

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
**«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»**
(Н И У « Б е л Г У »)

ИНСТИТУТ ИНЖЕНЕРНЫХ ТЕХНОЛОГИЙ И ЕСТЕСТВЕННЫХ НАУК
КАФЕДРА ПРИКЛАДНОЙ ИНФОРМАТИКИ И ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ

**РАЗРАБОТКА ПРИКЛАДНОГО ПАКЕТА АНАЛИЗА И
ПРОГНОЗИРОВАНИЯ НА ОСНОВЕ ВРЕМЕННЫХ РЯДОВ**

Выпускная квалификационная работа
обучающегося по направлению подготовки 09.03.03 «Прикладная
информатика»
очной формы обучения, группы 07001405
Паршина Ивана Игоревича

Научный руководитель:
к.т.н. Путивцева Н.П.

БЕЛГОРОД 2018

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
1 Аналитическая часть.....	6
1.1 Обзор методов анализа временных рядов	6
1.2 Обзор эконометрических пакетов	17
1.3 Модель «Как есть», недостатки и проблемы	22
1.4 Постановка задачи	29
2 Проектная часть.....	31
2.1 Модель «Как должно быть».....	31
2.2 Структура программы	33
2.3 Алгоритм работы	35
3 Разработка и тестирование программного пакета и оценка эффективности	43
3.1 Разработка пакета.....	43
3.2 Контрольный пример.....	52
3.3 Оценка эффективности.....	58
ЗАКЛЮЧЕНИЕ	63
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	64
ПРИЛОЖЕНИЕ А	70
ПРИЛОЖЕНИЕ Б.....	71
ПРИЛОЖЕНИЕ В	75

ВВЕДЕНИЕ

В настоящее время дисциплина «Эконометрика» входит в учебные планы многих направлений подготовки. Одним из ключевых разделов данной дисциплины является «Временные ряды». В ходе изучения данного раздела обучающиеся получают основные сведения и навыки по анализу и прогнозированию показателей путём применения методов регрессионного анализа к временным рядам, содержащим соответствующие определённым моментам времени показатели. Однако, данный раздел эконометрики зачастую вызывает затруднения при изучении из-за сложности самостоятельной проверки корректности расчётов, а также низкой их наглядности при выполнении вычислений без привлечения средств электронной вычислительной техники. Данная проблема может быть решена путём использования специализированных пакетов прикладных программ.

Однако, имеющиеся на данный момент программные пакеты зачастую не позволяют произвести все операции, рассматриваемые в рамках раздела «Временные ряды», а, если и позволяют, то зачастую требуют дополнительного обучения для взаимодействия с данными системами. Также к основным недостаткам существующих систем, реализующих данный функционал, можно отнести отсутствие пояснений к получаемым результатам, что является ключевым в процессе обучения. Данные факторы обуславливают необходимость в разработке новой системы с учётом обозначенных недостатков существующих систем.

Объектом исследования является анализ временных рядов в процессе обучения дисциплине «Эконометрика».

Предметом исследования является программное сопровождение анализа временных рядов.

Целью данной работы является повышение эффективности процесса обучения в рамках раздела «Временные ряды» дисциплины «Эконометрика».

В рамках данной работы были поставлены следующие задачи:

- 1) провести анализ предметной области;
- 2) провести моделирование процесса обучения «Как есть» и «Как должно быть»;
- 3) осуществить проектирование программного пакета;
- 4) осуществить разработку программного пакета и оценить его эффективность.

Критериями достижения цели являются:

- значительное сокращение времени, затрачиваемого на выполнение лабораторных работ по данной дисциплине по сравнению с выполнением данных работ с использованием прикладного пакета MS Excel;
- высокая степень удовлетворённости пользователей, использующих программный продукт.

Данная работа содержит 3 раздела. Первый раздел данной работы содержит обзор методов анализа временных рядов и инструментария (статистических пакетов общего назначения и эконометрических пакетов) и их сравнение. Второй раздел работы содержит: модель «Как должно быть», структуру программы, алгоритм работы. В третьем разделе представлены: процесс разработки системы, функционал окон разрабатываемого программного пакета, результаты его тестирования (контрольный пример), а также оценку его эффективности.

В приложениях изложены: таблица значений критериев Стьюдента, обзор некоторых эконометрических методов и увеличенные пояснения программы

Пояснительная записка изложена на 75 страницах и содержит 48 рисунков, 3 таблицы и 3 приложения.

1 Аналитическая часть

1.1 Обзор методов анализа временных рядов

Временной ряд (ряд динамики) – обобщённые статистические данные за определённый период времени, отражающие значение тех или иных параметров рассматриваемого процесса [5]. Главным отличием ряда динамики от простой выборки данных при анализе является учёт связи его измерений с моментом времени, а не только статистические характеристики [6]. Элемент временного ряда называется отсчетом или измерением, также допустимо название его уровнем в том случае, если вместе с ним указан момент времени. Для каждого отсчета в ряду динамики должен указан соответствующий ему временной период или порядковый номер его измерения [8].

Анализ временных рядов включает следующие операции [13]:

- анализ ряда на стационарность;
- определение структуры ряда;
- аппроксимация;
- построение тренда;
- сглаживание.

При анализе ряда на стационарность применяются различные параметрические, непараметрические и полупараметрические тесты. К наиболее часто применяемым относятся следующие методы и критерии [13]:

- критерий KPSS;
- метод выборочной медианы;
- критерий Стьюдента;
- тест Фишера.

Критерий KPSS (критерий Квятковского-Филлипса-Шмидта-Шина)

заключается в выдвижении двух гипотез:

- H_0 – временной ряд является стационарным;
- H_1 – временной ряд не является стационарным.

Если рассматриваемый ряд вида:

$$Y_t = c_t + \delta * t + e_t, \quad (1.1)$$

$$c_t = c_{t-1} + \vartheta_t, \quad (1.2)$$

где δ – коэффициент тренда;

e_t – стационарный процесс;

ϑ – некоторый независимый и одинаково распределенный с e_t процесс с математическим ожиданием 0 и дисперсией σ^2

то при $\sigma^2=0$ подтверждается гипотеза H_0 , если $\sigma^2 \neq 0$, то – гипотеза H_1 , при этом σ^2 вычисляется по формуле 1.3.

$$\sigma^2 = \frac{\sum_{t=1}^T S_t^2}{s^2 * T^2}, \quad (1.3)$$

где T – размер выборки;

$$S_t = e_1 + e_2 + \dots + e_t;$$

s^2 – Стандартная ошибка в форме Нью-Уеста.

Метод выборочной медианы – метод анализа временного ряда на стационарность, основанный на выборе медианы от временного ряда.

Если имеется временной ряд $y(t)$, то он сортируется в порядке возрастания. Далее, если в ряде нечетное число элементов, то в качестве медианного значения x_{med} берётся средний элемент отсортированного ряда, если – чётное, то x_{med} рассчитывается как среднее арифметическое значение от двух средних элементов. Следующим шагом является построение ряда $y'(t) = x_t - x_{med}$. После этого строится ряд серий $s(t)$ по следующим правилам:

- если $y'_t < 0$, то $s(t) = -1$;
- если $y'_t > 0$, то $s(t) = 1$;
- если $y'_t = 0$, то $s(t)$ пропускается.

После этого происходит подсчёт числа серий v и максимальной длины серии l . При этом под серией понимается непрерывная последовательность одинаковых элементов ряда $s(t)$. Если последующий элемент меняет знак или был пропущен, то серия прерывается.

Последним этапом является проверка двух условий:

- $v > \frac{1}{2} * (n + 1 - 1,96 * \sqrt{n - 1})$;
- $l < 3,3 * \log(n + 1)$.

Если оба данных условия выполняются, то ряд стационарен и тренда не имеет. Если нет – ряд нестационарен и имеет тренд. Данный метод прост в реализации и применении и обладает достаточной точностью. К недостаткам можно отнести невозможность расчёта вероятности, с которой гипотеза подтверждается или опровергается.

Критерий Стьюдента – параметрический критерий стационарности ряда [2]. Суть данного метода заключается в вычислении коэффициента Стьюдента k в его последующим сравнением с табличным значением, соответствующим $\eta - 2$ степеням свободы, где $\eta = n - 2$ – число элементов ряда.

Вначале выдвигаются две гипотезы:

- H_0 – временной ряд является стационарным;

– H_1 – временной ряд не является стационарным.

При расчёте значения критерия Стьюдента для временного ряда происходит разбиение его на 2 части, состоящие из T_1 и T_2 элементов соответственно. После чего происходит вычисление коэффициента Стьюдента k по формуле 1.4.

$$k = \frac{\overline{M}_1 - \overline{M}_2}{\sqrt{\frac{s_1^2}{T_1} + \frac{s_2^2}{T_2}}}, \quad (1.4)$$

где \overline{M}_1 – среднее значение первого субряда;

\overline{M}_2 – среднее значение второго субряда;

s_1^2 – стандартное отклонение первого субряда

s_2^2 – стандартное отклонение второго субряда.

После того как коэффициент был вычислен происходит его сравнение с $\eta - 2$ элементов таблицы значений критериев Стьюдента (Приложение А).

Если $k > \eta - 2$, то гипотеза H_0 отбрасывается и принимается гипотеза H_1 , что означает нестационарность ряда и наличие тренда. Если же $k < \eta - 2$, то ряд стационарен и тренда не имеет. Данный метод имеет высокую точность, однако является довольно сложным в применении и реализации. Кроме того, этот метод рекомендуется использовать только для рядов с нормальным распределением, что не всегда возможно.

Тест Фишера – параметрический тест проверки постоянности дисперсии (стационарности) временного ряда. Суть метода заключается в сравнении расчётного значения критерия Фишера F с табличным значением $F(p^*, n_1, n_2)$, где p^* – заданный уровень доверительной вероятности, $n_1 = n - 1$ – число

степеней свободы (n – число разбиений), $n_2 = \sum_{i=1}^n (T_i) - n$, где T_i – число элементов на i – ом интервале разбиения.

В начале интервал наблюдения временного ряда $y(t), t = 1, 2, \dots, T$ разбивается на n интервалов разбиения, после чего происходит расчёт значения критерия Фишера по формуле 1.5.

$$F = \frac{\frac{1}{n-1} * \sum_{j=1}^n T_j * (\bar{y}_j - \bar{y})^2}{s^2(n)}, \quad (1.5)$$

где n – число частей разбиения интервала;

T_j – количество элементов, входящих в интервал разбиения j ;

\bar{y} – среднее значение ряда;

s^2 – среднее значение дисперсии, рассчитываемое по формуле 1.6.

$$s^2 = \frac{1}{T-1} * \sum_{j=1}^n (T_j - 1) * \bar{s}_j^2, \quad (1.6)$$

где \bar{s}_j^2 – дисперсия j – го интервала.

Если $F < F(p^*, n_1, n_2)$, то ряд является стационарным с вероятностью p^* . К достоинствам этого метода отнести возможность подбора вероятности, с которой ряд будет стационарен. К недостаткам можно отнести высокую трудоёмкость и неоднозначность результата.

Для определения структуры ряда используется расчёт коэффициентов автокорреляции. Коэффициент автокорреляции показывает тесноту связи между уровнями ряда, сдвинутыми на 1 или более шаг. Расчёт коэффициента k – ого порядка по формуле 1.7.

$$r_k = \frac{\sum_{t=k+1}^n (y_t - \overline{y_{1+2*(k-1)}}) * (y_{t-k} - \overline{y_{2*k}})}{\sqrt{\sum_{t=k+1}^n (y_t - \overline{y_{1+2*(k-1)}})^2 * \sum_{t=k+1}^n (y_{t-k} - \overline{y_{2*k}})^2}}, \quad (1.7)$$

где n – число элементов в временном ряду y ;

r_k – коэффициент автокорреляции r_k порядка;

$$\overline{y_{1+2*(k-1)}} = \frac{1}{n-1} * \sum_{t=k+1}^n y_t$$

$$\overline{y_{2*k}} = \frac{1}{n-1} * \sum_{t=k+1}^n y_{t-1}$$

При определении структуры ряда, как правило, ограничиваются расчётом коэффициентов автокорреляции для $\frac{n}{4}$ порядка. После того как коэффициенты были рассчитаны возможны 3 варианта:

- r_1 – наибольший – ряд имеет линейный тренд;
- $r_1 = r_\tau$ – наибольшие – ряд имеет линейный тренд с сезонными колебаниями периодичности τ ;
- r_τ – наибольший – ряд имеет нелинейный тренд с сезонными колебаниями периодичности τ .

Аппроксимация временных рядов может быть подразделена на 2 вида: линейную и нелинейную.[15]

Линейная аппроксимация предполагает сведение временного ряда к линейному тренду $y'(t) = b + a * t + \varepsilon_t$. Данная операция подразумевает расчёт коэффициентов b и a .

Коэффициент a вычисляется по формуле 1.8.

$$a = \frac{n \cdot \sum_{i=1}^n t_i \cdot y_i - (\sum_{i=1}^n t_i) \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n t_i^2 - (\sum_{i=1}^n t_i)^2}, \quad (1.8)$$

где n – число элементов ряда;

t_i – i – й момент времени;

y_i – уровень временного ряда $y(t)$ в i – й момент времени.

Коэффициент b рассчитывается по формуле 1.9.

$$b = \frac{n \cdot \sum_{i=1}^n y_i - a \cdot \sum_{i=1}^n t_i}{n}, \quad (1.9)$$

После того как значения коэффициентов были найдены, происходит расчёт значений $y'(t)$ для всех моментов времени t .

Нелинейная аппроксимация заключается в сведении временного ряда к одному из нелинейных трендов. Наиболее часто используемыми при аппроксимации временного ряда трендами являются:

- экспоненциальный тренд;
- гиперболический тренд;
- степенной тренд;
- полиномальный тренд второго порядка.

Экспоненциальный тренд представляет собой уравнение вида $y(t) = b \cdot e^{a \cdot t + \varepsilon}$. Для вычисления коэффициентов a и b производят преобразование нелинейной функции $y(t)$ в линейную функцию $y'(t) = b' + a' \cdot t'$, где :

- $t' = t$;
- $y' = \ln(y)$;

- $a = a'$;
- $b = e^{b'}$.

Преобразование заключается в перерасчёте моментов времени t и значений временного ряда y в соответствии с указанными преобразованиями. После чего происходит расчёт коэффициентов a' и b' по формулам 1.8 и 1.9 соответственно. Далее коэффициентов a' и b' преобразовываются в коэффициенты a и b . Последним этапом является расчёт значений временного ряда y для всех моментов времени.

Гиперболический тренд представляет собой уравнение вида $y(t) = b + \frac{a}{t} + \varepsilon$. Для вычисления коэффициентов a и b выполняют преобразование нелинейной функции $y(t)$ в линейную функцию $y'(t) = b' + a' * t'$, где :

- $t' = \frac{1}{t}$;
- $y' = y$;
- $a = a'$;
- $b = b'$.

Преобразование заключается в перерасчёте моментов времени t и значений временного ряда y в соответствии с указанными преобразованиями. После чего происходит расчёт коэффициентов a' и b' по формулам 1.7 и 1.8 соответственно. Далее коэффициентов a' и b' преобразовываются в коэффициенты a и b . Последним этапом является расчёт значений временного ряда y для всех моментов времени.

Степенной тренд представляет собой уравнение вида $y(t) = b * t^a * e^\varepsilon$. Для вычисления коэффициентов a и b выполняется преобразование нелинейной функции $y(t)$ в линейную функцию $y'(t) = b' + a' * t'$, где :

- $t' = \ln(t)$;
- $y' = \ln(y)$;
- $a = a'$;
- $b = e^{b'}$.

Преобразование заключается в перерасчёте моментов времени t и значений временного ряда y в соответствии с указанными преобразованиями. После чего происходит расчёт коэффициентов a' и b' по формулам 1.7 и 1.8 соответственно. Далее коэффициентов a' и b' преобразовываются в коэффициенты a и b . Последним этапом является расчёт значений временного ряда y для всех моментов времени.

Полиномиальный тренд представляет собой уравнение $y(t) = a + b_1 * t + b_2 * t^2 + \varepsilon$.

Для нахождения коэффициентов a , b_1 и b_2 составляется система уравнений:

$$\begin{aligned} a * n + b_1 * \sum t + b_2 * \sum t^2 &= \sum y \\ a * \sum t + b_1 * \sum t^2 + b_2 * \sum t^3 &= \sum y * t \\ a * \sum t^2 + b_1 * \sum t^3 + b_2 * \sum t^4 &= \sum y * t^2 \end{aligned},$$

после чего происходит решение данной системы уравнений методом Крамера. Для этого рассчитываются следующие определители:

$$\Delta = \begin{vmatrix} n & \sum t & \sum t^2 \\ \sum t & \sum t^2 & \sum t^3 \\ \sum t^2 & \sum t^3 & \sum t^4 \end{vmatrix};$$

$$- \Delta_1 = \begin{vmatrix} \sum y & \sum t & \sum t^2 \\ \sum y * t & \sum t^2 & \sum t^3 \\ \sum y * t^2 & \sum t^3 & \sum t^4 \end{vmatrix};$$

$$- \Delta_2 = \begin{vmatrix} n & \sum y & \sum t^2 \\ \sum t & \sum y * t & \sum t^3 \\ \sum t^2 & \sum y * t^2 & \sum t^4 \end{vmatrix};$$

$$- \Delta_3 = \begin{vmatrix} n & \sum t & \sum y \\ \sum t & \sum t^2 & \sum y * t \\ \sum t^2 & \sum t^3 & \sum y * t^2 \end{vmatrix}.$$

Далее происходит расчёт коэффициентов $a = \frac{\Delta_1}{\Delta}$, $b_1 = \frac{\Delta_2}{\Delta}$ и $b_2 = \frac{\Delta_3}{\Delta}$.

Последним этапом является расчёт значений временного ряда y_t для всех моментов времени.

Для определения наиболее подходящего вида тренда из тех, к которым был сведён временной ряд, выполняется расчёт средней ошибки аппроксимации и коэффициента детерминации.

Средняя ошибка аппроксимации показывает среднее отклонение расчетных значений от фактических и вычисляется по формуле 1.10.

$$\bar{A} = \frac{1}{n} * \sum_{i=1}^n \left| \frac{y_i - y'_i}{y_i} \right|, \quad (1.10)$$

Коэффициент детерминации является статистической мерой согласия, с помощью которой можно определить, насколько уравнение регрессии соответствует реальным данным и рассчитывается по формуле 1.11.

$$R^2 = \frac{\sum_{i=1}^n (y'_i - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})}, \quad (1.11)$$

где \bar{y} – среднее значение изначального ряда динамики;
 y'_i – расчётное значение ряда для момента времени i .

Сглаживание временного ряда – важный процесс, позволяющий уменьшить воздействие значений, слишком сильно отличающихся от среднего. Это позволит снизить влияние случайных факторов на результаты анализа и прогнозирования. Наиболее используемыми методами сглаживания являются метод скользящего среднего и метод экспоненциально взвешенного скользящего среднего.[17]

При сглаживании Методом скользящего среднего сглаженное значение временного ряда рассчитывается по формуле 1.12.

$$y'(t) = \sum_{k=-m}^m w_k * y_{t+k}, \quad (1.12)$$

где w – весовые коэффициенты нумерация которых начинается с $-m$ и заканчивается m с шагом 1.

Наиболее часто метод скользящего среднего применяют при значениях $m=1,2,3$.

При $m=1$ используют ряд весовых коэффициентов $w = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$.

При $m=2$ - ряд весовых коэффициентов $w = \{-\frac{3}{35}, \frac{12}{35}, \frac{17}{35}, \frac{12}{35}, -\frac{3}{25}\}$.

При $m=3$ - ряд весовых коэффициентов $w = \{-\frac{2}{21}, \frac{3}{21}, \frac{6}{21}, \frac{7}{21}, \frac{6}{21}, \frac{3}{21}, -\frac{2}{21}\}$.

Главным недостатком данного метода является потеря крайних $2*m$ значений.

Суть метода экспоненциально взвешенного скользящего среднего (Метод Брауна) при сглаживании заключается в уделении большего внимания недалеко отстоящим по времени значениям, т.е. принимается, что недавние значения временного ряда будут иметь большое влияние на последующие. Расчёт сглаженного методом Брауна значения временного ряда осуществляется по формуле 1.13.

$$y'(t) = \frac{1-\lambda}{1-\lambda^{t+1}} * y_t + \lambda * \frac{1-\lambda}{1-\lambda^{t+1}} * y'_{t-1}, \quad (1.13)$$

где $t = 2, 3 \dots n$ – момент времени.

При этом $y'_1 = y_1$. Данный метод отличается от МСС в лучшую сторону тем, что при сглаживании не теряются крайние значения. Недостатком же является необходимость подбора коэффициента λ .

1.2 Обзор эконометрических пакетов

К наиболее часто используемым для анализа временных рядов программным средствам можно отнести [17]:

- Пакет MS Excel;
- SPSS;
- Statistica;
- Eviews;
- Stata.

Пакет SPSS для Windows на данный момент одним из лидеров среди универсальных статистических пакетов. SPSS предоставляет широкий набор средств, позволяющих осуществлять деятельность на всех стадиях анализа [1].

К возможностям и особенностям программного пакета SPSS можно отнести:

- удобный доступ к информации;
- удобный сбор и ввод данных;
- эффективный data - менеджмент;
- широкий спектр статистических операций для анализа данных и построения наиболее достоверных моделей;
- возможность публикации результатов в сети.

Обзор остальных эконометрических пакетов представлен в приложении Б.

Так как сопоставить качества данных систем и закладываемые в разрабатываемую систему функциональные возможности между собой достаточно сложно, было принято решение осуществить их сравнение методом анализа иерархий Саати. Данный метод позволяет осуществить выбор оптимальной альтернативы, основываясь не только на количественных, но и на качественных критериях. Однако данный метод достаточно трудоёмок [35]. Использование программного средства позволило бы ускорить анализ с помощью данного метода. Программа, обеспечивающая необходимый функционал была создана в рамках курса «Методы и системы поддержки и принятия решений».

В рамках данного анализа можно выделить несколько наиболее существенных критериев:

- стоимость пакета;
- функциональность;
- удобство интерфейса;
- удобность загрузки/выгрузки данных;
- понятность пояснений.

Данные критерии и альтернативы были внесены как исходные данные в разработанной программе, что проиллюстрировано рисунком 1.1.

