МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение

высшего профессионального образования

**«Сибирский федеральный университет»**

Институт

Кафедра

Современные компьютерные технологии в науке

Курсовая работа

Разработка лабораторного комплекса в пакете SPSS для решения задачи анализа временных рядов

Выполнил:

Проверил: к.т.н., доцент

Панфилов И.А.

Красноярск, 20\_\_\_\_

# Содержание

[Постановка задачи 3](#_Toc358041349)

[Теоретические сведения 4](#_Toc358041350)

[Выполнение лабораторной работы 7](#_Toc358041351)

[Вывод 13](#_Toc358041352)

[Список литературы 14](#_Toc358041353)

# Постановка задачи

Разнообразные данные в экономике, социологии, коммерции и других сферах человеческой деятельности поступают в форме временных рядов. Из необходимости работать с такими данными и некоторыми их особенностями возник целый раздел математической статистики, называемый анализ временных рядов.

Дано: статистические данные, расположенные в хронологическом порядке: . Например, данные об урожайности пшеницы в разные годы в РФ, представленные в таблице 1. Данные взяты с сайта [2].

Цель работы – познакомиться с основными этапами анализа временных рядов и построить различные модели для прогнозирования данных, представленных в таблице 1.

Таблица 1. Данные об урожайности (гектограмм на гектар) пшеницы в РФ по годам

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Год | Урожайность (гг/га) | Год | Урожайность (гг/га) | Год | Урожайность (гг/га) |
| 1992 | 19030,36 | 1999 | 15689,64 | 2006 | 19491,82 |
| 1993 | 17654,86 | 2000 | 16141,43 | 2007 | 21007,21 |
| 1994 | 14498,14 | 2001 | 20576,67 | 2008 | 24458,91 |
| 1995 | 13945,78 | 2002 | 20675,59 | 2009 | 23181,78 |
| 1996 | 15486,87 | 2003 | 17047,38 | 2010 | 19181,14 |
| 1997 | 18426,58 | 2004 | 19813,66 | 2011 | 22645,00 |
| 1998 | 13602,39 | 2005 | 19324,34 |  |  |

# Теоретические сведения

Введем понятие временного ряда, которое будет использовано на протяжении всей лабораторной работы.

Ряд наблюдений анализируемой случайной величины , произведенных в последовательные моменты времени , называется **временным рядом** [1]. Тут будут рассмотрены временные ряды с равностоящими моментами наблюдений, тогда их удобно представлять в виде . Принципиальные отличия временного ряда от последовательности наблюдений случайной величины, не образующей временной ряд: (1) в отличие от элементов случайной выборки члены временного ряда не являются статистически независимыми; (2) члены временного ряда не являются одинаково распределенными.

Выделяют четыре типа факторов, под воздействием которых формируются значения элементов временного ряда: (А) Долговременные, формирующие общую тенденцию изменения анализируемого признака, описываемую неслучайной функцией тренда; (Б) Сезонные, формирующие периодически повторяющиеся в определенное время года колебания анализируемого признака (описывается неслучайной периодической функцией ); (В) Циклические, формирующие изменения анализируемого признака, обусловленные действием долговременных цикло разной природы (описывается неслучайной функцией ); (Г) Случайные, не поддающиеся учету и регистрации (обозначим ). Для определенности примем аддитивную структурную схему влияния упомянутых факторов на формирование значений :

**(1)**

Основные задачи анализа временных рядов:

По имеющейся траектории анализируемого временного ряда требуется: (а) Определить какие из неслучайных функций-факторов присутствуют в разложении (1); (б) построить «хорошие» оценки для тех неслучайных функций, которые присутствуют в разложении (1); (в) подобрать модель, адекватно описывающую поведение «случайных остатков» , и статистически оценить параметры этой модели. Успешное решение этих задач является основой для решения задачи кратко- и среднесрочного прогноза значений временного ряда.

