

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ИЗМЕНЕНИЯ ПАРАМЕТРОВ МОДЕЛИ ARIMA НА КАЧЕСТВО ПРОГНОЗА ДЛЯ КОРОТКИХ НАБОРОВ ДАННЫХ

М.Н. Фелькер, В.В. Чеснов

*Пермский национальный исследовательский политехнический университет,
Березниковский филиал, г. Березники, Россия*

Временные ряды – это данные, собираемые в различные отрезки времени, которые, в свою очередь, в зависимости от задачи могут отличаться. Временные ряды используются для принятия решений. Анализ временных рядов позволяет получить некоторый результат, который определит характер принимаемого решения. Анализом временных рядов занимались в очень давние времена, например, следствием анализа стали различные календари. Позднее анализ временных рядов применялся для исследования и прогнозирования экономических, социальных и других систем. Временные ряды появились очень давно. Когда-то древнеави-лонские астрономы, изучая положение звезд, обнаружили периодичность затмений, что позволило в дальнейшем прогнозировать их появление. Позднее анализ временных рядов подобным образом привел к созданию различных календарей, например, урожайных. В дальнейшем помимо естественных областей добавились социальные и экономические. **Цель исследования:** поиск классификационных признаков временных рядов, позволяющих понять, можно ли для их краткосрочного (3 отсчета) прогноза применять модель *ARIMA*. **Материалы и методы.** Разработано специальное программное обеспечение, реализующее модели семейства *ARIMA* и необходимые интерфейсы. В работе были исследованы 59 наборов годовых данных с малой длиной, менее 20 значений. Данные обрабатывались с помощью *Python* библиотек *Statsmodels*, *Pandas*. Для определения стационарности ряда использовался тест Дики – Фуллера. Стационарность временного ряда позволяет более качественно строить прогнозы. Для выбора наилучшей модели применялся информационный критерий Акаике. Получены рекомендации по обоснованному подбору параметров настройки *ARIMA*-моделей. Показана зависимость настроек от категории годовых рядов. **Заключение.** После обработки данных были выделены четыре категории, или шаблона, годовых рядов. В зависимости от категории были подобраны диапазоны параметров для настройки *ARIMA*-моделей. Предлагаемые диапазоны позволят определить начальные параметры для исследования аналогичных наборов данных. Даны рекомендации по улучшению качества постпрогноза и прогноза при помощи *ARIMA*-модели за счет подбора настроек.

Ключевые слова: *ARIMA*, эконометрика, анализ, временные ряды.

Введение

Временные ряды появляются в результате измерения неких показаний, полученных от технических, природных, социальных, экономических и других систем [1, 2–7]. В зависимости от задачи, рисков и требуемой скорости принятия решения могут быть различные промежутки между сборами [1, 8]. Например, системы с высоким риском, такие как биржи, собирают статистику с секундным периодом, так как в любой момент (секунду) времени может поменяться тенденция биржевого курса. Менее рискованные системы используют часовые или суточные сборы. Для безрисковых систем или систем со слабым риском необходимость в частых сборах отсутствует и для таких систем используется ежемесячный, ежеквартальный, ежегодный период сбора.

Анализ временных рядов позволяет получить некоторый результат, который каким-либо образом повлияет на принятие решения. Опять же на примере биржи: анализ ряда приведет к быстрому принятию решения, что в свою очередь позволит увеличить доход или минимизировать потери.

1. Модель *ARIMA*

Интегрированная модель авторегрессии – скользящего среднего (*ARIMA*) – одна из моделей, предназначенных для анализа временных рядов. Основана на модели *ARMA* с тем отличием, что

позволяет работать с нестационарными временными рядами, которые можно сделать стационарными, используя разность некоторого порядка от исходного ряда [1, 8, 9].

Стационарность – свойство процесса сохранять свои статистические характеристики с течением времени. Должны сохраняться постоянство мат. Ожидания, дисперсии и независимость от времени. Стационарность позволяет более точно строить прогноз [1, 8, 9].

Модель ARMA состоит из нескольких моделей – модель авторегрессии (AR) и модель скользящего среднего (MA) [10].

AR основывается на линейной зависимости текущих значений ряда от предыдущих [11, 12]:

$$X_t = c + \sum_{i=1}^p a_i X_{t-i} + \varepsilon_t, \quad (1)$$

где a_i – параметры модели; c – константа; ε_t – белый шум.

MA используется для моделирования случайных ошибок временных рядов [13]:

$$X_t = c + \sum_{j=1}^q a_j \varepsilon_{t-j}, \quad (2)$$

где a_i – параметры модели; ε_t – белый шум.

Для анализа временных рядов с помощью модели ARIMA используются три параметра: p , d , q [9].

График (диаграмма) функции корреляции – коррелограмма [9].

P – порядок модели AR: порядок, определяемый с помощью коррелограммы функции частичной автокорреляции (PACF). Здесь используется корреляция между y_t и y_{t-k} :

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_k y_{t-k} + \varepsilon_t; \quad (3)$$

d – порядок интегрированного ряда (I): порядок, определяемый количеством разностей для получения из нестационарного ряда стационарного;

Q – порядок модели MA: порядок, определяемый с помощью коррелограммы функции автокорреляции (ACF).

$$r_k = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}, \quad k = 1, 2, \dots \quad (4)$$

Для того чтобы определить порядок (d) интегрированного временного ряда, проводится тест Дики – Фуллера [1, 14]:

$$\Delta y = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t, \quad (5)$$

где α – константа, предполагается, что единичный корень один ($\alpha = 1$); β – коэффициент тренда; p – лаг процесса авторегрессии.

Данный тест проводится для поиска единичных корней, т. е. характеризующих свойств нестационарных временных рядов. Если исходный ряд не стационарен, то его можно сделать стационарным путем взятия разности временного ряда с исходным с некоторым сдвигом:

$$\Delta y = y_t - y_{t-1}. \quad (6)$$

Например, если первая разность ряда стационарна, то это интегрированный ряд первого порядка.

Другой способ определения – на основе коррелограмм ACF, PACF.