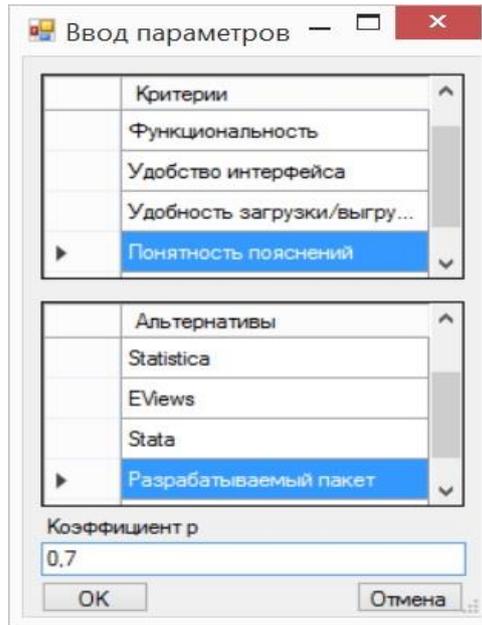


Рисунок 1.1 – Ввод параметров

Далее была задана матрица парных сравнений критериев, представленная на рисунке 1.2.

	Стоимость	Функцио	Удобство интерфейса	Удобность загрузки/в данных	Понятность Пояснений	ВПП
Стоимость	1	3	6	9	4	3,65
Функциональность	1/3	1	4	7	1/4	1,18
Удобство интерфейса	1/6	1/4	1	8	1/6	0,56
Удобность загрузки/выгрузки	1/9	1/7	1/8	1	1/9	0,19
Понятность Пояснений	1/4	4	6	9	1	2,22

Рисунок 1.2 – Матрица парных сравнений критериев

Далее была задана матрица парных сравнений альтернатив для критерия «Стоимость пакетов», представленная на рисунке 1.3.

Метод Саати

Матрица парных сравнений критериев Матрицы парных сравнений альтернатив Итоги

	MS Excel	SPSS	Statistica	EViews	Stata	Разрабатываемый пакет
MS Excel	1	5	5	3	2	1/9
SPSS	1/5	1	1/3	1/2	1/6	1/9
Statistica	1/5	3	1	1/3	1/2	1/9
EViews	1/3	2	3	1	2	1/9
Stata	1/2	6	2	1/2	1	1/9
Разрабатываемый пакет	9	9	9	9	9	1

Критерий: Стоимость Рассчитать для текущего критерия OK

Lmax: 6,682 IC: 0,136

Рисунок 1.3 – Матрица парных сравнений альтернатив для критерия «Стоимость пакетов»

Далее была задана матрица парных сравнений альтернатив для критерия «Удобность загрузки/выгрузки данных», представленная на рисунке 1.4.

Метод Саати

Матрица парных сравнений критериев Матрицы парных сравнений альтернатив Итоги

	MS Excel	SPSS	Statistica	EViews	Stata	Разрабатываемый пакет
MS Excel	1	9	3	3	5	6
SPSS	1/9	1	1/2	1/6	1/2	1/2
Statistica	1/3	2	1	1/2	1/5	3
EViews	1/3	6	2	1	1/5	4
Stata	1/5	2	5	5	1	1/2
Разрабатываемый пакет	1/6	2	1/3	1/4	2	1

Критерий: Удобность загр Рассчитать для текущего критерия OK

Lmax: 7,554 IC: 0,311

Рисунок 1.4 – Матрица парных сравнений альтернатив для критерия «Удобность загрузки/выгрузки данных»

Далее была задана матрица парных сравнений альтернатив для критерия «Удобство интерфейса», представленная на рисунке 1.5.

	MS Excel	SPSS	Statistica	EViews	Stata	Разрабатываемый пакет
MS Excel	1	8	4	5	9	3
SPSS	1/8	1	1/3	1/3	2	1/5
Statistica	1/4	3	1	1/2	4	1/3
EViews	1/5	3	2	1	4	1/3
Stata	1/9	1/2	1/4	1/4	1	1/6
Разрабатываемый пакет	1/3	5	3	3	6	1

Критерий: Удобство интер
 Lmax: 6,242 ИС: 0,048

Рисунок 1.5 – Матрица парных сравнений альтернатив для критерия «Удобство интерфейса»

На последнем этапе было произведено подведение итоговой анализа, заключающийся в расчёте веса каждой альтернативы, что проиллюстрировано рисунком 1.6.

	Стоимость	Функцию	Удобство интерфейса	Удобность загрузки/выгр данных	Понятность Пояснений	Вес альтернативы
MS Excel	0,16	0,04	0,46	0,46	0,03	0,131
SPSS	0,03	0,3	0,05	0,05	0,13	0,1
Statistica	0,05	0,19	0,1	0,1	0,08	0,0834
EViews	0,08	0,34	0,12	0,15	0,19	0,1532
Stata	0,08	0,11	0,03	0,17	0,06	0,0764
Разрабатываемый пакет	0,61	0,02	0,24	0,08	0,52	0,4537

Рисунок 1.6 – Итог анализа

Максимальным весом среди рассмотренных альтернатив обладает разрабатываемый пакет. Данный результат обусловлен акцентированием внимания на удобство интерфейса, понятность пояснений и стоимости. Преимущество разрабатываемого программного пакета в удобности интерфейса обусловлено его русифицированностью и закладываемым минимализмом. При разработке предполагается, что при выполнении анализа пояснения будут на русском языке, а также будут давать пользователю программы понятие о том, каким образом и по каким формулам было рассчитано то или иное значение или принято решение. Преимущество в цене обусловлено разработкой пакета в рамках ВКР, а также нацеленность пакета на выполнение не слишком широкого спектра задач, рассматриваемым в рамках раздела «Временные ряды» дисциплины «Эконометрика».

1.3 Модель «Как есть», недостатки и проблемы

Контекстная диаграмма процесса обучения дисциплине Эконометрика «Как есть» представлена на рисунке 1.7.

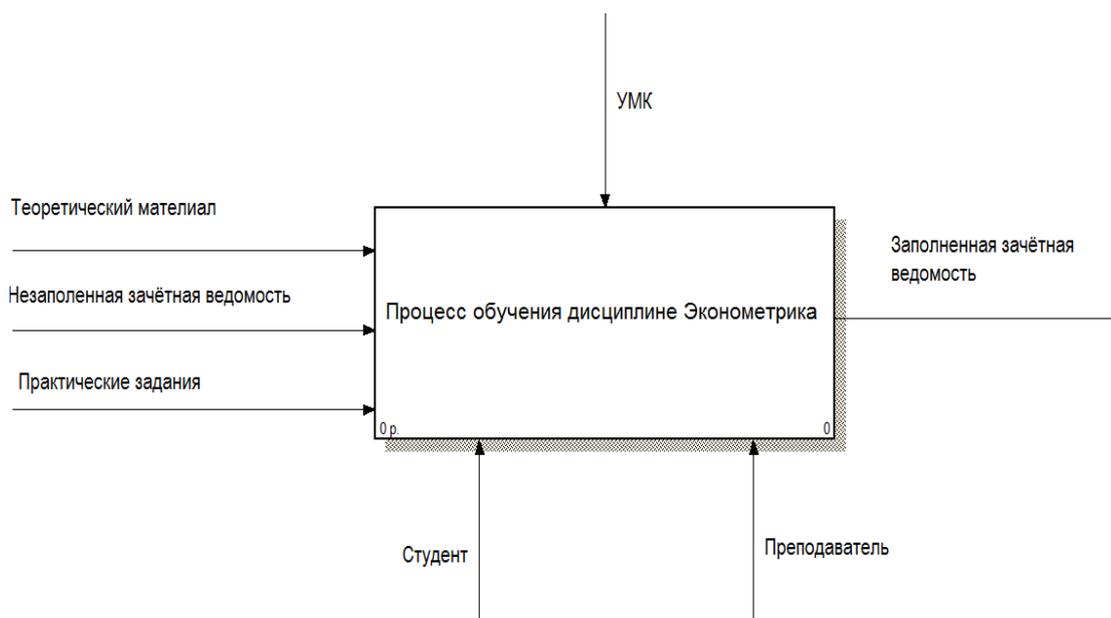


Рисунок 1.7 – Контекстная диаграмма «Процесс обучения дисциплине Эконометрика»

Процесс обучения дисциплине Эконометрика можно разбить на следующие подпроцессы:

- теоретическое обучение;
- практическое обучение;
- контроль обучения;
- сглаживание тренда.[22]

Взаимосвязь перечисленных подпроцессов была изложена на диаграмме декомпозиции процесса «Процесс обучения дисциплине Эконометрика», представленной на рисунке 1.8.

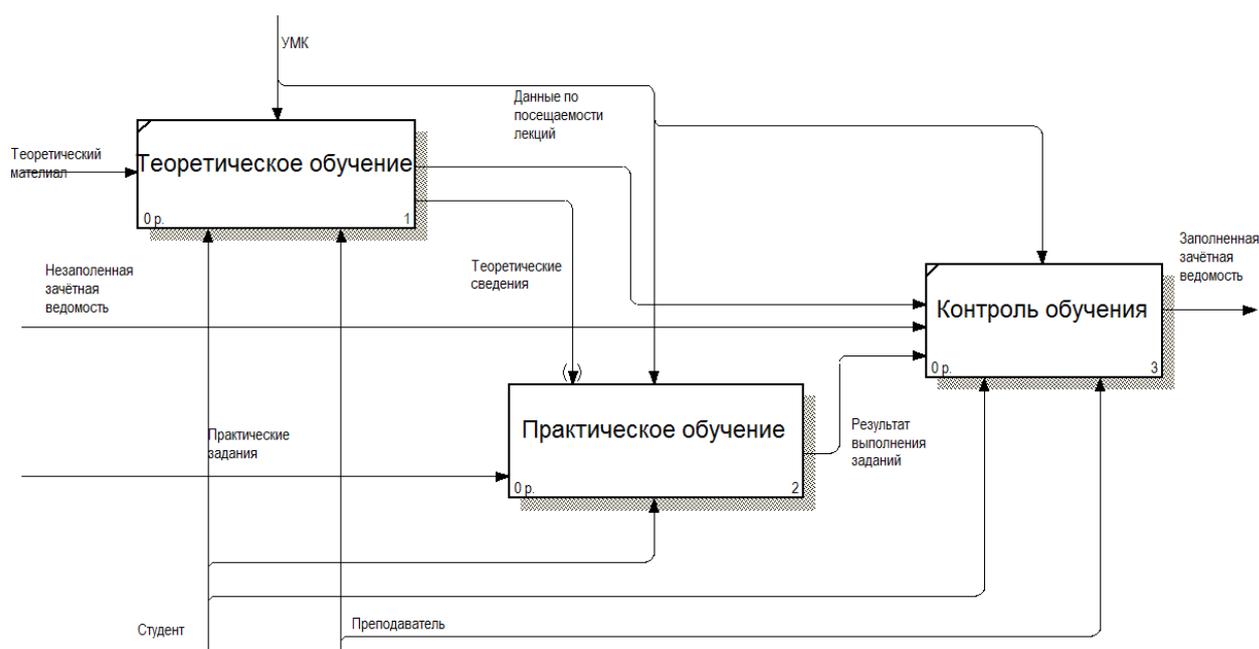


Рисунок 1.8 – Диаграмма декомпозиции процесса «Процесс обучения дисциплине Эконометрика»

Так как в данной работе рассматривается улучшение практической части обучения, то дальнейшая декомпозиция процессов «Теоретическое обучение» и «Контроль обучения» не требуется.

Процесс практического обучения в рамках курса «Эконометрика» предполагает выполнения практических заданий по следующим темам:

- корреляционно – регрессионный анализ;
- временные ряды;
- системы одновременных уравнений;
- вычисление значения коэффициента Автокорреляции первого порядка по ряду из логарифмов.

Взаимосвязь перечисленных подпроцессов была изложена на диаграмме декомпозиции процесса «Практическое обучения», представленной на рисунке 1.9.

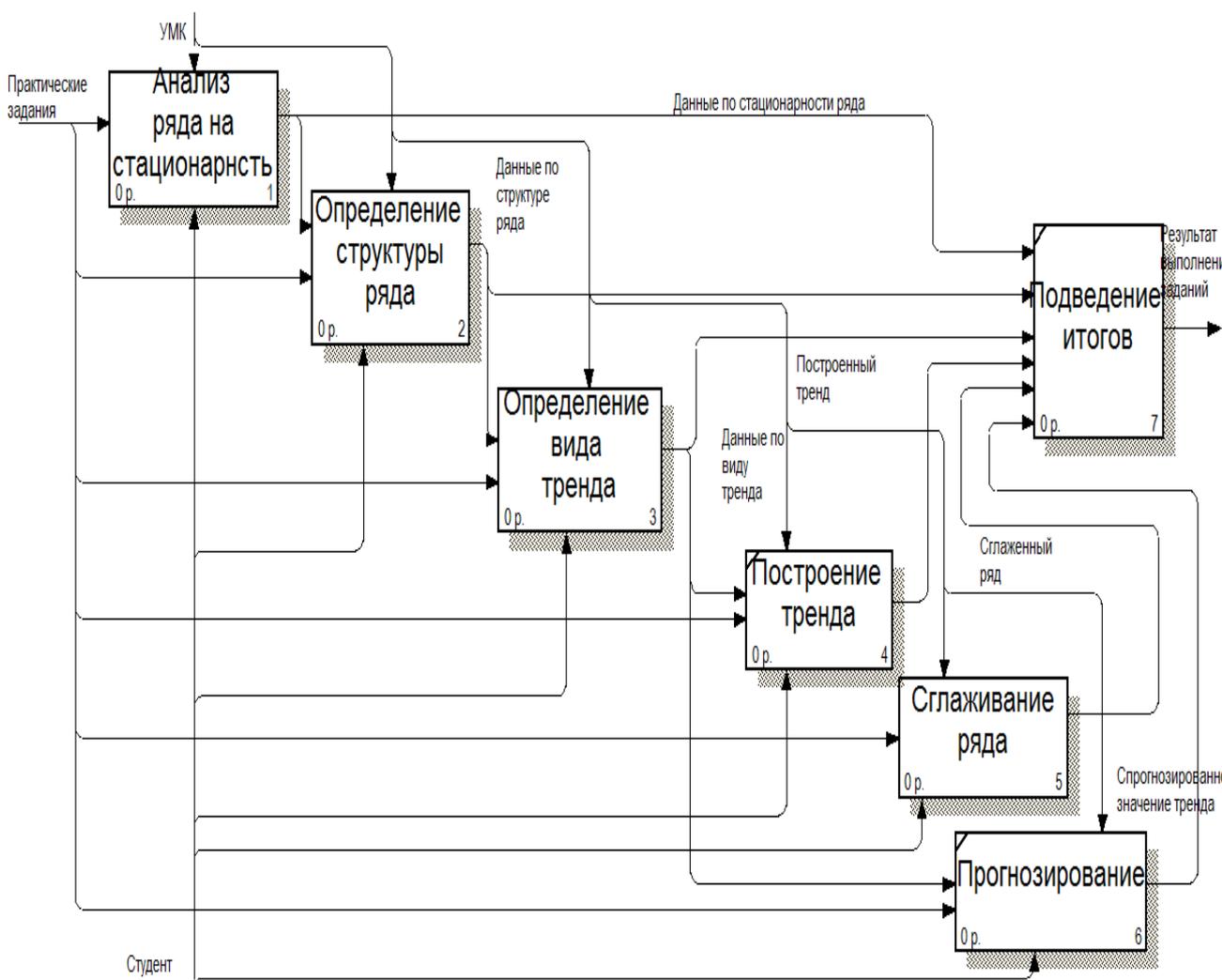


Рисунок 1.9 – Диаграмма декомпозиции процесса «Временные ряды»

Анализ ряда на стационарность может быть разбит на следующие подпроцессы:

- определение медианного значения ряда;
- построение ряда из разностей $X_i - X_{med}$;
- построение серий;
- проверка по критерию выборочной медианы.

Взаимосвязь перечисленных подпроцессов была изложена на диаграмме декомпозиции процесса «Анализ ряда на стационарность», представленной на рисунке 1.10.

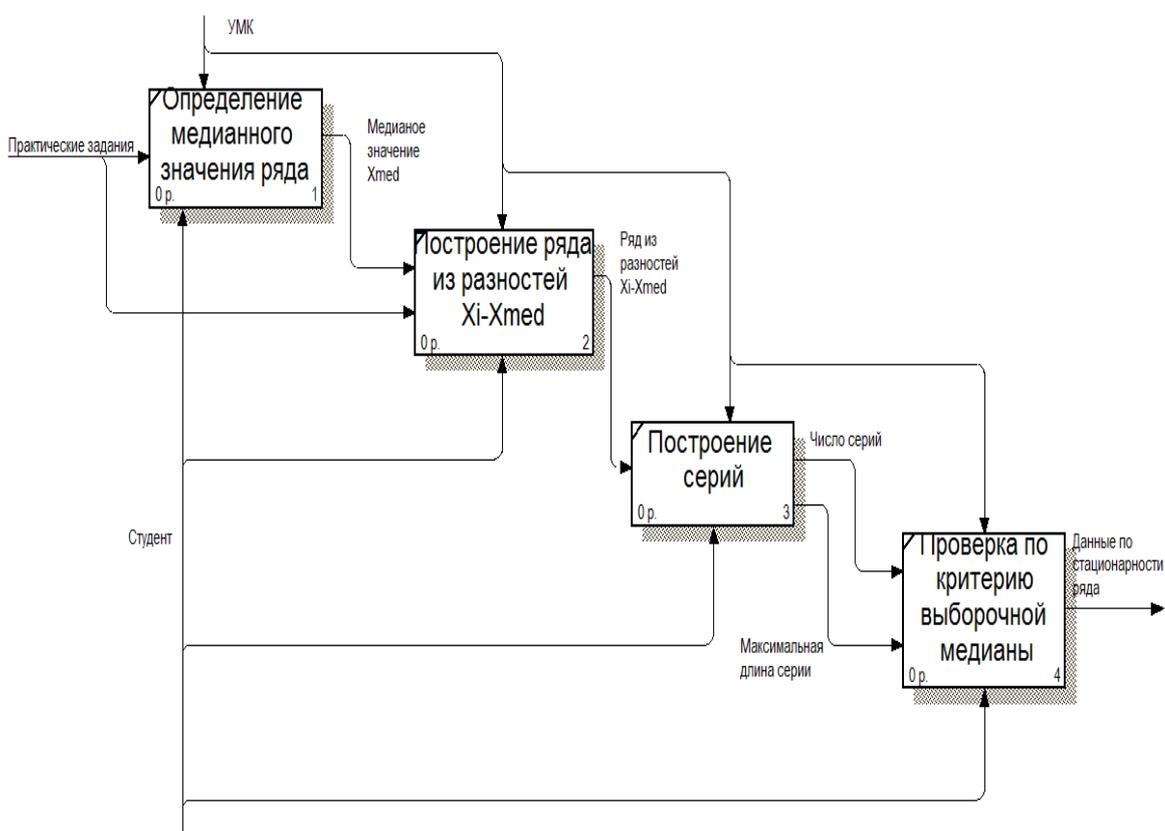


Рисунок 1.10 – Диаграмма декомпозиции процесса «Анализ ряда на стационарность»

Определение структуры тренда может быть разбито на следующие подпроцессы:

- вычисление коэффициентов автокорреляции;
- сравнение значений коэффициентов автокорреляции.

Взаимосвязь перечисленных подпроцессов была изложена на диаграмме декомпозиции процесса «Определение структуры ряда», представленной на рисунке 1.11.

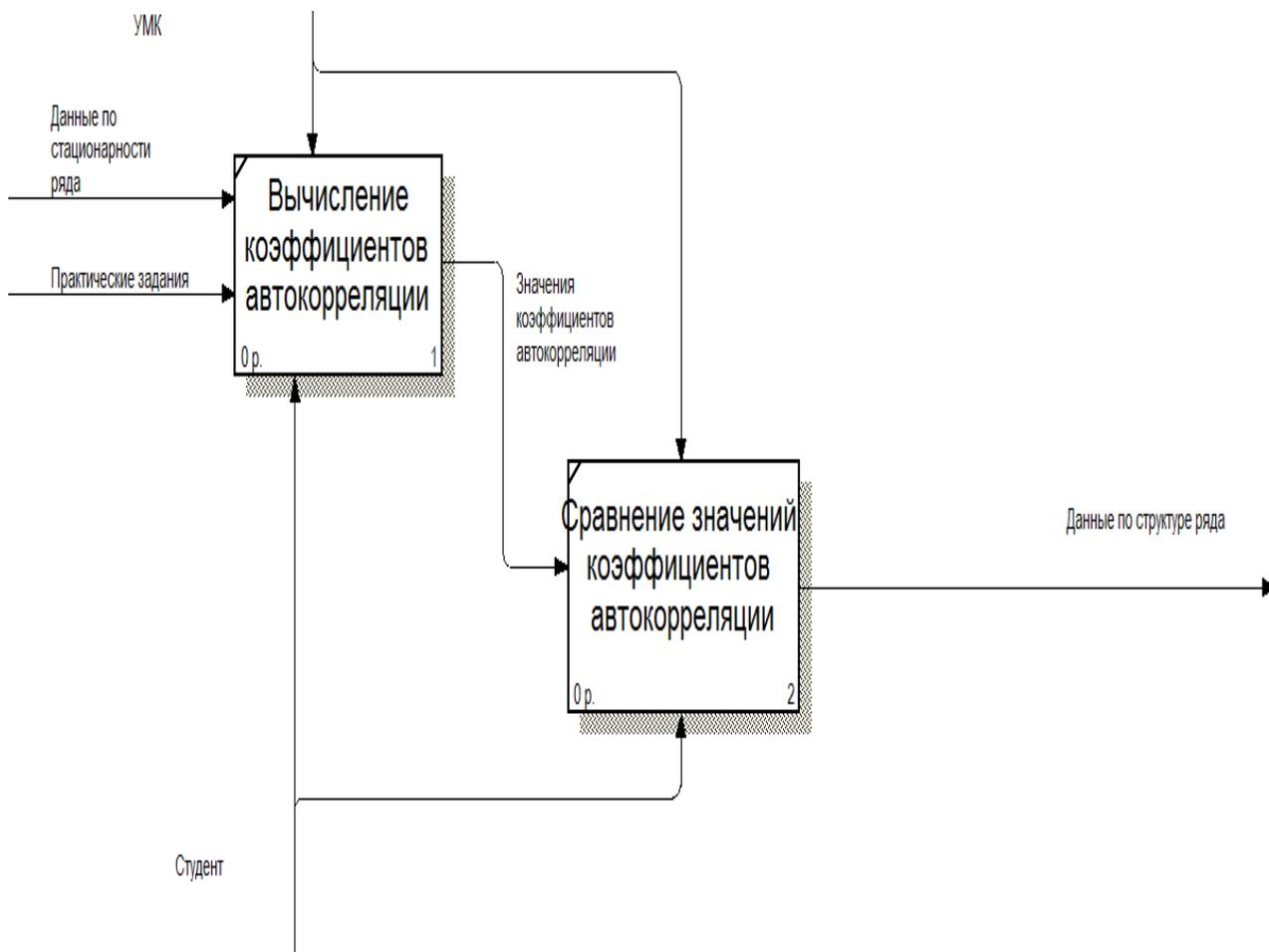


Рисунок 1.11 – Диаграмма декомпозиции процесса «Определение структуры ряда»

Определение структуры тренда может быть разбито на следующие подпроцессы:

- вычисление коэффициентов автокорреляции;
- сравнение значений коэффициентов автокорреляции.

