



RUSSIAN TECHNOLOGICAL JOURNAL

**РОССИЙСКИЙ
ТЕХНОЛОГИЧЕСКИЙ
ЖУРНАЛ**

*Information systems.
Computer sciences.
Issues of information security*

*Multiple robots (robotic centers) and systems.
Remote sensing and nondestructive testing*

Modern radio engineering and telecommunication systems

*Micro- and nanoelectronics.
Condensed matter physics*

Analytical instrument engineering and technology

Mathematical modeling

*Economics of knowledge-intensive and high-tech enterprises and industries.
Management in organizational systems*

Product quality management. Standardization

Philosophical foundations of technology and society



RUSSIAN TECHNOLOGICAL JOURNAL

**РОССИЙСКИЙ
ТЕХНОЛОГИЧЕСКИЙ
ЖУРНАЛ**

- Information systems. Computer sciences. Issues of information security
- Multiple robots (robotic centers) and systems. Remote sensing and nondestructive testing
- Modern radio engineering and telecommunication systems
- Micro- and nanoelectronics. Condensed matter physics
- Analytical instrument engineering and technology
- Mathematical modeling
- Economics of knowledge-intensive and high-tech enterprises and industries. Management in organizational systems
- Product quality management. Standardization
- Philosophical foundations of technology and society

- Информационные системы. Информатика. Проблемы информационной безопасности
- Роботизированные комплексы и системы. Технологии дистанционного зондирования и неразрушающего контроля
- Современные радиотехнические и телекоммуникационные системы
- Микро- и нанoeлектроника. Физика конденсированного состояния
- Аналитическое приборостроение и технологии
- Математическое моделирование
- Экономика наукоемких и высокотехнологичных предприятий и производств. Управление в организационных системах
- Управление качеством продукции. Стандартизация
- Мировоззренческие основы технологии и общества

Russian Technological Journal
2026, Vol. 14, No. 4

Russian Technological Journal
2026, том 14, № 4

Russian Technological Journal
2026, Vol. 14, No. 4

Publication date July 31, 2026.

The peer-reviewed scientific and technical journal highlights the issues of complex development of radio engineering, telecommunication and information systems, electronics and informatics, as well as the results of fundamental and applied interdisciplinary researches, technological and economical developments aimed at the development and improvement of the modern technological base.

Periodicity: six times a year.

The journal was founded in December 2013. The titles were «Herald of MSTU MIREA» until 2016 (ISSN 2313-5026) and «Rossiiskii tekhnologicheskii zhurnal» from January 2016 until July 2021 (ISSN 2500-316X).

Founder and Publisher:

Federal State Budget
Educational Institution of Higher Education
«MIREA – Russian Technological University»
78, Vernadskogo pr., Moscow, 119454 Russia.

The journal is included in the White List (Level 1),
List of peer-reviewed science press of the State
Commission for Academic Degrees and Titles of Russian
Federation, Russian Science Citation Index (RSCI),
Russian State Library (RSL), Science Index,
eLibrary, Directory of Open Access Journals (DOAJ),
Directory of Open Access Scholarly Resources (ROAD),
Google Scholar, Ulrich's International Periodicals Directory.

Editor-in-Chief:

Alexander S. Sigov, Academician at the Russian Academy
of Sciences, Dr. Sci. (Phys.–Math.), Professor,
President of MIREA – Russian Technological University
(RTU MIREA), Moscow, Russia. Scopus Author ID 35557510600,
ResearcherID L-4103-2017, sigov@mirea.ru.

Editorial staff:

Managing Editor	Cand. Sci. (Eng.) Galina D. Seredina
Scientific Editor	Dr. Sci. (Eng.), Prof. Gennady V. Kulikov
Executive Editor	Anna S. Alekseenko
Technical Editor	Darya V. Trofimova
Head of Department	Elizaveta I. Kuricheva

86, Vernadskogo pr., Moscow, 119571 Russia.
Phone: +7 (499) 600-80-80 (#31288).
E-mail: seredina@mirea.ru.

The registration number ПИ № ФС 77 - 81733
was issued in August 19, 2021
by the Federal Service for Supervision
of Communications, Information Technology,
and Mass Media of Russia.

The subscription index of *Pressa Rossii*: 79641.

<https://www.rtfj-mirea.ru>

Russian Technological Journal
2026, том 14, № 4

Дата опубликования 31 июля 2026 г.

Научно-технический рецензируемый журнал освещает вопросы комплексного развития радиотехнических, телекоммуникационных и информационных систем, электроники и информатики, а также результаты фундаментальных и прикладных междисциплинарных исследований, технологических и организационно-экономических разработок, направленных на развитие и совершенствование современной технологической базы.

Периодичность: 6 раз в год.

Журнал основан в декабре 2013 года. До 2016 г. издавался под названием «Вестник МГТУ МИРЭА» (ISSN 2313-5026), а с января 2016 г. по июль 2021 г. под названием «Российский технологический журнал» (ISSN 2500-316X).

Учредитель и издатель:

федеральное государственное бюджетное
образовательное учреждение высшего образования
«МИРЭА – Российский технологический университет»
119454, РФ, г. Москва, пр-т Вернадского, д. 78.

Журнал включен в Единый государственный перечень
научных изданий — «Белый список» (уровень 1),
Перечень ведущих рецензируемых научных журналов
ВАК Российской Федерации, входит в РИНЦ,
Russian Science Citation Index (RSCI), РГБ, eLibrary,
Directory of Open Access Journals (DOAJ),
Directory of Open Access Scholarly Resources (ROAD),
Google Scholar, Ulrich's International Periodicals Directory.

Главный редактор:

Сигов Александр Сергеевич, академик РАН,
доктор физ.-мат. наук, профессор, президент ФГБОУ ВО
МИРЭА – Российский технологический университет
(ПТУ МИРЭА), Москва, Россия. Scopus Author ID
35557510600, ResearcherID L-4103-2017, sigov@mirea.ru.

Редакция:

Зав. редакцией	к.т.н. Г.Д. Середина
Научный редактор	д.т.н., проф. Г.В. Куликов
Выпускающий редактор	А.С. Алексеенко
Технический редактор	Д.В. Трофимова
Начальник отдела	Е.И. Куричева

119571, г. Москва, пр-т Вернадского, 86, оф. П-108.
Тел.: +7 (499) 600-80-80 (#31288).
E-mail: seredina@mirea.ru.

Регистрационный номер и дата принятия решения
о регистрации СМИ ПИ № ФС 77 - 81733 от 19.08.2021 г.
СМИ зарегистрировано Федеральной службой по надзору
в сфере связи, информационных технологий и массовых
коммуникаций (Роскомнадзор).

Индекс по объединенному каталогу «Пресса России» 79641.

Editorial Board

Stanislav A. Kudzh	Dr. Sci. (Eng.), Professor, Rector of RTU MIREA, Moscow, Russia. Scopus Author ID 56521711400, ResearcherID AAG-1319-2019, https://orcid.org/0000-0003-1407-2788 , rector@mirea.ru
Juras Banys	Habilitated Doctor of Sciences, Professor, Vice-Rector of Vilnius University, Vilnius, Lithuania. Scopus Author ID 7003687871, juras.banys@ff.vu.lt
Vladimir B. Betelin	Academician at the Russian Academy of Sciences (RAS), Dr. Sci. (Phys.-Math.), Professor, Supervisor of Scientific Research Institute for System Analysis, RAS, Moscow, Russia. Scopus Author ID 6504159562, ResearcherID J-7375-2017, betelin@niisi.msk.ru
Alexei A. Bokov	Dr. Sci. (Phys.-Math.), Senior Research Fellow, Department of Chemistry and 4D LABS, Simon Fraser University, Vancouver, British Columbia, Canada. Scopus Author ID 35564490800, ResearcherID C-6924-2008, http://orcid.org/0000-0003-1126-3378 , abokov@sfu.ca
Sergey B. Vakhrushev	Dr. Sci. (Phys.-Math.), Professor, Head of the Laboratory of Neutron Research, A.F. Ioffe Physico-Technical Institute of the RAS, Department of Physical Electronics of St. Petersburg Polytechnic University, St. Petersburg, Russia. Scopus Author ID 7004228594, ResearcherID A-9855-2011, http://orcid.org/0000-0003-4867-1404 , s.vakhrushev@mail.ioffe.ru
Yury V. Gulyaev	Academician at the RAS, Dr. Sci. (Phys.-Math.), Professor, Academic Supervisor of V.A. Kotelnikov Institute of Radio Engineering and Electronics of the RAS, Moscow, Russia. Scopus Author ID 35562581800, gulyaev@cplire.ru
Dmitry O. Zhukov	Dr. Sci. (Eng.), Professor of the Department of Telecommunications, Institute of Radio Electronics and Informatics, RTU MIREA, Moscow, Russia. Scopus Author ID 57189660218, zhukov_do@mirea.ru
Alexey V. Kimel	PhD (Phys.-Math.), Professor, Radboud University, Nijmegen, Netherlands, Scopus Author ID 6602091848, ResearcherID D-5112-2012, a.kimel@science.ru.nl
Sergey O. Kramarov	Dr. Sci. (Phys.-Math.), Professor, Surgut State University, Surgut, Russia. Scopus Author ID 56638328000, ResearcherID E-9333-2016, https://orcid.org/0000-0003-3743-6513 , mavoo@yandex.ru
Dmitry A. Novikov	Academician at the RAS, Dr. Sci. (Eng.), Director of V.A. Trapeznikov Institute of Control Sciences, Moscow, Russia. Scopus Author ID 7102213403, ResearcherID Q-9677-2019, https://orcid.org/0000-0002-9314-3304 , novikov@ipu.ru
Philippe Pernod	Dr. Sci. (Electronics), Professor, Dean of Research of Centrale Lille, Villeneuve-d'Ascq, France. Scopus Author ID 7003429648, philippe.pernod@ec-lille.fr
Mikhail P. Romanov	Dr. Sci. (Eng.), Professor, Academic Supervisor of the Institute of Artificial Intelligence, RTU MIREA, Moscow, Russia. Scopus Author ID 14046079000, https://orcid.org/0000-0003-3353-9945 , m_romanov@mirea.ru
Viktor P. Savinykh	Academician at the RAS, Dr. Sci. (Eng.), Professor, President of Moscow State University of Geodesy and Cartography, Moscow, Russia. Scopus Author ID 56412838700, vp@miigaik.ru
Andrei N. Sobolevski	Professor, Dr. Sci. (Phys.-Math.), Director of Institute for Information Transmission Problems (Kharkevich Institute), Moscow, Russia. Scopus Author ID 7004013625, ResearcherID D-9361-2012, http://orcid.org/0000-0002-3082-5113 , sobolevski@iitp.ru
Li Da Xu	Academician at the European Academy of Sciences, Russian Academy of Engineering (formerly, USSR Academy of Engineering), and Armenian Academy of Engineering, Dr. Sci. (Systems Science), Professor and Eminent Scholar in Information Technology and Decision Sciences, Old Dominion University, Norfolk, VA, the United States of America. Scopus Author ID 13408889400, https://orcid.org/0000-0002-5954-5115 , lxu@odu.edu
Yury S. Kharin	Academician at the National Academy of Sciences of Belarus, Dr. Sci. (Phys.-Math.), Professor, Director of the Institute of Applied Problems of Mathematics and Informatics of the Belarusian State University, Minsk, Belarus. Scopus Author ID 6603832008, http://orcid.org/0000-0003-4226-2546 , kharin@bsu.by
Yuri A. Chaplygin	Academician at the RAS, Dr. Sci. (Eng.), Professor, Member of the Departments of Nanotechnology and Information Technology of the RAS, President of the National Research University of Electronic Technology (MIET), Moscow, Russia. Scopus Author ID 6603797878, ResearcherID B-3188-2016, president@miet.ru
Vasily V. Shpak	Cand. Sci. (Econ.), Deputy Minister of Industry and Trade of the Russian Federation, Ministry of Industry and Trade of the Russian Federation, Moscow, Russia; Associate Professor, National Research University of Electronic Technology (MIET), Moscow, Russia, mishinevaiv@minprom.gov.ru

Редакционная коллегия

Кудж Станислав Алексеевич	д.т.н., профессор, ректор РТУ МИРЭА, Москва, Россия. Scopus Author ID 56521711400, ResearcherID AAG-1319-2019, https://orcid.org/0000-0003-1407-2788 , rector@mirea.ru
Банис Юрас Йонович	хабилированный доктор наук, профессор, проректор Вильнюсского университета, Вильнюс, Литва. Scopus Author ID 7003687871, juras.banys@ff.vu.lt
Бетелин Владимир Борисович	академик Российской академии наук (РАН), д.ф.-м.н., профессор, научный руководитель Федерального научного центра «Научно-исследовательский институт системных исследований» РАН, Москва, Россия. Scopus Author ID 6504159562, ResearcherID J-7375-2017, betelin@niisi.msk.ru
Боков Алексей Алексеевич	д.ф.-м.н., старший научный сотрудник, химический факультет и 4D LABS, Университет Саймона Фрейзера, Ванкувер, Британская Колумбия, Канада. Scopus Author ID 35564490800, ResearcherID C-6924-2008, http://orcid.org/0000-0003-1126-3378 , abokov@sfu.ca
Вахрушев Сергей Борисович	д.ф.-м.н., профессор, заведующий лабораторией нейтронных исследований Физико-технического института им. А.Ф. Иоффе РАН, профессор кафедры Физической электроники СПбГПУ, Санкт-Петербург, Россия. Scopus Author ID 7004228594, ResearcherID A-9855-2011, http://orcid.org/0000-0003-4867-1404 , s.vakhrushev@mail.ioffe.ru
Гуляев Юрий Васильевич	академик РАН, д.ф.-м.н., профессор, научный руководитель Института радиотехники и электроники им. В.А. Котельникова РАН, Москва, Россия. Scopus Author ID 35562581800, gulyaev@cplire.ru
Жуков Дмитрий Олегович	д.т.н., профессор кафедры телекоммуникаций Института радиоэлектроники и информатики РТУ МИРЭА, Москва, Россия. Scopus Author ID 57189660218, zhukov_do@mirea.ru
Кимель Алексей Вольдемарович	к.ф.-м.н., профессор, Университет Радбауд, г. Наймерген, Нидерланды. Scopus Author ID 6602091848, ResearcherID D-5112-2012, a.kimel@science.ru.nl
Крамаров Сергей Олегович	д.ф.-м.н., профессор, Сургутский государственный университет, Сургут, Россия. Scopus Author ID 56638328000, ResearcherID E-9333-2016, https://orcid.org/0000-0003-3743-6513 , mavoo@yandex.ru
Новиков Дмитрий Александрович	академик РАН, д.т.н., директор Института проблем управления им. В.А. Трапезникова РАН, Москва, Россия. Scopus Author ID 7102213403, ResearcherID Q-9677-2019, https://orcid.org/0000-0002-9314-3304 , novikov@ipu.ru
Перно Филипп	Dr. Sci. (Electronics), профессор, Центральная Школа г. Лилль, Франция. Scopus Author ID 7003429648, philippe.pernod@ec-lille.fr
Романов Михаил Петрович	д.т.н., профессор, научный руководитель Института искусственного интеллекта РТУ МИРЭА, Москва, Россия. Scopus Author ID 14046079000, https://orcid.org/0000-0003-3353-9945 , m_romanov@mirea.ru
Савиных Виктор Петрович	академик РАН, Дважды Герой Советского Союза, д.т.н., профессор, президент Московского государственного университета геодезии и картографии, Москва, Россия. Scopus Author ID 56412838700, vp@miigaik.ru
Соболевский Андрей Николаевич	д.ф.-м.н., директор Института проблем передачи информации им. А.А. Харкевича, Москва, Россия. Scopus Author ID 7004013625, ResearcherID D-9361-2012, http://orcid.org/0000-0002-3082-5113 , sobolevski@iitp.ru
Сюй Ли Да	академик Европейской академии наук, Российской инженерной академии и Инженерной академии Армении, Dr. Sci. (Systems Science), профессор, Университет Олд Доминион, Норфолк, Соединенные Штаты Америки. Scopus Author ID 13408889400, https://orcid.org/0000-0002-5954-5115 , lxu@odu.edu
Харин Юрий Семенович	академик Национальной академии наук Беларуси, д.ф.-м.н., профессор, директор НИИ прикладных проблем математики и информатики Белорусского государственного университета, Минск, Беларусь. Scopus Author ID 6603832008, http://orcid.org/0000-0003-4226-2546 , kharin@bsu.by
Чаплыгин Юрий Александрович	академик РАН, д.т.н., профессор, член Отделения нанотехнологий и информационных технологий РАН, президент Института микроприборов и систем управления им. Л.Н. Преснухина НИУ «МИЭТ», Москва, Россия. Scopus Author ID 6603797878, ResearcherID B-3188-2016, president@miet.ru
Шпак Василий Викторович	к.э.н., зам. министра промышленности и торговли Российской Федерации, Министерство промышленности и торговли РФ, Москва, Россия; доцент, Институт микроприборов и систем управления им. Л.Н. Преснухина НИУ «МИЭТ», Москва, Россия, mishinevaiv@minprom.gov.ru

Contents

Information systems. Computer sciences. Issues of information security

- 7** *Vagid A. Kadymov, Sergey M. Dumansky, Evgeny V. Vasiliev*
Algorithm for the formation and functioning of a standard information and analytical system for multidisciplinary medical centers
- 21** *Sergey A. Lesko, Vladimir N. Kalinin, Dmitry O. Zhukov*
Estimating the overload risk of a Kubernetes node based on the Fokker–Planck equation and the first-to-critical-bound problem

Multiple robots (robotic centers) and systems. Remote sensing and non-destructive testing

- 36** *Yuri A. Bykovtsev, Valery M. Lokhin, Sergey V. Manko, Mikhail P. Romanov*
Fault diagnostics in intelligent control systems of autonomous robots: An ontological approach
- 53** *Vladimir P. Dashevsky, Maxim M. Bizin, Yuri V. Galykin*
Design and manufacturing technology of a folding conveyor for a robotic platform for interflight servicing of multicopters

Modern radio engineering and telecommunication systems

- 68** *Boris M. Vovshin, Alexander A. Pushkov*
Synthesis of matched multichannel processing of orthogonal MIMO radar signals
- 82** *Nguyen Van Dung, Phan Trong Hanh*
Hybrid neural network architecture ResBiLSTM-BiGRU for efficient noise suppression in radio channels with digitally modulated signals
- 96** *Sergey S. Shadinov, Mihail S. Kostin*
Tracking-based formation of vector radio images

Micro- and nanoelectronics. Condensed matter physics

- 106** *Stanislava V. Doronkina, Anastasia S. Babayants, Ivan V. Lisitsa, Anna E. Myasnikova*
Photoemission modeling of a two-liquid charge-carrier system formed at strong electron–phonon coupling

Mathematical modeling

- 116** *Kristina V. Gubareva, Evgenii Yu. Prosviryakov, Anton V. Eremin*
Numerical simulation of Couette flow in a micropolar fluid using the projection method
- 125** *Vladimir A. Ivanov, Sergey G. Paramonov, Ekaterina S. Zhigacheva, Ekaterina A. Pozdnyakova*
Modernization of a statistical data processing algorithm for integrated background monitoring
- 137** *Alexey N. Nazarov*
Stochastic properties of the dynamics of the first derivative in Cauchy problems

Содержание

Информационные системы. Информатика. Проблемы информационной безопасности

- 7** *В.А. Кадымов, С.М. Думанский, Е.В. Васильев*
Алгоритм формирования и функционирования типовой информационно-аналитической системы для многопрофильных медицинских центров
- 21** *С.А. Лесько, В.Н. Калинин, Д.О. Жуков*
Оценивание риска перегрузки вычислительного узла в Kubernetes на основе уравнения Фоккера – Планка и задачи первого достижения критической границы

Роботизированные комплексы и системы. Технологии дистанционного зондирования неразрушающего контроля

- 36** *Ю.А. Быковцев, В.М. Лохин, С.В. Манько, М.П. Романов*
Диагностика неисправностей в интеллектуальных системах управления автономными роботами на основе онтологического подхода
- 53** *В.П. Дашевский, М.М. Бизин, Ю.В. Галыкин*
Конструкция и технология изготовления складного конвейера роботизированной платформы межполетного обслуживания мультикоптеров

Современные радиотехнические и телекоммуникационные системы

- 68** *Б.М. Вовшин, А.А. Пушков*
Синтез согласованной многоканальной обработки ортогональных сигналов ММО радиолокационных станций
- 82** *В.З. Нгуен, Ч.Х. Фан*
Гибридная архитектура нейронной сети ResBiLSTM-BiGRU для эффективного подавления шума в радиоканале с цифровыми модулированными сигналами
- 96** *С.С. Шадинов, М.С. Костин*
Трекинг-формирование векторных радиоизображений

Микро- и наноэлектроника. Физика конденсированного состояния

- 106** *С.В. Доронкина, А.С. Бабаянц, И.В. Лисица, А.Э. Мясникова*
Моделирование фотоэмиссии двухжидкостной системы носителей заряда, формирующейся при сильном электрон-фононном взаимодействии

Математическое моделирование

- 116** *К.В. Губарева, Е.Ю. Просвираков, А.В. Еремин*
Численное моделирование течения Куэтта микрополярной жидкости методом проекции
- 125** *В.А. Иванов, С.Г. Парамонов, Е.С. Жигачева, Е.А. Позднякова*
Модернизация алгоритма статистической обработки данных сети комплексного фонового мониторинга
- 137** *А.Н. Назаров*
Стохастические свойства динамики первой производной в задачах Коши

UDC 004.043

<https://doi.org/10.32362/2500-316X-2026-14-4-7-20>

EDN BUFYXO



RESEARCH ARTICLE

Algorithm for the formation and functioning of a standard information and analytical system for multidisciplinary medical centers

Vagid A. Kadymov¹, Sergey M. Dumansky^{2, @}, Evgeny V. Vasiliev²¹ MIREA – Russian Technological University, Moscow, 119454 Russia² State Research Center – Burnasyan Federal Medical Biophysical Center of Federal Medical Biological Agency, Moscow, 123098 Russia

@ Corresponding author, e-mail: s.dumansky.m@gmail.com

• Submitted: 23.12.2025 • Revised: 25.02.2026 • Accepted: 26.05.2026

Abstract

Objectives. The widespread use of computed tomography (CT) for diagnostic and treatment purposes can lead to excessive cumulative doses of ionizing radiation received by patients in medical institutions. Currently, there are no effective methods of visualization and control of cumulative doses of patients in medical institutions. Thus, there is a need for operational monitoring of the dose load from ionizing radiation. The purpose of this work is to present a developed algorithm for the creation and functioning of an information and analytical system (IAS) for the retrospective accounting, analysis and visualization of cumulative doses of ionizing radiation received by patients undergoing radiation diagnosis and treatment.

Methods. When forming visual output forms for accounting, analyzing and identifying critical values of the dose load of ionizing radiation and the number of CT scans of patients, structured query language (SQL) was used to select the necessary data in order to apply them in summary tables and volumetric diagrams.

Results. An algorithm has been developed for the formation and functioning of IAS based on client-server technology using summary tables and volumetric diagrams in a client application. The algorithm can be used to visualize the analysis and accounting of cumulative doses of ionizing radiation in patients undergoing treatment and medical examination in large multifunctional medical research hospitals, identify the types of CT scans for reducing the number of sessions and limit the growth of cumulative doses of ionizing radiation in patients, as well as to detect errors in data entry by medical professionals into medical information systems and provide timely correction. In addition, an unlimited number of forms can be created for viewing, analyzing, and accounting for data.

Conclusions. The developed methodology for the formation of IAS can be scaled up to other multidisciplinary medical centers for tracking critical values of cumulative doses of ionizing radiation in patients in order to apply the necessary medical solutions in radiation diagnosis.

Keywords: information and analytical systems, client-server technologies, pivot tables, SQL

For citation: Kadymov V.A., Dumansky S.M., Vasiliev E.V. Algorithm for the formation and functioning of a standard information and analytical system for multidisciplinary medical centers. *Russian Technological Journal*. 2026;14(4):7–20. <https://doi.org/10.32362/2500-316X-2026-14-4-7-20>, <https://elibrary.ru/BUFYXO>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Алгоритм формирования и функционирования типовой информационно-аналитической системы для многопрофильных медицинских центров

В.А. Кадымов ¹, С.М. Думанский ^{2, @}, Е.В. Васильев ²

¹ МИРЭА – Российский технологический университет, Москва, 119454 Россия

² ФГБУ ГНЦ ФМБЦ им. А.И. Бурназяна ФМБА России, Москва, 123098 Россия

@ Автор для переписки, e-mail: s.dumansky.m@gmail.com

• Поступила: 23.12.2025 • Доработана: 25.02.2026 • Принята к опубликованию: 26.05.2026

Резюме

Цели. Широкое применение компьютерной томографии (КТ) в целях диагностики и лечения может приводить к избыточным кумулятивным дозам ионизирующего излучения, получаемым пациентами лечебных учреждений. В медицинских учреждениях на данный момент отсутствуют эффективные методы визуализации и контроля за кумулятивными дозами пациентов. Существует необходимость оперативного контроля за дозовой нагрузкой от ионизирующего излучения. Целью данной работы является разработка алгоритма функционирования информационно-аналитической системы (ИАС) для ретроспективного учета, анализа и визуализации кумулятивных доз ионизирующего излучения, получаемых пациентами, проходящими лучевую диагностику и лечение.

Методы. Используются SQL¹-запросы для выборки необходимых данных с целью применения их в сводных таблицах и объемных диаграммах при формировании выходных наглядных форм учета, анализа и выявления критических величин дозовой нагрузки ионизирующего излучения и количества КТ-исследований пациентов.

Результаты. Разработан алгоритм формирования и функционирования ИАС на основе клиент-серверной технологии с использованием сводных таблиц и объемных диаграмм в клиентском приложении, которая обеспечивает, во-первых, визуализацию анализа и учета кумулятивных доз ионизирующего излучения пациентов, проходящих лечение и диспансеризацию в крупных многофункциональных лечебно-исследовательских стационарах. Во-вторых, данный алгоритм позволяет выявлять виды КТ-исследований, по которым возможно уменьшить число применяемых сеансов с целью ограничения роста кумулятивных доз ионизирующего излучения пациентов. В-третьих, алгоритм дает возможность обнаруживать ошибки ввода данных медицинскими работниками в медицинские информационные системы и вовремя их корректировать. В-четвертых, можно формировать неограниченное число форм для просмотра данных с целью проведения их анализа и учета.

Выводы. Разработанную методику формирования ИАС возможно масштабировать в другие медицинские многопрофильные центры для отслеживания критических величин кумулятивных доз ионизирующего излучения пациентов в целях применения необходимых медицинских решений при лучевой диагностике пациентов.

Ключевые слова: информационно-аналитические системы, клиент-серверные технологии, сводные таблицы, SQL-запросы

¹ Structured query language – язык структурированных запросов.

Для цитирования: Кадымов В.А., Думанский С.М., Васильев Е.В. Алгоритм формирования и функционирования типовой информационно-аналитической системы для многопрофильных медицинских центров. *Russian Technological Journal*. 2026;14(4):7–20. <https://doi.org/10.32362/2500-316X-2026-14-4-7-20>, <https://elibrary.ru/BUFYXO>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Multidisciplinary medical centers and hospitals utilize medical information systems (MIS) to record patient data across various medical procedures and treatment areas. Due to the inability of MIS to retrospectively record and analyze data over long time periods, significant volumes of accumulated information are not fully utilized. The present work presents an algorithm for the development and operation of an information and analytical system (IAS) on the example of computed tomography (CT). The algorithm is used for organizing retrospective recording and analysis of cumulative doses of ionizing radiation received by patients along with the number of CT sessions during radiation diagnostics.

Due to the widespread use of CT, the radiation dose to patients from ionizing radiation has increased significantly, as reflected in numerous publications on this topic. In [1], it is noted that higher doses received from CT in children throughout life lead to an increased estimated risk of death from cancer compared to adults. In [2], it is indicated that intravenous administration of a contrast enhancer during CT for the detection of kidney stones increases the total radiation dose, which must be taken into account when determining the radiation dose to patients. In [3], an increased risk of cancer in children associated with the use of CT for diagnostic purposes is identified due to frequent repeat examinations, as well as the fact that children are more sensitive to the carcinogenic effects of ionizing radiation than adults and a longer life expectancy. In [4], an increased risk of hematological neoplasms in children and adolescents is identified as a result of CT irradiation along with the need for careful justification of CT in this age group. The work [5] provides information on long-term oncological consequences among residents of the city of Ozersk (Chelyabinsk oblast), who were examined using CT for a comprehensive assessment of the effects of low doses associated with diagnostic exposure in a cohort study. In [6], an attempt was made to answer the question of whether patients who were repeatedly examined with a CT scanner for indications related to the diagnosis of oncological pathology are at risk of developing late effects during subsequent years of observation. In [7], the need for centralized recording

of ionizing radiation doses in the electronic medical records of patients undergoing diagnostic and radiation therapy is indicated. The article [8] describes simulation modeling methods used to determine the organs and tissues exposed to the greatest radiation exposure during CT scanning of the chest, abdomen, and head. In [9], an overview of key Russian and foreign studies of the possible adverse effects of medical radiation on human health is presented along with a consideration of existing methods for assessing the possible harm to human health as a result of medical radiological diagnostic procedures. The work [10] indicates that the impact of ionizing radiation during CT scanning is a source of potential risk associated with the development of radiation-induced carcinogenic effects. In [11], methods of using information technology for retrospective analysis of the frequency of detection of neoplasms in patients are described using the example of a large multidisciplinary medical center. In [12], the experience of using computer technologies in visualizing data on the medical examination of patients who underwent CT diagnostics and radiation therapy is presented. The work [13] focuses on the need to develop centralized data warehouses to account for the accumulated dose load of patients in such a way that each person knows their load and can make appropriate decisions on each type of upcoming radiation diagnostics considering the associated risks. In [14], the need for strict personalized accounting of the dose load when using CT is indicated. The article [15] presents the experience of using summary tables and volumetric diagrams for data analysis. In [16] it is asserted that, since the highest doses of medical diagnostic radiation exposure of patients in Russia are caused by CT procedures, radiation risk control becomes especially relevant when using CT.

Since, according to the no-threshold radiation risk model, any dose of radiation increases the risk of cancer, it is essential to monitor received doses. Depending on the location of the exposure, the dose may be occupational, medical, natural, etc. Recently, diagnostic imaging using radionuclides to enhance contrast, X-rays, and other methods have become widespread in medicine. For example, radiotherapy (teletherapy and brachytherapy) is widely used to treat cancer. The dose received by a patient from a CT scan per session

varies from 0.3 to 63 mSv depending on the procedure. However, visualization, analysis, and recording of patient doses, particularly those undergoing diagnostic imaging, are not organized to the required standard. Medical information systems (MIS) used in healthcare institutions, such as MEDMIS², Medesk³, ArchiMed+⁴, KMIS⁵, Medialog⁶, and others, are not designed to directly implement these functions, as they operate in online transaction processing (OLTP) mode, which utilizes highly normalized linked tables in the database schema. Visualization, analysis, and accounting of medical data require a different database scenario—online analytical processing (OLAP). This requires uploading data entered into the MIS to other databases, particularly OLAP. Using the example of the Medialog MIS, a system for exporting from the MIS to a database was developed at one of the medical centers of the Federal Medical and Biological Agency of Russia⁷, which implements the above-mentioned functions.

1. ALGORITHM FOR THE FORMATION AND FUNCTIONING OF THE IAS

The algorithm consists of three stages. In the first stage, a data request is generated in the MIS, each record of which corresponds to one CT session of a specific patient with the corresponding attributes (characteristics): patient identifier, type of examination, equipment used, received ionizing radiation dose, examination date, patient date of birth, gender, ICD-10⁸ nosology code, etc. In the second stage, the resulting sample for a specified period is exported from the MIS first to an *Excel* table, from where it is imported into a structured query language (SQL) server database. Figure 1 shows a screenshot of this *Excel* table. The 116269-row imported table in the SQL server database named “Export_CT\$” and covers the period from January 1, 2010, to August 31, 2025. The method has been practically tested using a large multifunctional medical center as an example. A total of 50527 patients were examined, including 27333 women, 23181 men, and 13 patients whose

gender was not specified. The periodically repeated addition of data from the MIS to the SQL server database to Stages 1 and 2 from subsequent time periods keeps it up-to-date.

At the third stage, in order to visualize data on the number of CT scans and the dose received for a given nosology, the *Access* database management system (DBMS) is used as a client application to connect to a SQL server database using an open database connectivity (ODBC) driver. This client-server approach can be used simultaneously by an unlimited number of research clients. In the *Access* application, a query is generated to retrieve data from the SQL server table “Export_CT\$.” The query is executed server-side, which speeds up the execution process, and the result is transferred to the *Access* DBMS application as a table view named “Export_CT_PerDose,” a screenshot of which is shown in Fig. 2.

2. DESCRIPTION OF THE STRUCTURE OF THE SQL QUERY USED IN THE IAS

SQL is a structured query language based on the logic of the relational database model. A relational database model is a logical system consisting of interconnected objects that can be manipulated according to specific rules. The relational database model was founded by mathematician and IBM⁹ employee E.F. Codd, who in the 1960s also created relational algebra, in which he described a number of specific elements and operations on them. This relational algebra forms the theoretical basis of SQL. Introducing the concept of a relation (a flat data table), Codd identified eight relational operations: union, intersection, subtraction, Cartesian product, selection, projection, join, and division.

The structure of an SQL query used in the IAS consists of an outer query and several nested uncorrelated subqueries. It is presented in the table.