Коррелограмма ACF стационарного временного ряда быстро убывает с ростом k (значение коррелограммы после 1, 2, 3-го лага стремится к нулю). Аналогичным образом стационарность определяется коррелограммой PACF.

Еще один метод – наглядный. Нестационарный ряд можно определить по графику. Если на графике виден тренд, то временной ряд нестационарный.

После определения порядка (d) интегрированного временного ряда определяются порядки (p и q) компонент AR и MA.

В различных источниках [9, 15, 16] упоминаются два способа определения значений p , q .

Первый способ – подсчет количества лагов на коррелограммах.

Для q – количество значимых лагов на коррелограмме ACF, после которых значения функции резко падает, пример на рис. 1.

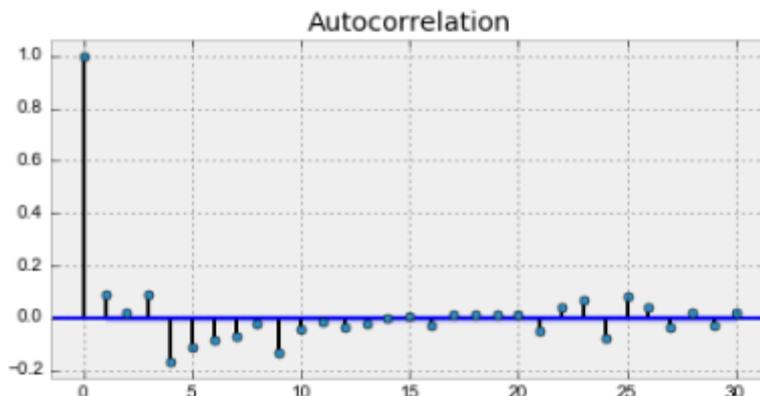


Рис. 1. Пример коррелограммы ACF
Fig. 1. An example of an ACF correlogram

На данной диаграмме видно, что после первого лага происходит резкое снижение значений функции.

Для p – количество лагов на коррелограмме PACF, сильно отличных от нуля, пример на рис. 2.

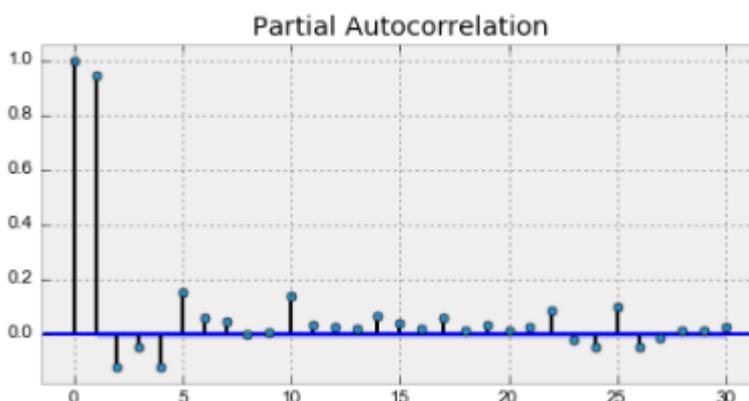


Рис. 2. Пример коррелограммы PACF
Fig. 2. An example of a PACF correlogram

На данной диаграмме видно, что только два лага сильно отличны от нуля.

Далее полученные значения p , d , q подставляются в модель ARIMA и применяется для исследуемого временного ряда.

Второй – модификация первого, для параметров берутся значения p и q с запасом. Для всех перестановок значений p и q в диапазонах $[0; p + n]$ и $[0; q + m]$, n и m больше или равно 1, и известной d вычисляется наилучшая модель для исследуемого временного ряда.

Наилучшая модель выбирается с помощью некоторого информационного критерия. В данном исследовании при выборе лучшей модели был задействован критерий Акаике для малых наборов данных (AICc) [16, 17]:

$$AICc = AIC + 2k \frac{k+1}{n-k-1} = 2k - 2\ln(L) + 2k \frac{k+1}{n-k-1}, \quad (7)$$

где k – число параметров модели; L – максимум функции правдоподобия; n – объем выборки.

На основании полученной модели можно спрогнозировать дальнейшее поведение системы, описываемой временным рядом. Для проверки качества модели проводится постпрогноз. Постпрогноз – метод, при котором известные последние значения ряда реакции принимаются неизвестными, после прогнозируют их с использованием модели, оснащенной на сокращенном интервале, и исходных рядов факторов, а затем сравнивают найденные значения с исходными.

Цель работы – поиск классификационных признаков временных рядов, позволяющих понять, можно ли для их краткосрочного (3 отсчета) прогноза применять модель ARIMA.

2. Анализ наборов данных и выделение основных категорий

Для данной работы были исследованы 59 наборов годовых рядов данных о больших социально-экономических системах и объектах, в том числе численность персонала, занятого научными исследованиями и разработками в России, уровень преступности в России, анализ продуктовой корзины и другие. Длина наборов данных – от 5 до 18 значений. На основе графиков наборов данных были выявлены четыре категории шаблонов.

Первая категория – наборы данных, график которых напоминает прямую и имеет явный тренд (рис. 3).

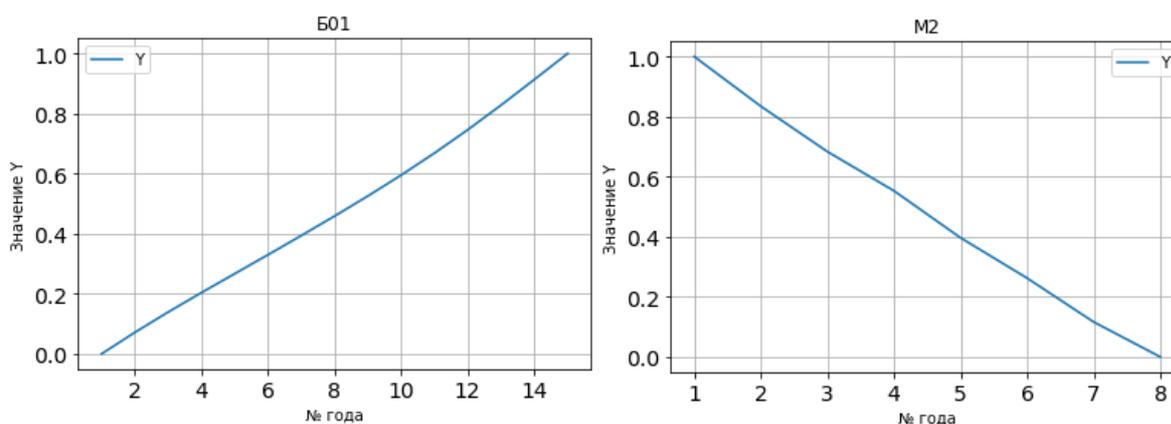


Рис. 3. Примеры графиков первой категории
Fig. 3. Examples of graphs of the first category