Взаимосвязь перечисленных подпроцессов была изложена на диаграмме декомпозиции процесса «Подведение итогов», представленной на рисунке 1.12.

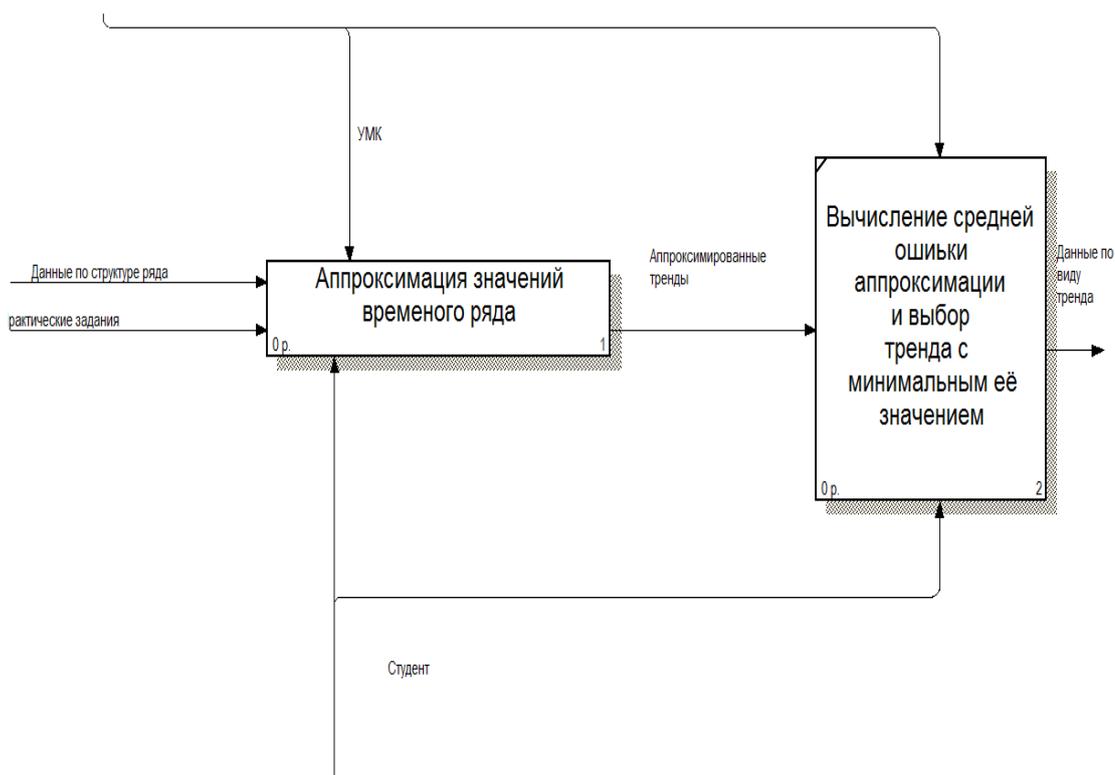


Рисунок 1.12 – Диаграмма декомпозиции процесса «Определение вида тренда»

Аппроксимация значений временного ряда может быть разбита на следующие подпроцессы:

- линейная аппроксимация;
- экспоненциальная аппроксимация;
- гиперболическая аппроксимация;
- степенная аппроксимация;
- полиномиальная аппроксимация второго порядка.

Взаимосвязь перечисленных подпроцессов была изложена на диаграмме декомпозиции процесса «Аппроксимация значений временного ряда», представленной на рисунке 1.13.

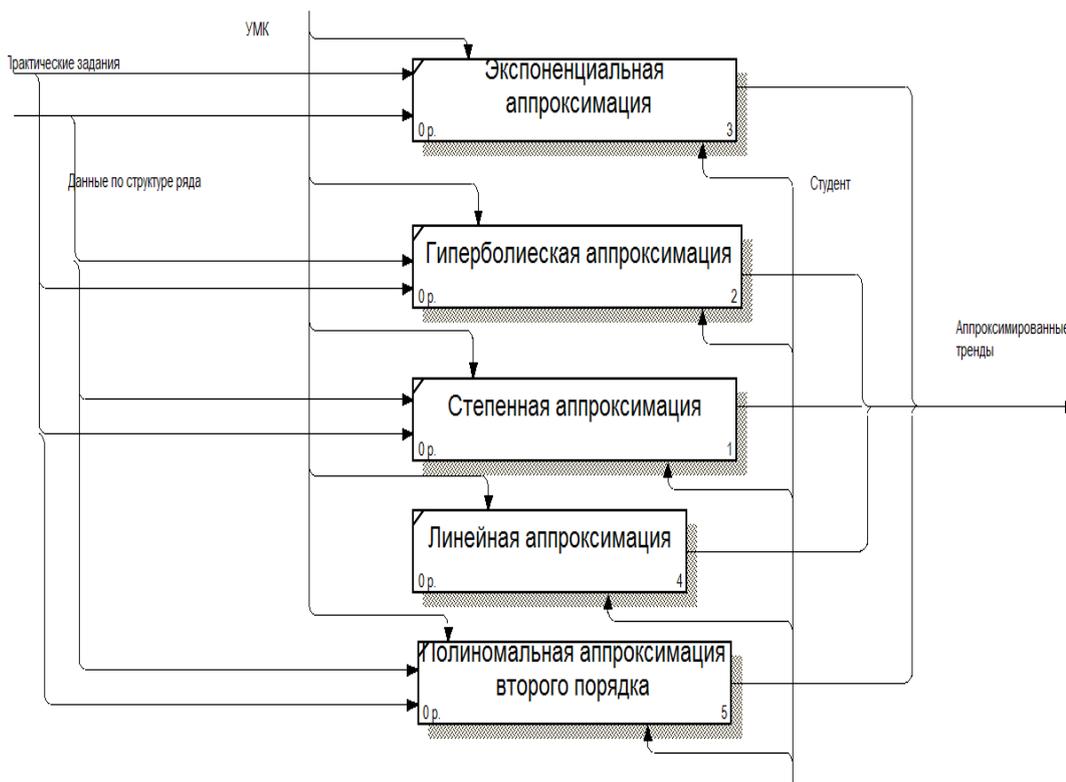


Рисунок 1.13 – Диаграмма декомпозиции процесса «Аппроксимация значений временного ряда»

В процессе обучения предполагается применять следующие виды сглаживания:

- сглаживание по алгоритму МСС $m=1$;
- сглаживание по алгоритму МСС $m=2$;
- сглаживание по алгоритму МСС $m=3$;
- сглаживание по методу Брауна;

При этом предполагается выполнять одновременно не более 1 вида сглаживания, что подразумевает выбор между ними. Для отображения этого, диаграмма декомпозиции процесса «Сглаживание ряда» была выполнена в нотации IDEF3. Данная диаграмма представлена на рисунке 1.14.

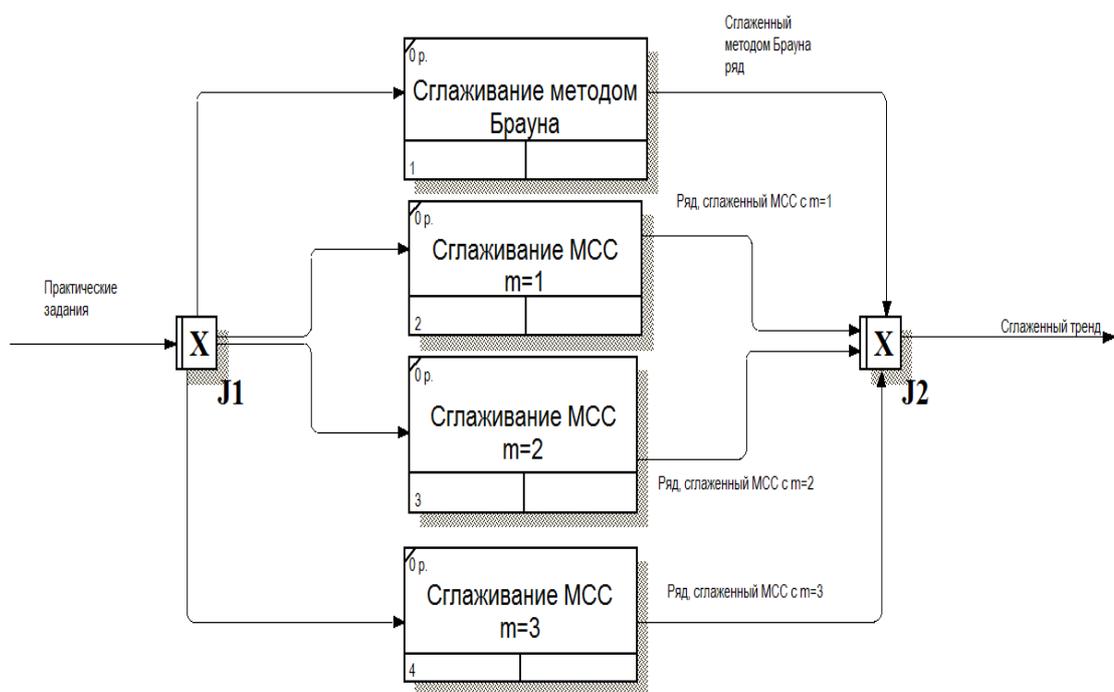


Рисунок 1.14 – Диаграмма декомпозиции процесса «Сглаживание ряда»

Описанная схема выполнения анализа и прогнозирования значений временных рядов достаточно минималистична, однако, при выполнении этих операций вручную или при помощи не специализированных программных пакетов встаёт вопрос о пояснении всех выполняемых операций, так как простое шаблонное повторение действий без вывод неэффективно с точки зрения закрепления полученных навыков, что является ключевым в процессе обучения.

1.4 Постановка задачи

По результатам анализа существующих эконометрических пакетом можно прийти к выводу, что они не в полной мере отвечают критериям, являющимся ключевыми при обучении в университете. В большей части

эконометрических пакетов не предусмотрено русской локализации, что усложняет их освоение. Также следствием широких функциональных этих пакетов являются большие затраты при разработке, что обуславливает их высокую стоимость при приобретении. Данный функционал является избыточным при обучении раздела «Временные ряды» базового курса дисциплины «Эконометрика». Однако ключевым недостатком данных пакетов в рамках процесса обучения является отсутствие пояснений полученных результатов.

Все перечисленные факторы делают разработку нового программного пакета оптимальным решением, обеспечивающим сопровождение процесса обучения. Это, в свою очередь, ставит задачу по реализации следующего функционала:

- загрузка и сохранение данных в файл;
- анализ временного ряда на стационарность;
- анализ структуры ряда;
- определение вида тренда;
- построение тренда;
- сглаживание ряда методом Брауна, МСС при $m=1,2,3$;
- прогнозирование значений ряда по оптимальному тренду;
- вывод пояснений для каждой стадии анализа.

Выводы по первому разделу

В рамках данного раздела был произведён обзор основных методов анализа временных рядов, рассмотрены и сравнены по существенным в рамках процесса обучения критериям между собой и с разрабатываемым программным пакетом. Было осуществлено построение модели процесса обучения «Как есть» и выявлены его недостатки. Также была осуществлена постановка задачи и её обоснование.

2 Проектная часть

2.1 Модель «Как должно быть»

Контекстная диаграмма «Как должно быть» представлена на рисунке 2.1.



Рисунок 2.1 – Контекстная диаграмма

Диаграмма декомпозиции процесса «Процесса обучения дисциплине Эконометрика» (Как должно быть) представлена на рисунке 2.2.



Рисунок 2.2 – Диаграмма декомпозиции процесса «Процесса обучения дисциплине Эконометрика» (Как должно быть)

Диаграмма декомпозиции процесса «Практическое обучение» как должно быть, представлена на рисунке 2.3.

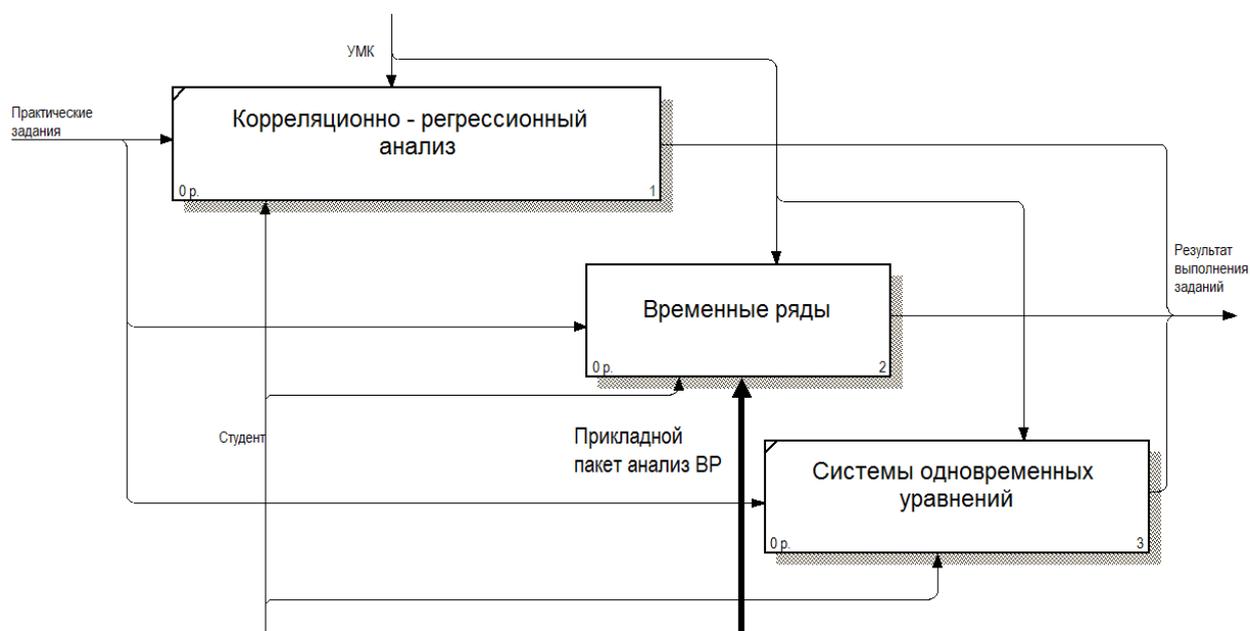


Рисунок 2.3 – Диаграмма декомпозиции процесса «Практическое обучение» (Как должно быть)

В программа предполагается иметь 8 блоков, блок загрузки и выгрузки данных: блок анализа на стационарность, блок определения структуры ряда. блок определения вида тренда. блок подведения итогов, блок сглаживания значений ряда, блок прогнозирования значения тренда.[23]

Данная структура программы проиллюстрирована рисунком 2.5.

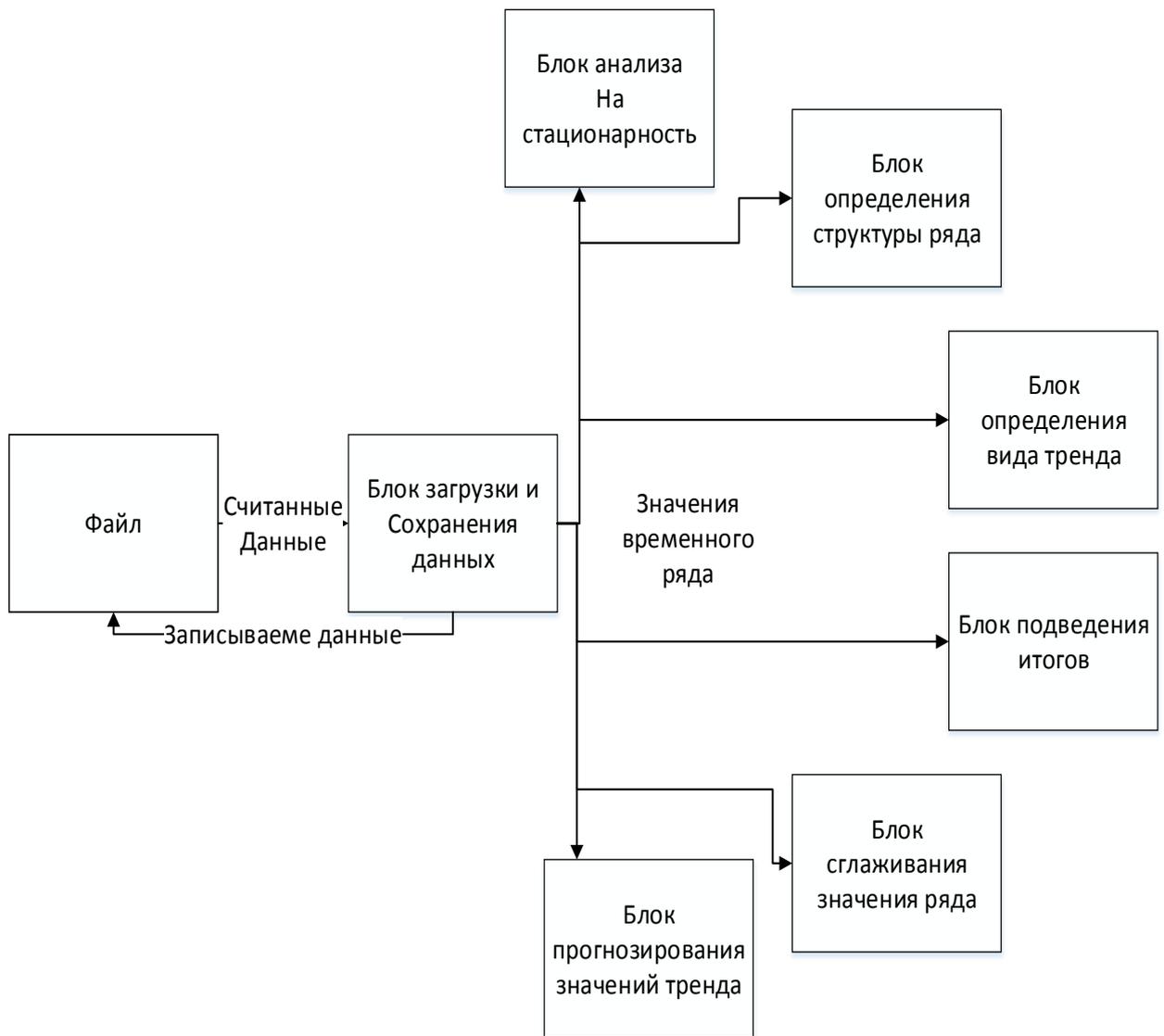


Рисунок 2.5 – Структура программы

При работе программы её взаимодействие с внешними файлами осуществляется через специализированный блок работы с файлами. Данный

блок осуществляет чтение и сохранение временного ряда из файла, а также передачу данных в остальные блоки программы.

Блок анализа на стационарность осуществляет проверку ряда на стационарность методом выборочной медианы.

Блок определения структуры ряда осуществляет вычисление коэффициентов автокорреляции и осуществляет выводы по структуре ряда, основываясь на том, коэффициент какого порядка является максимальным.

Блок определения вида тренда осуществляет сведение исходного временного ряда к линейному, экспоненциальному, гиперболическому, степенному и полиномиальному.

В блоке построения тренда осуществляется построение графиков по сведённым трендам.

В блоке подведения итогов вывод данных по стационарности и структуре ряда, а также рассчитываются коэффициенты автокорреляции и детерминации.

В блоке сглаживания значения тренда осуществляется сглаживание методом Брауна, методом скользящей средней с $m=1$, методом скользящей средней с $m=2$, методом скользящей средней с $m=3$.

В блоке прогнозирования осуществляется расчёт прогнозных значений временного ряда по определённому оптимальному тренду и вывод рассчитанных значений в виде графика.

2.3 Алгоритм работы

Общий алгоритм работы программы представлен на рисунке 2.6.

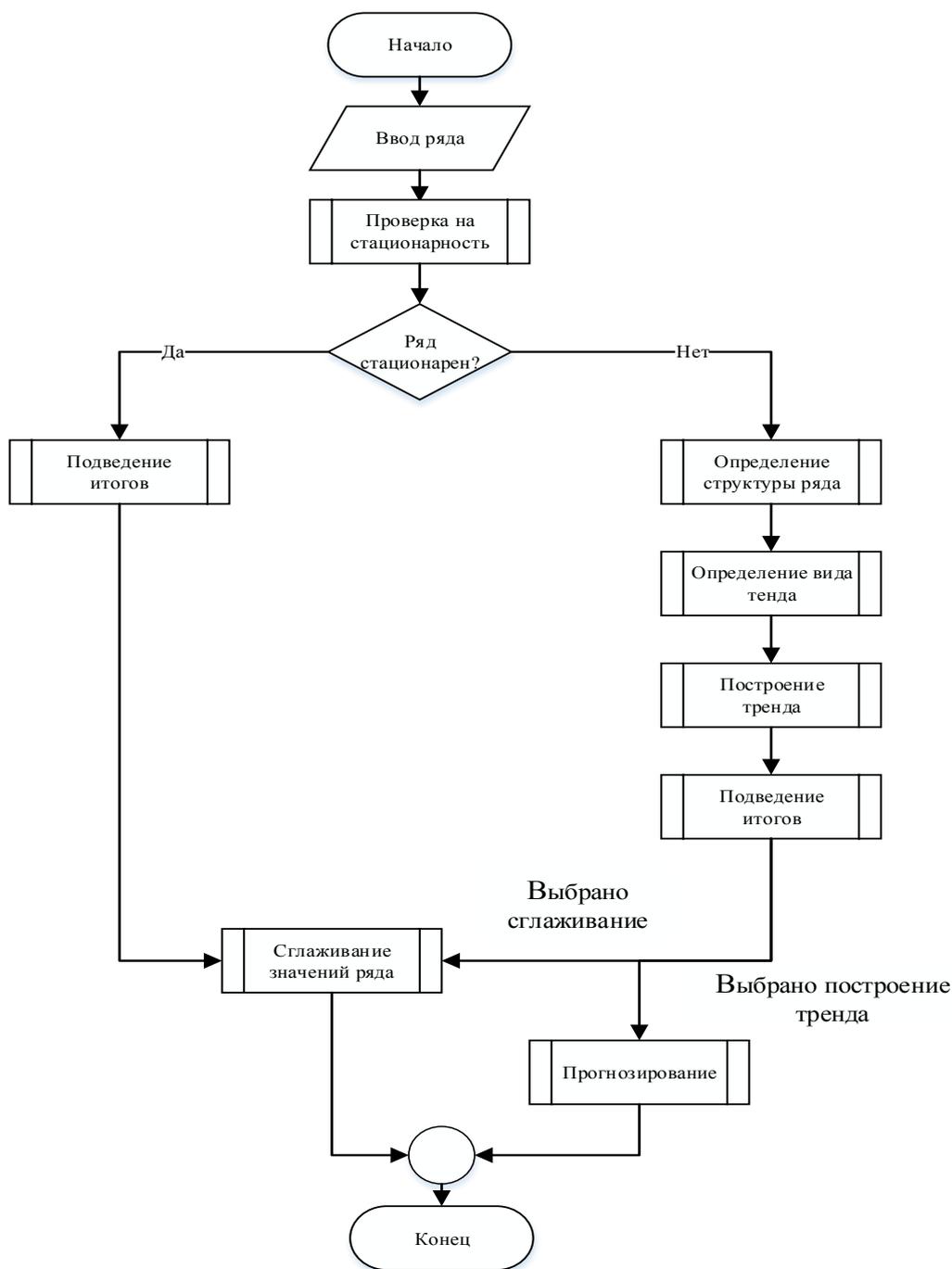


Рисунок 2.6 – Общий алгоритм работы

В соответствии с общим алгоритмом работы прикладной пакет в ходе своей работы выполняет следующие действия: осуществляется ввод значений анализируемого временного ряда, после чего происходит проверка на стационарность временного ряда на стационарность. Если проверка ряда на стационарность показала, что ряд стационарен, то сразу же происходит вывод итогов анализа, а также даётся возможность провести сглаживание значений

временного ряда. Если же ряд оказался нестационарным, то следующим этапом становится определение структуры ряда. После определения структуры ряда осуществляется определение вида тренда и его построение. После происходит подведение итогов с условием, что ряд не является стационарным, после чего становится возможным осуществить сглаживание значений временного и/или значений тренда.

