

Научно-практический
рецензируемый журнал

СТАТИСТИКА И ЭКОНОМИКА
Том 19. № 2. 2022

Учредитель:
РЭУ им. Г.В. Плеханова

Главный редактор
Виталий Григорьевич Минашкин

Зам. главного редактора
Елена Алексеевна Егорова
Павел Александрович Смелов

Ответственный редактор
Никита Дмитриевич Эпштейн

Технический редактор
Елена Ивановна Аникеева

Журнал издается с 2004 года.
Свидетельство о регистрации СМИ:

ПИ № ФС77-65889

от 27.05.16 г.

ISSN 2500-3925 (Print)

Все права на материалы,
опубликованные
в номере, принадлежат журналу
«Статистика и экономика».
Перепечатка материалов,
опубликованных в журнале, без
разрешения редакции запрещена.
При цитировании материалов ссылка
на журнал «Статистика и экономика»
обязательна.

Мнение редакции может не совпадать
с мнением авторов

Журнал включен ВАКом в перечень
периодических научных изданий.

Тираж журнала
«Статистика и экономика»
1500 экз.

Адрес редакции:
117997, г. Москва,
Стремянный пер., 36, корп. 6, офис 345
Тел.: (499) 237-83-31, (доб. 18-04)
E-mail: Smelov.PA@rea.ru
Адрес сайта: www.statecon.rea.ru

Подписной индекс журнала
в каталоге «РОСПЕЧАТЬ»: 80246

© ФГБОУ ВО
«РЭУ им. Г. В. Плеханова», 2020

Подписано в печать 22.04.22.
Формат 60x84 1/8. Цифровая печать.
Печ. л. 7,5. Тираж 1500 экз.
Заказ

Напечатано в ФГБОУ ВО
«РЭУ им. Г.В. Плеханова».
117997, Москва, Стремянный пер., 36

СОДЕРЖАНИЕ

МЕТОДОЛОГИЯ СТАТИСТИКИ

- Т.В. Золотова, Д.А. Волкова*
Методы интеллектуальной обработки данных для
коррекции атипичных значений котировок акций 4

НАЦИОНАЛЬНЫЕ СЧЕТА И МАКРОЭКОНОМИЧЕСКАЯ СТАТИСТИКА

- Н.С. Айюбова*
Эконометрический анализ и моделирование динамики
развития платежного баланса в Азербайджане 14

ДЕМОГРАФИЧЕСКАЯ СТАТИСТИКА

- А.В. Алпатов*
Смертность в Волгоградской области на фоне пандемии
Covid-19 23

СОЦИАЛЬНАЯ СТАТИСТИКА

- Д.А. Гучмазова*
Мировые тенденции дифференциации доходов населения 36

- Ю.С. Пиньковецкая*
Оценка доли мужчин в численности педагогических
работников в школах 43

СТАТИСТИКА И МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ЭКОНОМИКЕ

- А.А. Микрюков, М.Е. Мазуров*
Комбинированная когнитивная модель прогнозирования
деятельности университета 52

Scientific and practical reviewed
journal

STATISTICS AND ECONOMICS
Vol. 19. № 2. 2022

Founder:
Plekhanov Russian University of
Economics

Editor in chief
Vitaliy G. Minashkin

Deputy editor
Elena A. Egorova
Pavel A. Smelov

Executive editor
Nikita D. Epshtein

Technical editor
Elena I. Anikeeva

Journal issues since 2004.
Mass media registration certificate:
ΦC77-65889 от 27.05.16.
ISSN 2500-3925 (Print)

All rights for materials published in the
issue belong to the journal
«Statistics and Economics».

Reprinting of articles published in the
journal, without the permission of the
publisher is prohibited.

When citing a reference to the journal
«Statistics and Economics» is obligatory.

Editorial opinion may be different from
the views of the authors

The journal is included in the list of VAK
periodic scientific publications.
Journal articles are reviewed.
The circulation of the journal
«Statistics and Economics» –
1,500 copies.

Editorial office:
117997, Moscow,
Streymanny lane. 36, Building 6, office 345
Tel.: (499) 237-83-31 (18-04)
E-mail: Smelov.PA@rea.ru
Web: www.statecon.rea.ru

Subscription index of journal
in catalogue «ROSPECHAT»: 80246

© Plekhanov Russian University of
Economics, 2020

Signed to print 22.04.22.
Format 60x84 1/8. Digital printing.
Printer's sheet 7.5. 1500 copies.
Order

Printed in Plekhanov Russian University
of Economics,
Streymanny lane. 36, Moscow, 117997,
Russia

CONTENTS

METHODOLOGY OF STATISTICS

- Tatiana V. Zolotova, Daria A. Volkova*
Intelligent Data Processing Methods for the Atypical Values
Correction of Stock Quotes 4

NATIONAL ACCOUNTS AND MACROECONOMIC STATISTICS

- Natavan S. Ayyubova*
Econometric analysis and modeling of the dynamics of the
balance of payments' development in Azerbaijan 14

DEMOGRAPHIC STATISTICS

- Alexey V. Alpatov*
Mortality in the Volgograd Region Against the COVID-19
Pandemic 23

SOCIAL STATISTICS

- Daniya A. Huchmazova*
Global Trends in Differentiation of Population Income..... 36
- Yulia S. Pinkovetskaya*
Assessment of the Proportion of Men in the Number of
Teaching Staff in Schools 43

STATISTICAL AND MATHEMATICAL METHODS IN ECONOMICS

- Andrey A. Mikryukov, Mikhail E. Mazurov*
Combined Cognitive Model for Forecasting University
Activities 52

Редакционная коллегия

АСТАШОВА Ирина Викторовна, д.ф.-м.н., профессор, профессор кафедры дифференциальных уравнений, Московский государственный университет имени М.В. Ломоносова, Москва, Россия

АРХИПОВА Марина Юрьевна, д.э.н., профессор, факультет экономических наук, Департамент статистики и анализа данных, Высшая школа экономики — национальный исследовательский университет, Москва, Россия

БАКУМЕНКО Людмила Петровна, д.э.н., профессор, заведующая кафедрой прикладной статистики и информатики, Марийский государственный университет, Йошкар-Ола, Россия

ВОЛКОВА Виолетта Николаевна, д.э.н., профессор, профессор кафедры системного анализа и управления, Санкт-Петербургский государственный политехнический университет, Санкт-Петербург, Россия

ГЕВОРКЯН Эдуард Аршавирович, д.ф.-м.н., профессор кафедры Высшей математики, Российский экономический университет имени Г.В. Плеханова, Москва, Россия

ГЛИНКИНА Светлана Павловна, д.э.н., профессор, заведующая кафедрой общей экономической теории Московской школы экономики, Московский государственный университет имени М.В. Ломоносова, Москва, Россия

ЕЛИСЕЕВА Ирина Ильинична, д.э.н., профессор, член-корреспондент РАН, Заслуженный деятель науки Российской Федерации, заведующая кафедрой статистики и эконометрики, Санкт-Петербургский государственный экономический университет, г. Санкт-Петербург, Россия

ЗАРОВА Елена Викторовна, д.э.н., профессор, начальник отдела обработки и анализа статистической информации, Департамент экономической политики и развития города Москвы, руководитель Центрально-Евразийского представительства Международного статистического института, Москва, Россия

КАРМАНОВ Михаил Владимирович, д.э.н., профессор, профессор кафедры отраслевой и бизнес-статистики, Российский экономический университет имени Г.В. Плеханова, Москва, Россия

КУЧМАЕВА Оксана Викторовна, д.э.н., профессор, профессор кафедры народонаселения экономического факультета Московского государственного университета им. М.В. Ломоносова, Москва, Россия.

КЮРКЧАН Александр Гаврилович, д.ф.-м.н., профессор, заведующий кафедрой теории вероятностей и прикладной математики, Московский технический университет связи и информатики, Москва, Россия

ЛАЙКАМ Константин Эмильевич, д.э.н., заместитель руководителя Федеральной службы государственной статистики Российской Федерации, Москва, Россия

ЛУЛА Павел, доктор наук, доцент, заведующий кафедрой вычислительных систем, Краковский экономический университет, Краков, Польша

МОТОРИН Руслан Миколайович, д.э.н., профессор кафедры статистики и эконометрии, Киевский национальный торгово-экономический университет, Киев, Украина

МХИТАРЯН Владимир Сергеевич, д.э.н., профессор, заведующий отделением статистики, анализа данных и демографии, заведующий кафедрой статистических методов, Высшая школа экономики — национальный исследовательский университет, Москва, Россия

САДОВНИКОВА Наталья Алексеевна, д.э.н., профессор, заведующая кафедрой статистики, Российский экономический университет имени Г.В. Плеханова, Москва, Россия

САЖИН Юрий Владимирович, д.э.н., профессор, заведующий кафедрой статистики, эконометрики и информационных технологий в управлении, Национальный исследовательский Мордовский государственный университет имени Н. П. Огарева, Саранск, Россия

УПАДХАЯ Шьям, руководитель статистического отдела ЮНИДО, Организация Объединённых Наций по промышленному развитию, Вена, Австрия

ШУВАЛОВА Елена Борисовна, д.э.н., профессор, начальник управления аттестации научных кадров, Российский экономический университет имени Г.В. Плеханова, Москва, Россия

Editorial Board

Irina V. ASTASHOVA, Dr. Sci. (Phys.-Math.), Professor, Professor of the Differential Equations Department, Lomonosov Moscow State University, Moscow, Russia

Marina Yu. ARKHIPOVA, Dr. Sci. (Economics), Professor, Faculty of Economic Sciences, Department of Statistics and Data Analysis, Higher School of Economics — National Research University, Moscow, Russia

Lyudmila P. BAKUMENKO, Dr. Sci. (Economics), Professor, Head of Applied Statistics and Informatics Department, Mari State University, Yoshkar-Ola, Russia

Violetta N. VOLKOVA, Dr. Sci. (Economics), Professor, Professor of System Analysis and Management Department, Saint Petersburg State Polytechnic University, Saint Petersburg, Russia

Eduard A. GEVORKYAN, Dr. Sci. (Phys.-Math.), Professor of the Department of Higher Mathematics, Plekhanov Russian University of Economics, Moscow, Russia

Svetlana P. GLINKINA, Dr. Sci. (Economics), Professor, Head of the General Economic Theory Department, Moscow School of Economics, Lomonosov Moscow State University, Moscow, Russia

Irina I. ELISEEVA, Dr. Sci. (Economics), Professor, Corresponding Member of Russian Academy of Sciences, Head of Statistics and Econometrics Department, Saint-Petersburg State University of Economics, Saint-Petersburg, Russia

Elena V. ZAROVA, Dr. Sci. (Economics), Professor, Head of the Department of Processing and Analysis of Statistical Information, Department of Economic Policy and Development of Moscow, Chair of ISI Central Eurasia Outreach Committee, Moscow, Russia

Mikhail V. KARMANOV, Dr. Sci. (Economics), Professor, Professor of the Department of Industrial and Business Statistics, Plekhanov Russian University of Economics, Moscow, Russia

Oksana V. KUCHMAEVA, Dr. Sci. (Economics), Professor, Professor of the Department of population, faculty of Economics, Moscow state University. M. V. Lomonosova, Moscow, Russia

Alexander G. KYURKCHAN, Dr. Sci. (Phys.-Math.), Professor, Head of the Theory of Probability and Applied Mathematics Department, Moscow Technical University of Communications and Informatics, Moscow, Russia

Konstantin E. LAYKAM, Dr. Sci. (Economics), Deputy Head, Federal State Statistics Service of the Russian Federation, Moscow, Russia

Pawel LULA, Dr. hab., Associate Professor, Head of the Department of Computational Systems, Cracow University of Economics, Cracow, Poland

Ruslan M. MOTORIN, Dr. Sci. (Economics), Professor of Statistics and Econometrics Department, Kiev National University of Trade and Economics, Kiev, Ukraine

Vladimir S. MKHITARYAN, Dr. Sci. (Economics), Professor, Head of the Department of Statistics, Data Analysis and Demography, Head of the Department of Statistical Methods, Higher School of Economics — National Research University, Moscow, Russia

Natalia A. SADOVNIKOVA, Dr. Sci. (Economics), Professor, Head of Statistics Department, Plekhanov Russian University of Economics, Moscow, Russia

Yury V. SAZHIN, Dr. Sci. (Economics), Professor, Head of the Department of Statistics, Econometrics and Information Technologies in Management, Ogarev Mordovia State University, Saransk, Russia

Shyam UPADHYAYA, Chief, UNIDO Statistics Unit, United Nations Industrial Development Organization, Vienna, Austria

Elena B. SHUVALOVA, Dr. Sci. (Economics), Professor, Head of the Department of Scientific Personnel Certification, Plekhanov Russian University of Economics, Moscow, Russia

Методы интеллектуальной обработки данных для коррекции атипичных значений котировок акций

Цель исследования. Цель исследования состоит в проведении сравнительного анализа различных методов коррекции атипичных значений статистических данных на фондовом рынке и выработке рекомендаций для их использования.

Материалы и методы. В статье проведен анализ российской и зарубежной библиографии по проблеме исследования. Предлагается рассмотрение методов машинного обучения для обнаружения и коррекции выбросов во временных рядах. Математическую основу методов машинного обучения составляют метод Z-score, метод изолирующего леса, метод опорных векторов для обнаружения выбросов и методы винсоризации и множественного вменения для коррекции выбросов. Для построения моделей использован программный инструмент Jupyter Notebook, поддерживающий язык программирования Python. Для реализации методов машинного обучения используются данные котировок акций Московской биржи.

Результаты. Продемонстрированы результаты работы алгоритмов машинного обучения для наборов реальных статистических данных, представляющих собой цены закрытия акций трех российских компаний «Сбербанк», «Аэрофлот», «Газпром» в период с 01.12.2019 по 30.11.2020, полученные с сайта с сайта Инвестиционной компании «ФИНАМ». Проведен сравнительный анализ методов обнаружения и коррекции выбросов по среднеквадратическому отклонению. Статистический метод Z-score позволяет точно определить расстояние от подозрительного наблюдения до центра распределения, что является преимуществом. Недостаток данного метода — это влияние выбросов на среднее значение и стандартное отклонение, которое может способствовать маскировке выбросов или некорректному их обнаружению. Метод изолирующего леса распознает выбросы различных видов, а при реализации метода отсутствуют параметры, требующие подбора; но недостатком является

более низкая скорость обнаружения выбросов по сравнению с другими методами. Метод опорных векторов считается очень быстрым методом и сводится к решению задачи квадратичного программирования, которая всегда имеет единственное решение. Метод винсоризации для коррекции выбросов снижает влияние выбросов на среднее значение и дисперсию, что является преимуществом, но может вносить систематическую ошибку из-за подбора пороговых значений для разделения наблюдений в выборке. Метод множественного вменения создает для каждого пропущенного значения не одно, а множество вменений, что позволяет избежать систематической ошибки, но за счет высоких вычислительных затрат. Для использованных в работе исходных данных лучший результат показала реализация алгоритма множественного вменения по обнаруженным выбросам методом опорных векторов.

Заключение. В теории анализа данных нет универсального метода обнаружения и/или устранения выбросов. В общем случае определение выбросов субъективно, и решение принимается индивидуально для каждого конкретного набора данных с учетом его особенностей или имеющегося опыта в данной области. Реализованные в работе методы обнаружения и устранения выбросов могут найти применение при вычислении более точных значений показателей в различных сферах деятельности, например, для построения более точного прогноза цены акции. В перспективе планируется исследование влияния параметров методов обнаружения и коррекции выбросов на результаты работы моделей, а также оптимизация этих параметров.

Ключевые слова: обнаружение выбросов, коррекция выбросов, язык программирования Python, среднеквадратическое отклонение.

Tatiana V. Zolotova¹, Daria A. Volkova²

¹Financial University under the Government of the Russian Federation, Moscow, Russia

²Peoples' Friendship University of Russia, Moscow, Russia

Intelligent Data Processing Methods for the Atypical Values Correction of Stock Quotes

Purpose of the study. The purpose of the study is to carry out a comparative analysis of various methods for correcting atypical values of statistical data on the stock market and to develop recommendations for their use.

Materials and methods. The article analyzes Russian and foreign bibliography on the research problem. Consideration of machine learning methods for detecting and correcting outliers in time series is proposed. The mathematical basis of machine learning methods is the Z-score method, the isolation forest method, support vector method for outlier detection, and winsorization and multiple imputation methods for outlier correction. To create the models, the Jupyter Notebook software tool, which supports the Python programming language, was used. To implement machine-learning methods, data from stock quotes of the Moscow Exchange are used.

Results. The results of machine learning algorithms are demonstrated for sets of real statistical data representing the closing prices of shares of three Russian companies "Sberbank", "Aeroflot", "Gazprom" in

the period from 01.12.2019 to 30.11.2020, obtained from the website of the Investment Company "FINAM". A comparative analysis of methods for detecting and correcting outliers by standard deviation has been carried out. The Z-score statistical method allows you to accurately determine the distance from the suspicious observation to the distribution center, which is an advantage. The disadvantage of this method is the influence of outliers on the mean and standard deviation, which can contribute to the masking of outliers or their incorrect detection. The isolation forest method recognizes outliers of various types, and when implementing the method, there are no parameters that require selection; but the disadvantage is the slower detection rate of outliers compared to other methods. The support vector machine is a very fast method and is reduced to solving a quadratic programming problem, which always has a unique solution. The winsorization method for correcting outliers reduces the effect of outliers on the mean and variance, which is an advantage, but may introduce bias due to the selection of thresholds to separate

observations in the sample. The multiple imputation method creates for each missing value not one, but many imputations, which avoids a systematic error, but at the expense of high computational costs. For the initial data used in the work, the best result was shown by the implementation of the multiple imputation algorithm based on the detected outliers by the support vector method.

Conclusion. There is no universal method for detecting and/or eliminating outliers in data analysis theory. In general, the determination of outliers is subjective, and the decision is made individually for each specific dataset, considering its characteristics

or existing experience in this area. The practical implementation of the methods for detecting and eliminating outliers used in this work can be a tool for calculating more accurate indicators in any area, for example, to improve forecasting the stock price. As part of further work, it is possible to consider the optimization of the parameters used in the methods of detecting and correcting outliers to study their effect on the results of the models.

Keywords: outlier detection, outlier correction, Python programming language, standard deviation.

Введение

В простейшем случае выброс представляет собой наблюдение, которое существенно отличается от остальных наблюдений набора данных. Основная проблема выявления выбросов состоит в определении того, действительно ли наблюдения, несовместимые с остальными данными, являются выбросами. Также является проблемой и то, что в теории анализа данных нет универсального метода обнаружения и/или устранения выбросов. В качестве примера можно привести работу [6], к которой представлены алгоритмы для анализа выбросов, включая сведения о том, когда и почему конкретные алгоритмы работают эффективно; обсуждаются последние идеи в этой области, такие как ансамбли выбросов, матричная факторизация, ядерные методы и нейронные сети.

Проблема обнаружения выбросов на основе статистических методов и методов машинного обучения затрагивалась во многих отечественных и зарубежных работах. Некоторые методы обнаружения выбросов можно найти, например, в работах [2, 9, 10, 14, 15, 17, 18]. Описанные в этих работах подходы к обнаружению выбросов основаны на статистических методах и методах машинного обучения и представляют собой методы обнаружения выбросов с помощью кластеризации с учителем и без учителя, на основе искусственных нейронных сетей, способом построения профилей, представляющих реальное

и идеальное поведения, с использованием методов машинного обучения с построением самосовершенствующейся системы на основе предыдущих знаний, гибридные методы.

После определения выбросов, следующим шагом является их устранение, т. е. либо коррекция точек данных до их правильных значений, либо удаление таких наблюдений из набора данных. Для сглаживания атипичных данных предлагаются различные эвристические методы, вплоть до исключения атипичных точек (см., например, [3, 5, 7, 12, 15, 19]): исправление выброса до правильного значения; удаление выброса из данных; подход, требующий продолжать признавать наличие выброса, но ничего не делать с его значением; метод анализа результатов с выбросами и без них; винсоризация, предполагающая преобразование крайних значений в заданный процентиль данных; усечение, при котором происходит установка наблюдаемых значений в пределах правдоподобного диапазона и исключение других значений из набора данных; преобразование, предполагающим применение детерминированной математической функции (например, логарифмической функции) к каждому значению, чтобы не только сохранить выброс данных в анализе, но также уменьшить дисперсию ошибок; модификация, т.е. изменение значения выброса вручную на другое, менее экстремальное значение; двухэтапная робастная процедура; прямой робастный метод с использованием итеративно

взвешенных наименьших квадратов и некоторые другие.

Как известно, проблема выбросов возникает в задачах построения экономических трендов. В регрессионном анализе работа над выбросами традиционно состоит из трех этапов: обнаружение выбросов, устранение выбросов, построение регрессионной зависимости. В работе [13] предложен комбинированный подход к преобразованию данных в регрессионном анализе, объединяющий (после обнаружения выбросов) процедуру коррекции выбросов и построение регрессионной зависимости. При этом коррекция выбросов строится на основе линейных преобразований исходных данных.

Обнаружение и коррекция выбросов, т.е. значений данных, которые не соответствуют общему их распределению, является важным вопросом в том числе во многих экономических областях, т. к. выбросы могут информировать о критически важных событиях. Так, например, процедура обнаружения частого использования кредитной карты в финансовых транзакциях для противодействия мошенничеству в банке относится к работе с выбросами в данных; об этом идет речь, например, в работах [1, 4]. При этом в большинстве случаев злоумышленники приспосабливаются таким образом, чтобы наблюдения, являющиеся выбросами, казались нормальными, что затрудняет задачу определения аномального поведения. Таким образом, актуальность исследований, связанных с ана-

лизом выбросов, не вызывает сомнения.

Возникает вопрос считать ли выбросами те данные, которые на первый взгляд существенно не отличаются от остальных наблюдений набора данных. Обсуждение этого вопроса в статье проводится с использованием методов машинного обучения. В общем случае определение выбросов субъективно, и решение принимается индивидуально для каждого конкретного набора данных с учетом его особенностей или имеющегося опыта в данной области.

Задача исследования состоит в применении методов машинного обучения для анализа котировок акций на финансовых рынках и нахождения наилучшего подхода к решению вопроса обнаружения и коррекции выбросов по критерию среднего квадратического отклонения. В данной работе для обнаружения и коррекции выбросов рассматриваются методы машинного обучения, математическую основу которых составляют метод Z-score, метод изолирующего леса, метод опорных векторов для обнаружения выбросов и методы винсоризации и множественного вменения для коррекции выбросов (см., например, [8, 11, 12, 14, 15]).

Статистический метод Z-score проверяет каждый предполагаемый выброс и позволяет точно определить расстояние от подозрительного наблюдения до центра распределения, что является преимуществом. Но метод Z-score имеет недостаток, который заключается в том, что среднее значение и стандартное отклонение, которые используются в вычислении, подвержены влиянию выбросов, что может способствовать маскировке выбросов или некорректному их обнаружению.

Метод изолирующего леса распознает выбросы различных видов: как изолирован-

ные точки с низкой локальной плотностью, так и кластеры выбросов малых размеров. Также при реализации метода отсутствуют параметры, требующие подбора. Но недостатком является скорость обнаружения выбросов, которая не является его сильной стороной по сравнению с другими методами.

Метод опорных векторов считается очень быстрым методом и сводится к решению задачи квадратичного программирования, которая всегда имеет единственное решение. Недостатком метода является отсутствие устойчивости к шуму.

Использование метода винсоризации для коррекции выбросов снижает влияние выбросов на среднее значение и дисперсию, которые очень чувствительны к наличию выбросов, что является преимуществом, но в то же время может вносить систематическую ошибку из-за подбора пороговых значений для разделения наблюдений в выборке.

Метод множественного вменения является надежным и информативным методом, т. к. для каждого пропущенного значения создается не одно, а множество вменений и позволяет избежать систематической ошибки, но за счет высоких вычислительных затрат. Когда имеется большой объем данных множественное вменение сложно реализовать.

Для построения моделей использован программный инструмент Jupyter Notebook, поддерживающий язык программирования Python. Вообще, машинное обучение не привязано к конкретному языку программирования, но существуют языки, традиционно применяющиеся в машинном обучении и имеющие развитую систему библиотек, фреймворков и других инструментов. Одним из распространенных языков для данных целей является Python, в экосистеме

которого присутствуют все необходимые библиотеки для использования практически всех существующих алгоритмов и методов машинного обучения и анализа данных. В данной статье результаты работы алгоритмов машинного обучения продемонстрированы для наборов данных, представляющих собой котировки акций трех российских компаний. Проведен сравнительный анализ полученных результатов.

Реализация методов обнаружения выбросов

Проведем сначала описание и предварительную обработку данных, а затем применим метод Z-score, метод изолирующего леса, метод опорных векторов для обнаружения выбросов во временных рядах. Исходными данными являются ежедневные цены закрытия акций компаний «Сбербанк», «Аэрофлот», «Газпром» в период с 01.12.2019 по 30.11.2020, полученные с сайта www.finam.ru [20]. Проиллюстрируем более подробно проведенные исследования для данных компании «Сбербанк». На рис. 1 представлен график изменения цены закрытия акций компаний «Сбербанк» за указанный период.

Для построения моделей будем применять программный инструмент Jupyter Notebook, поддерживающий язык программирования Python. Первый шаг предполагает загрузку необходимых библиотек: numpy, pandas, matplotlib.pyplot и из sklearn — preprocessing, IsolationForest, svm, а также исходных данных через функцию `pd.read_csv`.

Для визуализации исходных данных будем использовать функцию `scatterplot`, аргументы которой есть `x_data` — значения по оси x, `y_data` — значения по оси y, `x_label` — название оси x, `y_label` — название оси y, `title` — заголовок графика.

Для визуализации выбросов мы определим функцию



Рис. 1. График цен закрытия акций «Сбербанк»

Fig. 1. Schedule of closing prices for Sberbank shares



Рис. 2. График цен закрытия акций «Сбербанк» с добавленными выбросами

Fig. 2. Schedule of closing prices of Sberbank shares with added outliers

scatterplot_with_color_coding, часть аргументов которой аналогична аргументам функции scatterplot, и добавлен аргумент «color_code_column» для изображения нормальных данных и выбросов по-разному.

Рассмотрим статистический метод *Z-score* [14]. Как известно, стандартизованная оценка значения x рассчитывается по формуле

$$z = \frac{x - \bar{x}}{\sigma_x},$$

где $\bar{x}$ — математическое ожидание (или среднее значение), σ_x — среднее квадратическое отклонение (СКО), вычисленные для множества данных $\{x_i\}$.

Выбросами будут считаться те значения данных, которые имеют расстояние от $\bar{x}$ более заданного, обычно 2 или 3 СКО. Для реализации метода в Jupyter Notebook воспользуемся функцией *z_anomaly_detection*, аргументами которой являются: *df* — таблица с данными, *column_name* — название столбца, в котором отыскиваются выбросы, *number_of_stdevs_away_from_mean* — количество СКО от среднего значения. Возьмем $z = 2$. Тогда все значения данных, отличающиеся от среднего значения более чем на 2 СКО, помечаются как выбросы.

Проверим работу метода, добавив во временной ряд пять аномальных значений (рис. 2), ожидая, что эти точки будут определены алгоритмом как выбросы. Для этого создадим словарь *anomaly_dictionary*, в котором ключом является индекс, а значением — цена.

Проверка алгоритма обнаружения с использованием функции *z_anomaly_detection* показала, что метод *Z-score* обнаружил четыре из пяти искусственно созданных выбросов. Обнаруженные выбросы в исходных данных представлены на графике (рис. 3), а также в табл. 1.

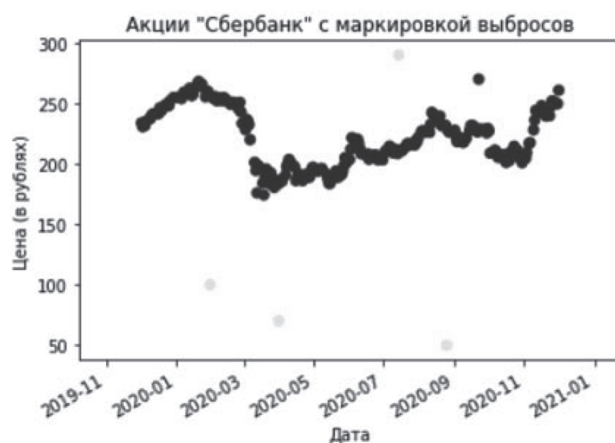


Рис. 3. Графики исходных данных с обнаруженными методом *Z-score* добавленными выбросами (слева) и выбросами в исходных данных (справа)

Fig. 3. Graphs of the original data with added outliers detected by the *Z-score* method (left) and outliers in the original data (right)

Выведем таблицу с обнаруженными выбросами в исходных данных (таблица 1).

Таблица 1 (Table 1)

Обнаруженные выбросы в исходных данных методом Z-score
Detected outliers in the original data by the Z-score method

	<DATE>	<CLOSE>	<CLOSE>_Z_Anomaly
32	2020-01-21	268.06	True
71	2020-03-18	174.27	True

Применим теперь метод *изолирующего леса* [11], идея которого основывается на случайном разбиении пространства признаков, таком что в среднем изолированные точки отсекаются от нормальных, кластеризованных данных. Окончательный результат усредняется по нескольким запускам стохастического алгоритма. Алгоритм изолирующего дерева (Isolation Tree) заключается в построении случайного бинарного решающего дерева. Корнем дерева является всё пространство признаков; в очередном узле выбирается случайный признак и случайный порог разбиения, сэмплированный из равномерного распределения на отрезке от минимального до максимального значения выбранного признака. Критерием останова является тожде-

ственное совпадение всех объектов в узле, то есть решающее дерево строится полностью. Ответом в листе, который также соответствует anomaly_score алгоритма, объявляется глубина листа в построенном дереве.

Для реализации данного метода используем модель IsolationForest из пакета scikit-learn, которую протестируем на нашем наборе данных. Для этого будем использовать функцию `isolation_forest_anomaly_detection`, аргументами которой являются `df` – таблица с данными, `column_name` – название столбца, в котором ищутся выбросы, `outliers_fraction` – доля допустимых выбросов.

Как и раньше проверим работу метода. Добавим к исходным данным пять аномальных значений, ожидая, что эти значения будут определены алгоритмом изолирующего леса как выбросы. Для этого создадим словарь `anomaly_dictionary`, в котором ключом является индекс, а значением – цена. Применяя функцию `isolation_forest_anomaly_detection` на исходных данных с добавленными значениями, получаем следующий результат: алгоритм изолирующего леса обнаруживает все искусственно созданные выбросы, но при этом метод классифицирует как выбросы не только

добавленные аномальные значения. Возможно, что данное явление относится к ложным срабатываниям, а, возможно, точки, не относящиеся к пяти добавленным значениям, действительно являются выбросами. Визуализируем исходные данные с обнаруженными выбросами (рис. 4).

Выведем таблицу с обнаруженными выбросами (табл. 2).

Таблица 2 (Table 2)

Обнаруженные выбросы в исходных данных методом изолирующего леса
Detected outliers in the original data by the isolating forest method

	<DATE>	<CLOSE>	<CLOSE>_Isolation_Forest
26	2020-01-13	262.40	True
30	2020-01-17	262.50	True
31	2020-01-20	266.28	True
32	2020-01-21	268.06	True
33	2020-01-22	266.54	True
34	2020-01-23	263.73	True
35	2020-01-24	265.49	True
38	2020-01-29	259.94	True
67	2020-03-12	175.91	True
71	2020-03-18	174.27	True
78	2020-03-27	180.38	True
79	2020-03-30	183.00	True
249	2020-12-01	260.81	True

Обнаружим выбросы с использованием одноклассного алгоритма *метода опорных векторов* (SVM). В основе метода SVM лежит идея перевода исходных векторов в пространство большей размерности и

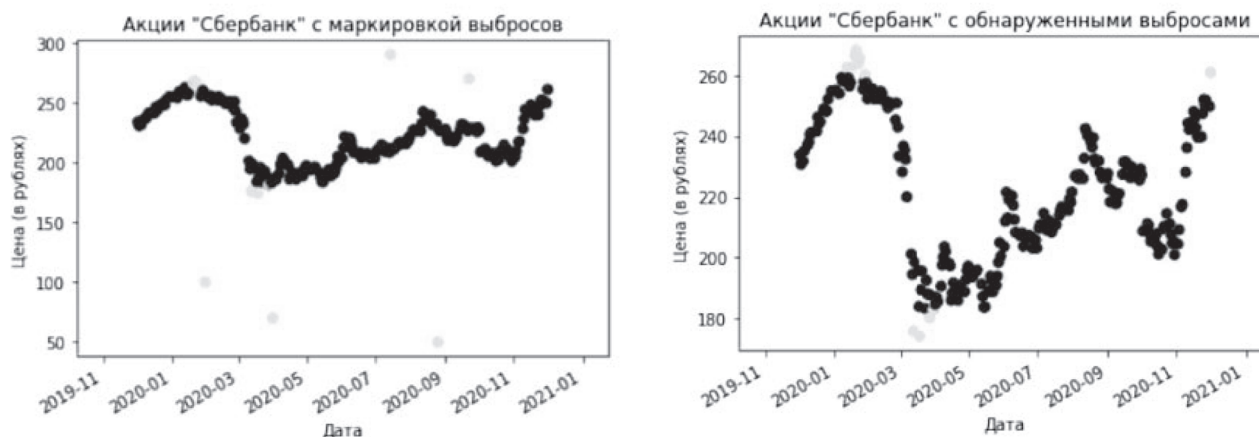


Рис. 4. Графики исходных данных с обнаруженными методом изолирующего леса добавленными выбросами (слева) и выбросами в исходных данных (справа)

Fig. 4. Graphs of the original data with added outliers detected by the isolation forest method (left) and outliers in the original data (right)

построения разделяющей гиперплоскости. После этого переходят к построению двух гиперплоскостей, параллельных разделяющей гиперплоскости так, чтобы опорные векторы (относящиеся к разным классам, ближайšie к разделяющей гиперплоскости) принадлежали им. При таком построении гиперплоскостей образуется полоса, свободная от данных точек. Получение оптимальной разделяющей гиперплоскости, т.е. такой разделяющей гиперплоскости, для которой ширина полосы максимальна, сводится к решению задачи квадратичного программирования [8].

Для реализации метода SVM используем модель OneClassSVM(), из пакета scikit-learn. Построим модель и протестируем ее на нашем наборе данных, используя функцию one_class_SVM_anomaly_detection, аргументами которой являются dataframe – таблица с данными, columns_to_filter – название столбца, в котором ищутся выбросы, outliers_fraction – доля допустимых выбросов.

Применяя функцию one_class_SVM_anomaly_detection, получаем график исходных данных с обнаруженными добавленными выбросами (рис. 5). Видно, что метод

опорных векторов обнаруживает не все искусственно созданные выбросы, однако также, как и изолирующий лес, он классифицирует и ложные срабатывания, которые возможно являются выбросами на самом деле. Для определения выбросов в исходных данных получим график исходных данных с обнаруженными выбросами.

