

ПРИМЕНЕНИЕ МОДЕЛЕЙ ARCH И GARCH ДЛЯ АНАЛИЗА ОБЪЕМА ДОБЫЧИ СЫРОЙ НЕФТИ В СТАВРОПОЛЬСКОМ КРАЕ

Димурина Н.Е., аспирант,
Северо-Кавказский федеральный университет

Аннотация: в данной статье рассмотрены теоретические аспекты построения моделей ARCH и GARCH, уделено внимание рассмотрению таких показателей, как куртозис и коэффициент эксцесса. При сравнении моделей ARCH и GARCH используются информационные критерии Акаике, Шварца, Ханнан-Куина и лучшей моделью будет та, которая имеет минимальные значения этих показателей. Для анализа объема добычи сырой нефти в Ставропольском крае сначала построен график временного ряда и коррелограммы ACF и PACF, в которых обнаруживается процесс случайного блуждания. ACF плавно убывает, стремясь к нулю с 12-го временного периода, а PACF – конечная, с обрывом после 1 периода. Оценены параметры изменения стоимости добычи сырой нефти для моделей ARCH (1) и GARCH (1,1). Коррелограмма квадратов остатков этих моделей показала, что остатки не коррелированы. Проведено сравнение результатов куртозиса, рассчитанного программой GRETЛ для модели GARCH (1,1) и теоретического куртозиса, рассчитанного по формуле. Сделал вывод, что модель ARCH(1) является наиболее подходящей для оценивания данного временного ряда, так как значения критериев AIC, SIK, HQIC у данной модели меньше, чем у модели GARCH (1,1).

Ключевые слова: временной ряд, ARCH- и GARCH-модели, математическое ожидание, дисперсия, куртозис, коррелограмма

На протяжении последних десятилетий моделирование и прогнозирование динамики различных показателей является очень актуальной темой, так как изменчивость финансов, в том числе цен, все сложнее предсказать. Основным из этих показателей можно считать волатильность, под которой в широком смысле понимают изменчивость, вариацию во времени величины финансовых активов, таких как ценные бумаги с фиксированным доходом, обменные курсы, индивидуальные акции и фондовые индексы. Волатильность имеет свойство кластеризации, характеризующееся тем, что, как правило, большие изменения порождают изменения такого же масштаба, а за малыми колебаниями следуют незначительные колебания, любого знака.

Применение таких традиционных моделей временных рядов, как ARMA, не позволяет правильно оценить характеристики финансовых активов в настоящее время. Поэтому, необходимо усовершенствование таких моделей.

Разработанные Р. Энглом (1982 г.) и Р. Боллерсломом (1986 г.) [3], авторегрессионная модель условной гетероскедастичности (ARCH-модель) и ее обобщенный вариант (GARCH-модель) предна-

значены для учета эмпирических закономерностей финансовых показателей. Они и до сих пор являются самыми известными и признанными, часто используемыми моделями для анализа временных рядов, так как учитывают периоды сильных колебаний, и периоды затишья. В последствии эти исследователи предложили и иные вариации этих моделей, учитывающие те или иные особенности. Процесс разработки и усовершенствования моделей временных рядов продолжается и до сих пор. Появляются новые модификации, такие как GARCH-M, EGARCH, AGARCH, TGARCH, GJR-GARCH и другие.

ARCH модель в эконометрике – это модель анализа финансовых динамических рядов, содержащих зависимость условной дисперсии ряда (по предыстории) от предыстории временного ряда, его дисперсий и иных показателей. Ее применяют для моделирования волатильности доходности финансовых активов.

Требуется разобрать ARCH модель q-го порядка (ARCH(q)) [2]:

$$\varepsilon_t = \sqrt{h_t} v_t, v_t \sim i. i. d. N(0,1)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 \quad (1)$$

В случае равенства математического ожидания модели нулю, условная дисперсия принимает значение h_t :

$$\begin{aligned} E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) &= \sqrt{h_t} E(v_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = 0 \\ D(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) &= \sqrt{h_t} D(v_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = h_t \end{aligned} \quad (2)$$

Следовательно, h_t есть условная дисперсия показателя ε_t относительно предыстории ε_t .

$$\alpha(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \dots + \alpha_q z^q$$

>1 и у процесса ARCH(q) имеется безусловная дисперсия.