Схожесть с прямой можно описать математически: в первую очередь необходимо найти ее уравнение. Для этого используется уравнение прямой по двум точкам:

$$\frac{y - y_{\min}}{y_{\max} - y_{\min}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (8)$$

Затем сравниваются значения исходных данных со значениями прямой. Схожесть графика с прямой – максимальное отклонение между значениями набора данных и прямой. Чем меньше отклонение, тем больше схожесть с прямой.

При прогнозировании наборов данных первой категории трудностей нет. Параметр $d = 1$, это можно определить любым способом, описанным ранее, однако в данном случае тренд очевиден. Первый порядок интегрированности временного ряда достаточен для дальнейшего использования. Параметры p и q для данной категории лежат в диапазоне от 1 до 3.

В качестве примеров приведены графики наборов данных B01 и M2. Для обоих наборов наилучшая ARIMA-модель найдена при использовании параметров $p = 1$, $d = 1$, $q = 1$. Графики постпрогноза и прогноза представлены на рис. 4.

Вторая категория – это наборы данных, график которых не попадает в первую категорию, однако по-прежнему на графике можно наглядно определить направленный тренд. Примеры графиков таких наборов представлены на рис. 5.

При прогнозировании наборов данных второй категории начинаются трудности, они имеют более высокий порядок интегрированности $d \in [2, 3]$ из-за более сложной формы. Дальнейшее увеличение d приводит к изменению направления функции и появлению сильных неадекватных выбросов. Параметры p и q находятся в диапазоне $[1, 3]$.

В качестве примеров приведены графики наборов данных B07, B09, B24, M24. Для набора данных B07 найдена наилучшая модель ARIMA с параметрами: $p = 2$, $d = 2$, $q = 2$, для B09: $p = 1$, $d = 3$, $q = 1$, для B24: $p = 3$, $d = 2$, $q = 3$ и для M24: $p = 2$, $d = 2$, $q = 1$.

Графики постпрогноза и прогноза для B07, M24, представлены на рис. 6.

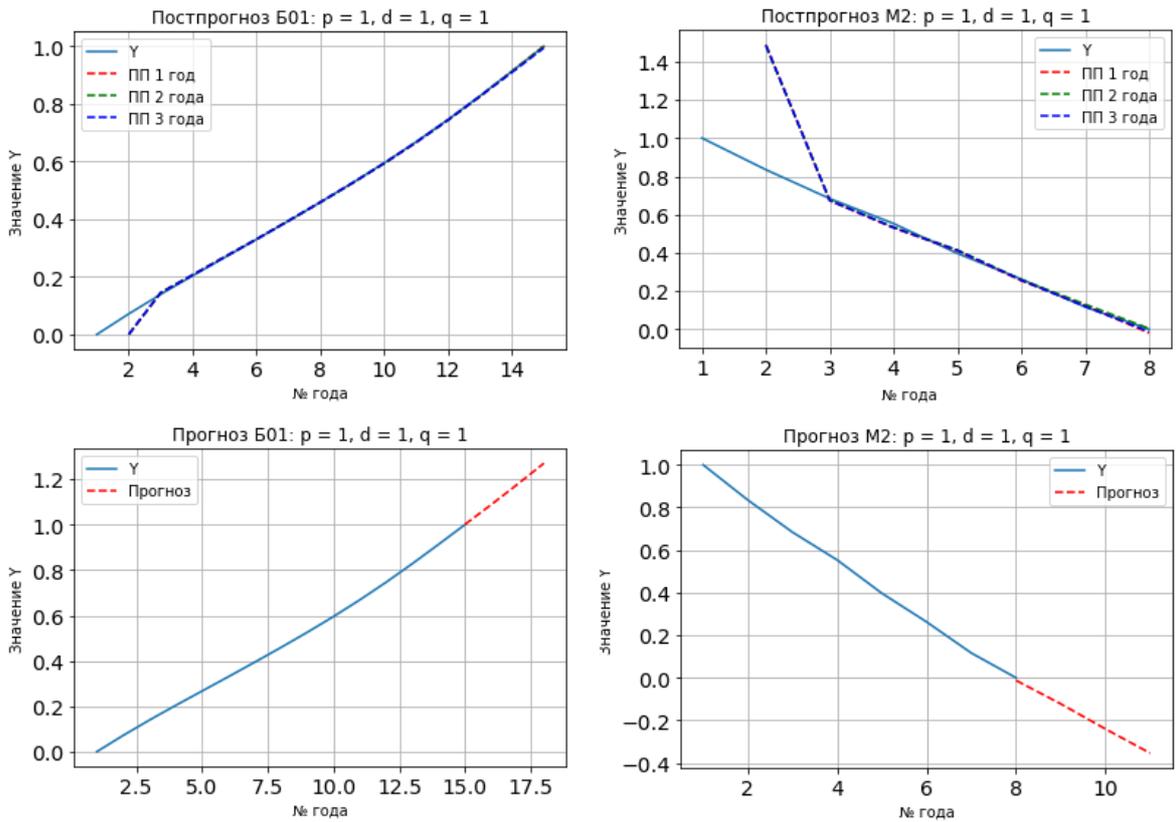


Рис. 4. Графики постпрогноза и прогноза наборов данных первой категории
 Fig. 4. Graphs of post-forecast and forecast of data sets of the first category