Выведем таблицу с обнаруженными выбросами (табл. 3).

Таблица 3 (Table 3)

Обнаруженные выбросы в исходных данных методом опорных векторов
Detected outliers in the original data by the support vector machines

	<DATE>	<CLOSE>	One_Class_SVM_Anomaly
14	2019-12-20	244.71	True
15	2019-12-23	248.80	True
16	2019-12-24	248.67	True
20	2019-12-30	254.75	True
21	2020-01-03	255.00	True
31	2020-01-20	266.28	True
35	2020-01-24	265.49	True
67	2020-03-12	175.91	True
71	2020-03-18	174.27	True
98	2020-04-24	188.91	True
99	2020-04-27	188.90	True
211	2020-10-07	210.60	True
249	2020-12-01	260.81	True

Проводились аналогичные расчеты с использованием ежедневных цен закрытия акций компаний «Аэрофлот» и «Газпром». Для данных этих

компаний метод Z-score обнаруживает все искусственно созданные выбросы; методы изолирующего леса и опорных векторов обнаруживают все искусственно созданные выбросы, но классифицируют и ложные срабатывания.

Предварительно удалив выбросы, проведем сравнение методов обнаружения выбросов в исходных данных по среднеквадратическому отклоне-

нию: $\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$, где x_i – i -й элемент множества данных, $\bar{x}$ – среднее значение, n – количество элементов множества данных.

Таблица 4 (Table 4)

Значения СКО для методов обнаружения выбросов
Standard deviation values for outlier detection methods

	Аэрофлот	Газпром	Сбербанк
Z-score	15,06	24,65	22,87
Isolation Forest	15,78	26,8	21,56
SVM	15,35	26,55	22,29

Согласно таблице 4 получаем, что для данных компаний «Аэрофлот» и «Газпром» метод Z-score приводит к меньшим значениям СКО и, тем самым лучше выявляет выбросы, а для данных компании «Сбербанк» лучшим оказался метод изолирующего леса. Таким

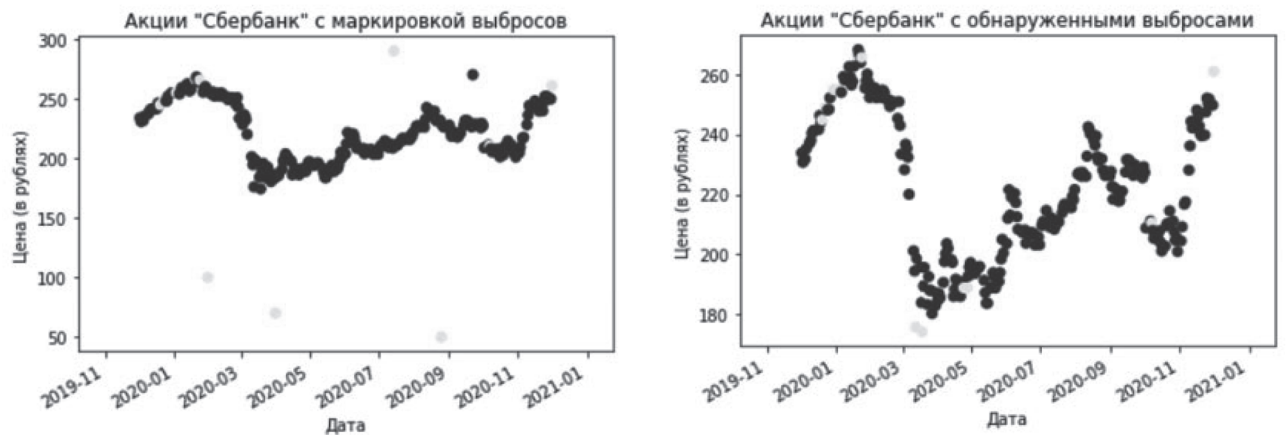


Рис. 5. Графики исходных данных с обнаруженными методом опорных векторов добавленными выбросами (слева) и выбросами в исходных данных (справа)
Fig. 5. Graphs of original data with added outliers detected by the support vector machines (left) and outliers in the original data (right)

образом, на данных выборках ни один из методов не является наилучшим, т. к. каждый метод может давать неточные результаты обнаружения выбросов. Поэтому, чтобы точно говорить являются ли определенные экземпляры данных выбросами, целесообразно на каждом наборе данных проверять несколько методов.

Методы винсоризации и множественного вменения для коррекции выбросов

После обнаружения выбросов в данных перейдем к следующему этапу работы с выбросами — устранению.

Основной проблемой устранения выбросов является то, что в большинстве методов применяется удаление обнаруженных выбросов, что на самом деле не всегда является верным. Например, удаление выбросов может привести к важным изменениям в результатах эксперимента. Поэтому далее будут рассмотрены методы устранения выбросов, которые не удаляют обнаруженные выбросы, а корректируют их значения. Основными такими методами являются винсоризация (winsorizing) и множественное вменение (multiple imputation).

Винсоризация — это процедура, которая уменьшает влияние выбросов на среднее значение и дисперсию [12]. Винсоризация включает определение пороговых значений. Выборочные наблюдения, значения которых лежат за пределами определенных заранее установленных значений, преобразуются, чтобы приблизить их к значению отсечения. Исходные данные, лежащие за пределами пороговых значений, преобразуются, чтобы они больше не считались выбросами. Однако следует отметить, что измененные значения являются искусственными и иногда могут быть неприемлемыми. Процесс винсоризации заключается в следующем: любое значение переменной

выше или ниже квантиля k на каждой стороне распределения переменных заменяется значением самого k -го квантиля. Например, если $k = 5$, то все наблюдения выше 95-го квантиля изменяются на значение этого 95-го квантиля, а значения ниже 5-го квантиля аналогично изменяются на соответствующее значение. По сравнению с простым удалением выбросов коррекцию с помощью винсоризации можно считать менее экстремальным вариантом, поскольку она «перекодирует» выбросы, а не удаляет их полностью. Часто берут $k = 5$ и, следовательно, оно используется как значение по умолчанию. В результате винсоризации мы получаем модифицированную выборку, в которой выбросы изменены нормальными значениями. Дальнейшие расчеты проводятся для модифицированного образца.

Для реализации метода винсоризации в Jupyter Notebook будем использовать функцию `winsorize` из библиотеки `scipy`, аргументами которой являются набор данных и `limits` — кортеж процентных значений, которые нужно вырезать на каждой стороне массива.

На рис. 6 представлены результаты применения метода для акций компании «Сбербанк».

Рассмотрим теперь *метод множественного вменения*. Метод вменения часто используется при обработке пропущенных значений; но он также может быть применен и при работе с выбросами. При этом выбросы удаляются (становятся как бы пропущенными значениями), а затем заменяются оценками, основанными на оставшихся данных.

Методы вменения можно разделить на две подгруппы: однократное вменение и множественное вменение. Большинство методов однократного вменения все выбросы или пропущенные значения заменяются средним значением, не учитывая неопределенность вмененных значений. Это означает, что методов однократного вменения распознают вмененные значения как фактические, не принимая во внимание стандартную ошибку, что приводит к смещению результатов. Поэтому лучшей альтернативой и более надежным методом вменения является множественное вменение. Множественное вменение — это метод вменения, основанный на статистике, когда для каждого пропущенного значения создается не одно, а множество вменений. Это означает многократное заполнение пропущенных значений, создание

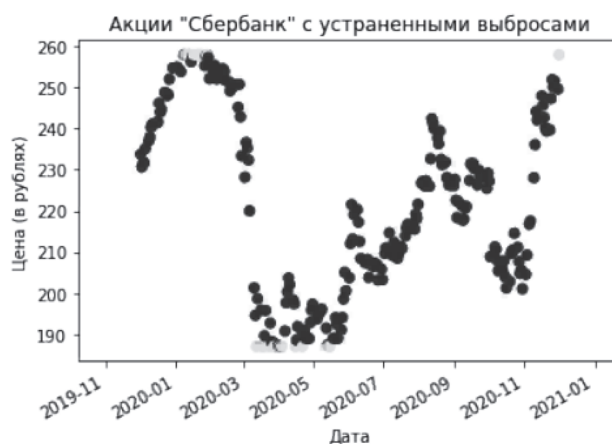


Рис. 6. График исходных данных с устраненными выбросами методом винсоризации

Fig. 6. Graph of the initial data with eliminated outliers by the winsorization method

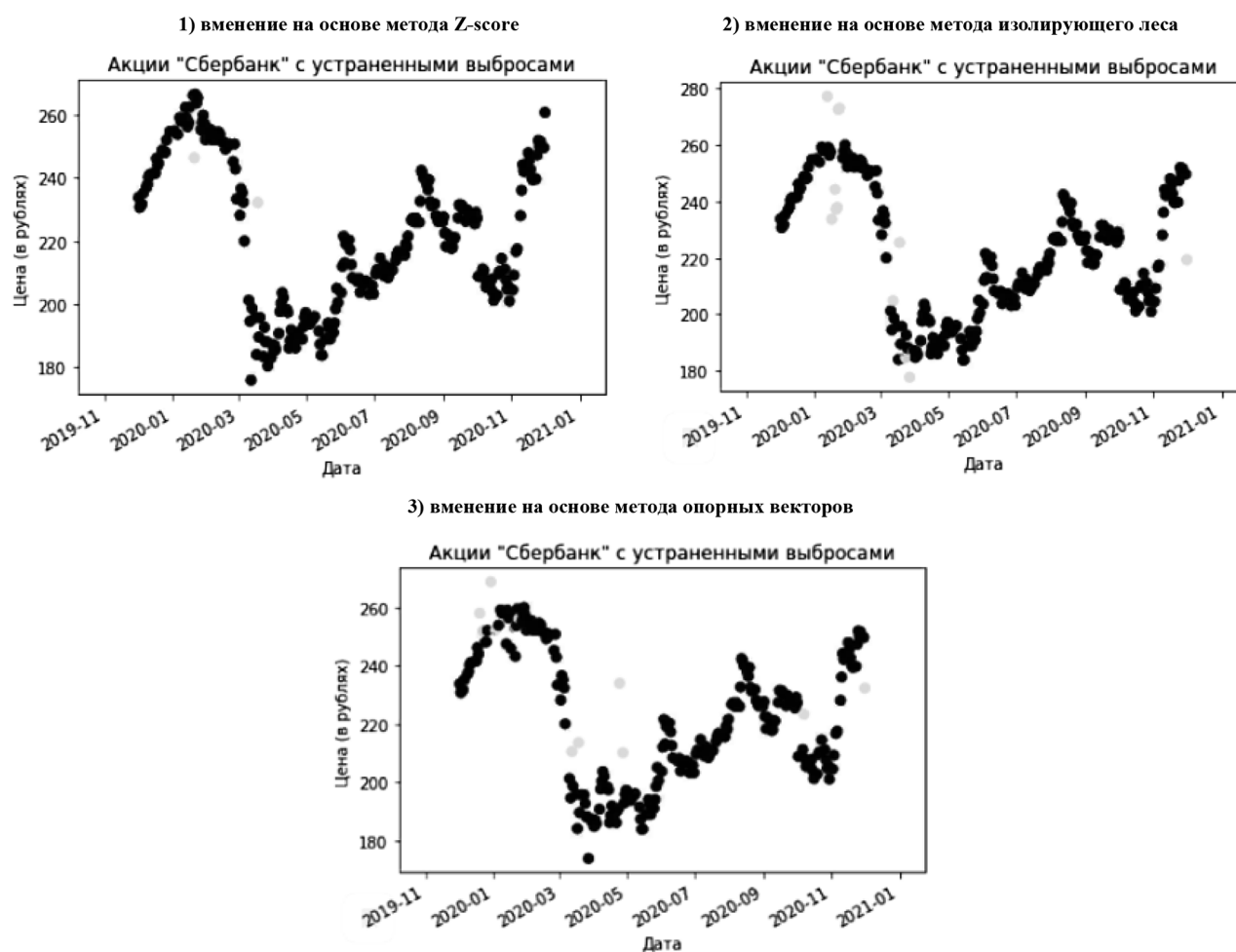


Рис. 7. Графики исходных данных с устраненными выбросами методом вменения на основе трех методов обнаружения выбросов

Fig. 7. Graphs of initial data with the eliminated outliers by imputation method based on three outlier detection methods

нескольких «полных» наборов данных [15]. Для множественного вменения был разработан ряд алгоритмов. Одним из хорошо известных алгоритмов является множественное вменение по цепному уравнению (Multiple Imputation by Chained Equations (MICE)) [15].

На практике множественное вменение в Python реализуется не так просто, однако в библиотеке `sklearn` есть функция `IterativeImputer`, которая может быть использована для множественного вменения. Главный ее аргумент, который мы будем использовать это `max_iter` — максимальное количество раундов вменения, которые необходимо выполнить перед возвратом вменений. Обычно выполняется 10 вменений и по умолчанию функция ис-

пользует `max_iter=10`, поэтому будем использовать это значение. И раз мы применяем множественное вменение, то аргумент `sample_posterior` надо изменить на `True`, который по умолчанию принимает значение `False`.

Обнаруженные ранее в исходных данных выбросы нужно удалить, сделав их пропущенными значениями. Так как мы применяли три метода обнаружения выбросов, то и вменение мы будем проводить для каждого набора выбросов соответствующих компаний.

На рис. 7 показаны результаты множественного вменения с обнаруженными различными методами выбросами для акций компании «Сбербанк».

Методы винсоризации и множественного вменения

были применены и для данных акций компаний «Аэрофлот» и «Газпром».

Результаты сравнения методов коррекции выбросов по среднеквадратическому отклонению (СКО) для акций рассматриваемых компаний представлены в таблице

Таблица 5 (Table 5)

Значения СКО для методов коррекции выбросов

Standard deviation values for outliers' correction methods

	Аэро-флот	Газпром	Сбер-банк
Винсоризация	16,234	27,371	22,456
Z-score и вменение	14,903	25,9	22,883
Isolation Forest и вменение	14,809	25,493	22,369
SVM и вменение	14,772	25,364	22,072

Как можно увидеть лучший результат дает совместное применение метода опорных векторов (SVM), который обнаруживает выбросы, и метода множественного вменения, который их устраняет.

Заключение

В работе рассмотрены методы обнаружения и коррекции выбросов, показаны их

возможности и реализация на реальных данных.

Для использованных в работе исходных данных лучший результат показала реализация алгоритма множественного вменения по обнаруженным методом опорных векторов выбросам.

Практическая реализация методов обнаружения и устранения выбросов, предложенная в данной работе, может

являться инструментом вычисления более точных показателей в любой сфере, например, для улучшения прогнозирования цены акции.

В рамках дальнейшей работы можно рассмотреть оптимизацию параметров, используемых в методах обнаружения и коррекции выбросов, для исследования их влияния на результаты работы моделей.

Литература

1. Ардан С. Д. Прогнозирование банкротства методами машинного обучения // Информационное общество. 2021. № 1. С. 56–67.
2. Десянин И.С. Предварительная обработка данных для машинного обучения // Фундаментальные и прикладные исследования в физике, химии, математике и информатике. 2021. С. 117–121.
3. Копырин А.С., Видищева Е.В. Оценка влияния аномалий на результаты анализа массивов экономических данных // Modern Economy Success. 2021. № 2. С. 235–240.
4. Чернова В.В. Применение методов машинного обучения для выявления аномалий с банковскими картами // МНСК-2021. 2021. С. 107–107.
5. Шибзуков З.М. О принципе минимизации эмпирического риска на основе усредняющих агрегирующих функций // Доклады Академии наук. 2017. Т. 476. № 5. С. 495–499.
6. Aggarwal C.C. Outlier analysis. Springer Science & Business Media. 2013. 455 с.
7. Aguinis H., Gottfredson R.K., Joo H. Best-practice recommendations for defining, identifying, and handling outliers // Organizational Research Methods. 2013. Т. 16. № 2. С. 270–301.
8. Chandola V., Banerjee A., Kumar V. Anomaly detection: A survey // ACM Computing Surveys. 2009. Т. 41. № 3. С. 15–58.
9. Cousineau D., Chartier S. Outliers detection and treatment: A review // International Journal of Psychological Research. 2010. Т. 23. № 1. С. 59–68.
10. Cunningham P., Cord M., Delan, S. J. Supervised Learning // Machine Learning Techniques for Multimedia. Springer Berlin Heidelberg. 2008. С. 21–49.

References

1. Ardan S.D. Bankruptcy forecasting using machine learning. Informationsnoye obshchestvo = Information society. 2021; 1: 56–67. (In Russ.)
2. Devyanin I.S. Preliminary data processing for machine learning. Fundamental'nyye i prikladnyye

11. Fei T.L., Kai M.T., Zhi-Hua Zhou. Isolation Forest // 2008 Eighth IEEE International Conference on Data Mining. 2008.
12. Frey B.B. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. 2018.
13. Gorelik V., Zolotova T. Method of Parametric Correction in Data Transformation and Approximation Problems // Lecture Notes in Computer Science (LNCS). Springer. Т. 12422. С. 122–133.
14. Hodge V., Austin J. A survey of outlier detection methodologies // Artificial Intelligence Review. 2004. Т. 22. № 2. С. 85–126.
15. Khan S.I., Hoque A.S. M.L. SICE: an improved missing data imputation technique. J Big Data. 2020. № 7. Article number 37.
16. Omar S., Ngadi A., Jebur H. Machine Learning Techniques for Anomaly Detection: An Overview // International Journal of Computer Applications. 2013. Т. 79. № 2. С. 33–41.
17. Patcha A., Park J. M. An overview of anomaly detection techniques: Existing solutions and latest technological trends // Computer Networks. 2007. Т. 51. № 12. С. 3448–3470.
18. Rousseeuw PJ, Hubert M. Robust statistics for outlier detection. Wiley Interdiscip Rev Data Min Knowl Discov. 2011. № 1(1). С. 73–9.
19. Zimek A., Filzmoser P. There and back again: Outlier detection between statistical reasoning and data mining algorithms // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2018. Т. 8. № 6. С. 73–79.
20. Инвестиционная компания «ФИНАМ» [Электрон. ресурс]. Режим доступа: <https://www.finam.ru/>.

issledovaniya v fizike, khimii, matematike i informatike = Fundamental and applied research in physics, chemistry, mathematics and informatics. 2021: 117–121. (In Russ.)

3. Kopyrin A.S., Vidishcheva Ye.V. Evaluation of the impact of anomalies on the results of the analysis of economic data arrays. Modern Economy

Success = Modern Economy Success. 2021; 2: 235–240. (In Russ.)

4. Chernova V.V. Application of machine learning methods to detect anomalies with bank cards. MNSK-2021 = MNSK-2021. 2021: 107–107. (In Russ.)

5. Shibzukov Z.M. On the principle of empirical risk minimization based on averaging aggregating functions. Doklady Akademii nauk = Reports of the Academy of Sciences. 2017; 476; 5: 495–499. (In Russ.)

6. Aggarwal C.C. Outlier analysis. Springer Science & Business Media. 2013. 455 p.

7. Aguinis H., Gottfredson R. K., Joo H. Best-practice recommendations for defining, identifying, and handling outliers. Organizational Research Methods. 2013; 16; 2: 270–301.

8. Chandola V., Banerjee A., Kumar V. Anomaly detection: A survey. ACM Computing Surveys. 2009; 41; 3: 15–58.

9. Cousineau D., Chartier S. Outliers detection and treatment: A review. International Journal of Psychological Research. 2010; 23; 1: 59–68.

10. Cunningham P., Cord M., Delan, S. J. Supervised Learning. Machine Learning Techniques for Multimedia. Springer Berlin Heidelberg. 2008: 21–49.

11. Fei T.L., Kai M. T., Zhi-Hua Zhou. Isolation Forest. 2008 Eighth IEEE International Conference on Data Mining. 2008.

12. Frey B.B. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. 2018.

13. Gorelik V., Zolotova T. Method of Parametric Correction in Data Transformation and Approximation Problems. Lecture Notes in Computer Science (LNCS). Springer; 12422: 122–133.

14. Hodge V., Austin J. A survey of outlier detection methodologies. Artificial Intelligence Review. 2004; 22; 2: 85–126.

15. Khan S.I., Hoque A.S. M.L. SICE: an improved missing data imputation technique. J Big Data. 2020; 7. Article number 37.

16. Omar S., Ngadi A., Jebur H. Machine Learning Techniques for Anomaly Detection: An Overview. International Journal of Computer Applications. 2013; 79; 2: 33–41.

17. Patcha A., Park J. M. An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer Networks. 2007; 51; 12: 3448–3470.

18. Rousseeuw PJ, Hubert M. Robust statistics for outlier detection. Wiley Interdiscip Rev Data Min Knowl Discov. 2011; 1(1): 73–9.

19. Zimek A., Filzmoser P. There and back again: Outlier detection between statistical reasoning and data mining algorithms. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2018; 8; 6: 73–79.

20. Investitsionnaya kompaniya «FINAM» = Investment company «FINAM» [Internet]. Available from: <https://www.finam.ru/>. (In Russ.)

Сведения об авторах

Татьяна Валерьяновна Золотова

Д.ф.-м.н., профессор

Финансовый университет

при Правительстве РФ, Москва, Россия

Эл. почта: tzolotova@fa.ru

Дарья Александровна Волкова

Российский университет дружбы народов,

Москва, Россия

Эл. почта: 174557@edu.fa.ru

Information about the authors

Tatiana V. Zolotova

Dr. Sci. (Physics and Mathematics), Professor

Financial University under the Government of the

Russian Federation, Moscow, Russia

E-mail: tzolotova@fa.ru

Daria A. Volkova

Peoples' Friendship University of Russia,

Moscow, Russia

E-mail: 174557@edu.fa.ru

Эконометрический анализ и моделирование динамики развития платежного баланса в Азербайджане

Цель исследования. Исследование посвящено эконометрическому анализу и моделированию динамики развития платежного баланса Азербайджана, формированию математико-статистического тренда, способной дать перспективную оценку развития платежного баланса. В соответствии с целью были поставлены задачи выбора наилучшего состава объясняющих факторов для модели, с помощью характеристик и критериев корреляционного и регрессионного анализа, эконометрических тестов расчет оценок характера и тесноты связи между объясняющими факторами, зависимым и независимыми факторами, проверки стационарности ряда.

Материалы и методы. Использованы официальные статистические данные Государственного Комитета Статистики и Центрального Банка Азербайджана, научные труды и исследования ученых, специалистов, как азербайджанских, так и зарубежных, в областях экономики и математико-экономического моделирования. Для эмпирического анализа нестационарных временных рядов в работе применены статистические методы обработки информации, для проверки адекватности и тестирования многомерной модели использованы соответствующие критерии и современные эконометрические процедуры с учетом воздействия экзогенных факторов. Для расчетов использованы пакеты прикладных программ, таких как Excel и Eviews 8.

Результаты. Создана многомерная регрессионная модель, позволяющая проводить экономико-статистический анализ динамики счета текущих операций платежного баланса; определены форма и направления функциональной зависимости между зависимыми и независимыми переменными, оценена изменчивость переменных, проанализированы результаты многомерного регрессионного анализа по эконометрическим методикам; измерены и интерпретированы количественные характеристики механизмов влияния объясняющих факторов на платежный баланс; в модели исследованы корреляционные зависимости для причинно-следственных зависимостей, выполнен тест Грейнджера и выявлены факторы, достоверно

объясняющие исход с высокими вероятностями на основе критерия Фишера; стационарность модели измерялась на основе теста Дики-Фуллера. При разностях первой и второй степени стационарность модели авторегрессии определялась на основе критерия Стьюдента путем изменения величины лага.

В процессе моделирования изначально построенная модель, охватывающая 1995–2017-е годы с 5-ю факторами как, иностранные инвестиции, экспорт, импорт, курс маната, общие инвестиции, показала недостаточную адекватность, то есть не стационарность ряда текущего счета платежного баланса. Курс национальной валюты, который участвует в модели как объясняющий фактор, подверг значения зависимого ряда большим колебаниям, росту дисперсии в остатках, что создала не стационарность и которое можно объяснить деноминацией национальной валюты в 2006 году. В последующем шаге был исследован период охватывающий 2006–2017-е годы. Также в процессе исследования в модель были добавлены независимые факторы, как дефицит государственного бюджета и валютные резервы. В результате была построена многофакторная эконометрическая модель.

Заключение. Построенная авторегрессионная модель достаточно адекватна, демонстрирует стационарность для временного ряда зависимой переменной и может считаться пригодной для прогнозных значений текущего счета платежного баланса.

Для выработки конкретных рекомендаций перспективного развития платежного баланса, полученные результаты исследования, обоснованные проведенным анализом динамики развития платежного баланса, дают возможность выявить реальные тенденции платежного баланса Азербайджана по текущему счету и определить его взаимозависимость с другими макроэкономическими переменными.

Ключевые слова: платежный баланс, текущий счет, эконометрическая модель, описательная статистика, вариабельность, стационарность, тест Дики Фуллера, тест Грейнджера.

Natavan S. Ayyubova

Baku State University, Baku, Azerbaijan

Econometric analysis and modeling of the dynamics of the balance of payments' development in Azerbaijan

Purpose of the study. The study is devoted to econometric analysis and modeling of the dynamics of the balance of payments' development of Azerbaijan, the formation of a mathematical and statistical trend that can give a perspective assessment of the development of the balance of payments. In accordance with the goal, the tasks of choosing the best composition of explanatory factors for the model were set, using the characteristics and criteria of correlation and regression analysis, econometric tests, calculating estimates of the nature and closeness of the relationship between the explanatory factors, dependent and independent factors, testing the stationarity of the series.

Materials and methods. The official statistical data of the State Statistics Committee and the Central Bank of Azerbaijan, scientific

works and studies of scientists, specialists, both Azerbaijani and foreign, in the fields of economics, mathematical and economic modeling were used. For the empirical analysis of non-stationary time series, statistical methods of information processing are used in the work; to check the adequacy and test the multivariate model, the appropriate criteria and modern econometric procedures are used, taking into account the impact of exogenous factors. For calculations, application packages such as Excel and Eviews 8 were used.

Results. A multivariate regression model has been created that makes it possible to conduct an economic and statistical analysis of the dynamics of the current account of the balance of payments; the form and directions of the functional relationship between dependent

and independent variables were determined, variability of variables was estimated, the results of multivariate regression analysis using econometric methods were analyzed; the quantitative characteristics of the mechanisms of influence of explanatory factors on the balance of payments were measured and interpreted; correlation dependencies for causal dependencies were investigated in the model, the Granger test was performed and factors were identified that reliably explain the outcome with high probabilities based on the Fisher criterion; the stationarity of the model was measured based on the Dickey-Fuller test. With differences of the first and second degree, the stationarity of the autoregressive model was determined based on the Student's criterion by changing the lag value.

In the process of modeling, the initially constructed model, covering the years 1995–2017 with five factors such as foreign investment, exports, imports, manat exchange rate, general investments, showed insufficient adequacy, that is, non-stationarity of the current account series of the balance of payments. The exchange rate of the national currency, which is involved in the model as an explanatory factor, subjected the values of the dependent series to large fluctuations, an increase in the variance in the residue, which created non-stationarity

and which can be explained by the denomination of the national currency in 2006. In the next step, the period covering 2006–2017 was examined. In addition, in the process of research, independent factors were added to the model, such as state budget deficit and foreign exchange reserves. As a result, a multifactorial econometric model was created.

Conclusion. The constructed autoregressive model is quite adequate, demonstrates stationarity for the time series of the dependent variable and can be considered suitable for predictive values of the current account of the balance of payments. To develop specific recommendations for the long-term development of the balance of payments, the results of the study, substantiated by the analysis of the dynamics of the development of the balance of payments, make it possible to identify real trends in the balance of payments of Azerbaijan on the current account and determine its interdependence with other macroeconomic variables.

Keywords: balance of payments, current account, econometric model, descriptive statistics, variability, stationarity, Dickey Fuller test, Granger test.

Introduction

The dynamic integration of Azerbaijan into the world economic system, accompanied by the activation of transboundary flows of capital, the characteristics of economic development of the country's trading partners, changes in prices on world markets and other external factors has an absolute effect to the processes in the sphere of economic activity of the country.

In order to conduct analytical and scenario-forecasting researches and make management decisions, it is important to take into account the interdependencies in the domestic economy, as well as to identify the factors that shape the country's foreign relations in a comprehensive and systematic manner[1,2,3,4,5,6]. It should be noted that modern modeling "Tools" of macroeconomic analysis and forecasting are less focused on a comprehensive consideration of the characteristics of foreign economic activity and the traditional system of macroeconomic indicators. The researches conducted by experts in this area focus on the key indicators of trade balance, the mechanisms of influence of exports and imports on the development of the national economy, the identification of threats that create high dependence of production on world markets conjuncture[7,8]. The issues such

as the systematic study of other aspects of foreign economic activity, as well as the establishment of econometric models, are rare.

The research on the regulation of the balance of payments in the late 1990s and early 2000s was a part of a general analysis of the transit economy. However, the recent researches have focused on the impact of national exchange rates on the balance of payments in the context of already established market relations[9,10]. By evaluating the impact of exchange rate fluctuations on the components of the balance of payments through the researches conducted, a model was developed, the theoretical and methodological bases of exchange rates were analyzed[11], methods for estimating devaluation expectations of financial market entities were developed taking into account the dynamics of the currency structure of investment assets, an economic mathematical models has been established to determine the dependence of exchange rates on inflation processes.

The dedication of the research to the topical issue is determined by the processes in the field of foreign economic activity, which are characterized by significant and growing impact on the development of the national economy and with the importance of taking them into account in assessing the country's economic development trends in the future.

The need for modeling of processes related to foreign economic activity, assessment of socio-economic results, development of adequate management decisions to prevent risks of the economic system forms a demand for methodological and modeling tools[12,13] of forecasting researches and expansion of analytical capabilities. Based on them, it is important to create the necessary opportunities to study and fully take into account the dependences between macroeconomic indicators of development and parameters of foreign economic activity. Along with the trade balance, the parameters of foreign economic activity include indicators such as the services balance of the balance of payments, transactions with financial instruments, and the capital account.

Although the general linear trend of the balance of payments of Azerbaijan for 2012–2018 is characterized by a decrease, the increase in all items abovementioned of the balance of payments from 2017 can be seen in fig.1.

The increase in revenues from the oil and gas sector has in the past again begun to ensure an increase in the total balance of payments surplus in Azerbaijan. In the second half of 2018, the trade balance surplus compared to the appropriate period in 2017 reached to \$ 3.665 mlrd. with an increase of \$ 1.462 billion, thus a

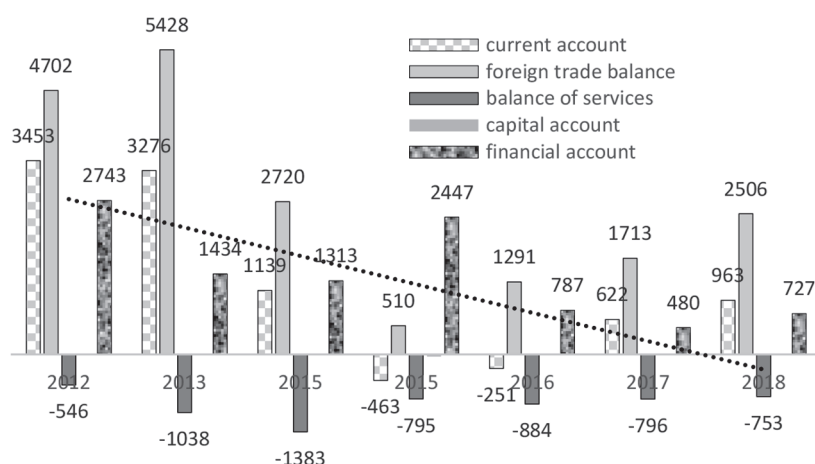


Рис. 1. Динамика ключевых показателей платежного баланса Азербайджана в 2012–2018 годы (млн долл. США)

Рис. 1. Динамика основных показателей платежного баланса Азербайджана в 2012–2018 гг. (млн долл. США)

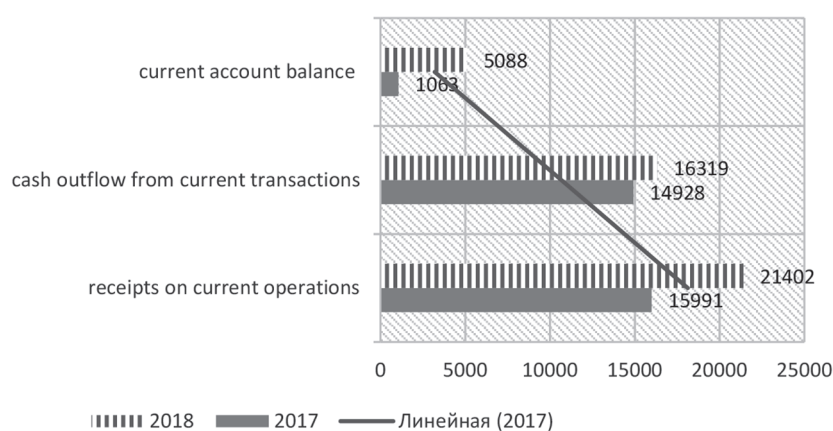


Рис. 2. Приток и отток денежных средств в страну в 2017 и 2018 гг. по текущим операциям (млн долл. США)

Fig. 2. Cash inflows and outflows to the country in 2017 and 2018 on current transactions (million US dollars)

Таблица 1 (Table 1)

Динамика изменения торгового баланса в 2016–2019 гг. по сравнению с предыдущим соответствующим периодом (тыс. долл. США и %)

Dynamics of trade balance changes in 2016–2019 compared to the previous relevant period (thousand US dollars and in%)

	2016		2017		2018		2019	
	Total	%-with	Total	%-with	Total	%-with	Total	%-with
Export	13210511	84,8	15152059	114,7	20793769	137,2	19868261	95,5
Import	9004176	92,1	9037316	100,4	10952441	121,2	11335316	103,5
Trade balance	4206335	-17,9	6114743	31,2	9841328	37,8	8532945	-15,3
Commercial turnover	22214687	-	16055795	-27,7	31746210	49,4	31203577	-1,73

Источник: Подготовлено автором на основе информации, полученной от Центрального банка Азербайджана (ЦБА)

Примечание: Знак «+» в табл. 1 указывает на увеличение, а знак «-» на уменьшение.

Source: Prepared by the author on the basis of information obtained from the Central Bank of Azerbaijan (CBA)

Note: The “+” sign in Table 1 indicates an increase, and the “-” sign indicates a decrease.

66% increase was achieved. During the period under review, revenues to the country reached to \$ 21.407 billion, and outflows to \$ 16.319 billion (see fig. 2). All statistics used in the analysis and graphs were taken from the official website of the Central Bank of Azerbaijan and State Statistical Committee of the Republic of Azerbaijan[14,15].