The developed technology proposes using uncorrelated subquery processing. A subquery in SQL is defined as a SELECT statement nested within another SELECT statement. With uncorrelated processing, the subquery is executed once, and its result (a list of patients satisfying the specified parameters) is stored in a temporary table, while the outer query retrieves data from the generated list of objects. Execution occurs from the bottom up: the subquery is executed at the bottom level, while a superstructure of other nested subqueries that use the resulting subset of objects is built from below. This enables the use of such SQL structures in various subject areas [17, 18].

² <https://www.medmis.ru> (in Russ.). Accessed May 24, 2026.

³ <https://www.medesk.net/ru> (in Russ.). Accessed May 24, 2026.

⁴ <https://archimed.pro> (in Russ.). Accessed May 24, 2026.

⁵ <https://k-integration.ru/vendors/kompleksnye-medicinskie-informacionnye-sistemy/> (in Russ.). Accessed May 24, 2026.

⁶ <https://www.medialog.ru/> (in Russ.). Accessed May 24, 2026.

⁷ <https://fmba.gov.ru/> (in Russ.). Accessed May 24, 2026.

⁸ The International Classification of Diseases, 10th revision, is a standard document that provides a generally accepted statistical classification of medical diagnoses.

⁹ International Business Machines is one of the world's largest manufacturers and suppliers of hardware and software.

A	B	C	D	E
PATIENTS_ID	RESEARCH_TYPE	NE_LE	GENDER	DOSE
163330	COMPUTED TOMOGRAPHY OF ABDOMINAL ORGANS	22.03.1988	0	15
39046	COMPUTED TOMOGRAPHY OF THE ABDOMINAL CAVITY AND RETROPERITONEAL SPACE	04.12.1974	1	5
56098	COMPUTED TOMOGRAPHY OF THE THORACIC CAVITY ORGANS	13.01.1944	1	3
31622	DENSITOMETRY	29.03.1953	1	
56776	COMPUTED TOMOGRAPHY OF THE PARANASAL SINUSES	20.03.1984	0	0.5
67175	COMPUTED TOMOGRAPHY OF THE THORACIC CAVITY ORGANS	22.10.1960	1	3
168173	COMPUTED TOMOGRAPHY OF THE THORACIC CAVITY ORGANS	13.12.1936	1	3
167682	DENSITOMETRY	03.08.1949	1	
168223	Computed tomography of the lumbosacral spine	28.12.1963	1	2
168706	COMPUTED TOMOGRAPHY OF ABDOMINAL ORGANS	30.04.1932	1	20.2
168668	Computed tomography of the lumbosacral spine	05.10.1946	1	2
22036	DENSITOMETRY	13.09.1952	1	
169293	COMPUTED TOMOGRAPHY OF THE THORACIC CAVITY ORGANS	13.11.1944	0	3
170153	COMPUTED TOMOGRAPHY OF THE BRAIN	23.10.1954	1	3
170439	COMPUTED TOMOGRAPHY OF THE PARANASAL SINUSES	14.09.1957	1	0.5

Fig. 1. A screenshot of an Excel spreadsheet. NE_LE—patient birth date; 0—male, 1—female;
DOSE is the dose received by the patient during one CT session in mSv

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
PATIENTS_ID	RESEARCH_TYPE	DATE_CONSULTATION	NameDisease	Disease_Category	NE_LE	Age	AgePeriod	GENDER	DOSE														
548594	MSCT OF THE THORACIC AND ABDOMINAL CAVITIES with intravenous contrast	24.01.2023	Malignant neoplasm of rectosigmoid junction	Neoplasms (C00-D48)	26.12.1963	60	> 50 и <= 60	Male	45.7														
548594	MSCT OF THE THORACIC AND ABDOMINAL CAVITIES with intravenous contrast	11.05.2022	Malignant neoplasm of rectosigmoid junction	Neoplasms (C00-D48)	26.12.1963	59	> 50 и <= 60	Male	19.6														
548594	MSCT OF THE THORACIC AND ABDOMINAL CAVITIES with intravenous contrast	30.11.2022	Malignant neoplasm of rectosigmoid junction	Neoplasms (C00-D48)	26.12.1963	59	> 50 и <= 60	Male	45.7														
548594	MSCT OF THE THORACIC AND ABDOMINAL CAVITIES with intravenous contrast	26.01.2022	Malignant neoplasm of rectosigmoid junction	Neoplasms (C00-D48)	26.12.1963	59	> 50 и <= 60	Male	45.7														
548594	MSCT OF THE THORACIC AND ABDOMINAL CAVITIES with intravenous contrast	14.09.2022	Malignant neoplasm of rectosigmoid junction	Neoplasms (C00-D48)	26.12.1963	59	> 50 и <= 60	Male	21.3														
548594	MSCT OF THE CORONARY ARTERIES	31.01.2022	Malignant neoplasm of rectosigmoid junction	Neoplasms (C00-D48)	26.12.1963	59	> 50 и <= 60	Male	45.7														
548594	MSCT OF THE THORACIC AND ABDOMINAL CAVITIES with intravenous contrast	27.07.2022	Malignant neoplasm of rectosigmoid junction	Neoplasms (C00-D48)	26.12.1963	59	> 50 и <= 60	Male	45.7														
548594	MSCT OF THE THORACIC AND ABDOMINAL CAVITIES with intravenous contrast	15.03.2022	Malignant neoplasm of rectum	Neoplasms (C00-D48)	26.12.1963	59	> 50 и <= 60	Male	45.7														

Fig. 2. Screenshot of the query result "Export_CT_PerDose." MSCT—multislice computed tomography

Table. Structure of an SQL query

	Query fragment	Explanation of the query fragment
1	SELECT * FROM	External query
2	<pre> (SELECT PATIENTS_ID, ISNULL(Export_CT\$. RESEARCH_TYPE, 'Not specified') AS RESEARCH_TYPE, Cast(Export_CT\$.DATE_CONSULTATION as datetime) as DATE_CONSULTATION,..., LABEL, CODE, ISNULL(SA_ICD_10_MIT_2021.WHO_Full_Desc, 'Not specified') AS NameDisease, ..., CASE WHEN SA_ICD_10_MIT_2021.ICD10_Code IN (SELECT SA_ICD_10_MIT_2021.ICD10_Code FROM Export_CT\$ WHERE SA_ICD_10_ MIT_2021.ICD10_Code LIKE 'A%' OR SA_ICD_10_MIT_2021.ICD10_Code LIKE 'B%') THEN 'Certain infectious and parasitic diseases (A00-B99)' WHEN SA_ICD_10_MIT_2021.ICD10_Code IN (SELECT SA_ICD_10_MIT_2021.ICD10_Code FROM Export_CT\$ WHERE SA_ICD_10_ MIT_2021.ICD10_Code LIKE 'C%' OR SA_ICD_10_MIT_2021.ICD10_Code LIKE 'D0%' OR SA_ICD_10_MIT_2021.ICD10_Code LIKE 'D1%' OR SA_ICD_10_MIT_2021.ICD10_ Code LIKE 'D2%' OR SA_ICD_10_MIT_2021.ICD10_Code LIKE 'D3%' OR SA_ICD_10_MIT_2021. ICD10_Code LIKE 'D4%') THEN 'Neoplasms (C00-D48)' WHEN SA_ICD_10_MIT_2021.ICD10_Code IN (SELECT SA_ICD_10_MIT_2021.ICD10_Code FROM Export_CT\$ WHERE SA_ICD_10_ MIT_2021.ICD10_Code LIKE 'D5%' OR SA_ICD_10_MIT_2021.ICD10_Code LIKE 'D6%' OR SA_ICD_10_MIT_2021.ICD10_Code LIKE 'D7%' OR SA_ICD_10_MIT_2021.ICD10_ Code LIKE 'D8%') THEN 'Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (D50-D89)' WHEN SA_ICD_10_MIT_2021.ICD10_Code IN (... ELSE 'Not specified ' END as Disease_Category..., </pre>	A nested subquery with a selection of attributes, converting data from symbolic to the format required for pivot tables, attaching the category and name of the disease from the ICD-10 directory of diseases
3	<pre> CASE WHEN EQUIPMENT = '0' THEN 'PHILIPS INGNUTY' WHEN EQUIPMENT = '1' THEN 'CANON ONE'... END AS EQUIPMENT_Name, ..., </pre>	The name of the equipment is programmatically assigned
4	<pre> DATEDIFF(yy, Cast(Export_CT\$.NE_LE as datetime), Cast(Export_CT\$.DATE_ CONSULTATION as datetime)) AS Age, CASE WHEN DATEDIFF(yy, Cast(Export_CT\$.NE_LE as datetime), Cast(Export_CT\$. DATE_CONSULTATION as datetime)) <=20 THEN '<= 20' WHEN DATEDIFF(yy, Cast(Export_CT\$.NE_LE as datetime), Cast(Export_CT\$. DATE_CONSULTATION as datetime)) <=30 AND DATEDIFF(yy, Cast(Export_CT\$.NE_LE as datetime), Cast(Export_CT\$.DATE_ CONSULTATION as datetime)) >20 THEN '> 20 and <=30'... ELSE 'Not specified ' END as AgePeriod, ..., </pre>	The “Age” and “AgePeriods” attributes are programmatically defined
5	<pre> CASE WHEN Export_CT\$.GENDER = 0 THEN 'M' WHEN Export_CT\$.GENDER = 1 THEN 'F' ELSE ' Not specified ' END as GENDER2, ..., </pre>	The value of the “GENDER” attribute is programmatically assigned
6	<pre> Try_Convert(float, Replace(IsNull(Export_CT\$.DOSE, 0), ',', '.')) as DOSE_ </pre>	The symbolic value is programmatically converted to a numeric value for the “DOSE_” attribute
7	<pre> FROM Export_CT\$ LEFT OUTER JOIN SA_ICD_10_MIT_2021 ON Export_CT\$.CODE = SA_ICD_10_ MIT_2021.ICD10_Code) W </pre>	The left outer connection of the tables Export_CT\$ and SA_ICD_10_MIT_2021
8	<pre> WHERE W.PATIENTS_ID IN (SELECT PATIENTS_ID FROM (SELECT PATIENTS_ID, SUM(Try_Convert(float, Replace(IsNull(Export_CT\$. DOSE, 0), ',', '.'))) SmDs FROM Export_CT\$ GROUP BY PATIENTS_ID) Q WHERE Q.SmDs >400 AND Q.SmDs <=700) </pre>	A nested uncorrelated subquery that creates a list of patients whose cumulative ionizing radiation dose falls within the specified range from 400 to 700 mSv (inclusively)

3. RESULTS

The application uses pivot table mode to display the data obtained as a result of executing a generated query. Figure 3 depicts this mode as an empty workspace with a list of pivot table fields. The locations from which pivot table fields can be dragged from the field list are indicated at the edges and in the center of the workspace.

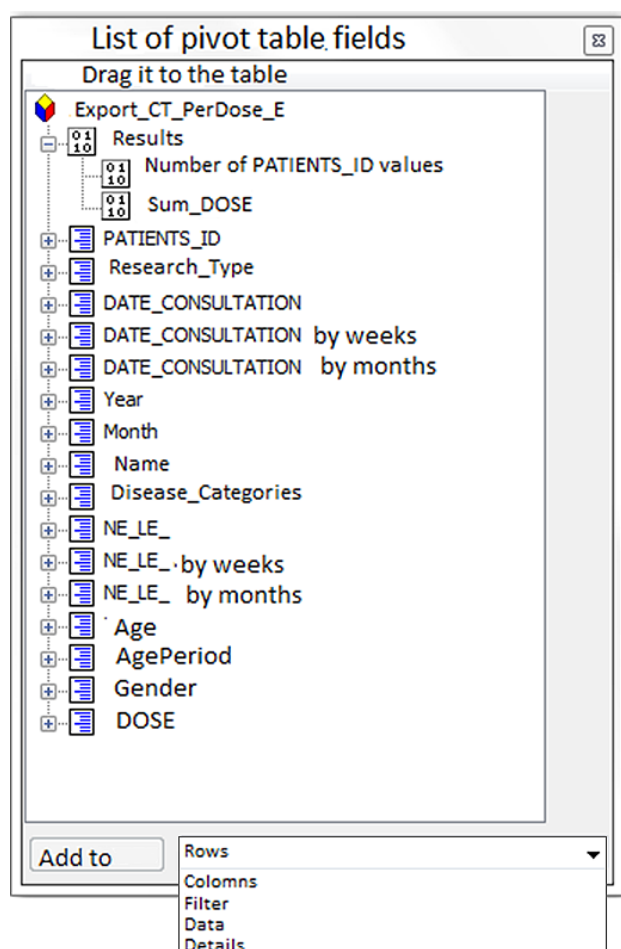


Fig. 3. View of an empty area of a pivot table

By dragging and dropping selected attributes from the field list or adding them to specified locations (rows, columns, filter, data, details), pivot tables can be created in the format required for research. For example, in order to find the patient ID who received specified cumulative doses of ionizing radiation over the past

few years, you would drag the “PATIENTS_ID” attribute from the pivot table field list to the rows area, “DATE_CONSULTATION” to the columns area, and “Sum_DOSE” to the data area. Figure 4 shows the resulting pivot table with the required data. The cumulative dose boundaries are specified in the query text in Section 9 of Table.

If, instead of “Sum_DOSE,” you drag the “Number of PATIENTS_ID values” attribute from the pivot table field list into the data area, you will get a similar pivot table, but the intersection of rows and columns will indicate the number of CT sessions received by patients in a specific time period. By varying the cumulative dose value boundaries in the query text (the last row of item 9 in Table) from 0 to 100 (Fig. 5), from 100 to 400 (Fig. 6), and from 400 to 700 (Fig. 7), and examining the frequency structure of CT sessions by disease category, it is possible to identify differences in the correlations between the number of CT sessions and the cumulative dose value. Thus, the order of disease categories in decreasing order of CT scan sessions for cumulative doses from 0 to 100 mSv differs from the order of disease categories for cumulative doses from 100 to 400 mSv and from 400 to 700 mSv. Neoplasms (C00-D48) ranks first in terms of the number of CT scan sessions in these groups based on cumulative dose load.

In order to view cumulative doses distributed among selected patients by disease category, the disease categories should be placed in Rows, patient IDs in Columns, and the “Sum_DOSE” attribute in Data area. Then the desired patients can be selected from the drop-down list. The resulting pivot table is shown in Fig. 8, while the corresponding 3D diagram appears in Fig. 9. If the “Number of values” attribute is placed in the Data area in this case, the intersection of the columns and rows of the pivot table will yield the corresponding number of CT sessions.

In order to study the distribution of cumulative doses of ionizing radiation by year for specific selected patients, categories can also be expanded by the names of the diseases they comprise. Figure 10 depicts a screenshot of the distribution of cumulative doses of ionizing radiation across all categories by all included diseases for patient ID 99111, who was found to have the highest cumulative dose.

		Year ▼														Overall results	
PATIENTS_ID ▼		2010	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025					
		Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE		
99111	+		20					40.2	431.2	47.8	79	5.2		623.4			
492869	+								126	92.1	255.1	103	15.2	591.4			
394588	+							491	10.6	11.3	48.2	18.3		579.4			
865195	+											516.9	59.7	576.6			
574199	+									345.1	174.5			519.6			
917367	+												515.6	515.6			
179363	+					20	40	113.5	147.6		47.6		144.7	513.4			
502603	+								319.5	182.4				501.9			
830626	+											495.5		495.5			
488894	+								398.6	21.1	38.9	2		460.6			
612961	+									82.5	371.1			453.6			
146442	+	224									225.022			449.022			
182157	+			4				59.6	95.8	29.1	132.3	127.8		448.6			
389560	+							66.9	350	30				446.9			
715676	+										404.4	38.9		443.3			
52627	+			10	55	41	30	47.2	93.6	18.8	48.1	29	63.4	436.1			
749577	+										415.5			415.5			
531085	+								410.8					410.8			
519626	+								344.2		19.3	44.6		408.1			
38717	+							104	302.9					406.9			
Overall results	+	224		14	75	61	70	922.4	3030.8	860.2	2259.022	1381.2	798.6	9696.222			

Fig. 4. Cumulative doses to patients from 400 to 700 mSv over the years

Disease_Category	Year												Overall results
	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	
Not specified	2036	2754	3270	2808	2808	4805	3305	3562	3321	2534	28395	28395	
Neoplasms (C00-D48)	2542	2774	2907	2188	2188	2956	2400	2728	2794	2073	23362	23362	
Diseases of the circulatory system (I00-I99)	729	763	782	622	622	1049	975	955	974	650	7499	7499	
Diseases of the digestive system (K00-K93)	424	505	556	367	367	635	611	671	594	375	4738	4738	
Injury, poisoning and certain other consequences of external causes (S00-T98)	321	360	322	136	136	228	208	444	702	640	3361	3361	
Diseases of the respiratory system (J00-J99)	244	283	295	194	194	410	293	221	275	152	2367	2367	
Diseases of the musculoskeletal system and connective tissue (M00-M99)	156	150	233	212	212	464	293	290	290	180	2268	2268	
Diseases of the genitourinary system (N00-N99)	77	71	223	237	237	464	360	280	280	187	2179	2179	
Diseases of the nervous system (G00-G99)	183	184	179	149	149	284	293	343	274	186	2075	2075	
Codes for special purposes (U00-U85)	73	79	53	30	30	133	106	83	92	59	708	708	
Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (D50-D89)	30	27	26	60	60	119	67	76	88	67	560	560	
Factors influencing health status and contact with health services (Z00-Z99)	25	29	16	12	12	76	65	56	34	40	353	353	
Endocrine, nutritional and metabolic diseases (E00-E90)	84	35	29	47	47	24	20	24	42	18	323	323	
Certain infectious and parasitic diseases (A00-B99)	36	47	19	20	20	29	14	11	24	16	216	216	
Diseases of the ear and mastoid process (H60-H95)	9	10	13	4	4	24	3	8	4	6	81	81	
Congenital malformations, deformations and chromosomal abnormalities (Q00-Q99)	1	1	9	4	4	12	13	18	7	7	65	65	
Diseases of the eye and adnexa (H40-H59)	5	1	2	1	1	29	7	6	7	2	60	60	
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99)	3	2	2	4	4	5	4	4	2	4	30	30	
Diseases of the skin and subcutaneous tissue (L00-L99)	6978	8075	8946	8124	8124	11773	9038	9780	9808	7189	79711	79711	
Pregnancy, childbirth and the puerperium (O00-O99)													
Overall results													

Fig. 5. Number of CT sessions by disease category for cumulative doses from 0 to 100 mSv

Disease_Category	Years												Overall results
	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	
Neoplasms (C00-D48)	252	373	669	834	834	2339	1346	1581	1474	1050	9918	9918	
Not specified	101	184	266	275	275	707	402	402	353	330	3020	3020	
Diseases of the digestive system (K00-K93)	43	85	134	191	191	386	236	265	232	161	1733	1733	
Injury, poisoning and certain other consequences of external causes (S00-T98)	17	31	49	92	92	166	169	171	126	92	913	913	
Diseases of the circulatory system (I00-I99)	41	60	43	63	63	126	102	102	64	67	631	631	
Diseases of the respiratory system (J00-J99)	10	9	9	11	11	102	21	31	5	13	211	211	
Diseases of the genitourinary system (N00-N99)	8	4	11	11	11	54	10	28	51	20	197	197	
Diseases of the nervous system (G00-G99)	3	8	5	28	28	62	25	13	22	19	185	185	
Factors influencing health status and contact with health services (Z00-Z99)	2	3	6	6	6	9	13	12	10	6	67	67	
Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (D50-D89)	2	3	5	15	15	12	8	6	6	6	58	58	
Certain infectious and parasitic diseases (A00-B99)	5	3	4	5	5	18	3	9	7	7	54	54	
Diseases of the musculoskeletal system and connective tissue (M00-M99)													
Endocrine, nutritional and metabolic diseases (E00-E90)	14	3	2	2	2	4	7	5	5	3	43	43	
Congenital malformations, deformations and chromosomal abnormalities (Q00-Q99)	2	1				2	12	12			17	17	
Codes for special purposes (U00-U85)						8	1				9	9	
Diseases of the skin and subcutaneous tissue (L00-L99)								1			4	4	
Diseases of the ear and mastoid process (H60-H95)													
Diseases of the eye and adnexa (H40-H59)													
Overall results	486	780	1214	1541	1541	4026	2353	2615	2389	1792	17196	17196	

Fig. 6. Number of CT sessions by disease category for cumulative doses from 100 to 400 mSv

Disease_Category	Years											
	Drag the filter fields here											
	2017	2018	2019	2020	2021	2022	2023	2024	2025	Overall results		
	+	+	+	+	+	+	+	+	+	+	+	+
	-	-	-	-	-	-	-	-	-	-	-	-
	Number of values	Number of values	Number of values	Number of values	Number of values	Number of values	Number of values	Number of values	Number of values	Number of values	Number of values	Number of values
Neoplasms (C00-D48)	2		5		8	50	17	54	44	15	139	195
Not specified	4		7		7	31	10	28	11	37	139	195
Diseases of the digestive system (K00-K93)	1		1		11	22	4	18	12	70	70	70
Injury, poisoning and certain other consequences of external causes (S00-T98)					1	25	14	3	15	6	64	64
Diseases of the circulatory system (I00-I99)						27	2	3	1	33	33	33
Diseases of the respiratory system (J00-J99)						12				12	12	12
Diseases of the genitourinary system (N00-N99)	1					6		1		8	8	8
Diseases of the skin and subcutaneous tissue (L00-L99)						1		3	2	6	6	6
Diseases of the nervous system (G00-G99)						4		1		5	5	5
Factors influencing health status and contact with health services (Z00-Z99)								2		2	2	2
Overall results	8	8	10	27	178	47	113	85	58	534	534	534

Fig. 7. Number of CT sessions by disease category for cumulative doses from 400 to 700 mSv

Disease_Category	PATIENTS_ID											
	Drag the filter fields here											
	99111	492869	394588	865195	574199	917367	519626	Overall results				
	+	+	+	+	+	+	+	+	+	+	+	+
	-	-	-	-	-	-	-	-	-	-	-	-
	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE
Neoplasms (C00-D48)	322.2	270.3								184.4	776.9	776.9
Not specified	80.2				59.7	95.6	515.6			751.1	751.1	751.1
Injury, poisoning and certain other consequences of external causes (S00-T98)				70.2	516.9	323	910.1			910.1	910.1	910.1
Diseases of the digestive system (K00-K93)		274.4	509.2		101		884.6			884.6	884.6	884.6
Diseases of the genitourinary system (N00-N99)	221						221			221	221	221
Diseases of the skin and subcutaneous tissue (L00-L99)							63.9			63.9	63.9	63.9
Diseases of the nervous system (G00-G99)							80.2			80.2	80.2	80.2
Diseases of the circulatory system (I00-I99)							79.6			79.6	79.6	79.6
Factors influencing health status and contact with health services (Z00-Z99)		46.7					46.7			46.7	46.7	46.7
Overall results	623.4	591.4	579.4	519.6	576.6	515.6	408.1			3814.1	3814.1	3814.1

Fig. 8. Cumulative doses to patients by disease category in the range from 400 to 700 mSv

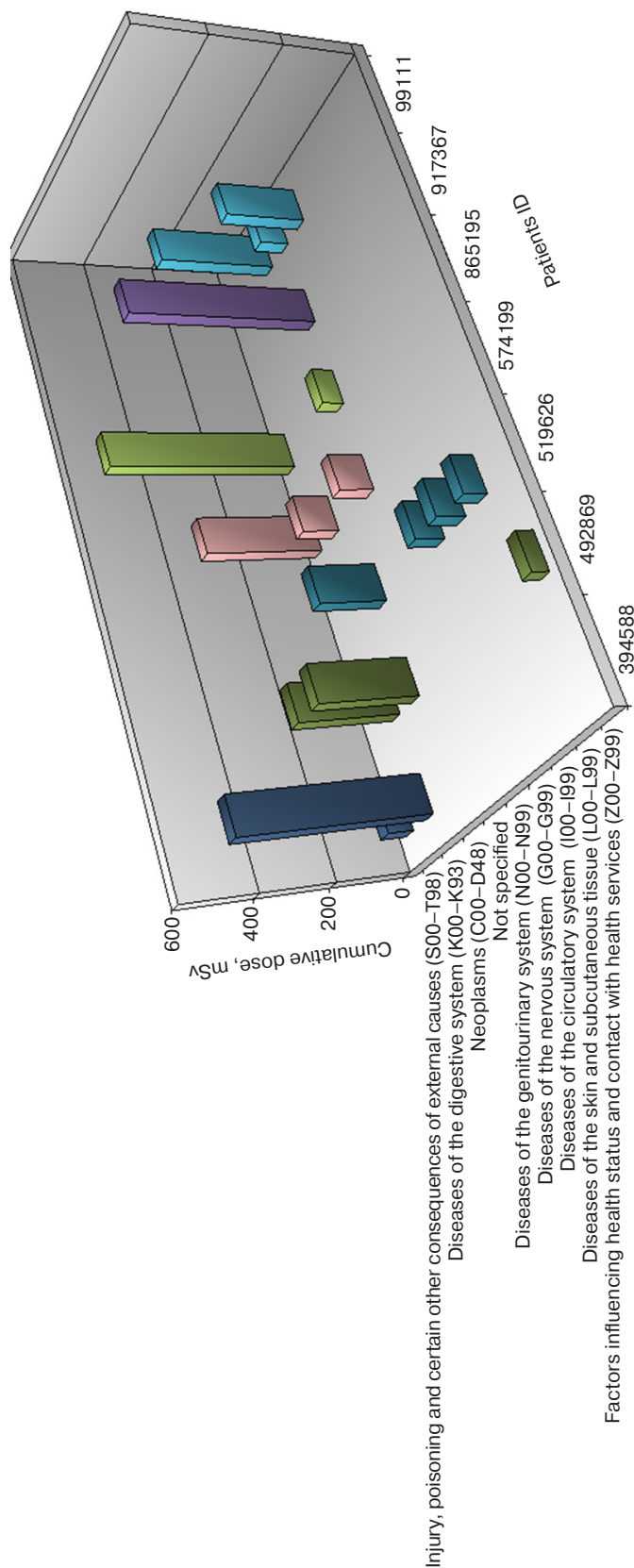


Fig. 9. Volumetric diagram of cumulative doses to patients by disease category in the range from 400 to 700 mSv

Export_CT_PerDose		PATIENTS_ID		Years		2017		2020		2021		2022		2023		2024		Overall results	
Disease_Category	PATIENTS_ID	All	99111			Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE	Sum_DOSE
NameDisease																			
Malignant neoplasm, overlapping lesion of retroperitoneum and peritoneum						226												226	
Rupture of bladder, nontraumatic						180.2												180.2	
Not specified							40.2							15				80.2	
Malignant neoplasms of independent (primary) multiple sites															58.2		5.2	63.4	
Malignant neoplasm, retroperitoneum														32.8				32.8	
Hydronephrosis with ureteral stricture, not elsewhere classified															20.8			20.8	
Acute tubulo-interstitial nephritis						20												20	
Overall results						20	40.2		431.2	47.8	79	5.2	623.4						

Fig. 10. Summary table of cumulative doses for patient ID 99111 for all disease categories and names by year

CONCLUSIONS

The proposed client-server technology using pivot tables and volumetric charts in the client application enables:

1. Visual analysis and accounting of cumulative doses of patients undergoing treatment and medical examination in large multifunctional treatment and research hospitals.
2. Identification of CT scans for which the number of sessions could be reduced to limit the increase in cumulative patient doses. For example, the maximum number of CT sessions identified for one patient over the entire observation period was 83.
3. Detection of data entry errors by medical workers in the MIS and their timely correction. For example,

the “Not specified” value in the tables above means that no data was entered in that column.

4. Creation of various pivot tables for data analysis. For example, all the data presented in the report was obtained using just one query, “Export_CT_PerDose.”

Thus, pivot tables and volumetric diagrams are shown to be a visual and effective tool for analyzing and studying the processes of disease progression across time and age series, for example, to determine the effectiveness of certain treatment procedures and the impact of the use of drugs on the results of treatment carried out in multidisciplinary clinical hospitals.

Authors' contributions

V.A. Kadymov—general guidelines for writing the article.

S.M. Dumansky—introduction, sections 1 and 2.

E.V. Vasiliev—section 3 and conclusion.

REFERENCES

1. Brenner D., Elliston C., Hall E., Berdon W. Estimated risks of radiation-induced fatal cancer from pediatric CT. *Am. J. Roentgenology*. 2001;176(2):289–296. <https://doi.org/10.2214/ajr.176.2.1760289>
2. Denton E.R., Mackenzie A., Greenwell T., Popert R., Rankin S.C. Unenhanced helical CT for renal colic – is the radiation dose justifiable? *Clin Radiol*. 1999;54(7):444–447. [https://doi.org/10.1016/S0009-9260\(99\)90829-2](https://doi.org/10.1016/S0009-9260(99)90829-2)
3. Kleinerman R.A. Cancer Risks Following Diagnostic and Therapeutic Radiation Exposure in Children. *Pediatr. Radiol*. 2006;36(Suppl.2):121–125. <https://doi.org/10.1007/s00247-006-0191-5>
4. Bosch de Basea M., Thierry-Chef I., Harbron R., et al. Risk of hematological malignancies from CT radiation exposure in children, adolescents and young adults. *Nat. Med*. 2023;29(12):3111–3119. <https://doi.org/10.1038/s41591-023-02620-0>
5. Osipov M.V., Shkarednykh V.Yu., Loginov V.S., Melnikov V.V., Druzhinina P.S., Sokolnikov M.E. Retrospective analysis of cancer morbidity among patients after computed tomography. *Radiatsionnaya Gygiena = Radiation Hygiene*. 2021;14(3): 80–90 (in Russ.). <https://doi.org/10.21514/1998-426X-2021-14-3-80-90>
6. Osipov M.V., Vazhenin A.V., Domozhirova A.S., Chernova O.N., Aksenova I.A. Computed tomography as a risk factor in cancer patients with professional radiation exposure. *Rossiiskii ehlektronnyi zhurnal luchevoi diagnostiki = Russian Electronic Journal of Radiology*. 2019;9(1):142–147 (in Russ.). <https://elibrary.ru/lzhdgh>
7. Bullard T.B., Falk J.L., Smith M.S., et al. Cumulative Radiation Exposure from Medical Imaging in Two Hospital Systems – Implications for Medical Record Portability. *Emergency Med*. 2014;4(211):5.
8. Kashcheev V.V., Pryakhin E.A., Menyailo A.N. A method for assessing the radiation risk of medical exposure after computed tomography scan of children and adolescents based on data from scanning protocols. *Radiatsiya i risk = Radiation and Risk*. 2023;32(3):5–13 (in Russ.). <https://doi.org/10.21870/0131-3878-2023-32-3-5-13>
9. Kashcheev V.V., Pryakhin E.A. Medical diagnostic imaging: radiation safety issues. Review. *Radiatsiya i risk = Radiation and Risk*. 2018;27(4):49–64 (in Russ.). <https://doi.org/10.21870/0131-3878-2018-27-4-49-64>
10. Osipov M.V. Computed tomography as a risk factor of cancer among the population of the nuclear city of Ozyorsk. *Radiatsiya i risk = Radiation and Risk*. 2023;32(3):109–121 (in Russ.). <https://doi.org/10.21870/0131-3878-2023-32-3-109-121>
11. Dumansky S.M., Vasiliev E.V., Kalinina M.V. The use of information technologies for a retrospective analysis of the frequency of neoplasm detection in patients using the example of a multidisciplinary medical center. In: *The 8th International Forum of Oncology and Radiotherapy “For Life – FOR LIFE”*: Collection of abstracts. Moscow: 2025. ISBN 978-5-00274-009-3. P. 90–91 (in Russ.). Available from URL: <https://drive.google.com/file/d/1hxnTQgBquJ0gBBoOyrXaMumUqJqWYmb6/view>. Accessed May 24, 2026.
12. Vasiliev E.V., Dumansky S.M., Kalinina M.V. On the need for personal centralized accounting and analysis of patients' dose load when using computed tomography. In: *Ilyinsky Readings 2025*: Collection of Materials of the International Scientific and Practical Forum of Young Scientists and Specialists. Moscow: 2025. P. 90–92 (in Russ.). <https://www.elibrary.ru/phvfoi>. Available from URL: <https://ilyinform.ru/materials>. Accessed May 24, 2026.
13. Dumansky S.M., Vasiliev E.V., Baranov L.I., Kalinina M.V. Visualization, analysis and accounting of the dose load of patients undergoing radiation diagnostics. In: *Medical and Ecological Effects of Ionizing Radiation*: Proceedings of the 8th International Scientific Conference, dedicated to the 25th anniversary of the Seversk Biophysical Research Center. Tomsk: 2025. P. 33–34 (in Russ.).

14. Dumansky S.M., Vasiliev E.V., Kalinina M.V. Computed tomography – contribution to cumulative medical dose Chronic radiation exposure: medical and biological effects of small doses: In: *Proceedings of the 8th International Scientific and Practical Conference*. September 23–25, 2025. Chelyabinsk: 2025. P. 88–89 (in Russ.).
15. Dumansky S.M., Vasiliev E.V., Kalinina M.V. Analysis of the prevalence of malignant neoplasms in the context of professions of current and former employees of energy enterprises according to medical examinations conducted in the medical unit system. In: *Proceedings of the 7th International Forum of Oncology and Radiotherapy*. September 16–20, 2024. Moscow: World Trade Center. P. 186–187 (in Russ.). Available from URL: <https://drive.google.com/file/d/1-8szW3yzzOdJ4Ex1ZnnydXkrX3-hmf40/view>. Accessed May 24, 2026.
16. Kashcheev V.V., Pryakhin E.A., Menyailo A.N., Ivanov V.K., Ivanov S.A., Kaprin A.D. Application of the new ICRP 2022 recommendations in calculating radiation risks of patients after computed tomography. *Radiatsiya i risk = Radiation and Risk*. 2025;34(2):39–49 (in Russ.). <https://doi.org/10.21870/0131-3878-2025-34-2-39-49>
17. Ben-Gan Itzik. *Microsoft SQL Server 2012. T-SQL Fundamentals*: transl. from Engl. Moscow: Eksmo; 2015, 400 p. (In Russ.). ISBN 978-5-699-73617-1
[Ben-Gan Itzik. *Microsoft SQL Server 2012 T-SQL Fundamentals*. Microsoft Press Series. Microsoft; 2012, 412 p.]
18. Nielsen P. *Microsoft SQL Server 2005. User's Bible*: transl. from Engl. Moscow: Vil'yams; 2008, 1232 p. (In Russ.). ISBN 978-5-84591-314-2
[Nielsen P. *SQL Server 2005. Bible*. Wiley & Sons; 2006, 1344 p.]

