**Динамика макроэкономических показателей РФ**

**Выбранный показатель**: Индекс реального объема сельскохозяйственного производства.

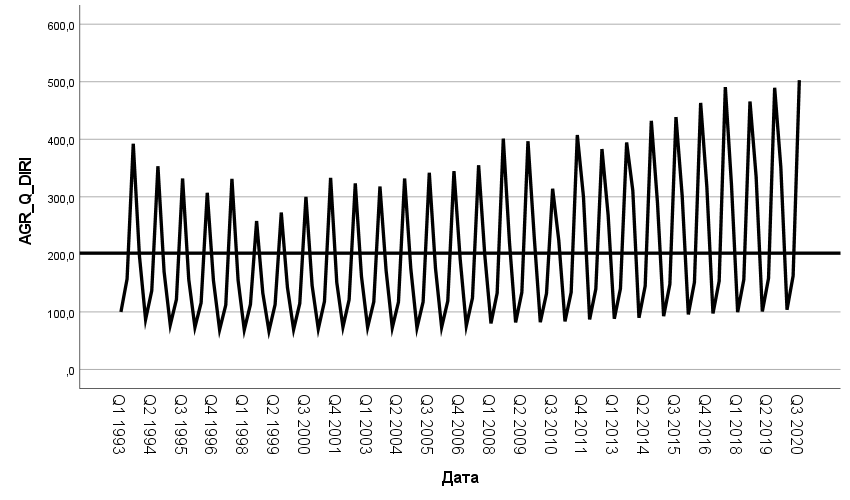
**Индекс реального объема сельскохозяйственного производства -** относительный показатель, характеризующий изменение объема произведенных продуктов растениеводства и животноводства в сравниваемых периодах.

**Период**: квартальные показатели за 1993-2020 год.

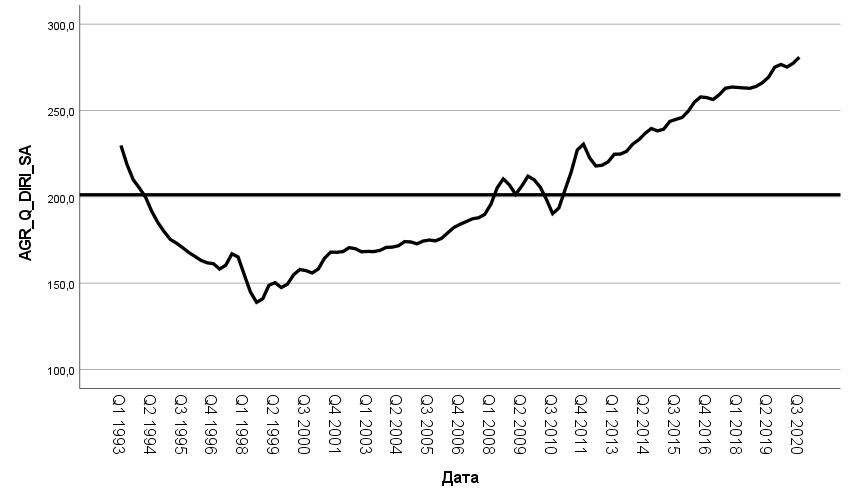
**Этапы работы:**

1. Загрузка данных и приведение в нужный вид (формат, создание колонки с датой, длина)
2. Разведочный анализ данных – построение диаграммы временного ряда, определение сезонности/тренда. (диаграммы последовательностей).
3. Анализ ряда на автокорреляцию.
4. Поиск оптимальной ARIMA модели с помощью мастера временных рядов. (на всем наборе данных).
5. Оценка полученных предсказаний.
6. Поиск оптимальной ARIMA модели с помощью мастера временных рядов. (на наборе данных с 2010 года).
7. Оценка полученных предсказаний.
8. Подбор моделей вручную.
9. Построение предсказания на лучшей модели по данным 2010-2020 годов.

**Анализ временных рядов:**



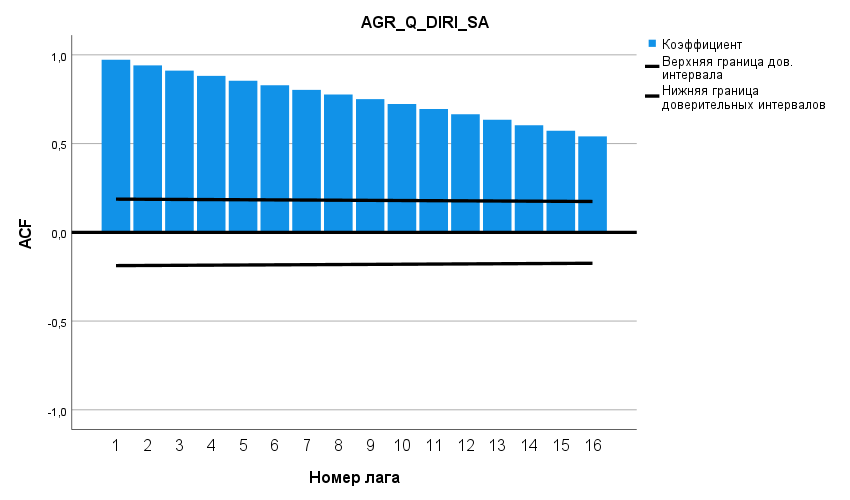
На графике видна ярко выраженная сезонность. При этом сложно определить наличие тренда. Для такого ряда коэффициент автокорреляции будет отрицательный. За счет наличия сезонности и тренда ряд можно считать стационарным, что логично, учитывая специфику рассматриваемого показателя. Для дальнейшего анализа необходимо избавиться от сезонной компоненты.