In 2018, compared to 2017, while the country's cash inflows from current operations increased in the amount of \$ 5.416 billion or 33.9%, the increase in outflows from the country was \$ 1.391 billion or 9.3%. As a result, cash inflows into the country are significantly higher than the outflows, the current account surplus has increased by almost 5 times and risen from \$ 1.063 billion to \$ 5.088 billion. According to experts, this factor played a decisive role in ensuring the stability of the country's national currency during the period under review.

Till the end of 2018, the trade turnover formed as a sum of import and export indicators reached a high level and increased to 31746 billion US dollars with an increase of 9.4%(see tab.1).

The main results of the study

In our initial research on modeling the dynamics of the balance of payments [16, 17], a regression analysis was conducted in order to conduct econometric analysis of the dependence of the current account of the balance of payments on total and foreign investments, exports and imports, the exchange rate of the Azerbaijani manat. Current account of y-balance of payments involved in the analysis, x_1 -foreign investment ($FORINV$), x_2 -export (EXP), x_3 -import (IMP), x_4 -exchange rate of manat to US dollar ($MANAT$), x_5 -total investment ($CENINV$) are dependent and explanatory variables covering the years of 1995–2017.

In the study, it is important to pay attention to the Darbin Watson coefficients with increased

and very low results, expressed by a very high multidimensional coefficient of determination in time series. Thus, relying on a high coefficient of determination and ignoring the low Durbin Watson coefficient, which points to autocorrelation, can lead to "false" regression and express incorrect dependencies[18].

According to the results of regression analysis with the above parameters (see tab. 2), the number of observations was 23; R^2 (determination coefficient) = 0.98; F -statistic (Fisher criterion) = 245.1; $prob.$ = 0.00; DW (Durbin Watson statistics) = 2.92. The results obtained are quite satisfactory. The coefficient of determination indicates that the independent variables included in the model explain the dependent variable by 98%. Criterion F received a fairly reliable estimate with a high probability. However, the result obtained for the DW criterion can't be considered satisfactory. The critical limits for the DW criterion with $n = 23$ and $k = 5$ are $D_L = 0.90$ and $D_U = 1.92$. Since the calculated value of the DW criterion is greater than 2, the value $4 - DW = 1.08$ is compared with critical values: $D_L < 1.08 < D_U$. Alternatively, we get an analogous result: $4 - D_U < DW < 4 - D_L$; because of being $2.08 < 2.92 < 3.1$, DW falls into the zone of uncertainty, and it is impossible to decide whether there is an autocorrelation.

The units of measurement of the independent variables x_1, x_2, x_3, x_5 included in the model for regression analysis are expressed in US dollars. Taking into account the denomination of the Azerbaijani manat in 2006, the research period was shortened due to the fact that the exchange rate of the x_4 manat, one of the explanatory factors was expressed in the national currency, which created a problem for the stability of the time series, the number of observations was reduced to 12 and covered the years of 2006-2017.

Таблица 2 (Table 2)

**Результаты регрессионного анализа за 1995–2017 гг.
(с пятью объясняющими факторами)**

Results of regression analysis for 1995-2017 (with five explanatory factors)

	FORINV	EXP	IMP	MANAT	CENINV	C
Coefficient	-1.041597	0.855962	0.040745	0.127631	-0.074593	-978.5116
Std. Error	0.214275	0.071390	0.054360	0.220403	0.144352	1126.953
t-Statistic	-4.861035	11.98994	0.749535	0.579079	-0.516739	-0.868281
Prob	0.0001	0.0000	0.4638	0.5701	0.6120	0.3973
R-squared		0.986318	Mean dependent var		4332.196	
Adjusted R-squared		0.982293	S.D. dependent var		6943.397	
S.E. of regression		923.9354	Akaike info criterion		16.71462	
Sum squared resid		14512163	Schwarz criterion		17.01084	
Log likelihood		-186.2181	Hannan-Quinn criter.		16.78912	
F-statistic		245.0926	Durbin-Watson stat		2.928609	
Prob(F-statistic)		0.000000				

Таблица 3 (Table 3)

**Результаты регрессионного анализа за 2006–2017 гг.
(с пятью объясняющими факторами)**

Results of regression analysis for 2006-2017 (with five explanatory factors)

	FORINV	EXP	IMP	MANAT	CENINV	C
Coefficient	-0.043786	0.940172	0.075568	-3669.255	-0.505909	-365.0545
Std. Error	0.851259	0.130751	0.097305	3126.497	0.394494	3458.028
t-Statistic	-0.051436	7.190547	0.776614	-1.173600	-1.282425	-0.105567
Prob	0.9606	0.0004	0.4669	0.2850	0.2470	
R-squared		0.977251	Mean dependent var		9106.875	
Adjusted R-squared		0.958293	S.D. dependent var		6616.478	
S.E. of regression		1351.229	Akaike info criterion		17.56227	
Sum squared resid		10954912	Schwarz criterion		17.80472	
Log likelihood		-99.37361	Hannan-Quinn criter.		17.47250	
F-statistic		51.54956	Durbin-Watson stat		2.922200	
Prob(F-statistic)		0.000075				

Number of observations according to the results (see Table 3) of the 2nd regression analysis 12; R^2 (determination coefficient) = 0.97; F -statistic (Fisher criterion) = 51.5; $prob.$ = 0.000075; (Durbin Watson statistics) = 2.92. According to the results, no significant change is observed, so the quality of the model does not increase or decrease significantly, and DW statistics still fall into the zone of uncertainty. This doesn't tell us whether there is an autocorrelation over time. In such cases, steps such as extending the time series and editing the explanatory factors in the model can be used to improve the quality of the model.

In general, the current account deficit formed under the influence of the trade balance can be financed by capital inflows in the following forms: Foreign borrow-

ings from other countries, the International Monetary Fund, the World Bank; Assets sold to foreign investors; Direct investments that provide foreign exchange inflows into the country for the establishment of new production facilities; Foreign exchange reserves.

The application of these measures supports the reduction of foreign assets of the state. However, if the government increases its foreign debt, which significantly exceeds the current account deficit, then the country is in danger of a foreign debt crisis with the balance of payments. Proper regulation of these financial processes is very important for the balance of payments and the dynamic development of the country's economy in general.

Thus, in order to improve the quality of the model, we have ad-

justed it and included in the model an important macroeconomic financial indicator of the country, the factor of state budget deficit. The indicators we received from the Central Bank of Azerbaijan for the budget deficit cover the years of 2006-2017 and the unit of measurement is million manat: -85.5; 78.6; -79.6; -12.2; -178; -363.5; 306; -135; 352.8; -308.4; -286.5; -241.2. This indicator is included in the model as x_6 -budget deficit (*BD*). The changes in the regression model are shown in Table 4 below.

In the next step, the model includes a macroeconomic indicator of international foreign exchange reserves. Statistical data for official international foreign exchange reserves for 2006-2017 (in USD million) were obtained from the Central Bank and included in the model as follows: x_7 -foreign exchange reserves(*REZ*) 1967.3; 4015.3; 6137; 5161.7; 6407.6; 10481.5; 11694.8; 14152; 13758.3; 5016.7; 3974.4; 5334.6.

The developmental dynamics of the explanatory factors in the model are described in a complex way in fig.3.

The model is formed as follows:

$$Y = -949.6 + 0,017FORINV + 0,99EXP + 0,12IMP - 3185,8MANAT - 0,55CENINV + 2,64BD - 0,12REZ$$

According to the results of the last regression analysis, the number of observations was 12; $R^2 = 0.98$; F statistic = 30.95; $prob.$ = 0.0025; DW = 2.66. The critical limits for the DW criterion with $n = 12$ and $k = 7$ are $D_L = 0.103$ and $D_U = 3.033$. As the price of DW statistics is closer to 2 than the previous prices, the quality of the model is improving.

In order to build a successful model, it is important that the independent variables in the model, the regressors, have a sufficiently wide range of variations. The range of variation can be measured on the basis of variability

Таблица 4 (Table 4)
Результаты регрессионного анализа за 2006–2017 гг.
(с шестью объясняющими факторами)
Results of regression analysis for 2006-2017(with six explanatory factors)

	FORINV	EXP	IMP	MANAT	CENINV	C
Coefficient	0.185987	0.995924	0.122918	-3744.041	-0.684332	-405.5565
Std. Error	0.862494	0.138169	0.104898	3074.143	0.420372	3399.490
t-Statistic	0.215638	7.208013	1.171791	-1.217914	-1.627923	-0.119299
Prob	0.8378	0.0008	0.2941	0.2776	0.1645	0.9097
R-squared		0.981681	Mean dependent var		9106.875	
Adjusted R-squared		0.959698	S.D. dependent var		6616.478	
S.E. of regression		1328.277	Akaike info criterion		17.51235	
Sum squared resid		8821597.	Schwarz criterion		17.79521	
Log likelihood		-98.07410	Hannan-Quinn criter.		17.40762	
F-statistic		44.65689	Durbin-Watson stat		2.715850	
Prob(F-statistic)		0.000348				

Таблица 5 (Table 5)
Результаты регрессионного анализа за 2006–2017 гг.
(с семью объясняющими факторами)
Results of regression analysis for 2006-2017 (with seven explanatory factors)

	FORINV	EXP	IMP	MANAT	CENINV	C
Coefficient	0.017248	0.997093	0.122752	-3185.879	-0.554208	2.640400
Std. Error	1.263255	0.153776	0.116669	4368.352	0.787641	2.618642
t-Statistic	0.013653	6.484073	1.052139	-0.729309	-0.703630	1.008309
Prob	0.9898	0.0029	0.3521	0.5062	0.5205	0.3703
R-squared		0.981872	Mean dependent var		9106.875	
Adjusted R-squared		0.950148	S.D. dependent var		6616.478	
S.E. of regression		1477.297	Akaike info criterion		17.66854	
Sum squared resid		8729625	Schwarz criterion		17.99181	
Log likelihood		-98.01122	Hannan-Quinn criter.		17.54885	
F-statistic		30.95050	Durbin-Watson stat		2.660883	
Prob(F-statistic)		0.002511				

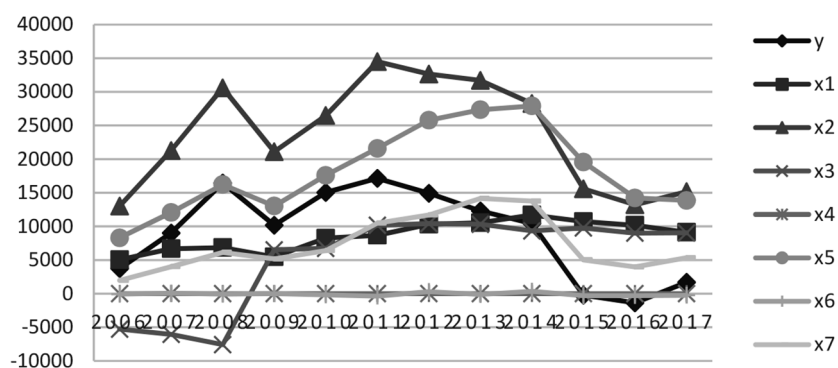


Рис. 3. Динамика развития экономических параметров, входящих в модель на 2006–2017 гг.

Fig. 3. Development dynamics of economic parameters included in the model for 2006-2017

(coefficient of variation), defined as the specific gravity of the ratio of standard variations of parameters to mathematical expectations:

$$v_{x_i} = \frac{\sigma_{x_i}}{x_i} 100\%.$$

$$\sigma_{x_i} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}.$$

The results of the descriptive statistics in tab.6 were used to analyze the overall statistics and variability of the model. All re-

sults obtained should not be less than 10% to ensure variability, otherwise if this condition isn't met for any variable, it may be more appropriate to remove the variable from the model or replace it with another variable.

There is no need to replace all variables as the required condition is met.

A multicollinear analysis which expressing the correlation dependences between the explained and explanatory coefficients was also performed based on the Pearson coefficients. The calculated double-line correlation coefficients in tab.8 demonstrate strongly straight dependence between y and x_5 ($r = 0,74$) and y and x_2 ($r = 0,92$), strongly inverse dependence to y and x_4 ($r = -0,74$), dependence at straight medium density between y and x_7 ($r = 0,55$), weakly straight dependence between y and x_6 ($r = 0,32$), and very weak inverse dependence between y and x_1 ($r = -0,11$) and x_3 ($r = -0,08$). All coefficients were calculated without taking into account the lags.

To evaluate the cause-and-effect relationships between the variables in our model, we performed the Granger test based on the Fisher criterion. According to the test, the H_0 hypothesis rejects the existence of causal dependencies between all possible pairs in the model with *Prob.(F-statistic)* assumption[19].

Among the many results we obtained, the number of observations, *F-Statistic* and *Prob.(F-statistic)* is indicated in the tab.9 which can be a one-sided reason for the current account of the balance of payments and formulate based on the corresponding results that reject the H_0 hypothesis. According to the *F-Statistic* criterion for the current account of the balance of payments, the pairs that confirm the H_1 hypothesis with 95% probability are y - x_1 (foreign investments), y - x_2 (exports), y - x_5 (total investments), y - x_7 (foreign exchange reserves). Given the small number of obser-

Таблица 6 (Table 6)

Результаты описательной статистики

Results of descriptive statistics

	y	x_7	x_6	x_5	x_4	x_3	x_2	x_1
Mean	9106.875	7341.767	-79.37500	18121.17	0.977250	5201.833	23623.33	8626.508
Median	10298.20	5735.800	-110.2500	16906.70	0.812000	9020.500	23872.50	8897.200
Maximum	17146.10	14152.00	352.8000	27907.50	1.720000	10417.00	34494.00	11697.70
Minimum	-1363.400	1967.300	-363.5000	8300.400	0.784000	-7574.000	13014.00	5052.800
Std. Dev.	6616.478	4091.665	229.6012	6393.992	0.329809	7062.213	8045.514	2190.651
Skewness	-0.383181	0.595558	0.729514	0.279494	1.579096	-1.071375	-0.086722	-0.330311
Kurtosis	1.710498	1.931446	2.479581	1.891123	3.802741	2.306700	1.456093	1.819013
Jarque-Bera	1.125064	1.280283	1.199800	0.771038	5.309284	2.536019	1.206866	0.915576
Probability	0.569765	0.527218	0.548866	0.680098	0.070324	0.281391	0.546931	0.632682
Sum	109282.5	88101.20	-952.5000	217454.0	11.72700	62422.00	283480.0	103518.1
Sum Sq. Dev.	4.82E+08	1.84E+08	579883.9	4.50E+08	1.196514	5.49E+08	7.12E+08	52788467
Observations	12	12	12	12	12	12	12	12

Таблица 7 (Table 7)

Коэффициенты вариации объясняющих переменных (в %)

Coefficients of variation on explanatory variables (in%)

x_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7
v_{x_i}	25,4	34	135,7	33,7	35,2	289,2	55,7

Таблица 8 (Table 8)

Корреляционная матрица

Correlation matrix

	y	GENINV x_5	FORINV x_1	EXP x_2	IMP x_3	MANAT x_4	BD x_6	REZ x_7
y	1	0.74	-0.11	0.92	-0.08	-0.77	0.32	0.55
GENINV x_5	0.74	1	0.07	0.76	0.26	-0.81	0.34	0.62
FORINV x_1	-0.11	0.07	1	0.23	0.74	0.21	0.09	0.65
EXP x_2	0.92	0.76	0.23	1	0.16	-0.68	0.32	0.79
IMP x_3	-0.08	0.26	0.74	0.16	1	0.22	-0.12	0.54
MANAT x_4	-0.77	-0.81	0.21	-0.68	0.22	1	-0.46	-0.42
BD x_6	0.32	0.34	0.09	0.32	-0.12	-0.46	1	0.4
REZ x_7	0.55	0.62	0.65	0.79	0.54	-0.42	0.4	1

Таблица 9 (Table 9)

Результаты тестов причинности Грейнджера (лаги: 1)

Results of Granger Causality Tests (lags: 1)

	Null Hypothesis:	Obs	F-Statistic	Prob.
x_1	FORINV does not Granger Cause y	11	8.37980	0.0201
x_2	EXP does not Granger Cause y	11	7.20144	0.0278
x_3	IMP does not Granger Cause y	11	1.09125	0.3267
x_4	MANAT does not Granger Cause y	11	0.03827	0.8498
x_5	CENINV does not Granger Cause y	11	9.34623	0.0156
x_6	BD does not Granger Cause y	11	0.48063	0.5078
x_7	REZ does not Granger Cause y	11	6.41873	0.0351

vations (2006-2017, annual), the test doesn't allow for more lags, so we were able to conduct the test with only 1 lag. This necessitates an increase in the number of observations. Due to the current

technical problems with the provision of statistical information, we plan to increase the number of our observations in order to address this problem in our future research. For $Y-X3(import)$, y - x_4

(*manat*) and y - x_6 (*budget deficit*), the test results are not very satisfactory, because the H_1 hypothesis can be accepted with very little probability.

In general, if autocorrelation is determined at the time series levels in the research process, it is important to eliminate it in some way before applying the regression equation for the forecast. If there is a strong autocorrelation between the levels, then it is better to use their differences in the calculations instead of quantitative indicators of the series. Differences in the form of $y_t = y_t - y_{t-1}$; $x_{t1} = x_{t,1} - x_{t-1,1}$; ... $x_{t7} = x_{t,7} - x_{t-1,7}$ [20] are applied instead of the y , x_1 , x_2 , x_3 , x_4 , x_5 , x_6 , x_7 variables in the model. The autoregression model for first-order differences is applied in the form of the following equation:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t.$$

When the parameters of this model are calculated by the method of the smallest squares, the model is formed in a form of $y_t = \alpha_0 + \alpha_1 y_{t-1}$ which is suitable for forecasting. It should be noted that it is purposeful to apply the method of differences on the basis of preliminary data when DW statistics approach 0 or 4 [18].

The stableness of the time series was tested on the basis of an extended Dickey-Fuller (DF) test. In this case, the H_0 hypothesis accepts the assumption that the time series under study has a single root. In the initial stage of the test (see tab.10), the H_0 hypothesis is accepted that the time series has a single root for the current account balance (y). The stationary nature of the time series is not confirmed, so the value of the t -Statistic criterion is obtained with a very small probability $t = -3.107$ with $p = 0.1646$. This indicates that the result obtained was observed with large errors, and the hypothesis H_0 , which confirms the single root, is not allowed to be rejected. At the same time, the fact that the test result for t -Statistic at the 1%, 5%, and 10% significance levels is

Таблица 10 (Table 10)

Расширенный тест Дики-Фуллера (с различиями 1-й степени)
Extended Dickey-Fuller test (with 1st degree differences)

t-Statistic	-3.107422	Log likelihood	-85.97255
1% level	-5.521860	F-statistic	6.662377
5% level	-4.107833	Prob(F-statistic)	0.033771
10% level	-3.515047	Mean dependent var	-487.3889
Prob.*	0.1646	S.D. dependent var	8078.168
R-squared	0.799897	Akaike info criterion	19.99390
Adjusted R-squared	0.679835	Schwarz criterion	20.08156
S.E. of regression	4570.878	Hannan-Quinn criter.	19.80474
Sum squared resid	1.04E+08	Durbin-Watson stat	1.736959

Таблица 11 (Table 11)

Расширенный тест Дики-Фуллера (с различиями 2-й степени)
Extended Dickey-Fuller test (with 2nd degree differences)

t-Statistic	-3.550973	Log likelihood	-78.52002
1% level	-4.582648	F-statistic	17.29797
5% level	-3.320969	Prob(F-statistic)	0.005666
10% level	-2.801384	Mean dependent var	2238.012
Prob.*	0.0370	S.D. dependent var	13326.02
R-squared	0.873724	Akaike info criterion	20.38000
Adjusted R-squared	0.823214	Schwarz criterion	20.40979
S.E. of regression	5603.045	Hannan-Quinn criter.	20.17908
Sum squared resid	1.57E+08	Durbin-Watson stat	1.570826

to the left of the critical values for t -Statistic brings the time series closer to stationary. $t = -3.107$ value is to the right of three critical values, $-5.521860(1\%)$; $-4.107833(5\%)$; $-3.515047(10\%)$; It should be noted that the extended DF test was performed on the basis of the autoregression model (AR) and *Schwarz* criteria with constant and trend, 1st degree differences. The maximum number of lags was taken as 3. Due to this, the number of observations decreased to 9.

By changing the test parameters, the test in the next stage was conducted based on the autoregression model (AR) with 2nd degree differences, stable, without trends and here the maximum number of lags is 2 (see tab.11).

From the results obtained in tab. 11, it can be seen that the stationary nature of the time series for the current account of the balance of payments can be accepted according to the results obtained. The value of the t -Statistic criterion is obtained with the probability $t = -3.55$ with $p = 0.037$. This indicates that the

result occurred with minimal error and allows the H_0 hypothesis to be rejected. The result obtained for t -Statistic is to the left respectively of -3.320969 ; -2.801384 critical values at the 5% and 10% significance levels. The adequacy of the autoregression model is also quite satisfactory. $R^2 = 0,87$ indicates that the overall quality of the model is high. The corrected determination coefficient is 82%, the value F -statistic = 17.3 was obtained with a high $Prob(F$ -statistic) = 0.005 assumption. $DW = 1.57$ and is close to 2.

The main results.

A multidimensional regression model has been established that allows economic and statistical analysis of the dynamics of the current account of the balance of payments;

The form and directions of functional dependence between dependent and independent variables were determined, variability was assessed, the results of multivariate regression analysis were

analyzed according to econometric methodologies; quantitative characteristics of the mechanisms of influence of explanatory factors on the balance of payments were measured and interpreted;

Correlation dependencies for cause-and-effect dependencies in the model were investigated, a Granger test was performed, and factors that significantly explained the outcome with high probabilities were identified based on the Fisher criterion;

The stability of the model was measured based on the Dick-

ey-Fuller test. With the first and second degree differences, the stationarity of the autoregression model was determined on the basis of the Student's criterion by changing the lag sizes.

The built-in autoregression model demonstrates sufficient adequacy, the time series for the dependent variable is stationary, and can be considered suitable for the forecast values of the current account of the balance of payments.

The results of the research case provide an opportunity to

identify real development trends in the balance of payments of Azerbaijan at the present stage and to determine its interdependence with other macroeconomic variables based on the analysis of the dynamics of the balance of payments and develop specific recommendations for balance of payments. The model also allows analyzing and forecasting the dynamics of Azerbaijan's national currency exchange rate and foreign economic activity within the balance of payments.

Литература

1. Mendoza E., Uribe M. The Business Cycles of Balance-of-Payments Crises: A Revision of a Mundellian Framework. NBER Working Paper. 1999. № 7045.
2. Mundell R. A. The Monetary Dynamics of International Adjustment under Fixed and Flexible Exchange Rates // The Quarterly Journal of Economics. 1960. T. 74. № 2. С. 227–257.
3. Thirlwall A. P., Hussain M. N. The Balance of Payments Constraint, Capital Flows and Growth Rate Differences between Developing Countries // Oxford Economic papers. 1982. T. 34. № 3. С. 498–510.
4. Moreno-Brid J. C. On Capital Flows and the Balance-of-Payments Constrained Growth Model // Journal of Post Keynesian Economics. 1998. T. 21. № 2. С. 283–298.
5. Kumhof M., Li S., Yan I. Balance of Payments Crises under Inflation Targeting // Journal of International Economics. 2007. T. 72. № 1. С. 242–264.
6. Cespedes L., Chang R., Velasco A. Balance Sheets, Exchange Rate Regimes, and Credible Monetary Policy [Электрон. ресурс] // Harvard University and NBER. 2001. Режим доступа: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.203.1042&rep=rep1&type=pdf>.
7. Кравцов М.К., Пашкевич А.В., Бурдыко Н.М., Гаспарец О.И. Система эконометрических моделей для анализа и краткосрочного прогнозирования основных макроэкономических показателей Республики Беларусь // Экономика и управление. 2007. № 3. С. 69–80.
8. Харемза В.В., Харин Ю.С., Макарова С.Б., Малюгин В.И., Гурин А.С., Раскина Ю.В. О моделировании Экономики России и Беларуси на основе эконометрической модели LAM-3 // Прикладная эконометрика. 2006. № 2. С. 124–139.
9. Chaldaeveva L.A., Chinaeva T.I., Bogopolskiy A.S. Analysis of financial and economic indicators, characterizing the activities of organizations in the oil and gas industry // Statistics and Economics. 2020. № 17(1). С. 69–78. DOI: 10.21686/2500-3925-2020-1-69-78/.
10. Ovcharov A.O., Terekhov A.M. Econometric analysis of the use of biological assets in agricultural

organizations // Statistics and Economics. 2020. № 17(1). С. 79–87. DOI: 10.21686/2500-3925-2020-1-79-87.

11. Pilnik N. P., Shaikhutdinova M. F. Modeling of the Balance of Payments State in Russia // Financial Journal. 2017. № 5. С. 84–101.

12. Engle, R. F. and Yoo, B. S. Cointegrated Economic Time Series: An Overview with New Results in R. F. Engle and C. W. J. Granger, eds., Long-Run Economic Relationships. Oxford: Oxford University Press, 1991. С. 237–266.

13. Johansen S. and Juselius K. Identification of the Long-run and the Short-run Structure: An Application to the ISLM Model // Journal of Econometrics. 1994. № 63. С. 7–36.

14. Official website of the Central Bank of Azerbaijan [Электрон. ресурс]. Режим доступа: <https://www.cbar.az>; <https://www.cbar.az/page-41/macro-economic-indicators>.

15. State Statistical Committee of the Republic of Azerbaijan [Электрон. ресурс]. Режим доступа: <https://www.cbar.az/page-41/macro-economic-indicators>.

16. Orudzhev Elshar G., Ayyubova Natavan S. Empirical analysis of the factors affecting the balance of payments in Azerbaijan // Aktual'ni Problemy Ekonomiky. Kiev. 2016. T. 181. С. 400–411.

17. Айюбова Н.С., Исазаде А.Э. Вопросы моделирования и анализа взаимосвязей валютного курса маната и детерминантов национальной экономики Азербайджана // Евразийский Союз Ученых (ЕСУ), Ежемесячный научный журнал. 2021. Т. 6. № 1(82). С. 4–11. Серия: Экономические науки. ISSN 2411-6467. DOI: 10.31618/ESU.2413-9335.2021.6.82.

18. Dickey D.A., Fuller W.A. Distribution of Estimators for Autoregressive Time Series with a Unit Root // Journal of the American Statistical Association. 1979. № 74. С. 427–431.

19. Granger C.W.J., Newbold P. Spurious Regressions in Econometrics // Journal of Econometrics. 1974. T. 2. С. 111–117.

20. Kantorovich Grigory. Time series analysis // Higher School of Economics Economic Journal. 2002. T. 6. № 2. С. 498–523

References

1. Mendoza E., Uribe M. The Business Cycles of Balance-of-Payments Crises: A Revision of a Mundellian Framework. NBER Working Paper. 1999; 7045.
2. Mundell R. A. The Monetary Dynamics of International Adjustment under Fixed and Flexible Exchange Rates. The Quarterly Journal of Economics. 1960; 74; 2: 227–257.
3. Thirlwall A. P., Hussain M. N. The Balance of Payments Constraint, Capital Flows and Growth Rate Differences between Developing Countries. Oxford Economic papers. 1982; 34; 3: 498–510.
4. Moreno-Brid J. C. On Capital Flows and the Balance-of-Payments Constrained Growth Model. Journal of Post Keynesian Economics. 1998; 21; 2: 283–298.
5. Kumhof M., Li S., Yan I. Balance of Payments Crises under Inflation Targeting. Journal of International Economics. 2007; 72; 1: 242–264.
6. Cespedes L., Chang R., Velasco A. Balance Sheets, Exchange Rate Regimes, and Credible Monetary Policy [Internet]. Harvard University and NBER. 2001. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.203.1042&rep=rep1&type=pdf>.
7. Kravtsov M.K., Pashkevich A.V., Burdyko N.M., Gaspardarets O.I. System of econometric models for analysis and short-term forecasting of the main macroeconomic indicators of the Republic of Belarus. Ekonomika i upravleniye = Economics and Management. 2007; 3: 69–80. (In Russ.)
8. Kharemza V.V., Kharin YU.S., Makarova S.B., Malyugin V.I., Gurin A.S., Raskina YU.V. On modeling the Economy of Russia and Belarus based on the LAM-3 econometric model. Prikladnaya ekonometrika = Applied Econometrics. 2006; 2: 124–139. (In Russ.)
9. Chaldaeva L.A., Chinaeva T.I., Bogopolskiy A.S. Analysis of financial and economic indicators, characterizing the activities of organizations in the oil and gas industry. Statistics and Economics. 2020; 17(1):69–78. DOI: 10.21686/2500-3925-2020-1-69-78/. (In Russ.)
10. Ovcharov A.O., Terekhov A.M. Econometric analysis of the use of biological assets in agricultural organizations. Statistics and Economics. 2020; 17(1): 79–87. DOI: 10.21686/2500-3925-2020-1-79-87. (In Russ.)
11. Pilnik N. P., Shaikhutdinova M. F. Modeling of the Balance of Payments State in Russia. Financial Journal. 2017; 5: 84–101.
12. Engle, R. F. and Yoo, B. S. Cointegrated Economic Time Series: An Overview with New Results in R. F. Engle and C. W. J. Granger, eds., Long-Run Economic Relationships. Oxford: Oxford University Press; 1991: 237–266.
13. Johansen S. and Juselius K. Identification of the Long-run and the Short-run Structure: An Application to the ISLM Model. Journal of Econometrics. 1994; 63: 7–36.
14. Official website of the Central Bank of Azerbaijan [Internet]. Available from: <https://www.cbar.az/>; <https://www.cbar.az/page-41/macro-economic-indicators>.
15. State Statistical Committee of the Republic of Azerbaijan [Internet]. Available from: <https://www.cbar.az/page-41/macro-economic-indicators>.
16. Orudzhev Elshar G., Ayyubova Natavan S. Empirical analysis of the factors affecting the balance of payments in Azerbaijan. Aktual'ni Problemy Ekonomiky. Kiev. 2016; 181: 400–411.
17. Ayyubova N.S., Isazade A.E. Issues of modeling and analysis of interrelations between the exchange rate of the manat and the determinants of the national economy of Azerbaijan. Yevraziyskiy Soyuz Uchenykh (YESU), Yezhemesyachnyy nauchnyy zhurnal = Eurasian Union of Scientists (ESU), Monthly scientific journal. 2021; 6; 1(82): 4–11. Series: Economic sciences. DOI: 10.31618/ESU.2413-9335.2021.6.82. (In Russ.)
18. Dickey D.A., Fuller W.A. Distribution of Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association. 1979; 74: 427–431.
19. Granger C.W.J., Newbold P. Spurious Regressions in Econometrics. Journal of Econometrics. 1974; 2: 111–117.
20. Kantorovich Grigory. Time series analysis. Higher School of Economics Economic Journal. 2002; 6; 2: 498–523.

Сведения об авторе

Натаван Солтан Айюбова

К.э.н., доцент кафедры Математической экономики
Бакинский Государственный Университет,
Баку, Азербайджан
Эл. почта: neyyubova@mail.ru

Information about the author

Natavan S. Ayyubova

Cand. Sci. (Economics), Associate Professor of the
Department of Mathematical Economics,
Baku State University,
Baku, Azerbaijan
E-mail: neyyubova@mail.ru

Смертность в Волгоградской области на фоне пандемии Covid-19

Пандемия COVID-19, которая началась в России в марте 2020 года, оказала огромное воздействие на социально-экономические процессы. В многочисленных исследованиях, посвященных анализу смертности, вызванной коронавирусной инфекцией, делается вывод о занижении числа умерших. Высокая заболеваемость и смертность, обусловленная коронавирусной инфекцией, имеет далеко идущие последствия для экономики регионов и страны в целом: ухудшение здоровья, снижение численности трудоспособного населения, изменение структуры потребления товаров и услуг и т.д. В связи с этим **актуальным** является анализ процессов, связанных со смертностью от коронавирусной инфекции.

Целью исследования является выявление основных тенденций в нозологической и половозрастной структуре смертности населения Волгоградской области в годы, предшествующие пандемии COVID-19, оценка вклада смертности от коронавирусной инфекции в общую смертность в 2020 году. Оценка избыточной смертности проводилась с учетом динамики возрастных коэффициентов смертности.

Материалы и методы. Основными источниками информации для исследования смертности была Российская база данных по рождаемости и смертности и данные Росстата. В работе при анализе смертности от COVID-19 использовались также данные оперативного штаба. Анализ динамики смертности проводился с помощью таких показателей как средняя ожидаемая продолжительность жизни при рождении, общий коэффициент смертности, возрастные показатели смертности в абсолютном и относительном (на 1000 человек) выражении. Обработка

статистических данных проводилась с помощью пакета прикладных программ Microsoft Excel и библиотек matplotlib, pandas, numpy (язык программирования Python), pyramid (язык программирования R).

Результаты. В 2020 году число умерших в Волгоградской области оказалось больше, чем в 2019 году на 6647 человек. Если бы тенденции в динамике интенсивности смертности сохранились бы и в год пандемии, то общее количество умерших по Волгоградской области оказалось бы равным 32044 человек. При этом избыточная смертность составила бы 7368 человек.

Заключение. В результате исследования было выявлено, что существенный рост числа умерших в Волгоградской области в период пандемии объясняется Росстатом в качестве причины коронавирусной инфекцией только на 33,2%. Данное несоответствие может являться следствием неправильного учета смертности от коронавирусной инфекции. Другим фактором увеличения смертности в период пандемии может быть снижение качества медицинского обслуживания. Произошла переориентация работы медицинских учреждений на лечение пациентов с коронавирусной инфекцией, увеличилась нагрузка на скорую медицинскую помощь.

Ключевые слова: COVID-19, коронавирусная инфекция, смертность, общий коэффициент смертности, средняя ожидаемая продолжительность жизни при рождении, Волгоградская область, избыточная смертность.

Alexey V. Alpatov

Volgograd State Technical University, branch of the Volga Polytechnical Institute, Volzhskiy, Russia

Mortality in the Volgograd Region Against the COVID-19 Pandemic

The COVID-19 pandemic, which began in Russia in March 2020, had a huge impact on socio-economic processes. In numerous studies analyzing mortality caused by coronavirus infection, it is concluded that the number of deaths is underestimated. The high morbidity and mortality caused by coronavirus infection has far-reaching consequences for the economy of the regions and the country as a whole: deterioration in health, a decrease in the working-age population, a change in the structure of consumption of goods and services, etc. In this regard, it is relevant to analyze the processes associated with mortality from coronavirus infection.

The purpose of the study is to identify the main trends in the nosological and age-sex structure of mortality in the Volgograd region in the years preceding the COVID-19 pandemic, to assess the contribution of mortality from coronavirus infection to total mortality in 2020. Estimation of excess mortality was carried out taking into account the dynamics of age-specific mortality rates.