СПИСОК ЛИТЕРАТУРЫ

1. Brenner D., Elliston C., Hall E., Berdon W. Estimated risks of radiation-induced fatal cancer from pediatric CT. *Am. J. Roentgenology*. 2001;176(2):289–296. <https://doi.org/10.2214/ajr.176.2.1760289>
2. Denton E.R., Mackenzie A., Greenwell T., Popert R., Rankin S.C. Unenhanced helical CT for renal colic – is the radiation dose justifiable? *Clin Radiol*. 1999;54(7):444–447. [https://doi.org/10.1016/S0009-9260\(99\)90829-2](https://doi.org/10.1016/S0009-9260(99)90829-2)
3. Kleinerman R.A. Cancer Risks Following Diagnostic and Therapeutic Radiation Exposure in Children. *Pediatr. Radiol*. 2006;36(Suppl.2):121–125. <https://doi.org/10.1007/s00247-006-0191-5>
4. Bosch de Basea M., Thierry-Chef I., Harbron R., et al. Risk of hematological malignancies from CT radiation exposure in children, adolescents and young adults. *Nat. Med*. 2023;29(12):3111–3119. <https://doi.org/10.1038/s41591-023-02620-0>
5. Осипов М.В., Шкаредных В.Ю., Логинов В.С., Мельников В.В., Дружинина П.С., Сокольников М.Э. Ретроспективный анализ онкологической заболеваемости пациентов после проведения компьютерной томографии. *Радиационная гигиена*. 2021;14(3):80–90. <https://doi.org/10.21514/1998-426X-2021-14-3-80-90>
6. Осипов М.В., Важенин А.В., Доможирова А.С., Чернова О.Н., Аксенова И.А. Компьютерная томография как фактор риска у онкологических пациентов при наличии профессионального облучения. *Российский электронный журнал лучевой диагностики*. 2019;9(1):142–147. <https://elibrary.ru/lzhdgh>
7. Bullard T.B., Falk J.L., Smith M.S., et al. Cumulative Radiation Exposure from Medical Imaging in Two Hospital Systems – Implications for Medical Record Portability. *Emergency Med*. 2014;4(211):5.
8. Кашеев В.В., Пряхин Е.А., Меньяло А.Н. Метод оценки радиационного риска медицинского облучения при прохождении компьютерной томографии детей и подростков на основе данных протоколов сканирования. *Радиация и риск*. 2023;32(3):5–13. <https://doi.org/10.21870/0131-3878-2023-32-3-5-13>
9. Кашеев В.В., Пряхин Е.А. Медицинское диагностическое облучение: проблема радиационной безопасности. Обзор. *Радиация и риск*. 2018;27(4):49–64. <https://doi.org/10.21870/0131-3878-2018-27-4-49-64>
10. Осипов М.В. Компьютерная томография как фактор риска злокачественных новообразований среди населения города атомной промышленности Озёрск. *Радиация и риск*. 2023;32(3):109–121. <https://doi.org/10.21870/0131-3878-2023-32-3-109-121>
11. Думанский С.М., Васильев Е.В., Калинина М.В. Использование информационных технологий для ретроспективного анализа частоты выявления новообразований у пациентов на примере многопрофильного медицинского центра. В сб.: *VIII Международный Форум онкологии и радиотерапии «Ради жизни – FOR LIFE»*: Сборник тезисов. Москва. 15–19 сентября 2025 г. ISBN 978-5-00274-009-3. С. 90–91. URL: <https://drive.google.com/file/d/1hxnTQgBquJ0gBB0YrXaMumUqJqWYmb6/view>. Дата обращения 24.05.2026.
12. Васильев Е.В., Думанский С.М., Калинина М.В. О необходимости персонального централизованного учета и анализа дозовой нагрузки пациентов при использовании компьютерной томографии. В сб.: *«Ильинские чтения 2025»*: Сборник материалов международного научно-практического форума молодых ученых и специалистов. Москва: 2025. С. 90–92. <https://www.elibrary.ru/phvfoi>, URL: <https://ilyinform.ru/materials>. Дата обращения 24.05.2026.
13. Думанский С.М., Васильев Е.В., Баранов Л.И., Калинина М.В. Визуализация, анализ и учет дозовой нагрузки пациентов, проходящих лучевую диагностику. В сб.: *Медицинские и экологические эффекты ионизирующего излучения*: материалы VIII Международной научно-практической конференции, посвященной 25-летию образования Северского биофизического научного центра. 09–11 сентября 2025 г., Томск: 2025. С. 33–34.
14. Думанский С.М., Васильев Е.В., Калинина М.В. Компьютерная томография – вклад в кумулятивную медицинскую дозу. В сб.: *Хроническое радиационное воздействие: медико-биологические эффекты малых доз*: материалы VIII Международной научно-практической конференции. 23–25 сентября 2025 г. Челябинск: 2025. С. 88–89.

15. Думанский С.М., Васильев Е.В., Калинина М.В. Анализ распространенности злокачественных новообразований в разрезе профессий действующих и бывших работников энергетических предприятий по данным диспансеризаций, проводимых в системе медсанчастей. В сб.: *VII Международный форум онкологии и радиотерапии*. 16–20 сентября 2024 г. Москва: Центр международной торговли. С. 186–187. URL: <https://drive.google.com/file/d/1-8szW3yzzOdJ4Ex1ZnnydXkrX3-hmf40/view>. Дата обращения 24.05.2026.
16. Кашеев В.В., Пряхин Е.А., Меньяло А.Н., Иванов В.К., Иванов С.А., Каприн А.Д. Применение новых рекомендаций МКРЗ 2022 г. при расчете радиационных рисков пациентов после компьютерной томографии. *Радиация и риск*. 2025;34(2):39–49. <https://doi.org/10.21870/0131-3878-2025-34-2-39-49>
17. Бен-Ган И. *Microsoft SQL Server 2012. Основы T-SQL*: пер. с англ. М.: Эксмо; 2015, 400 с. ISBN 978-5-699-73617-1
18. Нильсен П. *Microsoft SQL Server 2005. Библия пользователя*: пер. с англ. М.: Вильямс; 2008, 1232 с. ISBN 978-5-84591-314-2

About the Authors

Vagid A. Kadymov, Dr. Sci. (Phys.-Math.), Professor, Higher Mathematics Department – 3, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: kadymov@mirea.ru. Scopus Author ID 6507691680, ResearcherID E-8760-2014, RSCI SPIN-code 4351-2884, <https://orcid.org/0009-0008-4550-8082>

Sergey M. Dumansky, Cand. Sci. (Econ.), Senior Researcher, A.I. Burnazyan Federal Medical Biophysical Center (23, Marshala Novikova ul., Moscow, 123098 Russia). E-mail: s.dumansky.m@gmail.com. RSCI SPIN-code 7655-7200, <https://orcid.org/0009-0002-9067-3876>

Evgeny V. Vasiliev, Senior Researcher, A.I. Burnazyan Federal Medical Biophysical Center (23, Marshala Novikova ul., Moscow, 123098 Russia). E-mail: e.vasilyev.v@gmail.com. <https://orcid.org/0000-0003-2133-7297>

Об авторах

Кадымов Вагид Ахмедович, д.ф.-м.н., профессор, кафедра высшей математики, Институт перспективных технологий и промышленного программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: kadymov@mirea.ru. Scopus Author ID 6507691680, ResearcherID E-8760-2014, SPIN-код РИНЦ 4351-2884, <https://orcid.org/0009-0008-4550-8082>

Думанский Сергей Митрофанович, к.э.н., старший научный сотрудник, ФГБУ ГНЦ Федеральный медицинский биофизический центр имени А.И. Бурназяна ФМБА России (ФГБУ ГНЦ ФМБЦ им. А.И. Бурназяна ФМБА России) (123098, Россия, Москва, ул. Маршала Новикова, д. 23). E-mail: s.dumansky.m@gmail.com. SPIN-код РИНЦ 7655-7200, <https://orcid.org/0009-0002-9067-3876>

Васильев Евгений Владимирович, старший научный сотрудник, ФГБУ ГНЦ Федеральный медицинский биофизический центр имени А.И. Бурназяна ФМБА России (ФГБУ ГНЦ ФМБЦ им. А.И. Бурназяна ФМБА России) (123098, Россия, Москва, ул. Маршала Новикова, д. 23). E-mail: e.vasilyev.v@gmail.com. <https://orcid.org/0000-0003-2133-7297>

Translated from Russian into English by L. Bychkova

Edited for English language and spelling by Thomas A. Beavitt

UDC 004.75; 004.4

<https://doi.org/10.32362/2500-316X-2026-14-4-21-35>

EDN ENTOLE



RESEARCH ARTICLE

Estimating the overload risk of a Kubernetes node based on the Fokker–Planck equation and the first-to-critical-bound problem

Sergey A. Lesko[®], Vladimir N. Kalinin, Dmitry O. Zhukov

MIREA – Russian Technological University, Moscow, 119454 Russia

[®] Corresponding author, e-mail: lesko@mirea.ru

• Submitted: 01.04.2026 • Revised: 15.04.2026 • Accepted: 08.06.2026

Abstract

Objectives. In containerized computing infrastructures, workload placement relies on static parameters such as *requests* and *limits*. However, the stochastic nature of actual resource consumption, involving rare but significant peak bursts, can lead either to excessive resource reservation or an increased risk of overload. Existing approaches to resource allocation focus on deterministic node selection criteria or point forecasts of workloads, but do not provide direct probabilistic estimates of critical state thresholds. The probabilistic model for the overload risk of *Kubernetes* compute nodes developed in this paper formalizes overload as the first instance of a critical boundary being reached by a stochastic load process.

Methods. The normalized node load is modeled as a one-dimensional stochastic process, the probability density of which evolves according to the Fokker–Planck equation. An analytical solution to the boundary-value problem is derived for constant coefficients and mixed Robin–Dirichlet boundary conditions. For variable coefficients estimated from telemetry data, a numerical method based on the Crank–Nicolson scheme is employed. The drift and diffusion coefficients are estimated from monitoring data.

Results. The model was validated using a discrete-event simulator for a 170-node cluster under four workload scenarios. A verification of the developed numerical solution against an analytical solution on six test cases produced a discrepancy of less than 0.001. A systematic comparison with four baseline methods (Persistence, Linear trend, AR(1), and Historical exceedance) across 904800 test windows demonstrates the superiority of the FP model in terms of the Brier score (0.0245 vs 0.0372 for the best baseline) for all prediction horizons (1, 2, 5, and 10 min).

Conclusions. The proposed model provides a calibrated, probabilistic estimate of *Kubernetes* node overload risk over a finite prediction horizon, which outperforms baseline methods in terms of both calibration and discrimination metrics. These results pave the way for the development of risk-aware node ranking mechanisms and the integration of probabilistic risk assessment into container cluster management systems.

Keywords: Kubernetes, stochastic dynamics, Fokker–Planck equation, boundary-value problem, overload probability, first-hitting time, Crank–Nicolson scheme, telemetry

For citation: Lesko S.A., Kalinin V.N., Zhukov D.O. Estimating the overload risk of a Kubernetes node based on the Fokker–Planck equation and the first-to-critical-bound problem. *Russian Technological Journal*. 2026;14(4):21–35. <https://doi.org/10.32362/2500-316X-2026-14-4-21-35>, <https://elibrary.ru/ENTOLE>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Оценивание риска перегрузки вычислительного узла в Kubernetes на основе уравнения Фоккера – Планка и задачи первого достижения критической границы

С.А. Лесько[@], В.Н. Калинин, Д.О. Жуков

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[@] Автор для переписки, e-mail: lesko@mirea.ru

• Поступила: 01.04.2026 • Доработана: 15.04.2026 • Принята к опубликованию: 08.06.2026

Резюме

Цели. В контейнерных вычислительных инфраструктурах размещение нагрузки опирается на статические параметры – запрашиваемые ресурсы (requests) и предельные ресурсы (limits), тогда как фактическое потребление ресурсов носит стохастический характер с редкими, но существенными пиковыми выбросами, что ведет к избыточному резервированию или риску перегрузки. Существующие подходы ориентированы на детерминированные критерии выбора узла или точечный прогноз нагрузки и не дают прямой вероятностной оценки достижения критического состояния. Цель работы – построение вероятностной модели риска перегрузки вычислительного узла *Kubernetes*¹, формализующей событие перегрузки как первое достижение критической границы стохастическим процессом загрузки узла.

Методы. Нормированная загрузка узла описывается как одномерный стохастический процесс, эволюция плотности вероятности задается уравнением Фоккера – Планка (Fokker–Planck, FP). Для краевой задачи получено аналитическое решение при постоянных коэффициентах и смешанных граничных условиях Робина – Дирихле. Для переменных коэффициентов, оцениваемых по телеметрии, реализован численный метод на основе схемы Кранка – Николсона. Коэффициенты дрейфа и диффузии оцениваются по данным мониторинга.

Результаты. Численное решение верифицировано на шести тестовых случаях с расхождением менее 0.001 относительно аналитического. Модель проверена на дискретно-событийном имитаторе кластера из 170 узлов в четырех сценариях нагрузки. Систематическое сравнение с четырьмя базовыми методами (наивный прогноз, линейный тренд, авторегрессия первого порядка, историческая частота) на 904800 тестовых окнах показало преимущество FP-модели по критерию Брайера (0.0245 против 0.0372 у лучшей базовой модели) на всех горизонтах прогнозирования (1, 2, 5 и 10 мин).

Выводы. Предложенная модель обеспечивает вероятностную оценку риска перегрузки узла *Kubernetes* на конечном горизонте прогноза и превосходит базовые методы по метрикам калибровки и дискриминации. Результаты создают основу для построения риск-ориентированных механизмов ранжирования узлов и интеграции вероятностной оценки в систему управления контейнерными кластерами.

Ключевые слова: Kubernetes, стохастическая динамика, уравнение Фоккера – Планка, краевая задача, вероятность перегрузки, первое достижение границы, схема Кранка – Николсон, телеметрия

¹ Открытое программное обеспечение для оркестрации контейнеризированных приложений. <https://kubernetes.io/> Дата обращения 11.05.2026. [Open-source software for orchestrating containerized application. <https://kubernetes.io/>. Accessed May 11, 2026.]

Для цитирования: Лесько С.А., Калинин В.Н., Жуков Д.О. Оценивание риска перегрузки вычислительного узла в Kubernetes на основе уравнения Фоккера – Планка и задачи первого достижения критической границы. *Russian Technological Journal*. 2026;14(4):21–35. <https://doi.org/10.32362/2500-316X-2026-14-4-21-35>, <https://elibrary.ru/ENTOLE>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

In *Kubernetes*² container orchestrators, the deployment of pods³ on compute nodes relies on declarative values, such as requested resources (requests) and limiting resources (limits) [1, 2]. However, actual resource consumption is stochastic and nonstationary. While this architecture simplifies quality of service management and application isolation, it also leads to structural inefficiency: conservative redundancy reduces the average utilization of the cluster [3], while aggressive compaction increases the risk of local node overload [4]. The scale of the problem is confirmed by analysis of industrial clusters. According to the tracing data of the Google Borg cluster management system [5], the average utilization of processor time is 20–35%; according to the Alibaba Cluster Trace (a publicly available trace data set published by Alibaba company that contains records of resource consumption by containers and tasks on production clusters; it is used as a benchmark for evaluating the effectiveness of load planning methods), it is 20–45% [6].

In operational practice, it is essential not only to forecast (predict) the expected workload but also to evaluate the likelihood of hazardous conditions occurring within a specified timeframe. The substantial risk of reaching a critical workload level at least once during the forecast period can lead to performance deterioration, container eviction, and violations of service-level agreements. Most current approaches, which focus either on deterministic criteria for selecting a node or on forecasting the workload value at a specific time, do not directly provide a probabilistic assessment of the risk of overload.

In this paper, a probabilistic model of the risk of overload *Kubernetes* compute nodes is proposed. This model is based on interpreting loading as a one-dimensional stochastic process and overload events as those surpassing the first critical boundary [7, 8]. The evolution of the load distribution is described using the Fokker–Planck equation, which allows for transition

from a point forecast to an estimation of the probability of reaching the limit [9, 10].

LITERATURE REVIEW

Existing approaches to managing container load placement in *Kubernetes* and related orchestrators can be divided into the following areas.

The first of these is heuristic and combinatorial container placement methods, which are aimed at maximizing packing density, balancing node load, and reducing resource fragmentation [1, 2, and 11]. These methods use current or estimated residual node resources as a deterministic selection criterion to formulate the task as a variant of combinatorial optimization. The advantages of this class of approaches are computational simplicity, interpretability, and compatibility with the standard scheduler cycle. However, such methods generally do not adequately consider the instability and random emissions of the load, nor do they provide a direct probabilistic assessment of achieving the critical condition of the node. A recent review paper [11] confirms that, despite the growing number of solutions using forecasting and automatic scaling, a significant part of scheduling methods in *Kubernetes* is comprised of heuristic and rule-based schemes.

The second area covers work on expanding the standard *Kubernetes* scheduler and constructing custom node ranking mechanisms [2, 12]. Rather than focusing on a new mathematical load model, the emphasis is placed on integrating additional decision-making logic into the existing orchestration infrastructure through scheduler plug-ins, external decision-making tools, custom ranking functions and dynamic container reallocation mechanisms. In recent years, a number of studies have been published that use the scheduler extension not only for static node selection, but also for dynamic load balancing during execution [13, 14]. However, even in such settings, the decision is typically deterministic and uncertainty regarding future load dynamics is only considered indirectly, via current consumption metrics and empirical rules of choice.

The third area comprises methods for forecasting load time series that involve the proactive, automatic scaling of systems [15–18]. These studies employ classical statistical models, recurrent neural network

² Open-source software for orchestrating containerized application. <https://kubernetes.io/>. Accessed May 11, 2026.

³ A pod is a group of one or more containers that work together and use shared resources.

architectures, and hybrid neural network architectures, as well as adaptive decision-making schemes that consider service-level requirements. In the context of *Kubernetes*, significant development occurs in the areas of horizontal and vertical scaling. Future demand values are predicted based on historical observations, after which decisions are made to adjust the number of replicas or the allocated resources. Such approaches are capable of taking trends, seasonality, and short-term workload dynamics into account. However, since most of these settings focus on minimizing error in predicting either the load itself or the scaling quality metric, the likelihood of an overload is only considered indirectly. This is not usually associated with reaching the critical limit first, which may be unacceptable for mission-critical tasks.

A separate class consists of probabilistic and risk-oriented approaches to managing distributed computing systems [4, 19–21]. Within this framework, the study considers settings that account for demand uncertainty, the trade-off between utilization and reliability, excessive resource demand, and interference between jointly placed loads. The current trend in container orchestration is characterized by a shift from reactive scaling to policy-based management and reinforcement learning, including risk-based economic formulations [21]. While these results confirm the potential of a probabilistic interpretation of operational risk, they typically focus on optimizing cost, delays or average resource usage rather than considering the likelihood of the node loading process crossing a critical threshold at least once within a given timeframe. Consequently, risk is more frequently modelled as an aggregated loss function or the probability of service degradation than as a primary achievement event for a stochastic loading process.

Such statements are based on the theory of stochastic differential equations and the Itô integral [22], methods for analyzing Fokker–Planck equations [9, 10], and the theory of first-passage processes [7, 8]. These methods have long been used to describe unsteady stochastic dynamics for calculating the probability of a process exceeding an acceptable range. A systematic description of the modeling approach for stochastic processes in computing systems is provided in [23], including the derivation of generalized equations and solutions to boundary value problems for load models when balancing data transmission nodes. However, the problem of first passage within the framework of the Fokker–Planck equation has not previously been applied to assess the risk of overload a computing node in *Kubernetes*.

Existing approaches either solve the placement problem based on deterministic heuristics or predict the future load value. Others consider stochastic effects

without explicitly stating the risk as events of the first critical boundary. Therefore, the task of constructing a model that can directly estimate the probability of node overload at the final time horizon, while being computationally suitable for integration into the ranking and management mechanisms of containerized environments such as *Kubernetes*, remains open.

MATHEMATICAL MODEL AND SOLUTION OF THE OVERLOAD RISK ASSESSMENT PROBLEM

Task setup

Let us consider a cluster of N compute nodes. For the node j , determine the normalized load for the selected resource (processor time or RAM):

$$x_j(t) = \frac{u_j(t)}{C_j}, \quad x_j(t) \in [0, 1], \quad (2.1)$$

where $u_j(t)$ is the actual resource consumption; t is continuous time; C_j is the node capacity. The critical condition corresponds to reaching the $L = 1$ level. We denote the forecast horizon by $T_h > 0$.

Definition 1 (overload event). The overload of node j at the T_h horizon is defined as the first time the critical boundary is reached: $\tau_L = \inf\{t > 0: x_j(t) \geq L\}$. The risk of overload is defined as the probability $R_j(T_h) = \mathbb{P}(\tau_L \leq T_h | x_j(0) = x_0)$.

The task is to build a model that can calculate $R_j(T_h)$ given x_0 , T_h and the stochastic dynamics parameters, and to develop a computational method for determining this probability. The resulting risk assessment can be used for node ranking and load placement decision support mechanisms. However, this work focuses primarily on the model, the boundary value problem and the methods for solving it. It is important to note that this setting differs fundamentally from threshold criteria in that the same current load value can correspond to both low and high risk depending on the trend direction and intensity of fluctuations.

Stochastic loading model

In order to describe the evolution of the normalized node load, we introduce a discrete Markov model with small increments over a fixed time step. This formulation formally relates local random load changes to a continuous equation for the probability density to calculate the probability of reaching the critical limit.

Let us fix the discrete time step $\tau > 0$ and denote $t = k\tau$, $k \in \mathbb{N}_0$. Let the node be in the state $x \in [0, L]$ at step k . Three possible changes are permitted per time step: $x \rightarrow x + \varepsilon$, $x \rightarrow x - \xi$, and $x \rightarrow x$, where $\varepsilon, \xi > 0$ are the scales of loading increments. The probabilities of the three directions satisfy the normalization condition

$w_+(x) + w_-(x) + w_0(x) = 1$. The balance of probabilities is given by the following equation:

$$P_{k+1}(x) = w_+(x - \varepsilon)P_k(x - \varepsilon) + w_-(x + \xi)P_k(x + \xi) + w_0(x)P_k(x). \quad (2.2)$$

Remark 1. Since the intensity of requests and task processing typically varies, the asymmetry of increments ($\varepsilon \neq \xi$) is a natural feature of computational load. Therefore, the general form (2.2) is used below.

By approximating $P_k(x)$ with the density $\rho(x, t)$ and decomposing the right-hand side of equation (2.2) into a Taylor series with respect to the state variable x up to second order at $\varepsilon, \xi = O(\sqrt{\tau})$, $\tau \rightarrow 0$, we obtain the Fokker–Planck equation:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x}(A(x)\rho) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(B(x)\rho). \quad (2.3)$$

The drift coefficient $A(x)$ and the diffusion coefficient $B(x)$ are expressed in terms of the conditional moments of the increments:

$$A(x) = \lim_{\tau \rightarrow 0} \frac{\mathbb{E}[\Delta x | x]}{\tau}, \quad B(x) = \lim_{\tau \rightarrow 0} \frac{\mathbb{E}[(\Delta x)^2 | x]}{\tau}. \quad (2.4)$$

For scheme 2.2, the coefficients take the following form in the first approximation with respect to τ :

$$A(x) \approx \frac{\varepsilon w_+(x) - \xi w_-(x)}{\tau}, \quad B(x) \approx \frac{\varepsilon^2 w_+(x) + \xi^2 w_-(x)}{\tau}. \quad (2.5)$$

In the special case when the directional probabilities do not depend on the state and are equal to $w_+ = w_- = 1/2$, $w_0 = 0$ (symmetric transition scheme), the coefficients can be simplified as follows:

$$a = \frac{\varepsilon^2 + \xi^2}{4\tau}, \quad b = \frac{\varepsilon - \xi}{2\tau}, \quad (2.6)$$

where a corresponds to the diffusion coefficient $(1/2)B$ and b to the drift coefficient A .

While the coefficient $A(x)$ determines the average trend in load changes, the intensity of random fluctuations is determined by the coefficient $B(x)$. The consideration of non-linear effects is based on the dependence of these coefficients on the state, particularly increased dispersion near high node load. As fluctuations in processor time and RAM are often different in nature, the coefficients for each resource are evaluated independently.

Equation (2.3) corresponds to the Itô stochastic differential equation [22]:

$$dX_t = A(X_t)dt + \sqrt{B(X_t)}dW_t, \quad (2.7)$$

where W_t is the standard Wiener process. Equation (2.7) lends itself well to implementation using the Monte Carlo method, while equation (2.3) describes

the evolution of the density of the entire ensemble of trajectories simultaneously and is used to calculate the probability of overload.

Boundary value problem for the probability of overload

Let $x \in [0, L]$, $L = 1$. We define the initial condition as follows:

$$\rho(x, 0) = \delta(x - x_0), \quad x_0 \in (0, L), \quad (2.8)$$

where $\delta(\cdot)$ is the Dirac delta function.

The probability flow for equation (2.3) is given by:

$$J(x, t) = A(x)\rho(x, t) - \frac{1}{2} \frac{\partial}{\partial x}(B(x)\rho(x, t)). \quad (2.9)$$

Let us assume the following boundary conditions:

$$J(0, t) = 0 \text{ (reflection) and } \rho(L, t) = 0 \text{ (absorption)}. \quad (2.10)$$

The first condition prevents the process from entering the region $x < 0$, while the second condition means that when the critical limit is reached, the process leaves the computational area, and the corresponding probability mass is irreversibly lost.

Let us denote the probability that, by time t , the boundary L has not been reached:

$$P_{\text{surv}}(t) = \int_0^L \rho(x, t) dx. \quad (2.11)$$

Then there is an additional risk of overload:

$$R(t) = 1 - \int_0^L \rho(x, t) dx. \quad (2.12)$$

Note that $R(t)$ is determined not only by the current load level, x_0 , but also by the coefficients, $A(x)$ and $B(x)$. Therefore, two nodes with the same load may have different risk estimates. For two resources, the integral node risk is defined as $R_j(t) = \max_r R_{j,r}(t)$, $r \in \{\text{CPU}, \text{MEM}\}$, where CPU represents the central processing unit, while MEM represents memory (RAM).

Analytical solutions

Consider a special case of constant coefficients that allows an explicit analytical solution:

$$\frac{\partial \rho}{\partial t} = a \frac{\partial^2 \rho}{\partial x^2} - b \frac{\partial \rho}{\partial x}, \quad a > 0, \quad b \in \mathbb{R}. \quad (2.13)$$

The probability flow is expressed as $J = b\rho - a\rho_x$, where ρ_x is the partial derivative of $\rho/\partial x$ and conditions (2.10) are given as:

$$-a\rho_x(0, t) + b\rho(0, t) = 0, \quad \rho(L, t) = 0. \quad (2.14)$$

We introduce a substitution that eliminates the convective term:

$$\rho(x, t) = e^{cx} u(x, t), \quad c = \frac{b}{2a}. \quad (2.15)$$

Substituting (2.13) into the equation yields the equation for thermal conductivity with potential:

$$\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} - \gamma u, \quad \gamma = \frac{b^2}{4a}. \quad (2.16)$$

The conditions (2.14) become:

$$u_x(0, t) = cu(0, t), \quad u(L, t) = 0. \quad (2.17)$$

The initial condition for u takes the form: $u(x, 0) = e^{-cx_0} \delta(x - x_0)$.

By separating the variables $u(x, t) = \phi(x)e^{-\lambda t}$, we arrive at the Sturm–Liouville problem:

$$\phi''(x) + k^2 \phi(x) = 0, \quad \lambda = ak^2 + \gamma \quad (2.18)$$

with boundary conditions:

$$\phi'(0) = c\phi(0), \quad \phi(L) = 0. \quad (2.19)$$

Considering the first boundary condition, the general solution (2.18) can be expressed as follows:

$$\phi(x) = \cos(kx) + \frac{c}{k} \sin(kx). \quad (2.20)$$

From the condition $\phi(L) = 0$, we obtain a transcendental equation for the spectrum:

$$\tan(kL) = -\frac{k}{c}. \quad (2.21)$$

Let k_n , $n = 1, 2, \dots$, be the positive roots of equation (2.21). The eigenfunctions

$$\phi_n(x) = \cos(k_n x) + \frac{c}{k_n} \sin(k_n x) \quad (2.22)$$

are orthogonal on the segment $[0, L]$ and form a complete basis in $L_2(0, L)$.

Decomposing the initial condition by $\{\phi_n\}$ and returning to the variable ρ using formula (2.15), we obtain:

$$\rho(x, t) = e^{c(x-x_0)} \sum_{n=1}^{\infty} \frac{\phi_n(x_0)\phi_n(x)}{\|\phi_n\|^2} e^{-(ak_n^2 + \gamma)t}, \quad (2.23)$$

where $\|\phi_n\|^2 = \int_0^L \phi_n^2(x) dx$. Equation (2.23) provides an

exact solution to problems (2.13), (2.8), and (2.14). In practice, the series is truncated at $n \leq M$ and the roots of k_n are determined numerically.

We also consider an alternative formulation with Dirichlet conditions on both boundaries:

$$\rho(0, t) = 0, \quad \rho(L, t) = 0. \quad (2.24)$$

Here, zero load ($x = 0$) is considered as limiting and practically unattainable. This setting is natural for communication channels and nodes for which zero load is also a boundary condition.

Substituting $\rho = e^{cx}u$ and $c = b/(2a)$ results in (2.16) with homogeneous Dirichlet conditions $u(0, t) = u(L, t) = 0$. The eigenfunctions $\psi_n(x) = \sin(\pi nx/L)$ and eigenvalues $k_n = \pi n/L$ are known explicitly. The solution can be expressed as follows:

$$\rho(x, t) = \frac{2}{L} e^{\frac{b(x-x_0)}{2a}} \times \sum_{n=1}^M \sin\left(\frac{\pi nx_0}{L}\right) \sin\left(\frac{\pi nx}{L}\right) e^{-\left(\pi^2 n^2 a/L^2 + b^2/(4a)\right)t}. \quad (2.25)$$

The probability of not reaching the boundary, along with the associated risk, is calculated using formulas (2.11) and (2.12). Unlike in equation (2.23), the spectrum is clearly known, so a numerical solution to a transcendental equation is not required. The choice between the two statements depends on the physics of the problem: if the probability density at $x = 0$ can be nonzero, a reflective boundary is appropriate, whereas if zero loading is the limiting mode, the Dirichlet condition applies.

Numerical solution: Crank–Nicolson scheme

Since there is no closed solution for arbitrary $A(x)$ and $B(x)$, numerical methods must be applied [10, 24, and 25].

A uniform grid on $[0, L]$ can be introduced:

$$x_i = i\Delta x, \quad i = \overline{0, N_x}, \quad \Delta x = \frac{L}{N_x}, \quad (2.26)$$

where Δx is the grid pitch and N_x is the number of nodes in the spatial grid.

The diffusion term is approximated by the central difference:

$$\frac{\partial^2}{\partial x^2} (B\rho)|_{x_i} \approx \frac{(B\rho)_{i+1} - 2(B\rho)_i + (B\rho)_{i-1}}{(\Delta x)^2}, \quad (2.27)$$

and the convective term by the difference:

$$\frac{\partial}{\partial x} (A\rho)|_{x_i} \approx \frac{(A\rho)_{i+1} - (A\rho)_{i-1}}{2\Delta x}. \quad (2.28)$$

To approximate the condition $J(0, t) = 0$, we introduce a ghost point, x_{-1} . The derivative on the boundary is given by:

$$\frac{\partial}{\partial x} (B\rho)|_{x_0} \approx \frac{(B\rho)_1 - (B\rho)_{-1}}{2\Delta x}. \quad (2.29)$$

Therefore,

$$A_0 \rho_0 - \frac{1}{2} \cdot \frac{(B\rho)_1 - (B\rho)_{-1}}{2\Delta x} = 0, \quad (2.30)$$

and the value at the ghost point is:

$$(B\rho)_{-1} = (B\rho)_1 - 4\Delta x A_0 \rho_0. \quad (2.31)$$

By substituting this expression into the equation at node $i = 0$, we obtain the three-point pattern.

We can then use the Crank–Nicolson scheme to advance in time. Unconditional stability, second order in both variables:

$$\frac{\rho_i^{n+1} - \rho_i^n}{\Delta t} = \frac{1}{2} \left(\mathcal{L}(\rho^{n+1})_i + \mathcal{L}(\rho^n)_i \right), \quad (2.32)$$

where $\mathcal{L}$ is a discrete Fokker–Planck operator. Averaging between two-time layers yields an implicit scheme. Due to the three-point pattern, a system with a three-diagonal matrix emerges at each time step:

$$a_i \rho_{i-1}^{n+1} + b_i \rho_i^{n+1} + c_i \rho_{i+1}^{n+1} = d_i. \quad (2.33)$$

This is solved by the run-through method [26] in $O(N_x)$ operations.