Теперь необходимо вычислить куртозис для процесса ARCH (1), который описывается уравнениями [2]:

$$\begin{aligned} \varepsilon_t &= \sqrt{h_t} v_t \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \end{aligned} \quad (3)$$

или по-другому:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + w_t \quad (4)$$

$$\kappa = \frac{E(\varepsilon_t - E\varepsilon_t)^4}{(E(\varepsilon_t - E\varepsilon_t)^2)^2} = \frac{E(\varepsilon_t^4)}{(E\varepsilon_t^2)^2}, \text{ т. к. } E(\varepsilon_t) = 0 \quad (6)$$

Необходимо подробнее рассмотреть κ :

$$\kappa = \frac{3\alpha_0^2(1-\alpha_1^2)}{\frac{(1-\alpha_1)^2(1-3\alpha_1^2)}{\alpha_0^2}} = \frac{3(1-\alpha_1^2)}{(1-3\alpha_1^2)} \quad (7)$$

Кроме того, наравне с показателем «куртозис», используется коэффициент эксцесса γ [2, 3]. Их отличие в том, что измерение остроты пика распределения случайной величины рассчитывается,

когда $\kappa > 3$ при $\alpha_1^2 < \frac{1}{3}$, а $\gamma = \kappa - 3$. Для нормального

распределения γ должно быть равно 0. Следует рассчитать его для рассматриваемой модели:

$$\gamma = \frac{3(1-\alpha_1^2)}{(1-3\alpha_1^2)} - 3 = \frac{6\alpha_1^2}{1-3\alpha_1^2} \quad (8)$$

Из формулы (8), видно, что $\gamma > 0$. Этот факт указывает на остроту вершины.

$$E\varepsilon_t^4 = E(\varepsilon_t^2 * \varepsilon_t^2) = K(0) + E\varepsilon_t^2 E\varepsilon_t^2 = \frac{3(1+\alpha_1-\alpha_2)(1+\alpha_2)\alpha_0^2}{(-1+\alpha_1+\alpha_2)(-1+\alpha_2+3\alpha_2^2-3\alpha_2^3+3\alpha_1^2(1+\alpha_2))} \quad (12)$$

Тогда:

$$\begin{aligned} \kappa &= \frac{3(1+\alpha_1-\alpha_2)(1+\alpha_2)\alpha_0^2}{\frac{(-1+\alpha_1+\alpha_2)(-1+\alpha_2+3\alpha_2^2-3\alpha_2^3+3\alpha_1^2(1+\alpha_2))}{\alpha_0^2}} = \\ &= \frac{3(1+\alpha_1-\alpha_2)(1+\alpha_2)(-1+\alpha_1+\alpha_2)}{-1+\alpha_2+3\alpha_2^2-3\alpha_2^3+3\alpha_1^2(1+\alpha_2)} \end{aligned} \quad (13)$$

Куртозис κ по-новому выглядит так:

$$\kappa = \frac{3(-1+\alpha_1(-1+m))(1+\alpha_1 m)(-1+\alpha_1+\alpha_1 m)}{1-\alpha_1 m+3\alpha_1^3 m(-1+m^2)-3\alpha_1^3(1+m^2)} \quad (14)$$

Определим экстремумы функции $\kappa(\alpha_1)$:

Ряд ε_t^2 стационарен только при условии, что все корни уравнения $\alpha(z) = 0$,

Куртозис κ процесса ε_t (показатель, измеряющий остроту пика одномерного распределения и толщину его хвостов) выглядит следующим образом:

$$\kappa = \frac{\mu_4}{\sigma^4} \quad (5)$$

где μ_4 – четвертый центральный момент распределения, σ^4 – квадрат дисперсии. У куртозиса нормального распределение = 3. Поэтому:

Далее перейдем к описанию κ для процесса ARCH 2-го порядка, который описывается уравнениями [2]:

$$\begin{aligned} \varepsilon_t &= \sqrt{h_t} v_t \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 \end{aligned} \quad (9)$$

или иначе:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + w_t \quad (10)$$

Рассчитаем κ процесса ε_t :

$$\kappa = \frac{E(\varepsilon_t - E\varepsilon_t)^4}{(E(\varepsilon_t - E\varepsilon_t)^2)^2} = \frac{E(\varepsilon_t^4)}{(E\varepsilon_t^2)^2} \quad (11)$$