Алгоритм выполнения подпроцесса «Проверка на стационарность» представлен на рисунке 2.7.

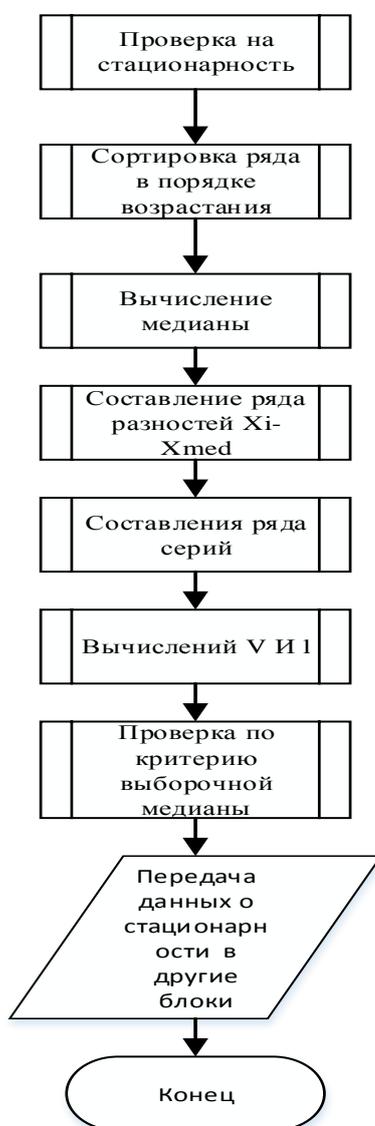


Рисунок 2.7 – Алгоритм проверки на стационарность

В соответствии с алгоритмом при проверке на стационарность осуществляется следующая последовательность действий: вначале происходит сортировка временного ряда по возрастанию. После чего осуществляется расчёт медианного значения временного, равного или срединному элементу отсортированного ряда, если количество членов ряда нечётное, либо среднему арифметическому от двух срединных элементов ряда, если количество членов ряда чётное. После это выполняется составление ряда из разниц $X_i - X_{med}$, что в дальнейшем необходимо для построения ряда серий. При составлении ряда серий значение соответствующего элемента ряда серии приравнивается к -1, если элемент ряда разниц меньше 0, а к 1, если – больше, и пропускается, если равно. После чего определяется число серий и их наибольшая длина. После этого по критерию выборочной медианы выносится заключение о стационарности или нестационарности временного ряда.

Алгоритм выполнения подпроцесса «Определение структуры ряда» представлен на рисунке 2.8.

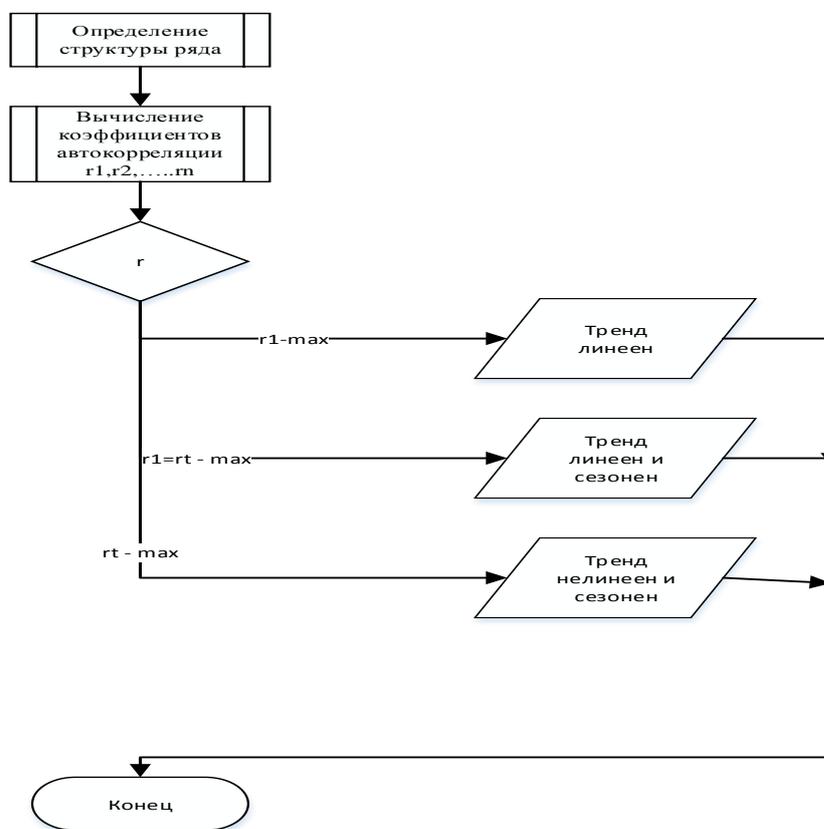


Рисунок 2.8 – Определение структуры ряда

В соответствии с алгоритмом при проверке на стационарность происходит вычисление коэффициентов автокорреляции вплоть до порядка $n/2$, где n – число элементов изначального ряда. После чего происходит определение максимального коэффициента автокорреляции. Если максимальным оказался коэффициент первого порядка, то выводится информация о том, что ряд имеет линейный тренд, если r_1 и r_t максимальны, то тренд тоже линейен, однако имеет сезонную составляющую, если же наибольшим оказывается r_t , то ряд имеет нелинейный тренд с сезонной составляющей или просто сезонные колебания, поэтому нужна дополнительная проверка на наличие нелинейного тренда.

Алгоритм выполнения подпроцесса «Определение вид тренда» представлен на рисунке 2.9.



Рисунок 2.9 – Определение вида тренда

В соответствии с алгоритмом при определении вида тренда последовательности происходят линейные, экспоненциальные, гиперболические, степенные и полиномиальные аппроксимации. После чего осуществляется построение и вывод этих трендов.

Алгоритм выполнения подпроцесса «Подведение итогов» представлен на рисунке 2.10.

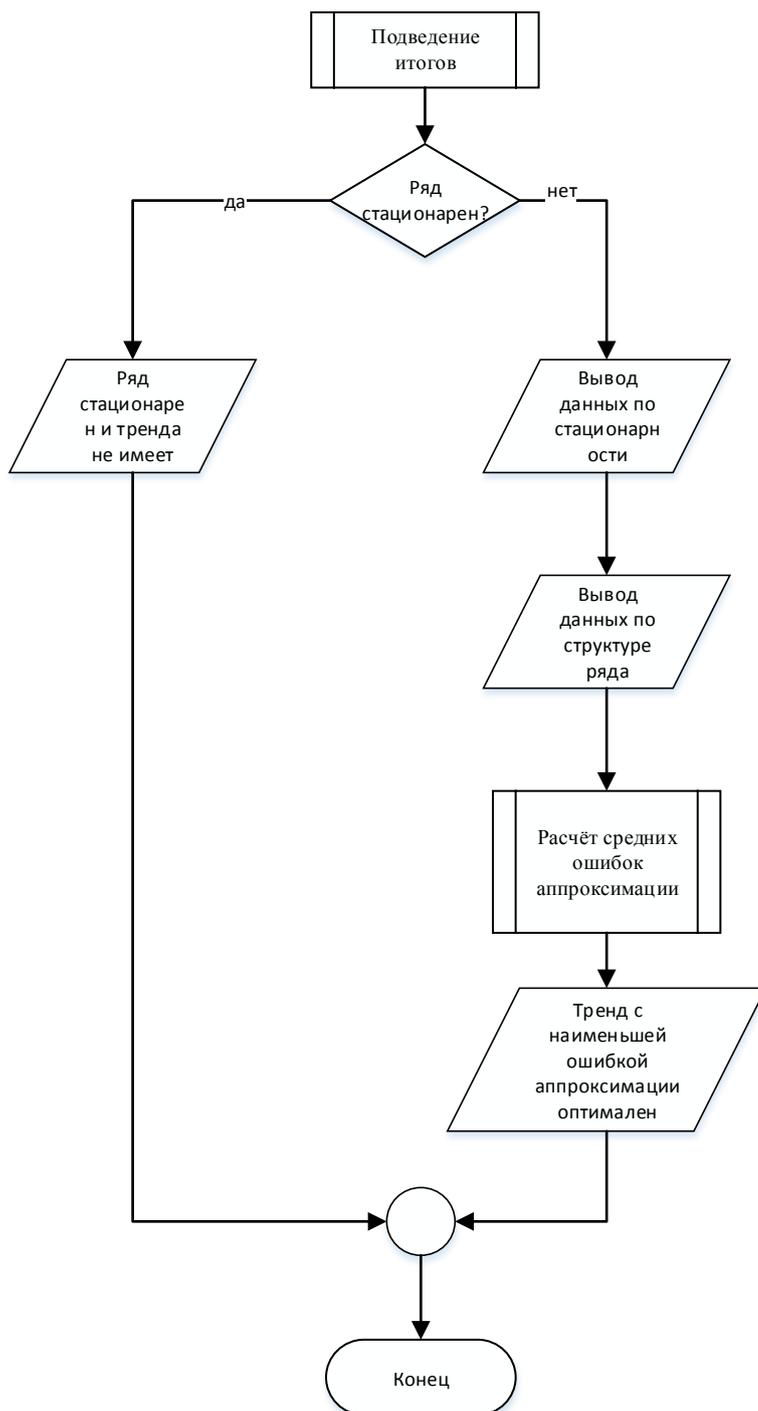


Рисунок 2.10 – Алгоритм подведения итогов

В соответствии с алгоритмом при подведении итогов происходит вывод информации о том, что ряд стационарен, если соответствующие данные пришли из блока анализа ряда на стационарность. Если же пришедшие данные говорят о нестационарности ряда, то данная информация выводится для пользователя. После чего осуществляется вывод информации по расчётной структуре ряда. Следующим этапом является расчёт средней ошибки аппроксимации для рядов, составленных в блоке определения вида тренда. После чего происходит вывод информации о тренде с наименьшей средней ошибкой аппроксимации как о наиболее точно описывающем данный случайный Процесс.

Алгоритм выполнения подпроцесса «Сглаживание ряда» представлен на рисунке 2.11.

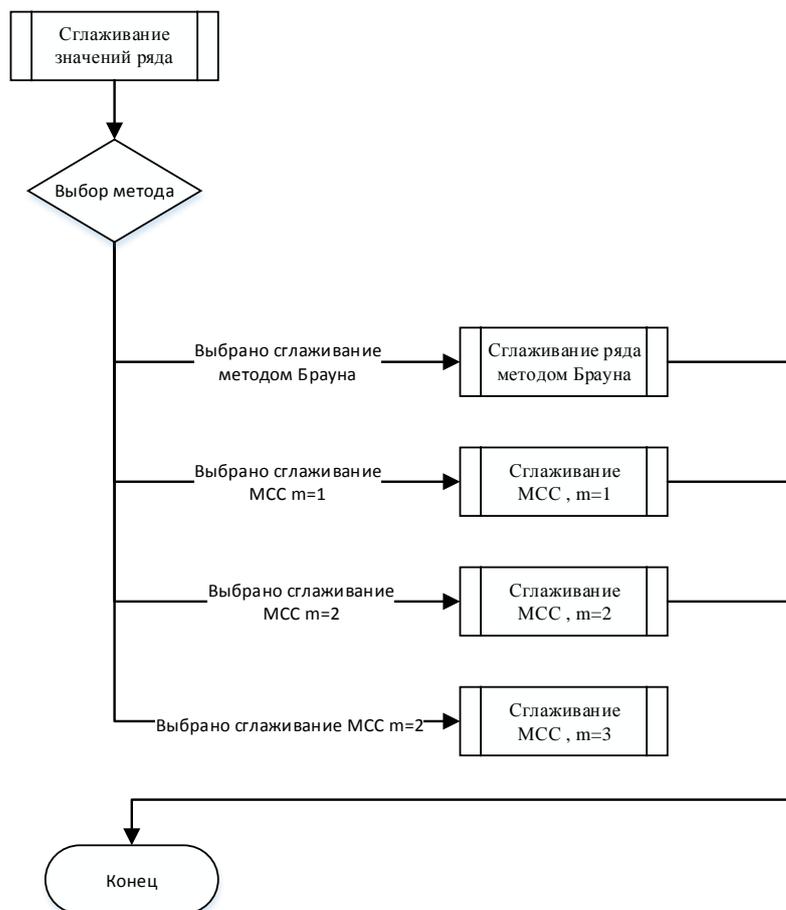


Рисунок 2.11 – Алгоритм сглаживания

В соответствии с алгоритмом при подведении проведения сглаживания существует возможность выбрать способ сглаживания из следующих: сглаживание методом Брауна, сглаживание МСС при $m=1$, сглаживание МСС при $m=2$, сглаживание МСС при $m=3$.

Последним из основных подпроцессов является процесс прогнозирования, алгоритм которого представлен на рисунке 2.12.

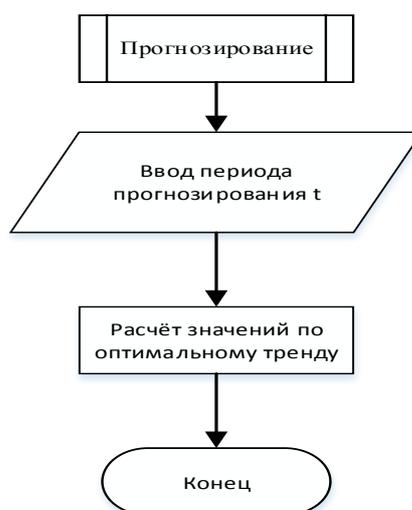


Рисунок 2.12 – Алгоритм сглаживания

В соответствии с алгоритмом при подведении проведения сглаживания сперва происходит ввод периода прогнозирования, после чего рассчитываются дополнительно t значений по оптимальному тренду, определённого при подведении итогов анализа.

Выводы по второму разделу.

В рамках данного раздела было проведено моделирование процесса обучения «Как должно быть», построение структуры программы и разработка алгоритма её работы, включая проверку ряда на стационарность, определение структуры ряда, определение вида тренда, построение тренда, подведение итогов, сглаживание по 4 – м алгоритмам и прогнозирование значений временного ряда.

3 Разработка и тестирование программного пакета и оценка эффективности

3.1 Разработка пакета

При разработке в соответствии со структурой программы на главной форме приложения были размещены вкладки:

- временной ряд;
- стационарность;
- структура ряда;
- вид тренда;
- итог анализа;
- сглаживание;
- прогнозирование.

Каждая из данных вкладок соответствует определённому блоку программы [3]. При запуске программы происходит открытие вкладки «временной ряд» разработанного приложения, что проиллюстрировано рисунком 3.1.



Рисунок 3.1 – Вкладка «временной ряд»

На данной вкладке располагается элемент DataGridView, позволяющий пользователю вводить и редактировать значения уровней ряда динамики. В контекстном меню имеется 2 пункта: файл и провести анализ. Пункт файл имеет подпункты «Открыть», «Сохранить» и «Сохранить как», позволяющие осуществлять загрузку и выгрузку данных [33]. При инициализации открытия пользователю становится доступно диалоговое окно, представленное на рисунке 3.2.

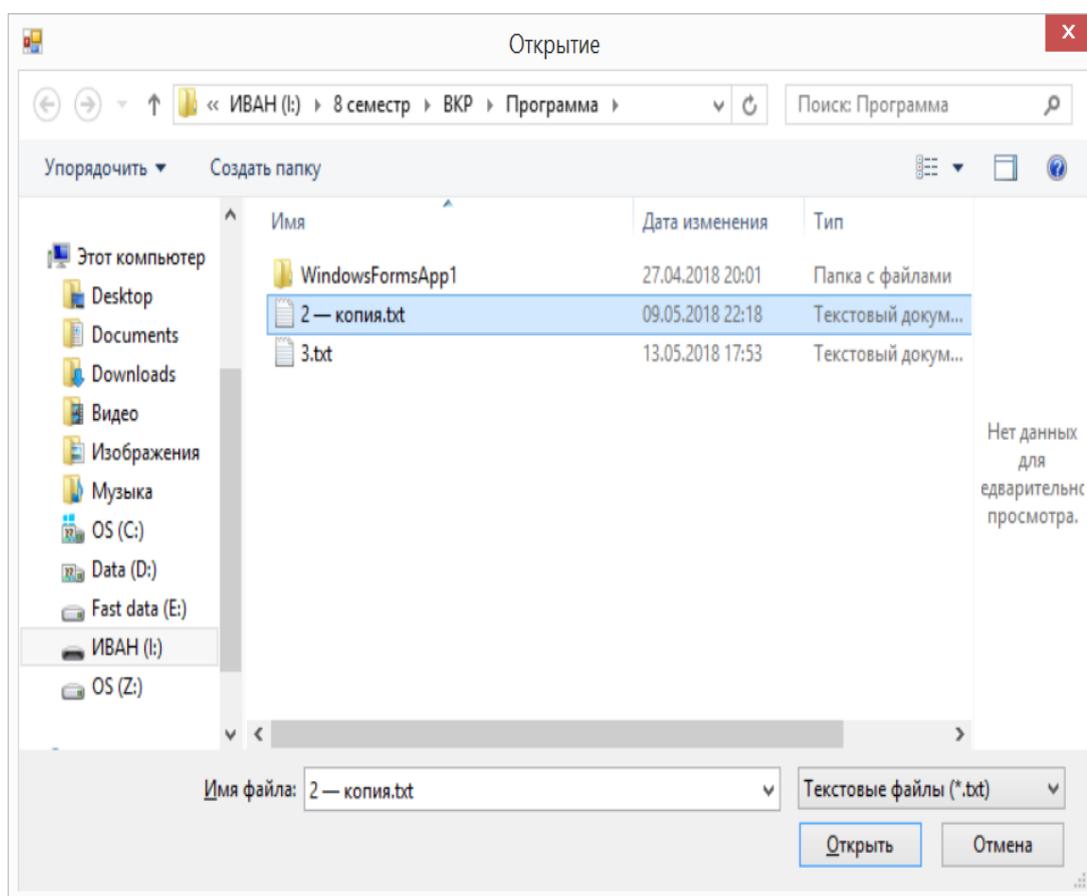


Рисунок 3.2 – Окно открытия файла

Данное окно позволяет выбрать txt-файл, в котором были записаны значения временного ряда. При успешном открытии файла меню приобретает вид, представленный на рисунке 3.3.

Момент времени	Значение
1	1
2	5
3	6
4	8
5	4
6	15
7	25
8	63
9	7
10	0,1

Рисунок 3.3 – Вкладка «Временной ряд» с загруженными данными

При сохранении файла выводится окно, схожее с окном на рисунке 3.2. При успешном сохранении выводится сообщение об успешной операции, представленное на рисунке 3.4.

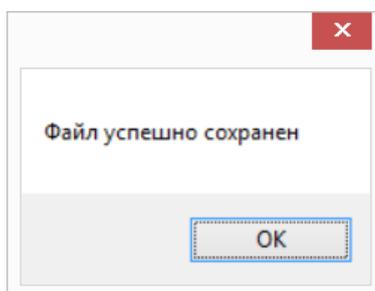


Рисунок 3.4 – Сообщение о сохранении

При инициализации анализа происходит проверка значений временного ряда. Если элемент временного ряда окажется не числом, то будет выведено соответствующее сообщение, представленное на рисунке 3.5.

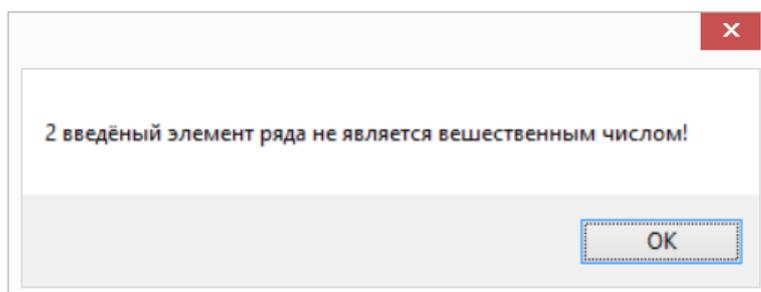


Рисунок 3.5 – Сообщение о несоответствии элемента ряда

После того как анализ был инициализирован, пользователю станут доступны вкладки «Стационарность», «Итоги» и «Сглаживание». Если при проверке на стационарность ряд окажется нестационарным, то пользователю станут доступны для просмотра и вкладки «Структура ряда», «Вид тренда», «Построение тренда» и прогнозирование. [4]

Вкладка «стационарность» при проведении анализа введённого временного ряд приняла вид, представленный на рисунке 3.6.