The probability mass flows monotonically over the absorbing boundary $x = L$; the $P_{\text{surv}}(t)$ integral decreases and $R(t) = 1 - P_{\text{surv}}(t)$ increases, reflecting the accumulation of risk. With grid parameters of $N_x = 200$ and $N_t = 500$, calculating $R(T_h)$ takes fractions of a millisecond, allowing it to be integrated into the scheduler cycle.

Estimation of coefficients based on monitoring data

Now, let us turn to the estimation of the coefficients $A(x)$ and $B(x)$ from monitoring data⁴. After fixing the aggregation step Δt , the segment $[0, 1]$ is divided into M_b uniform intervals, $\mathcal{B}_m$. The estimates are based on the definition of conditional moments:

$$A(x_m) = \frac{\mathbb{E}[\Delta x \mid x \in \mathcal{B}_m]}{\Delta t}, B(x_m) = \frac{\mathbb{E}[(\Delta x)^2 \mid x \in \mathcal{B}_m]}{\Delta t}, \quad (2.34)$$

with linear interpolation between the centers of the intervals and a lower bound of $B(x) \geq B_{\min} = 10^{-6}$, which guarantees the stability of the numerical scheme. In the experiments, the number of intervals (bins) used to divide the $[0, 1]$ interval is set to $M_b = 20$, while the aggregation step Δt coincides with the telemetry polling interval (15 s). The coefficients are recalculated every ten steps over a sliding window of one hundred observations.

EXPERIMENTAL ENVIRONMENT

The methodology was tested on a discrete-event *Kubernetes* cluster simulator. The environment parameters are shown in Table 1.

Table 1. Configuration of the experimental environment

Parameter	Value
Cluster	10 nodes $\times$ (8000m CPU, 16 GB RAM ⁵)
Number of modeling steps	3 600
Number of warm-up steps	100
Number of runs (seeds)	100
N_x (x-grid)	200
N_t (t-grid)	500
T_h (forecast horizon)	300
Number of pods	40–50

The simulator reproduces the main *Kubernetes* mechanisms, including proportional processor time allocation (the Completely Fair Scheduler, CFS), Kubelet out-of-memory eviction, QoS⁶ priority consideration, and the pod lifecycle finite automaton (pending $\rightarrow$ starting $\rightarrow$ running $\rightarrow$ termination). The load is generated using 11 profiles. These profiles are either stationary, pulsating (Poisson bursts with lognormal amplitude), returning to the average (Ornstein–Uhlenbeck), diffusion (a remote-control system with controlled drift), or a combination of these. Each experiment is repeated 100 times.

Comparing the model with empirical data

To verify the adequacy of the model, the distribution function of the time taken to reach critical load was compared with an empirical estimate (Fig. 1). According to the Kolmogorov–Smirnov metric, the deviation does not exceed 0.05 for time horizons of up to several minutes.

In a scenario involving realistic production loads (50 pods of six different types, including Poisson bursts, stationary profiles, and diffusion profiles), the risk $\bar{R} \approx 0.03$ corresponds to single stochastic displacements, and no system overload occurs. In contrast, a RAM shortage scenario involving 38 pods with total memory consumption exceeding the cluster capacity $\bar{R} \approx 0.26$ demonstrates the impossibility of providing uninterrupted operation without horizontal scaling.

⁴ Prometheus Documentation. Cloud Native Computing Foundation. <https://prometheus.io/docs/>. Accessed May 11, 2026.

⁵ Random access memory.

⁶ Quality of service.

Sensitivity of risk assessment to model parameters

The stability of the $R(T_h)$ estimate has been investigated by varying the T_h forecast horizon and the $R_{\max}$ threshold for acceptable risk.

With fixed drift and diffusion coefficients, the value of $R(T_h)$ remains close to zero with an average utilization of around 28% (moderate load) across all T_h values in the range $\{100, 200, 300, 500\}$. This is consistent with the process being far from the critical boundary. In the pulsating load scenario (70% stationary and 30% stochastically pulsating bursts), the $R_{\max}$ threshold significantly affects behavior: at $R_{\max} = 0.1$, the model initiates an average of 39.7 proactive migrations per run, whereas at $R_{\max} = 0.3$, this figure is fewer than 0.1 migrations per run. In both cases, the frequency of service requirement violations does not exceed 0.1%. The monotonic decrease in the number of migrations as the threshold increases confirms the adequate ranking of states by the model according to their proximity to overload.

Verifying the numerical solution

The accuracy of the numerical method was confirmed by comparing it with the analytical spectral solution (2.23) across six test cases involving various combinations of drift and diffusion coefficients (see Table 2). The initial condition is set as a localized distribution $\rho(x, 0) = \delta(x - x_0)$, which is approximated by a Gaussian function with a width parameter $\sigma = 2\Delta x$, where $\Delta x = L/N_x$ is the step of the spatial grid.

The high accuracy of the implemented scheme is indicated by the maximum discrepancy between the numerical and analytical solutions of less than 0.0007. The simulation confirms a monotonic decrease in the probability of not reaching the $P_{\text{surv}}(t)$ boundary, fulfilment of the absorbing condition at the right boundary, and convergence of the solution as the number of spatial grid nodes (N_x) increases.

Figure 2 shows the evolution of the probability density $\rho(x, t)$ at constant coefficients. As t increases, the distribution spreads under the influence of diffusion and loses mass at the absorbing boundary $x = L = 1$, reflecting the accumulation of the probability of overload.

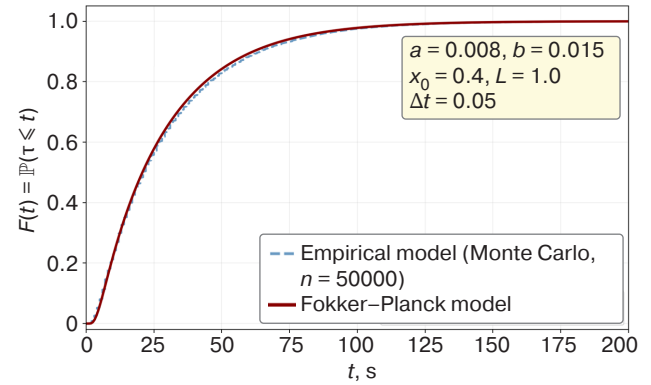


Fig. 1. Comparison of empirical (dotted line) and model (solid line) time distribution functions for reaching critical load

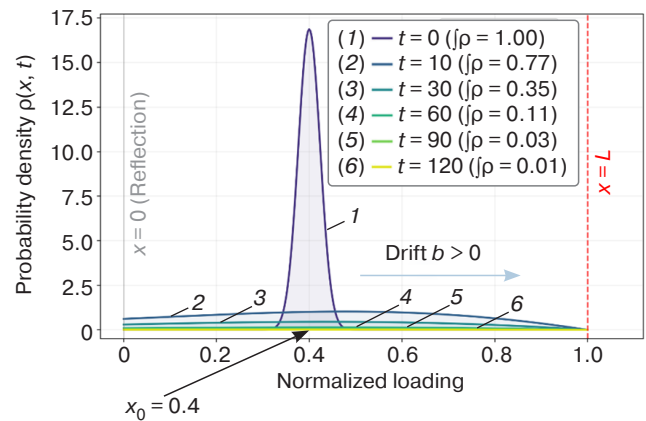


Fig. 2. Evolution of probability density $\rho(x, t)$ at constant drift and diffusion coefficients. As t increases, the distribution loses mass at the absorbing boundary $x = L = 1$

Table 2. Comparison of numerical (R_{num}) and analytical (R_{an}) solutions

N_x	a	b	x_0	T	R_{num}	R_{an}	$ \Delta $
1	0.005	0.01	0.5	100	0.0498	0.0503	0.0005
2	0.01	0.02	0.3	200	0.4412	0.4419	0.0007
3	0.01	0.0	0.5	150	0.0266	0.0268	0.0002
4	0.005	0.05	0.7	50	0.2877	0.2882	0.0005
5	0.02	0.01	0.4	300	0.4851	0.4855	0.0004
6	0.008	−0.005	0.6	80	0.0115	0.0118	0.0003

EXPERIMENTAL VERIFICATION OF RISK FORECASTING

To assess the practical applicability of the proposed model, a systematic, rolling-window, out-of-sample evaluation was conducted. This methodology involves the formation of training windows sequentially to estimate the model's coefficients, as well as test horizons to verify forecasts. Four widely used predictive approaches in computing cluster monitoring systems have been considered for basic comparison.

Experimental protocol

The discrete event simulator described above was used to generate the dataset. The simulator is run in four different load scenarios, each having a different dynamic nature. The four scenarios are linear growth (ramp pressure), RAM load (memory pressure), oscillating load (oscillating pressure), and a realistic profile that combines Poisson bursts with stationary and mean-reverting modes (mixed realistic profile). Each scenario is repeated with 20 different initialization values. To simulate real telemetry noise, interference with normally distributed variances of $\sigma_{\text{CPU}} = 0.02$ and $\sigma_{\text{MEM}} = 0.015$ has been added to all time series.

As a result, a cumulative dataset covering 170 unique nodes was obtained. Based on the coefficient of variation criterion (greater than 0.15), 44 of these nodes are classified as stable, while 126 are classified as nonstationary. The dataset comprises a total of approximately 500000 observations taken at 15-s intervals over 3600 simulation steps, which is equivalent to approximately 15 h.

For each node-resource pair, a series of sliding windows has been formed. A training window of 240 steps (60 min) has been used to evaluate the model's coefficients. The forecast quality has then been assessed at test horizons of $T_h \in \{4, 8, 20, 40\}$ steps (1, 2, 5, and 10 min) with a shift of 4 steps. The label $Y = 1$ is assigned to the window if the load reaches the threshold $L = 0.95$ at least once at the horizon. At this threshold, 904800 windows were formed, approximately 20.9% of which contained an overload event. This provides a sufficient balance of classes for the correct assessment of metrics.

Forecasting methods

The Fokker–Planck (FP) model. In each training window, the drift coefficients $A(x)$ and diffusion coefficients $B(x)$ are evaluated first. The segment $[0, 1]$ is divided into 30 uniform intervals in each of which the conditional transition moments are calculated. The resulting profile is then smoothed using the Savitzky–Golay filter. Next, the local values $A(x_0)$ and $B(x_0)$ are

extracted for the current load x_0 using Gaussian weighting over ± 2 adjacent intervals. This procedure increases the stability of estimates in areas where there is little data. After this, the Fokker–Planck equation is solved using the constant coefficients found on the space-time grid ($N_x = 100$, $N_t = 200$) according to the Crank–Nicolson scheme. Since the overload threshold L may differ from unity, scaling is performed before the solution is found: $x' = x/L$, $a' = a/L^2$, $b' = b/L$, which brings the absorbing boundary to the standard value. The

risk is then calculated as $R(T_h) = 1 - \int_0^1 \rho(x', T_h) dx'$. To

improve calibration, isotonic regression is applied to the obtained estimates, which are trained using a cross-validation scheme on two halves of the time series.

Persistence (naïve forecast). The simplest deterministic method is to set $P = 1$ for $x(t) \geq L$ and $P = 0$ otherwise. This reflects the assumption that the immutable current state is used as the lower limit of quality.

Linear trend. The linear regression $\hat{x}(t) = a + bt$ is fitted to the training window. The value of $\hat{x}(t + T_h)$ is then extrapolated to calculate the probability of overload using the normal distribution function: $P = 1 - \Phi((L - \hat{x}) / \sigma_{\text{res}})$, where σ_{res} is the standard deviation of the regression residuals.

Markov process of the first order autoregression (AR) (1). The model $x(t + 1) = c + \phi x(t) + \varepsilon$, where the parameters c , ϕ , and σ_ε^2 are estimated from the training window. Forecast for the horizon of h steps: $\mu_h = \mu_\infty + \phi^h (x_{\text{last}} - \mu_\infty)$, $\sigma_h^2 = \sigma_\varepsilon^2 (1 - \phi^{2h}) / (1 - \phi^2)$, $P = 1 - \Phi((L - \mu_h) / \sigma_h)$.

Historical exceedance. Empirical estimate: P is equal to the share of observations in the training window for which $x \geq L$; this is a nonparametric method that does not use information about the current state of x_0 .

Quality metrics

The quality of probabilistic forecasts is assessed using four metrics:

Brier score:

$$\text{BS} = \frac{1}{N} \sum (p_i - y_i)^2,$$

where p_i is the predicted probability, and $y_i \in \{0, 1\}$ is the actual outcome [27]. This measures the overall quality, including calibration and discrimination ability. The range is $[0, 1]$, and $\text{BS} = 0$ is ideal.

ROCAUC^7 : the area under the <<True Positive Rate–False Positive Rate>> curve, which characterizes the

⁷ The area under the ROC curve. The receiver operating characteristic (ROC) is a graph showing the ratio between the share of objects correctly classified (True Positive Rate) and the share of objects incorrectly classified (False Positive Rate).

Table 3. Comparison of forecasting methods (threshold $L = 0.95$, all horizons)

Method	Brier Score ↓	Log Loss ↓	ROC AUC ↑	PR AUC ↑	ECE ↓
FP model	0.0245	0.0876	0.9915	0.9754	0.0028
Persistence	0.0372	1.2842	0.9139	0.8576	0.0372
AR(1)	0.0383	0.1285	0.9945	0.9824	0.0447
Historical exceedance	0.0412	0.1606	0.9879	0.9627	0.0376
Linear trend	0.0419	0.1545	0.9908	0.9682	0.0394

Table 4. Brier score of the FP model and the best baseline method over horizons

T_h , min	FP	Best base method	ΔBS	Method
1	0.0126	0.0197	−0.0071	Persistence
2	0.0154	0.0272	−0.0118	Persistence
5	0.0261	0.0423	−0.0162	Persistence
10	0.0438	0.0551	−0.0113	Historical exceedance

ability of the model to rank positive examples above negative ones, and does not depend on calibration.

PR AUC⁸: the area under the <<Precision – Recall>> curve. Unlike the ROC AUC, it is sensitive to class imbalances and penalizes false positives.

ECE (Expected calibration error)⁹:

$$ECE = \sum_m \frac{n_m}{N} |\bar{p}_m - \bar{y}_m|,$$

where n_m is the number of observations that fall within the m th interval of predicted probabilities, and the summation is carried out over 10 uniform intervals of predicted probabilities. This measures the consistency between predicted probabilities and observed frequencies, where $ECE = 0$ indicates perfect calibration.

RESULTS

Table 3 presents the results of an aggregated comparison across all horizons, where the arrows ↓/↑ indicate the direction of improvement and Log Loss is a logarithmic loss function.

⁸ The precision–recall area under the curve is the area under the PR curve, where P stands for precision and R stands for recall.

⁹ The Expected Calibration Error (ECE) metric measures the degree of correspondence between the probabilities predicted by the model and the actual probabilities of outcomes.

The FP model achieves the best Brier score (0.0245). The difference with the nearest competitor (Persistence: 0.0372) equates to a 34.1% reduction in error. The calibration of the model ($ECE = 0.0028$) is an order of magnitude higher than that of the basic methods, indicating high reliability of the predicted probabilities.

Table 4 shows the results for different forecasting horizons.

The FP model outperforms the basic methods on all four horizons. The greatest advantage is observed at $T_h = 5$ min ($\Delta BS = -0.0162$), where the basic methods fail to cope with nonlinear dynamics. On a short horizon ($T_h = 1$ min), the FP model achieves a Brier score of 0.0126, close to ideal accuracy. Even without post-processing calibration, this model demonstrates high discriminatory ability (ROC AUC = 0.9840, PR AUC = 0.9582), surpassing all basic methods except AR(1). However, a noticeable calibration violation ($ECE = 0.1526$) leads to an increase in the Brier score up to 0.1013. This is because the solution to the boundary value problem introduces a systematic error by approximating the drift and diffusion coefficients to systematically overestimate the predicted probabilities at moderate loads. Isotonic regression corrects this by monotonically rebuilding the scale of predicted probabilities while maintaining—and even slightly improving—the ranking (ROC AUC: 0.9840 → 0.9915). Therefore, the FP model’s primary

Table 5. Brier score by subgroup (resource × node type)

Subgroup	FP	Best base method	Method	Is FP better?
CPU / stable	0.0486	0.0677	Historical exceedance	Yes
CPU / nonstationary	0.0044	0.0062	Persistence	Yes
MEM / stable	0.0546	0.0915	AR(1)	Yes
MEM / nonstationary	0.0264	0.0366	AR(1)	Yes

contribution consists in its ability to distinguish between windows with impending overload and safe ones, while isotonic calibration ensures that the predicted probabilities are consistent with the actual frequencies.

Subgroup analysis (Table 5) reveals the advantages of the model.

The FP model consistently outperforms the basic methods across all four subgroups. Even in the CPU subgroup on nonstationary nodes, where the Persistence method achieves a high level of quality (BS = 0.0062), the isotonic-calibrated FP model still achieves a 29% improvement with BS = 0.0044. The FP model's greatest advantage is observed on stable nodes, where load dynamics do not follow simple

patterns, but can be well described by a stochastic process. For MEM/stable, $\Delta BS = -0.0369$ (a 40% improvement).

Figure 3 shows calibration curves that reflect the consistency of the predicted probabilities with the observed frequencies. Figure 4 illustrates the model's dynamics by presenting three load scenarios alongside a combined load time series, $x(t)$, and a risk assessment, $R(t)$, at a horizon of $T_h = 5$ min. In each case, $R(t)$ reflects changes in $x(t)$: with an increasing load, the risk rises in advance to approach ≈ 1 at the $L = 0.95$ threshold; with a fluctuating load, the risk rises and falls in sync; with a low load, it stays close to zero. This confirms the alignment of the predicted risk with the observed data.

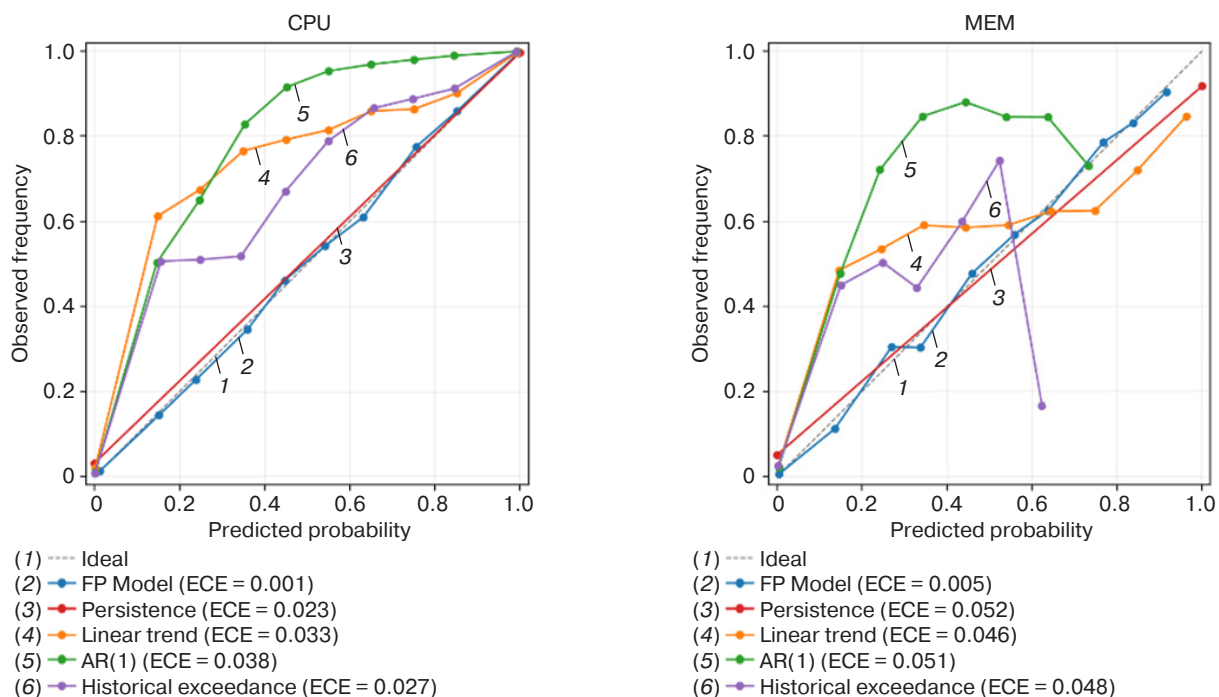


Fig. 3. Calibration curves: predicted overload probability and observed frequency ($L = 0.95$).
The diagonal line represents the ideal calibration

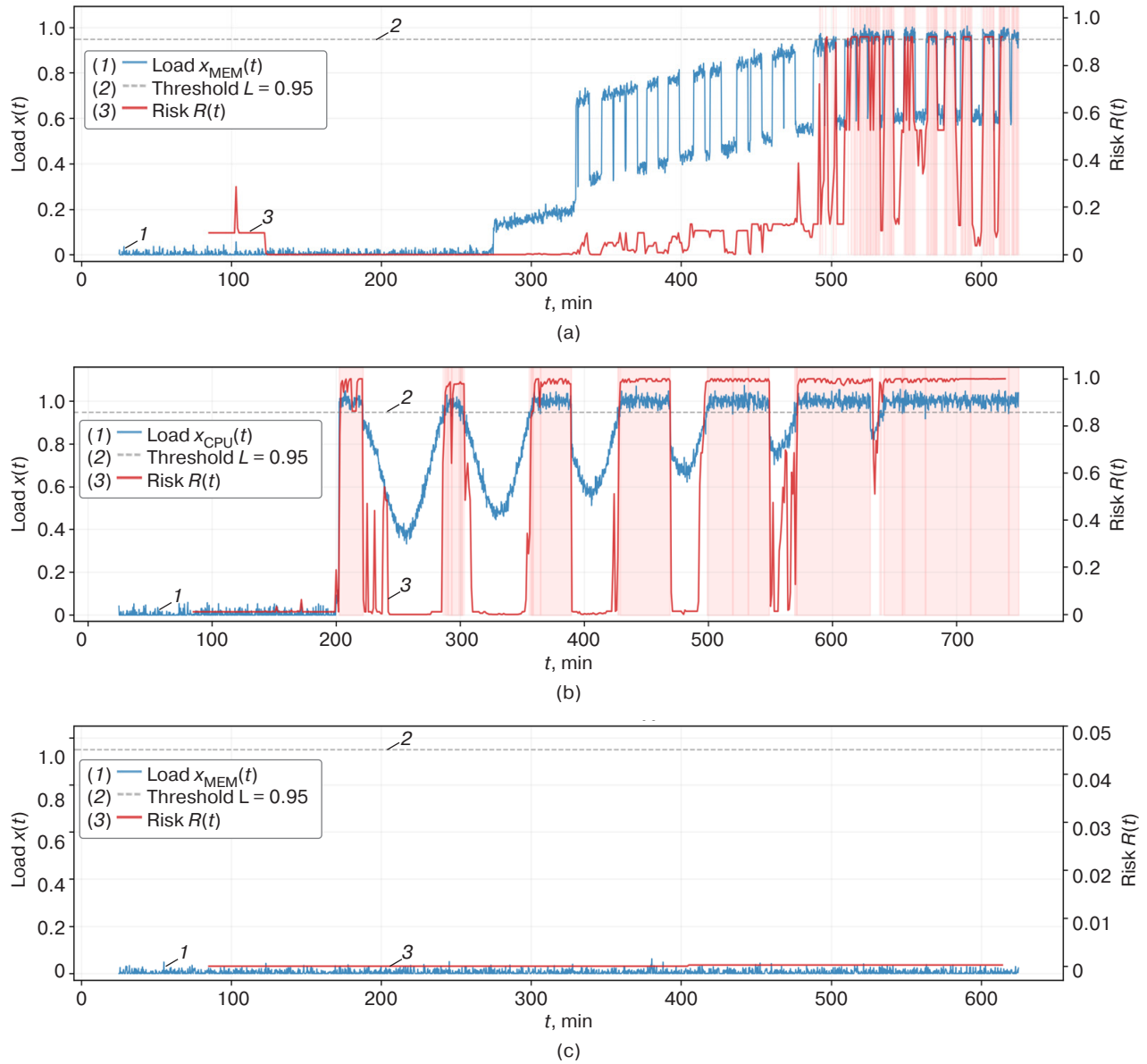


Fig. 4. Dynamics of node load and risk assessment for three characteristic fragments:
(a) increasing MEM load to the threshold; (b) oscillating CPU load; (c) consistently low load

ANALYSIS OF THE RESULTS

The high performance of the FP model can be attributed to three factors:

1. *Consideration of the stochastic load structure.* Unlike basic methods such as linear trend and AR(1), which reduce uncertainty to Gaussian noise with fixed variance, the Fokker–Planck equation clearly distinguishes between two sources of dynamics: a directional trend (drift A) and the intensity of random deviations (diffusion B). Since coefficients $A(x_0)$ and $B(x_0)$ are extracted locally, the FP model adapts to the specific operating mode of the node in the vicinity of the current load level to capture an increase in volatility near full

load, which the basic approaches systematically underestimate.

2. *Correct formulation of the overload problem.* The fundamental difference lies in the definition of the estimated value itself. Basic methods calculate the probability that the load will exceed the threshold at the fixed point in time $\mathbb{P}(x(t + T_h) \geq L)$, whereas the FP model estimates the probability of at least one threshold crossing over the entire forecast horizon: $\mathbb{P}(\exists s \in (0, T_h]: x(s) \geq L)$. Since even short-term loads exceeding the acceptable limit are sufficient to cause performance degradation, the second formulation corresponds with the practical understanding of risk by providing more conservative—and, at the same time, more accurate—estimates.

3. *The role of forecast calibration.* The practical value of a probabilistic model is determined both by its discriminative ability and the extent to which the predicted probabilities are consistent with the observed frequencies. By comparing the results of the FP model before and after isotonic calibration, the contribution of the actual physical model (discrimination) can be separated from that of postprocessing (calibration). Even without calibration, the FP model demonstrates high ranking quality (ROC AUC = 0.9840, PR AUC = 0.9582), surpassing most basic methods in these indicators. However, the “raw” probabilities are systematically overestimated (ECE = 0.1526), resulting in a significant deterioration in the Brier score (0.1013 vs 0.0245). Isotonic regression corrects the calibration (ECE: 0.1526 → 0.0028) without altering the ranking order and reduces the final Brier score by 76%. Particular attention should be given to the CPU subgroup on nonstationary nodes, where the persistence method traditionally produces high-quality results (BS = 0.0062) due to an almost deterministic two-level loading structure. Nevertheless, these results also demonstrate the potential of an ensemble approach. In future work, it would be promising to enable each node to automatically select a forecasting method based on the characteristics of its time series, thereby combining the advantages of stochastic and deterministic models.

CONCLUSIONS

This paper proposes a probabilistic model for the risk of overload *Kubernetes* compute nodes based on interpreting overload as the first event of reaching a critical boundary. Normalized node loading is described as a one-dimensional stochastic process, while the evolution of the probability density is formulated within the framework of the Fokker–Planck equation. The choice of boundary conditions is justified by the physics of the container environment.

A boundary value problem is formulated for the constructed model, for which an analytical solution is obtained in the case of constant coefficients and mixed Robin–Dirichlet boundary conditions. An alternative formulation is considered having Dirichlet conditions on both boundaries. For the case of variable coefficients, a numerical method based on the Crank–Nicolson scheme with ghost-point approximation for the reflecting boundary is proposed. Over six test cases, the discrepancy between the numerical and analytical solutions does not exceed 0.001 at a computational cost of less than a millisecond per solution.

The developed method for estimating drift and diffusion coefficients, which is based on telemetry data, has a controlled frequency of recalculation. Initial testing on a discrete event simulator using a cluster of 10 nodes with realistic load profiles confirms the adequacy of the model estimates: the value of $\bar{R} \approx 0.03$ corresponds to isolated incidents under normal load, while $\bar{R} \approx 0.26$ correctly signals a systemic shortage of resources.

Systematic experimental verification using the sliding window method on an extended dataset (170 nodes, 4 load scenarios, and 904800 test windows) confirms the practical advantage of the FP model over four basic forecasting methods. The FP model achieves the best Brier score (0.0245), surpassing all the basic methods at each of the four forecasting horizons (1, 2, 5, and 10 min) and in all 4 subgroups of nodes. Indicators of discrimination (ROC AUC = 0.9915) and calibration (ECE = 0.0028) confirms the high quality of the probabilistic forecasts, which are suitable for use in decision-making systems.

The scientific novelty of the work is as follows:

1. The proposed model is suitable for estimating the risk of overloading a *Kubernetes* node in terms of the probability of surpassing the critical load threshold for the first time. The evolution of the load distribution is described by the Fokker–Planck equation. The choice of boundary conditions is justified by the specifics of the container environment.
2. A method has been developed for estimating drift and diffusion coefficients from telemetry data with a controlled recalculation frequency.
3. The model has been experimentally verified on a discrete-event cluster simulator with realistic load profiles.
4. A systematic comparison has been carried out with four basic forecasting methods using the sliding window method. This showed the advantage of the proposed approach in terms of calibration and discrimination metrics.

These results form a basis for developing risk-oriented node ranking mechanisms and integrating probabilistic risk assessment into container cluster management tools. The adaptive choice of a forecasting method based on the characteristics of a time series is a promising area for further research. Promising areas for further research include the ensemble approach, multidimensional load modeling for several resources simultaneously, and model evaluation based on industrial cluster data using Prometheus telemetry.

Authors' contributions

S.A. Lesko—research concept, methodology, development of the mathematical model, scientific supervision, problem formulation, interpretation of results, and manuscript editing.

V.N. Kalinin—development of the mathematical model and algorithms, software implementation, data processing, computational experiments, verification and analysis of results.

D.O. Zhukov—research concept, methodological oversight, scientific consulting, verification of the correctness of results, critical editing, and approval of the final version of the manuscript.

REFERENCES

1. Ahmad I., Alfaiakawi M.Gh., AlMutawa A., Alsalmán L. Container scheduling techniques: A survey and assessment. *J. King Saud Univ. Comput. Inf. Sci.* 2022;34(7):3934–3947. <https://doi.org/10.1016/j.jksuci.2021.03.002>
2. Carrión C. Kubernetes scheduling: Taxonomy, ongoing issues and challenges. *ACM Comput. Surv.* 2023;55(7):138.
3. Bashir N., Deng N., Rzađca K., Irwin D., Kodak S., Jnagal R. Take it to the limit: Peak prediction-driven resource overcommitment in datacenters. In: *Proceedings of the 16th European Conference on Computer Systems (EuroSys '21)*. New York: ACM; 2021. P. 556–573.
4. Cohen M.C., Keller P., Mirrokni V., Zadimoghaddam M. Overcommitment in cloud services: Bin packing with chance constraints. *Management Sci.* 2019;65(7):3255–3271.
5. Tirmazi M., Barker A., Deng N., Haque M.E., Qin Z.G., Hand S., et al. Borg: the next generation. In: *Proceedings of the 15th European Conference on Computer Systems (EuroSys '20)*. New York: ACM; 2020. P. 1–14.
6. Lu C., Ye K., Xu G., Xu C.-Z., Bai T. Imbalance in the cloud: An analysis on Alibaba cluster trace. In: *2017 IEEE International Conference on Big Data (BigData)*. IEEE; 2017. P. 2884–2892. <https://doi.org/10.1109/BigData.2017.8258257>
7. Redner S. *A Guide to First-Passage Processes*. Cambridge: Cambridge University Press; 2001, 312 p.
8. Artime O., Khalil N., Toral R., San Miguel M. First-passage distributions for the one-dimensional Fokker–Planck equation. *Phys. Rev. E*. 2018;98(4):042143. <https://doi.org/10.1103/PhysRevE.98.042143>
9. Gardiner C.W. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. 4th ed. Berlin: Springer; 2009, 447 p.
10. Risken H. *The Fokker–Planck Equation: Methods of Solution and Applications*. 2nd ed. Berlin: Springer; 1996, 472 p.
11. Senjab K., Abbas S., Ahmed N., et al. A survey of Kubernetes scheduling algorithms. *J. Cloud Comput.* 2023;12(1):87. <https://doi.org/10.1186/s13677-023-00471-1>
12. Burns B., Grant B., Oppenheimer D., Brewer E., Wilkes J. Borg, Omega, and Kubernetes. *Commun. ACM*. 2016;59(5):50–57.
13. Moritz de Carvalho Neto P., Castro M., Siqueira F. Dynamic load balancing in Kubernetes Environments with Kubernetes Scheduling Extension (KSE). *Concurr. Comput. Pract. Exp.* 2025;37(3):e8344. <https://doi.org/10.1002/cpe.8344>
14. Menouer T., Cérin C., Darmon P. Reactive autoscaling of Kubernetes nodes. In: *FRAME'24: Proceedings of the 4th Workshop on Flexible Resource and Application Management on the Edge*. New York: ACM; 2024. P. 31–38.
15. Yuan H., Liao S. A time series-based approach to elastic Kubernetes scaling. *Electronics*. 2024;13(2):285. <https://doi.org/10.3390/electronics13020285>
16. Nunes J.P.K.S., Nejati S., Sabetzadeh M., Nakagawa E.Y. Self-adaptive, requirements-driven autoscaling of microservices. In: *19th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS 2024)*. IEEE/ACM; 2024. P. 168–174.
17. Guruge P.B., Priyadarshana Y.H.P.P. Time series forecasting-based Kubernetes autoscaling using Facebook Prophet and Long Short-Term Memory. *Front. Comput. Sci.* 2025;7:1509165. <https://doi.org/10.3389/fcomp.2025.1509165>
18. Santos J., Reppas E., Wauters T., Volckaert B., De Turck F. Gwydion: Efficient auto-scaling for complex containerized applications in Kubernetes through reinforcement learning. *J. Netw. Comput. Appl.* 2025;234:104067. <https://doi.org/10.1016/j.jnca.2024.104067>
19. Yu L., Chen L., Cai Z., Shen H., Liang Y., Pan Y. Stochastic load balancing for virtual resource management in datacenters. *IEEE Trans. Cloud Comput.* 2020;8(2):459–472. <https://doi.org/10.1109/TCC.2016.2525984>
20. Marbukh V. On systemic risk in the cloud computing model. In: *26th International Teletraffic Congress (ITC 26)*. Karlskrona: IEEE; 2014. <https://doi.org/10.1109/ITC.2014.6932938>
21. Chamberlain J., Zheng J., Zhu Z., Liu Z., Starobinski D. Exploiting Kubernetes autoscaling for economic denial of sustainability. In: *Proceedings of the ACM Measurement Analysis Computing Systems*. 2025;9(2):22.
22. Belopol'skaya Ya.I. *Stokhasticheskie differentsial'nye uravneniya. Prilozheniya k zadacham matematicheskoi fiziki i finansovoi matematiki (Stochastic Differential Equations. Applications to Problems of Mathematical Physics and Financial Mathematics)*. St. Petersburg: Lan; 2023, 308 p. (In Russ.). ISBN 978-5-507-47129-4
23. Lesko S.A., Zhukov D.O., Istratov L.A. Models of describing the dynamics of blocking nodes of computer networks by viruses based on the use of percolation, kinetic and stochastic methods. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika = Tomsk State University Journal of Control and Computer Science*. 2020;52:22–32 (in Russ.). <https://doi.org/10.17223/19988605/52/3>
24. Hwang H.J., Jang J., Jung J. The Fokker–Planck equation with absorbing boundary conditions in bounded domains. *SIAM J. Math. Anal.* 2018;50(2):2194–2232.
25. Samarskii A.A. *Teoriya raznostnykh skhem (Theory of Difference Schemes)*. Moscow: Nauka; 1989, 616 p. (In Russ.). ISBN 5-02-014576-9
26. Thomas L.H. *Elliptic Problems in Linear Difference Equations over a Network*. Watson Scientific Computing Laboratory Report. New York: Columbia University; 1949, 71 p.
27. Brier G.W. Verification of forecasts expressed in terms of probability. *Monthly Weather Rev.* 1950;78(1):1–3.