Распишем четвертый момент распределения $E\varepsilon_t^4$ в виде:

$$\kappa'(\alpha_1) = \frac{12\alpha_1^2(\alpha_1 + m - 2\alpha_1 m^2 + \alpha_1^2 m^3)}{(-1 + \alpha_1 m - 3\alpha_1^3 m(-1 + m^2) + 3\alpha_1^2(1 + m^2))^2} \quad (15)$$

Если упростить числитель, можно получить:

$$\alpha_1 + m - 2\alpha_1 m^2 + \alpha_1^2 m^3 = \alpha_1 + m(1 - 2\alpha_1 m + \alpha_1^2 m^2) = \alpha_1 + m(1 - \alpha_1 m)^2 \quad (16)$$

В случае, если $\alpha_1 > 0$, $\alpha_1 < \frac{1}{\sqrt{3}}$, $m > 0$, получим:

$$\kappa'(\alpha_1) > 0 \quad (17)$$

По результатам проведенных вычислений, можно сказать, что куртозис κ - это возрастающая функция.

При $\alpha_1 = 0$ и $m = 0$ значение $\kappa = 3$, следовательно, $\kappa > 3$, т.е. при:

$$0 < \alpha_1 < \sqrt{\frac{1 - \alpha_2 - 3\alpha_2^2 + 3\alpha_2^3}{1 + \alpha_2}} \text{ и } 0 < \alpha_2 < \frac{1}{\sqrt{3}} \quad (18)$$

А γ имеет вид:

$$\gamma = \frac{-6(-(-1 + \alpha_2)\alpha_2^2 + \alpha_1^2(1 + \alpha_2))}{-1 + \alpha_2 + 3\alpha_2^2 - 3\alpha_2^3 + 3\alpha_1^2(1 + \alpha_2)} \quad (19)$$

При $\kappa > 3$, видно, что $\gamma > 0$.

Основная цель модели ARCH заключается в кластеризации волатильности, то есть если величина ε_t возрастает, она влияет на рост значений условной дисперсии h_t , а при высокой h_t более вероятно возникновение больших значений ε_t . И наоборот, если значения ε_t в течение определенного времени приближались к нулю, то это ведет к уменьшению h_t в дальнейшем. Также при низкой h_t более вероятно возникновение малых значений ε_t .

Кроме того ARCH-процессы обладают еще одним свойством финансовых динамических рядов,

состоящем в безусловном распределении процесса ε_t имеющем более высокое значение куртозиса κ .

Подытоживая вышесказанное, отметим, что ARCH модель имеет некоторые преимущества перед традиционными моделями временных рядов, проявляющиеся в анализе кластерности и тяжёлых хвостов волатильности.

На следующем этапе необходимо рассмотреть модели GARCH.

Более обобщенно процесс зависимости условной дисперсии h_t от предыстории ε_i^2 можно записать в виде [2, 3]:

$$h_t = \alpha_0 + \pi(L)\varepsilon_t^2 \quad (20)$$

где πL - лаговый оператор, $LX_t = X_{t-1}$.

Запишем $\pi(L)$ в виде отношения:

$$\pi(L) = \frac{\alpha(L)}{1 - \beta(L)} = \frac{\alpha_1 L^1 + \alpha_2 L^2 + \dots + \alpha_m L^m}{1 - \beta_1 L^1 - \beta_2 L^2 - \dots - \beta_r L^r} \quad (21)$$

Из формулы (21) посчитаем, что корни уравнения $1 - \beta(z) = 0$ больше 1. Тогда, умножив h_t на $1 - \beta(L)$, можно получить:

$$[1 - \beta(L)]h_t = [1 - \beta(L)]\alpha_0 + \alpha(L)\varepsilon_t^2 \quad (22)$$

Или

$$h_t = \alpha_0^* + \beta_1 h_{t-1} + \beta_2 h_{t-2} + \dots + \beta_r h_{t-r} + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_m \varepsilon_{t-m}^2 \quad (23)$$

где

$$\alpha_0^* = [1 - \beta_1 - \beta_2 - \dots - \beta_r]\alpha_0 \quad (24)$$

Это представление условной дисперсии h_t и называется обобщенной авторегрессионной моделью условной гетероскедастичности (generalized ARCH, GARCH). Память ARCH(q) процесса ограничена q периодами и при ее использовании часто необходим длинный лаг q и большое число параметров. GARCH модель же

имеет бесконечную память и допускает более экономную параметризацию.