Materials and methods. The main sources of information for the study of mortality were the Russian database on fertility and mortality and Rosstat data. In the work, when analyzing mortality from COVID-19, data from the operational headquarters were also used. The analysis of the mortality dynamics was carried out using such indicators as the average life expectancy at birth, the crude death-rate, age-specific

mortality rates in absolute and relative (per 1000 people) terms. The processing of statistical data was carried out using the Microsoft Excel application package and matplotlib, pandas, numpy (Python programming language), pyramid (R programming language) libraries.

Results. In 2020, the number of deaths in the Volgograd region turned out to be more than in 2019 by 6647 people. If the trends in the dynamics of the intensity of mortality would persist in the year of the pandemic, then the total number of deaths in the Volgograd region would be equal to 32044 people. In this case, the excess mortality would have amounted to 7368 people.

Conclusion. As a result of the study, it was revealed that a significant increase in the number of deaths in the Volgograd region during the pandemic is explained by Rosstat as the cause of coronavirus infection by only 33.2%. This discrepancy may be the result of incorrect accounting of deaths from coronavirus infection. Another factor in the increase in mortality during a pandemic may be a decrease in the quality of medical care. There has been a reorientation of the work of medical institutions to the treatment of patients with coronavirus infection; the burden on emergency medical care has increased.

Keywords: COVID-19, coronavirus infection, mortality, crude death-rate, life expectancy at birth, Volgograd region, excess mortality.

Введение

Пандемия COVID-19, которая началась в России в марте 2020 года, существенно повлияла на характер социально-экономических процессов. Произошла перестройка практически всех сфер человеческой деятельности. Наибольшее негативное воздействие пандемия оказала на протекание демографических процессов: увеличение смертности и снижение рождаемости [1].

В научных исследованиях, в которых рассматривается влияние коронавирусной инфекции на смертность, наибольшее внимание уделяется следующим вопросам: подсчет числа умерших от причин, связанных с COVID-19, их структура по возрасту и полу, влияние пандемии на изменение общих показателей смертности, выявление факторов, которые увеличивают риск смерти от COVID-19.

В публикациях [2,3,4] делается вывод о занижении числа умерших вследствие коронавирусной инфекцией. Например, в работе [2] отмечается, что в России в 2020 году итоговая смертность оказалась выше по сравнению с 2019 годом на 323,7 тыс. человек или на 18%. При этом число умерших от COVID-19 по данным Росстата в избыточной смертности составляет 50,2%. В регионах России доля смертей, обусловленной коронавирусной инфекцией в избыточной смертности, существенно различается. В Москве доля ковидных смертей почти совпадает с избыточной смертностью и составляет 98,4%. Довольно высокий уровень в Еврейской АО (92,4%) и Санкт-Петербурге (91,0%). К регионам, в которых избыточная смертность меньше всего была объяснена коронавирусной инфекцией относятся Ленинградская область (5,8%), Рязанской область (4,5%), Чеченская Республика (3,9%). В резуль-

тате сопоставления данных о смертности между регионами России и другими странами в период пандемии автор в работе [2] делает предположение о «существенном расхождении в методологии учета смертности в различных субъектах Российской Федерации». Это как минимум, а как максимум «речь может идти о намеренном искажении данных».

При международном сравнении смертности от коронавирусной инфекции Россия показывает не самые лучшие результаты. В работе [5] была проведена оценка снижения средней ожидаемой продолжительности жизни при рождении (далее – ОПЖ) в 2020 году на фоне пандемии COVID-19 для 37 стран мира с уровнем жизни выше среднего и для которых есть надежные и полные данные о смертности в период с 2005 по 2020 год. Уровень снижения ОПЖ определялся как разница между фактическим значением ОПЖ и прогнозным. При прогнозировании ОПЖ использовалась модель Ли-Картера [6]. Наихудший результат был зафиксирован в России. Снижение ОПЖ у мужчин составило 2,33 лет, у женщин – 2,14 лет. Второй по степени снижения оказались США (мужчины – 2,27, женщины – 1,37). В таких странах как Дания, Исландия и Южная Корея уменьшение ОПЖ не выявлено.

В работах [7,8,9] исследовались факторы, влияющие на повышение смертности в условиях пандемии в России. К экономическим показателям, которые влияют на смертность, относятся валовой региональный продукт на душу населения [8], уровень безработицы [7]. Среди демографических показателей наибольшее влияние оказывает доля городского населения [9], а также возрастная структура населения: уровень смертности повышается с увеличением возраста. Это оценивалось с помощью доли

пенсионеров [8], доли лиц старше трудоспособного возраста [7].

Научные публикации, в которых анализируется смертность в регионах России в период пандемии COVID-19, нацелены прежде всего на выявление общих закономерностей. При этом не уделяется существенное внимание особенностям отдельных регионов. Понимание демографических процессов в настоящее время крайне важно при планировании социально-экономического развития. Высокая заболеваемость и смертность, обусловленная коронавирусной инфекцией, имеет далеко идущие последствия: ухудшение здоровья, снижение численности трудоспособного населения, изменение структуры потребления товаров и услуг и т.д. Анализ публикаций на момент написания статьи показал отсутствие подробных демографических исследований по Волгоградской области. В связи с этим актуальным является анализ процессов, связанных со смертностью от коронавирусной инфекции на региональном уровне.

Целью настоящей работы является выявление основных тенденций в нозологической и половозрастной структуре смертности населения Волгоградской области в годы, предшествующие пандемии COVID-19, оценка вклада смертности от коронавирусной инфекции в общую смертность в 2020 году. Измерение избыточной смертности проводилось с учетом динамики возрастных коэффициентов смертности.

1. Статистические данные и методы исследования

Основным источником информации для исследования смертности была Российская база данных по рождаемости и смертности [10]. Оперативные данные по учету числа умер-

ших за 2019, 2020 и 2021 год, а также динамика смертности по основным причинам были взяты с сайта Федеральной службы государственной статистики [11]. В работе при анализе смертности от COVID-19 использовались также данные оперативного штаба [12].

Анализ динамики показателей смертности населения Волгоградской области проводился при сопоставлении с общероссийскими показателями. Смертность исследовалась в целом по региону, в разрезе мужчины/женщины, по основным классам причин смерти. Использовались такие показатели как ОПЖ, общий коэффициент смертности (далее — ОКС), возрастные показатели смертности в абсолютном и относительном (на 1000 человек) выражении.

Расчет ОПЖ по Волгоградской области и в целом по России проводился автором на основе однолетних таблиц смертности. Если сравнивать эти результаты с данными, опубликованными Росстатом, то есть небольшие отличия. Для мужчин среднее арифметическое различий в показателе ОПЖ равно $\Delta\text{ОПЖ} = 0,00$ со средним квадратическим отклонением $\sigma = 0,04$. Для женщин, соответственно, $\Delta\text{ОПЖ} = -0,02$, $\sigma = 0,06$.

Оценка влияния пандемии, вызванной COVID-19, на смертность в 2020 году осуществлялась следующим образом. Сначала были рассчитаны среднегодовые темпы роста возрастных коэффициентов смертности для мужчин и женщин во временном диапазоне с 2015 по 2019 год. Затем на основе среднегодовых темпов роста был сделан прогноз по каждой возрастной группе: каким было бы число умерших, если бы тенденции в сокращении смертности сохранились. Разница между прогнозным значением числа умерших в 2020 году и фактическим количеством умерших в 2019 яв-

ляется мерой оценки влияния пандемии на итоговую смертность в регионе.

2. Динамика смертности в Волгоградской области

В настоящее время одной из острейших проблем Волгоградской области является продолжающаяся депопуляция населения. Сокращение численности населения происходит вследствие естественной убыли, а также за счет миграционного оттока населения в другие регионы России. Руководство страны и области предпринимает попытки для улучшения демографической ситуации. Принимаются законы и проекты, которые должны стимулировать рождаемость и снижать смертность.

Существенное влияние на увеличение показателей рождаемости в период с 2007 по

2017 год оказала благоприятная половозрастная структура населения — наличие большой когорты женщин фертильного возраста, которые родились во второй половине 80-х годов прошлого века. Кроме того, принятие Федерального закона № 256-ФЗ от 29.12.2006 «О дополнительных мерах государственной поддержки семей, имеющих детей» повлияло на календарный сдвиг рождаемости. Это проявилось в росте суммарного коэффициента рождаемости в 2007–2015 годах [13]. В ближайшее время число родившихся будет сокращаться, поскольку в возраст с наибольшей интенсивностью деторождений будут вступать малочисленные поколения женщин, родившихся в середине 90-х годов. Это можно увидеть из рисунка 1, на котором показана возрастная структура населения Вол-

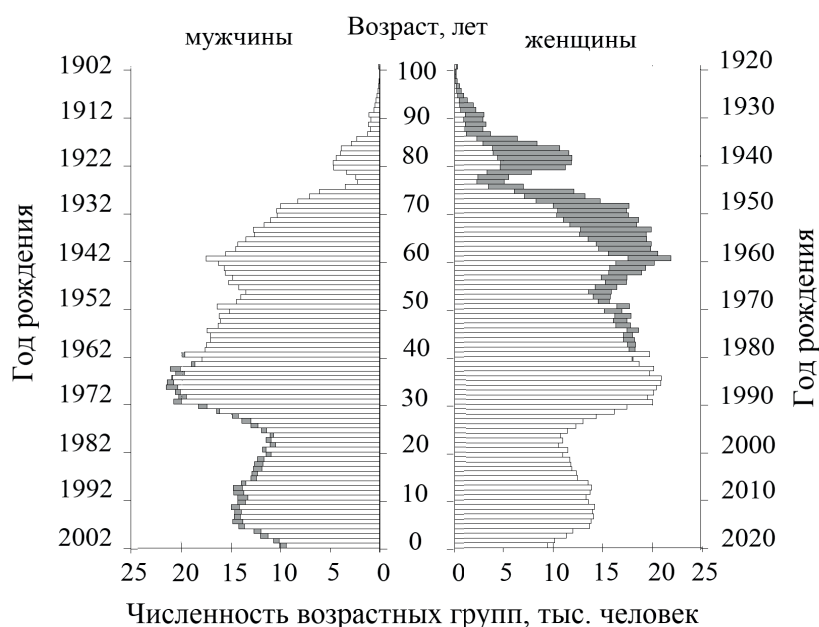


Рис. 1. Возрастно-половая структура населения Волгоградской области за 2021 год. Темная заливка слева показывает перевес мужчин над женщинами в данной возрастной группе, справа показывает, соответственно перевес женщин над мужчинами.

Источник: рассчитано автором на основе данных [11].

Fig. 1. Sex and age structure of the population of the Volgograd region for 2021. The dark shading on the left shows the predominance of men over women in this age group, on the right shows the predominance of women over men, respectively.

Source: calculated by the author based on the data [11].

гоградской области на начало 2021 года. Возрастно-половая структура имеет волнообразную форму, что является характерным для России и для большей части ее регионов. На рисунке приведена дополнительная шкала с указанием года рождения поколения.

Наиболее существенные результаты, связанные с демографическими процессами в регионе, были достигнуты в области снижения смертности. ОПЖ в период с 2003 по 2015 год выросла у всего населения Волгоградской области на 5,8 лет и составила 71,9 лет [14].

На рис. 2. показана динамика числа умерших по региону в период с 1989 по 2020 год (оба пола). Можно отметить, что в течение рассматриваемого периода данный показатель существенно варьируется. Значения изменяются от 29,72 тыс. умерших в 1989 году до 42,97 тыс. в 2002. Такая противоречивая динамика числа умерших обусловлена волнообразной половозрастной структурой населения (см. рис. 1), а также высокой смертностью в 90-х годах XX столетия, прежде всего среди мужчин трудоспособного возраста, что стало предметом многочисленных исследований

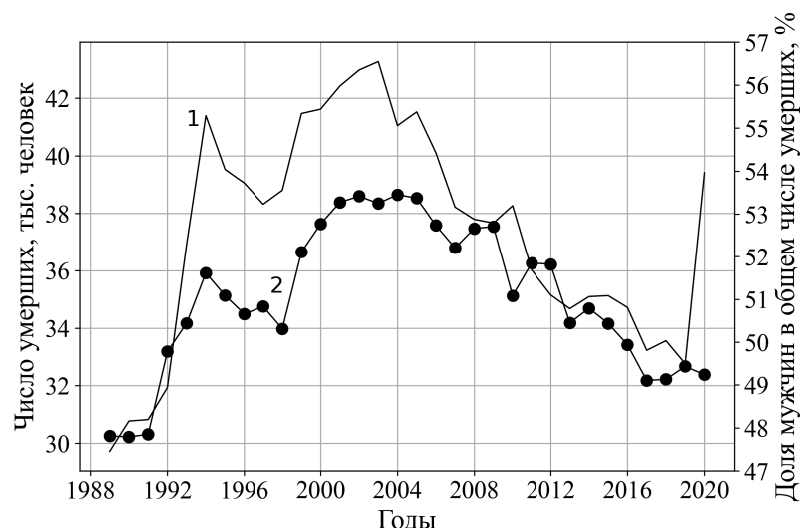


Рис. 2. Динамика общего числа умерших всего населения Волгоградской области (кривая 1) и доли мужчин в общем числе умерших (кривая 2).

Источник: рассчитано автором на основе данных [10].

Fig. 2. Dynamics of the total number of deaths of the entire population of the Volgograd region (curve 1) and the proportion of men in the total number of deaths (curve 2).

Source: calculated by the author based on the data [10].

как на уровне России в целом, так и на региональном уровне [14].

Кривая, показывающая на рис. 2 динамику доли мужчин в общем числе умерших, довольно хорошо повторяет характер временного ряда абсолютного показателя смертности. Это во многом отражает тот факт, что существенный вклад в об-

щую смертность вносит сверхсмертность мужчин, особенно в сложные периоды развития России. С 1993 по 2016 год доля умерших в течение года мужчин составляла более 50%. С 2017 года наметился гендерный разворот в сторону преобладания женщин в структуре общей смертности. Данный факт, очевидно, обусловлен

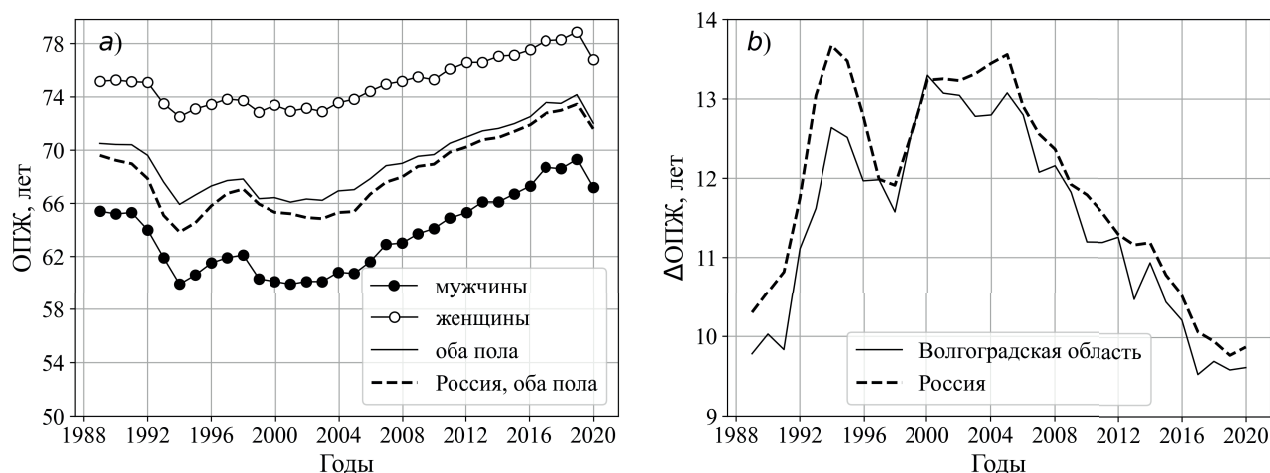


Рис. 3. Динамика средней ожидаемой продолжительности жизни при рождении ОПЖ, лет (а) и разница в ОПЖ между женщинами и мужчинами ΔОПЖ (b).

Источник: рассчитано автором на основе данных [10].

Fig. 3. Dynamics of average life expectancy at birth, years (a) and the difference in life expectancy between women and men (b).

Source: calculated by the author based on the data [10].

тем, что длительный период вымирания мужчин привел к существенной гендерной диспропорции в старших возрастных группах.

На рис. 3а показана динамика интегрального показателя смертности – средней ожидаемой продолжительности жизни, дифференцированной по полу. Для сравнения приведены временные ряды ОПЖ всего населения Волгоградской области и России в целом. Общее улучшение социально-экономической ситуации в стране позитивно отразилось и на динамике ОПЖ. С 2004 по 2019 год включительно наблюдался практически стабильный рост данного показателя у обоих полов. При этом ОПЖ мужчин выросла с 60,1 лет (2003 год) до 69,3 (2019 год), т.е. на 15,3%, а у женщин в течение рассматриваемого временного интервала рост составил 8,2% (с 72,9 лет в 2003 до 78,9 лет в 2019 году).

Таким образом, мужское население существенно сократило свое отставание в средней ожидаемой продолжительности жизни. Это особенно видно из рисунка 3б, на котором показана разница в ОПЖ между женщинами и мужчинами –

ΔОПЖ. Наибольшая гендерная разница в ОПЖ по региону наблюдалась в 2000 году и составляла 13,3 лет. К 2020 году вследствие того, что снижение уровня смертности у мужчин происходило более высокими темпами, ΔОПЖ уменьшилась и составила 9,6 лет.

При сравнении ОПЖ населения Волгоградской области с общероссийскими значениями видно, что кривая для населения региона на протяжении всего исследуемого периода остается выше значений показателя ОПЖ для России в целом. Из рис. 3б следует, что величина ΔОПЖ по Волгоградской области меньше чем средний показатель по России. Это различие особенно существенно в 1994 году (ΔОПЖ для России было равным 13,7 лет) и в 2005 году (ΔОПЖ = 13,3 лет), когда по Волгоградской области уже несколько лет происходило сокращение ΔОПЖ.

Необходимо отметить, что большой гендерный разрыв в показателе ОПЖ является характерной особенностью России и ряда государств постсоветского пространства [15]. Одним из самых существенных факторов, который влияет на сверхсмертность мужчин

в целом по стране и в Волгоградской области в частности, является бытовое пьянство. Подробные исследования о связи смертности и потребления алкоголя проводятся Немцовым А. В., например, [16, 17]. В работе [17] отмечается, что с 2004 года происходит заметное уменьшение потребления алкоголя в России. Это согласуется с кривыми, представленными на рис. 3.

Снижение количества умерших, которое с 2004 года продолжалось до 2019 года, обусловлено, как было показано выше, уменьшением смертности. В течение данного периода население Волгоградской области сократилось с 2673,1 до 2507,5 тыс. человек. Это является дополнительным фактором снижения числа умерших. Однако общая положительная динамика была остановлена в 2020 году. Это год начала пандемии, вызванной коронавирусной инфекцией. Число умерших в 2020 году в регионе составило 39410 человек, а это на 6647 человек больше, чем в предыдущем году. Такого однолетнего прироста числа умерших не было даже в самые кризисные 90-е годы, тем более на фоне ежегодного

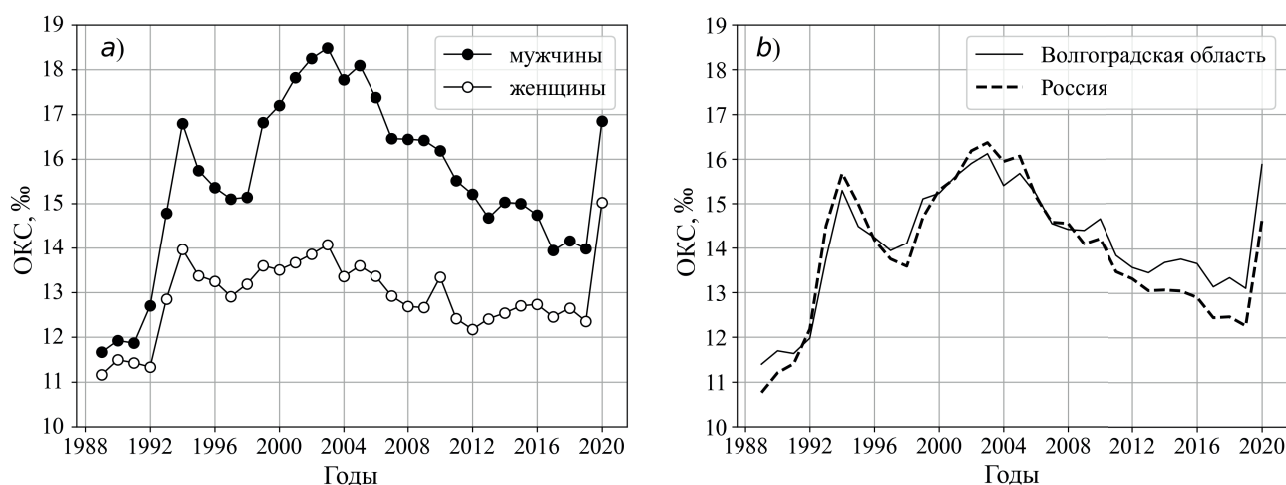


Рис. 4. Динамика ОКС: а) мужчин и женщин по Волгоградской области; б) всего населения Волгоградской области и России.

Источник: рассчитано автором на основе данных [10].

Fig. 4. Dynamics of the crude death-rate: a) men and women in the Volgograd region; b) the entire population of the Volgograd region and Russia.

Source: calculated by the author based on the data [10]

сокращения населения в регионе. ОПЖ в Волгоградской области за 2020 год уменьшилась и у мужчин, и у женщин, соответственно, на 1,9 и 1,8 лет (см. рис. 3а). Такая же тенденция наблюдается и в целом по стране. На резкий рост смертности, безусловно, повлияла эпидемия. Однако, данный вопрос будет более детально рассмотрен в следующем разделе работы. Сейчас остановимся на некоторых аспектах дина-

Таблица 1 (Table 1)

Оперативные данные по числу умерших в течение четырех месяцев для населения Волгоградской области (оба пола), человек
Operational data on the number of deaths within four months for the population of the Volgograd region (both sexes), people

Месяц	2009	2010	2011
июнь	3246	3416	2901
июль	3584	3427	3156
август	2667	4673	3229
сентябрь	2809	2841	2680

Источник: составлено автором на основе данных [11]

Source: compiled by the author based on the data [11]

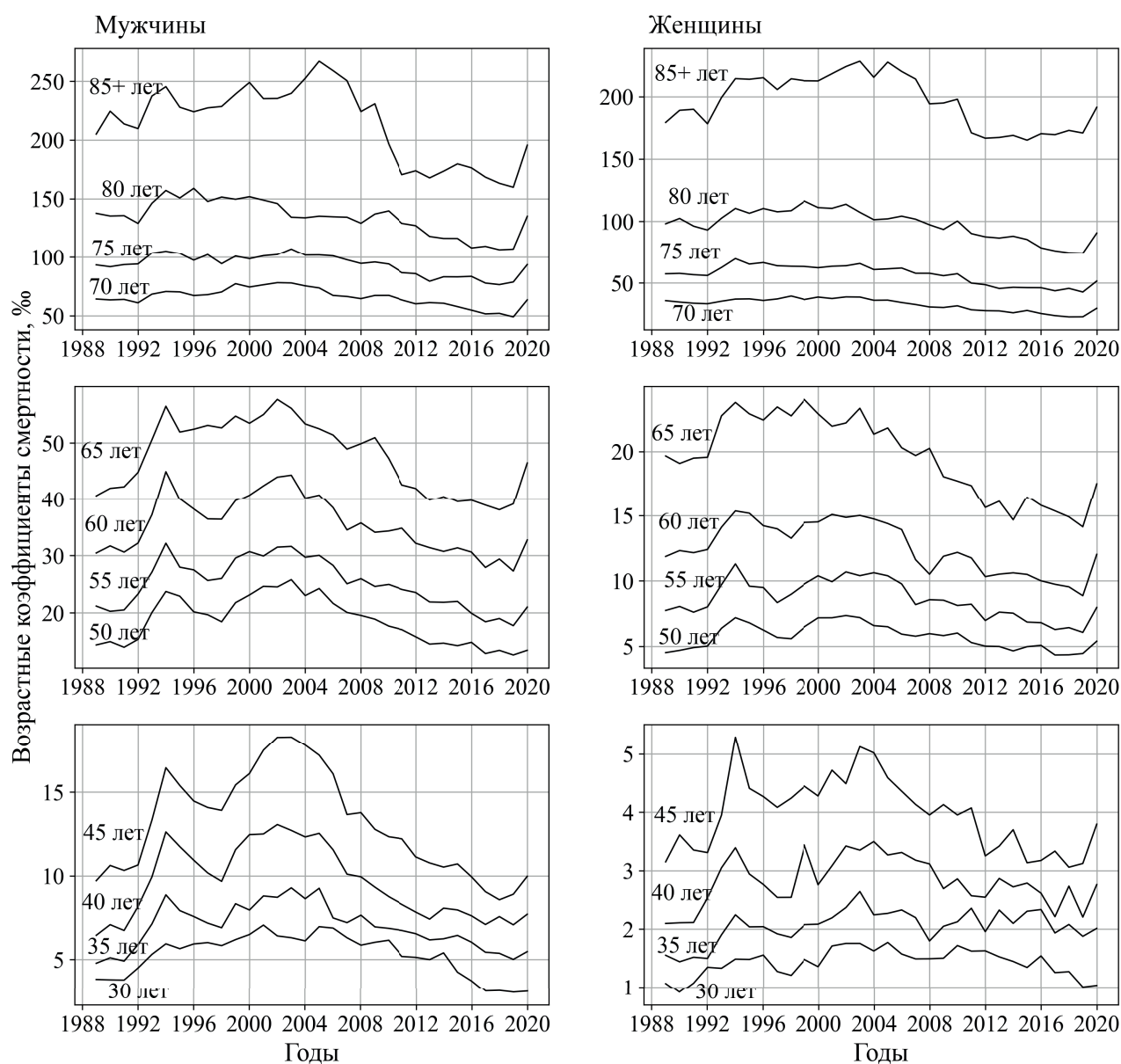


Рис. 5. Динамика пятилетних возрастных коэффициентов смертности населения Волгоградской области старше 30 лет в разрезе мужчин/женщин.

Источник: рассчитано автором на основе данных [10].

Fig. 5. Dynamics of five-year age-specific mortality rates of the population of the Volgograd region over 30 years old for men and women.

Source: calculated by the author based on the data [10]

мики смертности в течение 10–15 лет, которые предшествовали началу пандемии.

На рис. 4 показана динамика общего коэффициента смертности для мужского и женского населения Волгоградской области, а также для обоих полов по региону и по России в целом. Общий коэффициент смертности применяется для сопоставления данных о смертности в различных статистических совокупностях. Показатель ОКС не зависит от общей численности населения, а зависит от возрастной структуры населения, а также от распределения интенсивности смертей по возрастным группам. Характер кривых с 2003 по 2019 год для мужчин и женщин несколько отличается друг от друга. В частности, динамика ОКС у населения мужского пола в целом положительная. Выпадают из общей тенденции только 2013 и 2017 годы, в которых снижением ОКС было несколько завышенным. На кривой ОКС для женского населения есть два

основных участка: 2003–2012 и 2012–2019. На первом из них коэффициент смертности снижается, а на втором демонстрирует довольно продолжительный рост.

Обращает на себя внимание аномально высокое значение ОКС в 2010 году для населения женского пола. Это согласуется и с другими данными: на рис. 2 в 2010 доля смертей мужчин существенно сократилась, на рис. 3 можно увидеть небольшое снижение ОПЖ женщин. Лето 2010 года был очень жарким во многих регионах России, в том числе и в Волгоградской области. Существует ряд исследований, посвященных влиянию климатических условий на динамику смертности. Увеличение смертности летом 2010 года в России анализировалось, например, в работе [18]. В частности, было отмечено, что в августе 2010 года по сравнению с августом 2009 года, смертность выросла в целом по стране на 27,4%. В таблице 1 приведены оперативные данные по числу умерших в течение четырех ме-

сяцев за 2009, 2010 и 2011 год по Волгоградской области для обоих полов. В июле-августе 2010 года регион попал в зону с аномально высокой температурой. Результатом воздействия неблагоприятных факторов окружающей среды стала высокая смертность в августе 2010 по сравнению с аналогичным периодом в 2009 и 2011 годах.

Анализ динамики возрастных коэффициентов смертности женщин (см. рис. 5) показывает, что в 2010 году наблюдается всплеск интенсивности смертности в возрастных группах старше 70 лет. Данные смерти и внесли наиболее существенный вклад в снижение ОПЖ. У мужчин снижения ожидаемой продолжительности жизни не произошло. При этом каких-либо существенных трансформаций в возрастном профиле коэффициентов смертности не выявлено.

На рис. 6. показана динамика общих коэффициентов смертности по основным классам смертности (оба пола). Как видно из рисунка характер

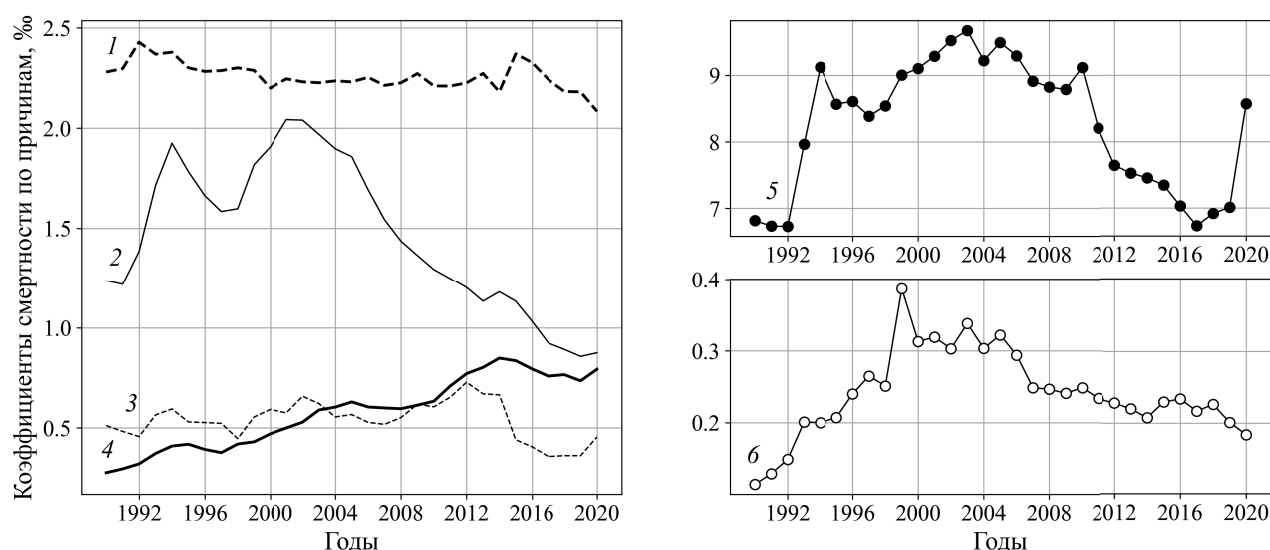


Рис. 6. Динамика коэффициентов смертности по основным классам причин смерти (оба пола), на 1000 человек: 1 – новообразования, 2 – внешние причины, 3 – болезни органов дыхания, 4 – болезни органов пищеварения, 5 – болезни системы кровообращения, 6 – инфекционные и паразитарные болезни.

Источник: рассчитано автором на основе данных [11].

Fig. 6. Dynamics of mortality rates by main classes of causes of death (both sexes), per 1000 people: 1 – neoplasms, 2 – external causes, 3 – diseases of the respiratory system, 4 – diseases of the digestive system, 5 – diseases of the circulatory system, 6 – infectious and parasitic diseases.

Source: calculated by the author based on the data [11].

динамики для различных классов причин смерти отличается. Наибольший вклад в смертность на протяжении всего исследуемого периода вносят болезни системы кровообращения (кривая 5). Можно отметить, что во временном диапазоне с 1993 по 2010 год именно данный класс причин обеспечивал высокий уровень смертности, который наблюдался в Волгоградской области. После 2010 года наблюдается снижения смертности от болезней системы кровообращения. В 2020 году произошел резкий рост соответствующего коэффициента смертности с 7,01‰ до 8,57‰.

Коэффициент смертности от новообразований (кривая 1) в течение исследуемого периода времени оставался практически на одном уровне и лишь в последние шесть лет намечилось снижение данного показателя с 2,37‰ в 2015 до 2,08‰ в 2020 году. Практически в этот же период произошло снижение смертности от болезней органов дыхания (кривая 3): с 0,67‰ в 2014 до 0,36‰ в 2019 году. Пандемия, вызванная коронавирусной инфекцией, привела к тому, что 2020 году анализируемый показатель смертности вырос до 0,46‰ (на 25,7%).

Наиболее существенно варьировался в течение исследуемого периода коэффициент смертности от внешних причин (кривая 2). В 2001 году данный показатель достиг своего пикового значения и составил 2,04‰. Начиная с 2002 года смертность от внешних причин стала сокращаться и в 2020 коэффициент смертности был уже равен 0,88‰. С 2000 года позитивную тенденцию демонстрирует смертность вследствие инфекционных и паразитарных болезней (кривая 6). Коэффициент смертности снизился с 0,39% в 1999 году до 0,18% в 2020.

На динамику смертности по причине болезней органов пищеварения (кривая 4) не ока-

зывали существенного влияния ни кризисные годы, ни периоды социально-экономического подъема. Коэффициент смертности от болезней органов пищеварения практически весь период рос. Начиная с 2015 года можно отметить стабилизацию данного коэффициента и даже небольшое снижение. В 2020 году смертность от болезней органов пищеварения увеличилась по сравнению с 2019 годом на 22,2%.

3. Вклад COVID-19 в смертность за 2020 год

Перейдем теперь к рассмотрению наиболее важного в настоящее время вопроса — влиянию пандемии на смертность в Волгоградской области. В таблице 2 приведены общие данные по числу умерших от коронавирусной инфекции оперативного штаба и Росстата за 2020 год. Наблюдается существенное расхождение между этими данными как в целом по России, так и по Волгоградской области.

Избыточная смертность в 2020 году по сравнению с 2019 годом для Волгоградской области составила 6647 человек, а по России 340286 человек. Ранее было отмечено, что в целом уровень смертности в регионе снижался. При этом уменьшалась и общая численность населения. Объяснить такое увеличение числа умер-

ших только изменениями в возрастной структуре населения не получится. Поэтому избыточную смертность можно связать с пандемией. Однако, если разделить смертность от COVID-19 по данным Росстата на избыточную смертность, то окажется, что по Волгоградской области доля составляет 36,8%, а по России 42,5%. Таким образом, более 50% избыточных смертей остаются необъясненными. Данное расхождение характерно практически для всех регионов. В работе [2] показано, что доля смертности от COVID-19 в избыточной смертности в декабре 2020 года варьируется по регионам России от 98,4% (г. Москва) до 3,9% (Чеченская республика). В данном «рейтинге» Волгоградская область занимает среднее положение.