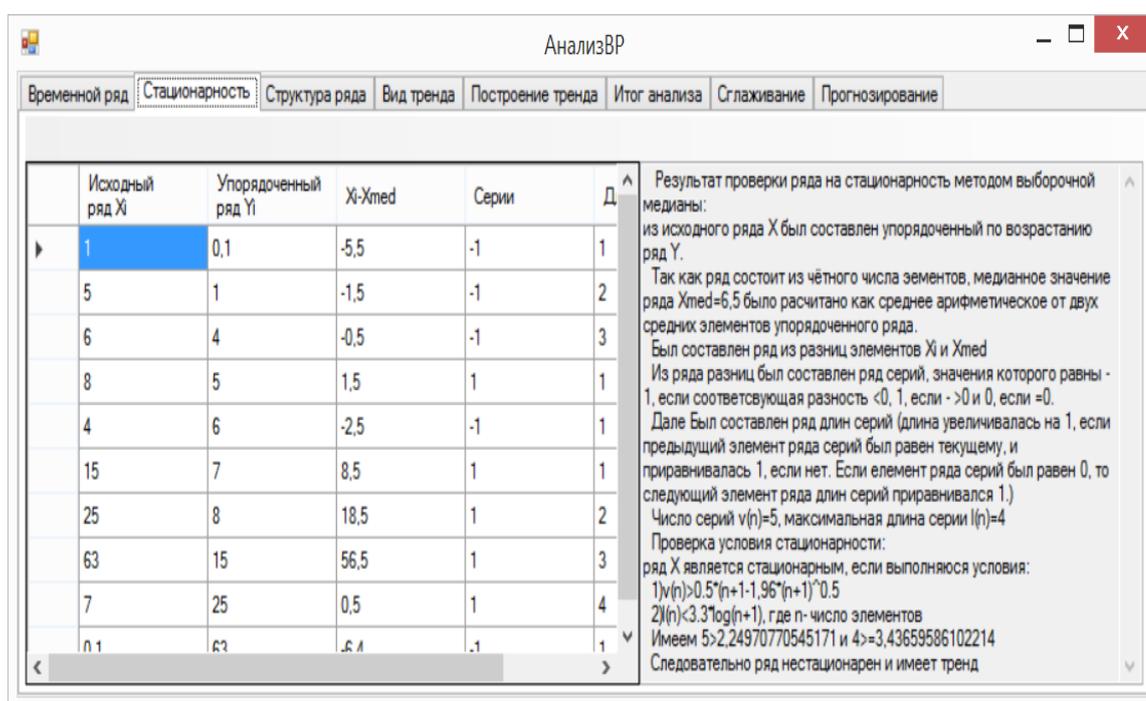


Рисунок 3.6 – Вкладка «Стационарность»

При проведении анализа в таблицу записываются значения исходного и упорядоченного ряда, также разница элемента и медианы, ряд серий, а также длина непрерывной серии для текущего элемента серии. В пояснении выводится информация о медианном значении и о результате проверки ряда на стационарность методом выборочной медианы. Вкладка «Структура ряда» имеет вид, представленный на рисунке 3.7.

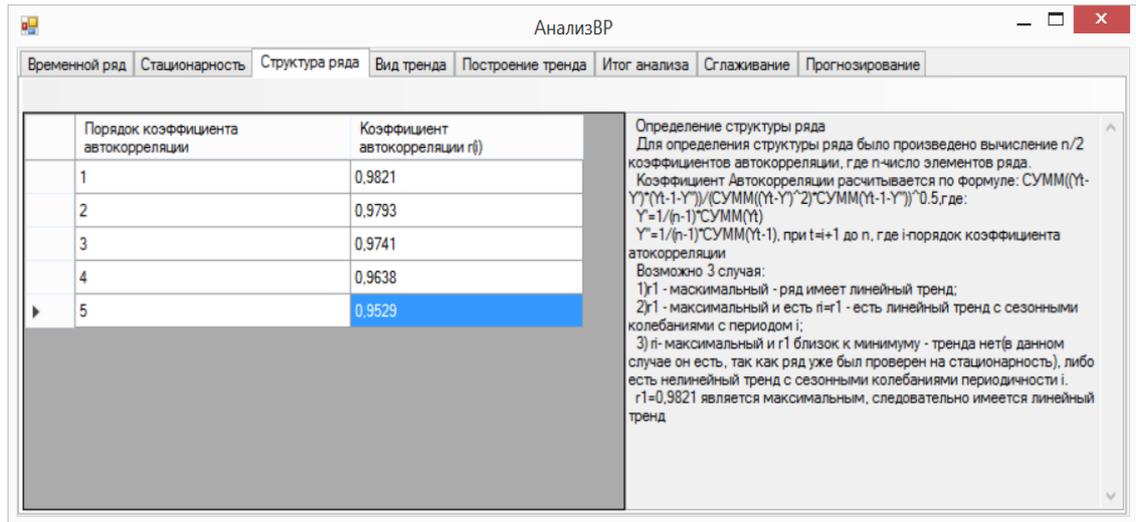


Рисунок 3.7 – Вкладка «Структура ряда»

Данная вкладка содержит информацию о $n/2$ коэффициентах автокорреляции и информацию о структуре ряда, выведенных в соответствии со значениями данных коэффициентов.

Вкладка «Вид тренда» имеет вид, представленный на рисунке 3.8.

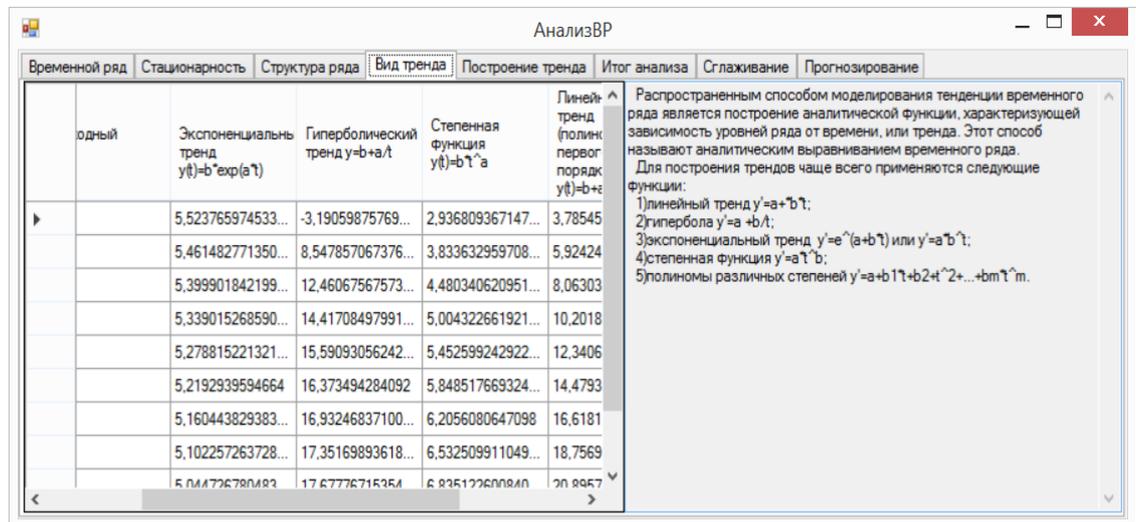


Рисунок 3.8 – Вкладка Вкладка «Вид тренда»

Данная вкладка содержит аппроксимированные значения временного ряда.

Построение данных трендов происходит на вкладке «Построение тренда», представленной на рисунке 3.9.

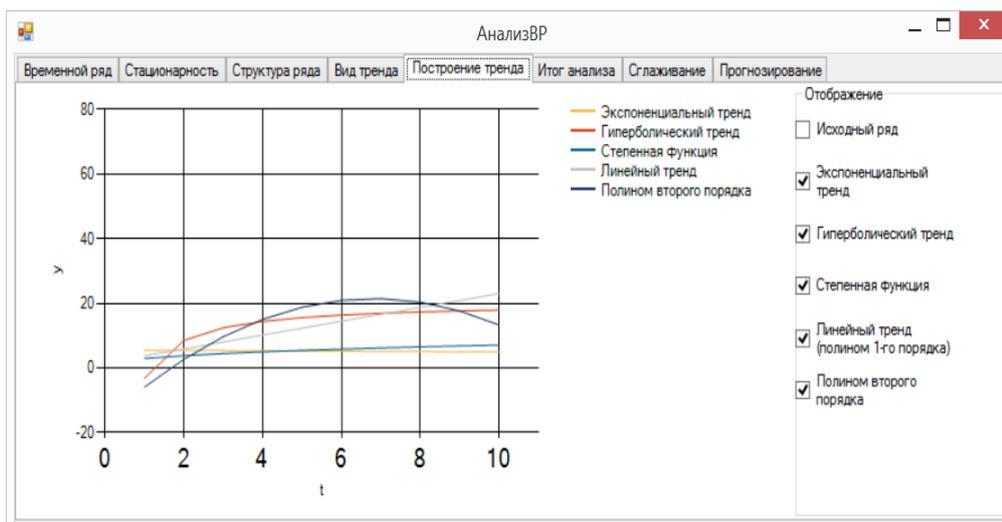


Рисунок 3.9 – Вкладка «Построение тренда»

Данная вкладка содержит графики всех аппроксимированных трендов и исходного ряда. При этом имеется возможность выбирать, какие графики должны быть отражены.

Итог анализа подводится на одноимённой вкладке, представленной на рисунке 3.10.

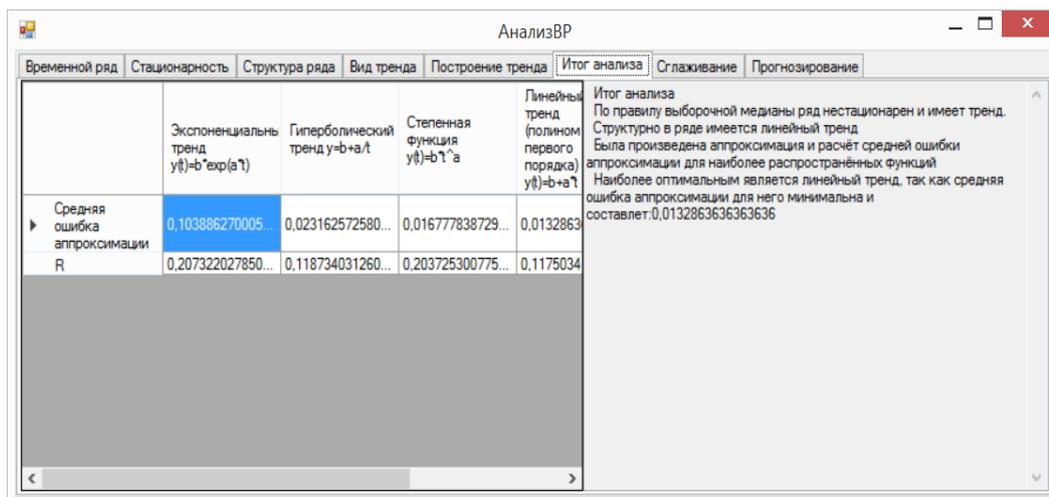


Рисунок 3.10 – Итог анализа

Данная вкладка содержит информации о средней ошибке аппроксимации и значении коэффициентов детерминации. Также на ней происходит выбор наиболее подходящего тренда, соответствующего минимальному значению средней ошибки аппроксимации.

При открытии вкладка «Сглаживание» имеет вид, представленный на рисунке 3.11.

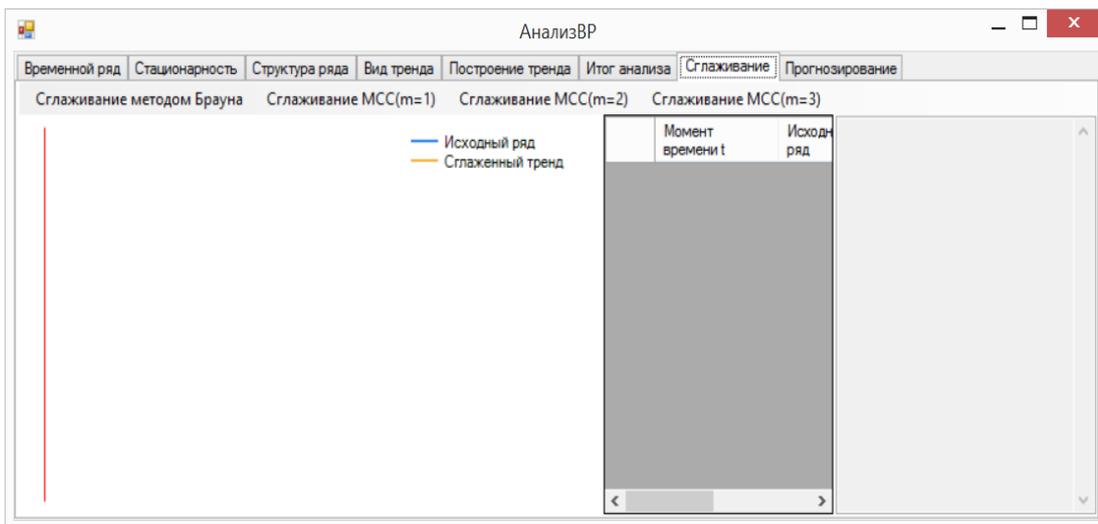


Рисунок 3.11 – Вкладка «Сглаживание»

Данная вкладка позволяет выбрать один из четырёх способов сглаживания значений временного ряда при нажатии на соответствующий элемент меню. Ряд, сглаженный методом Брауна, представлен на рисунке 3.12.

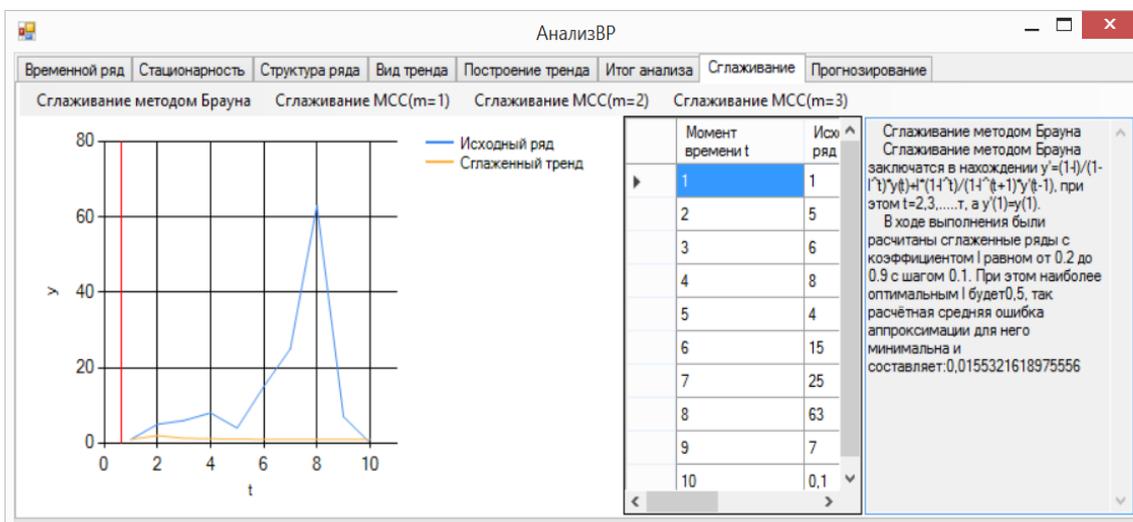


Рисунок 3.12 – Метод Брауна

Ряд, сглаженный МСС при $m=1$, представлен на рисунке 3.13.

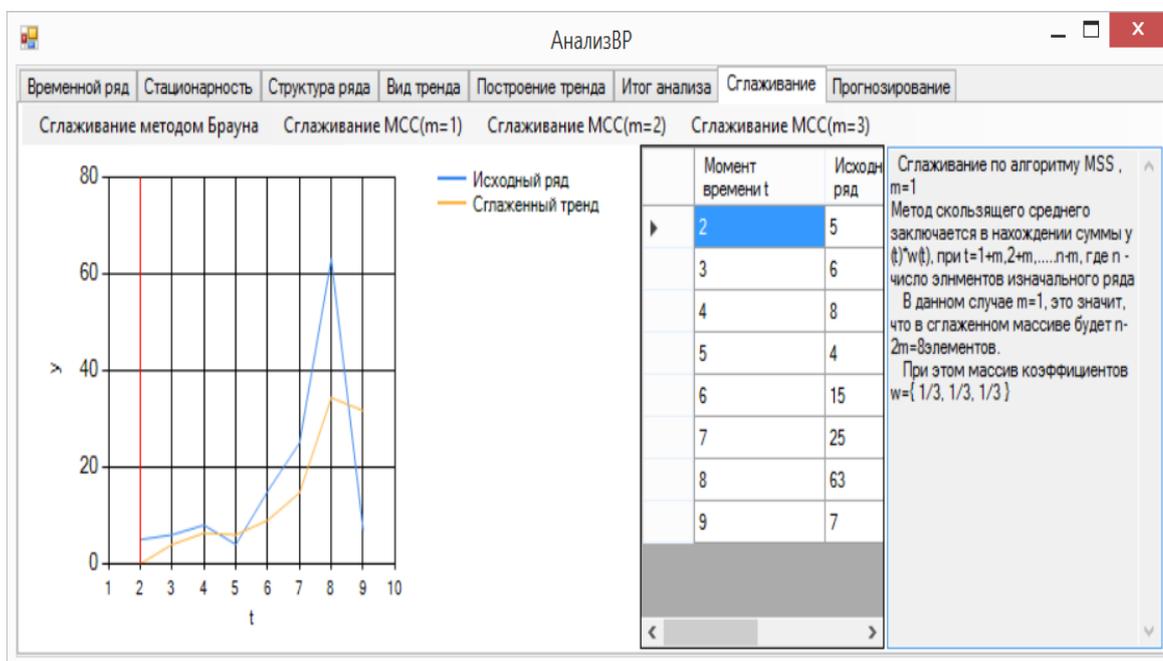


Рисунок 3.13 – Сглаживание МСС, $m=1$

Ряд, сглаженный МСС при $m=2$, представлен на рисунке 3.14.

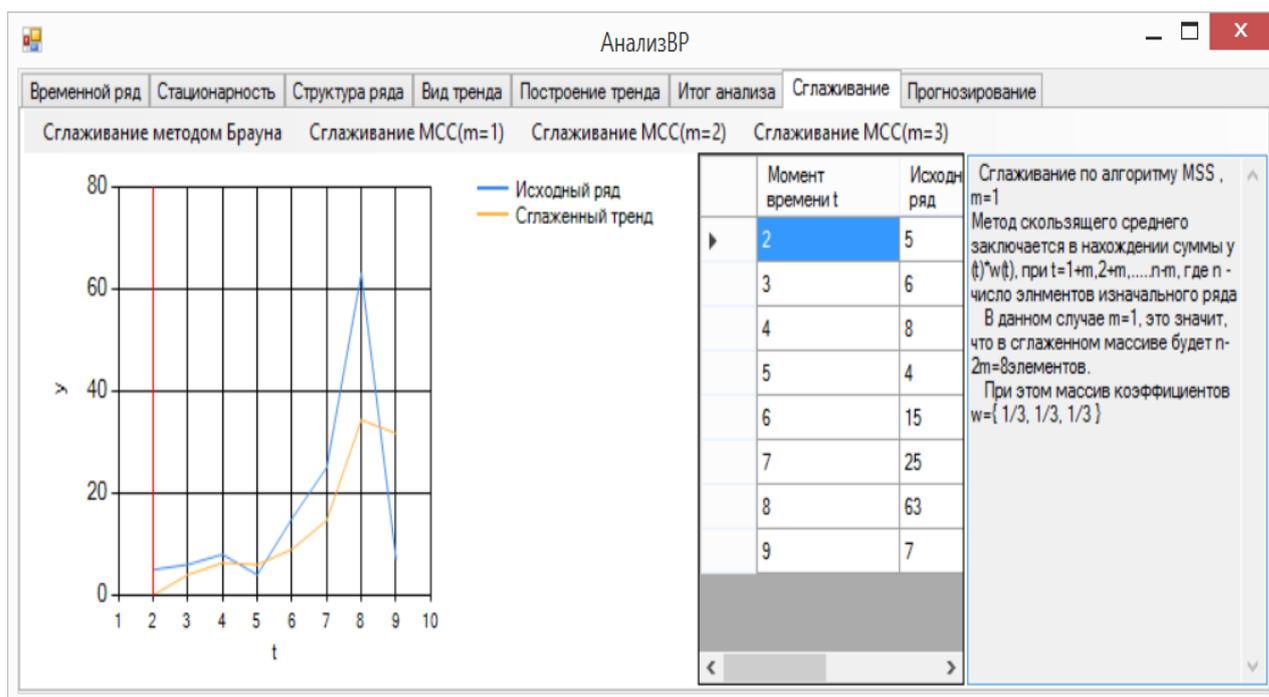


Рисунок 3.14 – Сглаживание МСС, $m=2$

Ряд, сглаженный МСС при $m=3$ представлен на рисунке 3.15.