About the Authors

Sergey A. Lesko, Dr. Sci. (Eng.), Associate Professor, Professor, Department of Information and Analytical Cybersecurity Systems, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: lesko@mirea.ru. Scopus Author ID 57189664364, ResearcherID AAF-5651-2019, <https://orcid.org/0000-0002-6641-1609>

Vladimir N. Kalinin, Assistant, Department of Telecommunications, Institute of Radioelectronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: kalinin_v@mirea.ru. Scopus Author ID 57562579000, <https://orcid.org/0000-0003-1365-4639>

Dmitry O. Zhukov, Dr. Sci. (Eng.), Professor, Department of Telecommunications, Institute of Radioelectronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: zhukov_do@mirea.ru. Scopus Author ID 57189660218, RSCI SPIN-code 1798-8891, <https://orcid.org/0000-0002-1211-5214>

Об авторах

Лесько Сергей Александрович, д.т.н., доцент, профессор кафедры «Информационно-аналитические системы кибербезопасности», Институт кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: lesko@mirea.ru. Scopus Author ID 57189664364, ResearcherID AAF-5651-2019, <https://orcid.org/0000-0002-6641-1609>

Калинин Владимир Николаевич, ассистент, кафедра телекоммуникаций, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: kalinin_v@mirea.ru. Scopus Author ID 57562579000, <https://orcid.org/0000-0003-1365-4639>

Жуков Дмитрий Олегович, д.т.н., профессор, профессор кафедры телекоммуникаций, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: zhukov_do@mirea.ru. Scopus Author ID 57189660218, SPIN-код РИНЦ 1798-8891, <https://orcid.org/0000-0002-1211-5214>

Translated from Russian into English by K. Nazarov

Edited for English language and spelling by Thomas A. Beavitt

Multiple robots (robotic centers) and systems. Remote sensing and non-destructive testing
Роботизированные комплексы и системы. Технологии дистанционного зондирования неразрушающего контроля

UDC 004.896

<https://doi.org/10.32362/2500-316X-2026-14-4-36-52>

EDN FJRJYK



RESEARCH ARTICLE

Fault diagnostics in intelligent control systems of autonomous robots: An ontological approach

Yuri A. Bykovtsev[@], Valery M. Lokhin, Sergey V. Manko, Mikhail P. Romanov

MIREA – Russian Technological University, Moscow, 119454 Russia

[@] Corresponding author, e-mail: bykovtsev@mirea.ru

• Submitted: 23.12.2025 • Revised: 23.01.2026 • Accepted: 24.05.2026

Abstract

Objectives. The work set out to develop diagnostic tools and methods for autonomous robots for the rapid detection of faults based on the analysis of cause-and-effect relationships between observed manifestations and the sources of failures.

Methods. The paper presents an analytical review of the current scientific and technical literature on the topic of the research, which includes ontology construction methods for developing a diagnostic model for autonomous robots, Boolean algebra methods for generating queries to the ontological diagnostic model, modifications to conditional and unconditional diagnostic methods for finding cause-and-effect relationships between failures and malfunctions based on the use of an ontological model, and machine experimentation methods for estimating the processing time of queries to the ontological diagnostic model.

Results. The presented ontological model and methods for conditional and unconditional diagnostics is used to establish cause-and-effect relationships between observed failures and the malfunctions that cause them.

Conclusions. Based on an ontological approach and taking into account the availability of advanced tool support, including ontology description languages, ontology modeling tools, and ontology-based inference and reasoning technologies, a new paradigm is identified for the development of a promising class of autonomous robotic systems with increased fault tolerance and, consequently, adaptability. The development of an autonomous robot presupposes the presence of an onboard intelligent control system built on a hierarchical principle that comprises executive, tactical, and strategic levels. The results of the ontological model provide an effective basis for activating and utilizing the intelligent control system's rich algorithm databases at all levels of the control hierarchy to support increased adaptability.

Keywords: autonomous robots, intelligent systems, self-diagnostics, fault tolerance, ontology

For citation: Bykovtsev Yu.A., Lokhin V.M., Manko S.V., Romanov M.P. Fault diagnostics in intelligent control systems of autonomous robots: An ontological approach. *Russian Technological Journal*. 2026;14(4):36–52. <https://doi.org/10.32362/2500-316X-2026-14-4-36-52>, <https://elibrary.ru/FJRJYK>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Диагностика неисправностей в интеллектуальных системах управления автономными роботами на основе онтологического подхода

Ю.А. Быковцев[@], В.М. Лохин,
С.В. Манько, М.П. Романов

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[@] Автор для переписки, e-mail: bykovtsev@mirea.ru

• Поступила: 23.12.2025 • Доработана: 23.01.2026 • Принята к опубликованию: 24.05.2026

Резюме

Цели. Целью работы является создание средств и методов диагностики автономных роботов для оперативного выявления неисправностей на основе анализа причинно-следственных связей между наблюдаемыми проявлениями и источниками возникновения отказов.

Методы. В ходе выполнения работы использовались методы аналитического обзора современной научно-технической литературы по тематике проводимых исследований; методы построения онтологий при разработке модели диагностики автономных роботов; методы булевой алгебры при формировании запросов к онтологической модели диагностики; модификации методов условной и безусловной диагностики в приложении к поиску причинно-следственных связей между отказами и неисправностями на основе использования онтологической модели; методы машинного эксперимента при оценке продолжительности обработки запросов к онтологической модели диагностики.

Результаты. Предложена онтологическая модель и методы условной и безусловной диагностики, позволяющие обеспечить установление причинно-следственных связей между наблюдаемыми отказами и вызывающими их неисправностями.

Выводы. Показано, что на основе онтологического подхода с учетом наличия развитой инструментальной поддержки, включая языки описания онтологий, средства их моделирования, а также технологии логического вывода и рассуждений на онтологиях, создается новая парадигма разработки перспективного класса автономных робототехнических комплексов, обладающих повышенной степенью отказоустойчивости и, соответственно, адаптивностью. Поскольку создание автономного робота предполагает наличие бортовой интеллектуальной системы управления, построенной по иерархическому принципу и содержащей исполнительный, тактический и стратегический уровни, повышение адаптивности может быть обеспечено за счет эффективного использования «богатых» баз алгоритмов интеллектуальной системы управления на всех уровнях иерархии управления, которые подключаются в зависимости от результатов работы онтологической модели.

Ключевые слова: автономные роботы, интеллектуальные системы, самодиагностика, отказоустойчивость, онтология

Для цитирования: Быковцев Ю.А., Лохин В.М., Манько С.В., Романов М.П. Диагностика неисправностей в интеллектуальных системах управления автономными роботами на основе онтологического подхода. *Russian Technological Journal*. 2026;14(4):36–52. <https://doi.org/10.32362/2500-316X-2026-14-4-36-52>, <https://elibrary.ru/FJRJYK>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

One of the top priorities for Russia's technological advancement in the 21st century involves the development of cutting-edge robotic systems. Intensive work on the development of remotely controlled, semi-automatic, and autonomous robots is being carried out by many teams at academic and industry-specific research institutes, specialized design bureaus, and the country's leading universities. Various types of developed robot models are beginning to find active application across a wide range of fields and purposes. However, an objective assessment of the operational experience with modern robotic complexes (RC) indicates the need for significant improvements in the level of their autonomy, adaptability, reliability, and fault tolerance when operating under extreme environmental conditions or in the presence of uncertainty factors and random disturbances.

One of the most important and pressing aspects of ensuring the fault tolerance of advanced autonomous robot prototypes involves the development of onboard self-diagnostic tools and methods for the rapid detection of emerging malfunctions to inform appropriate decisions regarding the full or partial restoration of lost functionality.

FEATURES OF THE DESIGN OF INTELLIGENT ON-BOARD CONTROL SYSTEMS FOR AUTONOMOUS ROBOTS

The autonomous operation of the RC is ensured by a combination of various factors, the most important of which is the presence of an onboard control system equipped with the necessary capabilities to plan behavior, appropriate actions, and movements based on the analysis of collected sensor data regarding changes in its own state and environmental conditions (Fig. 1). At the same time, the effectiveness of the full-scale implementation of the system's functional capabilities is largely determined by its adaptive properties in terms of adjusting its parameters or transforming its structure to maintain the required control quality in the face of environmental changes, disturbances, and a variety of other uncertainty factors.

Until recently, developers of robotics systems for specialized, industrial, and household applications had focused primarily on various aspects of enhancing the autonomy and adaptability of their designs through the comprehensive application of knowledge processing technologies [1, 2].

Active efforts in this field have led to significant progress on a number of challenging issues related

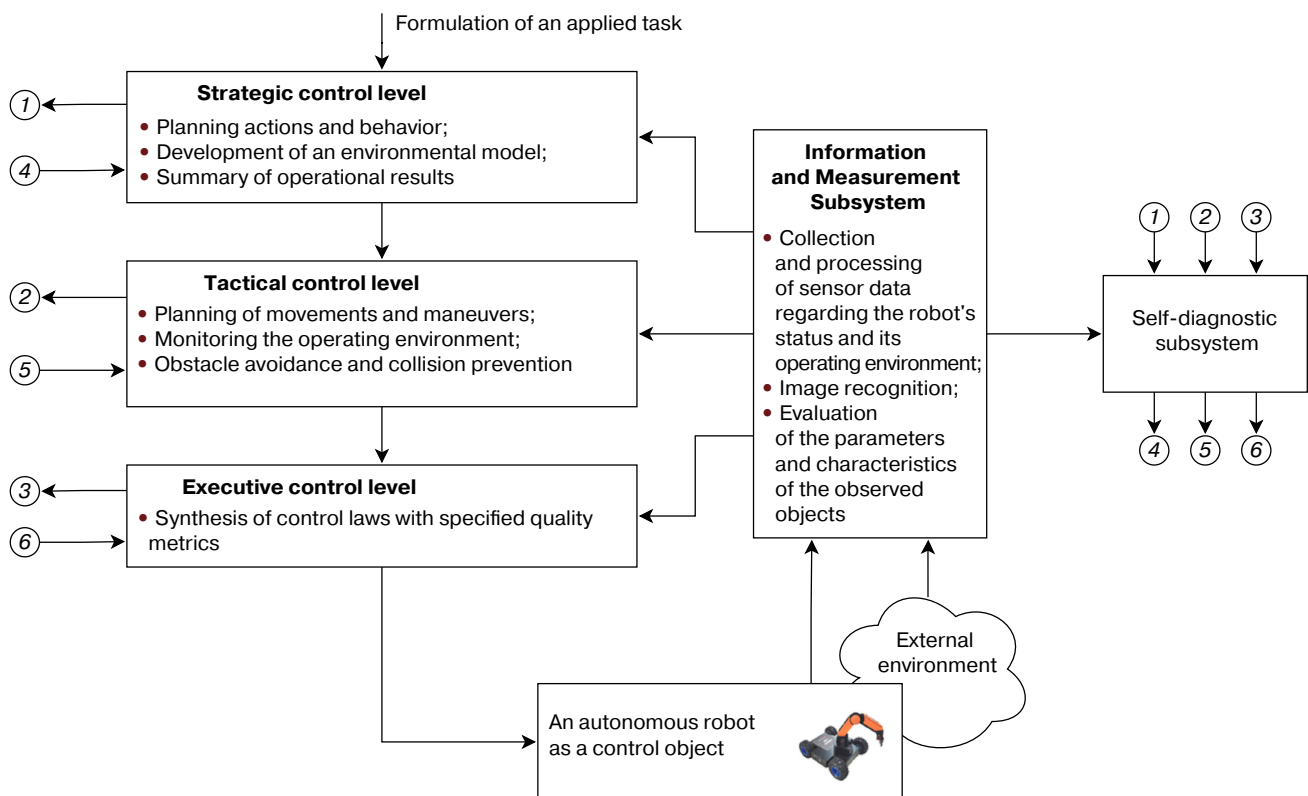


Fig. 1. Generalized structure of an autonomous robot's intelligent control system. Numbers 1, 2, and 3 denote data transmitted from the strategic, tactical, and executive levels to the diagnostic subsystem. Numbers 4, 5, and 6 denote signals from the diagnostic subsystem to the control system levels

to the intelligent control of robots operating autonomously under conditions of uncertainty. Such achievements include the development of methods for neural network recognition, intelligent planning, local navigation and mapping, deep data analysis, and self-learning.

One of the key features of the modern approach to developing intelligent control systems for autonomous robots involves the identification of subsystems at the information-measurement, behavioral, trajectory, and actuator levels, the solution of specific tasks at each of which is ensured by knowledge processing mechanisms for the selective activation of necessary algorithms depending on the nature of current situations [3, 4]. However, the range of adaptive capabilities of the control system at all levels of its hierarchy is determined by the variety of algorithms used, which are contained in the corresponding libraries, as well as the possibility of switching between them and/or effecting situational adjustment of parameters.

At the same time, it is quite clear that the reliability—and, ultimately, the effectiveness—of autonomous robots is determined not only by the sum of their intellectual and adaptive capabilities, but also by their resilience to failures inevitably caused by extreme operating conditions and harsh environmental factors.

The specific nature of the challenge of ensuring a high level of fault tolerance and survivability in advanced autonomous robotic systems stems from the need to address a set of interrelated tasks: the rapid detection and localization of failures, fault diagnosis, and the assessment of the severity of expected consequences, along with the implementation of active measures to mitigate or compensate for them.

The main challenges of the comprehensive diagnostic phase for an autonomous robot stem from two key factors:

- the wide range of possible causes that can lead to the failure of a particular functional subsystem;
- the variety of ways in which the same malfunction can manifest itself, leading to simultaneous failures in multiple subsystems.

In turn, the success of the diagnostic phase will largely determine the effectiveness of attempts to partially or fully restore the robot's operational capability through the reconfiguration of the hardware and software of the intelligent onboard control system, including changes to operating modes, procedures, and/or objectives.

Thus, the present work focuses on issues related to the development of self-diagnostic tools for autonomous robots based on an ontological model of causal relationships between observed failures and the malfunctions that cause them.

MODERN METHODS OF TECHNICAL DIAGNOSTICS

The results of research actively conducted in the field of autonomous robot diagnostics in the USA, Western European countries, and a number of other nations since the late 20th century, which can be seen as part of a broader effort to develop fault-tolerant control systems [5], are beginning to find practical application in the first experimental designs. A typical example is the autonomous unmanned RAUVER submersible vessel (Ocean Systems Laboratory, Heriot-Watt University Edinburgh, United Kingdom), which is equipped with self-diagnostic capabilities. Following the completion of appropriate diagnostic processes, on-board planners adjust the initial mission objectives as necessary [6].

The experience in carrying out search-and-rescue and other special operations using modern robotic systems (whether autonomous, semi-autonomous, or remotely controlled) clearly demonstrates that the effectiveness of their practical application is significantly limited by an insufficient level of fault tolerance [7, 8]. However, the malfunctions, failures, and breakdowns of individual functional subsystems of robot devices are highly various in nature due to being caused by random external and internal disturbances, degradation of components, elements, and materials, as well as other factors. An analysis of available statistics allows us to classify observed such failures as being caused by malfunctions in actuators, sensor subsystems, control, power, and communication systems, as well as by a number of other circumstances, including operator errors and various external influences (Fig. 2).

Actual mean time between failure (MTBF) values vary widely, ranging from a few hours to several dozen hours (depending on the type of robot, the specific tasks assigned to it, and the conditions under which they are performed) [7–10]. As an example, Table 1 presents statistical data from the U.S. Center for Robot-Assisted Search and Rescue on the assessment of the probabilities of major failures for two types of special-purpose robots over a three-year period of operation.¹

The generally accepted conceptual interpretation of fault-tolerant control systems assumes that their operational reliability must be ensured through a comprehensive set of functional capabilities designed to address the following key tasks, diagnosing emerging faults and either eliminating or compensating for them [10–12]:

¹ Carlson J. Analysis of How Mobile Robots Fail in the Field (2004). USF Tampa Graduate Theses and Dissertations. <https://digitalcommons.usf.edu/etd/980>. Accessed April 30, 2026.

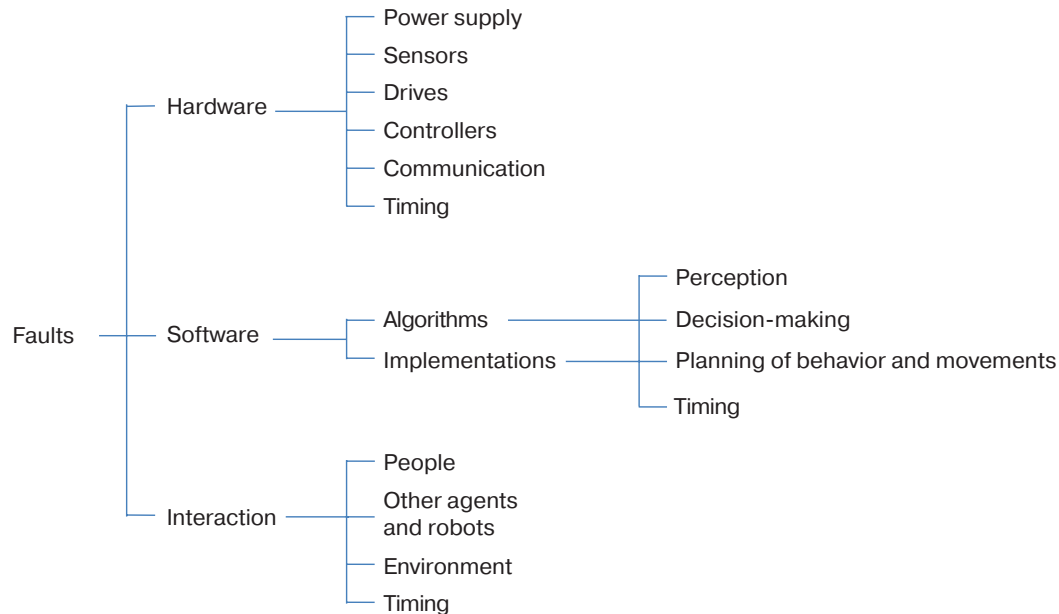


Fig. 2. Classification of faults in robotic systems

Table 1. Probability estimates for major failures in special-purpose mobile robots

Manufacturer	Actuators	Control system	Power supply system	Communication system	Sensor subsystem
Inuktun (Canada)	0.45	0.32	0.00	0.02	0.12
Irobot (USA)	0.22	0.30	0.15	0.22	0.11

- failure detection (i.e., determining that the system is behaving abnormally);
- fault localization (with the precision to identify the location and type of the failure that occurred);
- fault identification (i.e., determining the causes of a failure based on an analysis of its observable symptoms);
- assessing the severity of the identified faults and deciding whether to fully or partially restore functionality.

In the context of developing tools and methods for the automatic diagnosis of faults and the restoration of operational capability in autonomous robots, it is also necessary to distinguish between the following critically important types of failures:

- total and partial, the former of which result in the loss of an object's functionality, while the latter result in its limitation;
- defects that can be repaired (until the object regains its original functionality) and defects that can be compensated for (enabling lost functions to be restored to a limited extent).

In turn, the severity of the faults identified during the diagnostic process is assessed on a four-category scale:

- low level—at which the robot, by adjusting certain operational parameters, can continue to carry out its assigned mission at the same level of autonomy;
- average level—at which the robot can continue to perform a shortened version of the specified target mission while maintaining the same level of operational autonomy;
- significant level—at which the robot cannot continue to perform its assigned mission while maintaining the same level of operational autonomy;
- fatal level—at which the robot cannot continue to perform its assigned mission.

As a review of the literature [10–21] shows, three fundamentally different approaches serve as the conceptual foundation for the development of methods for fault detection and diagnosis; these approaches are associated, respectively, with the use of analytical models [11, 19], data analysis [13, 17], and knowledge processing [6, 16, 18]. The corresponding classification of diagnostic methods is shown in Fig. 3.

Model-based diagnostic methods involve identifying faults based on deviations between the actual observed behavior of a system and its expected behavior. The main drawback of this approach stems from the difficulty of

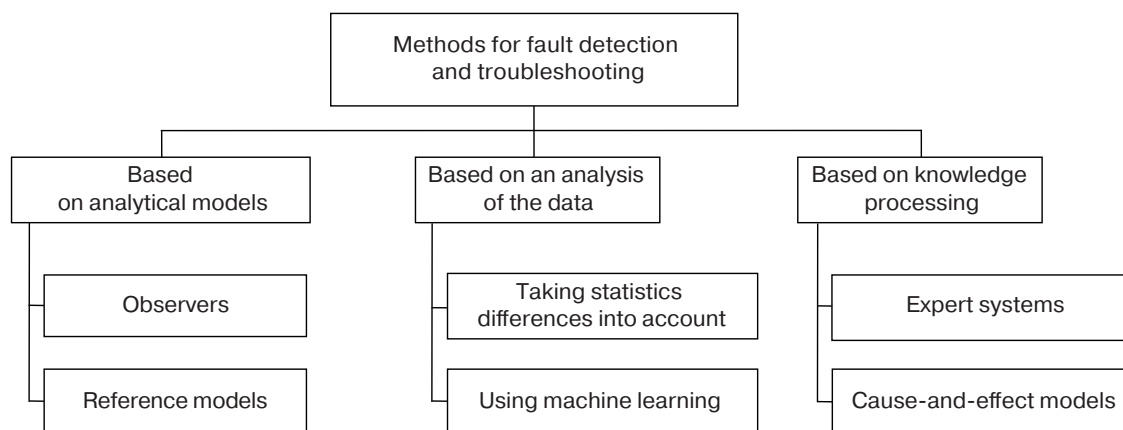


Fig. 3. Classification of methods for fault detection and diagnosis

creating adequate mathematical models of the diagnostic objects [15]. Nevertheless, model-based diagnostic methods are widely used to detect faults that arise during the operation of RC [11, 14, 18, 19, 22–27]. Within this class, methods using a reference model and an observer, which are implemented using various modifications of the Kalman filter, have become the most widespread due to their effectiveness.

A fundamentally different approach based on the analysis of data streams and the rapid detection of anomalies in the changes of monitored signals is actively used for the real-time diagnosis of robots, as well as other types of complex dynamic systems, without requiring the prior creation of mathematical models of system behavior [20, 21]. In this case, deviations of current data from statistically expected values are considered an indicator of a malfunction. To extract the necessary features that enable the detection of faults during the processing of the incoming data stream, various methods of signal analysis are applied in the time, frequency, or time-frequency domains, as well as those based on machine learning [20, 28].

For example, analyzing signals in the time domain yields a range of statistically significant characteristics, including the mean and root-mean-square values, the area under the curve, the slope and shape factor, the variance, entropy, and so on.

In turn, the analysis of signals in the frequency domain, performed using Fourier transforms, also employs relevant statistical characteristics [29, 30], the main ones being the center frequency, root mean square frequency, root of the variance frequency, and a number of others [31].

Time-frequency analysis, performed using the wavelet transform or the short-time Fourier transform, provides a detailed picture of the signal's characteristics in both the time and frequency domains [32, 33].

A set of features against which current data can be compared to detect failures or diagnose malfunctions

can be generated through machine learning using neural networks or Bayesian networks, cluster and regression analysis, or a number of other alternative mechanisms, using examples from a pre-labeled sample or operational observations from a “sliding window” [20, 34].

The main drawback of data-driven diagnostic methods is the challenge of processing large volumes of information in real time.

Knowledge-based intelligent diagnostics assumes that fault detection can be performed either at the level of expert representations, which encapsulate the experience of highly qualified specialists, or using causal models that describe the diagnostic object in terms of its structure² and the characteristics of the applied tasks being performed, as well as the manifestations of possible failures and their probable causes [6, 16, 18, 35–44].

The main drawback of intelligent diagnostic methods lies in their ability to detect and identify only a limited set of faults from among the types predefined in the knowledge base.

Various forms of knowledge representation are used to construct intelligent diagnostic models, including first-order predicates, production rules, semantic networks, frame-based structures, ontologies, etc. [16, 35, 36, 41, 45–48]. The chosen form of knowledge representation determines the set of tools and methods required for their logical processing during intelligent diagnostics. As presented in Table 2, existing approaches to fault diagnosis are often complementary in terms of their capabilities [16, 20].

² GOST R 27.302-2009. National Standard of the Russian Federation. *Reliability in Engineering: Fault Tree Analysis* (approved and put into effect by Order No. 1249-st of the Federal Agency for Technical Regulation and Metrology dated December 15, 2009). Moscow: Standartinform; 2011, 22 p. (In Russ.).

Table 2. Key features of modern approaches to fault diagnosis

Approaches	Benefits	Problems of practical application
Based on models	• Real-time processing; • The ability to detect unknown faults	• Modeling of complex systems; • Modeling of context
Based on data analysis	• Does not require a model; • The ability to detect unknown faults	• Real-time processing
Based on knowledge	• Real-time processing; • Low computational load	• Detecting of the unknown faults

For example, model-based diagnostic methods are used to detect faults in actuators, drives, internal sensors, and other system components, whose behavior can be described in terms of analytical models.

In turn, data-driven diagnostic methods are widely used in the processing of external sensor readings to detect and locate failures caused by both internal system malfunctions and environmental factors.

An effective approach to diagnosing faults in external sensors within multisensor systems involves the use of various data fusion methods, including those based on the Kalman or Dempster–Shafer filters.

The most significant advantage of a knowledge-based approach to fault detection lies in its ability to integrate data-driven and model-based diagnostic approaches within a hybrid monitoring system to identify causal relationships between observed failure symptoms and specific malfunctions or emergency situations.

Diagnosis and troubleshooting of malfunctions related to various types of software and hardware failures within the onboard computing system are carried out using established and proven methods in accordance with the relevant plan.

A summary of the results of the review leads to a number of key conclusions:

- The comprehensive use of modern technical diagnostic methods ensures the prompt detection of malfunctions that occur during the operation of an autonomous robot.
- In order to reliably identify the types of malfunctions, the vast majority of which negatively impact the performance of multiple robot subsystems simultaneously, specialized diagnostic models are required for establishing causal relationships between observed failure symptoms and their root causes.
- The use of an ontological approach enables the creation of a self-diagnostic model for an autonomous robot that captures the full set of cause-and-effect

relationships underlying its potential failures and specific malfunctions, taking into account the hierarchical structure of intelligent onboard control systems.

Depending on the severity of the detected faults, system recovery can be achieved in several different ways, including adjusting control algorithm parameters, reconfiguring hardware and software, dynamically modifying task execution plans, or adjusting operational objectives [10].

The development of tools and methods for self-diagnosis and recovery of intelligent onboard control systems opens up real prospects for creating a new generation of robotic systems characterized by a high degree of fault tolerance, autonomy, and operational reliability [20, 23].

ONTOLOGICAL MODEL FOR INTELLIGENT DIAGNOSTICS OF AN AUTONOMOUS ROBOT

The most promising approach to creating fault-tolerant control systems for autonomous robots with built-in onboard self-diagnostic capabilities involves the use of ontological knowledge representations [45, 46, 48]. Here, an ontology is understood as a formal description of a conceptual model of a domain within a specific context [49, 50].

It is important to note that the main advantage of an ontological representation of knowledge lies in its applicability to solving various functional and practical problems, both at the level of an individual robot and at the level of a robotic group.

Formally, an ontology is defined by the following tuple:

$$O = \langle X, R, F \rangle,$$

where X is a finite set of concepts in the domain; R is a finite set of relations between concepts; F is a finite set of interpretation functions defined on concepts and/or relations.

In turn, the ontological model, which allows us to describe the relationships between ontologies at different levels, is also defined as a tuple:

$$Z = \langle O, P, M \rangle,$$

where O is a top-level ontology (meta-ontology) that contains general concepts and relationships independent of the domain; P is a set of subject-area ontologies and task ontologies (taking user preferences into account); M is the inference engine model of the ontological system (for example, to modify search relevance criteria or repository formation criteria).

The use of ontologies enables a wide range of functional capabilities, the main ones being:

- formalized knowledge representation, including the organization of general and domain-specific terminology, the description of domain areas for use by various intelligent agents, and the indexing, filtering, and classification of formalized knowledge;
- acquiring and organizing knowledge;
- semantic integration of information resources (creation of a unified online knowledge space);
- enabling users to find the information they need (indexing, filtering, and classifying formalized knowledge).

The creation and processing of ontological knowledge representation models is supported by a well-developed set of tools, including ontology description languages (SHOE³, XOL⁴, OIL⁵, RDFS⁶, DAML⁷+OIL, and OWL⁸) [51], ontology modeling tools (*Apollo*⁹, *WebOnto*¹⁰, *WebODE*¹¹, *OntoEdit*¹², and *Protégé*¹³) [52], as well as technological procedures for logical inference and reasoning on

ontologies (*Racer*¹⁴, *Pellet*¹⁵, *FaCT++*¹⁶, *FOWL*¹⁷, *KAON2*¹⁸, *Hoolet*¹⁹, *Jess*²⁰) [50, 53–58].

The effectiveness of ontological knowledge representation models using in fault detection and identification tasks is demonstrated by examples of real-world technical diagnostic systems for complex electromechanical equipment, CNC machine tools, forklifts, autonomous robots, and other types of dynamic objects [36, 48, 50, 59, 60].

Ontological models are also becoming the foundational framework for the emergence of a new paradigm in the development of the next generation of autonomous mobile robots [45, 46, 61–63]. A striking illustration of this trend is the 2015 publication by the IEEE Robotics and Automation Society Ontologies for Robotics and Automation Working Group (IEEE Robotics and Automation Society Ontologies for Robotics and Automation Working Group, IEEE ORA WG) of the first-ever special standard, IEEE 1872-2015²¹, which defines a set of ontologies related to robotics and automation (R&A) [64]. From the perspective of current trends, this concept aligns with the idea of integrating intelligent control and diagnostic systems for autonomous robots.

The development of any ontology, including an ontological model for the diagnostics of an autonomous robot, is an iterative process that involves the following main stages [50], whose sequence is given in Fig. 4 in the form of a flow diagram:

- identifying the entities involved in the model (decomposition);
- creation of individual entities (identification);
- creation of classes that correspond to groups of entities and assigning entities to classes (classification);
- determining ways to express information about the characteristics and relationships of entities (properties description);
- assigning values to properties, creating relationships (values, relationships).

This approach allows the developer to focus on identifying the core concepts and objects to be incorporated into the model, and only then to define the general properties or characteristics according to which these objects can be grouped into classes to organize

³ Simple HTML ontology extensions is the ontology description language based on the HTML data markup language (hypertext markup language).

⁴ XML-based ontology exchange language is the ontology description language designed for exchanging ontology descriptions (XML—eXtensible Markup Language).

⁵ Ontology interchange language.

⁶ RDF Schema is the ontology description language, a semantic extension of the RDF language (resource description framework).

⁷ Digital asset modeling language is the open-source smart contract language.

⁸ Web ontology language is the language for describing ontologies for the semantic web.

⁹ <https://www.apollographql.com/>. Accessed April 30, 2026.

¹⁰ <http://kmi.open.ac.uk/projects/webonto/>. Accessed April 30, 2026.

¹¹ <https://oeg.fi.upm.es/index.php/en/old-technologies/60-webode/index.html>. Accessed April 30, 2026.

¹² <https://www.softpedia.com/get/Science-CAD/OntoStudio.shtml>. Accessed April 30, 2026.

¹³ <https://protege.stanford.edu/>. Accessed April 30, 2026.

¹⁴ <https://www.chai.uni-hamburg.de/~moeller/racer/>. Accessed April 30, 2026.

¹⁵ <http://pellet.owldl.com/>. Accessed April 30, 2026.

¹⁶ <http://owl.cs.manchester.ac.uk/tools/fact/>. Accessed April 30, 2026.

¹⁷ <https://fowl.sourceforge.net/>. Accessed April 30, 2026.

¹⁸ <https://km.aifb.kit.edu/projects/owltools/>. Accessed April 30, 2026.

¹⁹ <http://owl.man.ac.uk/hoolet/>. Accessed April 30, 2026.