В работе [19] исследовалась заболеваемость, вызванная COVID-19 в Волгоградской области в период с 24.03.2020 по 18.10.2020. Отмечается, что все пациенты со смертельными случаями имели несколько сопутствующих заболеваний. При этом «в подавляющем большинстве случаев летальные исходы от COVID-19 были ассоциированы с заболеваниями сердечно-сосудистой системы (65,3%) и патологией эндокринной и сердечно-сосудистой системы (21,1%)». Это коррелирует с ростом смертности вследствие болез-

Таблица 2 (Table 2)

Смертность от COVID-19 за 2020 год
Deaths from COVID-19 in 2020

Регион	Число умерших от COVID-19, человек		Общее число умерших в течение года, человек		Избыточная смертность по сравнению с 2019 годом, человек	Доля смертей, связанной с COVID-19, %
	Данные оперштаба	Данные Росстата	2019	2020		
Россия	56271	144699	1798307	2138586	340279	42,5
Волгоградская область	458	2448	32763	39410	6647	36,8

Источник: составлено автором на основе данных [10, 11, 12].

Source: compiled by the author based on the data [10, 11 and 12].

Таблица 3 (Table 3)

**Оперативные данные по числу умерших в Волгоградской области
(оба пола)**

**Operational data on the number of deaths in the Volgograd region
(both sexes)**

Месяц	Число умерших, человек			Индекс, %	
	2019	2020	2021	2020/2019	2021/2020
Январь	3041	2 934	3 838	96,5	130,8
Февраль	2680	2 655	3 165	99,1	119,2
Март	2847	2 643	3 263	92,8	123,5
Апрель	2834	2 793	3 323	98,6	119,0
Май	2924	3 119	3 237	106,7	103,8
Июнь	2608	2 753	3 145	105,6	114,2
Июль	2693	3 566	3 883	132,4	108,9
Август	2637	2 813	4 543	106,7	161,5
Сентябрь	2338	3 220	4384	137,7	136,1
Октябрь	2912	4 139	—	142,1	—
Ноябрь	2634	4 137	—	157,1	—
Декабрь	2704	4 342	—	160,6	—

Источник: составлено автором на основе данных [11]

Source: compiled by the author based on the data [11]

ней системы кровообращения (см. рис. 6) в 2020 году по сравнению с 2019 годом. В относительном выражении рост коэффициента смертности составил 22,2%. Сильнее всего вырос коэффициент смертности по классу «Болезни органов дыхания» — на 25,7%. Рост коэффициентов смертности по данным причинам был отмечен и в целом по России.

В 2020 году в России были две волны роста заболеваемости от COVID-19. Периоды развития этих волн в разных регионах отличались. Раньше всех эпидемия затронула Москву: пик заболеваемости пришелся на начало мая. В Волгоградской области первая волна началась в середине апреля и продолжилась до конца июля. Следующий всплеск заболеваемости произошел в октябре и продлился до конца 2020 года. Вторая волна сопровождалась увеличением числа заболевших и умерших по сравнению с первой волной [19]. Пиковые значения ежедневно выявляемых заболевших во время второй волны превышали примерно в три раза число заболевших во время первой волны. В таблице 3 приведены оперативные данные по количеству умерших в течение месяца по Волгоградской области за 2019, 2020 и 2021 годы. В пятом столбце приведены индексы, которые показывают отношение числа умерших в те-

чение месяца в 2020 году к числу умерших за аналогичный период в 2019 году. Видно, что до апреля включительно в регионе наблюдалось ежемесячное снижение умерших по сравнению с 2019 годом. Однако уже с началом первой волны пандемии наблюдается рост смертности: в мае — на 6,7%, июне — 5,6%, июле — 32,7%. Во время второй волны прирост числа умерших был существенно выше, например, в декабре он составил 60,6%. Таким образом, оперативные данные по учету числа умерших согласуются с особенностями протекания пандемии COVID-19.

Проанализируем половозрастной профиль умерших в Волгоградской области от коронавирусной инфекции по данным Росстата. Отметим, что до возраста 35 лет число умерших по данной причине небольшое. При этом избыточная смертность была отрицательной. Поэтому в первую возрастную группу включены умершие в возрасте до 35 лет; интервал следующих возрастных групп составил 10 лет. Наибольшее число смертей от COVID-19 у мужчин приходится на возрастную группу 65–74 лет, а у женщин — 75–84 лет. В этих же возрастных катего-

Таблица 4 (Table 4)

**Половозрастной профиль умерших от COVID-19
Sex and age profile of those, who died from COVID-19**

Возраст, лет	Число умерших от COVID-19, человек			Избыточная смертность в 2020 году по сравнению с 2019 годом, человек			Доля умерших в данной возрастной группе от COVID-19 в избыточной смертности, %		
	мужчины	женщины	оба пола	мужчины	женщины	оба пола	мужчины	женщины	оба пола
0–34	11	6	17	–3	–7	–10	—	—	—
35–44	30	21	51	119	67	186	25,2	31,4	27,4
45–54	96	85	181	145	134	279	66,2	63,5	64,9
55–64	256	233	489	651	477	1128	39,3	48,9	43,3
65–74	383	360	743	1281	1036	2316	29,9	34,8	32,1
75–84	305	420	725	607	1084	1691	50,2	38,8	42,9
85+	94	148	242	410	647	1057	22,9	22,9	22,9
Всего	1175	1273	2448	3210	3438	6647	36.6	37.0	36.8

Источник: составлено автором на основе данных [10]

Source: compiled by the author based on the data [10]

риях наблюдается наибольшая избыточная смертность. Доля умерших от COVID-19 в избыточной смертности выше всего в возрастном интервале 45–54 лет и составляет 66,2% для мужчин и 63,5% для женщин. В меньшей степени избыточная смертность объясняется коронавирусной инфекцией в возрастных группах 35–44 и 85+.

Можно отметить, что женщин умерло от COVID-19 больше, чем мужчин. При этом до 75-летнего возраста перевес был на стороне мужского населения. Итоговое превышение числа умерших женщин над мужчинами произошло благодаря возрастным группам от 75 лет и старше. Такое половозрастное распределение числа умерших от коронавирусной инфекции обусловлено особенностями возрастной структуры населения региона.

Чтобы оценить интенсивность смертности, вызванной COVID-19 в различных половозрастных группах были рассчитаны коэффициенты смертности. На рис. 7 показаны возрастные коэффициенты смертности вследствие коронавирусной инфекции (АМС) в разрезе мужчины/женщины, а также доля АМС в возрастном коэффициенте смертности от всех причин КАМ.

Условные обозначения: АМС — возрастные коэффициенты смертности от коронавирусной инфекции, где А — возраст (age); М — смертность (mortality); С — COVID-19. КАМ — доля АМС в возрастных коэффициентах смертности по всем причинам.

Интенсивность смертности от COVID-19 у мужчин выше, чем у женщин. При увеличении возраста это различие еще больше усиливается. В возрастной группе 80–84 лет с наивысшей интенсивностью смертности у мужчин АМС равен 9,8‰, а у женщин — 5,2‰. В ряде исследований также был выявлен более высокий уровень смертности от коронавирусной инфекции у мужчин по сравнению с женщинами. Например, в работе [1] отмечается, что «Мужчины по сравнению с женщинами потеряли на 45% больше лет жизни. Гендерная разница в потерянных годах потенциальной жизни связана с двумя закономерностями: во-первых, от COVID-19 умирает больше мужчин и, во-вторых, мужчины умирают в более молодом возрасте».

Однако, несмотря на то, что возрастные коэффициенты смертности от COVID-19 у мужчин больше, чем у женщин следует отметить, что пандемия оказала существен-

ное влияние и на смертность женщин. Коронавирусная инфекция вносит существенный вклад в интенсивность смертности в возрастных группах от 45 до 75 лет и составляет порядка 10–12% (см. рис. 7). У мужчин этот вклад в процентном соотношении ниже: 6–7% в возрасте 45 лет и выше.

Общая избыточная смертность в Волгоградской области за 2020 год по сравнению с 2019 составляет, как было отмечено выше, 6647 человек. Данное значение не учитывает динамику возрастных коэффициентов смертности. В годы предшествующие пандемии смертность во всех возрастных группах в целом снижалась (см. рис. 5). Косвенно на снижение смертности указывают также оперативные данные за первые четыре месяца 2020 года: количество зарегистрированных умерших оказалось меньше, чем за аналогичный период предыдущего года (см. таблица 3). В 2020 году у мужчин и женщин старше 45 лет возрастные коэффициенты смертности заметно увеличились (см. рис. 5).

В данной работе была проведена оценка избыточной смертности с учетом динамики пятилетних возрастных коэффициентов смертности во временном интервале с 2015 по

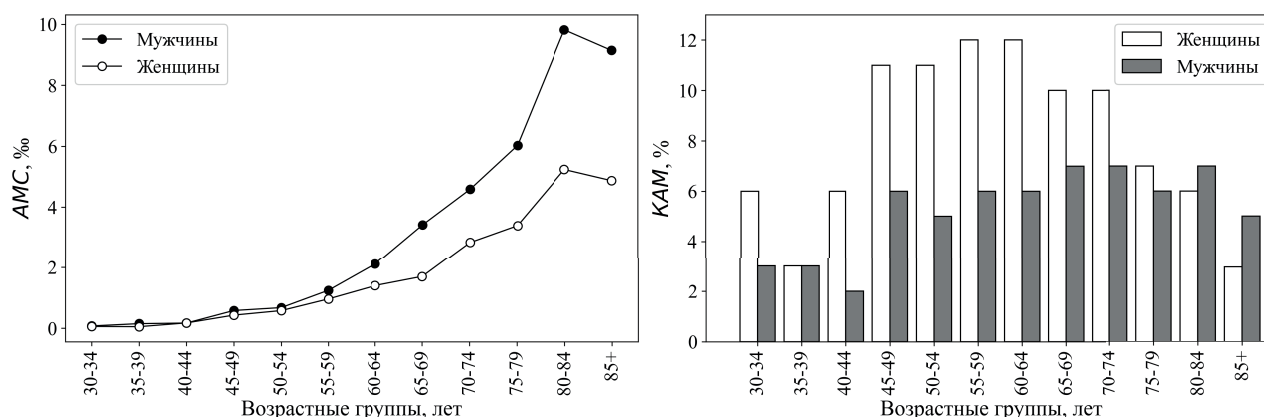


Рис. 7. Возрастные коэффициенты смертности от коронавирусной инфекции за 2020 год по Волгоградской области (АМС), ‰ и доля АМС в возрастных коэффициентах смертности по всем причинам (КАМ), %.

Источник: рассчитано автором на основе данных [10].

Fig. 7. Age-specific mortality rates from coronavirus infection in 2020 for the Volgograd region (АМС: А — age; М — mortality; С — COVID-19), ‰ and the share of АМС in age-specific mortality rates for all causes (КАМ), %.

Source: calculated by the author based on the data [10].

Результаты оценки избыточной смертности с учетом динамики возрастных коэффициентов смертности по Волгоградской области

Results of excess mortality assessment taking into account the dynamics of age-specific mortality rates in the Volgograd region

Возраст, лет	Прогноз числа умерших в 2020 году на основе среднего темпа роста m_x , человек			Разница между прогнозным числом умерших в 2020 году и фактическим числом умерших в 2019, человек		
	мужчины	женщины	оба пола	мужчины	женщины	оба пола
0–34	586	247	833	71	21	92
35–44	1073	364	1437	151	85	236
45–54	1577	628	2205	206	146	352
55–64	3294	1420	4714	806	555	1361
65–74	4154	2814	6968	1095	934	2029
75–84	3494	5561	9055	889	1411	2300
85+	1596	5236	6832	416	582	998
Всего	15774	16270	32044	3634	3734	7368

Источник: составлено автором на основе данных [10]

Source: compiled by the author based on the data [10]

2019 год. В дальнейшем избыточную смертность, оцененную на основе динамики возрастных коэффициентов смертности, будем называть потенциальной избыточной смертностью. В таблице 5 приведены результаты оценки данного показателя.

Итоговые результаты следующие. Если бы тенденции в динамике интенсивности смертности сохранились бы и в 2020 году, то общее количество умерших по Волгоградской области составило бы 32044 человек. Таким образом, избыточную смертность можно оценить в 7368 человек. Доля умерших от COVID-19 в потенциальной избыточной смертности равна 33,2%.

Сравнение возрастного профиля потенциальной избыточной смертности с фактической (таблица 4) показывает, что до 65-летнего возраста потери населения оказались гораздо существенней, чем это следует из фактического распределения избыточной смертности. Данный факт обусловлен тем, что смертность снижалась прежде всего у населения трудоспособного возраста.

На момент подготовки статьи (декабрь 2021) Россия переживает четвертую волну заболеваемости от коронавирусной инфекции. Она характеризуется более высоким уровнем заболеваемости, тяжелым про-

теканием болезни и высокой смертностью. Очевидно, что за 2021 год уровень смертности вследствие пандемии окажется выше, чем в 2020 году. На основе оперативных данных учета смертей, приведенных в таблице 3, можно оценить сколько умрет по Волгоградской области в 2021 году. Последние данные – за сентябрь. Если предположить, что в оставшиеся до конца года три месяца численность умерших будет ежемесячно фиксироваться на уровне сентября, то общее число умерших за 2021 год составит 45933. Это значение существенно выше, чем в 2020 году, в котором было учтено 39410 умерших. В условиях снижения рождаемости и продолжающейся миграционной убыли в Волгоградской области получится демографический кризис сопоставимый с кризисом, который был в 90-е годы прошлого столетия.

Заключение

В результате данного исследования было выявлено, что существенный рост числа умерших в Волгоградской области в период пандемии объясняется Росстатом в качестве причины коронавирусной инфекцией только на 33,2%. Естественным образом возникает вопрос: чем объяснить остальные 2/3 смер-

тей? Ведь до начала пандемии показатели смертности снижались. В 2020 году имеет место значительный рост смертности по таким классам как болезни системы кровообращения и болезни органов дыхания. В ряде исследований отмечается, что умершие от коронавирусной инфекции часто имеют сопутствующие заболевания, связанные с сердечно-сосудистой системой, хроническими болезнями нижних дыхательных путей [19, 20].

Таким образом, возможной причиной анализируемого несоответствия является неправильный учет смертности от коронавирусной инфекции. Другим фактором увеличения смертности в период пандемии может быть снижение качества медицинского обслуживания. Произошла переориентация работы медицинских учреждений на лечение пациентов с коронавирусной инфекцией, увеличилась нагрузка на скорую медицинскую помощь. Наиболее чувствительными к качеству медицинских услуг являются люди преклонного возраста. В таблице 4 приведено половозрастное распределение доли умерших от COVID-19 в избыточной смертности. В возрастной группе 85+ данный показатель имеет наименьшее значение и составляет 22,9%, что в некоторой степени подтверждает сделанное выше предположение.

Литература

1. Семеко Г.В. Демографическое развитие в условиях пандемии COVID-19: вызовы для экономики // Экономические и социальные проблемы России. 2021. № 3. С. 123–140. DOI: 10.31249/espr/2021.03.07.
2. Смирнов А.Ю. Анализ смертности от коронавирусной инфекции в России // Народонаселение. 2021. Т. 24. № 2. С. 76–86. DOI: 10.19181/population.2021.24.2.7.
3. Рязанцев С.В., Иванова А.Е., Архангельский В. Н. Демографическая ситуация в тюменской области и вклад пандемии COVID-19 в ее трансформацию // Человеческий капитал. 2021. № 9(153). С. 81–92. DOI: 10.25629/HC.2021.09.08.
4. Дружинин П.В., Молчанова Е.В., Подлевских Ю. Л. Влияние пандемии COVID-19 на смертность населения российских регионов // Труды Карельского научного центра РАН. 2021. № 7. С. 116–128. DOI: 10.17076/them1421.
5. Nazrul Islam, Dmitri A Jdanov, Vladimir M Shkolnikov, Kamlesh Khunti, Ichiro Kawachi, Martin White, Sarah Lewington, Ben Lacey. Effects of covid-19 pandemic on life expectancy and premature mortality in 2020: time series analysis in 37 countries // BMJ. 2021. DOI: 10.1136/bmj-2021-066768.
6. Lee R.D., Carter L.R. Modeling and Forecasting U. S. Mortality // Journal of the American Statistical Association. Т. 87. № 419. С. 659–671. DOI: 10.2307/2290201.
7. Makarova M.N., Pyshmintseva O.A. Excess mortality in Russian regions during the COVID-19 pandemic // R-economy. 2021. № 7(4). С. 225–234. DOI: 10.15826/recon.2021.7.4.020.
8. Дружинин П.В. Молчанова Е.В. Смертность населения российских регионов в условиях пандемии COVID-19 // Регионология. 2021. Т. 29. № 3. С. 666–685. DOI: 10.15507/2413-1407.116.029.202103.666-685.
9. Кашепов А.В. Ковидный мультипликатор смертности или новый методический подход к анализу избыточной смертности населения в 2020–2021 гг. // Социально-трудовые исследования. 2021. № 44(3). С. 54–64. DOI: 10.34022/2658-3712-2021-44-3-54-64.
10. Российская база данных по рождаемости и смертности // Центр демографических исследований Российской экономической школы [Электрон. ресурс]. Режим доступа: http://demogr.nes.ru/index.php/ru/demogr_indicat/data. (Дата обращения: 01.11.2021).
11. Данные Федеральной службы государственной статистики [Электрон. ресурс]. Режим доступа: <https://rosstat.gov.ru/>. (Дата обращения: 01.11.2021).
12. Оперативные данные. Коронавирус COVID-19 [Электрон. ресурс] // Официальная информация о коронавирусе в России. Режим доступа: <https://стопкоронавирус.рф/information/>. (Дата обращения: 01.11.2021).
13. Алпатов А.В., Рубинштейн Е.Ю. Статистический анализ динамики рождаемости в Волгоградской области // Бизнес. Образование. Право. Вестник Волгоградского института бизнеса. 2016. № 3 (36). С. 41–49.
14. Алпатов А.В., Татаренко А.А. Динамика ожидаемой продолжительности жизни населения Волгоградской области // Бизнес. Образование. Право. Вестник Волгоградского института бизнеса. 2017. № 2 (39). С. 150–155.
15. Тарко А.М. Новые результаты анализа развития России на основе динамики ее демографических и технологических параметров // Пространство и время. 2012. № 4(10). С. 108–118.
16. Немцов А.В. Российская смертность в свете потребления алкоголя // Демографическое обозрение. 2015. Т. 2. № 4. С. 111–135. DOI: 10.17323/demreview.v2i4.1770.
17. Немцов А.В., Симонов А.Н., Фаттахов Т. А., Гридин Р.В. Избыточная смертность в России в праздничные дни // Демографическое обозрение. 2021. Т. 8. № 1. С. 16–43. DOI: 10.17323/demreview.v8i1.12392.
18. Ревич Б.А. Волны жары и смертность населения [Электрон. ресурс] // Демоскоп. 2010. № 439–440. Режим доступа: www.demoscope.ru/weekly/2010/0439/tema04.php. (Дата обращения: 23.11.2021).
19. Удовиченко С.К., Жуков К.В., Никитин Д. Н., Топорков А.В., Викторов Д.В., Зубарева О.В., Климина И.А., Таратутина М.Н. Эпидемические проявления COVID-19 на территории Волгоградской области: промежуточные итоги // Вестник ВолГМУ. 2020. № 4(76). С. 30–36. DOI: 10.19163/1994-9480-2020-4(76)-30-36.
20. Сабгайда Т.П., Зубко А.В., Семенова В.Г. Изменение структуры причин смерти во второй год пандемии COVID-19 в Москве [Электрон. ресурс] // Социальные аспекты здоровья населения. 2021. № 67(4). С. 1. Режим доступа: <http://vestnik.mednet.ru/content/view/1280/30/lang,ru/>.

References

1. Semeko G.V. Demographic development in the context of the COVID-19 pandemic: challenges for the economy. *Ekonomicheskiye i sotsial'nyye problemy Rossii* = Economic and social problems of Russia. 2021; 3: 123–140. DOI: 10.31249/espr/2021.03.07. (In Russ.)
2. Smirnov A.Yu. Analysis of mortality from coronavirus infection in Russia. *Narodonaseleniye* = Population. 2021; 24; 2: 76–86. DOI: 10.19181/population.2021.24.2.7. (In Russ.)
3. Ryazantsev S.V., Ivanova A.Ye., Arkhangel'skiy V.N. The demographic situation in the Tyumen region and the contribution of the COVID-19 pandemic to its transformation. *Chelovecheskiy kapital* = Human capital. 2021; 9(153): 81–92. DOI: 10.25629/HC.2021.09.08. (In Russ.)
4. Druzhinin P.V., Molchanova Ye.V., Podlevskikh Yu.L. The impact of the COVID-19 pandemic on the mortality of the population of Russian regions. *Trudy Karel'skogo nauchnogo tsentra RAN* = Proceedings of the Karelian Scientific

Center of the Russian Academy of Sciences. 2021; 7: 116–128. DOI: 10.17076/them1421. (In Russ.)

5. Nazrul Islam, Dmitri A Jdanov, Vladimir M Shkolnikov, Kamlesh Khunti, Ichiro Kawachi, Martin White, Sarah Lewington, Ben Lacey. Effects of covid-19 pandemic on life expectancy and premature mortality in 2020: time series analysis in 37 countries. *BMJ*. 2021. DOI: 10.1136/bmj-2021-066768.

6. Lee R.D., Carter L.R. Modeling and Forecasting U. S. Mortality. *Journal of the American Statistical Association*. 87; 419: 659–671. DOI: 10.2307/2290201.

7. Makarova M.N., Pyshmintseva O.A. Excess mortality in Russian regions during the COVID-19 pandemic. *R-economy*. 2021; 7(4): 225–234. DOI: 10.15826/recon.2021.7.4.020.

8. Druzhinin P. V., Molchanova Ye. V. Mortality of the population of Russian regions in the context of the COVID-19 pandemic. *Regionologiya = Regionology*. 2021; 29; 3: 666–685. DOI: 10.15507/2413-1407.116.029.202103.666-685. (In Russ.)

9. Kashepov A.V. Covid mortality multiplier or a new methodological approach to the analysis of excess mortality in the population in 2020–2021. *Sotsial'no-trudovyye issledovaniya = Social and labor research*. 2021; 44(3): 54–64. DOI: 10.34022/2658-3712-2021-44-3-54-64. (In Russ.)

10. Russian database on fertility and mortality. Tsentr demograficheskikh issledovaniy Rossiyskoy ekonomicheskoy shkoly = Center for Demographic Research of the Russian Economic School [Internet]. Available from: http://demogr.nes.ru/index.php/ru/demogr_indicat/data. (cited 01.11.2021). (In Russ.)

11. Dannyye Federal'noy sluzhby gosudarstvennoy statistiki = Data of the Federal State Statistics Service [Internet]. Available from: <https://rosstat.gov.ru/>. (cited 01.11.2021). (In Russ.)

12. Operational data. Coronavirus COVID-19 [Internet]. Ofitsial'naya informatsiya o koronavirusе v Rossii = Official information about the coronavirus in Russia. Available from: <https://stopkoronavirus.rf/information/>. (cited 01.11.2021). (In Russ.)

13. Alpatov A.V., Rubinshteyn Ye.Yu. Statistical analysis of the dynamics of fertility in the Volgograd

region. *Biznes. Obrazovaniye. Pravo. Vestnik Volgogradskogo instituta biznesa = Business. Education. Right. Bulletin of the Volgograd Institute of Business*. 2016; 3 (36): 41–49. (In Russ.)

14. Alpatov A.V., Tatarenko A.A. Dynamics of life expectancy of the population of the Volgograd region. *Biznes. Obrazovaniye. Pravo. Vestnik Volgogradskogo instituta biznesa = Business. Education. Right. Bulletin of the Volgograd Institute of Business*. 2017; 2 (39): 150–155. (In Russ.)

15. Tarko A.M. New results of the analysis of the development of Russia based on the dynamics of its demographic and technological parameters. *Prostranstvo i vremya = Space and Time*. 2012; 4(10): 108–118. (In Russ.)

16. Nemtsov A.V. Russian mortality in the light of alcohol consumption. *Demograficheskoye obozreniye = Demographic Review*. 2015; 2; 4: 111–135. DOI: 10.17323/demreview.v2i4.1770. (In Russ.)

17. Nemtsov A.V., Simonov A.N., Fattakhov T.A., Gridin R.V. Excessive mortality in Russia during the holidays. *Demograficheskoye obozreniye = Demographic Review*. 2021; 8; 1: 16–43. DOI: 10.17323/demreview.v8i1.12392. (In Russ.)

18. Revich B.A. Heat waves and population mortality [Elektron. resurs]. *Demoskop*. 2010: 439–440. Available from: www.demoscope.ru/weekly/2010/0439/tema04.php. (cited 23.11.2021). (In Russ.)

19. Udovichenko S.K., Zhukov K.V., Nikitin D.N., Toporkov A.V., Viktorov D.V., Zubareva O.V., Klimina I.A., Taratutina M.N. Epidemic manifestations of COVID-19 on the territory of the Volgograd region: intermediate results. *Vestnik Volgogradskogo gosudarstvennogo meditsinskogo universiteta = Bulletin of the Volgograd State Medical University*. 2020; 4(76): 30–36. DOI 10.19163/1994-9480-2020-4(76)-30-36. (In Russ.)

20. Sabgayda T.P., Zubko A.V., Semenova V.G., Change in the structure of causes of death in the second year of the COVID-19 pandemic in Moscow [Internet]. *Sotsial'nyye aspekty zdorov'ya naseleniya = Social aspects of public health*. 2021; 67(4): 1. Available from: <http://vestnik.mednet.ru/content/view/1280/30/lang,ru/>. (In Russ.)

Сведения об авторе

Алексей Викторович Алпатов

К.ф.-м.н., доцент кафедры «Информатика и технология программирования»

Волжский политехнический институт (филиал), Волжский, Россия

Эл. почта: alpatov80@mail.ru

Information about the author

Alexey V. Alpatov

Cand. Sci. (Physics and Mathematics), Associate Professor of the Department of «Computer science and programming technology»

Volgograd State Technical University, branch of the Volga Polytechnical Institute, Volzhskiy, Russia

E-mail: alpatov80@mail.ru

Мировые тенденции дифференциации доходов населения

Цель исследования. Выявление мировых тенденций неравенства в распределении доходов населения. В соответствии с целью, поставлены следующие задачи: 1) изучить современные международные исследования, в которых рассматривается проблема неравенства распределения доходов населения; 2) оценить дифференциацию доходов населения на глобальном и региональном уровнях; 3) на основе индексов Джини и Тейла проанализировать динамику неравенства населения по доходам внутри и между странами мира.

Материалы и методы. В процессе подготовки статьи автором использованы данные международных докладов и отчетов, аналитические статистические материалы, научные труды российских и зарубежных ученых. В работе были использованы научные методы познания: анализ (для оценки изменения показателей доходного неравенства населения), синтез (для определения взаимосвязи межстранового и внутристранового неравенства населения по доходам), графический (для построения графиков, отражающих динамику изменения распределения национального дохода и активов среди населения, коэффициента Джини, индекса Тейла). Указанные методы позволили выявить масштабы и тенденции дифференциации доходов населения в мире.

Результаты. Была исследована проблема неравномерности распределения доходов населения. Установлено, что доходное

неравенство существенно отличается между регионами мира, а уровень неравенства населения по доходам внутри стран значительно выше, чем уровень неравенства между странами. Дана оценка текущего состояния и выявлены тенденции дифференциации доходов населения в мире на основе индекса Джини и индекса Тейла.

Заключение. Установлено, что проблема дифференциации доходов населения находится в центре внимания как научного сообщества, так и международных организаций, а именно: Организации Объединенных наций, Организации экономического сотрудничества и развития, Всемирного банка, Оксфам. Уровень дифференциации населения по доходам между регионами мира существенно отличается. Масштабы глобального доходного неравенства населения на данный момент достигли уровня, который наблюдался при расцвете западного империализма. С помощью индексов Джини и Тейла было выявлено, что внутристрановое неравенство существенно больше, чем межстрановое неравенство населения по доходам.

Ключевые слова: доходы населения; неравенство в распределении доходов; дифференциация доходов; мировая экономика; показатели доходного неравенства.

Daniya A. Huchmazova

Donetsk National University, Donetsk, Donetsk People's Republic

Global Trends in Differentiation of Population Income

Purpose of the study. Identification of global trends in inequality in the distribution of the population income. In accordance with the goal, the following tasks are set: 1) to examine current international research that addresses the problem of income distribution inequality of the population; 2) assess the differentiation of the population income at the global and regional levels; 3) on the basis of the Gini and Theil indexes, to analyze the dynamics of income inequality of the population within and between countries of the world.

Materials and methods. In the process of preparing the article, the author used data from international reports, analytical statistical materials, scientific works of Russian and foreign scientists. The scientific methods of cognition were used in the work: analysis (to assess changes in indicators of income inequality of the population), synthesis (to determine the relationship between inter-country and intra-country income inequalities of the population), graphical (to build graphs that reflect the dynamics of changes in the distribution of national income and assets among the population, Gini coefficient, Theil index). These methods made it possible to identify the scale and trends in the differentiation of the population income in the world.

Results. The problem of uneven distribution of the population income was investigated. It has been established that inequality in the

population income differs significantly between regions of the world, and the level of inequality of the population in terms of income within countries is much higher than the level of inequality between countries. An assessment of the current state is given and trends in the differentiation of the population income in the world based on the Gini index and the Theil index are revealed.

Conclusion. It has been established that the problem of income differentiation of the population is in the focus of attention of both the scientific community and international organizations, namely: the United Nations, the Organization for Economic Cooperation and Development, the World Bank, Oxfam. The level of differentiation of the population by income between regions of the world differs significantly. The scale of global income inequality of the population now has reached the level that was observed during the heyday of Western imperialism. With the help of the Gini and Theil indexes, it was revealed that intra-country inequality is significantly greater than the inter-country income inequality of the population.

Keywords: income of the population; inequality in income distribution; income differentiation; world economy; indicators of income inequality.

Введение

Дифференциация доходов населения представляет собой одну из ключевых проблем распределительных отношений. Эксперты Лаборатории мирового неравенства полагают, что прозрачная, надёжная и открытая информация о неравенстве доходов населения является глобальным общественным благом [1]. Следовательно, объективная оценка таких статистических данных имеет принципиально важное значение для регулирования дифференциации доходов населения на различных уровнях хозяйствования.

В последние годы в рамках научного сообщества вышло несколько крупных работ, посвященных проблеме неравенства в распределении доходов: Т. Пикетти «Капитал в XXI веке» (2013 г.), Дж. Стиглиц «Цена неравенства. Чем расхождение общества угрожает нашему будущему» (2015 г.), Б. Миланович «Глобальное неравенство: новый подход для эпохи глобализации» (2017 г.), Э. Аткинсон «Неравенство: как с ним быть?» (2018 г.).

Французский экономист Т. Пикетти на основе анализа статистических данных о росте концентрации богатства в Европе и США, начиная с XVIII в., пришел к выводу, что в настоящее время уместно говорить о «потомственном капитализме», когда богатство концентрируется в одних руках и в основном передается по наследству, а не приобретает трудом или заслугами [2].

Лауреат Нобелевской премии Дж. Стиглиц в работе «Цена неравенства. Чем расхождение общества угрожает нашему будущему» акцентировал внимание на угрозах, которые несет чрезмерное неравенство в доходах. Он подчеркнул, что «неравенство достигло той стадии, на которой оно перестало быть эффективным и превратилось в серьезную помеху для экономического развития» [3, с. 13].

Сербско-американский экономист Б. Миланович в работе

«Глобальное неравенство: новый подход для эпохи глобализации» отмечает, что в начале XXI в. глобальное неравенство в доходах между людьми не только увеличилось в масштабах, по сравнению с показателями более чем 150-летней давности, но и полностью изменилась его структура: из неравенства, примерно в равной мере определяемого различиями в классовой принадлежности и стране проживания, оно превратилось в неравенство, определяемое исключительно страной проживания [4].

В работе Э. Аткинсона «Неравенство: как с ним быть?» основное внимание было уделено решению задачи по уменьшению неравенства в распределении доходов, которое в настоящее время приобрело невероятные масштабы [5]. Основными факторами увеличения доходного неравенства, с точки зрения ученого, являются: технологические изменения, увеличение сектора финансовых услуг, глобализация, изменения в нормах оплаты труда, ослабление роли профсоюзов, сворачивание политики перераспределения доходов. Несмотря на достаточно большой объем научных исследований по данной проблеме, системная оценка масштабов и тенденций дифференциации доходов населения в мире отсутствует.

В данной работе предлагается выявить мировые тенденции неравенства в распределении доходов населения.

Основная часть

Проблема неравномерности распределения доходов находится в центре внимания не только научного сообщества, но и международных организаций. Ряд международных научно-статистических проектов посвящен проблеме неравенства доходов в мире: Мировая база данных неравномерности распределения доходов ООН (United Nations World Income Inequality Database); База дан-

формируемая ОЭСР (OECD Income Distribution Database); статистика Всемирного банка (World Bank PovcalNet) о населении, которое живет в нищете; Лаборатория мирового неравенства (World Inequality Lab, WIL).

Вышеперечисленные международные организации выпустили серию докладов, посвященных проблеме мирового неравенства — его масштабам, различиям между странами, тенденциям, негативным последствиям растущего неравенства и разработке рекомендаций по управлению им. Среди них: доклад неправительственной некоммерческой организации Oxfam International «Экономика для 1 %. Как привилегии и власть в экономике приводят к радикальному неравенству и как это можно остановить»; доклад Всемирного банка «Бедность и всеобщее процветание: борьба с неравенством»; ежегодные доклады Credit Suisse о распределении мирового богатства; Доклады о мировом неравенстве (World Inequality Report 2018; World Inequality Report 2022), выпущенные Лабораторией мирового неравенства.

Итогом работы 109-й июньской сессии Международной конференции труда от 2021 г. стал IV доклад «Неравенство и сфера труда» [6]. На основе представленного в докладе глобального обзора данных и результатов исследований можно отметить следующее:

а) уровень неравенства доходов различен внутри стран и достигает максимума в некоторых развивающихся странах [7];

б) Латинская Америка и Африка являются двумя регионами мира с самым высоким уровнем доходного неравенства [8; 9; 10];

в) с 1980-х годов доходное неравенство увеличилось в большинстве стран мира, особенно в странах с высоким уровнем дохода [11]. Неравенство в доходах за 1990–2015 гг. увеличилось в 77 из 140 стран [12];

г) в большинстве стран с высоким уровнем доходов тен-

денция к повышению уровня доходного неравенства сопровождалась снижением доли среднего класса [13].

Одним из крупнейших проектов, работающим над проблемой неравенства в мире, является Лаборатория мирового неравенства. Проект объединяет свыше 100 исследователей из более 70 стран, которые формируют Мировую базу данных о богатстве и доходах.