**Модель Брауна** относится к моделям экспоненциального сглаживания и используется для целей краткосрочного прогнозирования нестационарных процессов. Общий вид модели: , в классическим варианте . Если перегруппировать слагаемые в модели , то становятся очевидны адаптивные свойства модели. В классических границах изменения постоянной сглаживания модель Брауна инерционна.

**Модель Холта** является некоторой модификацией модели Брауна для краткосрочного прогнозирования рядов, имеющих ярко выраженную тенденцию роста. Модель используем допущение о том, что ряд данных имеет некоторую тенденцию к изменяющемуся во времени линейному росту (либо снижению). Вид модели: , - прогнозное значение ряда, сделанное на шагов вперед, - расчетные значения коэффициентов, соответственно характеризующие начальное значение динамического ряда и пропорциональность между и . Расчет значений коэффициентов: , , . Коэффициенты подбираются программными пакетами при помощи численных методов с каким-либо условием (например, минимизация суммы квадратов отклонений фактических значений от расчётных).

**Модель АРПСС** (авторегрессии проинтегрированного скользящего среднего) предложена Дж. Боксом и Г. Дженкинсом, так же известна как «модель Бокса-Дженкинса» или ARIMA (AutoRegressive Integrated Moving Average). Она предназначена для описания нестационарных временных рядов. Модель включает в себя два процесса: процесс авторегрессии, отражающий последовательную зависимость элементов ряда друг от друга, и процесс скользящего среднего, отражающий подверженность каждого элемента ряда суммарному воздействию предыдущих ошибок [4]. Модель АРПСС(*p, d, q*) включает параметры авторегрессии (*p*), порядок разности (*d*), параметры скользящего среднего (*q*). Для модели АРПСС необходимо, чтобы ряд был **стационарным** (т.е. его среднее постоянно, а выборочные дисперсия и автокорреляция не меняются во времени), поэтому необходимо брать разности ряда до тех пор, пока он не станет стационарным. Число разностей, которые были взяты, чтобы достичь стационарности, определяются параметром *d*. Также для построения этой модели необходимо решить, как много параметров авторегрессии (*p*) и скользящего среднего (*q*) должно присутствовать в эффективной и экономной модели процесса. Модель АРПСС(*p, d, q*) анализируемого процесса может быть записана в виде: . Для оценки адекватности этой модели применяется анализ остатков.

# Выполнение лабораторной работы

Работа производилась в русифицированном статистическом пакете SPSS Statistics 17.0.

**Этап 1:** Запуск и добавление данных.

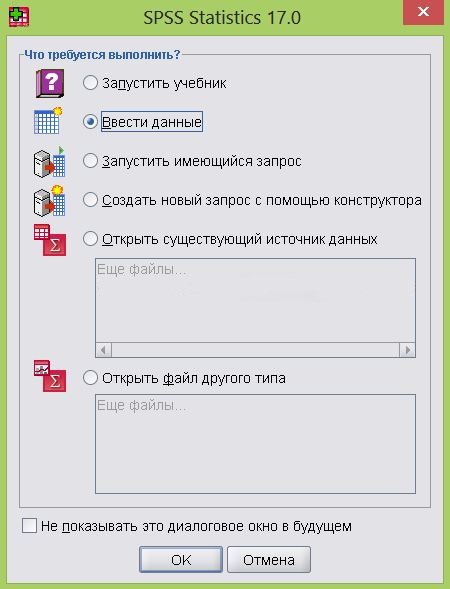


Рисунок 1 – Добавление данных

Выбрав соответствующую позицию (рисунок 1), вводим данные из таблицы 1.