²⁰ <https://www.ida.liu.se/~hener77/JessTab/>. Accessed April 30, 2026.

²¹ <https://standards.ieee.org/ieee/1872/5354/>. Accessed April 30, 2026.

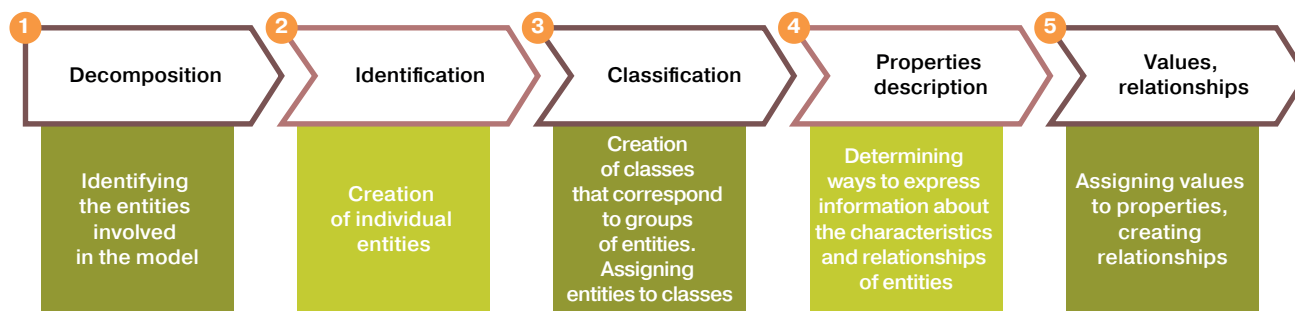


Fig. 4. Stages in the development of a diagnostic model ontology

causal, logical, quantitative, and other relationships between individual classes. At the same time, the set of basic concepts will depend directly not only on the subject area or the specific object under study, but also on the task that the ontological model is designed to solve.

One example is the development of an ontological model for diagnosing the mobile platform of an autonomous RC, which includes the following main components:

- actuators that integrate a motor, feedback sensors, and a dedicated drive control system (CS) controller;
- a control system for the platform's executive level, comprising a set of actuators, a machine vision system (MVS), and a global navigation satellite system (GNSS), as well as a central computer that generates control signals (to the actuator control system);
- power supply system with a controller;
- communication module for the tactical level of the RC control system.

In this case, the primary entities of the ontological model being created (Fig. 5) should not be the individual components of the mobile platform, but rather their potential failures and malfunctions..

For example, in a drive control system, situations such as a complete motor shutdown, a loss of power to the drive, or CS sensor failures can be classified under the separate category of "*Drive Control System Failure*."

By applying similar reasoning to each specific component or subsystem of the mobile platform, it is possible to identify the following set of classes:

- executive-level failure;
- executive level sensor error;
- drive control system failure;
- drive control unit error;
- no power supply;
- no orders from the tactical level;
- a mechanical fault;

- failure caused by interaction with the medium;
- overload.

For the identified classes of potential malfunctions in specific platform subsystems that could lead to errors in its movement, a more general class, "*Error Causes*," is additionally introduced. At the same time, the errors themselves, represented by the subclasses "*Movement Blocked*" and "*Incorrect Movement*," are grouped into the "*Errors*" class. The relationships between faults and their manifestations at the level of observable failures are defined using the pointer properties "*entails*" and "*is caused by*," which will determine the corresponding relationships between the "*Errors*" and "*Error Causes*" classes.

In this context, as properties-values characterizing the possible states of the defined entities (true or false (false)) within the ontological model, two parameters are considered: "*Error State*" and "*Suspected Error*." The former is defined for objects in both the "*Errors*" class and the "*Error Causes*" class, while the latter applies only to objects in the latter class.

It is clear that the ontological model (Fig. 5) can be expanded to account for a broader range of malfunctions and failures, which occur not only at the level of the mobile platform, but also within the RC system as a whole.

ALGORITHMS FOR UNCONDITIONAL AND CONDITIONAL FAULT DETECTION

Fault diagnosis and analysis of their potential consequences based on observed failures using an ontological model requires the formulation of queries.

In accordance with the unconditional diagnosis algorithm, when the property-value or property-pointer of an individual ontology object changes, the logical inference mechanism processes each query from its knowledge base independently of the results of processing previous queries.

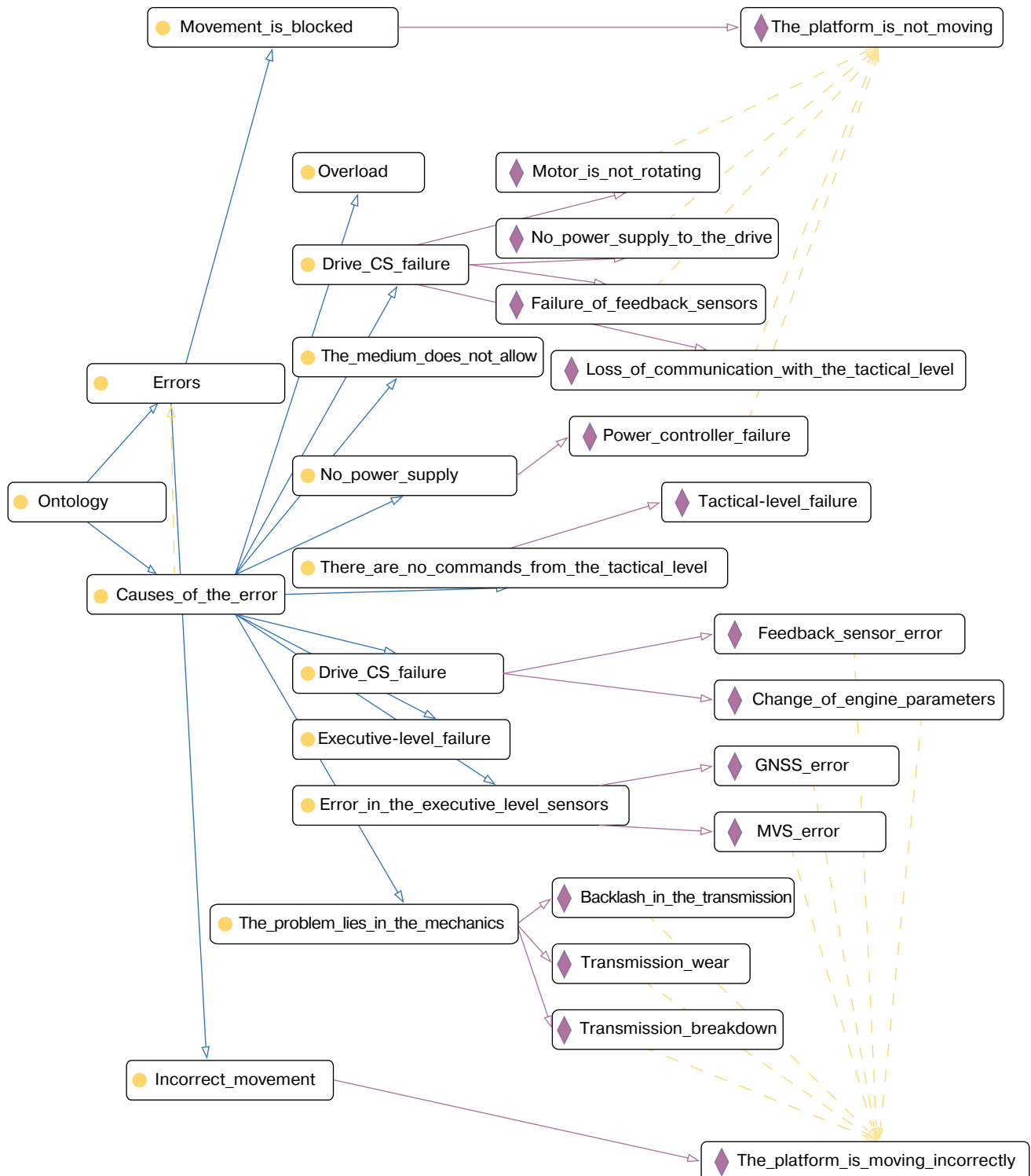


Fig. 5. Ontological model for diagnosing a mobile autonomous RC platform

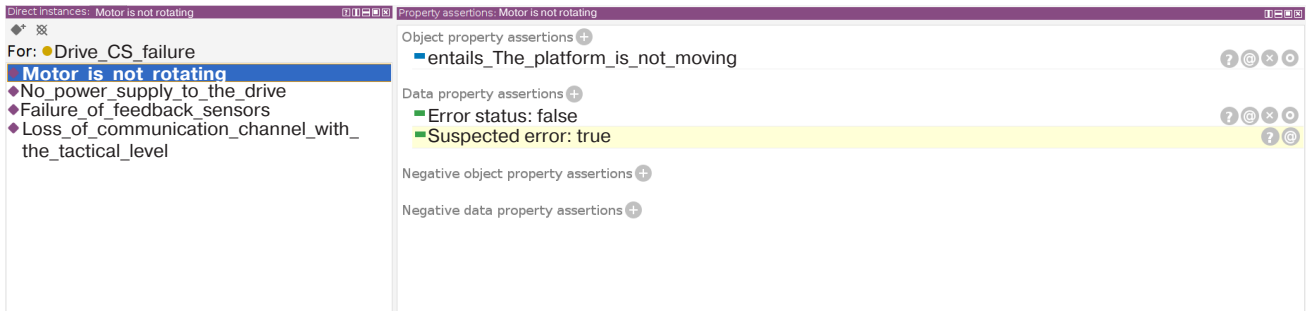


Fig. 6. Results of a simulation experiment on processing an “upstream” query regarding the potential consequences of a malfunction

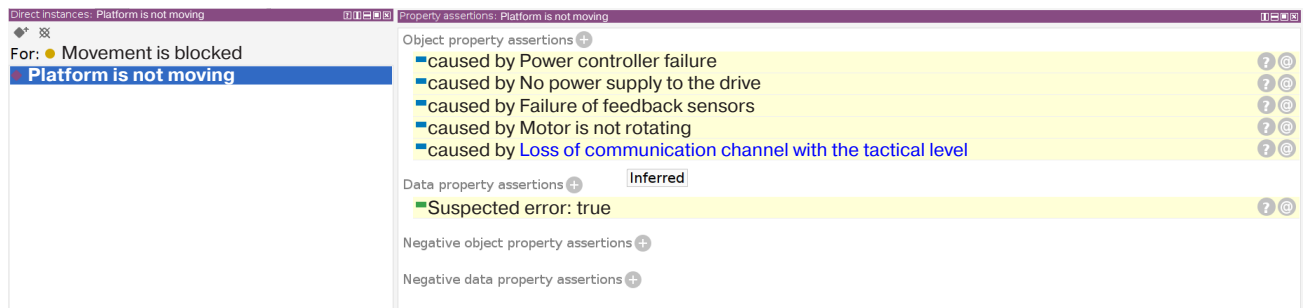


Fig. 7. Results of a simulation experiment on processing a “downstream” query regarding the possible causes of the observed failure

The general structure of queries formulated in SWRL²² (as the most feature-rich and widely used language) is as follows:

$$\begin{aligned} & \text{Class_1}(x) \wedge \text{property-value_1}(x, \text{value_1}) \wedge \\ & \wedge \text{property-pointer}(x, y) \wedge \text{Class_2}(y) \rightarrow \\ & \rightarrow \text{property-value_2}(y, \text{value_2}), \end{aligned}$$

where x and y are the objects of classes 1 and 2, respectively; property-value_1 and property-value_2 are the values (numeric, string, Boolean, etc.) of certain parameters associated with the objects of classes 1 and 2, respectively; property-pointer is the relationship linking the objects of classes 1 and 2.

This query structure enables drawing conclusions about changes in the state of an object of one class based on information about the state of objects of another class.

Specifically, an “upstream” query type, which enables tracing the manifestation of a failure at the platform level back to the occurrence of an error in a specific component, will take the form of:

$$\begin{aligned} & \text{Drive_CS_failure}(?c) \wedge \text{Error_status}(?c, \text{true}) \wedge \\ & \wedge \text{entails}(?c, ?y) \rightarrow \text{Error_status}(?y, \text{true}). \end{aligned}$$

The processed query is handled by a logical inference procedure that sequentially executes a series of relevant steps, including:

- selecting those components (c) from the “*Drive Control System Failure*” class for which the “*Error Status*” is true;
- find all objects (y) associated with the selected property-indicator “*entails*” and set the property-value “*Error Status*” for the objects y .

Thus, if for one or more objects belonging to the “*Drive Control System Failure*” class (e.g., “*Motor Not Rotating*”), the “*Error Status*” property-value takes the value “true” (true), then as a result of the request being processed by the logical inference mechanism, the property-value “*Error Status*” for the object “*Platform Not Moving*” is also assigned the value “true” (Fig. 6).

A “downstream” type of query, which enables fault diagnosis based on its manifestation at the platform failure level, is generated in the same way:

$$\begin{aligned} & \text{Movement_blocked}(?c) \wedge \text{Error_state}(?c, \text{true}) \wedge \\ & \wedge \text{caused by}(?c, ?y) \rightarrow \text{Suspected_error}(?y, \text{true}). \end{aligned}$$

The results of a simulation experiment on processing a “downstream” query regarding the possible causes of the observed failure are shown in Fig. 7.

While the absence of conditions for processing queries significantly simplifies the analysis of the results produced by the ontological diagnostic model, the increased size of both the ontology itself and the query database can lead to a critical increase in the diagnostic system’s response time when a fault occurs. In this regard, the implementation of conditional fault search appears

²² Semantic web rule language is the rule description language for the publicly accessible global Semantic web.

to be highly promising, since enabling the selection of queries to be processed from the knowledge base at the current step based on the results of the previous step, thereby reducing the overall duration of the diagnostic process.

A distinctive feature of the proposed algorithm is that it processes only those queries from the knowledge base that can lead to a change in the state of the ontological model. Here, the state of the model is understood as the set of property-value pairs of the ontology's entities, while queries are selected based on an analysis of their antecedent part.

Let $\mathbf{X}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,n})$ be the status vector of the ontological model at the current step (here, a "state" refers to the properties and values of objects); $\mathbf{X}_{i-1} = (x_{i-1,1}, x_{i-1,2}, \dots, x_{i-1,n})$ is the status vector at the previous step; $\mathbf{X} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$ is the new status associated with the occurrence of a fault; $\hat{\mathbf{X}} = (\mathbf{X}_i - \mathbf{X}_{i-1})$ is the status change vector.

In this case, the formal description of the conditional fault-finding algorithm will take the following form:

1. Accept $\mathbf{X}_i = \tilde{\mathbf{X}}$;
2. Calculate $\hat{\mathbf{X}}$;
3. If $\hat{\mathbf{X}} \neq 0$, then for each nonzero component of the vector, find a query in the knowledge base that contains that component in its antecedent. Otherwise, the algorithm terminates;
4. Accept $\mathbf{X}_{i-1} = \mathbf{X}_i$;
5. Generate a new vector $\mathbf{X}_i$ based on the results of processing the selected queries using a logical inference mechanism;
6. Go to step 2.

It is clear that the potential applicability of the diagnostic algorithms under development is largely determined by their computational complexity and average time complexity, which are primarily associated with processing queries to the ontological model, and depend not only on its structure and dimensionality, but also on the characteristics of the logical inference mechanism.

It should be noted that one of the most effective tools for constructing semantic reasoning is the Pellet logical inference engine, which, unlike many well-known alternatives (*Hermit*²³, *RacerPro*²⁴, *FaCT++*, etc.), supports working with ontologies not only at the level of classes and property-indicators, but also at the level of individual objects. This is of fundamental importance for solving diagnostic problems using unconditional and conditional search algorithms.

Estimates of the computational complexity of the Pellet logical inference mechanism as presented

in a number of studies [65, 66] follow a quadratic relationship ($O(n^2 + m^2)$) with respect to the number of classes n and relations m , and a linear relationship $O(k)$ with respect to the number of queries k . Based on these data, it can be argued that the expansion of the diagnostic model (if necessary) should be carried out by increasing the number of queries rather than by refining the structure of the ontology itself, which is oriented toward finding causal relationships between observed failures and the malfunctions causing them. At the same time, monitoring the correct operation of individual robot components, nodes, and subsystems should be carried out using local diagnostic tools and methods based on models and data analysis. A series of experiments conducted on a personal computer with a quad-core i5-7200U central processing unit operating at a clock speed of 2.50 GHz showed that the maximum processing time for a query to the presented ontological diagnostic model is 61 ms, which is a perfectly acceptable figure in the context of fault-tolerant control of autonomous robots.

CONCLUSIONS

Ensuring fault tolerance, representing one of the most critical requirements for next-generation autonomous robots, requires the development of mechanisms for rapid fault detection and the reconfiguration of the onboard control system to mitigate the consequences of such faults. Despite the wide variety of modern technical diagnostic methods, a comprehensive check of a robot's operational capability remains a highly challenging task, whose complexity stems from the need to identify the true causes of the observed malfunctions in the behavior of its functional subsystems.

The ontological model and self-diagnostic algorithms for an autonomous robot proposed in this article is designed to identify specific malfunctions based on an analysis of the cause-and-effect relationships between a set of actual symptoms and the sources of potential failures.

Unlike existing approaches based on analytical models and data analysis, ontological modeling technology can serve as a constructive basis for building a comprehensive self-diagnostic system that reflects the entire set of cause-and-effect relationships among possible failures and specific malfunctions of an autonomous robot, taking into account the hierarchical structure of the intelligent onboard control system and the specific tasks of each of its levels.

Experimental studies assessing the processing time for queries to the diagnostic ontology model have confirmed that the processing time for ten queries is approximately 60 ms, which is quite acceptable given the inertial parameters of mobile robots.

²³ <http://www.hermit-reasoner.com/>. Accessed April 30, 2026.

²⁴ https://franz.com/agraph/racer/racer_features.lhtml. Accessed April 30, 2026.

The use of an ontological model also opens up long-term prospects for improving the survivability and capability of autonomous robots operating under extreme conditions by reconfiguring the hardware and software components of the intelligent control system based on the results of self-diagnostic tests.

Authors' contributions

Yu.A. Bykovtsev—development of an example ontology model and simulation.

V.M. Lokhin—general guidance and review.

S.V. Manko—literature review and analysis.

M.P. Romanov—problem statement.

REFERENCES

1. Jang D.S., Cho D.H., Lee W.C., Ryu S.K., Jeong B., Hong M., Jung M., Kim M., Lee M., Lee S., Choi H.L. Unlocking Robotic Autonomy: A Survey on the Applications of Foundation Models. *International Journal of Control, Automation, and Systems (IJCAS)*. 2024;22(8):2341–2384. <https://doi.org/10.1007/s12555-024-0438-7>
2. Zghair N.A., Al-Araji A.S. A One Decade Survey of Autonomous Mobile Robot Systems. *International Journal of Electrical and Computer Engineering (IJECE)*. 2021;11(6):4891–4906. <http://doi.org/10.11591/ijece.v11i6.pp4891-4906>
3. Makarov I.M., Lokhin V.M. (Eds.). *Intellektual'nye sistemy avtomaticheskogo upravleniya (Intelligent Automatic Control Systems)*. Moscow: Fizmatlit; 2001, 576 p. (In Russ.). ISBN 5-9221-0162-5
4. Quan N.D., Ha H.H., Son L.D., An T.V., Hung N.T., Linh C.V. Control system architecture of an intelligent humanoid robot. *Vietnam Journal of Science and Technology*. 2022;60(6):1162–1178.
5. Zhang Y., Jiang J. Bibliographical Review on Reconfigurable Fault-Tolerant Control System. *Ann. Rev. Contr.* 2008;32(2): 229–252. <https://doi.org/10.1016/j.arcontrol.2008.03.008>
6. Hamilton K., Lane D., Taylor N., Brown K. Fault diagnosis on autonomous robotic vehicles with RECOVERY: an integrated heterogeneous-knowledge approach. In: *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation*. IEEE. 2001. V. 4. P. 3232–3237. <https://doi.org/10.1109/ROBOT.2001.933116>
7. Carlson J., Murphy R.R. How UGVs physically fail in the field. *IEEE Trans. Robotics*. 2005;21(3):423–437. <https://doi.org/10.1109/TRO.2004.838027>
8. Carlson J., Murphy R. Reliability analysis of mobile robots. In: *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*. IEEE. 2003. V. 1. P. 274–281. <https://doi.org/10.1109/ROBOT.2003.1241608>
9. Khalastchi E., Kaminka G.A., Kalech M., Lin R. Online anomaly detection in unmanned vehicles. In: *10th International Conference on Autonomous Agents and Multiagent Systems*. 2011. V. 1–3. P. 105–112.
10. Crestani D., Godary-Dejean K., Lapierre L. Enhancing Fault Tolerance of Autonomous Mobile Robots. *Robot. Autonom. Syst.* 2015;68:140–155. <https://doi.org/10.1016/j.robot.2014.12.015>
11. Hasan A., Tahavori M., Midtiby H. Model-Based Fault Diagnosis Algorithms for Robotic Systems. *IEEE Access*. 2023;11:2250–2258. <https://doi.org/10.1109/ACCESS.2022.3233672>
12. Abdilatif Alobaidy M., Alobaidy A., Abdul-Jabbar J., Saeed S. Faults Diagnosis in Robot Systems: A Review. *Al-Rafidain Engineering Journal (AREJ)*. 2020;25:164–175.
13. Golombek R., Wrede S., Hanheide M., Heckmann M. Online data-driven fault detection for robotic systems. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2011. P. 3011–3016. <https://doi.org/10.1109/IROS.2011.6095034>
14. Isermann R. Model-based fault-detection and diagnosis—Status and applications. *Ann. Rev. Contr.* 2005;29(1):71–85. <https://doi.org/10.1016/j.arcontrol.2004.12.002>
15. Chandola V., Banerjee A., Kumar V. Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*. 2009;41(3):Article 15.
16. Xu S. A Survey of Knowledge-Based Intelligent Fault Diagnosis Techniques. *J. Phys.: Conf. Ser.* 2019;1187:032006. <https://doi.org/10.1088/1742-6596/1187/3/032006>
17. Khalastchi E., Kalech M., Rokach L. Sensor fault detection and diagnosis for autonomous systems. In: *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-Agent Systems*. 2013. P. 15–22. <https://doi.org/10.65109/RKJM1183>

18. Xu F., Liu X., Chen W., Zhou C., Cao B. Ontology-Based Method for Fault Diagnosis of Loaders. *Sensors*. 2018;18(3):729. <https://doi.org/10.3390/s18030729>
19. Alikhani A., Sharifi G. Application of a Model-Based Fault Detection Approach on a Spacecraft. *International Journal of Reliability, Risk and Safety: Theory and Application (IJRRS)*. 2020;3(2):19–26.
20. Khalastchi E., Kalech M. On Fault Detection and Diagnosis in Robotic Systems. *ACM Computing Surveys (CSUR)*. 2018;51(1):1–24. <https://doi.org/10.1145/3146389>
21. Quamar M., Nasir A. Review on Fault Diagnosis and Fault-Tolerant Control Scheme for Robotic Manipulators: Recent Advances in AI, Machine Learning, and Digital Twin. *arXiv preprint*. arXiv:2402.02980. 2024. <https://doi.org/10.48550/arXiv.2402.02980>
22. Brandstötter M., Hofbaur M., Steinbauer-Wagner G., Wotawa F. Model-based fault diagnosis and reconfiguration of robot drives. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2007. P. 1203–1209. <https://doi.org/10.1109/IROS.2007.4399092>
23. Pettersson O. Execution monitoring in robotics: A survey. *Robot. Auton. Syst.* 2005;53(2):73–88. <https://doi.org/10.1016/j.robot.2005.09.004>
24. Aguiar A.P. Multiple-model adaptive estimators: Open problems and future directions. In: *Proceedings of ECC'07 – European Control Conference*. IEEE. 2007. P. 5544–5545. <https://doi.org/10.23919/ECC.2007.7069028>
25. Lao L., Ellis M., Christofides P.D. Proactive fault-tolerant model predictive control. *AIChE J.* 2013;59(8):2810–2820. <https://doi.org/10.1002/aic.14074>
26. Yao J., Wang D., Cai D., Xu Y., Zhao Y. Fault-tolerant strategy and experimental study on compliance assembly of a redundant parallel six-component force sensor. *Sens. Actuators A: Phys.* 2018;272:114–124. <https://doi.org/10.1016/j.sna.2017.12.036>
27. Du Pu, Venkidasalapathy J.A., Venkateswaran S., Wilhite B., Kravaris C. Model-Based Fault Diagnosis and Fault Tolerant Control for Safety-Critical Chemical Reactors: A Case Study of an Exothermic Continuous Stirred-Tank Reactor. *Ind. Eng. Chem. Res.* 2023;62(34):13554–13571. <https://doi.org/10.1021/acs.iecr.3c01205>
28. Mogal S., Lalwani D. A brief review on fault diagnosis of rotating machineries. *Appl. Mech. Mater.* 2014;541-542:635–640. <https://doi.org/10.4028/www.scientific.net/AMM.541-542.635>
29. Li S., Xin Y., Li X., Wang J., Xu K. A review on the signal processing methods of rotating machinery fault diagnosis. In: *Proceedings IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*. IEEE. 2019. P. 1559–1565. <https://doi.org/10.1109/ITAIC.2019.8785572>
30. Niu G., Dong X., Chen Y. Motor Fault Diagnostics Based on Current Signatures: A Review. *IEEE Trans. Instrum. Meas.* 2023;72:Article 3520919.
31. Niu G. Data Acquisition and Preprocessing. In: *Data-Driven Technology for Engineering Systems Health Management: Design Approach, Feature Construction, Fault Diagnosis, Prognosis, Fusion and Decisions*. Singapore: Springer; 2017. P. 49–99. https://doi.org/10.1007/978-981-10-2032-2_4
32. Islam A., Amin R., Sharan M.H., Ushno J.A., Islam K.M.T., Hossen D., Hossain M.S., Alam M.S. Designing an autonomous framework to detect and classify the fabric defects using wavelet transform and neural network. *AIP Conf. Proc.* 2023;2788:050003. <https://doi.org/10.1063/5.0148618>
33. Bai Y., Cheng W., Wen W., Liu Y. Application of time-frequency analysis in rotating machinery fault diagnosis. *Shock and Vibration*. 2023;2023:Article 9878228. <https://doi.org/10.1155/2023/9878228>
34. Surucu O., Gadsden S., Yawney J. Condition Monitoring using Machine Learning: A Review of Theory, Applications, and Recent Advances. *Expert Syst. Applicat.* 2023;221:19738. <https://doi.org/10.1016/j.eswa.2023.119738>
35. Manzoor S., Rocha Y.G., Joo S.H., Bae S.H., Kim E.J., Joo K.J., Kuc T.Y. Ontology-Based Knowledge Representation in Robotic Systems: A Survey Oriented toward Applications. *Appl. Sci.* 2021;11(10):4324. <https://doi.org/10.3390/AP11104324>
36. Diab M., Pomarlan M., Borgo S., et al. FailRecOnt – An ontology-based framework for failure interpretation and recovery in planning and execution. In: *Proceedings of the 2nd International Workshop on Ontologies for Autonomous Robotics*. 2021. P. 1–14. Available from URL: <https://hdl.handle.net/2117/357471>. Accessed July 19, 2025.
37. Yan B., Zhang T., Xie C. Fuzzy expert system for fault diagnosis of robotic assembly. In: *Proceedings of the World Congress on Intelligent Control and Automation*. IEEE. 2002. P. 445–449. <https://doi.org/10.1109/WCICA.2002.1022148>
38. De la Sen M., Minambres J.J., Garrido A.J., Almansa A., Soto J.S., Basic theoretical results for expert systems: application to the supervision of adaptation transients in planar robots. *Artif. Intell.* 2004;152(2):173–211. [https://doi.org/10.1016/S0004-3702\(03\)00136-X](https://doi.org/10.1016/S0004-3702(03)00136-X)
39. Lamine K.B., Kabanza F. History checking of temporal fuzzy logic formulas for monitoring behavior-based mobile robots. In: *Proceedings of the 12th IEEE International Conference on Tools with Artificial Intelligence*. IEEE. 2000. P. 312–319. <https://doi.org/10.1109/TAI.2000.889888>

40. Lamine K.B., Kabanza F. Using temporal fuzzy logic for monitoring behavior based mobile robots. In: *Proceedings of the IASTED Conference in Robotics and Applications*. 2000. P. 116–121.
41. Gao D., Wu Ch., Zhang B., Ma X. Signed Directed Graph and Qualitative Trend Analysis Based Fault. *Chinese J. Chem. Eng.* 2010;18(2):265–276. [https://doi.org/10.1016/S1004-9541\(08\)60352-3](https://doi.org/10.1016/S1004-9541(08)60352-3)
42. Berman A.F., Pavlov N.Yu., Nikolaichuk O.A. Method for synthesis and analysis of “failure trees” based on the concepts of mechanism and events kinetics. *Problemy analiza riska = Issues of Risk Analysis*. 2018;15(3):62–77 (in Russ.). <https://www.elibrary.ru/xwpjif>
43. Byun S., Papaalias M., Márquez F.P.G., Lee D. Fault-Tree-Analysis-Based Health Monitoring for Autonomous Underwater Vehicle. *J. Mar. Sci. Eng.* 2022;10(12):1855. <https://doi.org/10.3390/jmse10121855>
44. Bai B., Xie C., Liu X., Li W., Zhong W. Application of integrated factor evaluation–analytic hierarchy process–T-S fuzzy fault tree analysis in reliability allocation of industrial robot systems. *Appl. Soft Comp.* 2022;115:108248. <https://doi.org/10.1016/j.asoc.2021.108248>
45. Manzoor S., Rocha Y.G., Joo S.H., Bae S.H., Kim E.J., Joo K.J., Kuc T.Y. Ontology-Based Knowledge Representation in Robotic Systems: A Survey Oriented toward Applications. *Appl. Sci.* 2021;11(10):4324. <https://doi.org/10.3390/app11104324>
46. Olivares-Alarcos A., Bessler D., Khamis A., et al. A review and comparison of ontology-based approaches to robot autonomy. *The Knowledge Engineering Review*. 2019;34:e29. <https://doi.org/10.1017/S0269888919000237>
47. Gribova V.V., Shalfeeva E.A. The Ontology of processes diagnosis. *Ontologiya proektirovaniya = Ontology of Designing*. 2019;4(34):449–461 (in Russ.). <https://doi.org/10.18287/2223-9537-2019-9-4-449-461>
48. Timoshenko A.A., Zuev A.V., Mursalimov E.S., Gribova V.V., Inzartsev A.V. Description and diagnosis of malfunctions in autonomous uninhabited underwater vehicles based on ontologies. *Ontologiya proektirovaniya = Ontology of Designing*. 2022;3(45):310–324 (in Russ.). <https://doi.org/10.18287/2223-9537-2022-12-3-310-324>
49. Gruber T.R. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*. 1993;5(2):199–220. <https://doi.org/10.1006/knac.1993.1008>
50. Xu Y., Zhang S., Li Y., Liu J. Ontology-based Fault Diagnosis and Maintenance Process Generation of Electromechanical System. *Int. J. Performability Eng.* 2019;15(2):454–463. <https://doi.org/10.23940/ijpe.19.02.p10.454463>
51. Taye M. Understanding Semantic Web and Ontologies: Theory and Applications. *J. Computing*. 2010;2(6):182–192.
52. Su X.M., Ilebrikke L. A Comparative Study of Ontology Languages and Tools. In: Pidduck A.B., Ozsu M.T., Mylopoulos J., Woo C.C. (Eds.). *Advanced Information Systems Engineering (CAISE)*. 2002. P. 761–765. https://doi.org/10.1007/3-540-47961-9_62
53. Haarslev V., Moller R. Racer: A Core Inference Engine for the Semantic Web. In: *Proceedings of the 2nd International Semantic Web Conference (ISWC2003)*. 2003. P. 292–297.
54. Sirin E., Parsia B., Grau B.C., et al. Pellet: A Practical OWL-DL Reasoner. *J. Web Semantics*. 2007;5(2):51–53. <https://doi.org/10.1016/j.websem.2007.03.004>
55. Tsarkov D., Horrocks I. FaCT++ Description Logic Reasoner: System Description. In: *Proceedings of the 3rd International Joint Conference on Automated Reasoning*. 2006. P. 292–297. https://doi.org/10.1007/11814771_26
56. Tsarkov D., Riazanov A., Bechhofer S., Horrocks I. Using Vampire to Reason with OWL. In: *Proceedings of the 3rd International Semantic Web Conference (ISWC2004)*. 2004. P. 471–485. https://doi.org/10.1007/978-3-540-30475-3_33
57. Abburu S. A Survey on Ontology Reasoners and Comparison. *Int. J. Computer Applications*. 2012;57(17):33–39. Available from URL: <https://www.ijcaonline.org/archives/volume57/number17/9208-3748/>. Accessed July 19, 2025.
58. Rubin D.L., Noy N.F., Musen M.A., Protégé: A Tool for Managing and using Terminology in Radiology Applications. *J. Digit. Imaging*. 2007;20(1):34–46. <https://doi.org/10.1007/s10278-007-9065-0>
59. Shen B., Zhao S.Y., Wang J.H. Ontology-Based Fault Diagnosis Knowledge Representation of CNC Machine Tool. *Appl. Mech. Mater.* 2013;427–429:1372–1375. <https://doi.org/10.4028/www.scientific.net/AMM.427-429.1372>
60. Wang D., Tang W.H., Wu Q.H. Ontology-based fault diagnosis for power transformers. In: *IEEE PES General Meeting*. IEEE. 2010. P. 1–8. <https://doi.org/10.1109/PES.2010.5589575>
61. Dogmus Z., Erdem E., Patoglu V. RehabRobo-Onto: Design, development and maintenance of a rehabilitation robotics ontology on the cloud. *Robot. Comput. Integr. Manuf.* 2015;33:100–109. <https://doi.org/10.1016/j.rcim.2014.08.010>
62. Sokolov S.M. An ontological approach to the creation of robotic complexes with an increased degree of autonomy. *Izvestiya YuFU. Tekhnicheskie nauki = Izvestiya SFedU. Engineering Sciences*. 2022;1(225):42–58 (in Russ.). <https://doi.org/10.18522/2311-3103-2022-1-42-59>
63. Lebedev S.L., Panteleyev M.G. Ontology-driven design approach to development of situation awareness subsystem of intelligent agents. *Ontologiya proektirovaniya = Ontology of Designing*. 2016;3(21):297–316 (in Russ.).
64. Bayat B., Bermejo-Alonso J., Carbonera J., et al. Requirements for building an ontology for autonomous robots. *Industrial Robot*. 2016;43(5):469–480. <https://doi.org/10.1108/IR-02-2016-0059>

