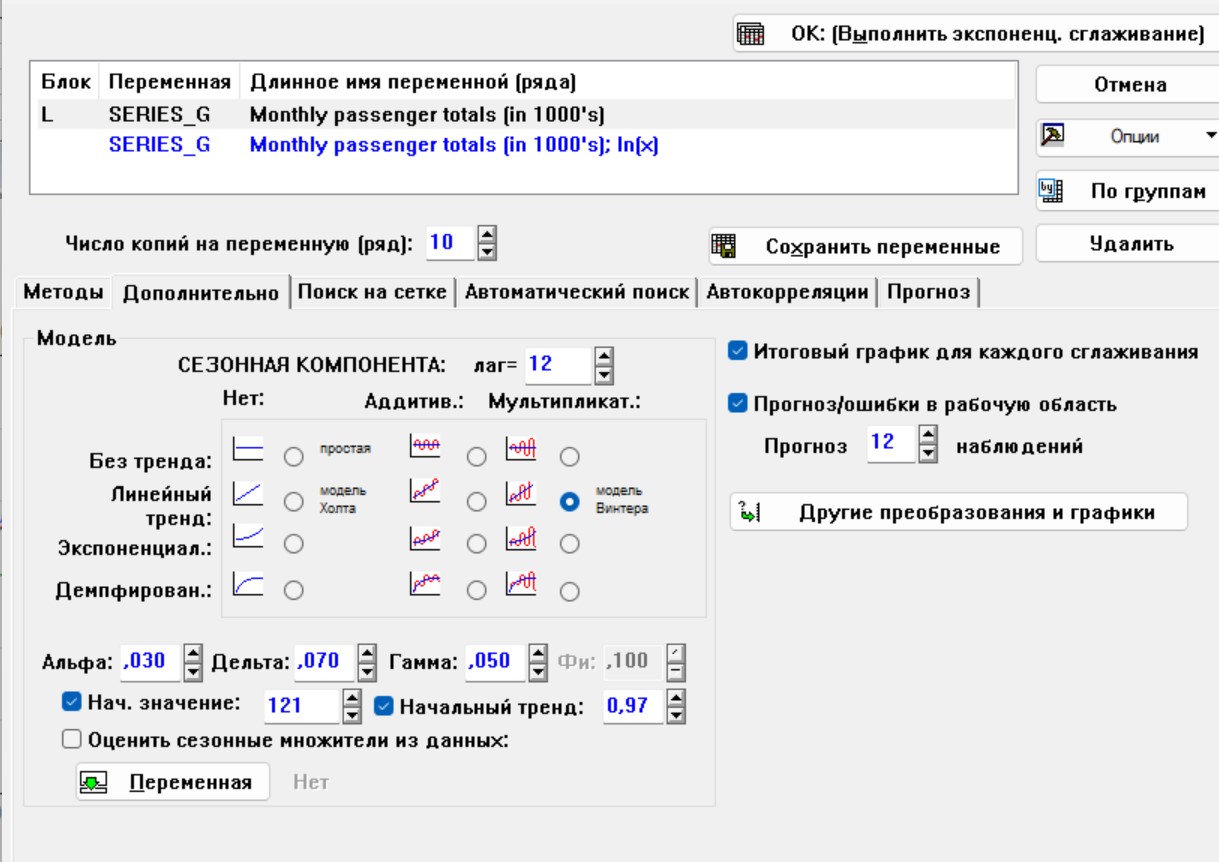


Рисунок 15 - Заполненная панель для временного ряда Винтерса

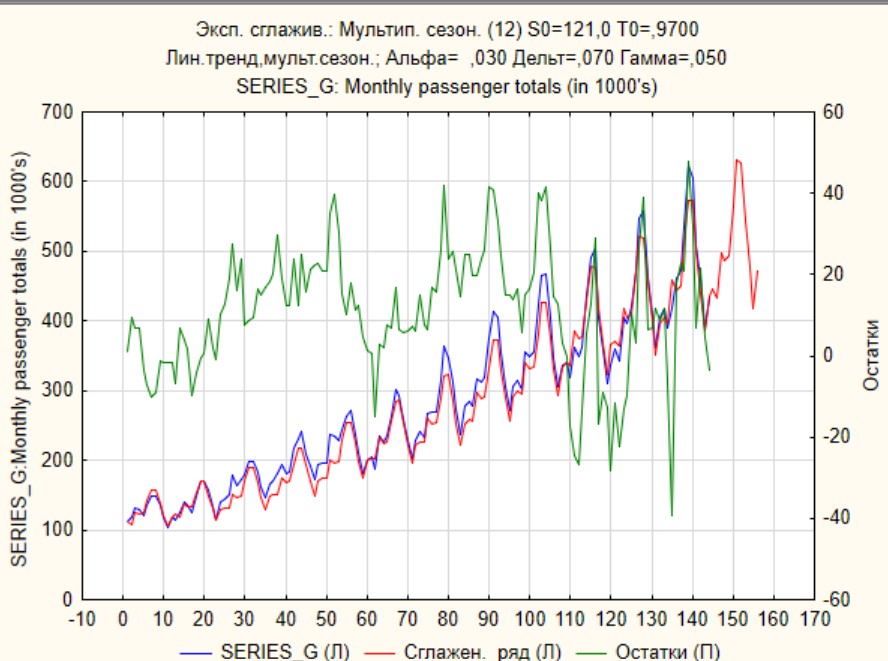


Рисунок 16 - Прогноз сезонного цикла

Возьмём временной ряд Винтерса и скопируем на основную книгу. Далее удалим все кроме исходных данных и логарифмического масштаба из выбранных переменных и переключаем в соседнее поле. Из предыдущего шага берём 4,8 и 0,007 и заполним их в поля нач. значение = 121 и начальный тренд, затем выберем логарифмированные данные и нажимает кнопку экспоненциально сгладить.