В «Докладе о неравенстве в мире 2022» (World Inequality Report 2022) Лабораторией мирового неравенства дан обобщённый обзор международной исследовательской деятельности по отслеживанию глобального доходного неравенства населения [14]. Данные свидетельствуют о том, что:

1) *В настоящее время существует огромное неравенство в уровнях дохода и богатства в мире.* В исследовании население мира разделено на три группы по уровню доходов (рис. 1): первая – население с низкими доходами (50%); вторая – население со средними доходами (40%); третья – население с высокими доходами (10%).

В данной схеме распределения мирового дохода в 2021 г. один человек из группы населения с высокими доходами получает в среднем 122 100 долл. США (или 87 200 евро) в год, а представитель группы населения с низкими доходами – 3 920 долл. США (или 2 800 евро) в год. Годовой доход среднестатистического взрослого гражданина (рассчитанный по паритету покупательной способности) составил 23 380 долл. США (или 16 700 евро), а средний размер сбережений достиг 102 600 долл. США (72 900 евро).

За этими средними показателями скрывается доходное неравенство как между странами, так и внутри страны (рис. 2).

В 2021 г. группа населения мира с низкими доходами (50%, или около 2,5 млрд чел.) получала всего 8,5% мирового дохода и владела 2% активов, а группа населения мира с высокими доходами (10%, или 517 млн чел.) получила 52% всех



Рис. 1. Группы населения мира по уровню доходов

Fig. 1. Groups of the world's population by income level



Рис. 2. Глобальное неравенство в уровнях дохода и активов населения в 2021 г. [14]

Fig. 2. Global income and asset inequality in 2021 [14]

доходов и владела 76% всех активов (по паритету покупательской способности). Необходимо отметить, что владельцы самых больших активов не всегда выступают получателями максимальных доходов.

Глобальное неравенство в благосостоянии (имущественное неравенство) проявляется ещё сильнее, чем доходное неравенство. В 2021 г. на группу населения мира с низкими доходами (50%) приходилось только 2% активов, а на группу населения мира с высокими доходами (10%) – 76% активов в мире. Так, в среднем взрослый человек из группы насе-

ния с низкими доходами (50%) владеет 4 100 долл. США (или 2 900 евро), а представитель, входящий в группу населения с высокими доходами (10%), владеет 771 300 долл. США (или 550 900 евро).

2) *Ближний Восток и Северная Африка – регион мира, который характеризуется наиболее высоким показателем доходного неравенства населения, а в Европе данные показатели самые низкие.* На рис. 3 показаны уровни неравенства в доходах населения между регионами мира в 2021 г.

Уровень доходного неравенства населения значитель-

но отличается между регионами мира. Так, в Европе доходы группы населения мира с низкими доходами (50%) и высокими доходами (10%) составляют, соответственно, 19% и 36% национального дохода. В то время как в Ближнем Востоке и Северной Африке, на группу населения мира с низкими доходами (50%) и группу населения мира с высокими доходами (10%) приходится, соответственно, 9% и 58% национального дохода.

В других регионах мира модели распределения доходов среди населения различаются. В Восточной Азии, Северной Азии, России и Центральной Азии на группу населения мира с низкими доходами (50%) приходится, соответственно, 14%, 13% и 15%, а на группу населения мира с высокими доходами (10%), соответственно, 43%, 46% и 47% всего национального дохода в 2021 г.

В Южной и Юго-Восточной Азии, Латинской Америке и Странах Африки к югу от Сахары на группу населения мира с низкими доходами (50%) приходится, соответственно, 12%, 10% и 9%, а на группу населения мира с высокими доходами (10%), соответственно, 55%, 55% и 56% всего национального дохода в 2021 г.

3) *Показатель национального дохода на душу населения не объективно отражает реальную ситуацию с уровнем неравенства населения в доходах внутри самой страны.* Среди стран с высоким национальным доходом на душу населения некоторые страны отличаются крайне высоким уровнем неравенства населения по доходам (например, США), в то время как в других (например, в Швеции) доходы населения распределяются сравнительно равномерно. Аналогичная ситуация характерна для стран с низким и средним уровнем национального дохода на душу населения: для одних стран характерно крайнее доходное неравенство (например, в Бразилии и Индии), для вторых — довольно высокий уровень

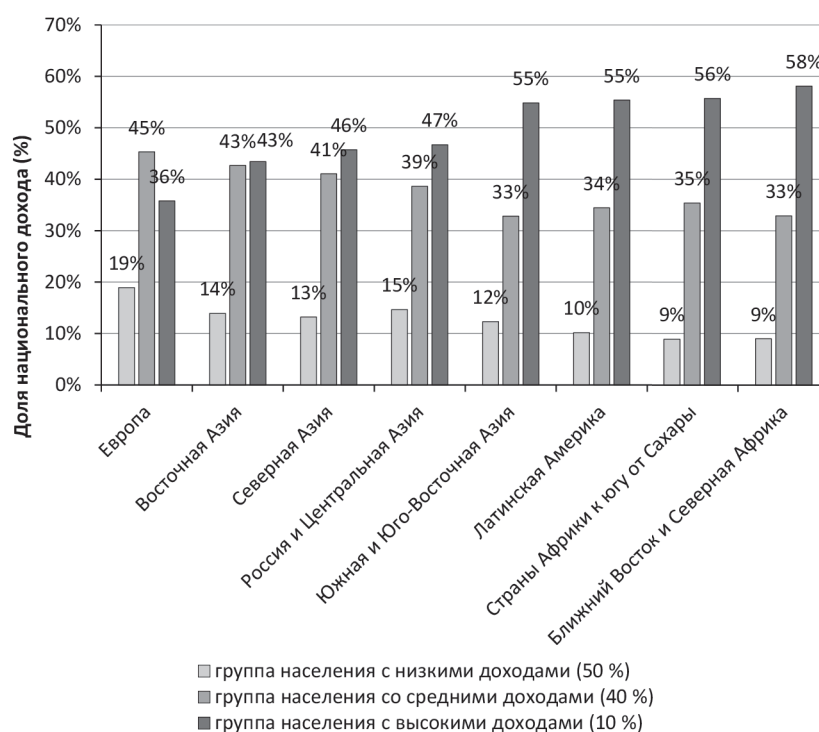


Рис. 3. Распределение национального дохода среди населения по регионам мира в 2021 г. [14]

Fig. 3. Distribution of national income among the population by regions of the world in 2021 [14]



Рис. 4. Глобальное неравенство в уровнях дохода населения за 1820–2020 гг.: отношение среднего дохода 10 % наиболее обеспеченного населения к среднему доходу 50 % наименее обеспеченного населения [14]

Fig. 4. Global income inequality of the population for 1820–2020: ratio of the average income of the top 10% of the population to the average income of the bottom 50% of the population [14]

доходного неравенства (например, в Китае), а для третьих — умеренный и сравнительно низкий уровни неравенства населения по доходам (например, в Малайзии, Уругвае). Так, в Бразилии и во Франции группа населения мира с низ-

кими доходами (50%) получает доходы, соответственно, в 29 и 7 раз меньше, чем группа населения мира с высокими доходами (10%) [14].

4) *Современные уровни доходного неравенства населения приближаются к уровням XX-го века*

на пике западного империализма. При том, что за последние двадцать лет во многих странах мира уровень неравенства населения по доходам возрос, глобальное неравенство между странами снизилось. Так, разрыв между средним доходом группы населения мира с высокими доходами (10%) и средним доходом группы населения мира с низкими доходами (50%) уменьшился: приблизительно 50-кратный разрыв сократился до менее чем 40-кратного разрыва (рис. 4).

Глобальное доходное неравенство, измеренное на основе отношения среднего дохода группы населения мира с высокими доходами (10%) к среднему доходу группы населения с низкими доходами (50%), возросло более чем в два раза в период между 1820 г. и 1910 г. (с менее чем 18 до около 41), а в период между 1910 г. и 2020 г. стабилизировалось на уровне 40 раз. Разрыв достиг рекордно высокого уровня в 1980 г. (в 53 раза) и в 2000 г. (в 50 раз). В мире в 2020 г. средний доход группы населения мира с высокими доходами (10 %) в 38 раз превышал средний доход группы населения с низкими доходами (50%). Глобальное неравенство населения по доходам, измеряемое индексом Джини, выросло примерно с 0,6 в 1820 г. до 0,72 в 1910 г., а затем стабилизировалось на уровне около 0,7 в период с 1910 г. по 2020 г. (рис. 5)

С 2008 г. наблюдается тенденция уменьшения неравенства населения по доходам в мире (что в основном обусловлено финансовым кризисом 2008 г.). Поэтому, как утверждают эксперты Лаборатории мирового неравенства, говорить о сохранении данной тенденции преждевременно [14]. При этом разрыв между средним доходом группы населения мира с высокими доходами (10%) и средним доходом группы населения с низкими доходами (50%) внутри стран практически удвоился: с 8,5 до 15 раз [15]. Это означает, что уровень неравенства насе-

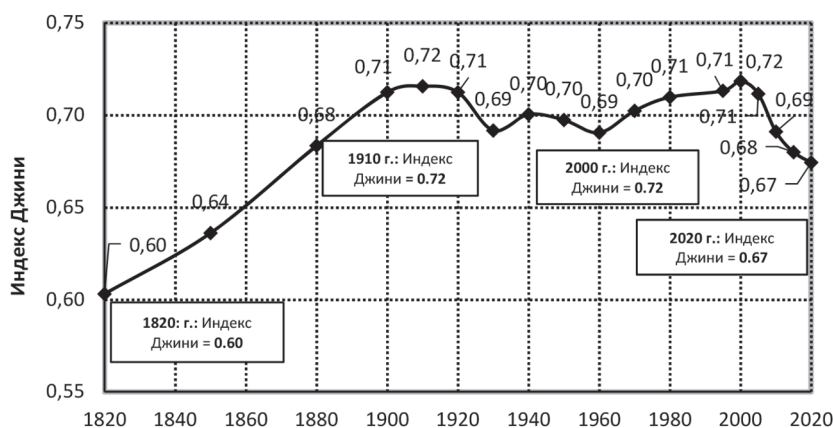


Рис. 5. Динамика неравенства населения по доходам в мире за 1820–2020 гг.: индекс Джини [14]

Fig. 5. Dynamics of income inequality in the world for 1820–2020: Gini index [14]

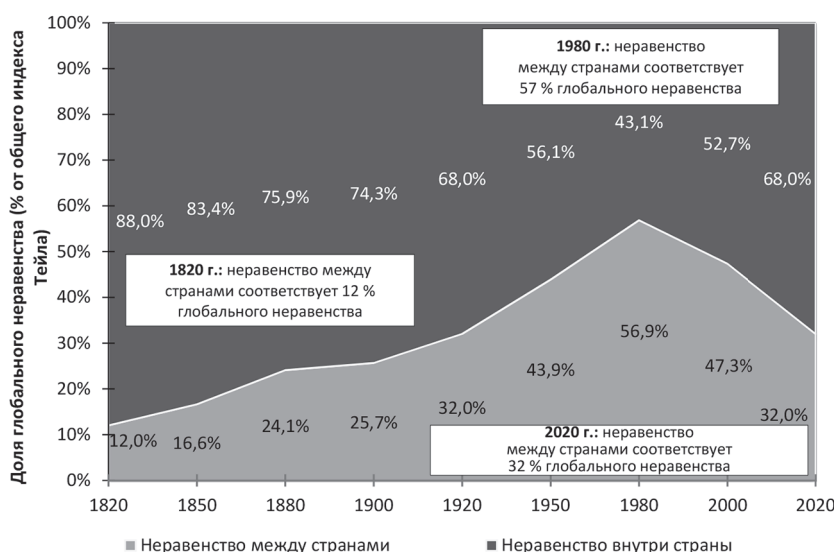


Рис. 6. Глобальное неравенство в уровнях дохода: неравенство между странами по сравнению с неравенством внутри стран (индекс Тейла) за 1820–2020 гг. [14]

Fig. 6. Global income inequality: inequality between countries versus inequality within countries (Theil index) for 1820–2020 [14]

ния по доходам внутри стран в настоящее время значительнее, чем уровень неравенства между странами (рис. 6).

Значение доходного неравенства между странами в общем глобальном неравенстве, измеренное с помощью индекса Тейла, выросло в период между 1820 г. (12%) и 1980 г. (57%) а затем значительно снизилось. В 2020 г. неравенство по доходам между странами соответствует примерно 1/3 глобального неравенства между отдельными членами общества. Остальная часть соответствует внутристрано-

вому доходному неравенству. Доля мирового дохода, получаемая группой населения мира с высокими доходами (10%), в период между 1820 г. и 2020 г. колебалась от 50% до 60% (в 1820 г. — 50%; в 1910 г. — 60%; в 1980 г. — 56%; в 2000 г. — 61%; в 2020 г. — 55%). А доля мирового дохода, получаемая группой населения мира с низкими доходами (50%), составляла примерно 10% и ниже (в 1820 г. — 14%; в 1910 г. — 7%; в 1980 г. — 5%; в 2000 г. — 6%; в 2020 г. — 7%). Глобальное неравенство населения по доходам всегда нахо-

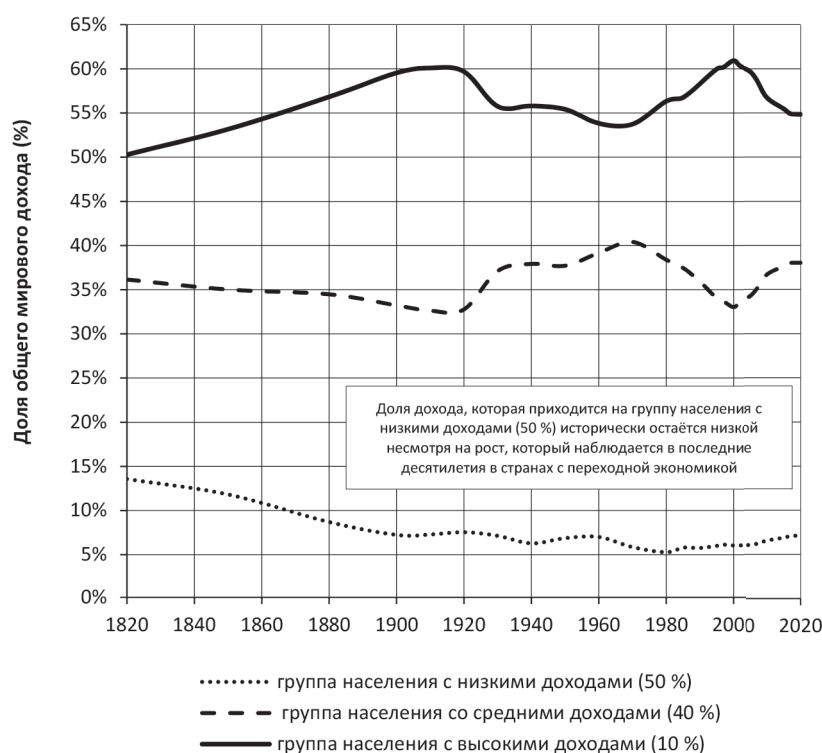


Рис. 7. Глобальное неравенство населения в уровнях дохода за 1820–2020 гг. [14]

Fig. 7. Global population inequality in income levels for 1820-2020 [14]

дилось на высоком уровне.

Необходимо отметить, что в 2020 г. доля мирового дохода, получаемая группой населения с низкими доходами (50%), практически в 8 раз меньше той, которую получила группа населения с высокими доходами (10%) (соответственно, 7% и 55%).

Вышеизложенное подтверждает, что глобальные

уровни неравенства населения по доходам практически достигли такого уровня, который наблюдался при расцвете западного империализма в н. XX-го века (рис. 7).

Заключение

Проведенное исследование позволяет сделать следующие выводы:

1. Выявлено, что в последние годы на международном уровне возрос общественный интерес к проблеме неравенства в распределении доходов. Так, множество международных научно-статистических проектов и докладов (ООН, ОЭСР, ВБ, Оксфам и др.) были посвящены проблеме неравенства населения в доходах;

2. На основе оценки неравенства в распределении доходов на международном уровне было определено, что: уровни неравенства в доходах населения между регионами мира существенно отличаются; уровень доходного неравенства населения внутри стран значительно выше, чем уровень неравенства между странами; доля доходов, приходящаяся на группу населения с низкими доходами (50 %) исторически остаётся низкой;

3. Установлено, что, несмотря на достижения многих стран мира в экономической сфере, масштабы дифференциации доходов населения на данный момент представляют серьезную общественную проблему. Так, в 2020 г. средний доход группы населения с высокими доходами (10%) в 38 раз превышал средний доход группы населения мира с низкими доходами (50%).

Литература

1. World inequality report 2022 [Электрон. ресурс]. Режим доступа: https://wir2022.wid.world/www-site/uploads/2021/12/WorldInequalityReport2022_Full_Report.pdf (Дата обращения: 14.01.2022).
2. Пикетти Т. Капитал в XXI веке / Пер. с фр. А. А. Дунаев, науч. ред. пер. А. Ю. Володин. М.: Ад Маргинем Пресс, 2015. 592 с.
3. Стиглиц Дж. Цена неравенства. Чем расхождение общества грозит нашему будущему. М.: Эксмо, 2015. 511 с.
4. Миланович Б. Глобальное неравенство. Новый подход для эпохи глобализации. пер. с англ. Д. Шестакова. М.: Институт Гайдара, 2017. 336 с.
5. Аткинсон Э.Б. Неравенство: как с ним быть? пер. с англ. О. Левченко; науч. ред. М. Добряковой. М.: Дело, 2018. 531 с.
6. Доклад IV Неравенство и сфера труда [Электрон. ресурс]. Режим доступа: https://www.ilo.org/wcmsp5/groups/public/---ed_norm/---relconf/documents/meetingdocument/wcms_792165.pdf.

org/wcmsp5/groups/public/---ed_norm/---relconf/documents/meetingdocument/wcms_792165.pdf. (Дата обращения: 14.01.2022).

7. Balestra C. Inequalities in Emerging Economies: Informing the Policy Dialogue on Inclusive Growth [Электрон. ресурс]. Режим доступа: [https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/DOC\(2018\)13&docLanguage=En](https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/DOC(2018)13&docLanguage=En). (Дата обращения: 14.01.2022).

8. Messina J., Silva J. Wage Inequality in Latin America: Understanding the Past to Prepare for the Future [Электрон. ресурс]. Режим доступа: <https://openknowledge.worldbank.org/handle/10986/28682>. (Дата обращения: 14.01.2022).

9. World Bank. Poverty and Shared Prosperity 2016: Taking on Inequality [Электрон. ресурс]. Режим доступа: <https://www.worldbank.org/en/publication/poverty-and-shared-prosperity-2016>. (Дата обращения: 14.01.2022).

10. Solt F. The Standardized World Income Inequality Database [Электрон. ресурс]. Режим доступа: https://www.researchgate.net/publication/46541564_The_Standardized_World_Income_Inequality_Database. (Дата обращения: 14.01.2022).

11. World Bank. Poverty and Shared Prosperity 2016: Taking on Inequality [Электрон. ресурс]. Режим доступа: <https://www.worldbank.org/en/publication/poverty-and-shared-prosperity-2016>. (Дата обращения: 14.01.2022).

12. World Inequality Report 2018 [Электрон. ресурс]. Режим доступа: <https://wir2018.wid.world/files/download/wir2018-full-report-english.pdf>. (Дата обращения: 14.01.2022).

13. Under Pressure: The Squeezed Middle Class [Электрон. ресурс]. Режим доступа: https://read.oecd-ilibrary.org/social-issues-migration-health/under-pressure-the-squeezed-middle-class_689afed1-en#page4. (Дата обращения: 14.01.2022).

14. World inequality report 2022 [Электрон. ресурс]. Режим доступа: https://wir2022.wid.world/www-site/uploads/2021/12/WorldInequalityReport2022_Full_Report.pdf. (Дата обращения: 14.01.2022).

15. Доклад о неравенстве в мире 2022. Краткое содержание [Электрон. ресурс]. Режим доступа: https://wir2022.wid.world/www-site/uploads/2021/12/Summary_WorldInequalityReport2022_Russian.pdf. (Дата обращения: 14.01.2022).

References

1. World inequality report 2022 [Internet]. Available from: https://wir2022.wid.world/www-site/uploads/2021/12/WorldInequalityReport2022_Full_Report.pdf. (cited 14.01.2022).

2. Piketti T. Kapital v XXI veke = Capital in the XXI century / Tr. from Eng. A. A. Dunayev, ed. A.YU.Volodin. Moscow: Ad Marginem Press; 2015. 592 p. (In Russ.)

3. Stiglits Dzh. Tsena neravenstva. Chem rassloyeniye obshchestva grozit nashemu budushchemu = The price of inequality. What stratification of society threatens our future. Moscow: Eksmo; 2015. 511 p. (In Russ.)

4. Milanovich B. Global'noye neravenstvo. Novyy podkhod dlya epokhi globalizatsii = Global inequality. A new approach for the era of globalization. Tr. from Eng. D. Shestakova. Moscow: Gaidar Institute; 2017. 336 p. (In Russ.)

5. Atkinson E.B. Neravenstvo: kak s nim byt'? = Inequality: how to deal with it? Tr. from Eng. O. Levchenko; Ed. M. Dobryakovoy. Moscow: Delo; 2018. 531 p. (In Russ.)

6. Doklad IV Neravenstvo i sfera truda = Report IV Inequality and the world of work [Internet]. Available from: https://www.ilo.org/wcmsp5/groups/public/---ed_norm/---relconf/documents/meeting-document/wcms_792165.pdf. (cited 14.01.2022).

7. Balestra C. Inequalities in Emerging Economies: Informing the Policy Dialogue on Inclusive Growth [Internet]. Available from: [https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/DOC\(2018\)13&docLanguage=En](https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/DOC(2018)13&docLanguage=En). (cited 14.01.2022).

8. Messina J., Silva J. Wage Inequality in Latin America: Understanding the Past to Prepare for

the Future [Internet]. Available from: <https://openknowledge.worldbank.org/handle/10986/28682>. (cited 14.01.2022).

9. World Bank. Poverty and Shared Prosperity 2016: Taking on Inequality [Internet]. Available from: <https://www.worldbank.org/en/publication/poverty-and-shared-prosperity-2016>. (cited 14.01.2022).

10. Solt F. The Standardized World Income Inequality Database [Internet]. Available from: https://www.researchgate.net/publication/46541564_The_Standardized_World_Income_Inequality_Database. (cited 14.01.2022).

11. World Bank. Poverty and Shared Prosperity 2016: Taking on Inequality [Internet]. Available from: <https://www.worldbank.org/en/publication/poverty-and-shared-prosperity-2016>. (cited 14.01.2022).

12. World Inequality Report 2018 [Internet]. Available from: <https://wir2018.wid.world/files/download/wir2018-full-report-english.pdf>. (cited 14.01.2022).

13. Under Pressure: The Squeezed Middle Class [Internet]. Available from: https://read.oecd-ilibrary.org/social-issues-migration-health/under-pressure-the-squeezed-middle-class_689afed1-en#page4. (cited 14.01.2022).

14. World inequality report 2022 [Internet]. Available from: https://wir2022.wid.world/www-site/uploads/2021/12/WorldInequalityReport2022_Full_Report.pdf. (cited 14.01.2022).

15. Doklad o neravenstve v mire 2022. Kratkoye soderzhaniye = World Inequality Report 2022. Summary [Internet]. Available from: https://wir2022.wid.world/www-site/uploads/2021/12/Summary_WorldInequalityReport2022_Russian.pdf. (cited 14.01.2022).

Сведения об авторе

Дания Анасовна Гучмазова

Аспирант кафедры коммерции и таможенного дела

Донецкий национальный университет,
Донецк, Донецкая Народная Республика
Эл. почта: d.guchmazova@donnu.ru

Information about the author

Daniya A. Huchmazova

Postgraduate student of the Department of Commerce and Customs

Donetsk National University,
Donetsk, Donetsk People's Republic
E-mail: d.guchmazova@donnu.ru

Оценка доли мужчин в численности педагогических работников в школах

Цель исследования. Одной из актуальных проблем развития регионов России является повышение эффективности региональных систем общего среднего образования. При этом необходимо проведение мероприятий по сокращению гендерного разрыва в кадровом персонале современных школ. Для понимания сложившегося в регионах гендерного разрыва компаративный анализ различий между регионами целесообразно проводить на основе удельных показателей. При этом необходимо учитывать имеющуюся дифференциацию доли педагогических работников мужского пола по возрасту. Целью нашего исследования была оценка показателей, характеризующих долю учителей-мужчин разных возрастных групп в общей численности педагогических работников школ в регионах России. Этими показателями были доли числа мужчин пяти возрастных групп в общем числе школьных учителей в этих возрастных категориях в 2020 году.

Материалы и методы. В исследовании использовался предложенный автором методический подход, базирующийся на рассмотрении удельных показателей, описывающих долю мужчин-учителей разных возрастов, работающих в общеобразовательных школах, в общей численности педагогического персонала таких школ. Исследование включало пять этапов. В исследовании использовалась официальная статистическая информация по всем регионам России. Эта информация отражает итоги федерального статистического обследования, которое проводилось Министерством образования Российской Федерации. В качестве моделей были использованы функции плотности нормального распределения. В исследовании использовались методы статистического анализа и в частности ANOVA. В

исследовании проводилось математическое моделирование распределения рассматриваемых показателей по регионам России.

Результаты. В ходе исследования было доказано, что в настоящее время отмечается существенная феминизация среднего образования в нашей стране. Проведенное моделирование эмпирических данных продемонстрировало, что значения показателей, характеризующих долю учителей-мужчин в общей численности педагогических работников школ, зависят от возраста учителей. Максимальное значение показателя было отмечено в возрастной группе до тридцати лет. При росте возраста значения показателей снижаются и достигают минимума в возрастной группе от сорока до сорока девяти лет. В более старшем возрасте доля учителей-мужчин увеличивается. В статье рассматриваются региональные особенности формирования преподавательского персонала школ. Были определены регионы с минимальными и максимальными значениями каждого из пяти показателей.

Заключение. Предлагаемый методический подход и полученные результаты обладают научной новизной, поскольку оценка территориальных особенностей гендерной структуры педагогических работников школ в регионах России ранее не проводилась. Представленный в статье методический подход к оценке гендерной структуры педагогического персонала общеобразовательных школ на основе предложенной методологии может применяться в дальнейших исследованиях.

Ключевые слова: среднее образование, гендерная структура, мужчины школьные учителя, регионы России, функции нормального распределения, возраст учителей.

Yulia S. Pinkovetskaya

Ulyanovsk state university, Ulyanovsk, Russia

Assessment of the Proportion of Men in the Number of Teaching Staff in Schools

The purpose of the study. One of the urgent problems of the development of the regions of Russia is to increase the efficiency of regional systems of general secondary education. At the same time, it is necessary to take measures to reduce the gender gap in the personnel of modern schools. In order to understand the gender gap that has developed in the regions, it is advisable to conduct a comparative analysis of differences between regions based on specific indicators. At the same time, it is necessary to take into account the existing differentiation of the proportion of male teaching staff by age. The purpose of our study was to evaluate the indicators characterizing the share of male teachers of different age groups in the total number of teaching staff of schools in the regions of Russia. These indicators were the shares of the number of men of five age groups in the total number of school-teachers in these age categories in 2020.

Materials and methods. The study used the methodological approach proposed by the author, based on the consideration of specific indicators describing the proportion of male teachers of different ages working in general education schools in the total number of teaching staff of such schools. The study included five stages. The study used official statistical information for all regions of Russia. This information reflects the results of a federal statistical survey conducted by the Ministry of Education of the Russian Federation. The density functions of the normal distribution were used as models. The study used methods of statistical analysis and, in particular, ANOVA. The

study carried out mathematical modeling of the distribution of the considered indicators by regions of Russia.

Results. The study proved that there is currently a significant feminization of secondary education in our country. The conducted modeling of empirical data demonstrated that the values of indicators characterizing the share of male teachers in the total number of teaching staff of schools depend on the age of teachers. The maximum value of the indicator was noted in the age group up to thirty years. With increasing age, the values of indicators decrease and reach a minimum in the age group from forty to forty-nine years. At an older age, the proportion of male teachers increases. The article discusses the regional features of the formation of the teaching staff of schools. Regions with minimum and maximum values of each of the five indicators were identified.

Conclusion. The proposed methodological approach and the results obtained have a scientific novelty, since the assessment of the territorial features of the gender structure of school-teachers in the regions of Russia has not been carried out before. The methodological approach presented in the article to the assessment of the gender structure of the teaching staff of secondary schools based on the proposed methodology can be used in further research.

Keywords: secondary education, gender structure, male school-teachers, regions of Russia, functions of normal distribution, age of teachers.

Введение

В конце двадцатого и начале двадцать первого века в большинстве развитых и развивающихся стран наблюдалась тенденция гендерного неравенства в преподавательском составе школ. Во многих научных публикациях последних лет, например [1-2], делаются выводы о нехватке учителей-мужчин в системе среднего образования. Нынешняя тенденция объясняется исследователями финансовыми проблемами, наличием социальных табу и этических проблем, а также безразличием мужчин к карьере школьных учителей [3-5].

Наше исследование было посвящено оценке существующих гендерных характеристик в структуре преподавательского состава средних школ. Изучение современного уровня этого явления важно как для государственных органов, регулирующих систему школьного образования, так и для потенциальных студентов педагогических вузов. Наша статья отвечает на призывы, прозвучавшие в ряде научных публикаций, например [6], изучить нехватку мужчин в общем числе школьных учителей. Несмотря на наличие исследований по проблеме гендерной структуры педагогических работников школ, региональным особенностям изучения этой проблемы по-прежнему уделяется недостаточное внимание. Однако, как показал ряд исследований, например [7], региональные различия в гендерной структуре школьных учителей играют значительную роль.

Целью нашего исследования была оценка показателей, характеризующих долю мужчин разных возрастных групп в общей численности педагогических работников школ в регионах России.

Структура этой статьи приведена ниже. В следующем

разделе представлен обзор научных публикаций, характеризующих гендерную структуру школьных учителей. Затем представлена методология, исходные данные и дизайн исследования. Результаты моделирования и их обсуждение представлены ниже. Последние разделы содержат выводы и библиографические ссылки.

Проблема нехватки мужчин, работающих учителями в школах, нашла отражение в научных публикациях. Обзор наиболее интересных из них, приведен далее. В статье [8] на примере двух городов, расположенных в Бразилии и Португалии, рассматриваются гендерные особенности структуры педагогических работников школ. Делается вывод о том, что присутствие учителей-мужчин в школах демонстрирует учащимся, что мужчины могут успешно выбирать эту деятельность, и доказывает, что у них есть возможность преподавать в средних школах. Статья [9] посвящена изучению проблем, с которыми сталкиваются учителя мужского пола. В этой статье был сделан вывод, что существующие социальные установки и устоявшиеся стереотипы приводят к прекращению образовательной деятельности некоторых молодых учителей-мужчин. Учитывая это, было предложено увеличивать роль преподавательских коллективов в поддержке своих начинающих коллег. По мнению автора статьи [10], основными четырьмя факторами, затрудняющими работу мужчин в качестве школьных учителей, являются социальный статус, относительно низкая заработная плата, работа в основном в женском коллективе, а также трудности в установлении контактов с детьми. Анализ различий между педагогическими работниками школ женского и мужского пола, как показало исследование [11], позволил сделать вывод, что учителя-женщины более за-

ботливы, а учителя-мужчины более спокойны. Опрос, проведенный автором этого исследования, подтвердил мнение о том, что в школах требуется больше мужчин, работающих учителями. Немецкий проект, результаты которого описаны в статье [12], был посвящен мотивациям работы молодых студентов педагогических вузов в качестве школьных учителей. Опрос этих студентов показал, что они находят работу с детьми увлекательной. Аналогичные выводы приведены в статье [13]. В ней отмечается положительная мотивация молодых учителей-мужчин к работе с детьми, многие из которых указали на правильность своего выбора профессии из-за хороших условий труда. В некоторых научных публикациях высказывается мысль о необходимости использования мужчин в качестве примеров для учащихся школ [14]. Особенно сильно мужчины влияют на продвижение позитивной мотивации мальчиков [5].

Определенное внимание этой проблеме уделяется в публикациях российских ученых [15-18]. В этих статьях отмечается преобладание женщин в российских школах и указывается на необходимость увеличения удельного веса мужчин в преподавательском персонале.

В нашей статье рассматриваются показатели, характеризующие долю мужчин, принадлежащих к пяти возрастным группам, в общей численности педагогических работников школ в регионах России. К этим показателям относятся:

- доля мужчин в численности педагогического персонала в возрасте менее тридцати лет (показатель 1);

- доля мужчин в численности педагогического персонала в возрасте от тридцати до тридцати девяти лет (показатель 2);

- доля мужчин в численности педагогического персонала в возрасте от сорока до сорока девяти лет (показатель 3);

— доля мужчин в численности педагогического персонала в возрасте от пятидесяти до пятидесяти девяти лет (показатель 4);

— доля мужчин в численности педагогического персонала в возрасте шестьдесят лет и старше (показатель 5).

Процесс исследования включал пять этапов. На первом этапе были сформированы исходные данные, описывающие численность и гендерную структуру преподавательского состава средних школ в регионах России в 2020 году. Для каждого из регионов были рассмотрены данные, характеризующие учителей, принадлежащих к пяти возрастным группам (моложе тридцати лет, от тридцати до тридцати девяти лет, от сорока до сорока девяти лет, от пятидесяти до пятидесяти девяти лет, шестьдесят лет и старше). На втором этапе были рассчитаны значения показателей, характеризующих долю мужчин в общей численности учителей по возрастным группам для каждого из регионов. На третьем этапе на основе разработки математических моделей было оценено распределение показателей по регионам. На четвертом этапе были определены средние значения показателей по регионам России, а также диапазоны, в которых находятся значения этих показателей для большинства из них. На пятом этапе были определены регионы, которые характеризовались максимальными и минимальными значениями показателей по данным за 2020 год.

В исследовании использовалась официальная статистическая информация [19].

В нашем исследовании были проверены следующие гипотезы:

гипотеза 1 — во всех регионах России в 2020 году среди педагогических работников школ преобладали женщины, то есть наблюдалась феминизация школьного образования;

гипотеза 2 — доля мужчин в общем числе школьных учителей зависит от того, к какой возрастной группе принадлежат эти учителя;

гипотеза 3 — значения показателей, характеризующих долю мужчин в общем числе педагогических работников школ, различаются по регионам, но коэффициент вариации для каждого из пяти показателей по регионам не очень значительный (то есть не превышает 33%);

гипотеза 4 — регионы, характеризующиеся максимальными и минимальными значениями доли мужчин в общем числе педагогических работников школ, расположены в разных федеральных округах. То есть территориальное расположение регионов не влияет на максимальные и минимальные значения каждого из пяти показателей.

Оценка значений пяти рассматриваемых показателей проводилась на основе математического моделирования исходных эмпирических данных. В качестве моделей мы использовали функции плотности нормального распределения, метод разработки которых для оценки значений относительных показателей был предложен автором. Некоторые аспекты использования методологии приведены в статьях [20–21]. Для оценки качества достигнутых функций, т.е. уровня аппроксимации эмпирических данных, мы использовали известные и хорошо зарекомендовавшие себя статистические тесты Пирсона, Колмогорова-Смирнова и Шапиро-Вилка.