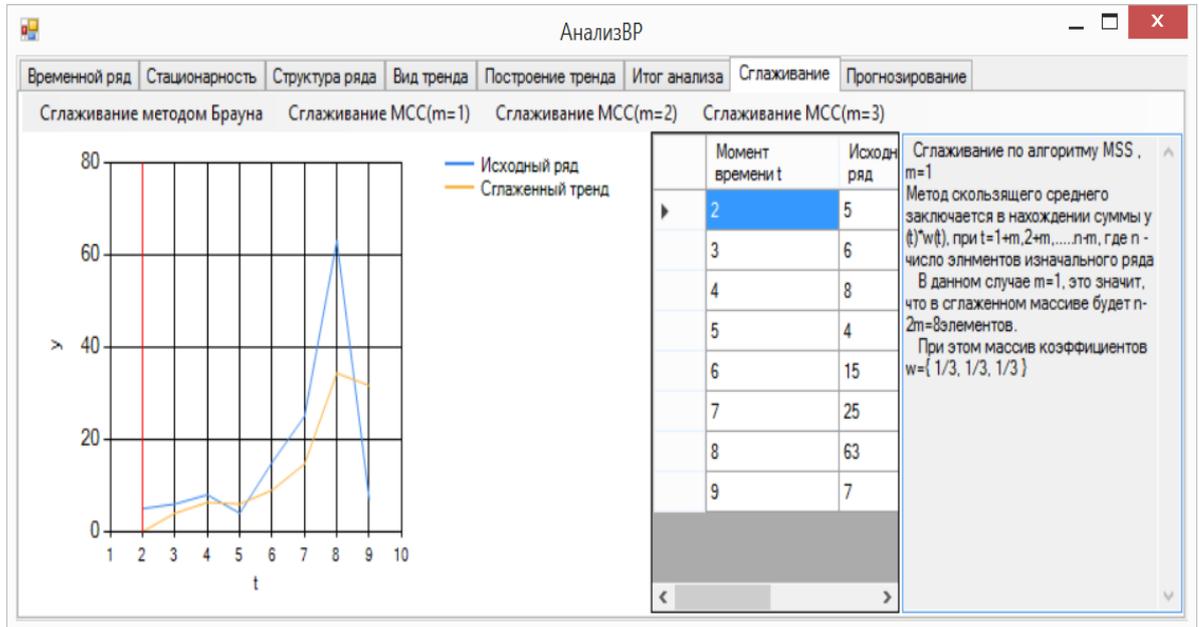
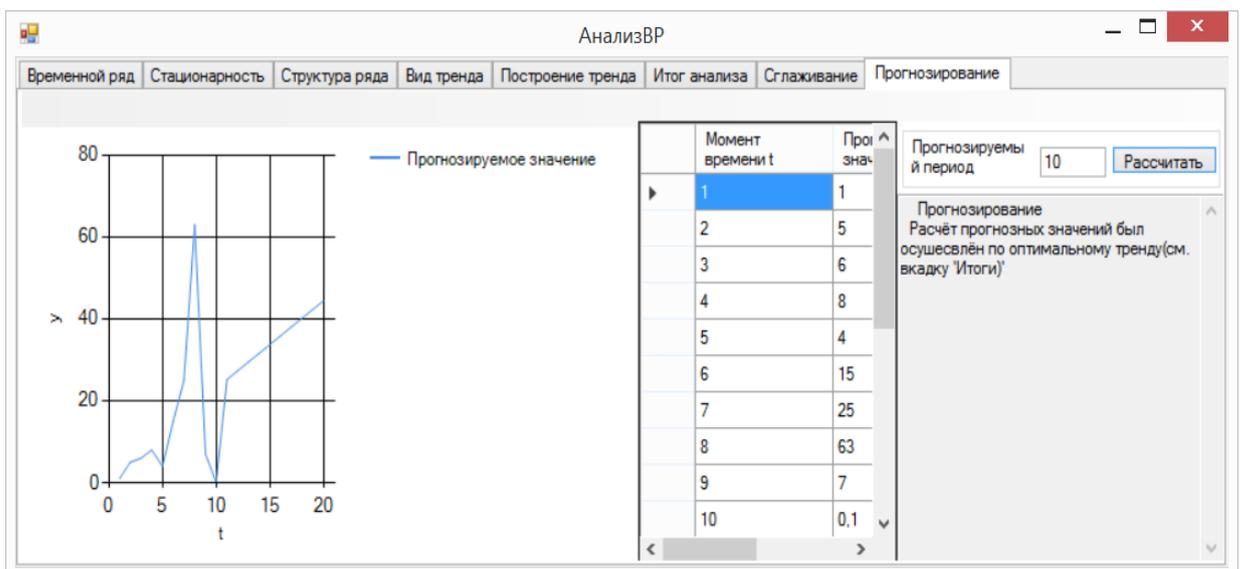


Рисунок 3.15 – Сглаживание МСС, $m=2$

Результат прогнозирования значения ряда для оптимального тренда представлен на рисунке 3.16.



3.16 – Вкладка «Итоги»

3.2 Контрольный пример

В качестве контрольного примера был проведён анализ временного ряда, представленного в таблице 3.1.

Таблица 3.1 – Временной ряд

Момент времени t	Значение временного ряда
1	11,672
2	16,52
3	11,535
4	8,773
5	4,809
6	14,159
7	19,465
8	16,728
9	14,692
10	14,671
11	14,009
12	11,592
13	14,193
14	23,388
15	25,105
16	25,934
17	34,922
18	35,309
19	41,581
20	50,693
21	54,869
22	49,28
23	54,719
24	46,118
25	47,368
26	43,159
27	49,565
28	46,989

29	50,268
30	49,289

Результат анализа этого ряда представлен на рисунке 3.17.

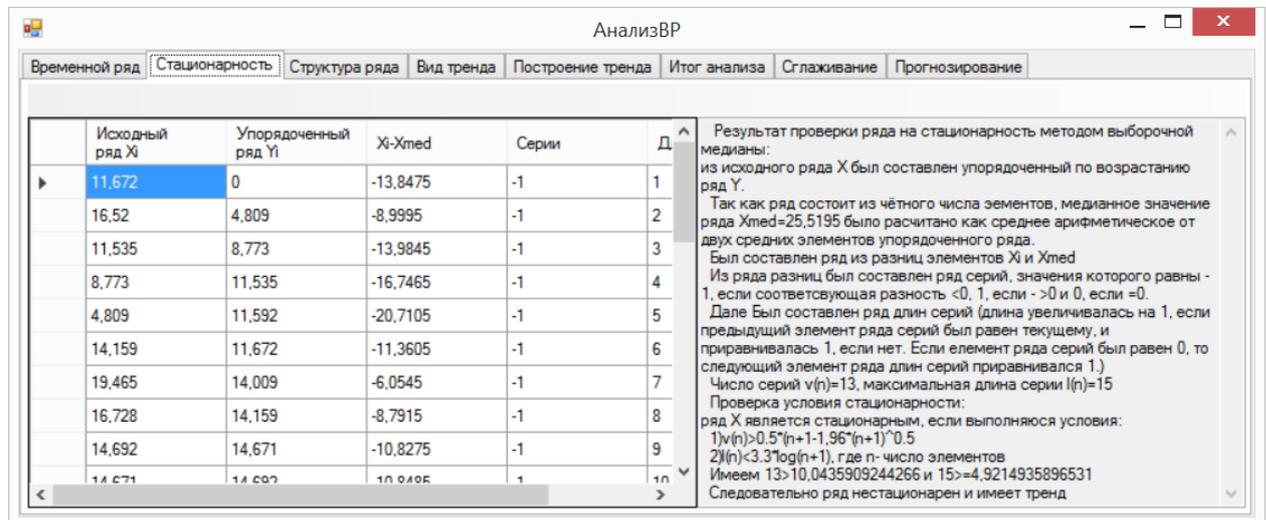


Рисунок 3.17 – Результат анализа ряда на стационарность

Так как второе условие стационарности по критерию выборочной медианы было нарушено, то ряд нестационарен и имеет тренд.

Анализ структуры ряда представлен на рисунке 3.18.

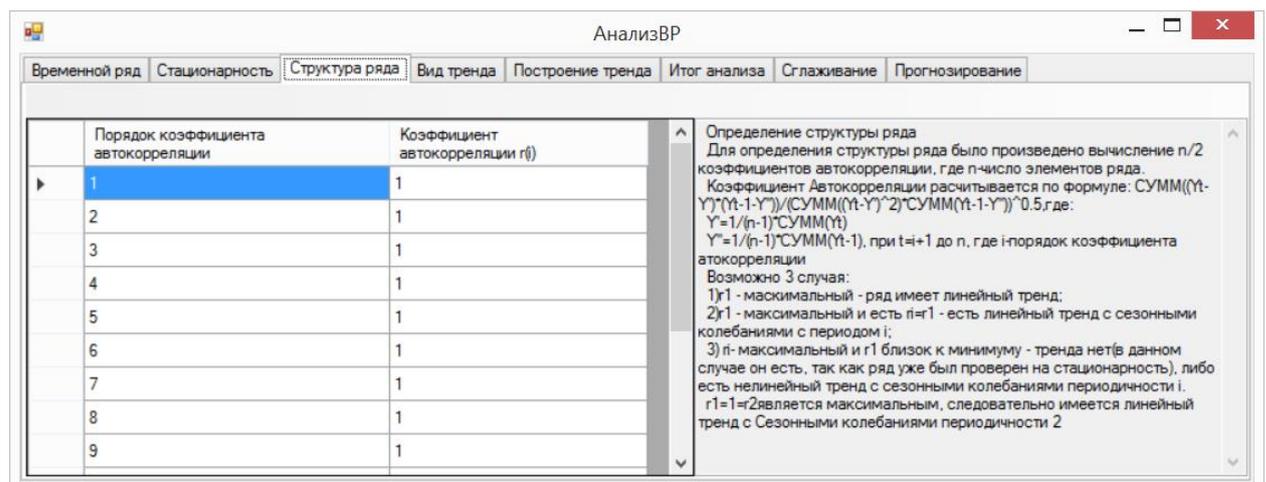


Рисунок 3.18 – Результат анализа структуры ряда

Так как коэффициент автокорреляции первого порядка является максимальным и равным коэффициенту автокорреляции второго порядка, то имеет место линейный тренд с сезонными колебаниями периодичности 2.

Результат определения вида тренда представлен на рисунке 3.19.

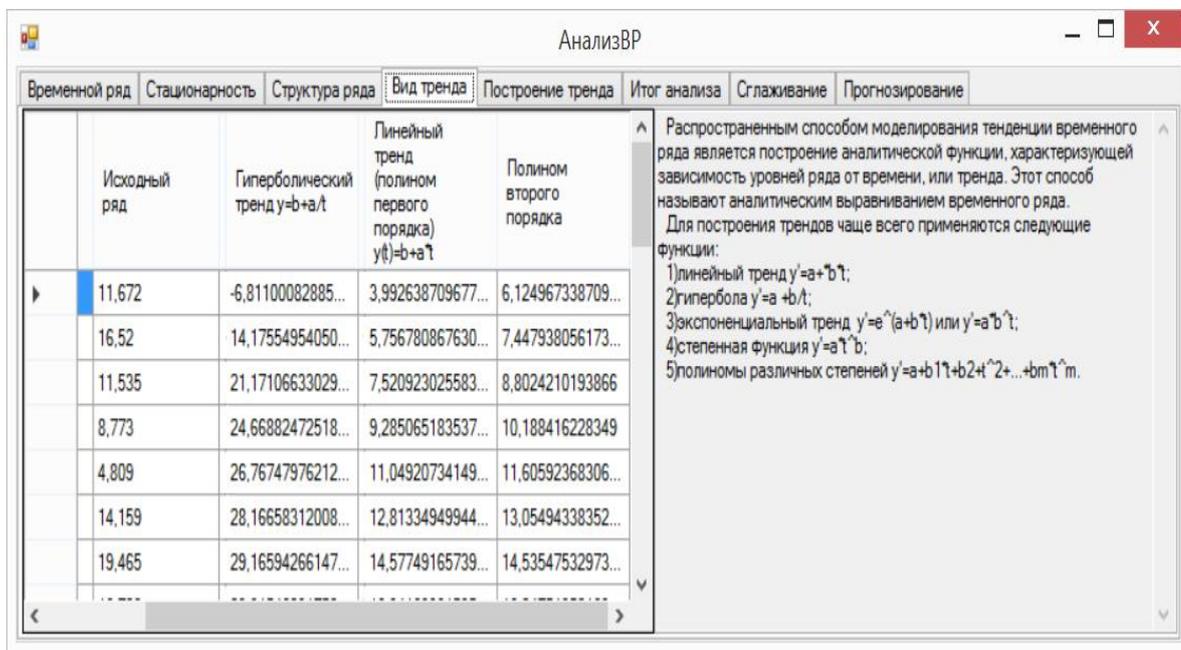


Рисунок 3.19 – Результат определения вида тренда

Результат построения аппроксимированных трендов представлен на рисунке 3.20.

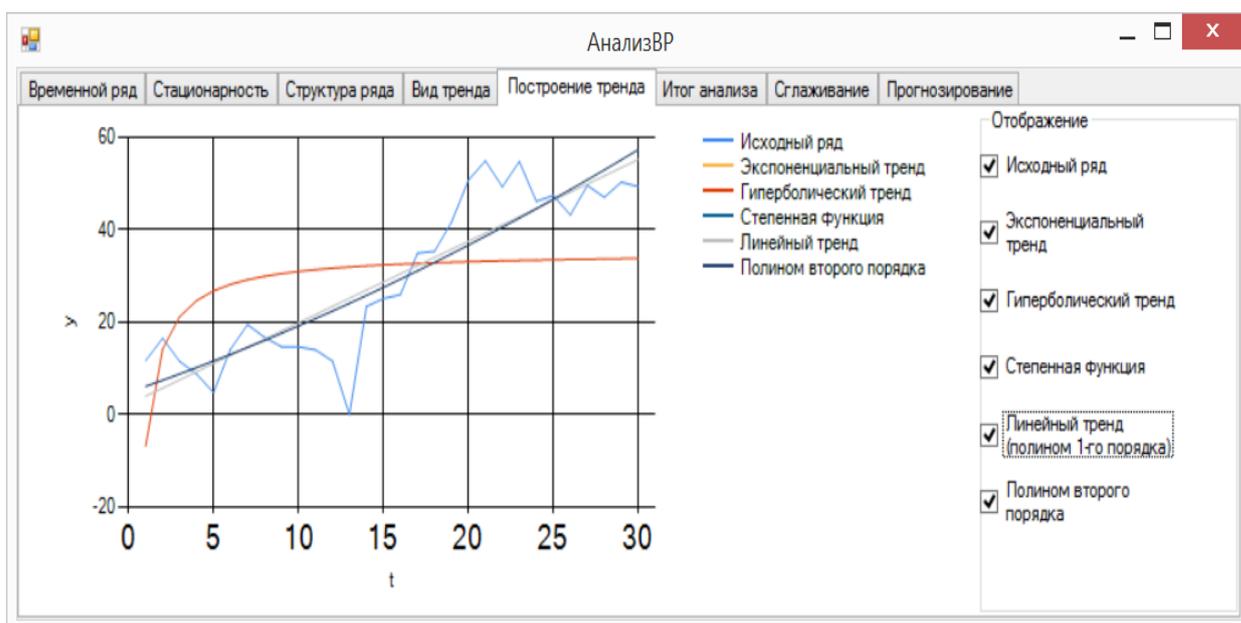


Рисунок 3.20 – Результат определения вида тренда

Итоги анализа представлены на рисунке 3.21.

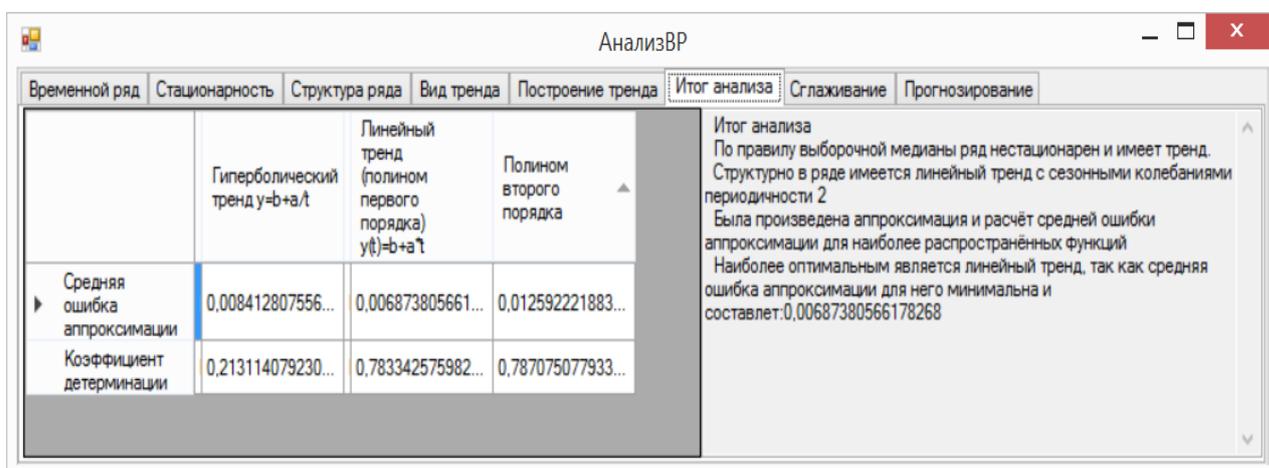


Рисунок 3.21 – Результат определения вида тренда

По итогам анализа ряд является нестационарным, имеет линейный тренд с сезонными колебаниями. Наиболее подходящим для описания временного ряда трендом является линейный, так средняя ошибка аппроксимации для него минимальна, а коэффициент детерминации достаточно близок к единице.

Результат сглаживания ряда методом Брауна представлен на рисунке 3.22.

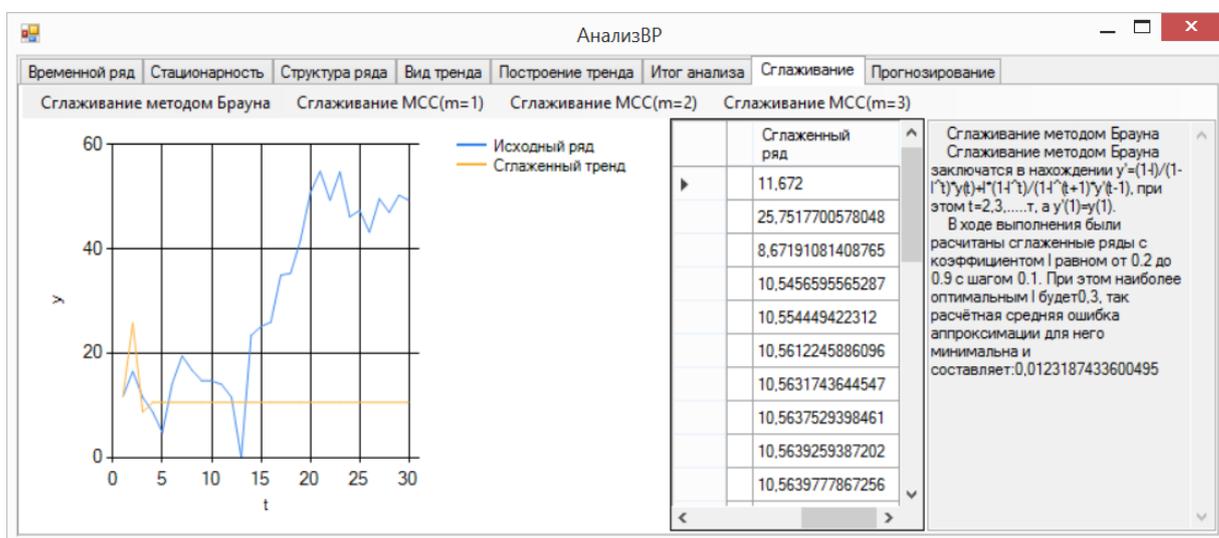


Рисунок 3.22 – Результат сглаживания методом Брауна

Результат сглаживания ряда МСС с $m=1$ представлен на рисунке 3.23.

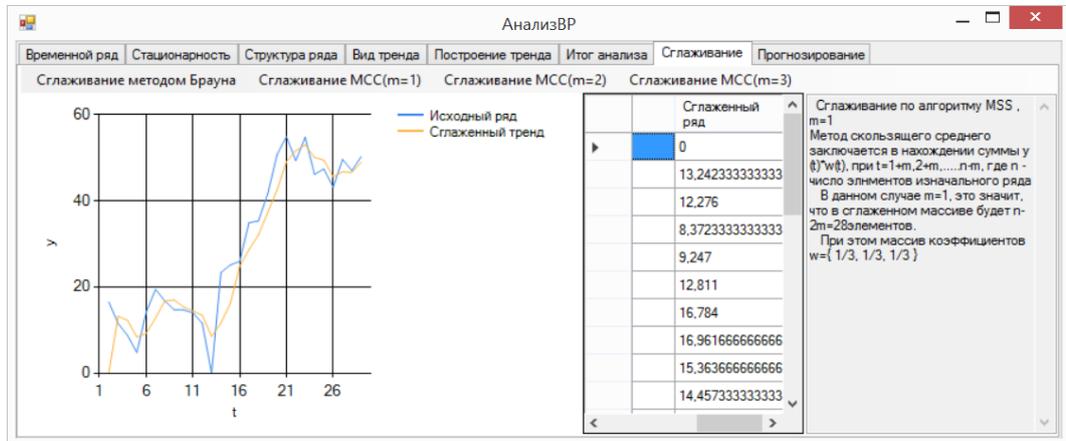


Рисунок 3.23 – Результат сглаживания МСС $m=1$

Результат сглаживания ряда МСС с $m=2$ представлен на рисунке 3.24.

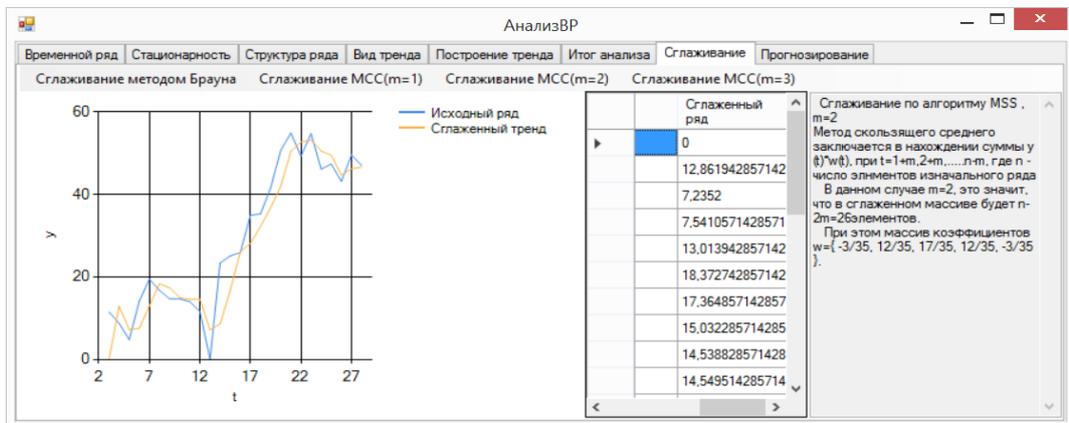


Рисунок 3.24 – Результат сглаживания МСС $m=2$

Результат сглаживания ряда МСС с $m=3$ представлен на рисунке 3.25.

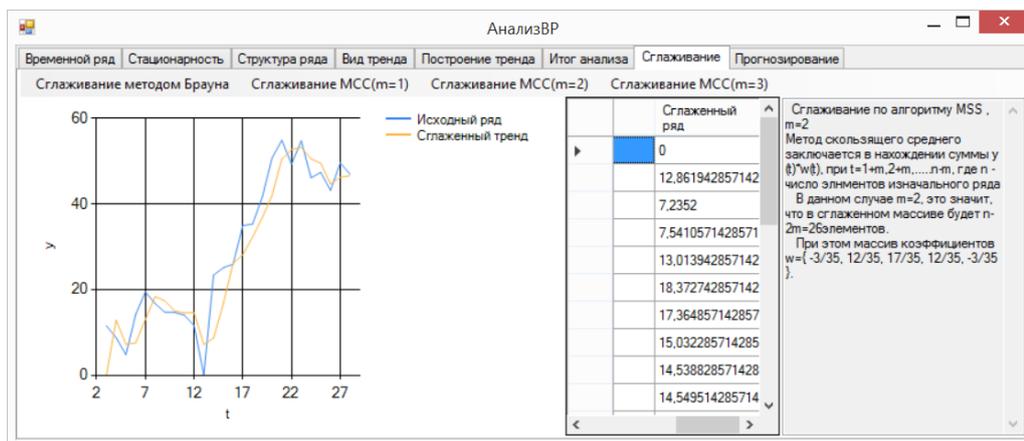


Рисунок 3.25 – Результат сглаживания МСС m=2

Из всех методов сглаживания МСС с $m=1$ является наиболее оптимальным, так как тренд, сглаженный этим методом, наиболее приближен к изначальному, а потери значений при этом минимальны.