Далее следует провести вычисление куртозиса κ для модели GARCH(1,1), который описывается уравнением [2]:

$$\varepsilon_t = \sqrt{h_t} v_t$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (25)$$

или иначе:

$$\varepsilon_t^2 = \alpha_0^* + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + w_t - \beta_1 w_{t-1}$$

Куртозис

$$K = \frac{3Eh_t^2}{(Eh_t)^2} = \frac{3 \frac{(1+\alpha_1+\beta_1)\alpha_0^2}{(-1+\alpha_1+\beta_1)(-1+3\alpha_1^2+2\alpha_1\beta_1+\beta_1^2)}}{\left(\frac{\alpha_0}{1-\alpha_1-\beta_1}\right)^2} \quad (26)$$

Получим

$$K = \frac{3(1+\alpha_1+\beta_1)(-1+\alpha_1+\beta_1)}{-1+3\alpha_1^2+2\alpha_1\beta_1+\beta_1^2} \quad (27)$$

Так как числитель этой дроби является отрицательным, поэтому и знаменатель должен быть того же знака.

Следует отметить, что GARCH-процесс, как и его частный случай ARCH-процесс имеет более высокое значение куртозиса k , чем нормальное распределение, при этом рассматривая отдельное наблюдение GARCH-процесса, можно сказать, что его безусловное распределение является

симметричным, поэтому все нечётные значения, начиная с третьего, равны нулю.

Для сравнения полученных моделей ARCH и GARCH можно использовать информационные критерии Акаике, Шварца, Ханнан-Куина. Главной целью является получение модели с минимальными значениями информационных критериев.

Критерий Акаике (AIC) описывается формулой [2, 4]:

$$AIC = \ln(\hat{\delta}^2) + \frac{2k}{T} \text{ или } AIC = \ln s^2(m) + \frac{2m}{n} \rightarrow \min \quad (28)$$

Критерий Шварца (SIC) выглядит следующим образом [2, 4]:

$$SIC = \ln(\hat{\delta}^2) + \frac{k}{T} \ln T \text{ или } SIC = \ln s^2(m) + \frac{m \ln n}{n} \rightarrow \min \quad (29)$$

Критерий Хеннана-Куина (HQIC) [2, 4]:

$$HQIC = \ln(\hat{\delta}^2) + \frac{2k}{T} \ln T \text{ или } HQ(m) = \ln s^2(m) + 2c \frac{m \ln \ln n}{n} \rightarrow \min,$$

$$c > 1 \quad (30)$$

где $\hat{\delta}^2$ – остаточная дисперсия, T – объем выборки.

Как подчеркивалось выше, выбор следует сделать в пользу модели с меньшим значением AIC, SIC и HQIC.

Проведем анализ динамики объема добычи сырой нефти (млн. руб.) с 1 полугодия 2012 г. по 1 полугодие 2019 г. в Ставропольском крае с использованием моделей ARCH и GARCH.

Проведем анализ модели ARCH (1). Информация об оценках показателей модели указана на рис. 1-3.

Сначала необходимо построить график временного ряда для оценивания динамики изменения объема добычи сырой нефти с 01.01.2012 г. до 01.07.2019 г. в Ставропольском крае [1].

Анализ автокорреляционной функции представлен на рис. 1 (максимальный лаг=14).

При обзорном анализе графика временного ряда и коррелограмм ACF и PACF обнаруживается процесс случайного блуждания. ACF плавно убывает, опускаясь ниже нуля в 6 временном периоде, и достигая пика в 11 временном периоде, а с 12-го периода начинает постепенный рост и стремится к нулю.