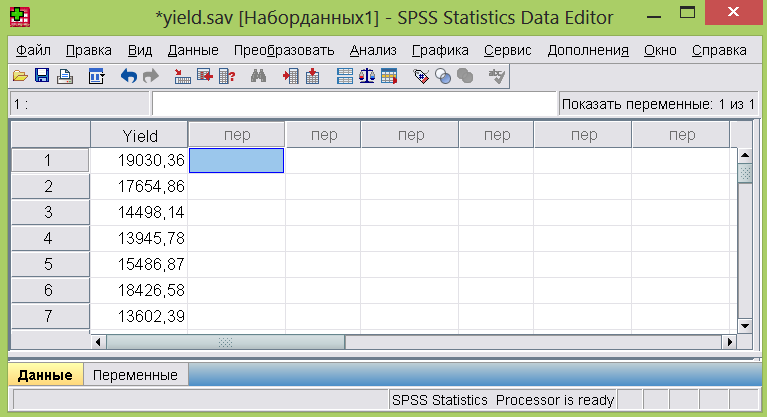


Рисунок 2 – Рабочее окно программы с данными

Программа обладает привычным Windows-интерфейсом. Сохранить введенные данные можно с помощью «Файл – Сохранить как…». Дважды щелкнув на заголовок «пер», можно перейти к изменению имени, типа, шкалы измерения добавленных переменных. Для удобства работы меняем поля имя (Year, Yield), десятичные (0, 2 соответственно), шкала (порядковая, количественная).

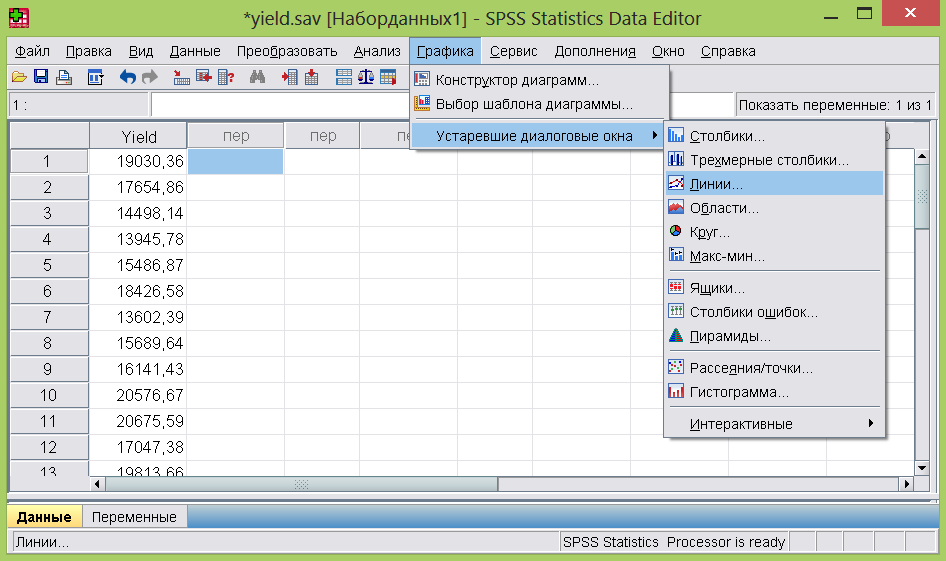


Рисунок 3 – Построение графического изображения данных

Следу указаниям, представленным на рисунках 3 и 4, можно построить график имеющихся данных (рисунок 5).

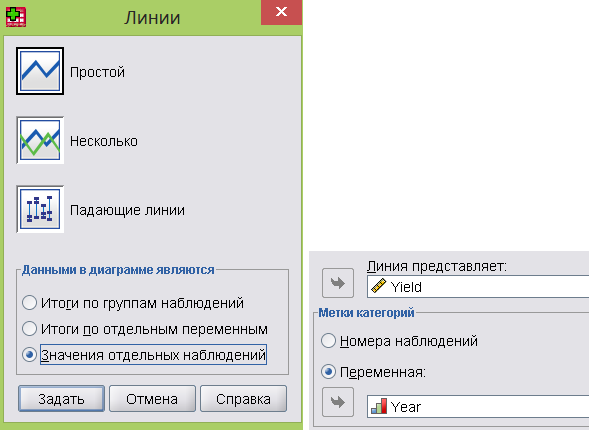


Рисунок 4 – Выбор параметров построения графика

Вывод результатов производится в новом окне, называемом SPSS Statistics Viewer.