Дисперсионный анализ показателей для регионов с минимальными и максимальными значениями показателей, проведенный на пятом этапе исследования для каждого из показателей, был основан на методе ANOVA [22]. Процедура однофакторного дисперсионного анализа включала опре-

деление соотношения между внутригрупповой дисперсией и межгрупповой дисперсией по группам с максимальными и минимальными значениями показателей. Дисперсионный анализ позволил нам проверить, насколько дисперсия, вызванная различием между группами, была больше по сравнению с дисперсией, вызванной внутригрупповой изменчивостью. То есть показать существенные различия между группами регионов с максимальными и минимальными значениями показателей.

Результаты исследования и их обсуждение

В ходе вычислительного эксперимента было проведено экономико-математическое моделирование на основе эмпирических данных. Модели, описывающие распределения ($y_1; y_2; y_3; y_4; y_5$) пяти показателей ($x_1, \%$; $x_2, \%$; $x_3, \%$; $x_4, \%$; $x_5, \%$) приведены далее:

— доля мужчин в численности педагогического персонала в возрасте менее тридцати лет

$$y_1(x_1) = \frac{201,57}{2,91 \times \sqrt{2\pi}} \cdot e^{-\frac{(x_1 - 14,89)^2}{2 \times 2,91 \times 2,91}}; \quad (1)$$

— доля мужчин в численности педагогического персонала в возрасте от тридцати до тридцати девяти лет

$$y_2(x_2) = \frac{213,43}{2,92 \times \sqrt{2\pi}} \cdot e^{-\frac{(x_2 - 12,14)^2}{2 \times 2,92 \times 2,92}}; \quad (2)$$

— доля мужчин в численности педагогического персонала в возрасте от сорока до сорока девяти лет

$$y_3(x_3) = \frac{207,50}{2,22 \times \sqrt{2\pi}} \cdot e^{-\frac{(x_3 - 8,10)^2}{2 \times 2,22 \times 2,22}}; \quad (3)$$

— доля мужчин в численности педагогического персонала в возрасте от пятидесяти до пятидесяти девяти лет

$$y_4(x_4) = \frac{193,67}{2,01 \times \sqrt{2\pi}} \cdot e^{-\frac{(x_4 - 9,06)^2}{2 \times 2,01 \times 2,01}}; \quad (4)$$

— доля мужчин в численности педагогического персона-

ла в возрасте шестьдесят лет и старше

$$y_5(x_5) = \frac{237,14}{3,59 \times \sqrt{2\pi}} \cdot e^{-\frac{(x_5 - 13,02)^2}{2 \times 3,59 \times 3,59}}. \quad (5)$$

Качество функций (1)–(5) мы проверили с использованием критериев Колмогорова-Смирнова, Пирсона и Шапиро-Вилка. Расчетные величины по критериям приведены в таблице 1.

Информация, приведенная в столбце 2 таблицы 1, показала, что все рассчитанные значения меньше критического значения по критерию Колмогорова-Смирнова (0,174) при уровне значимости, равном 0,05. Данные в столбце 3 меньше критического значения теста Пирсона (9,49), при уровне значимости 0,05. Данные в колонке 4 превышают критическое значение 0,93 теста Шапиро-Уилка при уровне значимости 0,05. Таким образом, вычислительный эксперимент показал, что все разработанные функции обладают высоким качеством.

На следующем этапе исследования мы произвели оценку рассматриваемых показателей на основе разработанных функций. Значения показателей, средние по странам, приведены в колонке 2 таблицы 2. Средние значения были определены на основе функций (1)–(5). В третьем столбце таблицы 2 указаны стандартные отклонения для изучаемых показателей. Значения показателей, характеризующих верхнюю и нижнюю границы интервалов, соответствующих большинству регионов, приведены в столбце 4. Нижние пределы вычисляются как разница между средним значением и стандартным отклонением, а верхние пределы – как сумма среднего значения и стандартного отклонения.

Приведенная выше информация показывает возможность оценки распределения показателей гендерной структуры педагогических работников школ

Таблица 1 (Table 1)

Рассчитанные значения критериев

Calculated criteria values

Показатели	Тесты		
	Колмогорова-Смирнова	Пирсона	Шапиро-Вилка
доля мужчин в численности педагогического персонала в возрасте менее тридцати лет	0,07	4,08	0,95
доля мужчин в численности педагогического персонала в возрасте от тридцати до тридцати девяти лет	0,04	0,47	0,98
доля мужчин в численности педагогического персонала в возрасте от сорока до сорока девяти лет	0,06	3,91	0,96
доля мужчин в численности педагогического персонала в возрасте от пятидесяти до пятидесяти девяти лет	0,05	3,78	0,97
доля мужчин в численности педагогического персонала в возрасте шестьдесят лет и старше	0,05	4,11	0,96

Источник: Данные в таблице основаны на разработанных функциях (1)–(5)

Source: The data in the table are based on the developed functions (1)–(5)

Таблица 2 (Table 2)

Значения показателей, характеризующих долю мужчин в общей численности школьных учителей по регионам России

Values of indicators characterizing the share of men in the total number of school-teachers by regions of Russia

Показатели	Средние по регионам значения	Значения стандартного отклонения	Значения по большинству регионов
доля мужчин в численности педагогического персонала в возрасте менее тридцати лет	14,89	2,91	11,98–17,80
доля мужчин в численности педагогического персонала в возрасте от тридцати до тридцати девяти лет	12,14	2,92	9,22–15,06
доля мужчин в численности педагогического персонала в возрасте от сорока до сорока девяти лет	8,10	2,22	5,88–10,32
доля мужчин в численности педагогического персонала в возрасте от пятидесяти до пятидесяти девяти лет	9,06	2,01	7,05–11,07
доля мужчин в численности педагогического персонала в возрасте шестьдесят лет и старше	13,02	3,59	9,43–16,61

Источник: Расчеты выполнены автором на основе функций (1)–(5).

Source: Calculations are made by the author based on functions (1)–(5).

по регионам с использованием функций плотности нормального распределения.

Официальная статистическая информация [19] показала, что в 2020 году мужчины работали школьными учителями во всех 82 регионах России. В среднем доля учителей-мужчин во всех школах России составляла 11%. Анализ данных,

приведенных во второй таблице, позволяет нам охарактеризовать долю учителей-мужчин в школах для каждой из пяти возрастных групп. В среднем по всем регионам доля мужчин в возрасте до тридцати лет в 2020 году в общем числе всех учителей этого возраста в школах составила 14,9%. То есть каждый седьмой учитель

принадлежал к этой гендерной страте. В большинстве регионов значения этого показателя колебались от 12,0% до 17,8%. В среднем по регионам доля мужчин в возрасте от тридцати до тридцати девяти лет в общем числе всех педагогических работников этой возрастной группы составила 12,1%. Следовательно, в этой возрастной категории из каждых восьми школьных учителей только один был мужчиной. В большинстве регионов значения соответствующего показателя варьировались от 9,2% до 15,1%. Среднее значение третьего показателя по регионам России в 2020 году составило 8,1%, то есть только каждый двенадцатый учитель был мужчиной. По большинству регионов интервал изменения значения доли мужчин в возрасте от сорока до сорока девяти лет находился в диапазоне от 5,9% до 10,3%. Среднее по регионам значение доли мужчин в возрасте от пятидесяти до пятидесяти девяти лет в 2020 году в общей численности всех педагогических работников школ этого возраста составило 9,1%. То есть каждый одиннадцатый школьный учитель принадлежал к этой гендерной страте. В большинстве регионов значения этого показателя колебались от 7,1% до 11,1%. Доля мужчин в возрасте шестидесяти лет и старше в общем числе всех учителей этой возрастной группы составила 13,0%. Следовательно, в этой возрастной категории из каждых восьми учителей только один был мужчиной. В большинстве регионов значения соответствующего показателя варьировались от 9,4% до 16,6%. В целом по всем возрастным категориям доля мужчин среди учителей была невелика и не превышала 15%. Приведенные выше данные подтверждают первую гипотезу, сформулированную ранее, о том, что среди педагогических работников школ преобладают женщины.

Представляет интерес сравнительный анализ значений удельного веса мужчин в пяти рассматриваемых возрастных группах педагогических работников школ. Этот анализ показывает, что по мере увеличения возраста учителей-мужчин их доля в общем числе учителей сначала уменьшается, а затем увеличивается. Минимальное значение соответствует группе учителей-мужчин в возрасте от сорока до сорока девяти лет. Максимальное значение соответствует возрастной группе менее тридцати лет. Чтобы объяснить этот феномен, необходимо обратиться к истории российского образования. Максимальное значение доли учителей-мужчин в возрасте менее тридцати лет обусловлено тем, что в последние годы Правительство приняло меры по повышению заработной платы учителей, работающих в средних школах. Кроме того, в большинстве регионов России, за исключением мегаполисов и территорий с большими запасами полезных ископаемых, наблюдается тенденция сложности трудоустройства в престижных отраслях экономики. Поэтому многие молодые люди мужского пола связали свою карьеру с преподаванием в школах, тем более что подготовка учителей по соответствующим программам в высших учебных заведениях обычно бесплатная. Относительно высокая доля учителей-мужчин в возрасте шестидесяти лет и старше объясняется тем фактом, что эти люди начали свою преподавательскую деятельность более тридцати пяти лет назад. В то время многие мужчины становились учителями, поскольку заработная плата учителей была на том же уровне, что и в других видах деятельности. Минимальное значение доли группы учителей в возрасте от сорока до сорока девяти лет объясняется тем, что они выбрали эту карьеру в начале девяностых го-

дов прошлого века, когда престиж профессии учителя был на низком уровне из-за низкой заработной платы. Сравнительный анализ возрастной структуры педагогических работников школ показал, что доля учителей-мужчин зависит от того, к какой возрастной группе они принадлежат. Таким образом, вторая гипотеза подтвердилась.

Данные второй таблицы позволяют сделать вывод о дифференциации значений показателей по регионам. Была проанализирована степень вариации каждого из показателей. Для этой цели использовались стандартные отклонения, указанные в столбце 3. Индексы вариации следующие: по первому показателю – 20%, по второму показателю – 24%, по третьему показателю – 27%, по четвертому показателю – 22%, по пятому показателю – 28%. Этот анализ показал, что в рассматриваемых регионах уровень дифференциации значений всех пяти показателей был ниже 33%, то есть не очень значителен. Таким образом, третья гипотеза подтвердилась.

Следующим этапом работы было определение регионов, в которых были отмечены максимальные и минимальные значения каждого показателя. Максимальными значениями являются те, которые превышают верхние пределы диапазонов, указанных в столбце 4 таблицы 2, а минимальными значениями являются те, которые меньше нижних пределов указанных диапазонов.

Ниже, в качестве примера, перечислены регионы, по которым были отмечены максимальные и минимальные значения первого показателя в 2020 году. Первый показатель имел максимальные значения в Курганской области (УФО), Новосибирской области, республике Алтай, республике Тыва, Томской области (СФО), республике Ингушетия, Чеченской республике,

республике Дагестан (СКФО), городе Москве (ЦФО), Магаданской области, Чукотском автономном округе, республике Саха, Сахалинской области (ДФО), республике Адыгея, республике Калмыкия (ЮФО). Минимальные значения первого показателя в 2020 году отмечались в городе Севастополе (ЮФО), Еврейской автономной области, Амурской области (ДФО), Архангельской области (СЗФО), Кемеровской области (СФО), республике Марий Эл, Пермском крае (ПФО), Брянской области, Тверской области (ЦФО).

В приведенных выше перечнях регионов в скобках после их названий указано их территориальное расположение. Проведенный анализ показал, что регионы с максимальными и минимальными значениями каждого из пяти показателей расположены в разных федеральных округах. То есть максимальные и минимальные значения показателей не связаны с территориальным расположением регионов. Следовательно, четвертая гипотеза подтвердилась.

Затем был проведен так называемый анализ ANOVA. При этом по каждому из рассматриваемых показателей сравнивались значения показателей для двух групп регионов соответственно с максимальными и минимальными значениями показателей. Результаты анализа ANOVA приведены в таблице 3. В нем содержатся статистические оценки для каждой из этих групп регионов. При этом в первой и второй строках таблицы указаны соответственно средние значения показателей по группам регионов с максимальными и минимальными значениями. Третья и четвертая строки показывают значения дисперсий для каждой из групп регионов с максимальными и минимальными значениями показателей. В пятой строке показаны межгрупповые дисперсии для

Таблица 3 (Table 3)

Статистические характеристики, описывающие группы регионов с максимальными и минимальными значениями показателей

Statistical characteristics describing groups of regions with maximum and minimum values of indicators

№	Статистические характеристики	Показатели				
		1	2	3	4	5
1	Среднее значение по группам регионов с максимальными значениями показателей	20,30%	17,06%	12,73%	13,16%	21,29%
2	Среднее значение по группам регионов с минимальными значениями показателей, %	10,57%	8,21%	5,40%	6,43%	8,39%
3	Дисперсия по группам регионов с максимальными значениями	8,30	3,35	3,63	10,38	20,18
4	Дисперсия по группам регионов с минимальными значениями	2,46	0,71	0,22	0,11	0,49
5	Дисперсия между группами стран с максимальными и минимальными значениями	567,85	470,47	281,34	293,89	1081,66
6	Дисперсия внутри групп регионов с максимальными и минимальными значениями	6,01	2,03	1,84	5,24	10,34
7	Расчетное значение критерия Фишера	94,44	231,79	153,17	56,05	104,65
8	Критическое значение критерия Фишера	4,28	4,30	4,38	4,26	4,26
9	Уровень значимости	less 0,001	less 0,001	less 0,001	less 0,001	less 0,001

Источник: Рассчитано автором на основе анализа ANOVA.

Source: Calculated by the author based on ANOVA analysis.

групп регионов с максимальными и минимальными значениями показателей. Шестая строка показывает дисперсию внутри групп регионов. Седьмая, восьмая и девятая строки таблицы демонстрируют результаты тестирования качества ANOVA.

Анализ данных, представленных в таблице 3, показывает, что для групп регионов, характеризующихся максимальными и минимальными значениями показателей, существуют относительно небольшие различия внутри каждой группы. Это свидетельствует о том, что в каждую из этих групп входят регионы с небольшими различиями в значениях показателей. Средние значения для групп регионов с максимальными значениями показателей существенно отличаются от средних значений для групп регионов с минимальными значениями. Дисперсия между

группами регионов с максимальными и минимальными значениями значительно больше, чем дисперсия, характерная для каждой из групп по всем рассматриваемым показателям. Данные, приведенные в таблице 3, показывают, что по каждому из рассматриваемых в статье показателей существуют существенные различия между группами регионов с максимальными значениями и минимальными значениями. Это следует из того факта, что соотношение между межгрупповыми и внутригрупповыми отклонениями по каждому из показателей, приведенных в седьмой строке таблицы, значительно больше единицы. Эти соотношения представляют собой рассчитанные значения критерия Фишера, которые больше табличных значений этого критерия, приведенных в восьмой строке таблицы. Необходимо отме-

титель, что уровень значимости составляет менее 0,001, то есть с вероятностью 99,9% наблюдаются существенные различия, характерные для групп регионов с максимальными и минимальными значениями показателей. Таким образом, статистические характеристики анализа ANOVA, основанного на межгрупповых различиях, а именно на критериях Фишера и уровне значимости, показали высокое качество полученных оценок. Следовательно, группы регионов с максимальными и минимальными значениями показателей имеют большие отличия между собой. Это положение относится ко всем пяти рассматриваемым показателям.

Заключение

Наше исследование вносит важный вклад в понимание гендерной структуры педагогических работников школ в регионах России. Цель исследования, связанная с оценкой показателей, характеризующих долю мужчин разных возрастных групп в общей численности школьных учителей в регионах России, была достигнута. Следующие выводы обладают научной новизной и оригинальностью. Во-первых, было установлено, что мужчины работали во всех российских школах, и их доля достигла 11% от числа всех педагогических работников школ. В ходе исследования были рассмотрены пять возрастных групп учителей-мужчин и предложен метод оценки пяти по-

казателей с использованием функций плотности нормального распределения. На основе предложенной методологии было оценено распределение этих показателей в 2020 году по всем регионам России. Результаты вычислительного эксперимента показали, что в среднем по российским регионам наблюдалась значительная феминизация преподавательского персонала школ. Доля учителей-мужчин в общей численности преподавательского состава школ составила соответственно 14,9% для возрастной группы младше тридцати лет, 12,1% для возрастной группы от тридцати до тридцати девяти лет, 8,1% для возрастной группы от сорока до сорока девяти лет, 9,1% для возрастной группы от пятидесяти до пятидесяти девяти лет и 13,0% для возрастной группы от шестидесяти лет и старше. Следовательно, ни в одной из возрастных групп доля мужчин не превышала 15%. В ходе исследования было доказано, что значения показателей, характеризующих долю мужчин в общей численности школьных учителей, зависят от их возраста. Максимальное среднее значение показателя было отмечено в возрастной группе до тридцати лет. Затем значения показателей снизились и достигли минимума в возрастной группе от сорока до сорока девяти лет. В более старшем возрасте доля учителей-мужчин увеличилась.

Наблюдалась определенная дифференциация значений пяти рассматриваемых пока-

зателей по регионам. В то же время он был не очень значительным, поскольку коэффициенты вариации значений показателей не превышали 28%. Были определены регионы, которые характеризовались максимальными и минимальными значениями пяти рассматриваемых показателей. Сравнительный анализ показал, что территориальное расположение регионов не влияет на максимальные и минимальные значения показателей.

Практическая значимость исследования для федеральных и региональных органов власти заключается в учете гендерных и территориальных особенностей формирования кадров среднего образования в России. Результаты работы могут быть использованы при обосновании необходимости дополнительного привлечения мужчин к преподаванию в средних школах. Для юношей абитуриентов особый интерес могут представлять данные об увеличении доли учителей-мужчин в школах. Полученные новые знания представляют интерес и могут быть использованы в образовательных программах высшего образования по соответствующим специальностям. Дальнейшие исследования могут быть направлены на оценку гендерной структуры школьных учителей в отдельных муниципальных образованиях, относящихся к каждому из регионов России. При этом может быть использована предложенная в статье методика.

Литература

1. Davis I., Hay S. Primary masculinities: how male teachers are regarded as employees within primary education – a global systematic literature review // *Sex Education*. 2018. № 18(1). С. 280–292.
2. Moosa S., Bhana D. Masculinity as care: men can teach young children in the early years // *Early Years*. 2020. № 40(1). С. 1–15.
3. Lahelma E. Gender awareness in Finnish

teacher education: an impossible mission? // *Education Inquiry*. 2011. №. 2(2). С. 263–276.

4. Martino W. Male Teachers as role models: addressing issues of masculinity, pedagogy and re-masculinization of schooling. The Ontario institute for studies in education of the University of Toronto. 2008. № 38(2). С. 189–223.

5. Saigol M., Danish S. Feminisation of teaching: factors affecting low male participation in early childhood teaching at private schools in

Pakistan // Journal of Education and Educational Development. 2016. № 3 (2). С. 147–178.

6. Nayoung H., Fitzpatrick B. Male teacher assignment and teacher turnover in elementary schools // AERA Open. 2021. № 7(1). С. 1–14.

7. Mcgrath K., Moosa S., Van Bergen P., Bhana D. The Plight of the Male Teacher: An Interdisciplinary and Multileveled Theoretical Framework for Researching a Shortage of Male Teachers // The Journal of Mens Studies. 2019. № 28(2). С. 1–16.

8. Rabelo A. Discriminated teachers: A study of male teachers in the early grades of primary school // Educação e Pesquisa. 2013. № 39(4). С. 907–925.

9. Cruickshank V., Kerby M., Baguley M. How do Pre-Service Male Primary Teachers Cope with Gender Related Challenges? // Australian Journal of Teacher Education. 2021. № 46 (1). С. 74–88.

10. Cushman P. Let's hear it from the males: Issues facing male primary school teachers // Teaching and Teacher Education. 2005. № 21(3). С. 227–240.

11. Wood T. Teacher Perceptions of Gender-Based Differences among Elementary School Teachers // International Electronic Journal of Elementary Education. 2012. № 4(2). С. 317–345.

12. Faulstich-Wieland H. Should Male Primary School Teachers be There Principally as Role Models for Boys? // Universal Journal of Educational Research. 2013. № 1(2). С. 65–73.

13. Mulholland J., Hansen P. Men who become primary school teachers: An early portrait // Asia-Pacific Journal of Teacher Education. 2003. № 31(3). С. 213–224.

14. Martino W., Kehler M. Male teachers and the «boy problem»: An issue of recuperative masculinity politics // McGill Journal of Education. 2006. № 41(2). С. 113–131.

15. Захарова Е.Ю., Порш Л.А. Феминизация российского образования: этапы становления // Гуманитарный вектор. 2016. № 11(1). С. 58–62.

16. Рязанцев А.А. Верните в школу мужчин // Наука и Школа. 2018. № 5. С. 189–192.

17. Семенова Л.Э. Особенности самопрезентации мужчин-педагогов средней школы // Психологическая наука и образование. 2008. № 13(1). С. 82–89.

18. Саенко Л.А., Суменко Л.В. Феминизация современной российской школы // Аспирантский вестник Поволжья. 2012. № 12(7–8). С. 137–142.

19. Отчет по форме федерального статистического обследования № ОО-1. Министерство образования Российской Федерации. 2022. [Электрон. ресурс]. Режим доступа: <https://docs.edu.gov.ru/document/ed3ca74f26a1dc055a313991f66d2fa3/>.

20. Pinkovetskaia I. Estimation of the share of women in the number of students of higher educational institutions: data by regions of Russia // Perspectivas em Diálogo. 2022. № 09(19). С. 7–20.

21. Pinkovetskaia I., Nuretdinova Y., Nuretdinov I., Lipatova N. Mathematical modeling on the base of functions density of normal distribution // REVISTA DE LA UNIVERSIDAD DEL ZULIA. 2021. № 12(33). С. 34–49.

22. Ostertagova E., Ostertag O. Methodology and Application of One-way ANOVA // American Journal of Mechanical Engineering. 2013. № 1(7). С. 256–261.

References

1. Davis I., Hay S. Primary masculinities: how male teachers are regarded as employees within primary education – a global systematic literature review. Sex Education. 2018; 18(1): 280–292.

2. Moosa S., Bhana D. Masculinity as care: men can teach young children in the early years. Early Years. 2020; 40(1): 1–15.

3. Lahelma E. Gender awareness in Finnish teacher education: an impossible mission? Education Inquiry. 2011; 2(2): 263–276.

4. Martino W. Male Teachers as role models: addressing issues of masculinity, pedagogy and re-masculinization of schooling. The Ontario institute for studies in education of the University of Toronto. 2008; 38(2): 189–223.

5. Saigol M., Danish S. Feminisation of teaching: factors affecting low male participation in early childhood teaching at private schools in Pakistan. Journal of Education and Educational Development. 2016; 3 (2): 147–178.

6. Nayoung H., Fitzpatrick B. Male teach-

er assignment and teacher turnover in elementary schools. AERA Open. 2021; 7(1): 1–14.

7. Mcgrath K., Moosa S., Van Bergen P., Bhana D. The Plight of the Male Teacher: An Interdisciplinary and Multileveled Theoretical Framework for Researching a Shortage of Male Teachers. The Journal of Mens Studies. 2019; 28(2): 1–16.

8. Rabelo A. Discriminated teachers: A study of male teachers in the early grades of primary school. Educação e Pesquisa. 2013; 39(4): 907–925.

9. Cruickshank V., Kerby M., Baguley M. How do Pre-Service Male Primary Teachers Cope with Gender Related Challenges? Australian Journal of Teacher Education. 2021; 46 (1): 74–88.

10. Cushman P. Let's hear it from the males: Issues facing male primary school teachers. Teaching and Teacher Education. 2005; 21(3): 227–240.

11. Wood T. Teacher Perceptions of Gender-Based Differences among Elementary School Teachers. International Electronic Journal of Elementary Education. 2012; 4(2): 317–345.

12. Faulstich-Wieland H. Should Male Primary School Teachers be There Principally as Role

Models for Boys? Universal Journal of Educational Research. 2013; 1(2): 65-73.

13. Mulholland J., Hansen P. Men who become primary school teachers: An early portrait. Asia-Pacific Journal of Teacher Education. 2003; 31(3): 213-224.

14. Martino W., Kehler M. Male teachers and the “boy problem”: An issue of recuperative masculinity politics. McGill Journal of Education. 2006; 41(2): 113-131.

15. Zakharova Ye.YU., Porsh L.A. Feminization of Russian education: stages of formation. Gumanitarnyy vektor = Humanitarian vector. 2016; 11(1): 58-62. (In Russ.)

16. Ryazantsev A.A. Return men to school. Nauka i Shkola = Science and School. 2018; 5: 189-192. (In Russ.)

17. Semenova L.E. Features of self-presentation of male teachers of secondary school. Psikhologicheskaya nauka i obrazovaniye = Psychological science and education. 2008; 13(1): 82-89. (In Russ.)

18. Sayenko L.A., Sumenko L.V. Feminization of the Modern Russian School. Aspirantskiy vest-

nik Povolzh'ya = Postgraduate Bulletin of the Volga Region. 2012; 12(7-8): 137-142. (In Russ.)

19. Otchet po forme federal'nogo statisticheskogo obsledovaniya № OO-1. Ministerstvo obrazovaniya Rossiyskoy Federatsii = Report on the form of the federal statistical survey No. OO-1. Ministry of Education of the Russian Federation. 2022. [Internet]. Available from: <https://docs.edu.gov.ru/document/ed3ca74f26a1dc055a313991f66d2fa3/>. (In Russ.)

20. Pinkovetskaia I. Estimation of the share of women in the number of students of higher educational institutions: data by regions of Russia. Perspectivas em Diálogo. 2022; 09(19): 7-20.

21. Pinkovetskaia I., Nuretdinova Y., Nuretdinov I., Lipatova N. Mathematical modeling on the base of functions density of normal distribution. REVISTA DE LA UNIVERSIDAD DEL ZULIA. 2021; 12(33): 34-49.

22. Ostertagova E., Ostertag O. Methodology and Application of One-way ANOVA. American Journal of Mechanical Engineering. 2013; 1(7): 256-261.

Сведения об авторе

Юлия Семеновна Пиньковецкая

К.э.н. доцент, доцент кафедры экономического анализа и государственного управления

Ульяновский государственный университет,

Ульяновск, Россия

Эл. почта: judy54@yandex.ru

Information about the author

Yulia S. Pinkovetskaya

Cand. Sci. (Economics), Associate Professor

Associate Professor of the Department of Economical Analysis and State Management, Ulyanovsk State

University, Ulyanovsk, Russia

E-mail: judy54@yandex.ru

Комбинированная когнитивная модель прогнозирования деятельности университета*

Цель исследования. Целью исследования является разработка модели прогнозирования показателей деятельности университета на основе когнитивного подхода, в основе которого лежит построение когнитивной карты, отражающей влияние на базовые показатели совокупности латентных факторов и обеспечивающей решение задачи сценарного прогнозирования. Степень достижения требуемых значений базовых показателей, определяющих рейтинг университета, зависит от величины приращения выявленных латентных факторов. Разработанная модель позволяет выбрать наиболее предпочтительный вариант сценарного прогнозирования показателей деятельности университета в условиях существующих ограничений на ресурсы, выделяемые на приращение латентных факторов.

Материалы и методы. Для достижения поставленной цели использованы методы когнитивного моделирования на основе серых нечетких когнитивных карт (НKK) в комбинации с методами интервальной математики и каузальной алгебры. Применение рассмотренного подхода позволило снизить неопределенность экспертных оценок силы взаимосвязи между концептами когнитивной карты за счет использования при описании взаимосвязей между концептами не точечных оценок, а специальных конструкций в виде интервальных оценок, что обеспечило повышение достоверности результатов моделирования. Разработанная модель построена на основе ансамбля серых НKK, что в свою очередь позволило повысить точность и достоверность прогнозной модели. Предложенный подход к решению задачи обеспечения прогнозирования деятельности

университета позволил разработать адекватную когнитивную модель.

Результаты. Разработанная когнитивная модель деятельности университета позволила анализировать динамику изменения факторов и их влияния на базовые показатели, а также динамику развития системы показателей в целом. Проведенный расчет позволил выбрать наиболее приемлемый с точки зрения затрат сценарий приращения значений латентных факторов для получения необходимого значения рейтинга университета в рамках международного институционального рейтинга университетов QS. Проведен сравнительный анализ результатов сценарного прогнозирования на основе обычной НKK, серой НKK и ансамбля серых НKK, который показал преимущество предложенного подхода.

Заключение. В ходе выполнения исследования разработана нечеткая когнитивная модель сценарного прогнозирования мероприятий по достижению требуемых значений целевых показателей деятельности университета в международном институциональном рейтинге QS на основе ансамбля серых НKK. Разработанная модель обеспечивает в условиях заданных ограничений получение наиболее приемлемого сценария планирования приращения базовых показателей до целевых значений за счет идентификации влияющих на них латентных факторов и расчета необходимых значений импульсных воздействий на латентные факторы.

Ключевые слова: нечеткое когнитивное моделирование, сценарное прогнозирование, ансамбль серых нечетких когнитивных карт, нечеткое интервальное множество.

Andrey A. Mikryukov, Mikhail E. Mazurov

Plekhanov Russian University of Economic, Moscow, Russia

Combined Cognitive Model for Forecasting University Activities

Purpose of the study. The purpose of the study is to develop a model for predicting university performance indicators based on a cognitive approach, which is based on the construction of a cognitive map that reflects the influence of a set of latent factors on the basic indicators and provides a solution to the problem of scenario forecasting. The degree of achievement of the required values of the basic indicators that determine the ranking of the university depends on the magnitude of the increment of the identified latent factors. The developed model makes it possible to choose the most preferable variant of scenario forecasting under the existing restrictions on the resources allocated for the increment of latent factors.

Materials and methods. To achieve this goal, cognitive modeling methods based on gray fuzzy cognitive maps (FCM) were used in combination with methods of interval mathematics and causal algebra. The application of the considered approach made it possible to reduce the uncertainty of expert estimates of the strength of the relationship between the concepts of the cognitive map due to the use of special constructions in the form of interval estimates rather than point estimates when describing the relationships

between the concepts, which ensured an increase in the reliability of the modeling results. The developed model is created based on an ensemble of gray FCMs, which, in turn, made it possible to increase the accuracy and reliability of the predictive model. The proposed approach to solving the problem of predicting the activities of the university made it possible to develop an adequate cognitive model.

Results. The developed cognitive model of the university's activities made it possible to analyze the dynamics of changes in factors and their influence on basic indicators, as well as the dynamics of the development of the system of indicators. The calculation made it possible to choose the most cost-effective scenario for incrementing the values of latent factors to obtain the required value of the university ranking in the framework of the QS international institutional ranking of universities. A comparative analysis of the results of scenario forecasting based on conventional FCM, gray FCM, and an ensemble of gray FCM was carried out, which showed the advantage of the proposed approach.

Conclusion. During the study, a fuzzy cognitive model was developed for scenario forecasting of measures to achieve the

* Работа выполнена при финансовой поддержке ФГБОУ ВО «РЭУ им. Г.В. Плеханова»

required values of university performance targets in the QS international institutional ranking based on an ensemble of gray FCMs. The developed model provides, under the given constraints, obtaining the most acceptable scenario for planning the increment of basic indicators to target values by identifying the latent factors

influencing them and calculating the required values of impulse effects on latent factors.

Keywords: fuzzy cognitive modeling, scenario forecasting, ensemble of gray fuzzy cognitive maps, fuzzy interval set.

Введение

В статье рассматривается актуальная задача, связанная с необходимостью повышения позиции РЭУ им. Г.В. Плеханова в авторитетном международном институциональном рейтинге Quacquarelli Symonds (QS), что должно обеспечить повышение конкурентоспособности университета при выполнении Программы развития университета на 2021–2030 годы в рамках реализации Программы стратегического академического лидерства «Приоритет-2030». Для решения поставленной задачи был проведен анализ существующих моделей, используемых для описания функционирования плохо формализуемых и слабоструктурированных социально-экономических и социально-технических систем, к классу которых относятся вышеуказанные учебные заведения.

При моделировании таких систем широкое распространение находят применение когнитивные модели на основе когнитивных карт, представленные в работах отечественных и зарубежных ученых [1–11]. Проведенный анализ литературных источников показал, что одним из подходов, достаточно хорошо показавшим себя при моделировании плохо формализуемых и слабоструктурированных систем, является нечеткое когнитивное моделирование (НKM), объединяющее методы их лингвистического, аналитического и статистического описания [3, 7–10].

Неоспоримым преимуществом НКМ в сравнении с другими подходами является возможность формального описания неизмеримых факторов, представление взаимосвя-

зисимостей между объектами и компонентами в виде нечетких отношений взаимовлияния с использованием методов нечеткой каузальной алгебры, а также использование неполной, нечеткой, в том числе противоречивой информации. НКМ позволяет наилучшим образом отразить неопределенность динамики, а также состояние концептов когнитивной карты и их взаимосвязей.

Модель НКК дает возможность описать поведение сложной системы в целом, а каждый концепт НКК – ее факторную характеристику. В настоящее время на основе НКМ базируется большинство современных систем моделирования и прогнозирования сложных слабоструктурированных систем, отличающихся высокой степенью неопределенности и стохастичности, к классу которых относятся социальные системы, социально-экономические и организационно-технические системы и др. Используемые при когнитивном моделировании инструменты могут обеспечить проведение разведочного и оценочного анализа различных стратегий поведения таких систем, позволяющего в том числе прогнозировать сценарии их развития [8].

Одной из распространенных проблем при построении когнитивных моделей является недостаточная степень адекватности модели объекту исследования, а также несоответствие требованиям по точности и достоверности, что не позволяет принимать на основе таких моделей рациональные управленческие решения. Как правило, указанные недостатки напрямую связаны с отсутствием необходимых достоверных данных, наличи-

ем большого количества неопределенностей и скрытых количественных и качественных закономерностей, присущих таким системам. Таким образом, задача обеспечения адекватности разрабатываемой модели, а также ее точности и достоверности является актуальной.

В статье предложен новый подход к построению когнитивной модели деятельности университета как плохо формализуемой и слабоструктурированной системы на основе использования разновидности обобщенных НКК, получивших название серых НКК, особенностью которых является возможность представления силы связей между концептами с помощью специальных конструкций в виде интервальных оценок [5, 12]. В этом случае существенно снижается погрешность в оценке силы связей между концептами, которая носит выраженный субъективный характер и не сводится к некоторой усредненной точечной числовой величине, а представляется в виде интервальных значений.

Для повышения точности и достоверности получения конечного результата прогнозирования на основе когнитивной модели разработан подход на основе ансамбля серых НКК, в котором реализовано несколько вариантов формализации знаний и опыта экспертов, а также метода взвешенного голосования, позволяющего получить более точный результат по сравнению с единичной НКК.