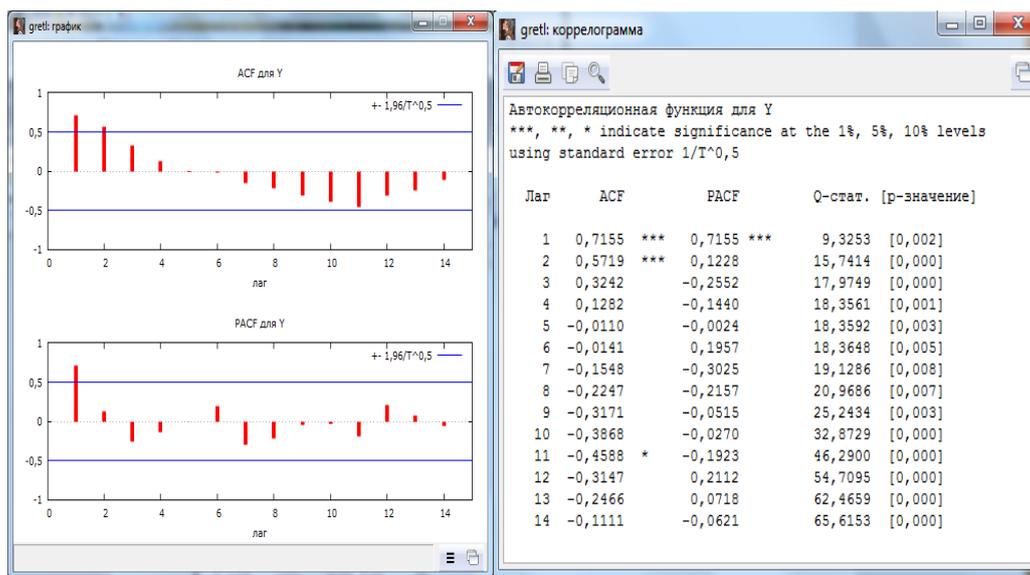


Рис. 1. Функция автокорреляции временного ряда

Из рисунка видно, что коррелограмма PACF – конечная, с обрывом после 1 периода.

Построим модель ARCH (1) (рис. 2).

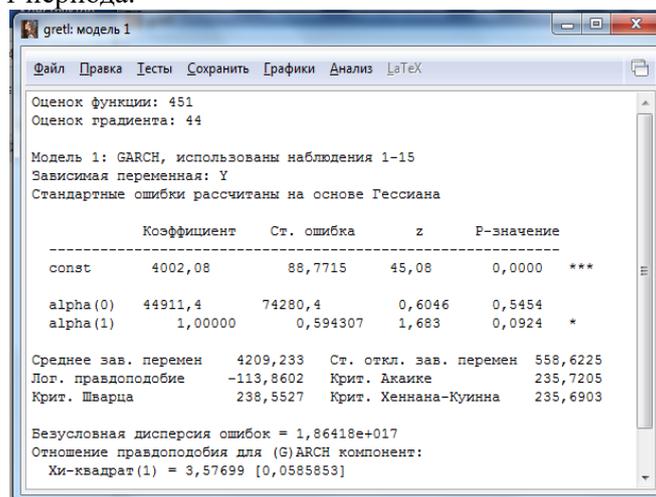


Рис. 2. Оценивание параметров изменения стоимости добычи сырой нефти для модели ARCH (1)

Значения коррелограммы квадратов остатков ε_t^2 выглядят следующим образом (рис. 3):

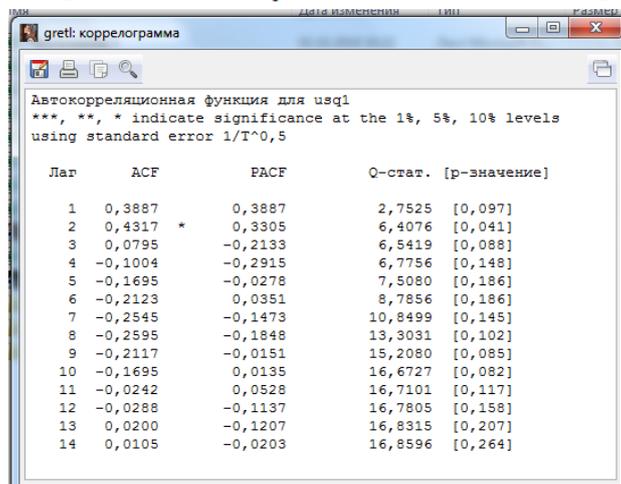


Рис. 3. Коррелограмма квадратов остатков ε_t^2 модели ARCH (1)

Из рисунка видно, что остатки не коррелированы.
 Далее рассмотрим модель GARCH (1,1). Ин-

формация об оценках показателей модели указаны на рис. 4-5.