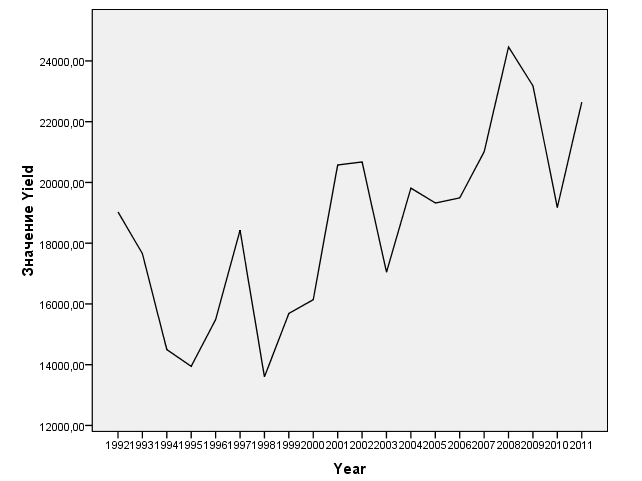


Рисунок 5 – График временного ряда из таблицы 1

**Этап 2:** Выделение неслучайной составляющей.

В SPSS это можно сделать, выбрав «Преобразовать – Создать временной ряд…» и выбрав следующие настройки (рисунок 6):

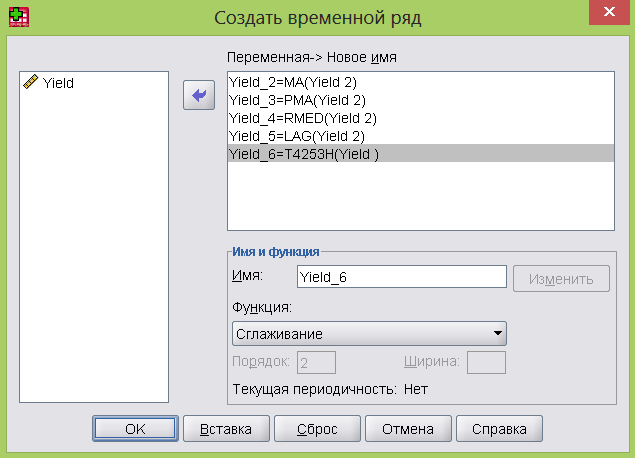


Рисунок 6 – Настройки при создании временного ряда

Здесь для создания новых переменных выбирались следующие функции: центрированное скользящее среднее, опережающее скользящее среднее, скользящие медианы, лаг, сглаживание. Стоит отметить, что при изменении имени, функции или ее параметров для сохранения этих изменений необходимо нажимать кнопку «Изменить».

**Этап 3:** Построение моделей и прогнозов.

Перейдем к конструктору построения моделей (рисунок 7), проделав следующий путь: «Анализ – Прогнозирование – Создание моделей…».