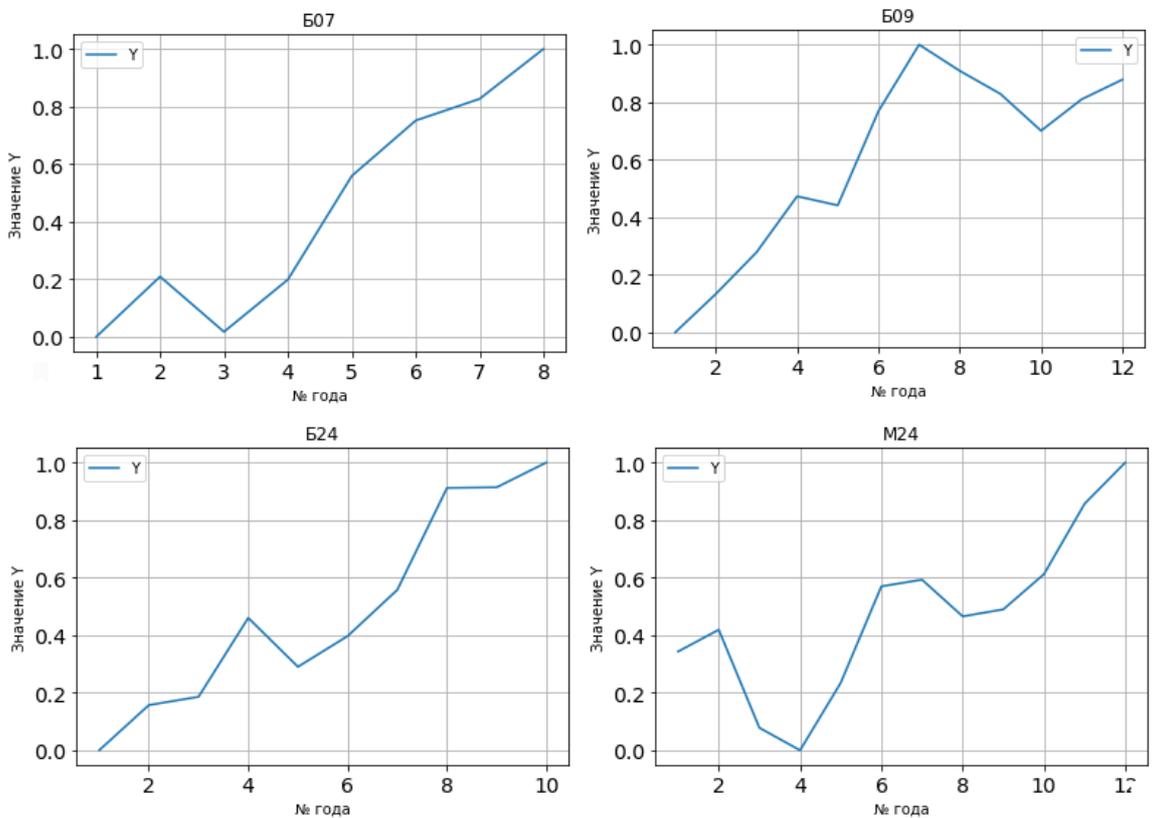


Рис. 5. Примеры графиков второй категории
 Fig. 5. Examples of graphs of the second category

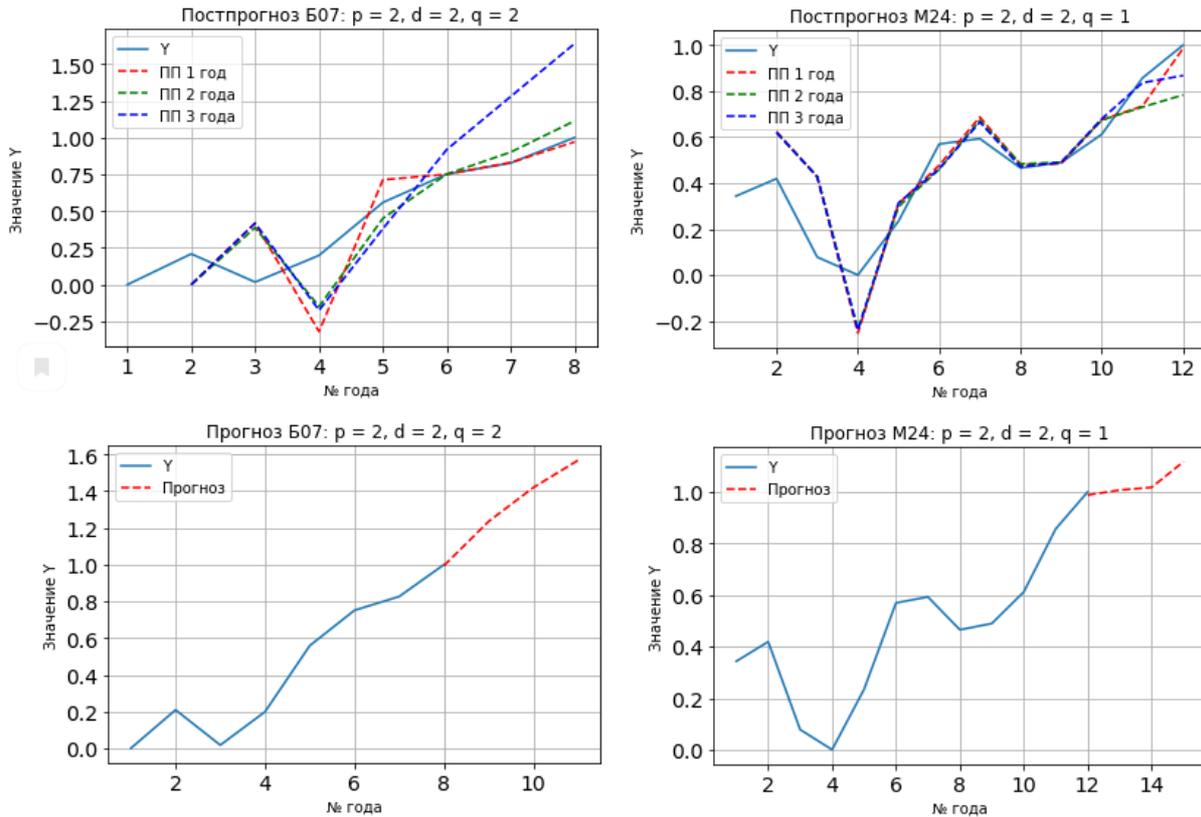


Рис. 6. Графики постпрогноза и прогноза наборов данных второй категории
Fig. 6. Graphs of post-forecast and forecast of data sets of the second category

В данном случае автор рекомендует находить лучшую ARIMA модель через перебор параметров p, d, q ($3 \cdot 2 \cdot 3$ итерации).