65. Sirin E., Parsia B., Grau B., Kalyanpur A., Katz Y. Pellet: a practical OWL-DL reasoner. *J. Web Semantics*. 2007;5(2):51–53. <https://doi.org/10.1016/j.websem.2007.03.004>
66. Alaya N., Ben Yahia S., Lamolle M. Towards Unveiling the Ontology Key Features Altering Reasoner Performances. *arXiv preprint*. arXiv:1509.087172015. <https://doi.org/10.48550/arXiv.1509.08717>

About the Authors

Yuri A. Bykovtsev, Cand. Sci. (Eng.), Assistant Professor, Department of Management Problems, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: bykovtsev@mirea.ru. Scopus Author ID 57302607300, ResearcherID KRQ-5339-2024, RSCI SPIN-code 9961-4437, <https://orcid.org/0009-0003-6671-5674>

Valery M. Lokhin, Dr. Sci. (Eng.), Professor, Department of Management Problems, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). Laureate of the State Prize of the Russian Federation in the field of science and technology. Russian Federation State Prize Laureate in the field of education. Member of the Scientific Council of the Russian Academy of Sciences on robotics and mechatronics. Honored Worker of Science of the Russian Federation. E-mail: kpu-mirea@yandex.ru. Scopus Author ID 6602931640, <https://orcid.org/0000-0001-6708-9124>

Sergey V. Manko, Dr. Sci. (Eng.), Professor, Department of Management Problems, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). Laureate of the State Prize of the Russian Federation in the field of science and technology. Russian Federation State Prize Laureate in the field of education. Member of the Scientific Council of the Russian Academy of Sciences on robotics and mechatronics. E-mail: manko@mirea.ru. Scopus Author ID 55761014700, RSCI SPIN-code 2070-1592, <https://orcid.org/0000-0002-6297-8894>

Mikhail P. Romanov, Dr. Sci. (Eng.), Professor, Academic Supervisor of the Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow 119454 Russia). E-mail: m_romanov@mirea.ru. Scopus Author ID 14046079000, RSCI SPIN-code 5823-8795, <https://orcid.org/0000-0003-3353-9945>

Об авторах

Быковцев Юрий Алексеевич, к.т.н., доцент, кафедра проблем управления, Институт искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: bykovtsev@mirea.ru. Scopus Author ID 57302607300, ResearcherID KRQ-5339-2024, SPIN-код РИНЦ 9961-4437, <https://orcid.org/0009-0003-6671-5674>

Лохин Валерий Михайлович, д.т.н., профессор, кафедра проблем управления, Институт искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). Лауреат государственной премии РФ в области науки и техники. Лауреат премии Правительства РФ в области образования. Член научного Совета Российской академии наук по робототехнике и мехатронике. Заслуженный деятель науки РФ. E-mail: kpu-mirea@yandex.ru. Scopus Author ID 6602931640, <https://orcid.org/0000-0001-6708-9124>

Манько Сергей Викторович, д.т.н., профессор, кафедра проблем управления, Институт искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). Лауреат премии Правительства Российской Федерации в области образования. Член научного Совета Российской академии наук по робототехнике и мехатронике. E-mail: manko@mirea.ru. Scopus Author ID 55761014700, SPIN-код РИНЦ 2070-1592, <https://orcid.org/0000-0002-6297-8894>

Романов Михаил Петрович, д.т.н., профессор, научный руководитель Института искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: m_romanov@mirea.ru. Scopus Author ID 14046079000, SPIN-код РИНЦ 5823-8795, <https://orcid.org/0000-0003-3353-9945>

Translated from Russian into English by L. Bychkova

Edited for English language and spelling by Thomas A. Beavitt

Multiple robots (robotic centers) and systems. Remote sensing and non-destructive testing
Роботизированные комплексы и системы. Технологии дистанционного
зондирования неразрушающего контроля

UDC 631.37

<https://doi.org/10.32362/2500-316X-2026-14-4-53-67>

EDN FJRWXM



RESEARCH ARTICLE

Design and manufacturing technology of a folding conveyor for a robotic platform for interflight servicing of multicopters

Vladimir P. Dashevsky[@], Maxim M. Bizin, Yuri V. Galykin

St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg,
199178 Russia

[@] Corresponding author, e-mail: dashevskiy.v@iias.spb.su

• Submitted: 27.10.2025 • Revised: 21.02.2026 • Accepted: 29.05.2026

Abstract

Objectives. The study aims to enhance the operational independence and uptime of unmanned aerial vehicles by developing a ground-based robotic service platform equipped with a unique conveyor system. The key technological requirement for the system is to create a lightweight, foldable conveyor that has a high throughput capacity to ensure the synchronization of all processes involved, including alignment after landing, movement, and maintenance operations.

Methods. Computer simulations and physical modeling approaches were used to evaluate the effectiveness of the proposed conveyor hardware and software solutions. In order to meet the specified operational requirements of the conveyor, a suitable system engineering methodology was employed. During the design phase, a comparative analysis of various kinematic drive systems based on chains, gears, and belts was carried out.

Results. The developed and tested prototype roller conveyor system for a drone port outperforms potential alternative systems, such as belt conveyors, gear-driven roller conveyors, and chain-driven roller conveyor systems, across a range of performance criteria. The main roller pulleys and belt support rollers are manufactured using 3D printing of engineering thermoplastics. The roller support bars are produced using 3-axis computer numerical control milling technology. The design was based on standard aluminum profiles and polymer pipes. Experimental tests of the toothed-belt-driven roller conveyor prototype under loads of up to 100 kg confirmed stable operation at speeds of 5–15 cm/s, including on inclines of up to 6°. Power consumption did not exceed 20–25 W at a working speed of 10 cm/s and a roller working area measuring 1900 × 900 mm.

Conclusions. A new design and manufacturing technology for a folding conveyor was developed and tested experimentally to meet the requirements of a robotic interflight service platform. Testing successfully confirmed the reliability of the conveyor, which is now ready to be integrated into the platform as a critical transport module. This will ensure high platform throughput while minimizing energy consumption.

Keywords: service maintenance, unmanned aerial vehicles, drone, landing, drone port, roller conveyor, ground mobile platforms, robotics

For citation: Dashevsky V.P., Bizin M.M., Galykin Yu.V. Design and manufacturing technology of a folding conveyor for a robotic platform for interflight servicing of multicopters. *Russian Technological Journal*. 2026;14(4):53–67. <https://doi.org/10.32362/2500-316X-2026-14-4-53-67>, <https://elibrary.ru/FJRWXM>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Конструкция и технология изготовления складного конвейера роботизированной платформы межполетного обслуживания мультикоптеров

В.П. Дашевский [@], М.М. Бизин, Ю.В. Галыкин

Санкт-Петербургский Федеральный исследовательский центр Российской академии наук,
Санкт-Петербург, 199178 Россия

[@] Автор для переписки, e-mail: dashevskiy.v@ias.spb.su

• Поступила: 27.10.2025 • Доработана: 21.02.2026 • Принята к опубликованию: 29.05.2026

Резюме

Цели. Целью исследования является повышение автономности функционирования и длительности бесперебойной работы беспилотных летательных аппаратов за счет разработки наземной роботизированной платформы сервисного обслуживания с оригинальной конвейерной системой транспортировки. Ключевым технологическим требованием к системе транспортировки является создание легкого складного конвейера высокой пропускной способности, обеспечивающего синхронизацию всех процессов: выравнивания после посадки, перемещения и выполнения сервисных операций.

Методы. Оценка эффективности разработанного комплекса программно-аппаратных средств конвейера осуществлялась с помощью методов компьютерного и натурного моделирования. Разработка конвейера осуществлялась в рамках методологии системного инженерного проектирования под заданные эксплуатационные требования. На этапе проектирования применялся сравнительный анализ кинематических схем (ленточный, цепной, шестеренчатый, зубчато-ременной).

Результаты. Создана и апробирована конструкция прототипа роликового конвейера дронопорта, которая по совокупности критериев превосходит возможные альтернативы: конвейер с лентой, роликовый конвейер с шестереночной передачей вращения, роликовый конвейер с цепной передачей. При изготовлении образца использовалась 3D-печать из конструкционных термопластов (полиэтилентерефталат-гликоль, PETG) для производства шкивов основных роликов и опорных роликов ремня, технологии трехосевых фрезерных станков с числовым программным управлением – для изготовления опорных планок для роликов, а также стандартные алюминиевые профили и полимерные трубы. Экспериментальные исследования прототипа роликового конвейера с зубчато-ременным приводом под нагрузкой до 100 кг подтвердили стабильную работу на скоростях 5–15 см/с, в т.ч. на подъеме до 6°. Энергопотребление не превышало 20–25 Вт при крейсерской скорости 10 см/с и габаритах рабочей области роликов 1900 × 900 мм.

Выводы. Разработаны и экспериментально апробированы новая конструкция и технология изготовления складного конвейера, отвечающие требованиям роботизированной платформы межполетного обслуживания. Секция конвейера успешно прошла испытания и готова к интеграции в платформу в качестве ключевого транспортирующего модуля, обеспечивающего высокую пропускную способность платформы при минимальном энергопотреблении.

Ключевые слова: сервисное обслуживание, беспилотные летательные аппараты, дрон, посадка, дроно-порт, роликовый конвейер, наземные мобильные платформы, робототехника

Для цитирования: Дашевский В.П., Бизин М.М., Галыкин Ю.В. Конструкция и технология изготовления складного конвейера роботизированной платформы межполетного обслуживания мультикоптеров. *Russian Technological Journal*. 2026;14(4):53–67. <https://doi.org/10.32362/2500-316X-2026-14-4-53-67>, <https://elibrary.ru/FJRWXM>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

The development of autonomous infrastructure for their maintenance, as described in studies on delivery logistics [1], urban transport systems [2], and automated drone ports [3, 4], is an essential element in enhancing the operational efficiency of unmanned aerial vehicles (UAVs).

Recent studies [1–5] have focused on designing mobile drone ports capable of automatically servicing UAVs. The process of servicing a UAV on a ground platform involves three main steps: landing, maintenance, and take-off. In addition, each step involves overcoming a set of difficulties related to the autonomous operation of ground-based and airborne equipment in three main areas: positioning, control, and energy [6–8].

To ensure a safe landing, the ground platform is equipped with additional on-site navigation systems that utilize optical and radio communication channels to facilitate data integration and improve reliability in windy conditions. This is particularly important when the UAV is running low on propulsive power or the lift-to-drag ratio is high. Such systems additionally enhance safety in low visibility scenarios, such as under evening light conditions [9, 10]. Aspects of physical interactions involving drones during the docking process and other aspects of object manipulation are discussed in [11–13].

The primary means for enhancing the endurance of UAVs involves the use of cable-based power transmission systems [14]. However, such systems are only viable for limited applications where the operational radius is restricted and the location of ground-based infrastructure does not hinder the installation of the cable connecting the UAV with the ground-based system. In other cases, approaches that involve charging and replacing UAV batteries at a ground station equipped with an onboard generator or connected to the electrical grid may be more effective.

Automated maintenance on a ground platform enhances the autonomous operational capacity of individual UAVs and groups of similar vehicles. To enable the servicing of multiple UAVs on a platform,

methods for controlling groups of robots that rely on mass service system theory must be employed [15–19].

The ability to perform such tasks is particularly relevant when involving groups of multicopters, such as agricultural drones, where providing increased flight duration directly impacts the economic viability of operations. Here the specific nature of an agricultural drone as a “vehicle” used to carry a payload (e.g., fluids) stipulates that aimless hovering during periods of waiting for servicing or while conducting operational circuits constitutes a highly inefficient use of valuable battery power, which must be replenished with additional time spent on servicing. Thus, in order to minimize time between flights, it is necessary to develop mobile service platforms capable of performing key operations (e.g., battery replacement and refueling) in parallel in field conditions [13].

Based on the analysis of current practices [4, 13], a multicopter typically lands on a stationary platform where maintenance takes place. This arrangement creates a bottleneck, since the platform is occupied for the duration of the process, thus significantly reducing the overall efficiency of the system during group operations.

The novelty of the approach discussed in the present work lies in the transition from a drone port based on a stationary landing pad to one that incorporates a conveyor system. Under this arrangement, landing, servicing, and take-off operations are separated spatially and carried out in parallel [5]. To implement this, a new component is required: a conveyor system for the internal movement of drones between different zones of the drone port.

In order to develop such a system, a number of technological challenges need to be addressed. In particular, precise navigation and control are essential for landing on a mobile platform [7, 10]. For this reason, interaction with the drone, such as securing it and managing its battery, relies on docking algorithms [11] and object capture techniques [12]. The optimization of a group of vehicles can be approached within the framework of multiagent systems theory [15, 16] and mass service theory [17].

The most critical and understudied component of the conveyor-based drone port concept is the

transport module. For a mobile platform operating in field conditions, the transport module should possess a unique set of properties, including foldability, low power consumption, and the ability to be repaired with a minimal set of tools in the field.

Thus, the present work sets out develop a design and manufacturing process for a foldable conveyor system that meets the specified requirements. The proposed design, which incorporates a modular construction, a group gear belt drive, and additive manufacturing technologies (3D printing), aims to ensure the technical viability of the concept [5] in terms of realizing its primary objective: maximum throughput of drones on servicing ports with minimal energy usage.

For clarity, the term “UAV” will hereinafter refer exclusively to a multicopter capable of vertical takeoff and landing. “Service maintenance” refers to the activities performed on a UAV in order to prepare it for its next mission, including replacing the battery and replenishing consumables such as the fuel tank of an agricultural drone.

1. MAIN OBJECTIVES OF AUTOMATED GROUND MAINTENANCE FOR UAVS

As part of the research conducted by the St. Petersburg Federal Research Center of the Russian Academy of Sciences¹ (SPB FRC RAS) into developing technological foundations and novel technical solutions for the seamless control of UAV groups in nondeterministic environments, a method for automating the ground infrastructure for UAV maintenance was proposed in the form of a model for autonomous robotic platforms for servicing agricultural drones [20]. The main objectives to be addressed through automated maintenance are as follows.

Automatic landing of the UAV on the platform. Traditional methods of positioning the UAV using satellite navigation signals do not always guarantee high accuracy in determining the ground coordinates. This can lead to difficulties in achieving precise landing within the designated landing area. The task of landing on a spatially restricted platform is more challenging than landing on the ground, since in the former case landing accuracy becomes paramount. The platform area, which can be transported on a trailer, has dimensions of approximately 2×2.5 m. It is preferable that any deviation from the center of the platform be limited to a maximum of 0.5–0.7 m when the UAV lands on it. To achieve this, the UAV’s positioning system must switch upon entering the platform area from using satellite navigation signals to using error signals

generated by local sources. These can include data from cameras monitoring the platform landing area or from a camera on the UAV trained on light beacons mounted on the platform and the maintenance area.

Geolocation using local radio beacons. While satellite navigation systems can provide accurate positioning for objects anywhere on Earth, they have limited resolution and can be disrupted by electronic countermeasures. Moreover, satellite navigation capabilities may be excessive for localizing UAVs within a single field. In order to ensure operation in the presence of interference, measuring coordinates using radio beacons deployed within a single field is a relevant task. The service platform naturally serves as a base station with known coordinates and a power source, enabling radio communication throughout the UAV’s entire flight duration.

Implementation of a unified data transmission network between the platform and the UAV’s onboard computer when UAV is in the landing area. Within the platform’s maintenance area, there is an increased need for information exchange between two systems. This is due to the activation of other landing support software modules, in addition to flight control tasks. Such modules may require a higher bandwidth communication channel, for example, in order to stream multiple video cameras in order to assess the landing point error. The platform can provide a Wi-Fi network in the immediate vicinity, to which a UAV approaching for landing or in the process of being serviced, will automatically be connected.

Transmission of some control commands from the ground station to the UAV flight controller. While the flight controller only supports operator control in the simplest landing mode, the control signals coming from the operator’s remote control must be replaced by signals generated by the platform’s onboard computer, which are based on an analysis of images received from the landing site surveillance camera.

Moving the drone along the conveyor belt in automatic mode. Since a drone generally lands on the platform with a positioning error, further mechanized operations are required to move it into the working area where manipulations are to be carried out. First, the drone must be levelled and moved onto the conveyor belt. Therefore, in order to move a heavy drone across the platform, the supports must include rollers that enable lateral movement. However, once the drone is on the conveyor, the rollers must be locked to prevent it from rolling off in the event of the platform tilting slightly. While the rollers may be physically controlled by the drone’s onboard computer, the command is transmitted via radio from the platform’s onboard computer.

Transmission of data on the UAV’s specifications and current status to adjust reenergizing instructions. During its operational life, the platform is likely to be

¹ <https://spcras.ru/>. Accessed May 24, 2026. (In Russ.).

used with different types of drone, each having different dimensions, fuel tank capacities, battery capacities, and so on. Additionally, a drone may return from a mission in emergency mode, having not used all consumables, or with malfunctioning automatization systems, requiring a switch to backup semi-automatic systems. This information must be transmitted to the platform and used in maintenance algorithms.

Optical monitoring to assess the condition of the UAV and determine whether human intervention is required during maintenance. Under normal operating conditions, personnel should avoid close proximity to the platform due to the risk of injury from ground-based engine noise and dust, and the risk of splashing from working fluids. However, in an abnormal situation, only a human operator can assess the feasibility of carrying out the necessary repairs. Thus, the optical monitoring system will identify potential abnormal situations and signal the need for corrective action.

Finally, it is also important *to record video from the platform and drone cameras, as well as real-time telemetry data, while the platform is connected to a single network.*

2. STRUCTURAL-FUNCTIONAL MODEL OF A ROBOTIZED MAINTENANCE PLATFORM

The platform is designed as a module mounted on a three-axle automotive trailer measuring 2×4 m. The key concept aimed at radically increasing throughput involves transitioning from a single area for landing, maintenance, and take-off to a conveyor-based system [5]. In this system, landing, maintenance (refueling and battery replacement), and take-off operations are spatially separated and performed in parallel.

To ensure the compactness of the mobile platform, the conveyor consists of four foldable flaps connected

by hinges and servo drives. When folded, these flaps form the floor, walls, and ceiling of the platform area on the trailer (Fig. 1).

Figure 2 shows the structural-functional model of the robotic service platform. When in the operating position, all the flaps unfold to form a unified flat platform that is divided into three areas: landing (1), maintenance (2), and take-off (3). The landing area is formed on the conveyor flap, which serves as the ceiling when folded. The maintenance area is located on a stationary flap that functions as the floor, which is mounted on the supporting frame of the platform; this, in turn, is attached directly to the trailer. A side flap connecting the floor and ceiling sections serves as a transition area. This is designed to move the landing area away from the body of the trailer, thereby reducing the risk of the UAV propeller contact with the trailer walls in case of an inaccurate landing. The take-off area is located in the remaining side panel, which folds backwards in the opposite direction.

After landing, the UAV is automatically transported to a maintenance and take-off area using the conveyor to enable all work areas to function in parallel. All operations are carried out either fully automatically or through remote control. The transition to this inter-flight UAV maintenance mode enhances the efficiency of platform resource utilization, as well as ensuring the safety of personnel by eliminating the risk of personnel being in the area covered by the UAV.

The maintenance area (2), which is located on the central fixed flap, houses automated systems used to prepare the UAV for its next mission. This area is equipped with a gantry-type manipulator that facilitates operations such as replacing heavy batteries or replenishing the UAV's solution tank. The maintenance area measures approximately 2 m by 2.8 m in size. Beyond this area, on the base of the platform, there is also a technical area that includes tanks containing

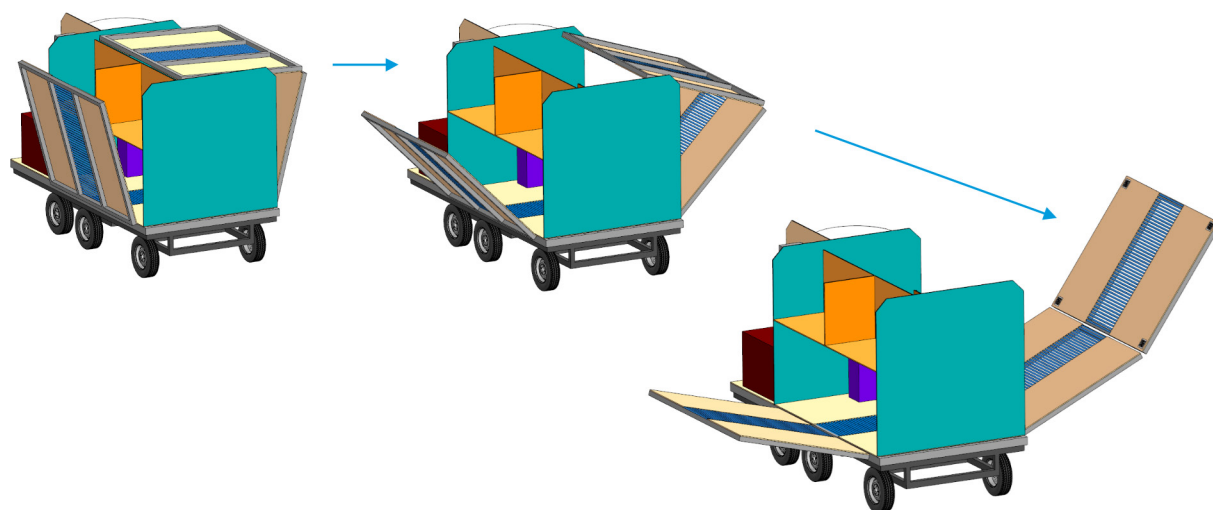


Fig. 1. Unfolding the conveyor of the mobile UAV maintenance platform

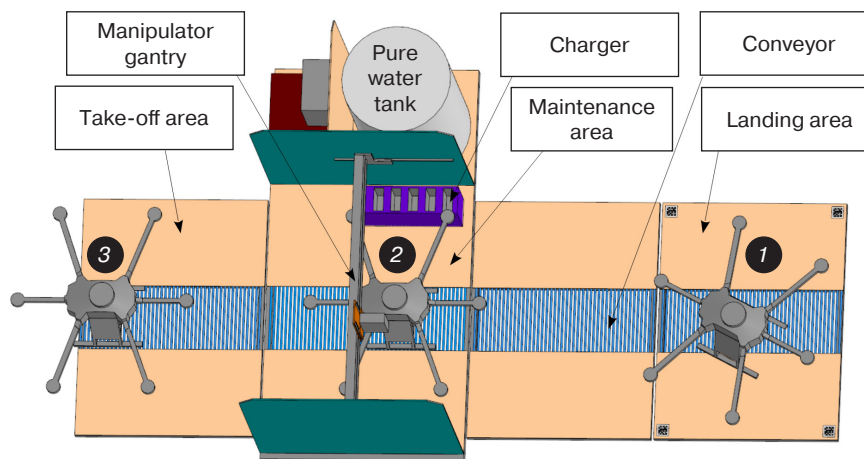


Fig. 2. Main components of the mobile service platform

concentrated working fluids, a solution preparation unit, a pump station for transferring fluids, a generator area, and an area for automating the entire platform's control.

The maintenance process proceeds as follows. Once the UAV has landed and been moved to the maintenance area, it is secured in place. The gantry-type manipulator then removes the discharged battery and installs a fully charged one. Simultaneously, the fluid replenishment system fills the UAV's tank with working fluids. Once all operations are complete, the conveyor transports the UAV which is ready for launching to the take-off area (3), while the vacated maintenance area is ready to receive the next UAV. This ensures a continuous service cycle for the platform.

The conveyor system plays a crucial role in integrating the various process areas into a unified service system, since the continuous and efficient operation of the conveyor directly impacts the platform's overall throughput capacity and economic efficiency in servicing the UAV fleet. However, the unique challenges associated with the platform's mobility, need for a folding design, and requirement for autonomous operation in field conditions present unique design requirements for the conveyor system that are not found in conventional stationary industrial solutions. These challenges will be discussed in the next section.

3. JUSTIFICATION OF REQUIREMENTS FOR THE TRANSPORT MODULE

Based on the platform's configuration, mobile nature and intended purpose of servicing a fleet of heavy agricultural drones, the following key requirements are formulated for the transport module:

1. Load capacity and dynamic load of up to 100 kg. This requirement is based on the take-off weight of the target UAVs (60–80 kg), taking into account emergency landings and the vertical speed that generates additional dynamic load.
2. A conveyor travel speed of 5–15 cm/s. With a travel distance of approximately 4 m between the centers of the take-off and maintenance areas, this speed ensures an acceptable travel time of 25–80 s. Taking into account a service time of 5–8 min, this fits well within the total inter-flight interval of 10–15 min of flight on a single battery, without creating excessive dynamic loads on the UAV structure or the conveyor itself.
3. Dimensions and adaptability. The conveyor width of 700–900 mm is sufficient to ensure all UAV landing gear legs (with a base dimension of approximately 600 × 600 mm) can fit onto the conveyor reliably. The variable section length is adaptable to the size of the conveyor flap (1600–2000 mm), which needs to be foldable to fit within the dimensions of a car trailer.
4. Energy efficiency is essential, since, due to the limited power supply available on an autonomous field platform, the energy consumed per flap drive must be within the range of ≤30–50 W. Preliminary estimates using a simplified model that considers rolling friction forces and inclines up to 10 cm/m have yielded values between 20 and 30 W for a model with a friction coefficient of 0.2 and a load of 100 kg.
5. The weight of the structure should be minimized. This is a primary consideration to ensure mobility and the ability to deploy the flaps using compact, low-power actuators.
6. Field maintainability. In the event of a failed component, it should be possible for a single or two operators to replace it without dismantling adjacent assemblies or requiring the use of specialized tools.
7. Fault tolerance. The system must continue to function if one or more adjacent components, such as individual conveyor rollers, fail.
8. Operation on uneven surfaces. The ability of the system to transport cargo on an incline of at

least 5° – 6° (approximately 10%) simplifies the requirements for precise platform installation on uneven ground in the field.

9. Manufacturing feasibility. The design should be such that it can be manufactured using readily available digital equipment, such as 3D printers and CNC (computer numerical control) machines [21], in order to reduce costs and facilitate adaptation to specific dimensions.

The described set of requirements fundamentally distinguishes the design of a mobile conveyor platform from that of fixed industrial conveyors. Fixed conveyors prioritize maximum throughput, durability for round-the-clock operations, and minimizing the unit cost of bulk cargo transportation through the use of heavy and high-strength structures, as well as high-power motors often measured in kilowatts. However, conventional calculation methods [22] do not account for critical parameters of mobile platforms, such as structural lightness, manufacturability, and maintainability in the field. Therefore, a more experimental approach has been adopted to directly verify the fulfilment of the formulated requirements. This involves a comparative analysis of kinematic schemes and fabrication of a prototype using adaptive technologies such as 3D printing and CNC (computer numerical control) machines, along with field measurements of key performance parameters under load. The following sections provide a detailed examination of the unsuitability of conventional conveyor designs and the outcomes of these experiments.

4. ANALYSIS OF DESIGN OPTIONS AND DRIVE SELECTION

During the design process for the conveyor system, two design options were explored: belt and roller conveyors. In this section, we examine both options and their respective advantages and disadvantages.

The belt conveyor consists of a flexible belt that is stretched over drive and tension rollers to serve as the carrying and propulsion mechanism. Detailed calculations for the parameters of various conveyor configurations can be found in [18]. The UAV rests on the belt, which is moved by friction between its internal surface and the rollers. The belt conveyor offers potential benefits, such as continuous support and smooth operation. However, it also has several significant drawbacks:

- The belt must be custom-manufactured to a specific length, requiring specialized equipment and incurring economic costs.
- To achieve the required strength and rigidity for a length of 2 m and a load capacity of approximately 100 kg, a heavy-duty frame and substantial drum assemblies are required. The conveyor belt itself

also has a significant weight, which contradicts requirement 5 (lightweight design).

- Replacement or repair of the belt (for example, in the event of a breakage or damage) necessitates complete removal of the belt from the entire segment. This is labor-intensive and contradicts requirement 6.
- Specialized tools (tensioning devices) are frequently required to adjust the tension of the belt after repair. This can be challenging to accomplish in the field. Belt tensioning is a critical and non-trivial procedure.
- Precise alignment and parallelism of the drums are essential to prevent the belt from running off the drums. Achieving this during on-site repairs by a team of 1–2 people in the field can be challenging.
- Belt conveyors require regular maintenance where the belt and deck come into contact, including cleaning and applying silicone lubricant to reduce friction and prevent wear. Under field conditions in which UAVs regularly stir up dust that settles on all structural components, the need for cleaning and lubrication is significantly higher, which will place a burden on the end user.
- Significant power is required to overcome the frictional forces between the entire belt surface and the stationary support surface (deck), as well as to lift the load uphill when the platform is inclined. Even compact, smaller-sized sports treadmills are equipped with motors rated at 1–3 kW. This is due to the need to overcome sliding friction across the entire belt surface.
- The belt creates a solid, continuous, and smooth surface across the entire conveyor area. When landing on the edge, such a platform forms a pronounced “aerodynamic cliff.” The air cushion effect ends abruptly, creating the greatest possible asymmetry and, consequently, a strong overturning moment. This leads to control instability during the final stage of landing and increases the risk of accidents.

In a roller conveyor system, cargo moves along a series of rotating rollers (shafts), which are powered by a common driving mechanism typically comprising a chain, belt, or direct gear connection via intermediate gears. The rollers provide a load-carrying function, while the driving mechanism performs the transportation function. This design offers the following advantages:

- Lightweight design and modularity. Since it does not need to bear significant tensile loads, the frame of the conveyor can be designed to be lighter and more compact than traditional belt conveyors. The individual rollers, which are made from lightweight materials such as aluminum or engineered plastics (PETG, PA6, PA12, and ASA²), are also

² Polyethylene terephthalate-glycol, polyamide 6, polyamide 12, acrylonitrile styrene-acrylate.

lightweight. Each roller forms an independent module; thus, if one fails, it can simply be removed from the bearing assembly and replaced without having to disassemble the entire conveyor section. In extreme cases, rollers that perform critical functions (for example, at the junctions between flaps) could be swapped with rollers that perform less critical functions.

- Assembly efficiency. All precision operations, including bearing installation and assembly of rollers with gears/pulleys, take place in the workshop using specialized workstations. During final assembly, pre-assembled units, such as a frame with installed rollers and a tensioned drive belt, are mounted onto the platform frame as a single unit. This reduces the need for adjustment to tension a single common belt, significantly simplifying the process compared to adjusting the entire belt system.
- The failure or absence of a single roller does not prevent the conveyor from operating. The drone can pass through the defective section, using adjacent rollers, with minimal force and without loss of stability.
- The roller conveyor utilizes rolling friction. The torque required to rotate each roller on its bearings is extremely low. Further testing has shown that a power input of only 25–30 W is sufficient to transport a 100-kg UAV, even when climbing inclines. Therefore, the roller conveyor represents a technologically advanced and highly

energy-efficient solution, critical for an autonomous platform with limited energy resources.

- According to its design, the roller conveyor has slots or perforations. This is because the air expelled by propellers does not create a continuous, stable cushion. Since a significant portion of the air dispersing downward through the spaces between the rollers, the turning moment is reduced to enhance the drone's stability during final landing.

A roller conveyor also has several drawbacks, including the relatively large number of components required and the need for them to be manufactured with great precision. While a chain-based solution makes it easy to adjust the chain length to match the required conveyor length, a chain is heavier and more expensive. Most importantly, it raises the issue of manufacturing sprockets in the correct size for each roller and securing them to the rollers. In the case of a belt drive, 3D-printed monolithic flanges that match the rollers and belt dimensions solve the problem of sprockets and attachment to the rollers, leaving only the issue of selecting a belt of the required length and adjusting its tension.

A comparison of the two conveyor designs (belt and roller) is presented in Table 1.