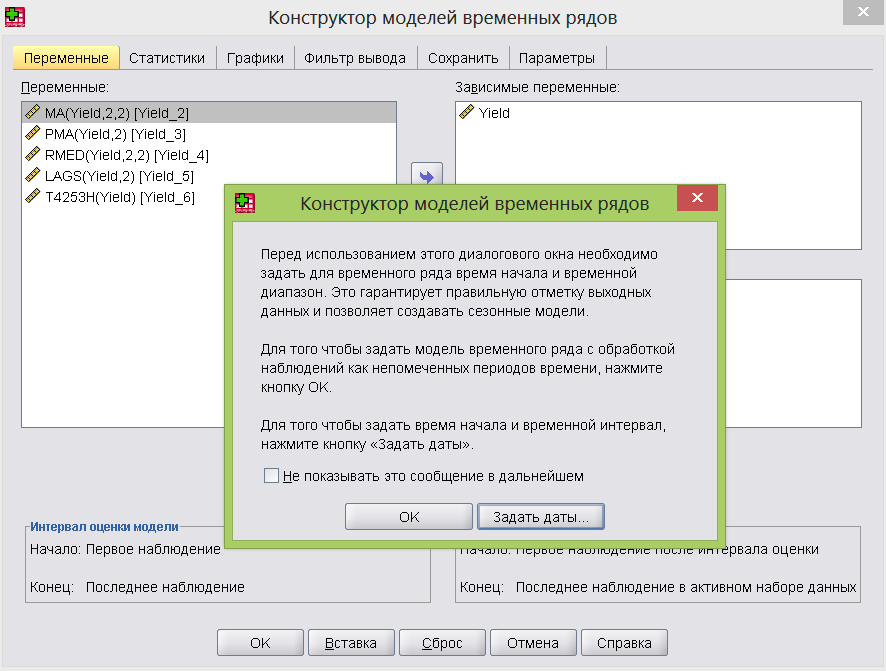


Рисунок 7 – Конструктор построения моделей временных рядов

Для начала работы необходимо задать даты (рисунок 8).

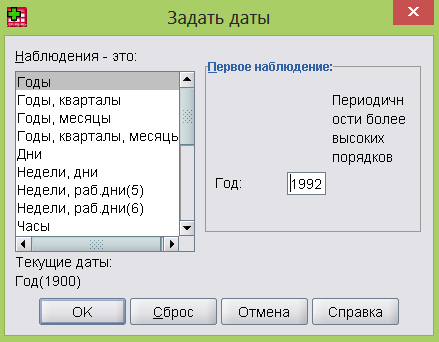


Рисунок 8 – Окно «Задать даты»

Задав даты наблюдений, опять перейдем к созданию моделей. В конструкторе моделей временных рядов все созданные ранее переменные переместим в зависимые, а переменную YEAR в независимые, выберем метод «Эксперт построения моделей» (рисунок 9). В окне «Критерии…» выберем «Все модели».

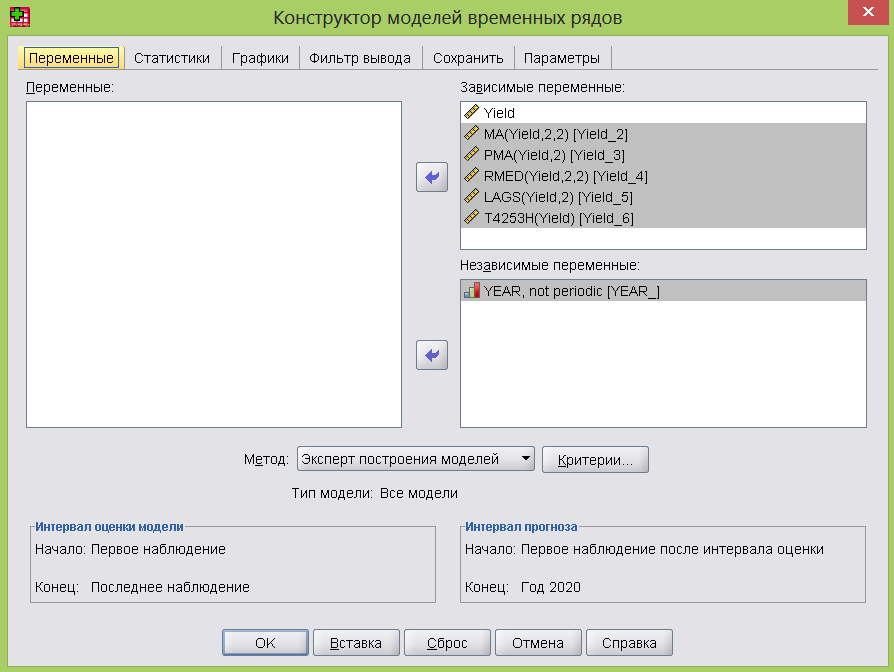


Рисунок 9 – Окно «Конструктор моделей временных рядов»

Во вкладке «Параметры» выберем интервал прогноза «От первого наблюдения после интервала оценки и до заданной даты» Зададим 2015 год. Во вкладке «Фильтр вывода» оставим «Включить в вывод все модели». Во вкладке «Графики» проверим, что стоят галочки напротив «Ряды», показать на каждом графике «Наблюдаемые значения» и «Прогнозные значения».