Третья категория, наборы данных, график которых сложно оценить и имеет слабый наглядный тренд (рис. 7).

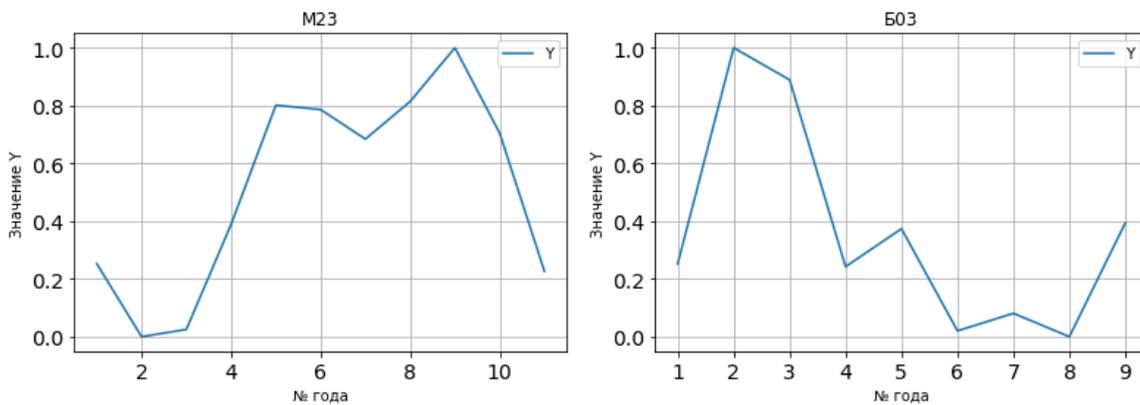


Рис. 7. Примеры графиков третьей категории
Fig. 7. Examples of graphs of the third category

Данная категория интересна тем, что значение d может принимать как малые, так и большие значения ($d \in [1, 5]$). В это же время коррелограммы будут показывать всего один лаг для p и q .

В данном случае автор рекомендует находить лучшую ARIMA модель через перебор параметров p, d, q с запасом ($2 \cdot 5 \cdot 2$ итерации). Качество постпрогноза и прогноза для этой категории оставляет желать лучшего. Автор предполагает, что в данном случае на ухудшении качества прогноза сказывается малая длина набора данных.

Информатика и вычислительная техника

В качестве примеров приведены графики наборов данных М23 и Б03. Для первого набора наилучшая *ARIMA*-модель найдена при использовании параметров $p = 1, d = 1, q = 1$, для второго: $p = 1, d = 2, q = 1$. Графики постпрогноза и прогноза представлены на рис. 8.

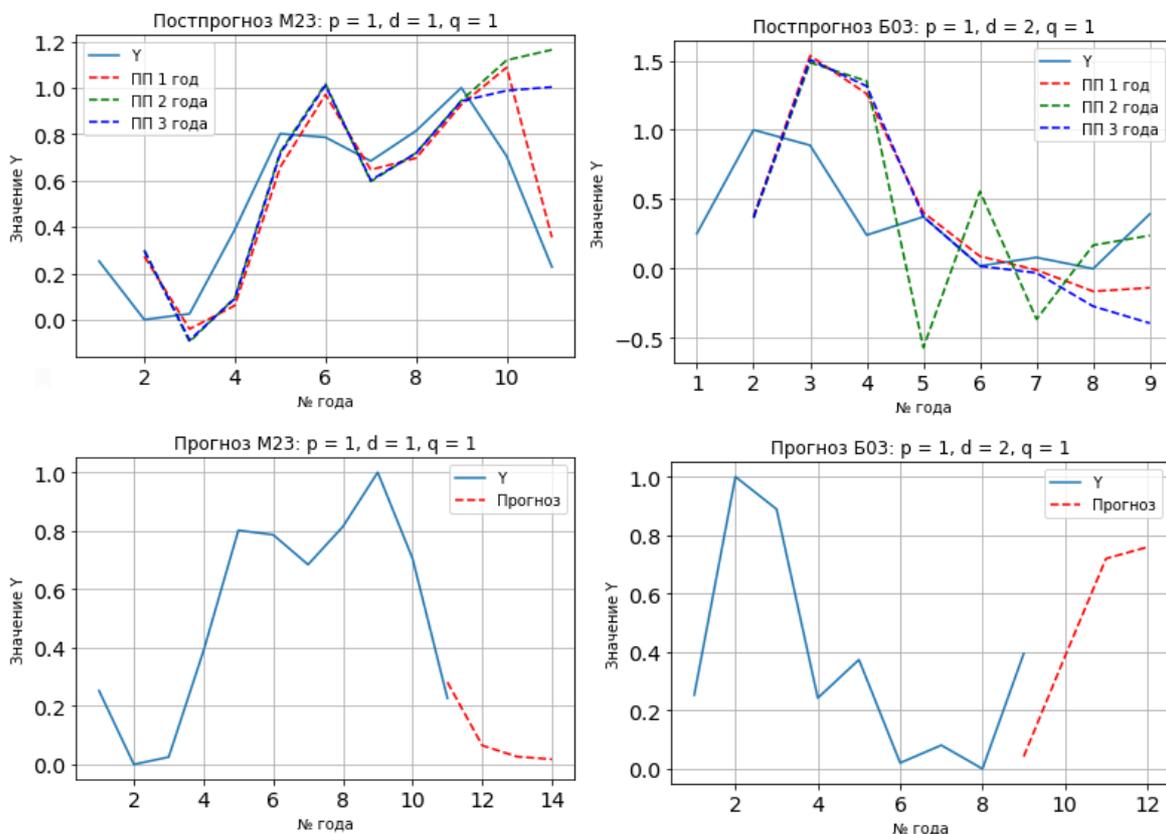


Рис. 8. Графики постпрогноза и прогноза наборов данных третьей категории
Fig. 8. Graphs of post-forecast and forecast of data sets of the third category

Четвертая категория, наборы данных, график которых сложно оценить, тренд отсутствует или сложно определить наглядно (рис. 9).