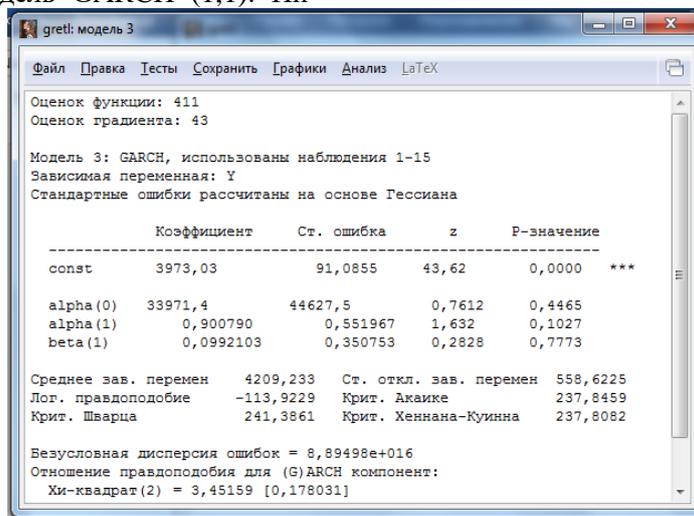


Рис. 4. Показатели модели GARCH (1,1) стоимости добычи сырой нефти

Коррелограмму ε_t^2 данной модели можно представить в виде:

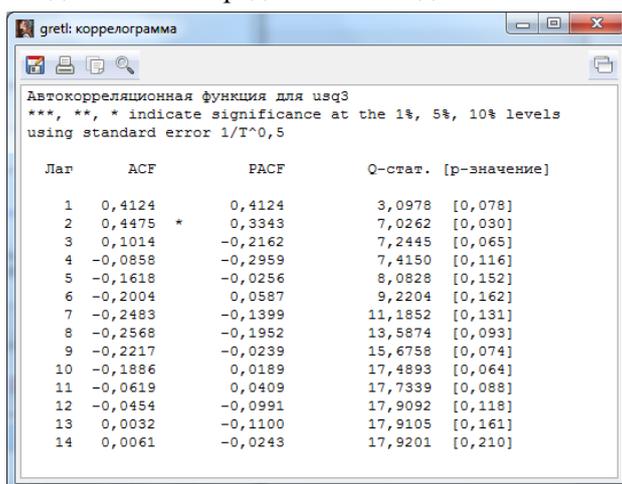


Рис. 5. Коррелограмма ε_t^2 модели GARCH (1,1) показателя добычи сырой нефти

Из рис. 5 можно сказать, что остатки не коррелированы.

GARCH (1,1) и теоретический куртозис, рассчитанный по формуле (рис. 6).

Теперь можно сравнить куртозис, рассчитанный программой GRETl для модели

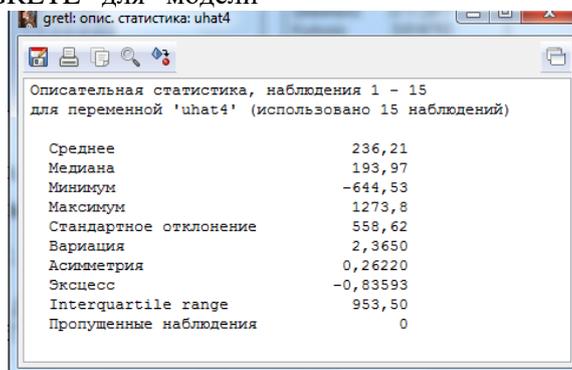


Рис. 6. Описательная статистика остатков модели GARCH (1,1)

Находим программный куртозис – -0,83593.

$\alpha_1=0,900790$, $\beta_1=0,0992103$

$$k = \frac{3(1 + \alpha_1 + \beta_1)(-1 + \alpha_1 + \beta_1)}{-1 + 3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2} = 0,000001109 \quad -$$

теоретический куртозис.

По результатам расчетов видно, что значения различаются. Этому есть свое объяснение. В теоретической части принималось $\varepsilon_t = 0$, в то же время

программой были оценены ошибки по-другому, и среднее значение из рис. 6 $\varepsilon_t=236,21$.