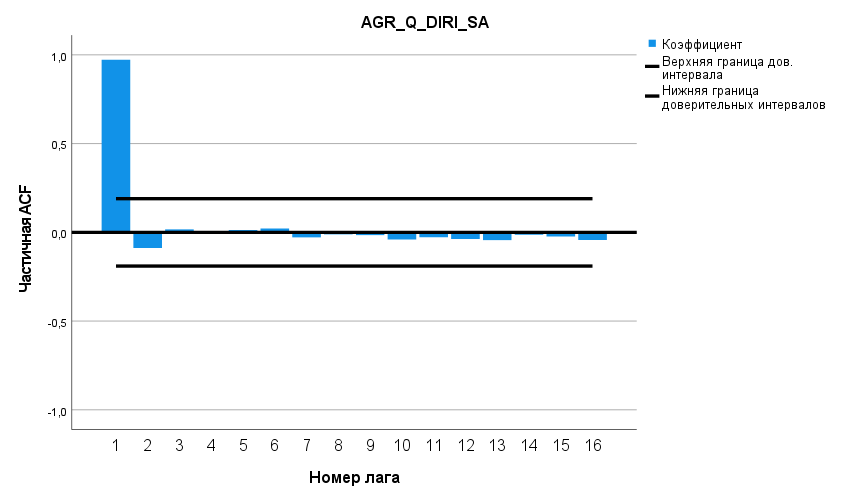


На данном этапе сразу заметен позитивный тренд с 1999 года.

Далее проанализируем наш ряд на наличие автокорреляции. Коэффициент автокорреляции показывает тесноту связи между уровнями ряда, сдвинутыми на 1 или более шаг.

На графике частных коэффициентов автокорреляции видно, что максимальное значение на 1 лаге, она равна приблизительно единице. Полученные значения коэффициента автокорреляции 1го и 2го порядков свидетельствует о тесноте зависимости между текущими уровнями ряда и уровнями ряда предыдущих периодов, а также свидетельствует о линейной тенденции.





Далее подбираем модель наиболее удача описывающую временной ряд. Воспользуемся мастером моделей временных рядов на полном наборе данных. Наилучшее качество у модели ARIMA (2,1,3), где 2-порядок модели, 1- порядок разности , 3- порядок скользящего среднего.

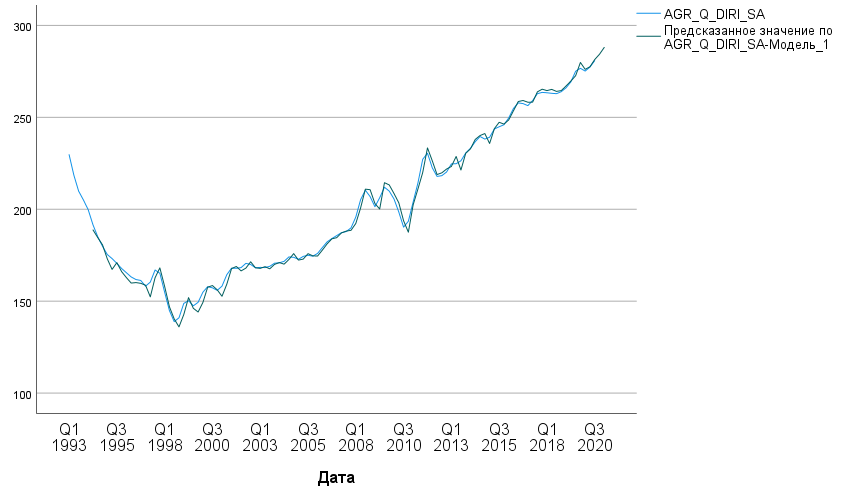
Рассмотрим характеристики модели-



Первые два значения указывают на то, что наша модель может хорошо описывать наш набор данных. Значения ошибок также не сильно большие для наших данных.



Значимость модели хорошая. Попробуем визуализировать насколько хорошо модель описывает временной ряд.