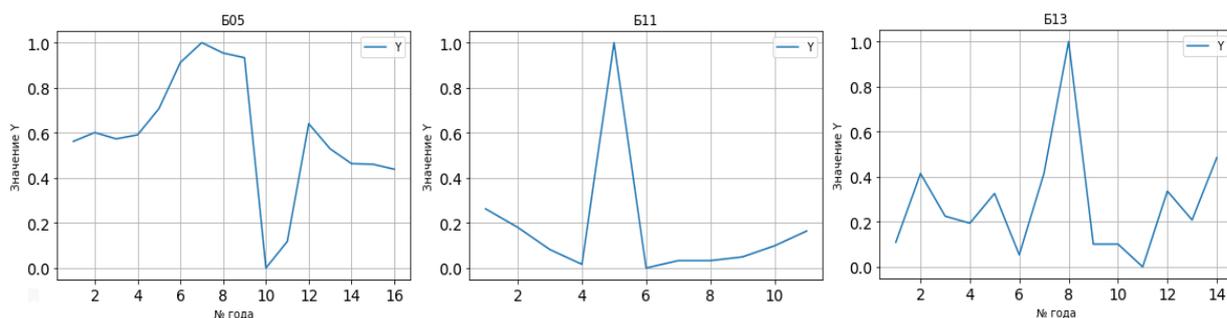


Рис. 9. Примеры графиков четвертой категории
Fig. 9. Examples of graphs of the fourth category

Данная категория интересна тем, что ряд может оказаться стационарным и модель *ARIMA* теряет свой смысл для наборов данных этой категории. Однако если взять малые значения ($d \in [1, 2]$) порядка интегрирования временного ряда, то модель может с нулевыми значениями p и q построить удовлетворительный прогноз. По некоторым опытам было обнаружено, что нулевые значения p и q дают более качественный прогноз. Автор рекомендует использовать малые значения параметров p, d, q .

В качестве примеров приведены графики наборов данных Б05, Б11, Б13. Для всех наборов наилучшая *ARIMA*-модель найдена при использовании параметров $p = 0, d = 2, q = 0$. Графики постпрогноза и прогноза представлены на рис. 10.

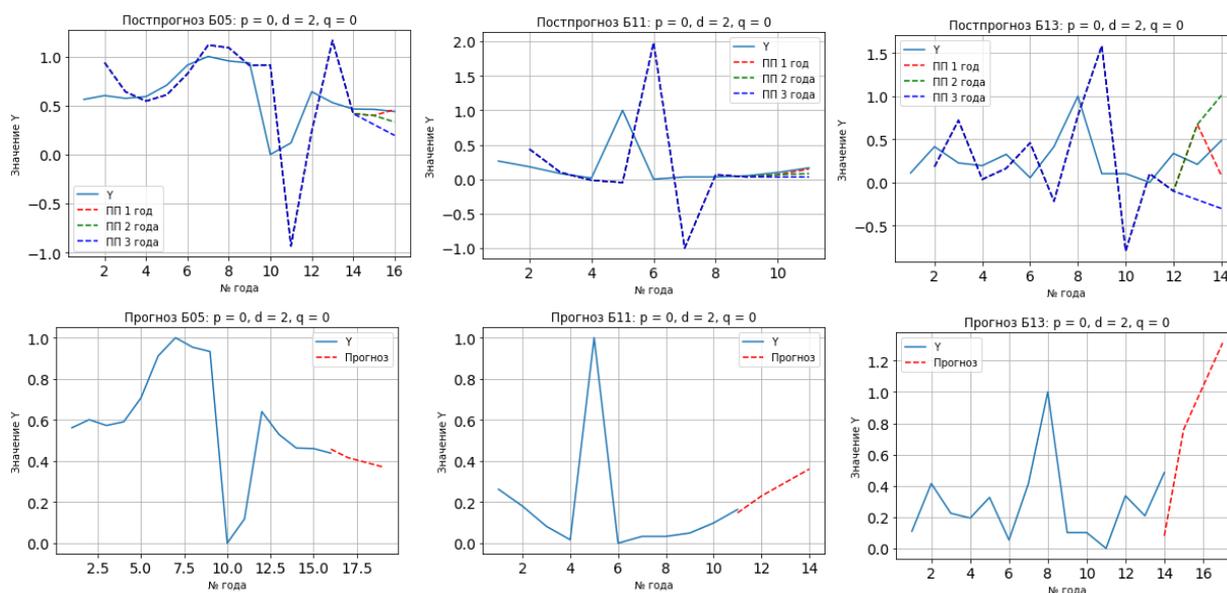


Рис. 10. Графики постпрогноза и прогноза наборов данных четвертой категории
Fig. 10. Graphs of post-forecast and forecast of data sets of the fourth category

Заключение

В результате исследования была достигнута цель работы – поиск классификационных признаков временных рядов, позволяющих понять – можно ли для их краткосрочного (3 отсчета) прогноза применять модель *ARIMA*, и с каким качеством.