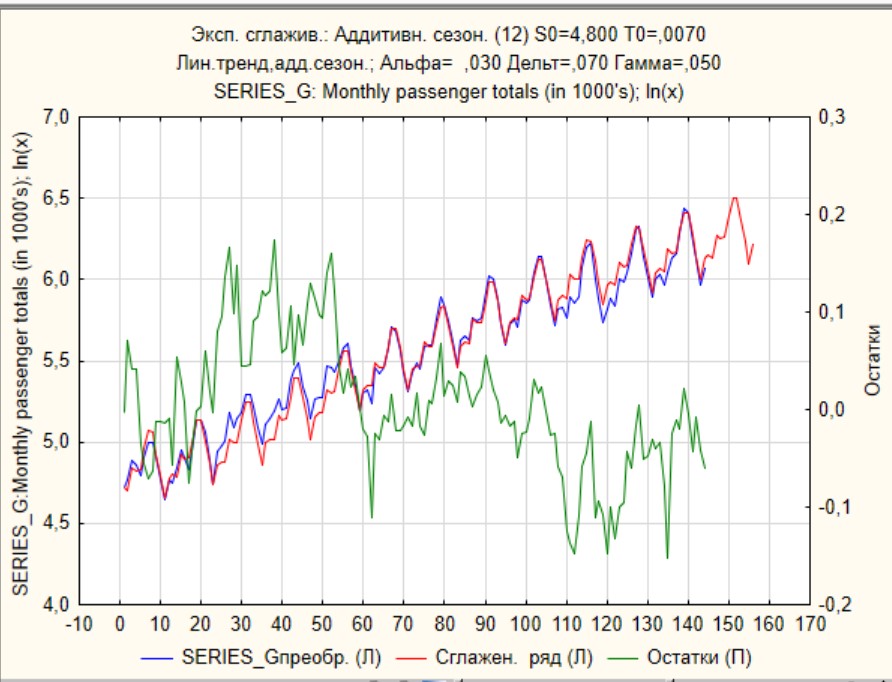


Рисунок 17 - Временной ряд Тейла-Веджа в логарифмическом масштабе.

Необходимо вернутся в исходный масштаб. По кнопке преобразование перейдём на панель второго уровня и выберем кнопку экспонента и нажать кнопку преобразовать и сохраняем результат из 3 столбца, после нажатия кнопки сохранить параметры.

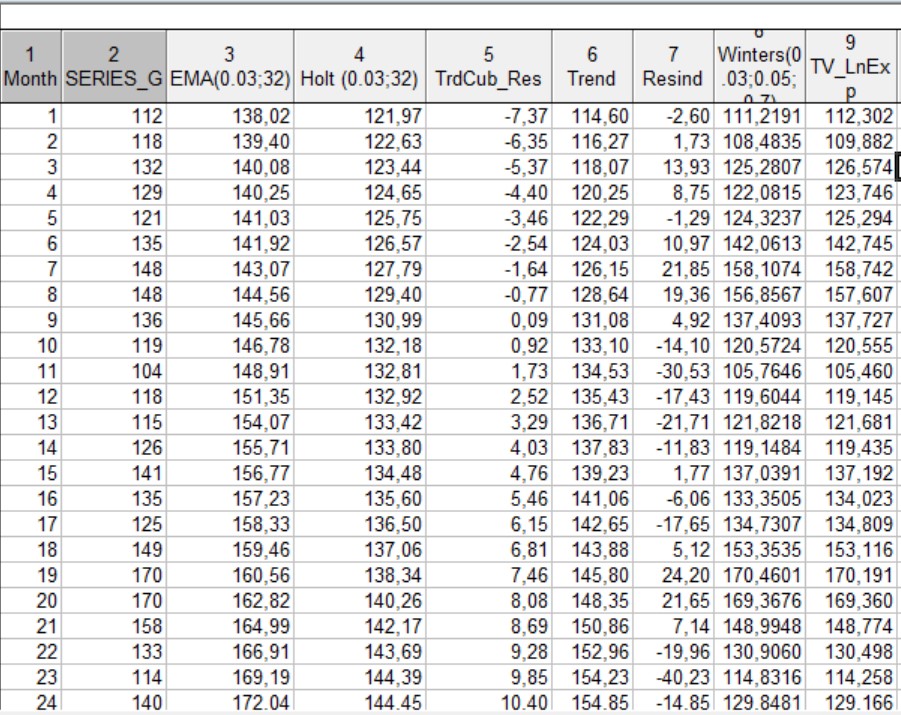


Рисунок 18 - Временной ряд Тейла-Веджа

Выберем начальные значения, финальный тренд и временные ряды Винтерса и Тейла-Вейджа на один график и сделаем шаг 6 с началом 108 и концом 168.