В окне вывода результатов увидим: описание модели (рисунок 10), параметры моделей экспоненциального сглаживания и моделей АРПСС (рисунок 11), графики наблюдаемых и прогнозных значений (рисунок 12), также таблица с прогнозируемыми значениями и границами доверительного интервала для каждой модели.

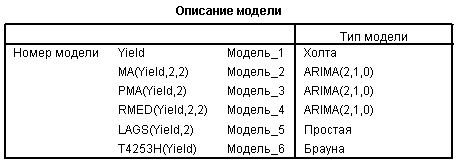


Рисунок 10 – Описание моделей



Рисунок 11 – Параметры моделей

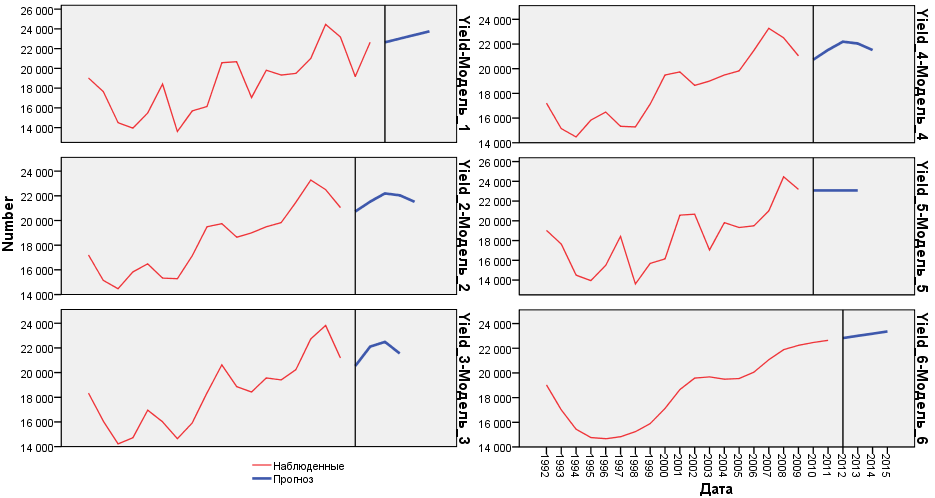


Рисунок 12 – Графики наблюдаемых значений и прогнозов

**Задание:** Проанализировать полученные результаты, учитывая корректность применения тех или иных моделей для построения оценок в данной задаче.

# Вывод

Статистический пакет SPSS является мощным инструментом при решении задач прикладной статистики, в том числе и для задач анализа временных рядов. Исследователь, используя этот пакет, имеет широкий круг возможностей для построения различных статистических оценок и задания необходимых ему параметров строящейся модели, также пакет предоставляет возможность для сравнения полученных моделей.

# Список литературы

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики: Учебник для вузов. – М.: ЮНИТИ, 1998. – 270 с.
2. FAOSTAT [Электронный ресурс]:<http://faostat3.fao.org>
3. Айвазян С.А. Прикладная статистика: Исследование зависимостей: Справ. Изд. / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин; Под ред. С.А. Айвазян. – М.: Финансы и статистика, 1985. – 487 с., ил.
4. StatSoft [Электронный ресурс]: <http://www.statsoft.ru/home/portal/applications/ForecastingAdvisor/Methods/ARPSS/arpss.htm>
5. [Электронный ресурс]: http://finance.yahoo.com/q/hp?s=AAPL+Historical+Prices