По обработанным данным были получены четыре категории графиков и найдены диапазоны значений параметров для настройки модели *ARIMA* для каждой категории:

- 1) график похож на прямую и имеет четкий тренд: $d = 1, p \in [1, 3], q \in [1, 3]$, прогнозируются без каких-либо проблем;
- 2) график с четким трендом: $d \in [2, 3], p \in [1, 3], q \in [1, 3]$ прогнозируются без каких-либо проблем;
- 3) график сложно оценить и имеет слабый тренд: $d \in [1, 5], p \in [1, 2], q \in [1, 2]$, качество прогноза может быть плохим, автор предполагает, что в данном случае на ухудшении качества прогноза сказывается малая длина набора данных;
- 4) график сложно оценить, тренд отсутствует или сложно определить наглядно: $d \in [1, 2], p \in [0, 2], q \in [0, 2]$, может возникнуть ситуация, при которой ряд окажется стационарным, в применении модели *ARIMA* не будет смысла, однако если ей воспользоваться, то при нулевых AR и MA можно получить качественные прогнозы.

Литература

1. Бокс, Дж. Анализ временных рядов, прогноз и управление / Дж. Бокс, Г. Дженкинс; пер. с англ. под ред. В.Ф. Писаренко. – М.: Мир, 1974. – Кн. 1. – 406 с.
2. Семенов, А.В. Анализ временных рядов в политической науке: возможности и ограничения / А.В. Семенов // Политическая наука. – 2021. – № 1. – С. 76–97.
3. Khan, F.M. ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India / F.M. Khan, R. Gupta // Journal of Safety Science and Resilience. – 2020. – Vol. 1, iss. 1. – P. 12–18. DOI: 10.1016/j.jnlssr.2020.06.007
4. Abonazel, M.R. Forecasting Egyptian GDP Using ARIMA Models / M.R. Abonazel, A.I. Abd-Elfтах // Reports on Economics and Finance. – 2019. – Vol. 5, no. 1. – P. 35–47. DOI: 10.12988/ref.2019.81023

5. Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks / M.J. Kane, N. Price, M. Scotch et al. // *BMC Bioinformatics* 15. – 2014. – Vol. 276. DOI: 10.1186/1471-2105-15-276

6. Верзунов, С.Н. Краткосрочное прогнозирование индекса качества воздуха на основе ARIMA-моделей / С.Н. Верзунов, Н.М. Лыченко // *Сборник материалов VII Международной научной конференции, посвященной памяти С.С. Ефимова*. – Омск, 2020. – С. 76–78.

7. Раднаев, Б.Б. ARIMA-модель пульсового сигнала / Б.Б. Раднаев, А.С. Цыбиков, Б.В. Хабитиев // *Вестник БГУ. Математика, информатика*. – 2017. – № 1. – С. 78–85.

8. Янченко, Т.В. Метод управления развитием социального ресурса региона на основе регрессионно-дифференциального моделирования / Т.В. Янченко, А.В. Затонский // *Управление большими системами: сб. тр.* – 2015. – № 54. – С. 86–113.

9. Затонский, А.В. Преимущества дифференциальной модели сложной экономической системы / А.В. Затонский, Н.А. Сиротина // *Образование. Наука. Научные кадры*. – 2012. – № 8. – С. 98–102.

10. Дуброва, Т.А. *Статистические методы прогнозирования: учеб. пособие для вузов* / Т.А. Дуброва. – М.: ЮНИТИ-ДАНА, 2003. – 206 с.

11. Wang, W. Autoregressive Model-Based Gear Fault Diagnosis / W. Wang, A.K. Wong // *Journal of Vibration and Acoustics*. – 2002. – Vol. 124 (2). – P. 172–179. DOI: 10.1115/1.1456905

12. Gupta, R. The impact of US uncertainty on the Euro area in good and bad times: evidence from a quantile structural vector autoregressive model / R. Gupta, C.K.M. Lau, M.E. Wohar // *Empirica*. – 2018. – Vol. 46. – P. 353–368. DOI: 10.1007/s10663-018-9400-3

13. Multistage fusion approaches based on a generative model and multivariate exponentially weighted moving average for diagnosis of cardiovascular autonomic nerve dysfunction / M. Mehedi Hassan, Sh. Huda, J. Yearwood et al. // *Information Fusion*. – 2018. – Vol. 41. – P. 105–118. DOI: 10.1016/j.inffus.2017.08.004

14. Optimized Dickey-Fuller Test Refines Sign and Boundary Problems Compare to Traditional Dickey-Fuller Test / I. Masudul, A. Afroza, M. Sirajum et al. // *International Journal of Statistics and Probability*. – 2018. – Vol. 7, no. 5. – P. 19–27. DOI: 10.5539/ijsp.v7n5p19

15. Forecasting E-Commerce Products Prices by Combining an Autoregressive Integrated Moving Average (ARIMA) Model and Google Trends Data / S. Carta, A. Medda, A. Pili et al. // *Future Internet*. – 2019. – Vol. 11 (1), 5. DOI: 10.3390/fi11010005

16. Akaike, H. A new look at the statistical model identification / H. Akaike // *IEEE Transactions on Automatic Control*. – 1974. – Vol. 19. – P. 716–723.

17. Akaike's Information Criterion, Cp and estimators of loss for elliptically symmetric distributions / A. Boisbunon, S. Canu, D. Fourdrinier et al. // *International Statistical Review*. – 2014. – Vol. 82 (3). – P. 422–439. DOI: 10.1111/insr.12052

Фелькер Мария Николаевна, канд. техн. наук, доцент кафедры автоматизации технологических процессов, Пермский национальный исследовательский политехнический университет, Березниковский филиал, г. Березники; eymarya@yandex.ru.

Чеснов Владислав Вадимович, магистрант, Пермский национальный исследовательский политехнический университет, Березниковский филиал, г. Березники; tazaqsp@mail.ru.

Поступила в редакцию 14 июня 2021 г.