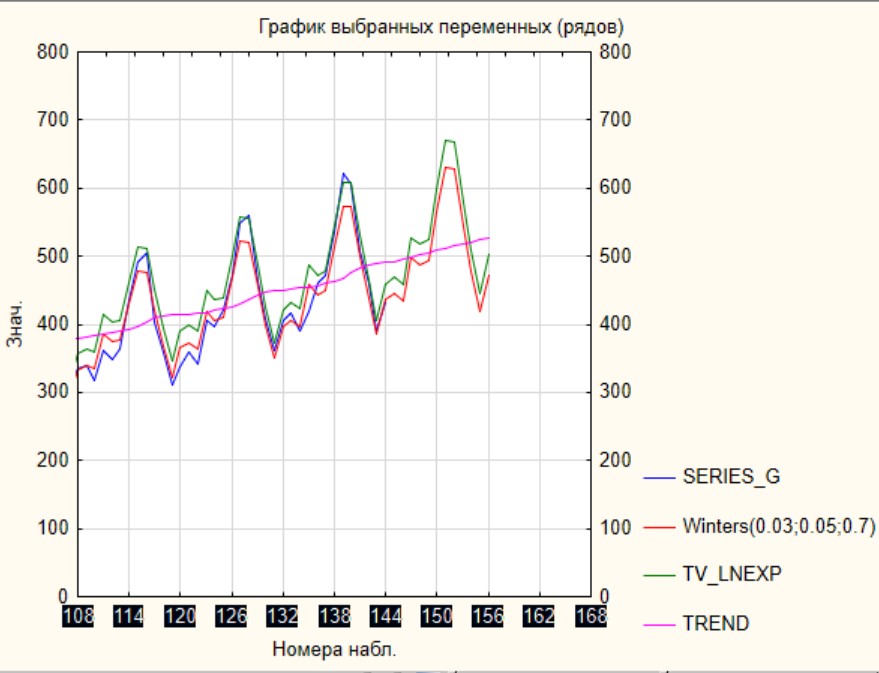


Рисунок 19 - График начальных значения, финального тренд и временные ряды Винтерса и Тейла-Вейджа

**5 Выводы**

Адаптивные модели экспоненциального сглаживания с достаточной степенью точности моделируют и прогнозируют динамику НВР. Модель линейного тепа изменения Хольта адекватно описывает тренд НВР. Наличие трендовой и квазипериодической составляющих НВР требует применения сезонных моделей Тейла – Вейджа или Винтерса в зависимости от аддитивного или мультипликативного характера поведения сезонности.

**Список литературы**

1. Цветков Э. И. Нестационарные случайные процессы и их анализ. – М.: Энергия, 1973. – 128 с.

2. Мирский Г. Я. Аппаратурное определение характеристик случайных процессов. – М.: Энергия, 1972. – 456 с.

3. Мирский Г. Я. Характеристики стохастической взаимосвязи и их измерения. – М.: Энергоиздат, 1982. – 320 с.

4. Ольшевский В. В. Основы теории статистический измерений. – Таганрог: ТРТИ, 1976. – 107 с.

5. Бендат Дж., Пирсол А. Прикладной анализ случайных данных: Пер. с англ. – М.: Мир, 1989. – 540 с.

6. Пугачев В. С. Теория случайных функций и ее применение к задачам автоматического управления. – М.: Физматгиз, 1962. – 883 с.

7. Романенко А. Ф., Сергеев Г. А. Аппроксимативные методы анализа случайных процессов. – М.: Энергия, 1974. – 176 с.

8. Прикладной анализ случайных процессов. Под ред. Прохорова С. А. / Самара: СНЦ РАН, 2007. – 582 с.

9. Голяндина Н. Э. Метод «Гусеница» - SSA: анализ временных рядов: Учеб. пособие. – СПб.: С. – Петерб. гос. ун-т, 2004. – 76 с.

10. Голяндина Н. Э. Метод «Гусеница» - SSA: прогноз временных рядов: Учеб. пособие. – СПб.: С. – Петерб. гос. ун-т, 2004. – 76 с.

11 Лабунец Л. В., Лебедева Н.Л. . Чижов М. Ю. Рекуррентные статистики нестационарных временных рядов // Радиотехника и электроника, 2011, т. 56, № 12, с. 1468 – 1489.

12. Боровиков В. П. Популярное введение в программу STATISTICA. - 2000. – 269с.

13. Боровиков В. П. STATISTICA. Искусство анализа данных на компьютере: для профессионалов. 2-е изд. - СПб.: Питер, 2003. – 688с.

14. Тихонов Э. Е. Методы прогнозирования. – Невинномысск, 2006. - 206с.

15. Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов: Учебное пособие. – М.: Финансы и статистика, 2003. – 416 с.