Результат прогнозирования значений временного ряда по оптимальному тренду на 10 моментов времени представлен на рисунке 3.26.

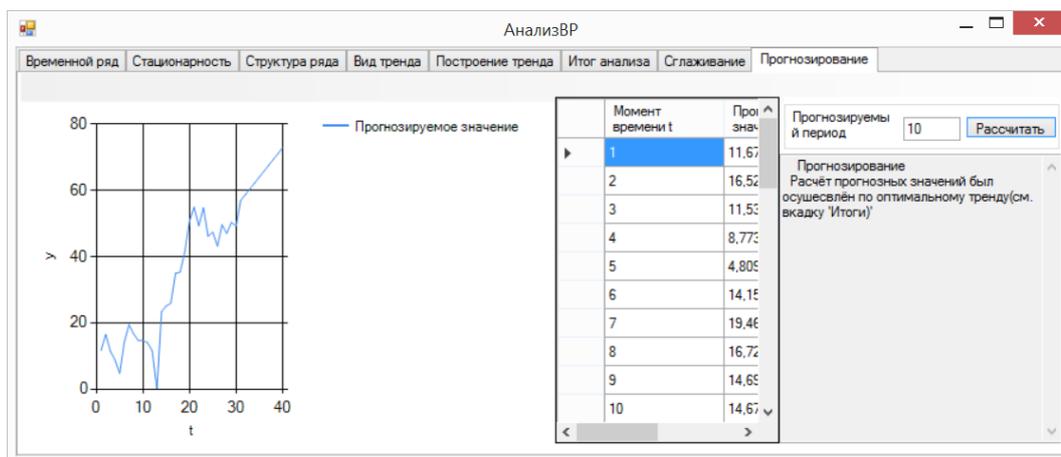


Рисунок 3.26 – Результат прогнозирования

Так как все расчёты, выполненные программой, корректны, то разработанный программный пакет обеспечивает необходимый функционал. Пояснения по работе каждой вкладки программы представлены в приложении В.

3.3 Оценка эффективности

В качестве критериев эффективности были выбраны:

- значительное сокращение времени, затрачиваемого на выполнение лабораторных работ по данной дисциплине;
- высокая степень удовлетворённости пользователей, использующих программный продукт.

Для их оценки наиболее подойдут опросы пользователей.

Для проверки первого критерия был проведён опрос среди уже использовавших программу студентов, содержащий 2 вопроса:

- сколько приблизительно времени вы затрачивали на выполнение лабораторных работ по разделу «Анализ временных рядов» дисциплины «Эконометрика» в минутах;
- сколько приблизительно времени вы затрачивали на выполнение этих же лабораторных работ по дисциплине «Эконометрика» при использовании программного пакета в минутах.[37]

Результат опроса представлен в таблице 3.2.

Таблица 3.2 – Результат опроса о затраченном времени.

Респондент	ЛРН№1		ЛРН№2		ЛРН№3		ЛРН№4		ЛРН№5	
	Время до, мин.	Время после, мин.								
1	240	80	100	40	260	140	260	140	360	120
2	200	120	220	60	310	210	310	210	300	210
3	250	180	180	130	400	180	340	210	270	200
4	300	120	260	140	480	240	270	200	360	120
5	120	80	310	210	360	120	240	100	120	80
6	400	240	220	155	390	120	300	180	300	180
7	320	140	350	160	420	280	370	140	360	120
8	180	60	360	120	360	120	180	120	300	180

9	210	150	300	180	300	180	360	120	320	180
10	120	90	120	45	420	240	340	230	270	200
11	410	180	220	90	280	140	270	145	250	180
12	250	140	300	220	360	200	360	120	360	200
13	330	90	300	150	330	90	330	90	330	90
Среднее	256,1	128	249	130	359	173,8	302,3	154,2	300	158,5
Соотноше- ние средних	1,994011976		1,905882353		2,066371681		1,960099751		1,893203883	
Среднее Соотношен ие средних	1,963913929									

При подсчете отношения между средними соотношениями средних было получено значение 1.96, что говорит об оценочном сокращении времени на выполнение лабораторных работ.

Для проверки второго критерия был создан опрос Google Forms и разослан той же группе респондентов.[38;39]

Данный опрос представлен на рисунке 3.16.

Удовлетворённость программным пакетом "АнализВР"

Описание

Насколько удобен пользовательский интерфейс пакета? *

- Удобен
- Скорее удобен
- Скорее неудобен
- Неудобен

Понятны ли вам пояснения пакета? *

- Понятны
- Скорее понятны
- Скорее непонятны
- Непонятны

Насколько функционал пакета соответствует содержанию раздела "Временные ряды" дисциплины "Эконометрика"? *

- Полностью соответствует
- Частично соответствует
- Полностью не соответствует

Насколько удобен текущий формат файлов, с которыми работает пакет? *

- Удобен
- Скорее удобен
- Скорее неудобен
- Неудобен

Рисунок 3.16 – Опрос об удовлетворённости

Соотношении ответов на вопрос «Насколько удобен пользовательский интерфейс пакета?» представлено на рисунке 3.17

Насколько удобен пользовательский интерфейс пакета?

13 ответов

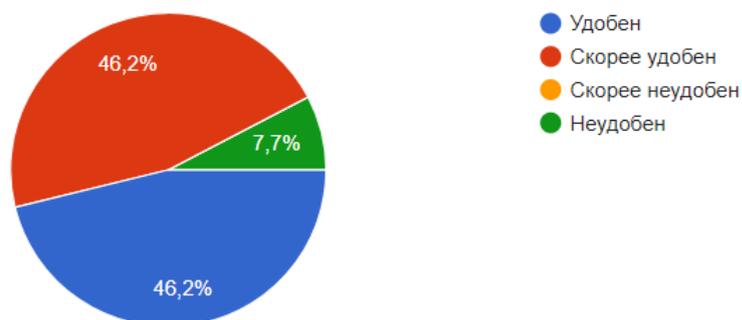


Рисунок 3.17 – Ответы на первый вопрос

Соотношении ответов на вопрос «Насколько удобен пользовательский интерфейс пакета?» представлено на рисунке 3.18

Понятны ли вам пояснения пакета?

13 ответов

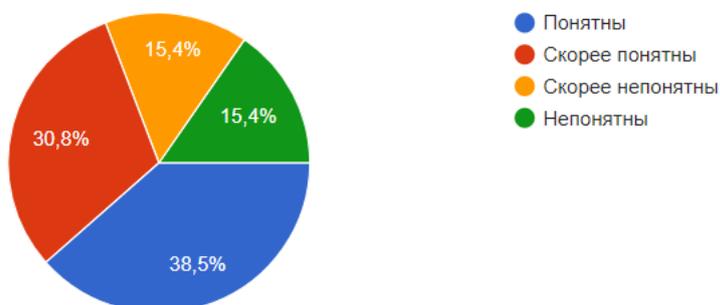


Рисунок 3.18 – Ответы на второй вопрос

Соотношении ответов на вопрос «Насколько функционал пакета соответствует содержанию раздела "Временные ряды" дисциплины "Эконометрика"?» представлено на рисунке 3.19

Насколько функционал пакета соответствует содержанию раздела "Временные ряды" дисциплины "Эконометрика"?

13 ответов

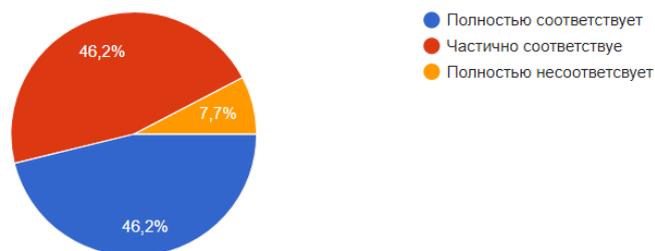


Рисунок 3.19 – Ответы на третий вопрос

Соотношении ответов на вопрос «Насколько функционал пакета соответствует содержанию раздела "Временные ряды" дисциплины "Эконометрика"?» представлено на рисунке 3.20

Насколько удобен текущий формат файлов, с которыми работает пакет?

13 ответов

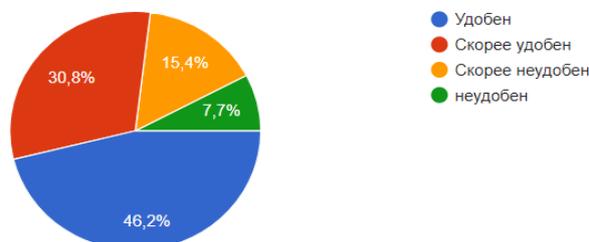


Рисунок 3.20 – Ответы на четвёртый вопрос

Анализируя ответы по данным четырём вопросам можно сделать вывод, что пользователи в целом удовлетворены прикладным пакетом, так как большинство ответов по ним составляют варианты «да» и «скорее да», однако сохраняется широкий простор для последующего улучшения путём добавления новых функциональных возможностей. Таким образом, в соответствии с результатами данных опросов можно сделать вывод, что критерии достижения цели достигнуты.

ЗАКЛЮЧЕНИЕ

В рамках работы был произведён анализ предметной области, включающий: обзор наиболее употребляемых при анализе временных рядов эконометрических методов и программных решений, сравнение их между собой и с разрабатываемым программным пакетом по критериям стоимости, удобства пользовательского интерфейса, удобства, а также моделирование процесса обучения «Как есть» и постановка задачи. Таким образом была выполнена задача по анализу предметной области.

В рамках проектирования были выполнены следующие операции: моделирование процесса обучения «Как должно быть», описание структуры и алгоритма работы программы. Это обеспечило выполнение задач по проведению моделирования процесса обучения «Как есть» и «Как должно быть» и по проектированию программного пакета.

В рамках разработки и тестирования было осуществлено описание процесса разработки, выполнен контрольный пример, а также произведена оценка эффективности разработанного программного пакета, что, в совокупности, обусловило выполнение задачи по разработке пакета и оценке его эффективности.

Так как в ходе оценки эффективности путём опроса было установлено, что оценочное время на выполнение лабораторных работ по разделу «Временные ряды» дисциплины «Эконометрика» существенно сократилось, а пользователи в целом удовлетворены разработанным программным продуктом, то можно сделать вывод, что цель выпускной квалификационной работы была достигнута по выбранным критериям.

Таким образом, все поставленные в рамках выпускной квалификационной работы задачи были решены, а цель – достигнута.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Multi-criteria selection of a corporate system by using paired comparison analysis [Текст] / V. V. Lomakin, N. P. Putivtseva, T. V. Zaitseva // Journal of Fundamental an Applied Sciences Vol 9, No 7S (2017).

2 Эконометрика (базовый курс) НИУ «БелГУ» [Текст]: учебное пособие / Н.П. Путивцева, Т.В. Зайцева, Е.В. Нестерова и др. – Белгород: ООО «ГИК», 2013. – 203 с.

3 Разработка специализированного пакета моделирования экономических процессов [Текст] / Н.П. Путивцева, Т.В. Зайцева, С.В. Игрунова и др. // Информатика: проблемы, методология, технологии: сборник материалов XVI международной научно-практической конференции, Воронеж, 11-12 февраля 2016 г. – Воронеж: Научно-исследовательские публикации, 2016.

4 О разработке пакета компьютерной поддержки принятия решений для выбора корреляционно-регрессионных моделей анализа и прогнозирования эмпирических данных [Текст] / Н.П. Путивцева, Т.В. Зайцева, С.В. Игрунова и др. // Научные Ведомости Белгородского Государственного Университета. Серия: ИСТОРИЯ. Политология. Экономика. Информатика. – Белгород: Изд-во НИУ «БелГУ», 2015. – 11 с.

5 Елисеева, И.И. Эконометрика. Учебник. [Текст] / И.И. Елисеева. – Москва: Финансы и статистика, 2003. – 344 с.

6 Артамонов, Н.В. Введение в эконометрику. [Текст] / Н.В. Артамонов. – Москва: МЦНМО, 2011. – 202 с.

7 Белолипецкий, А.А. Экономико-математические методы. [Текст] / А.А. Белолипецкий. – Москва: Академия, 2010. – 364 с.

8 Валентинов, В.А. Эконометрика. [Текст] / Валентинов В.А. – Москва: Дашков и К, 2010. – 448 с.

- 9 Колемаев, В.А. Эконометрика. [Текст] / В.А. Колемаев. – Москва: ИНФРА-М, 2010. – 160 с.
- 10 Гладилин, А.В. Эконометрика: Учебное пособие для студ. вузов по спец. "Математические методы в экономике" и другим экономическим специальностям. [Текст] / А.В. Гладилин. – Москва: КНОРУС, 2009. – 232 с.
- 11 Выявление структуры временного ряда [Электронный ресурс] / math.studopedia.ru – Режим доступа: https://studopedia.ru/4_78226_viyavlenie-strukturi-vremennogo-ryada.html, свободный.
- 12 Айвазян, С. А. Эконометрика [Текст]/ С.А. Айвазян, С.С. Иванова. – Москва: Маркет ДС, 2017. – 104 с.
- 13 Гладилин, А. В. Практикум по эконометрике / А.В. Гладилин, А.Н. Герасимов, Е.И. Громов. – Москва: Феникс, 2016. – 336 с.
- 14 Подкорытова, О.А. Анализ временных рядов. Учебное пособие для бакалавриата и магистратуры [Текст] / О.А. Подкорытова. – Москва: Юрайт, 2016. – 456 с.
- 15 Уравнение нелинейной регрессии [Электронный ресурс] / math.semestr.ru – Режим доступа: <https://math.semestr.ru/corel/noncorel.php>, свободный.
- 16 Эконометрическая модель как основа эконометрического моделирования. Типы данных и виды переменных в эконометрических исследованиях экономических явлений [Электронный ресурс] / studopedia.ru – Режим доступа: https://studopedia.ru/2_34934_ekonometriceskaya-model-kak-osnova-ekonometriceskogo-modelirovaniya-tipi-dannih-i-vidi-peremennih-v-ekonometriceskih-issledovaniyah-ekonomicheskikh-yavleniy.html, свободный.
- 17 Методы сглаживания временного ряда [Электронный ресурс] / math.studopedia.ru – Режим доступа: https://studopedia.ru/12_194491_metodi-sglazhivaniya-vremennogo-ryada-videlenie-nesluchaynoy-komponenti.html, свободный.
- 18 Мюллер, Г. Анализ временных рядов и прогнозирование. [Текст]/ Г. Мюллер и др. – Москва: Финансы и статистика, 2012. – 320 с.

- 19 Садовникова, Н.А. Анализ временных рядов и прогнозирование / Н.А. Садовникова, Р.А. Шмойлова. – Москва: МФПУ Синергия, 2016. – 152 с.
- 20 Афанасьев, В.Н. Анализ временных рядов и прогнозирование. [Текст] / В.Н. Афанасьев. – Москва: Финансы и статистика, 2012. – 320 с.
- 21 Бодров, О.А. Предметно-ориентированные экономические информационные системы[Текст] / О.А. Бодров. – Москва: ГЛТ, 2013. – 244 с.
- 22 Маторин, С.И. Теория систем и системный анализ: Учебное пособие [Текст] / С.И. Маторин, О.А. Зимовец. – Белгород: Изд-во НИУ «БелГУ», 2012. – 288 с.
- 23 Паттерсон, Д. Архитектура компьютера и проектирование компьютерных систем[Текст] / Д. Паттерсон. – Москва: Питер, 2012. – 784 с.
- 24 Чистов, Д. В. Проектирование информационных систем. Учебник и практикум. [Текст] / Д. В. Чистов. – Москва: Юрайт, 2016. – 258 с.
- 25 Регрессия в Excel [Электронный ресурс]: уравнение, примеры. Линейная регрессия/ fb.ru – Режим доступа: <http://fb.ru/article/322644/regressiya-v-excel-uravnenie-primeryi-lineynaya-regressiya>, свободный.
- 26 EViews [Электронный ресурс] / statmethods.ru – Режим доступа: <http://www.statmethods.ru/trainings/eviews.html>, свободный.
- 27 Statgraphics Plus for Windows [Электронный ресурс] / statmethods.ru – Режим доступа:https://studwood.ru/1940662/informatika/statgraphics_plus_windows_obschie_unikalnye_svoystva, свободный.
- 28 Возможности пакета spss для работы с данными [Электронный ресурс] / studfiles.net – Режим доступа: <https://studfiles.net/preview/5862667/page:6> , свободный.
- 29 Пакет Stata [Электронный ресурс] / studopedia.ru – Режим доступа: https://studopedia.ru/5_131342_paket-Stata.html, свободный.
- 30 Положение о выпускных квалификационных работах дипломированного специалиста, бакалавра, магистра (утв. 30.06.2014 г.) [Текст]. – Белгород: Изд-во БелГУ, 2014. - 20 с.

- 31 Виссер, Дж. Разработка обслуживаемых программ на языке С# [Текст] / Дж. Виссер. – Москва: ДМК Пресс, 2017. – 194 с.
- 32 Шарп, Дж. Microsoft Visual С#. Подробное руководство. 8-е издание [Текст] / Дж. Шарп. – Санкт – Петербург: Питер, 2017. – 848 с.
- 33 Ошероув, Р. Искусство автономного тестирования с примерами на С# [Текст] / Р. Ошероув. – Москва: ДМК Пресс, 2014. – 360 с.
- 34 Албахари, Д. С# 6.0. Справочник. Полное описание языка [Текст] / Д. Албахари. – Москва: Вильямс, 2016. – 1040 с.
- 35 Саати, Т. Принятие решений. Метод анализа иерархий. [Текст]. / Т. Саати; пер. с англ. Р.Г. Вачнадзе. – Москва: Радио и связь, 1993. – 278 с.
- 36 Сухарев, М.Г. Методы прогнозирования. Учеб. пособие – Москва: РГУ нефти и газа, 2009 г. – 208 с.
- 37 Батыгин, Г.С. Лекции по методологии социологических исследований [Текст] / Г.С. Батыгин. – Москва: Российский университет дружбы народов (РУДН), 2016. – 295 с.
- 38 Обработка опросов в Google формах [Электронный ресурс] / math.web-komfort.ru – Режим доступа: <http://web-komfort.ru/obrabotka-oprosov-v-google-formah>, свободный.
- 39 Все возможности Google forms [Электронный ресурс] / math.netology.ru – Режим доступа: <https://netology.ru/blog/google-formy>, свободный.
- 40 Обзор программного обеспечения эконометрического моделирования [Электронный ресурс] / scienceforum.ru – Режим доступа: <https://www.scienceforum.ru/2015/802/7954>, свободный.
- 41 ГОСТ 19.701-90 (ИСО 5807-85). ЕДИНАЯ СИСТЕМА ПРОГРАММНОЙ ДОКУМЕНТАЦИИ (ЕСПД). Схемы алгоритмов, программ, данных и систем. Обозначения условные и правила выполнения [Текст]. – Взамен ГОСТ 19.002-80, ГОСТ 19.003-80; Введ. 01.01.1992.
- 42 ГОСТ 7.32-2001. Отчет о научно-исследовательской работе. Структура и правила оформления [Текст]. - Взамен ГОСТ 7.32 - 91; Введ.

01.07.2002. – Москва: Стандартинформ, 2008. – 20 с. - (Система стандартов по информации, библиотечному и издательскому делу).

43 Фот, А.П. Оформление библиографических ссылок на использованные источники в научных работах [Электронный ресурс]: методические рекомендации соискателям ученых степеней / А.П. Фот. – Электрон. документ. – Оренбург: ГОУ «ОГУ», 2005. – 24 с. – Режим доступа: <http://www.osu.ru/docs/official/nauka/bibllinks.doc>, свободный.

44 Информационные технологии в эконометрических исследованиях [Электронный ресурс] / math. studopedia.ru – Режим доступа: https://studopedia.ru/19_246703_tema--informatisionnie-tehnologii-v-ekonometriceskih-issledovaniyah.html, свободный.

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ А

Таблица значений критерия Стьюдента

Таблица А.1 – Таблица значений критерия Стьюдента [36]

Степень свободы	k
1	6,31
2	2,92
3	2,35
4	2,13
5	2,02
6	1,94
7	1,9
8	1,86
9	1,83
10	1,81
11	1,8
12	1,78
13	1,77
14	1,76
15	1,75
16	1,75
17	1,74
18	1,73
19	1,73
20	1,73
21	1,72
22	1,72
23	1,71
24	1,71
25	1,71
26	1,71
27	1,7
28	1,7
29	1,7
30	1,7
40	1,68
∞	1,65

ПРИЛОЖЕНИЕ Б

Обзор эконометрических пакетов

Microsoft Excel предоставляет богатый набор инструментов для анализа статистических данных [25]. К наиболее часто употребляемым для анализа временных рядов функциям, реализованным в пакета Excel являются:

- ЛИНЕЙН – осуществляет расчёт массив данных, описывающих уравнение линейной множественной регрессии;
- РОСТ – осуществляет расчёт будущей уровней ряда в соответствии с экспоненциальной кривой;
- ЛИНЕЙН – осуществляет расчёт значений временных рядов, соответствующих уравнению линейной множественной регрессии.
- ЛГРФПРИБЛ – осуществляет расчёт значений временных рядов, соответствующих экспоненциального тренда;

Для более полного анализа используется вкладка анализ данных, представленная на рисунке Б.1.

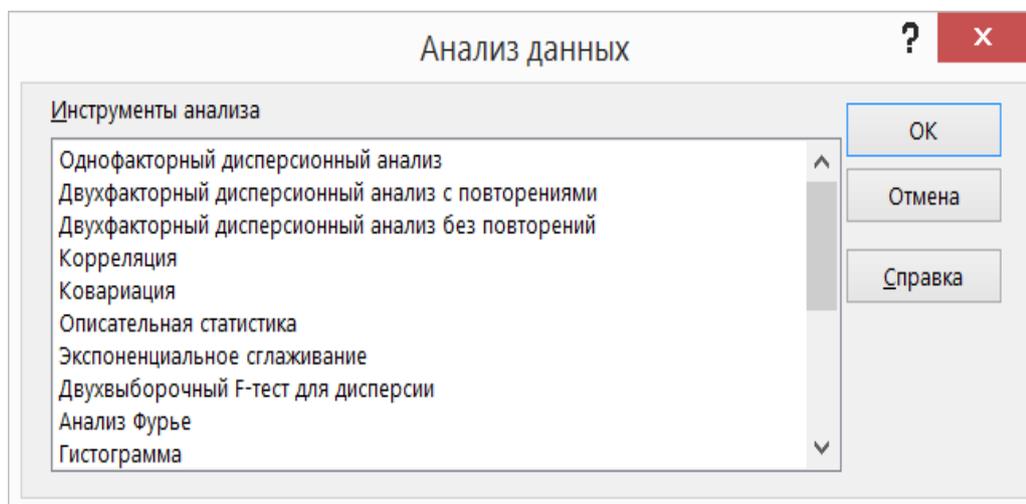


Рисунок Б.1 – окно «Анализ данных»

При этом существует возможность добавления линии тренда на сами графики. При этом можно построить экспоненциальный, степенной,

гиперболический, логарифмический и полиномиальный тренды, что проиллюстрировано рисунком Б.2.