Сравнение моделей ARCH(1) и GARCH (1,1) стоимости добычи сырой нефти (млн.руб.) с 1 полугодия 2012 г. по 1 полугодие 2019 г. в Ставропольском крае по информационным критериям AIC, SIK, HQIC представим в табл. 1:

Таблица 1

Анализ моделей ARCH (1) и GARCH (1,1) по информационным критериям AIC, SIK, HQIC

Модель	Значение AIC	Значение SIK	Значение HQIC
ARCH (1)	235,7205	238,5527	235,6903
GARCH (1,1)	237,8459	241,3861	237,8082

Из таблицы видно, что модель ARCH (1) является наиболее подходящей для оценивания временного ряда стоимости добычи сырой нефти, так как значения критериев AIC, SIK, HQIC у данной модели меньше, чем у модели GARCH (1,1).

Подытоживая вышесказанное, были рассмот-

рены свойства моделей анализа временных рядов ARCH и GARCH, был проведен анализ временного ряда стоимости добычи сырой нефти в Ставропольском крае с помощью данных моделей и выполнено сравнение и подбор лучшей из них.

Литература

1. *Официальный сайт управление Федеральной службы государственной статистики по Северо-Кавказскому федеральному округу [Электронный ресурс]. Режим доступа: URL: <https://stavstat.gks.ru> (дата обращения: 01.11.2019)*
2. *ARCH и GARCH модели временных рядов [Электронный ресурс]. Режим доступа: URL: <https://www.hse.ru/data/2014/06/09/1324317113/Диплом.pdf> (дата обращения: 05.11.2019)*
3. *Айвазян С.А., Фантаццини Д. Эконометрика-2: продвинутый курс с приложениями в финансах: учебник. М.: Магистр: Инфра-М, 2014. 944 с.*
4. *Информационные критерии [Электронный ресурс]. Режим доступа: URL: <https://metr-ekon.ru/index.php?request=full&id=399> (дата обращения: 15.11.2019)*

References

1. *Oficial'nyj sajt upravlenie Federal'noj sluzhby gosudarstvennoj statistiki po Severo-Kavkazskomu federal'nomu okrugu [Elektronnyj resurs]. Rezhim dostupa: URL: <https://stavstat.gks.ru> (data obrashcheniya: 01.11.2019)*
2. *ARCH i GARCH modeli vremennyh ryadov [Elektronnyj resurs]. Rezhim dostupa: URL: <https://www.hse.ru/data/2014/06/09/1324317113/Diplom.pdf> (data obrashcheniya: 05.11.2019)*
3. *Ajvazyan S.A., Fantaccini D. Ekonometrika-2: prodvinutyj kurs s prilozheniyami v finansah: uchebnik. M.: Magistr: Infra-M, 2014. 944 s.*
4. *Informacionnye kriterii [Elektronnyj resurs]. Rezhim dostupa: URL: <https://metr-ekon.ru/index.php?request=full&id=399> (data obrashcheniya: 15.11.2019)*

APPLICATION OF ARCH AND GARCH MODELS FOR ANALYSIS OF CRUDE OIL PRODUCTION VOLUME IN STAVROPOL REGION

*Dimurina N.E., Postgraduate,
North Caucasus Federal University*

Abstract: this article discusses the theoretical aspects of the construction of ARCH and GARCH models, paid attention to the consideration of indicators such as kurtosis and excess coefficient. When comparing ARCH and GARCH models, the information criteria of Akaike, Schwartz, Hannan-Quinn are used and the best model will be the one that has the minimum values of these indicators. To analyze the volume of crude oil production in the Stavropol region, a time series graph and ACF and PACF correlograms are first constructed, in which the process of random walk is detected. The ACF gradually decreases, tending to zero from the 12th time period, and the PACF is finite, with a break after the 1st period. The parameters of crude oil production cost change for ARCH (1) and GARCH (1,1) models are estimated. The correlogram of the squares of the residuals of these models showed that the residuals are not correlated. The results of the kurtosis calculated by the GRETl program for the GARCH (1,1) model and the theoretical kurtosis calculated by the formula are compared. It was concluded that the ARCH(1) model is the most suitable for estimating this time series, as the values of the AIC, SIK, HQIC criteria for this model are less than those of the GARCH(1,1) model.

Keywords: time series, ARCH- and GARCH- models, expectation, variance, kurtosis, correlogram