STUDY OF THE INFLUENCE OF CHANGING THE PARAMETERS OF THE ARIMA MODEL ON THE QUALITY OF THE FORECAST FOR SHORT DATA SETS

M.N. Fel'ker, axl_mail_box@mail.ru,
V.V. Chesnov, grisha@molniya-ufa.ru

Perm National Research Polytechnic University, Berezniki Branch, Berezniki, Russian Federation

Time series, i.e. data collected at various times. The data collection segments may differ depending on the task. Time series are used for decision making. Time series analysis allows you to get some result that will determine the format of the decision. Time series analysis was carried out in very ancient times, for example, various calendars became a consequence of the analysis. Later, time series analysis was applied to study and forecast economic, social and other systems. Time series appeared a long time ago. Once upon a time, ancient Babylonian astronomers, studying the position of the stars, discovered the frequency of eclipses, which allowed them to predict their appearance in the future. Later, the analysis of time series, in a similar way, led to the creation of various calendars, for example, harvest calendars. In the future, in addition to natural areas, social and economic ones were added. **Aim.** Search for classification patterns of time series, allowing to understand whether it is possible to apply the *ARIMA* model for their short-term (3 counts) forecast. **Materials and methods.** Special software with *ARIMA* implementation and all need services is made. We examined 59 data sets with a short length and step equal a year, less than 20 values in the paper. The data was processed using Python libraries: Statsmodels and Pandas. The Dickey–Fuller test was used to determine the stationarity of the series. The stationarity of the time series allows for better forecasting. The Akaike information criterion was used to select the best model. Recommendations for a reasonable selection of parameters for adjusting *ARIMA* models are obtained. The dependence of the settings on the category of annual data set is shown. **Conclusion.** After processing the data, four categories (patterns) of year data sets were identified. Depending on the category ranges of parameters were selected for tuning *ARIMA* models. The suggested ranges will allow to determine the starting parameters for exploring similar datasets. Recommendations for improving the quality of post-forecast and forecast using the *ARIMA* model by adjusting the settings are given.

Keywords: ARIMA, econometrics, analysis, time series.

References

1. Boks, Dzh. *Analiz vremennyh ryadov, prognoz i upravlenie* [Time series analysis, forecasting and management]. Moscow, Mir Publ., 1974. 406 p.
2. Semenov A.V. [Time series analysis in political science: opportunities and limitations]. *Political science*, 2021, no. 1, pp. 76–97. (in Russ.)
3. Khan F.M., Gupta R. ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. *Journal of Safety Science and Resilience*, 2020, vol. 1, iss. 1, pp. 12–18. DOI: 10.1016/j.jnlssr.2020.06.007
4. Abonazel M.R., Abd-Elftah A.I. Forecasting Egyptian GDP Using ARIMA Models. *Reports on Economics and Finance*, 2019, vol. 5, no. 1, pp. 35–47. DOI: 10.12988/ref.2019.81023
5. Kane M.J., Price N., Scotch M. Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics* 15, 2014, vol. 276. DOI: 10.1186/1471-2105-15-276
6. Verzunov S.N., Lychenko N.M. [Short-term forecasting of the air quality index based on ARIMA models]. *Collection of materials of the VII International Scientific Conference dedicated to the memory of S.S. Efimov*, Omsk, 2020, pp. 76–78. (in Russ.)
7. Radnaev B.B., Cybikov A.S., Habituev B.V. [ARIMA-pulse signal model]. *Bulletin of the BSU. Mathematics, computer science*, 2017, no. 1, pp. 78–85. (in Russ.)
8. Janchenko T.V., Zatonskij A.V. [The method of managing the development of the social resource

of the region on the basis of regression-differential modeling]. *Managing large systems: a collection of works*, 2015, no. 54, pp. 86–113. (in Russ.)

9. Zatonskij A.V., Sirotina N.A. [Advantages of the differential model of a complex economic system]. *Education. The science. Scientific personnel*, 2012, no. 8, pp. 98–102. (in Russ.)

10. Dubrova, T.A. *Statisticheskie metody prognozirovaniya* [Statistical methods of forecasting]. Moscow, YUNITI-DANA, 2003. 206 p.

11. Wang W., Wong A.K. Autoregressive Model-Based Gear Fault Diagnosis. *Journal of Vibration and Acoustics*, 2002, vol. 124 (2), pp. 172–179. DOI: 10.1115/1.1456905

12. Gupta R., Lau C.K.M., Wohar M.E. The impact of US uncertainty on the Euro area in good and bad times: evidence from a quantile structural vector autoregressive model. *Empirica*, 2018, vol. 46, pp. 353–368. DOI: 10.1007/s10663-018-9400-3

13. Mehedi Hassan M., Huda Sh., Yearwood J., Jelinek H. F., Almogren A. Multistage fusion approaches based on a generative model and multivariate exponentially weighted moving average for diagnosis of cardiovascular autonomic nerve dysfunction. *Information Fusion*, 2018, vol. 41, pp. 105–118. DOI: 10.1016/j.inffus.2017.08.004

14. Masudul I., Afroza A., Sirajum M., Salauddin Khan Md. & Monzur Murshed Md. Optimized Dickey-Fuller Test Refines Sign and Boundary Problems Compare to Traditional Dickey-Fuller Test. *International Journal of Statistics and Probability*, 2018, vol. 7, no. 5, pp. 19–27. DOI: 10.5539/ijsp.v7n5p19

15. Carta S., Medda A., Pili A., Reforgiato Recupero D., Saia R. Forecasting E-Commerce Products Prices by Combining an Autoregressive Integrated Moving Average (ARIMA) Model and Google Trends Data. *Future Internet*, 2019, vol. 11 (1), 5. DOI: 10.3390/fi11010005

16. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 1974, vol. 19, pp. 716–723.

17. Boisbunon A., Canu S., Fourdrinier D., Strawderman W., Wells M.T. Akaike's Information Criterion, Cp and estimators of loss for elliptically symmetric distributions. *International Statistical Review*, 2014, vol. 82 (3), pp. 422–439. DOI: 10.1111/insr.12052

Received 14 June 2021

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Фелькер, М.Н. Исследование влияния изменения параметров модели ARIMA на качество прогноза для коротких наборов данных / М.Н. Фелькер, В.В. Чеснов // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2021. – Т. 21, № 3. – С. 36–46. DOI: 10.14529/ctcr210304

FOR CITATION

Fel'ker M.N., Chesnov V.V. Study of the Influence of Changing the Parameters of the ARIMA Model on the Quality of the Forecast for Short Data Sets. *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics*, 2021, vol. 21, no. 3, pp. 36–46. (in Russ.) DOI: 10.14529/ctcr210304