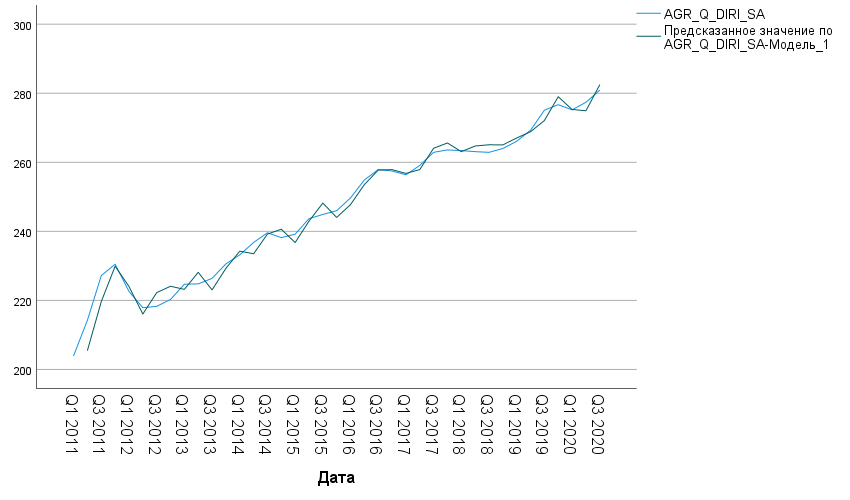


Видно, что на графике реальные данные и данные, предсказанные моделью, во многом совпадают.

Далее попробуем удалить данные до 2010 года и обучить модели на новом наборе, чтобы сравнить, будет ли модель с такими же параметрами выбрана мастером моделей временных рядов. Для урезанного набора данных мастер моделей выбрал ARIMA ( 2,1,0). Но если посмотреть на качество модели и ошибки, то становится очевидно, что качество модели упало. Среднее отклонение – 2,04.



**Реальные VS предсказанные данные. Модель ARIMA (2, 1, 0)**

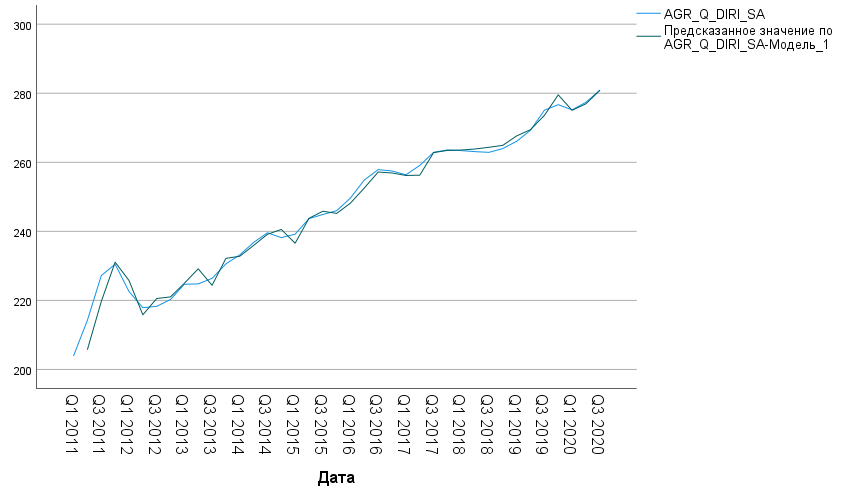


Попробуем построить модель на урезанных данных, выбранную на первом этапе, когда мы использовали весь датасет- ARIMA (2,1,3). Среднее отклонение- 1,5.



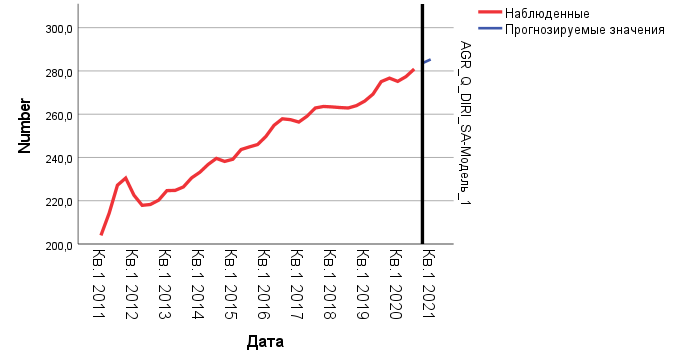
Очевидно, что она справляется лучше. Разница в рассматриваемых моделях в порядке скользящего среднего.

**Реальные VS предсказанные данные. Модель ARIMA (2, 1, 3)**



Далее были проделаны попытки вручную изменять коэффициенты ARIMA модели, но ни один вариант не дал результатов лучше, чем изначально выбранная ARIMA(2,1,3).

**Визуализация предсказания:**



**Источники:**

1. Единый архив экономических и социологических данных. URL: <http://sophist.hse.ru/hse/nindex.shtml>
2. Документация SPSS, Модуль Прогнозирование и Деревья решений. URL: https://www.ibm.com/support/knowledgecenter/ru/SSLVMB\_sub/statistics\_mainhelp\_ddita/spss/trends/idh\_idd\_tab\_vars.html