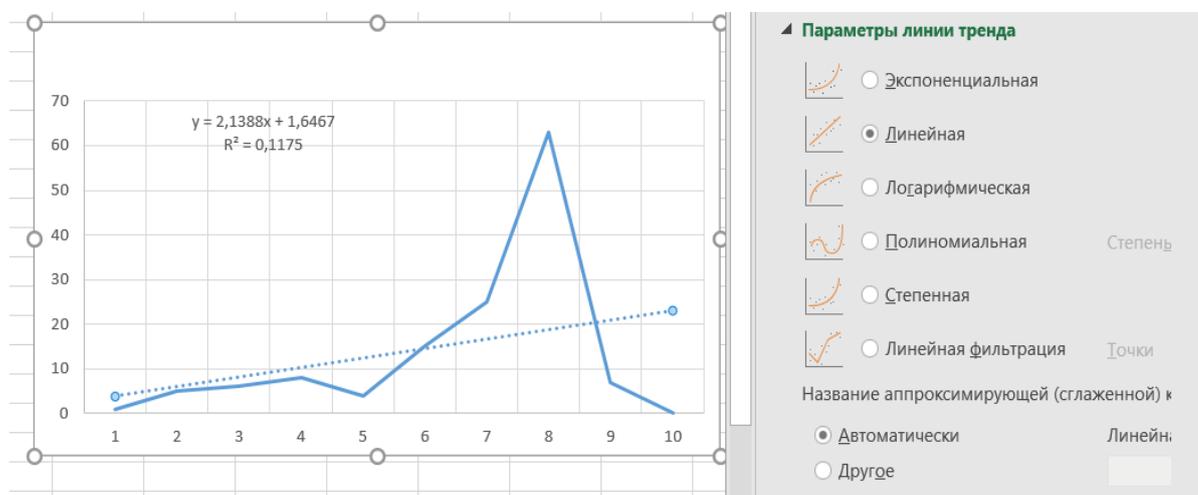


Рисунок Б.2 – построение тренда в excel.

К преимуществам регрессионного анализа посредством пакета MS Excel можно отнести его интегрированности для работы с самим табличным процессором и русифицированность интерфейса и результатов анализа. К недостаткам данного пакета можно отнести невозможность подробного просмотра процесса анализа.

К возможностям и особенностям программного пакета SPSS можно отнести:

- удобный доступ к информации;
- удобный сбор и ввод данных;
- эффективный data - менеджмент;
- широкий спектр статистических операций для анализа данных и построения наиболее достоверных моделей;
- возможность публикации результатов в сети.

STATGRAPHICS Plus for Windows включает более 250 статистических и системных процедур, применяющихся в бизнесе, экономике, маркетинге, медицине, биологии, социологии, психологии, на производстве и в других областях. [27]

Базовый пакет даёт большой выбор инструментальных средств для анализа и прогнозирования данных, при этом обладая хорошей оптимизацией и удобным интерфейсом.

Популярность пакета STATISTICA обуславливается продвинутыми методами визуализации результатов, высокая степень интеграции с прикладным пакетом MS Excel, дающая возможность удобного экспорта данных и обращения к функциям Excel прямо из меню пакета STATISTICA. Интерфейс пакета представлен на рисунке Б.3.

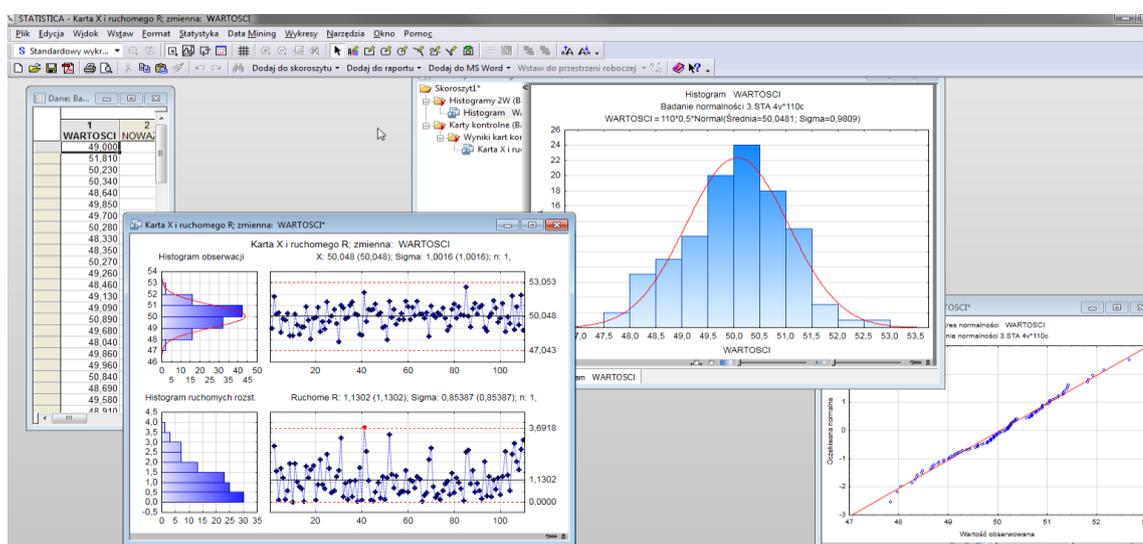


Рисунок Б.3 – Интерфейс пакета «Statistica»

Как видно из рисунка Б.3, интерфейс программы выполнен на английском языке, что создаёт некоторые трудности при работе. Также к недостаткам данного пакета можно отнести наличие модулей, излишних для анализа временных рядов на уровне предусмотренной учебной программой.

EViews – пакет статистического анализа, используемый, в основном, для анализа эконометрических данных временных рядов, моделирования панельных данных и построения регрессионных моделей. Интерфейс пакета представлен на рисунке Б.4. [25]

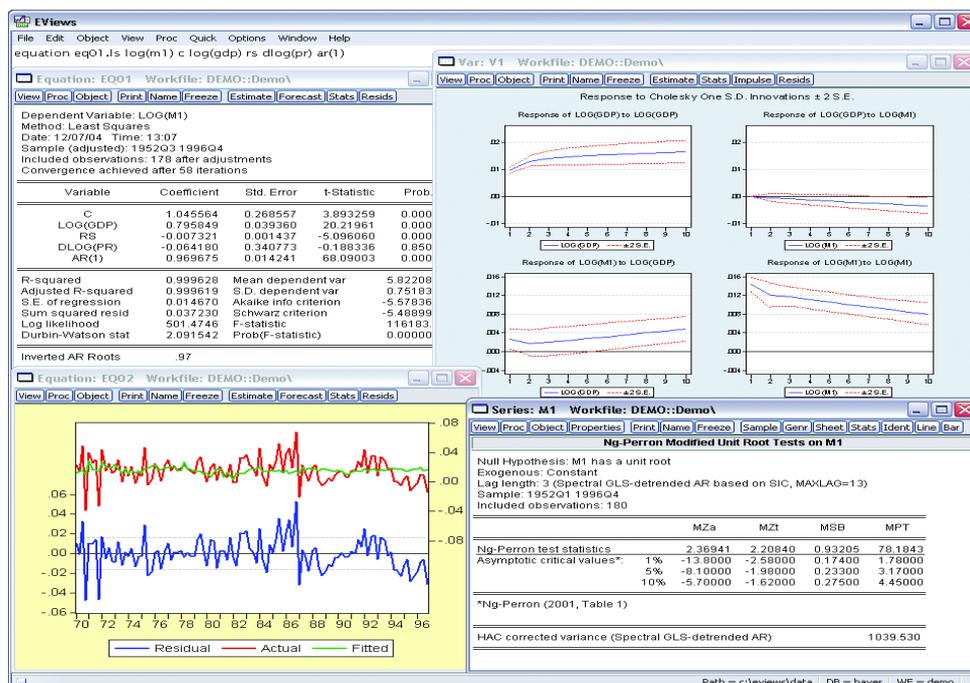


Рисунок Б.4 – интерфейс пакета «EViews»

Данный пакет обладает следующими достоинствами:

- включает в себя широкий выбор методов эконометрического анализа;
- хорошо проработанная система справки;
- хорошая оптимизация и высокая производительность;
- возможность разработки собственных скриптов.

Недостатком данного пакета являются отсутствие русифицированной версии.

Stata – пакет статистического анализа с широким спектром применения.

К достоинствам данной системы можно отнести:

- широкий выбор средств для эконометрического анализа;
- подробная система справки;
- доступная цена;
- наличие обучающих роликов.

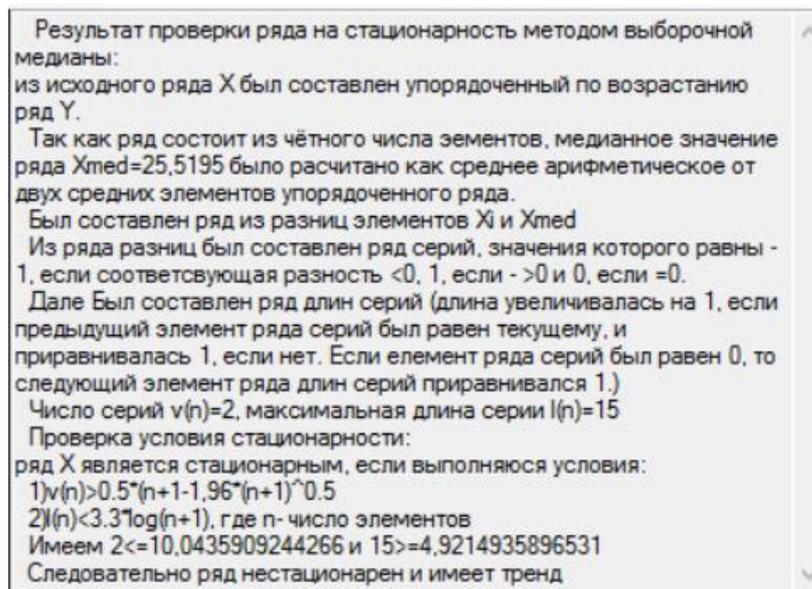
Недостатками данной системы являются:

- сложный командный синтаксис;
- нерусифицированность интерфейса и справки. [29]

ПРИЛОЖЕНИЕ В

Пояснения по вкладкам приложения

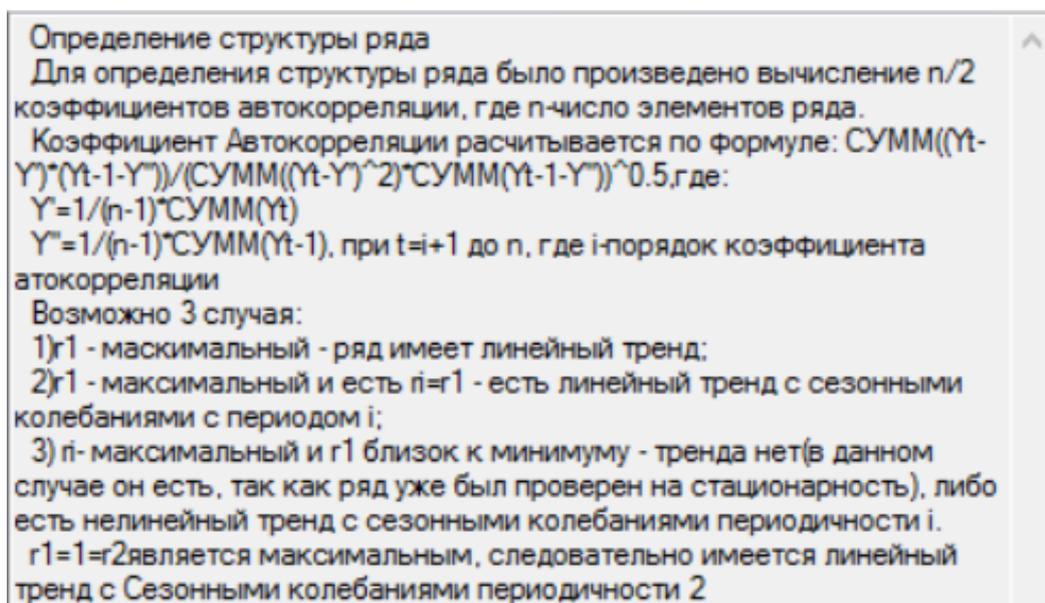
Пояснения по вкладке «Стационарность» представлены на рисунке В.1.



Результат проверки ряда на стационарность методом выборочной медианы:
из исходного ряда X был составлен упорядоченный по возрастанию ряд Y.
Так как ряд состоит из чётного числа элементов, медианное значение ряда $X_{med}=25,5195$ было рассчитано как среднее арифметическое от двух средних элементов упорядоченного ряда.
Был составлен ряд из разниц элементов X_i и X_{med}
Из ряда разниц был составлен ряд серий, значения которого равны -1, если соответствующая разность <0 , 1, если >0 и 0, если $=0$.
Далее был составлен ряд длин серий (длина увеличивалась на 1, если предыдущий элемент ряда серий был равен текущему, и приравнивалась 1, если нет. Если элемент ряда серий был равен 0, то следующий элемент ряда длин серий приравнивался 1.)
Число серий $v(n)=2$, максимальная длина серии $l(n)=15$
Проверка условия стационарности:
ряд X является стационарным, если выполняются условия:
1) $v(n) > 0.5 \cdot (n+1) \cdot 1.96 \cdot (n+1)^{0.5}$
2) $l(n) < 3.3 \cdot \log(n+1)$, где n- число элементов
Имеем $2 < 10,0435909244266$ и $15 > 4,9214935896531$
Следовательно ряд нестационарен и имеет тренд

Рисунок В.1 – Пояснения по вкладке «Стационарность»

Пояснения по вкладке «Структура ряда» представлены на рисунке В.2.



Определение структуры ряда
Для определения структуры ряда было произведено вычисление $n/2$ коэффициентов автокорреляции, где n-число элементов ряда.
Коэффициент Автокорреляции рассчитывается по формуле: $\text{СУММ}((Y_t - Y) \cdot (Y_{t-1} - Y)) / (\text{СУММ}((Y_t - Y)^2) \cdot \text{СУММ}((Y_{t-1} - Y)^2))^{0.5}$, где:
 $Y = 1/(n-1) \cdot \text{СУММ}(Y_t)$
 $Y^* = 1/(n-1) \cdot \text{СУММ}(Y_{t-1})$, при $t=i+1$ до n, где i-порядок коэффициента автокорреляции
Возможно 3 случая:
1) r_1 - максимальный - ряд имеет линейный тренд;
2) r_1 - максимальный и есть $\pi = r_1$ - есть линейный тренд с сезонными колебаниями с периодом i;
3) π - максимальный и r_1 близок к минимуму - тренда нет (в данном случае он есть, так как ряд уже был проверен на стационарность), либо есть нелинейный тренд с сезонными колебаниями периодичности i.
 $r_1 = 1 = r_2$ является максимальным, следовательно имеется линейный тренд с сезонными колебаниями периодичности 2

Рисунок В.2 – Пояснения по вкладке «Структура ряда»

Пояснения по вкладке «Вид тренда» представлены на рисунке В.3.

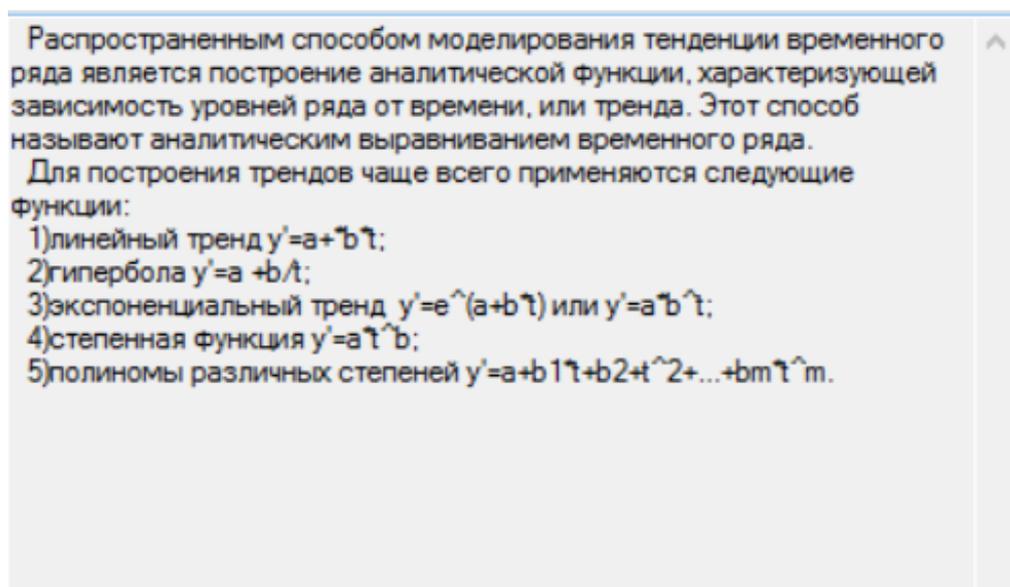


Рисунок В.3 – Пояснения по вкладке «Вид тренда»

Пояснения по вкладке «Подведение итогов» представлены на рисунке В.4.

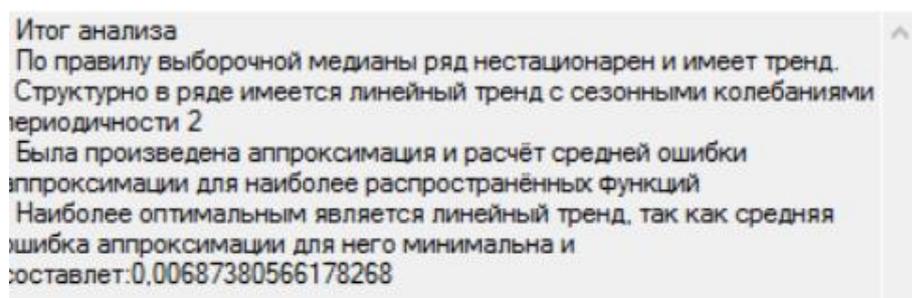


Рисунок В.4 – Пояснения по вкладке «Подведение итогов»

Пояснения по вкладке «Сглаживание» при применении сглаживания методом Брауна представлены на рисунке В.5.

Сглаживание методом Брауна заключается в нахождении $y' = (1-l)/(1-l^t) * y(t) + l * (1-l^t)/(1-l^{t+1}) * y'(t-1)$, при этом $t=2,3,\dots,t$, а $y'(1)=y(1)$.

В ходе выполнения были рассчитаны сглаженные ряды с коэффициентом l равном от 0.2 до 0.9 с шагом 0.1. При этом наиболее оптимальным l будет 0.3, так расчётная средняя ошибка аппроксимации для него минимальна и составляет: 0,0123187433600495

Рисунок В.5 – Пояснения по вкладке «Сглаживание» при применении сглаживания методом Брауна

Пояснения по вкладке «Сглаживание» при применении сглаживания МСС $m=1$ представлены на рисунке В.6.

Сглаживание по алгоритму МСС , $m=1$

Метод скользящего среднего заключается в нахождении суммы $y(t) * w(t)$, при $t=1+m, 2+m, \dots, n-m$, где n - число элементов изначального ряда

В данном случае $m=1$, это значит, что в сглаженном массиве будет $n-2m=28$ элементов.

При этом массив коэффициентов $w = \{ 1/3, 1/3, 1/3 \}$

Рисунок В.6 – Пояснения по вкладке «Сглаживание» при применении сглаживания МСС $m=1$

Пояснения по вкладке «Сглаживание» при применении сглаживания МСС $m=2$ представлены на рисунке В.7.

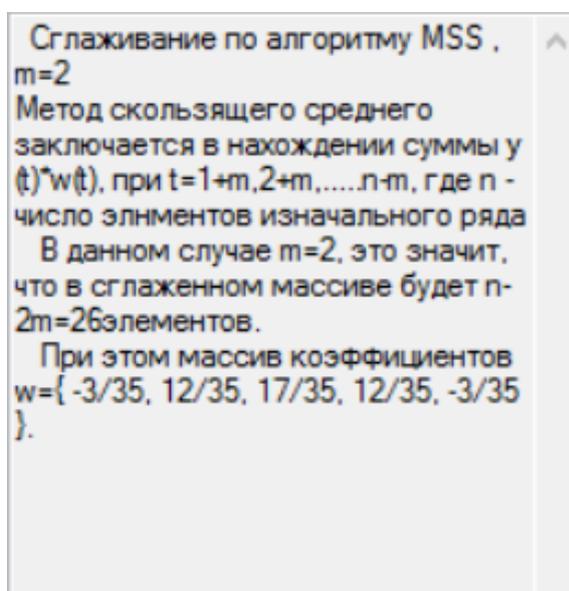


Рисунок В.7 – Пояснения по вкладке «Сглаживание» при применении сглаживания МСС $m=2$

Пояснения по вкладке «Сглаживание» при применении сглаживания МСС $m=3$ представлены на рисунке В.8.

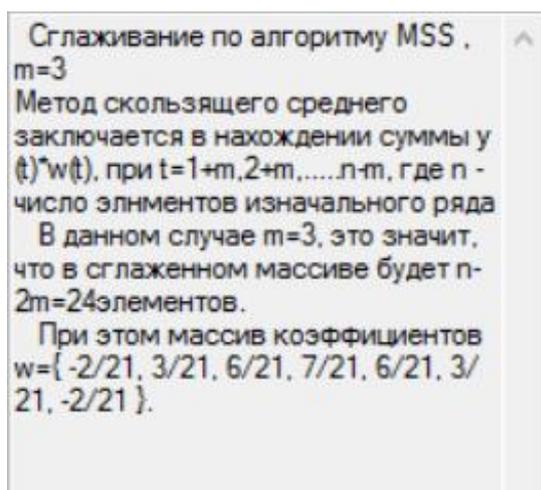


Рисунок В.8 – Пояснения по вкладке «Сглаживание» при применении сглаживания МСС $m=3$