Разработанная когнитивная модель позволила в условиях заданных ресурсных ограничений найти наиболее приемлемый сценарий поэтапного планирования приращения

базовых показателей деятельности университета до целевых значений.

Когнитивная модель сценарного прогнозирования показателей деятельности университета на основе серой нечеткой когнитивной карты

Для решения поставленной задачи предложен подход к сценарному прогнозированию на основе серой нечеткой когнитивной карты, отличительной особенностью которой является использование специальной конструкции, позволяющей снижать неопределенность (разброс) оценок экспертами состояний концептов когнитивной карты. НКК задается с помощью кортежа множеств [8]:

$$\text{НКК} = \langle C, F, W \rangle,$$

где $C = \{C_j\}$ – множество концептов – вершин графа, в качестве которых выступают факторы, являющиеся наиболее значимыми в рассматриваемой задаче;

$F = \{F_k\}$ – множество направленных дуг графа связей между концептами;

$W = \{W_{ij}\}$ – множество весов связей НКК (связи могут быть положительными (усиливающими, $W_{ij} > 0$) и отрицательными (ослабляющими, $W_{ij} < 0$) влияние концепта C_i на концепт C_j).

Значения весов W_{ij} задаются с помощью нечеткой лингвистической шкалы. Каждому значению лингвистической переменной ставится в соответствие некоторый интервал $[0, 1]$ для положительных связей или интервалу $[-1, 0]$ – для отрицательных связей.

В произвольный дискретный момент времени $t = 0, 1, 2, \dots, n$ состояние НКК описывается уравнением

$$X_i(t+1) = f \left[X_i(t) + \sum_{j=1}^n W_{ij} X_j(t) \right], (1)$$

$$i = 1, 2, \dots, n,$$

где $X_i(t)$ – значение переменной состояния i – го концепта C_j в момент времени $(t + 1)$; n – число концептов НКК, f – нелинейная функция концепта. Начальные условия для расчета определяются вектором $X(0) = (X_1(0), X_2(0), \dots, X_n(0))^T$.

Каждый концепт характеризуется терм-множеством лингвистической переменной

$$T_i = \{T_1^i, T_2^i, \dots, T_{m_j}^i\} \quad (2)$$

где m_j – число типовых состояний i -го концепта. Для описания каждого термина T_{ki} строится терм – множество с функцией принадлежности $\mu_{T_{ki}}(x)$. Связи между типовыми состояниями каждой пары концептов задаются нечеткими переменными, описываемыми соответствующими нечеткими множествами.

Значения весов (силы связей) W_{ij} задаются с помощью нечеткой лингвистической шкалы, представляющей собой упорядоченное множество лингвистических значений (термов) оценок силы связи, например вида: СИЛА_СВЯЗИ = {Не влияет; Слабая; Средняя; Сильная; Очень сильная}.

Каждому из этих значений ставится в соответствие некоторый числовой диапазон, принадлежащий отрезку $[0, 1]$ для положительных связей (табл. 1), или отрезку $[-1, 0]$ для отрицательных связей.

В общем случае взвешенный орграф с произвольными значениями весов $W_{ij} \in [-1, 1]$ описывается динамикой из-

менения его состояния во времени. Состояние орграфа (НКК) при этом определяется совокупностью состояний его концептов C_i , ($i = 1, 2, \dots, n$), каждое из которых описывается переменной состояния $X_i(t)$, принимающей значения в интервале $[0, 1]$.

С целью устранения указанного недостатка предложено использование подхода на основе специальных конструкций, реализованных в нечетких когнитивных моделях на основе серых НКК, позволяющих описывать веса связей между концептами когнитивной карты не точечными оценками, а интервальными числами, заданными на нечетких интервальных множествах [5].

В этом случае уравнение состояния НКК (1) может быть представлено выражением (3)

$$X_i(t+1) = f \left(X_i(t) \oplus \left(\bigoplus_{j=1, (j \neq i)}^n (W_{ji} \otimes X_j(t)) \right) \right) \quad (3)$$

$$i = 1, 2, \dots, n,$$

где: f – функция активации, а веса связей W_{ji} , а также переменные состояния $X_i(t+1)$, $X_j(t)$ представляют собой интервальные числа, которые представляют собой элементы нечетких интервальных множеств. Операции сложения $\oplus$ и умножения $\otimes$ интервальных чисел задаются на нечетких интервальных множествах.

Серое множество $A \subseteq X$ может быть представлено в виде

$$A = \{ \langle x, [\underline{x}, \bar{x}] \rangle | x \in X \}. \quad (4)$$

Элементы $x \in [\underline{x}, \bar{x}] \leq A$ серого множества могут при-

Таблица 1 (Table 1)

Оценка силы связи между концептами
Evaluation of the strength of the relationship between concepts

Лингвистическое значение	Числовой диапазон	Обозначение термина	Точечная оценка силы связей+
Не влияет	0	Z	0
Очень слабая	(0; 0,15)	VL	0,12
Слабая	(0,15; 0,35)	L	0,23
Средняя	(0,35; 0,60)	M	0,47
Сильная	(0,6; 0,85)	H	0,72
Очень сильная	(0,86; 1,0)	VH	0,93

нимать значения в диапазоне $[x, \bar{x}] \in [0, 1]$, где x и $\bar{x}$ – нижняя и верхняя граница серого числа, X – универсальное множество. Число $\delta x = (x, -\bar{x})$ носит название серости числа x , а число $x^0 = (x, -\bar{x})/2$ – «отбеленное» (центральное) значение этого числа.

Веса связей между концептами серой НКК задаются в виде серых чисел $[W_{ij}, \bar{W}_{ij}]$. Переменные состояния концептов также описываются серыми числами, значения которых лежат в интервале $[X_{ij}, \bar{X}_{ij}]$ и определяются уравнением (3).

В общем случае применение интервального подхода обладает рядом преимуществ [5]:

- не требуется знание вероятностных характеристик неопределенных факторов, которые редко бывают точно известны на практике;

- при минимаксном подходе получают строгие оценки для самих искомых величин, а не для вероятностей или математических ожиданий, что имеет важное значение при наличии малого числа измерений параметров и одной или нескольких реализаций;

- статистические характеристики не могут гарантировать определенный исход одного конкретного опыта;

- во всех случаях даются гарантированные двусторонние аппроксимации искомых решений.

Точность интервального результата полностью определяется следующими факторами: неопределенностью в задании исходных данных; значениями округлений при выполнении операций, изменяющих или порождающих интервальные объекты; приближенным характером используемого численного метода; а также степенью учета зависимостей между участвующими в вычислении интервальными объектами (переменными и константами).

На первом этапе исследования проведена идентификация и интерпретация скрытых

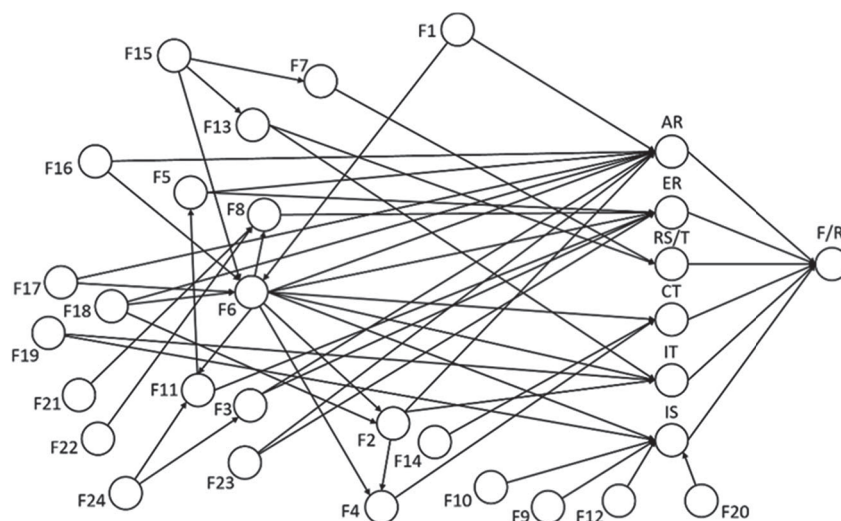


Рис. 1. Когнитивная карта на основе графа взаимосвязи факторов, базовых показателей и функционала

(Обозначения: F/R – функционал/рейтинг, (F1-F24) – латентные факторы)

Fig. 1. Cognitive map based on the graph of the relationship of factors, basic indicators and functional

(Denotations: F/R - functional/rating, (F1-F24) - latent factors)

(латентных факторов), влияющих на целевые показатели, и оценка их значимости с использованием методов факторного анализа [13-15].

На следующем этапе разработана когнитивная модель, включающая ансамбль серых НКК, построенных с учетом различных вариантов формализации знаний и опыта экспертов в указанной предметной области.

На рис. 1 представлена серая когнитивная карта в виде знакового ориентированного графа, отражающая влияние совокупности факторов на базовые показатели деятельности университета и рейтинговый показатель [16-18].

Когнитивная карта отражает взаимосвязи латентных факторов, базовых показателей и функционала с учетом полученных на основе факторного анализа в работе [15] значений корреляционных зависимостей между функционалом и базовыми показателями, а также экспертными оценками взаимовлияния латентных факторов и их влияния на целевые показатели.

Взаимосвязь базовых показателей и функционала

определяются формулой (5) в соответствии с правилами международного институционального рейтинга QS.

$$R = \sum_{i=1}^6 w_i x_i, \quad (5)$$

где w_i – вес соответствующего показателя; x_i – его значение.

На рис. 1 приняты следующие обозначения: F – функционал; R – рейтинг университета, Базовые показатели заданы правилами международного институционального рейтинга QS [19]: AP – академическая репутация; PP – репутация у работодателя; ОСП – отношение числа студентов к числу преподавателей; ЦП – показатель цитируемости преподавателей; МП – число международных преподавателей; МС – число международных студентов. Факторы представляют собой группы определенных переменных, коррелирующих между собой больше, чем с переменными, входящими в другой фактор. Таким образом, содержательный смысл факторов был выявлен путем исследования корреляционной матрицы исходных данных. Весовые значения силы связей между факторами заданы

на основе экспертных оценок с учетом интервальной шкалы (нижнее и верхнее значения, т.е. серость и «отбеленное» значение этого числа). F1 – фактор «Научные школы и диссертационные советы» (0,5-0,7; 0,6); F2 – фактор «Совместные научные проекты» (0,2-0,4; 0,3); F3 – Наличие базовых кафедр (0,1-0,3; 0,2); F4 – Число публикаций в БД Scopus (0,5-0,7; 0,6); F5 = Востребованные направления подготовки (0,3); F6 = уровень квалификации НПП (0,1-0,3; 0,2); F7 = Число НПП (0,5-0,7; 0,6); F8 = Уровень компетенции студентов (0,4-0,6; 0,5); F9 = НПП с языковой подготовкой (0,3-0,5; 0,4); F10 = Места в общежитии (0,1-0,3; 0,2); F11 = Востребованность выпускников у работодателей (0,2-0,4; 0,3); F12 = Площади для образовательной деятельности (0,2-0,4; 0,3); F13 – Уровень оплаты НПП (0,3-0,5; 0,4); F14 = Стимулирующие факторы (0,1-0,3; 0,2); F15 = Расширение соцпакета НПП (0,2-0,4; 0,3); F16, Изменение структуры занятости НПП (0,2-0,4; 0,3); F17, Доля НПП, планирующих строить интернациональную научную карьеру (0,1-0,3; 0,2); F18, Академическая мобильность НПП (0,2-0,4; 0,3); F19, Конвергенция образовательных программ с иностранными вузами (0,3-0,5; 0,4); F20, Иностранная абитуриентская компания (0,2-0,4; 0,3); F21, Увеличение

числа On-line курсов MOOCs (0,2-0,4; 0,3), F22, Внедрение индивидуальных образовательных траекторий (0,3-0,5; 0,4); F23, Внедрение дистанционных технологий (0,2-0,4; 0,3); F24, Теснота взаимосвязи с работодателем (0,3-0,5; 0,4).

Для повышения точности полученного решения когнитивная модель построена на основе прогнозирующего ансамбля, идея применения которого основана на комбинации результатов прогнозирования нескольких моделей когнитивных карт (рис. 2).

В работах [20,21] отмечается, что перспективным направлением повышения точности решений является объединение (композиция) множества отдельных алгоритмов в одну систему. В этом случае ошибки отдельных алгоритмов взаимно компенсируются. Ансамблевая организация рассматривается в ряде работ [22-24]. В работе [23] экспериментально доказана эффективность применения ансамблевой организации для распознавания изображений.

В настоящее время известно несколько методов построения прогнозирующих ансамблей. Преимущества ансамбля моделей по сравнению с отдельной моделью, входящей в ансамбль, обусловлено следующими причинами [23]:

1. Ансамбль уменьшает среднеквадратическую ошибку. Усреднение по множеству моделей, построенных на ос-

нове независимых обучающих множеств, всегда уменьшает ожидаемое значение среднеквадратической ошибки.

2. Ансамбли моделей, обученных на различных подмножествах исходных данных имеют больший шанс найти глобальный оптимум, так как ищут его из разных начальных точек.

Для формирования выходного значения ансамбля использован метод взвешенного голосования, достаточно хорошо зарекомендовавший себя. Каждой модели НКК по результатам тестирования на исторических данных присваивается весовой коэффициент с учетом среднеквадратической ошибки.

Выходное значение определяется в соответствии с формулами:

$$Y(x) = F(y_1(x), y_2(x), \dots, y_m(x)) = \frac{1}{m} \sum_{i=1}^m a_i(x) y_i(x),$$

$$\sum_{i=1}^m a_i(x) = 1, \forall x \in X, \quad (6)$$

где $x = (x_1, x_2, \dots, x_k)$ – входной вектор,

$y_i(x)$ – выходное значение i -ой НКК;

a_i – весовой коэффициент i -ой модели;

k – число НКК,

$Y(x) = (y_1, y_2, \dots, y_k)$ – вектор значений выходного сигнала прогнозирующего ансамбля, F – функция для получения результирующего решения.

Решаемая задача заключается в поиске на основе разработанной когнитивной модели наиболее приемлемого с точки зрения затрат сценария приращения значений латентных факторов для получения необходимого значения рейтинга университета в рамках международного институционального рейтинга университетов QS.

Наличие причинно-следственных связей между латентными факторами и базо-

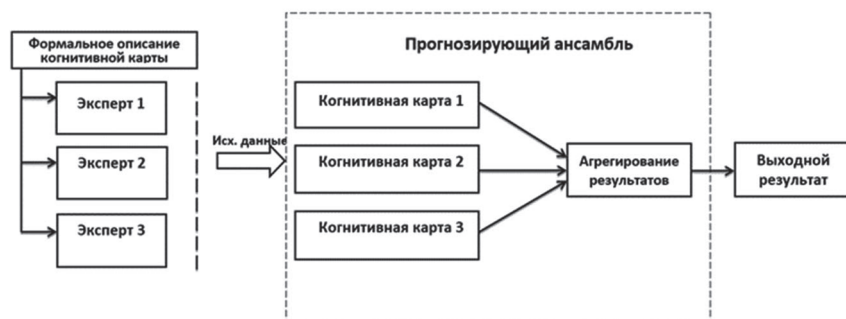


Рис. 2. Прогнозирующий ансамбль на основе композиции серых нечетких когнитивных карт

Fig. 2. Predictive ensemble based on the composition of gray fuzzy cognitive maps

выми показателями позволяют решить поставленную задачу сценарного прогнозирования в условиях заданных ограничений.

Подход к сценарному прогнозированию на основе когнитивной карты включает следующие этапы [17]:

- генерирование сценариев и оценка влияния приращений значений факторов на увеличение рейтинга университета,

- корректировка сценариев через дискретные интервалы времени (в нашем случае, ежегодно) с учетом достигнутых значений места университета в международном рейтинге QS.

По истечении очередного интервала времени строится новый сценарий для достижения нового целевого значения рейтинга.

Результаты сценарного прогнозирования показателей деятельности университета на основе нечеткой когнитивной модели

После описания взаимосвязей между факторами с помощью уравнений, задания интервальных значений весов их взаимовлияний и значений начальных приращений факторов, можно анализировать динамику изменения факторов и развития системы показателей в целом.

Для того чтобы сформировать возможные стратегии развития системы, необходимо в первую очередь прогнозировать ее саморазвитие, т. е. изучить динамику изменения значений базовых ориентиров при отсутствии внешних управляющих воздействий.

В результате такого прогноза получен вектор значений входных воздействий на латентные факторы $F = (F_1, F_2, \dots, F_{24})$ для достижения требуемого значения целевого фактора (рейтинга университета) в момент времени t с учетом требований по ограничению на ресурсы.

Значения исходных данных, не имеющих количественный характер (Уровень квалификации НПР, Уровень компетенций студентов, Уровень квалификации НПР и др.) определены экспертным путем. Предварительные значения интенсивности взаимовлияния между измеримыми факторами когнитивной модели были установлены на основе корреляционного и факторного анализа [15]. Далее коэффициенты были уточнены сообразно логике перехода системы из одного стационарного состояния в другое в результате внешних импульсных воздействий.

В табл. 2 представлены веса связей, заданные экспертами, между концептами в обычной, серой НКК, а также в ансамбле серых НКК, состоящего из трех разнородных НКК.

Веса связей между концептами ансамбля НКК заданы тремя разными экспертами. Сценарное прогнозирование на основе трех вариантов когнитивных моделей (классической НКК, серой НКК и ан-

самбля серых НКК) позволили получить в условиях заданных ограничений на ресурсы наиболее предпочтительные значения приращений факторов, влияющих на базовые показатели, необходимые для достижения требуемых значений последних на основе результатов анализа степени влияния каждого фактора на рейтинг. В табл. 3 представлены результаты эффектов в классической НКК. Аналогичные таблицы были построены для двух остальных вариантов когнитивной модели.

Последовательно задаваемые «слабые» приращения значений перечисленных выше факторов на уровне 10%, позволили оценить чувствительность целевого показателя (рейтинга университета) к управляющим воздействиям по данным направлениям регулирования, на основе чего была выбрана наиболее предпочтительная альтернатива сценарного прогнозирования.

Верификация когнитивной модели выполнена на основе критериев полноты и непроти-

Таблица 2 (Table 2)

Веса связей между концептами НКК
Link weights between FCM concepts

Вес связи W_{ij}	Обычная НКК W_{ij}	Серая НКК $[W_{ij}, \bar{W}_{ij}]$	Ансамбль серых НКК		
			НКК2 $[W_{ij}, \bar{W}_{ij}]$	НКК2 $[W_{ij}, \bar{W}_{ij}]$	НКК3 $[W_{ij}, \bar{W}_{ij}]$
F1-AR	0,6	[0.5-0.7]	[0.5-0.7]	[0.6- 0.8]	[0.3-0.5]
F1 – F6	0,7	[0.6-0.8]	[0.6-0.7]	[0.4-0.6]	[0.4-0.6]
F2-AR	0,3	[0.2-0.4]	[0.2-0.4]	[0.3-0.5]	[0.4-0.6]
F2 - IT	0,4	[0.3-0.5]	[0.3-0.5]	[0.4-0.6]	[0.3-0.5]
F2-F4	0,5	[0.4-0.6]	[0.4- 0.6]	[0.2-0.4]	[0.3-0.5]
...	...	...	...	...	...
...	...	...	...	...	...
F23-AR	0.3	[0.2-0.4]	[0.2-0.4]	[0.2-0.4]	[0.4-0.6]
F23-ER	0.3	[0.2-0.4]	[0.3-0.5]	[0.2-0.4]	[0.3-0.5]
F24-F11	0.4	[0.3-0.5]	[0.3-0.5]	[0.4-0.6]	[0.5-0.7]
F24- F3	0.2	[0.1-0.3]	[0.1-0.3]	[0.3-0.5]	[0.2-0.3]

Таблица 3 (Table 3)

Таблица эффектов нечеткой когнитивной карты
Table of effects of fuzzy cognitive map

Импульсное воздействие на латентный фактор	F1	F2	...	F22	F23	F24
Приращение рейтинга ΔR	0.001	0.0007	...	0.00004	0.00002	0.0006

Таблица 2 (Table 2)

Сравнительная оценка результатов сценарного прогнозирования
Comparative evaluation of the results of scenario forecasting

№ п\п	Тип когнитивной модели	Разброс оценок состояния целевых концептов НКК в %	Средняя ошибка прогнозирования в %
1.	НКК	17,3	14,1
2.	Серая НКК	14,6	11,7
3.	Ансамбль серых НКК	10,4	8,2

воречивости влияния факторов — причин на факторы — следствия и нарушения транзитивности каузальных явлений [19]. С целью проверки адекватности и точности когнитивной модели она была протестирована на ретроспективном периоде 2014–2021 гг. на основе имеющихся статистических данных по измеримым факторам модели. Общая корректность модели на этом этапе подтвердилась близостью рассчитанных на модели темпов прироста факторов к действительным темпам прироста.

Сравнительный анализ результатов сценарного прогнозирования на основе обычной НКК, серой НКК и ансамбля серых НКК, показал, что отклонение полученного результата моделирования — значения целевого концепта (рейтинга университета) от его реального значения на исторических данных является наименьшим у когнитивной модели на основе прогнозирующего ансамбля (табл. 4).

Средневзвешенная оценка состояния целевого концепта на основе ансамбля серых

НКК является более предпочтительной, т.к. она имеет меньшее значение разброса при моделировании различных воздействий на латентные факторы по сравнению с отдельной серой НКК и классической НКК за счет снижения влияния фактора субъективности экспертных оценок на результаты моделирования с использованием серых НКК, а также уменьшения ошибки моделирования за счет ансамблевой организации когнитивной модели.

Сценарный анализ вариантов развития ситуации позволил выбрать наиболее предпочтительный вариант, который в условиях заданных ограничений обеспечивает достижение требуемого планового значения целевого показателя при минимальных затратах ресурсов на приращение латентных факторов.

Заключение

В ходе выполнения исследования разработана нечеткая когнитивная модель сценарного прогнозирования мероприя-

тий по достижению требуемых значений целевых показателей деятельности университета в международном институциональном рейтинге QS на основе ансамбля серых НКК. Применение серых НКК позволило перейти от точечных оценок мнений экспертов о силе взаимосвязи между концептами к интервальным оценкам, что позволило более полно учесть имеющиеся данные и снизить фактор неопределенности при разбросе мнений экспертов.

Для идентификации наиболее значимых латентных факторов использован подход на основе методов факторного анализа. Для повышения точности и достоверности результатов прогнозирования разработана нечеткая когнитивная модель в виде прогнозирующего ансамбля серых НКК, отличительной особенностью которого является наглядность, высокие интерпретационные свойства процесса и промежуточных/итоговых результатов нечеткого когнитивного моделирования, а также адекватный учет различных типов неопределенностей в рамках единой когнитивной модели.

Практическая значимость полученных результатов заключается в возможности выбора на основе выполненных расчетов наиболее предпочтительного сценария приращения латентных факторов для достижения требуемых значений целевых показателей развития университета.

Литература

1. Аксельрод Роберт М. Структура решения: когнитивные карты политических элит. Принстон, Нью-Джерси: Princeton University Press, 1976. 404 с.
2. Ярушев С.А., Аверкин А.Н. Модульная система прогнозирования на основе нечетких когнитивных карт и нейронечетких сетей. В 7-й Всероссийской научно-практической конф. Нечеткие системы, мягкие вычисления и интеллектуальные технологии. Санкт-Петербург: Политехника — сервис, 2017. Т. 1. С. 180–189.

3. Кузнецов О.П. Когнитивное моделирование слабополуструктурированных ситуаций [Электрон. ресурс]. Режим доступа: <http://posp.raai.org/data/posp2005/Kuznetsov/kuznetsov.html>. (Дата обращения: 12.10.2021).

4. Робертс Ф.С. Дискретные математические модели с приложениями к социальным, биологическим и экологическим проблемам. М.: Наука, 1986. 312 с.

5. Carvalho J.P., Tom J.A.B.: Rule-Based Fuzzy Cognitive Maps — Fuzzy Causal Relationships. Computational Intelligence for Modeling, Control and Automation: Evolutionary Computing and

Fuzzy Logic for Intelligent Control, Knowledge, and Information Retrieval, edited by M. Mohammadyan, IOS Press. 1999. С. 102–119.

6. Силов В.Б. Принятие стратегических решений в нечеткой среде. М.: ИНПРО-РЭС, 1995. 228 с.

7. Горелова Г.В., Захарова Е.Н., Родченко С.А. Исследование малоструктурных проблем социально-экономических систем: когнитивный подход. Ростов н/д: РГУ, 2006. 332 с.

8. Борисов В.В., Круглов В.В., Федулов А.С. Нечеткие модели и сети. 2-е изд. стереотипы. М.: Горячая линия – Телеком, 2015. 284 с.

9. Борисов В.В., Луферов В.С. Метод многомерного анализа и прогнозирования состояния сложных систем и процессов на основе нечетких когнитивных темпоральных моделей // Системы управления, связи и безопасности. 2020. № 2. С. 1–23.

10. Kosko B. Fuzzy Cognitive Maps // International Journal of Man-Machine Studies. 1986. Т. 24. С. 65–75.

11. Chun Yan Miao., Xue Hong Tao., Zhi Qi Shen., Zhi Qiang Liu. Transformation of cognitive maps // IEEE Transactions on Fuzzy Systems. 2010. Т. 18. № 1. С. 114–124.

12. Salmeron J.L., Palos-Sanchez P.R.: Uncertainty propagation in fuzzy grey cognitive maps with Hebbian-like learning algorithms // IEEE transactions on cybernetics. 2017. № 49(1). С. 211–220.

13. Соколов Г.А. Введение в регрессионный анализ и планирование регрессионных экспериментов в экономике. М.: Инфра-М, 2016. 352 с.

14. Микрюков А.А., Гаспарян М.С., Карпов Д.С. Разработка предложений по продвижению вуза в международном институциональном рейтинге QS на основе методов статистическо-

го анализа // Статистика и экономика. 2020. № 17(1). С. 35–43.

15. Терещенко О.В., Курилович Е.И., Князева И.А. Многомерный статистический анализ данных в социальных науках. Минск: БГУ, 2012. 239 с.

16. Авдеева З.К., Коврига С.В., Макаренко Д.И., Максимов В.И. Когнитивный подход в управлении // Проблемы управления. 2007. Т. 3. С. 2–8.

17. Болотова Л.С. Системы искусственного интеллекта: модели и технологии, основанные на знаниях. М.: Финансы и статистика, 2012. 664 с.

18. Международный рейтинг университетов QS World University Rankings [Электрон. ресурс]. Режим доступа: <https://www.qs.com/ranking>. (Дата обращения: 12.10.2021).

19. Zhou Z. –H. Ensemble Methods: Foundations and algorithms // Chapman & Hall/Crc Machine Learning & Pattern Recognition. 2012. С. 123–146.

20. Kuncheva L.I. Combining Pattern Classifiers: Methods and algorithms. Hoboken, NJ: JohnWiley&Sons, 2004. С. 96–111.

21. Терехов С.А. Блестящие комитеты умных машин. В IX Российской научно-технической конференции «Нейроинформатика-2007»: Лекции по нейроинформатике. М.: МИФИ, 2007. С. 11–42.

22. Воронцов К.В. Лекции по алгоритмическим композициям [Электрон. ресурс]. Режим доступа: <http://www.ccas.ru/voron/download/Composition.pdf>. (Дата обращения: 16.10.2021).

23. Гончаров М. Модельные ансамбли [Электрон. ресурс]. Режим доступа: <http://www.businessdataanalytics.ru/download/ModelEnsembles.pdf>. (Дата обращения: 10.2021).

24. Боровиков В.П. Нейронные сети. Статистические нейронные сети. Методология и технологии современного анализа данных. 2-е изд. М.: Горячая линия – Телеком, 2008. 392 с.

References

1. Aksel'rod Robert M. Struktura resheniya: kognitivnyye karty politicheskikh elit = Decision structure: cognitive maps of political elites. Princeton, NJ: Princeton University Press; 1976. 404 p.

2. Yarushev S.A., Averkin A.N. Modular forecasting system based on fuzzy cognitive maps and neuro-fuzzy networks. V 7-y Vserossiyskoy nauchno-prakticheskoy konf. Nechetkiye sistemy, myagkiye vychisleniya i intellektual'nyye tekhnologii = In the 7th All-Russian Scientific and Practical Conf. Fuzzy systems, soft computing and intelligent technologies. St. Petersburg: Polytechnic – service; 2017; 1: 180–189. (In Russ.)

3. Kuznetsov O. P. Kognitivnoye modelirovaniye slabopolustrukturirovannykh situatsiy = Cognitive modeling of weakly semi-structured situations [Internet]. Available from: <http://posp.raai.org/data/posp2005/Kuznetsov/kuznetsov.html>. (cited 12.10.2021). (In Russ.)

4. Roberts F.S. Diskretnyye matematicheskiye modeli s prilozheniyami k sotsial'nyim, biologicheskim i ekologicheskim problemam = Discrete mathematical models with applications to social, biological and ecological problems. Moscow: Nauka; 1986. 312 p. (In Russ.)

5. Carvalho J.P., Tom J.A.B.: Rule-Based Fuzzy Cognitive Maps – Fuzzy Causal Relationships. Computational Intelligence for Modeling, Control and Automation: Evolutionary Computing and Fuzzy Logic for Intelligent Control, Knowledge, and Information Retrieval, edited by M. Mohammadyan, IOS Press. 1999: 102–119.

6. Silov V.B. Prinyatiye strategicheskikh resheniy v nechetkoy srede = Making strategic decisions in a fuzzy environment. Moscow: INPRO-RES; 1995. 228 p. (In Russ.)

7. Gorelova G.V., Zakharova Ye.N., Rodchenko S.A. Issledovaniye malostrukturnykh problem sotsial'no-ekonomicheskikh sistem: kognitivnyy pod-

khod = The study of low-structural problems of socio-economic systems: a cognitive approach. Rostov on Don: RGU; 2006. 332 p. (In Russ.)

8. Borisov V.V., Kruglov V.V., Fedulov A.S. Nechetkiye modeli i seti. 2-ye izd. Stereotypy = Fedulov A.S. Fuzzy models and networks. 2nd ed. stereotypes. Moscow: Hotline – Telecom; 2015. 284 p. (In Russ.)

9. Borisov V.V., Luferov V.S. The method of multidimensional analysis and forecasting of the state of complex systems and processes based on fuzzy cognitive temporal models. Sistemy upravleniya, svyazi i bezopasnosti = Control Systems. 2020; 2: 1–23. (In Russ.)

10. Kosko B. Fuzzy Cognitive Maps. International Journal of Man-Machine Studies. 1986; 24: 65–75.

11. Chun Yan Miao., Xue Hong Tao., Zhi Qi Shen., Zhi Qiang Liu. Transformation of cognitive maps. IEEE Transactions on Fuzzy Systems. 2010; 18; 1: 114–124.

12. Salmeron J.L., Palos-Sanchez P.R.: Uncertainty propagation in fuzzy grey cognitive maps with Hebbian-like learning algorithms. IEEE transactions on cybernetics. 2017; 49(1): 211–220.

13. Sokolov G.A. Vvedeniye v regressionnyy analiz i planirovaniye regressionnykh eksperimentov v ekonomike = Introduction to regression analysis and planning of regression experiments in economics. Moscow: Infra-M; 2016. 352 p. (In Russ.)

14. Mikryukov A.A., Gasparyan M.S., Karpov D.S. Development of proposals for the promotion of the university in the international institutional ranking QS based on the methods of statistical analysis. Statistika i ekonomika = Statistics and Economics. 2020; 17(1): 35–43. (In Russ.)

15. Tereshchenko O.V., Kurilovich Ye.I., Kn-yazeva I.A. Mnogomernyy statisticheskiy analiz dannykh v sotsial'nykh naukakh = Multivariate statistical data analysis in the social sciences. Minsk: BGU; 2012. 239 p.

16. Avdeyeva Z.K., Kovriga S.V., Makarenko D.I., Maksimov V.I. Cognitive approach in

management. Problemy upravleniya = Problems of management. 2007; 3: 2–8. (In Russ.)

17. Bolotova. L.S. Sistemy iskusstvennogo intell-ekta: modeli i tekhnologii, osnovannyye na znani-yakh = Artificial intelligence systems: models and technologies based on knowledge. Moscow: Finance and statistics; 2012. 664 p. (In Russ.)

18. Mezhdunarodnyy reyting universitetov QS World University Rankings = QS World University Rankings [Internet]. Available from: <https://www.qs.com/ranking>. (cited 12.10.2021).

19. Zhou Z. –H. Ensemble Methods: Foundations and algorithms. Chapman & Hall/Crc Machine Learning & Pattern Recognition; 2012: 123–146.

20. Kuncheva L.I. Combining Pattern Classifiers: Methods and algorithms. Hoboken, NJ: John-Wiley&Sons; 2004: 96–111.

21. Terekhov S.A. Blestyashchiye komitety umnykh mashin. V IX Rossiyskoy nauchno-tekhnicheskoy konferentsii “Neyroinformatika-2007”: Lekt-sii po neyroinformatike = Brilliant committees of intelligent machines. In the IX Russian scientific and technical conference “Neuroinformatics-2007”: Lectures on neuroinformatics. Moscow: MEPhI; 2007: 11–42. (In Russ.)

22. Vorontsov K.V. Leksii po algoritmicheskim kompozitsiyam = Lectures on algorithmic compositions [Internet]. Available from: <http://www.ccas.ru/voron/download/Composition.pdf>. (cited 16.10.2021). (In Russ.)

23. Goncharov M. Model'nyye ansambli = Model Ensembles [Internet]. Available from: <http://www.businessdataanalytics.ru/download/Model-ensembles.pdf>. (cited 10.2021). (In Russ.)

24. Borovikov V.P. Neyronnyye seti. Statistich-eskiye neyronnyye seti. Metodologiya i tekhnologii sovremennogo analiza dannykh. 2-ye izd = Neural networks. Statistical neural networks. Methodology and technologies of modern data analysis. 2nd ed. Moscow: Hot line - Telecom; 2008. 392 p. (In Russ.)

Сведения об авторах

Андрей Александрович Микрюков

К.т.н., доцент

Российский экономический университет
им. Г.В. Плеханова, Москва, Россия

Эл. почта: mikrakov.aa@rea.ru

Михаил Ефимович Мазуров

Д.ф.-м.н., профессор

Российский экономический университет
им. Г.В. Плеханова, Москва, Россия

Эл. почта: Mazurov37@mail.ru

Information about the authors

Andrey A. Mikryukov

Cand. Sci. (Engineering), Associate Professor
Plekhanov Russian University of Economics,
Moscow, Russia

E-mail: Mikrakov.aa@rea.ru

Mikhail E. Mazurov

Dr. Sci. (Physics and Mathematics), Professor
Plekhanov Russian University of Economics,
Moscow, Russia

E-mail: Mazurov37@mail.ru