The analysis indicates that for a mobile and foldable drone port, where the key requirements are lightweight, ease of assembly, and field serviceability, a roller conveyor system would be the optimal solution. Potential drawbacks, such as vibration and exposure, could be addressed during the design process:

Table 1. Comparison of conveyor implementations

Criteria	Belt conveyor	Roller conveyor
Cost	High	Low/moderate
Lightweight design	Low	High
Assembly efficiency	Low	High
Maintainability	Low	Very high
Fault tolerance	Low	High
Smooth running	High	Moderate
Environmental protection	High	Moderate/low
Energy consumption	Very high	Low
Maintenance requirements	High	Low
Resistance to abrasives	Low	Moderate
Aerodynamic stability during landing	Low	High



Fig. 3. Main roller shaft with stiffeners, flange, and pulley installed, along with the outer pipe prior to assembly

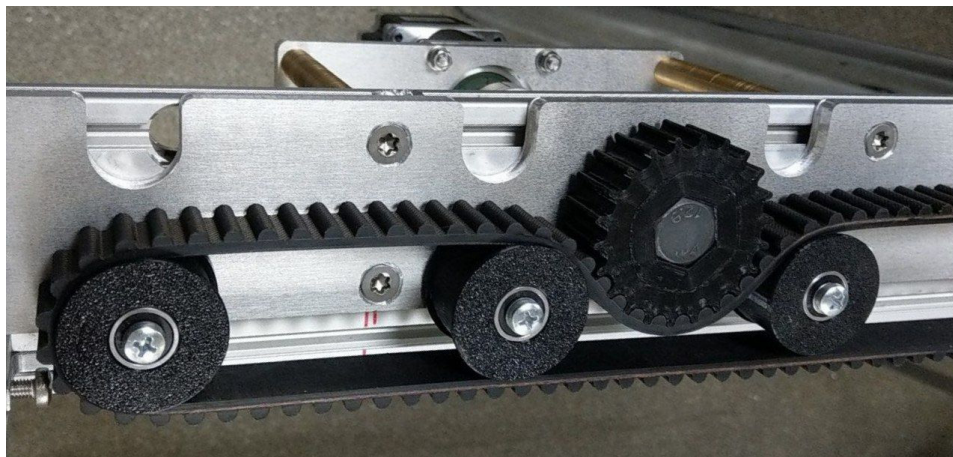


Fig. 4. Support guide with a toothed belt drive at the mounting point of the drive with a stepper motor and drive pulley

- By using 3D printing, high-precision rollers could be manufactured and easily customized to suit any drive belt specifications.
- The installation of easily removable protective covers or the use of sealed bearing assemblies could eliminate the risk of contamination.

Despite its apparent simplicity, a belt conveyor may not meet all critical requirements, particularly in terms of reparability and weight.

5. CONVEYOR DESIGN WITH GROUP GEAR-BELT DRIVE

The design selected for implementation relies on key principles such as modularity, ease of manufacturing, minimal weight, and maximum maintainability. A conveyor module is a self-contained unit that can be integrated into a folding platform flap. For the prototype, the length is 1880 mm, the width is 800 mm, and there are 20 rollers with a pitch of 96 mm.

The conveyor system is designed using a modular approach that separates precision workshop assembly from simpler field operations. The base consists of two aluminum profiles joined by struts to form a frame. Each

guide is supported by a set of standard panels made from aluminum sheets using CNC machines. These panels ensure precise positioning of all components.

Now, we examine the key components and hybrid manufacturing techniques involved. The primary roller assembly comprises (Fig. 3):

- shaft: 15 mm-diameter aluminum tube;
- body: 50 mm-diameter polypropylene (PP) tube;
- end elements: 3D-printed polyethylene terephthalate glycol (PETG) pulley with 14 teeth and a solid flange;
- stiffness enhancer: 3D-printed insert that ensures coaxiality between the shaft and the roller body, thus preventing deformation of the PP tube under load.

The timing belt drive assembly (Fig. 4), which provides a group drive, consists of:

- 20 mm wide, 8 mm pitch timing belt;
- belt support rollers, which are 3D-printed from PETG with press-fit bearings. These are installed under each main roller to press the belt against its pulley, ensuring engagement without slippage;
- drive unit consisting of a 3D-printed drive pulley with 22 teeth, a press-fit steel shaft based on a high-strength bolt, and a coupling connecting it to the motor gearbox.



Fig. 5. Mounting the motor via a plate allows it to be adapted to any pitch of the main rollers

The power unit includes a NEMA23/NEMA24³ stepper motor assembled with a planetary gearbox and mounted on a separate mounting plate, as shown in Fig. 5.

The developed assembly technology ensures serviceability. The assembly process is divided into workshop and field stages:

1. Workshop pre-assembly. All panels and belt support rollers are mounted on each support guide. The toothed belt is then tensioned and two complete modules, the “left” and “right” sides of the conveyor, are formed with the belt and drive pulley installed.
2. Frame assembly. The two guides are pre-installed into the platform and their adjustable fasteners are securely attached to the platform frame. They are then connected by cross-struts of a specific length that match the length of the main rollers.
3. Motor installation. The motor assembly including gearbox is mounted on a mounting plate and secured to support posts threaded through the load-bearing frame of the platform. The drive pulley shaft is connected to a coupling mounted on the output shaft of the gearbox to compensate for any alignment issues between the shafts.
4. Installation of rollers. The pre-assembled main rollers are inserted into the bearing housings.
5. Completion of assembly. This involves tensioning the belt, securing all fasteners on the platform frame, and connecting the motor to the control system.

The primary benefit of this technology is that the main roller can be easily replaced without affecting the belt tension or structural integrity. If the entire drive needs to be

replaced, only the pulley and motor assembly need to be removed. To remove the belt, all rollers must be removed and the supporting rollers loosened. Due to the light weight of the components, these operations can be performed by a single individual using minimal specialized tools.

6. EXPERIMENTAL RESULTS

The experimental phase was designed to comprehensively verify the performance characteristics of the developed conveyor section (Fig. 6) and confirm compliance with the specified requirements. Power consumption was monitored throughout the tests, which were conducted on a horizontal surface and on an incline under various loads (70–100 kg) and speeds (5–15 cm/s).

To test the conveyor’s fault tolerance and evaluate its behavior under abnormal conditions, experiments involving the deliberate removal of rollers from the working area were conducted (Fig. 7). The conveyor successfully transported a 70 kg load when one roller was missing, creating a gap of ~140 mm, and when two consecutive rollers were missing, creating a gap of ~240 mm. Although there was a slight increase in vibration as the load rolls over the gap, the system remained fully controllable without recorded control failures or motor stops. The load was confirmed to be effectively redistributed to the adjacent rollers.

Therefore, the design is confirmed to provide high fault tolerance. Since the conveyor can still operate even if up to two adjacent rollers fail, the selection of a modular roller system with group drive is validated. The creation of technological joints between conveyor sections, where the gap between the rollers needs to be increased due to supporting structures on the platform, is also confirmed as justified.

In order to test operation under less than ideal conditions, the conveyor was set up with an incline of 10 and 20 cm over a 2 m length. When equipped with

³ The NEMA23 and NEMA24 frame sizes (57 × 57 mm and 60 × 60 mm) differ slightly. However, they share the same mounting hole pattern (bolt circle), allowing them to be used interchangeably on the same bracket. This provides design flexibility for torque upgrades without mechanical modification.

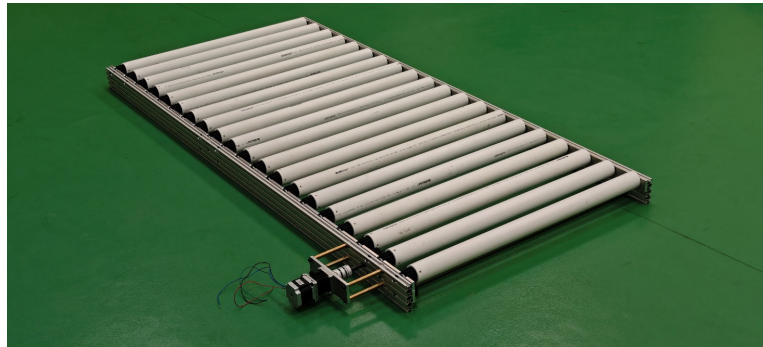


Fig. 6. Assembled prototype of the roller conveyor section



(a)



(b)

Fig. 7. Experiments on overcoming an obstacle formed by removed rollers:
(a) with one roller removed, (b) with two rollers removed

a power unit based on an NEMA23/NEMA24 motor and NEMA23 planetary gearbox (1 : 20), the conveyor reliably transported a 100 kg test load up an incline at speeds between 5 and 15 cm/s. Power consumption during this process did not exceed 25–30 W.

During testing of an agricultural drone having a weight of 70 kg, an additional issue was identified. This issue was not related to the performance of the drone's drive system, but rather due to insufficient traction between the soft supports of the drone and the polypropylene rollers on a steep slope. However, when the slope was reduced to 2.9°, this issue was resolved.

Thus, the mechanical components and the conveyor drive have sufficient torque reserve to operate on a significant incline, even when errors in platform

installation are taken into account. The main challenge lies in ensuring reliable contact between the transported apparatus and the rollers.

A key prior requirement was to minimize power consumption (≤ 30 –50 W). To meet this criterion, a comparative analysis of three drive configurations was conducted (see Table 2).

The results obtained lead to the following conclusions:

1. The underpowered unit (JK57HS56) has insufficient torque to operate stably, confirming the importance of having a torque reserve to overcome variable resistance in a long belt drive.
2. Drives offering increased torque (NEMA 23/24 with a 1 : 20 gear ratio) provide stable operation over

Table 2. Characteristics of the conveyor with three different power units

Linear speed, cm/s	Power supply for the power plant, W		
	JK57HS56-2804 + 1 : 10 (Changzhou Jkongmotor Co., China)	57AM23 + 1 : 20 (Shenzhen Rtelligent Technology Co., China)	60HB87F5AB + 1 : 20 (Changzhou Holry Electric Technology Co., China)
5	19–22	20–25	19–24
10	21–33, failure	18–22	18–23
15	Control failure	22–32	23–30

the entire speed range without increasing power consumption. At a working speed of 10 cm/s, their power consumption is 18–23 W.

3. Observed power fluctuations ($\pm 20\%$) can be attributed to variable friction in the multipoint transmission and dynamic changes in load on the roller bearings due to cargo movement.

Therefore, for the final design, a power unit based on a NEMA23/NEMA24 stepper motor with a phase current of at least 5A and a 1 : 20 planetary gearbox is recommended. This configuration provides the necessary reliability and stability, as well as a low power consumption of 20–25 W when transporting a 100 kg load. This is an order of magnitude lower than the energy consumption of belt conveyors.

All tests are distinguished by the following positive characteristics:

- Smooth and synchronized operation. No jerks; uniform rotation of all rollers in cruising mode.
- Low noise level. Operation is accompanied only by the steady hum of the motor and gearbox, slight vibration along the UAV's fuselage and slight swaying due to the flexibility of the rollers.
- Reliability of the belt drive. No instances of belt slippage or derailment were recorded.
- Control stability. No loss of control occurred when using powerful motors across the entire speed range. The UAV can be accelerated sharply to full speed without loss of control, accompanied by a jolt and slight slippage of the rollers under its supports.

Thus, the experimental tests have confirmed that the designed conveyor prototype meets all the essential requirements specified in Section 3 (Table 3).

The series of conducted tests thus not only demonstrate the functionality of the proposed design, but also quantitatively confirm its unique benefits for use in mobile autonomous platforms, such as extreme

energy efficiency, high reliability, and adaptability to suboptimal operating conditions.

CONCLUSIONS

This paper presents the results of the development and comprehensive experimental research of a folding conveyor representing a key module for a mobile robotic platform designed for inter-flight servicing of multicopter-type UAVs. The conducted research encompassed the full cycle, from formulating unique requirements stemming from the conditions of mobile deployment and autonomous field operation, through a comparative analysis of design schemes, to the fabrication of a prototype and its comprehensive field tests. The proposed approach to automating the ground infrastructure for servicing unmanned aerial vehicles minimizes flight intervals, increases the operational speed of aircraft, and completely eliminates the need for personnel to be in close proximity to operating machinery and hazardous materials.

The main scientific and technical results of the work are as follows:

1. A unique structural and functional model of a robotic service platform for multirotor UAVs has been developed. This platform features dedicated technological zones for UAV landing, servicing, and takeoff. It utilizes a folding roller conveyor with a toothed belt drive for automatic UAV transport between zones. This design increases the throughput of the entire platform during UAV group servicing, including battery replacement, refilling of operating fluid, and transport to the launch pad.
2. A new modular roller conveyor design has been implemented. This design utilizes a group toothed belt drive with a single drive pulley and increased torque from the power plant to overcome variable resistance in the extended belt drive when moving

Table 3. Main results and their compliance with requirements

Requirement	Target value	Experimentally confirmed result
Energy consumption	$\leq 30\text{--}50$ W	18–25 W (speed 10 cm/s, weight 70 kg)
Fault tolerance	Operation with roller failure	Confirmed (removal of 1–2 rollers)
Operation on an incline	Up to $5^\circ\text{--}6^\circ$	Confirmed for mechanical components and drive
Load capacity	100 kg	Confirmed
Speed	5–10 cm/s	Ensured (5–15 cm/s)
Maintainability	Roller replacement in the field	Implemented through design (modularity)

heavy UAVs. This design enables the formation of a four-section conveyor line that folds compactly for transport on the ground-based robotic service platform for agricultural multicopter UAVs.

3. The following critical operational advantages of the created prototype were experimentally proven:
 - Extremely low energy consumption. 18–25 W when transporting a load of 70–100 kg at a speed of 10 cm/s, which is an order of magnitude lower than the consumption of belt conveyors.
 - High fault tolerance. The system maintains operability even if one or two adjacent rollers fail.
 - Adaptability to non-ideal conditions. The conveyor remains operational on slopes up to 5° – 6° (a gradient of 10–20 cm/m), reducing requirements for the precision of platform leveling in the field.
 - Maintainability. The ability to quickly replace any main roller, motor, or drive pulley without disassembling the entire conveyor.
4. The optimal configuration of the power unit (a NEMA23/NEMA24 stepper motor with a 1 : 20 reduction gearbox) was determined such as to ensure stable operation of the extended belt transmission with minimal energy consumption.

The practical significance of the work lies in the creation and experimental validation of a transport module, representing a key element of a drone servicing platform. Without an efficient internal transport system capable of organizing a continuous servicing cycle for a group of aircraft, operations become highly inefficient: a multicopter waiting for its turn to land is forced to either waste its primary limited resource (battery

charge) by hovering near the landing area or to land off-platform, which reduces operational efficiency. The proposed conveyor scheme eliminates this drawback by solving the problem of non-productive energy losses and providing the basis for implementing full-featured mobile platforms capable of servicing groups of multicopter-type UAVs with minimal inter-flight intervals.

Thus, the results confirm the feasibility and high efficiency of the proposed design and technological solution. Further research should focus on integrating the conveyor with automatic battery replacement and refueling systems, as well as optimizing control algorithms to minimize time costs within the entire service cycle.

ACKNOWLEDGMENTS

This study was conducted under the state assignment No. FFZF-2025-0012.

Authors' contributions

V.P. Dashevsky—study concept, formulating the main idea and key technical requirements, structural design, selecting components, analytical calculations, developing the experimental setup and conducting experiments, writing and editing the manuscript.

M.M. Bizin—fabrication of mechanical components, assembly of the prototype, experimental testing and measurements, data collection, literature review.

Yu.V. Galykin—CAD design and 3D modeling, component selection and procurement, fabrication of parts, assembly of the prototype, experimental testing, literature review.

REFERENCES

1. Kong J., Xie M., Wang H. Integrating Autonomous Vehicles and Drones for Last-Mile Delivery: A Routing Problem with Two Types of Drones and Multiple Visits. *Drones*. 2025;9(4):280. <https://doi.org/10.3390/drones9040280>
2. Rinaldi M., Primates S. Comprehensive Task Optimization Architecture for Urban UAV-Based Intelligent Transportation System. *Drones*. 2024;8(9):473. <https://doi.org/10.3390/drones8090473>
3. Sun W., Luo Z., Huang K., Shi J. Joint Deployment and Coverage Path Planning for Capsule Airports with Multiple Drones. *Drones*. 2023;7(7):457. <https://doi.org/10.3390/drones7070457>
4. Blaha L., Severa O., Goubey M., Myslivec T., Reitinger J. Automated Drone Battery Management System—Droneport: Technical Overview. *Drones*. 2023;7(4):234. <https://doi.org/10.3390/drones7040234>
5. Saveliev A.I., Lebedeva V.V., Lebedev I.V., Kamynin K.V., Kuznetsov L.D., Ronzhin A.L. UAV group management when working out of crisis flight situations in solving transport problems. *Izvestiya YuFU. Tekhnicheskie nauki = Izvestiya SFedU. Engineering Sciences*. 2022;1(225):110–120 (in Russ.). <https://doi.org/10.18522/2311-3103-2022-1-110-120>
6. Rodionov A.S., Matkurbanov T.A. Planning UAV flight trajectories for monitoring large areas. *Informatika i avtomatizatsiya = Informatics and Automation*. 2025;24(3):791–827 (in Russ.). <https://doi.org/10.15622/ia.24.3.3>
7. Kuleshov S.V., Kvasnov A.V., Zaytseva A.A., Ronzhin A.L. Multipurpose approach to visual navigation based on landscapes for UAVs operating in GNSS-unavailable conditions. *Izvestiya YuFU. Tekhnicheskie nauki = Izvestiya SFedU. Engineering Sciences*. 2025;2(244):269–278 (in Russ.). <https://doi.org/10.18522/2311-3103-2025-2-269-278>
8. Muslimov T.Z. Collision Avoidance in Circular Motion of a Fixed-Wing Drone Formation Based on Rotational Modification of Artificial Potential Field. *Informatika i avtomatizatsiya = Informatics and Automation*. 2025;24(1):72–98 (in Russ.). <https://doi.org/10.15622/ia.24.1.4>
9. Beliaikov M.E., Diane S. Algorithms for the visual analysis of an environment by an autonomous mobile robot for area cleanup. *Russian Technological Journal*. 2023;11(4):26–35. <https://doi.org/10.32362/2500-316X-2023-11-4-26-35>

10. Antipov A.S., Kokunko Yu.G., Wolf D.A., Shirokov A.S. Synthesis of Combined Quadrocopter Attitude and Altitude Control Based on Block Approach with Sigmoidal Feedbacks. *Informatika i avtomatizatsiya = Informatics and Automation*. 2025;24(3):745–790 (in Russ.). <https://doi.org/10.15622/ia.24.3.2>
11. Golubov V.V., Manko S.V. Automation of autonomous mobile robot docking based on the counter growth rapidly exploring random tree method. *Russian Technological Journal*. 2024;12(1):7–14. <https://doi.org/10.32362/2500-316X-2024-12-1-7-14>
12. Voronkov A.D., Diane S.A. Continuous genetic algorithm for grasping an object of a priori unknown shape by a robotic manipulator. *Russian Technological Journal*. 2023;11(1):18–30. <https://doi.org/10.32362/2500-316X-2023-11-1-18-30>
13. Ngo K.T., Solenaya O.Ya., Ronzhin A.L. Analysis of mobile robotic platforms used to batteries service of unmanned aerial vehicles in autonomous missions. *Trudy MAI*. 2017;95:11 (in Russ.). <https://elibrary.ru/zhnxxyh>
14. Davlyud I.I. Calculation of the main operational characteristics of a tethered high-altitude ship-based system. *Russian Technological Journal*. 2024;12(4):40–50. <https://doi.org/10.32362/2500-316X-2024-12-4-40-50>
15. Manko S.V., Lokhin V.M., Diane S.K. Prototype multi-agent robotic debris removal system: principles of development and experimental studies. *Russian Technological Journal*. 2022;10(6):28–41. <https://doi.org/10.32362/2500-316X-2022-10-6-28-41>
16. Le V.N., Ronzhin A.L. Review of intelligent control systems and robotic tasks in aquaculture production. *Morskije intellektual'nye tekhnologii = Marine Intellectual Technologies*. 2024;1-1(63):171–180 (in Russ.). <https://doi.org/10.37220/MIT.2024.63.1.021>
17. Arkhipov I.S., Erofeeva V.A., Granichin O.N., Kiselev V.A., Chernov A.O. Accelerated Local Voting Protocol for Single Remote Controller Robot Swarm System. *Informatika i avtomatizatsiya = Informatics and Automation*. 2025;24(4):1029–1058 (in Russ.). <https://doi.org/10.15622/ia.24.4.2>
18. Saveliev A., Anikin D., Ronzhin A., Erokhin G., Agafonov V. System for Planing Seismic Sensors Based on Actions of UAVs Group with Optimized Flight Plan. In: Ronzhin A., Savage J., Meshcheryakov R. (Eds.). *Interactive Collaborative Robotics. Lecture Notes in Computer Science*. Cham: Springer; 2024. V. 14898. P. 334–343. https://doi.org/10.1007/978-3-031-71360-6_25
19. Goncharenko V.A., Khomenko A.D., Abu Hasan R. A Compositional Approach to the Simulation of Queuing Systems with Random Parameters. *Informatika i avtomatizatsiya = Informatics and Automation*. 2024;23(6):1577–1608 (in Russ.). <https://doi.org/10.15622/ia.23.6.1>
20. Ronzhin A.L., Dashevsky V.P., Krestovnikov K.D. Interaction of aerial drones with ground-based robotic vehicles. In: *Proceedings of the 18th All-Russian Multiconference on Control Problems (MKPU – 2025)*. V. 1. *Robotics and Mechatronics (RM – 2025)*. Tula; 2025. P. 145–148 (in Russ.). <https://elibrary.ru/ibktbu>
21. Galykin Yu.V., Dashevsky V.P. Automation of the manufacturing of device enclosures using a specialized CNC machine for the production of fastening bars. *Vestnik MGTU “Stankin” = Vestnik MSUT “Stankin”*. 2024;1(68):119–128 (in Russ.). <https://elibrary.ru/mcymqz>
22. Poskrebyshchev V.A., Zinov'ev A.A., Lokhova N.A., Is'ko A.B., Belykh S.A. *Mekhanicheskoe oborudovanie dlya proizvodstva stroitel'nykh materialov i izdelii (Mechanical Equipment for the Production of Building Materials and Products)*. Bratsk; 2009, 378 p. (In Russ.). ISBN 978-5-8166-0261-7

СПИСОК ЛИТЕРАТУРЫ

1. Kong J., Xie M., Wang H. Integrating Autonomous Vehicles and Drones for Last-Mile Delivery: A Routing Problem with Two Types of Drones and Multiple Visits. *Drones*. 2025;9(4):280. <https://doi.org/10.3390/drones9040280>
2. Rinaldi M., Primates S. Comprehensive Task Optimization Architecture for Urban UAV-Based Intelligent Transportation System. *Drones*. 2024;8(9):473. <https://doi.org/10.3390/drones8090473>
3. Sun W., Luo Z., Huang K., Shi J. Joint Deployment and Coverage Path Planning for Capsule Airports with Multiple Drones. *Drones*. 2023;7(7):457. <https://doi.org/10.3390/drones7070457>
4. Blaha L., Severa O., Goubey M., Myslivec T., Reitering J. Automated Drone Battery Management System—Droneport: Technical Overview. *Drones*. 2023;7(4):234. <https://doi.org/10.3390/drones7040234>
5. Савельев А.И., Лебедева В.В., Лебедев И.В., Камынин К.В., Кузнецов Л.Д., Ронжин А.Л. Управление группой БПЛА при отработке кризисных полетных ситуаций в решении транспортных задач. *Известия ЮФУ. Технические науки*. 2022;1(225):110–120. <https://doi.org/10.18522/2311-3103-2022-1-110-120>
6. Родионов А.С., Маткурбанов Т.А. Планирование траектории полета БПЛА при мониторинге большой области. *Информатика и автоматизация*. 2025;24(3):791–827. <https://doi.org/10.15622/ia.24.3.3>
7. Кулешов С.В., Кваснов А.В., Зайцева А.А., Ронжин А.Л. Комплексный подход к визуальной навигации по естественным ориентирам для БПЛА, работающих в условиях недоступности ГНСС. *Известия ЮФУ. Технические науки*. 2025;2(244):269–278. <https://doi.org/10.18522/2311-3103-2025-2-269-278>
8. Муслимов Т.З. Предотвращение столкновений при круговом движении группы дронов самолетного типа на основе вращательной модификации искусственного потенциального поля. *Информатика и автоматизация*. 2025;24(1):72–98. <https://doi.org/10.15622/ia.24.1.4>
9. Беляков М.Э., Диане С. Алгоритмы визуального анализа внешней среды автономным мобильным роботом в задаче уборки территории. *Russian Technological Journal*. 2023;11(4):26–35. <https://doi.org/10.32362/2500-316X-2023-11-4-26-35>
10. Антипов А.С., Кокуньо Ю.Г., Вольф Д.А., Широков А.С. Синтез комбинированного управления ориентацией и высотой квадрокоптера на основе блочного подхода с сигмовидными обратными связями. *Информатика и автоматизация*. 2025;24(3):745–790. <https://doi.org/10.15622/ia.24.3.2>

11. Голубов В.В., Манько С.В. Автоматизация стыковки автономных мобильных роботов на основе развития метода поисковых случайных деревьев со встречным ростом. *Russian Technological Journal*. 2024;12(1):7–14. <https://doi.org/10.32362/2500-316X-2024-12-1-7-14>
12. Воронков А.Д., Диане С.А. Непрерывный генетический алгоритм в задаче захвата манипуляционным роботом объекта априорно неизвестной формы. *Russian Technological Journal*. 2023;11(1):18–30. <https://doi.org/10.32362/2500-316X-2023-11-1-18-30>
13. Нго К.Т., Соленая О.Я., Ронжин А.Л. Анализ подвижных роботизированных платформ для обслуживания аккумуляторов беспилотных летательных аппаратов. *Труды МАИ*. 2017;95:11. <https://elibrary.ru/zhnxylh>
14. Давлюд И.И. Расчет основных эксплуатационных характеристик привязной высотной системы корабельного базирования. *Russian Technological Journal*. 2024;12(4):40–50. <https://doi.org/10.32362/2500-316X-2024-12-4-40-50>
15. Манько С.В., Лохин В.М., Диане С.К. Принципы построения и экспериментальные исследования прототипного образца многоагентной робототехнической системы для разбора завалов. *Russian Technological Journal*. 2022;10(6):28–41. <https://doi.org/10.32362/2500-316X-2022-10-6-28-41>
16. Ле В.Н., Ронжин А.Л. Обзор интеллектуальных систем управления и робототехнических задач в производстве аквакультуры. *Морские интеллектуальные технологии*. 2024;1-1(63):171–180. <https://doi.org/10.37220/MIT.2024.63.1.021>
17. Архипов И.С., Ерофеева В.А., Граничин О.Н., Киселев В.А., Чернов А.О. Ускоренный протокол локального голосования для группы роботов с одним пультом дистанционного управления. *Информатика и автоматизация*. 2025;24(4):1029–1058. <https://doi.org/10.15622/ia.24.4.2>
18. Saveliev A., Anikin D., Ronzhin A., Erokhin G., Agafonov V. System for Planing Seismic Sensors Based on Actions of UAVs Group with Optimized Flight Plan. In: Ronzhin A., Savage J., Meshcheryakov R. (Eds.). *Interactive Collaborative Robotics*. Lecture Notes in Computer Science. Cham: Springer; 2024. V. 14898. P. 334–343. https://doi.org/10.1007/978-3-031-71360-6_25
19. Гончаренко В.А., Хомоненко А.Д., Абу Хасан Р. Композиционный подход к имитационному моделированию систем массового обслуживания со случайными параметрами. *Информатика и автоматизация*. 2024;23(6):1577–1608. <https://doi.org/10.15622/ia.23.6.1>
20. Ронжин А.Л., Дашевский В.П., Крестовников К.Д. Взаимодействие воздушных дронов с наземными роботизированными средствами. В сб.: *XVIII Всероссийская мультиконференция по проблемам управления (МКПУ – 2025): Материалы мультиконференции*. Т. 1. *Робототехника и мехатроника (РиМ – 2025)*. Тула: Изд-во ТулГУ; 2025. С. 145–148. <https://elibrary.ru/ibktbu>
21. Галыкин Ю.В., Дашевский В.П. Автоматизация производства приборных корпусов с применением специализированного станка с ЧПУ для изготовления крепежных брусков. *Вестник МГТУ «Станкин»*. 2024;1(68):119–128. <https://elibrary.ru/mcsumqz>
22. Поскрёбышев В.А., Зиновьев А.А., Лохова Н.А., Исько А.Б., Белых С.А. *Механическое оборудование для производства строительных материалов и изделий*. Братск: ГОУ ВПО «БрГУ»; 2009, 378 с. ISBN 978-5-8166-0261-7

About the Authors

Vladimir P. Dashevsky, Cand. Sci. (Eng.), Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences (39, Vasil'evskii ostrov, 14-ya liniya, St. Petersburg, 199178 Russia). E-mail: dashevskiy.v@ias.spb.su. Scopus Author ID 6603052314, RSCI SPIN-code 7675-2463, <https://orcid.org/0000-0002-2285-1936>

Maxim M. Bizin, Leading Engineer, St. Petersburg Federal Research Center of the Russian Academy of Sciences (39, Vasil'evskii ostrov, 14-ya liniya, St. Petersburg, 199178 Russia). E-mail: bizin@ias.spb.su. RSCI SPIN-code 5499-4214. <https://orcid.org/0009-0006-2314-9715>

Yuri V. Galykin, Junior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences (39, Vasil'evskii ostrov, 14-ya liniya, St. Petersburg, 199178 Russia). E-mail: yuri.galykin@strategic-it.ru. <https://orcid.org/0009-0000-1274-9754>

Об авторах

Дашевский Владимир Павлович, к.т.н., старший научный сотрудник, ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (199178, Россия, Санкт-Петербург, 14-я линия Васильевского острова, д. 39). E-mail: dashevskiy.v@ias.spb.su. Scopus Author ID 6603052314, SPIN-код РИНЦ 7675-2463, <https://orcid.org/0000-0002-2285-1936>

Бизин Максим Михайлович, ведущий инженер, ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (199178, Россия, Санкт-Петербург, 14-я линия Васильевского острова, д. 39). E-mail: bizin@ias.spb.su. SPIN-код РИНЦ 5499-4214, <https://orcid.org/0009-0006-2314-9715>

Галыкин Юрий Валерьевич, младший научный сотрудник, ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (199178, Россия, Санкт-Петербург, 14-я линия Васильевского острова, д. 39). E-mail: yuri.galykin@strategic-it.ru. <https://orcid.org/0009-0000-1274-9754>

Translated from Russian into English by K. Nazarov

Edited for English language and spelling by Thomas A. Beavitt

Modern radio engineering and telecommunication systems
Современные радиотехнические и телекоммуникационные системы

UDC 621.396.96

<https://doi.org/10.32362/2500-316X-2026-14-4-68-81>

EDN IZWLXU



RESEARCH ARTICLE

Synthesis of matched multichannel processing of orthogonal MIMO radar signals

Boris M. Vovshin^{1, 2, @}, **Alexander A. Pushkov**¹¹ NPO Almaz, Moscow, 127411 Russia² MIREA – Russian Technological University, Moscow, 119454 Russia@ Corresponding author, e-mail: boris@eleron.net

• Submitted: 31.07.2024 • Revised: 03.02.2026 • Accepted: 01.06.2026

Abstract

Objectives. In contemporary radar studies, increasing attention is being paid to MIMO (Multiple-Input, Multiple-Output) radars, which are predicted to offer significant advantages over traditional phased-array radars. MIMO radars increase spatial degrees of freedom by simultaneously transmitting coherent, orthogonal radio waveforms in parallel. The main challenge for MIMO radars is to optimize space-time processing (STP) of the entire set of echo signals against a white Gaussian noise background according to the Neyman–Pearson criterion. In order to ensure optimal detection performance, the work statistically synthesizes STP in MIMO radars to derive variants of STP structural schemes for identifying their features as compared to STP in traditional phased-array radars, and analyze the angular resolution of the MIMO radar with the resulting STP structure.

Methods. Statistical optimization methods for signal detection against interference using the Neyman–Pearson criterion were used in accordance with approaches informed by the theory of beamforming in multibeam antenna arrays.

Results. Two STP variants were synthesized to generate optimal pre-threshold statistics for a MIMO radar detector with an M -element transmitting and N -element receiving antenna array. The structural diagrams of these variants were found to differ in their sequences of spatial and temporal processing. Due to the increased size of the equivalent aperture of the receiving antenna array, the angular resolution of the MIMO radar is demonstrated to be significantly better than that of a traditional radar at the same signal-to-noise ratio.

Conclusions. The structural diagrams of the synthesized MIMO radar detectors confirm their statistically optimal detection performance.

Keywords: MIMO radar, multichannel space-time processing, fast Fourier transform algorithm, multi-beam antenna array, detector pre-threshold statistics, angular resolution

For citation: Vovshin B.M., Pushkov A.A. Synthesis of matched multichannel processing of orthogonal MIMO radar signals. *Russian Technological Journal*. 2026;14(4):68–81. <https://doi.org/10.32362/2500-316X-2026-14-4-68-81>, <https://elibrary.ru/IZWLXU>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Синтез согласованной многоканальной обработки ортогональных сигналов ММО радиолокационных станций

Б.М. Вовшин^{1, 2, @}, А.А. Пушков¹

¹ ПАО «НПО «Алмаз», Москва, 127411 Россия

² МИРЭА – Российский технологический университет, Москва, 119454 Россия

@ Автор для переписки, e-mail: boris@eleron.net

• Поступила: 31.07.2024 • Доработана: 03.02.2026 • Принята к опубликованию: 01.06.2026

Резюме

Цели. В современной теории радиолокации все большее внимание уделяется ММО (англ. «много входов – много выходов») радиолокационных станций (РЛС), для которых прогнозируются важные преимущества перед традиционными РЛС с фазированными антенными решетками (ФАР). ММО РЛС имеют дополнительную степень свободы за счет одномоментного излучения системы когерентных ортогональных радиосигналов, что однозначно определяет параллельный обзор пространства. Для ММО РЛС возникает актуальная задача синтеза оптимальной по критерию Неймана – Пирсона пространственно-временной обработки (ПВО) всей совокупности эхо-сигналов на фоне белого гауссова шума. Целями данной статьи являются: статистический синтез ПВО в ММО РЛС, обеспечивающий оптимальные показатели качества обнаружения; получение вариантов структурных схем ПВО и выявление их особенностей по сравнению с ПВО традиционных РЛС с ФАР; анализ углового разрешения ММО РЛС с полученной структурой ПВО.

Методы. Используются методы статистической оптимизации процедуры обнаружения сигналов на фоне помех по критерию Неймана – Пирсона, а также методы теории диаграммообразования в многолучевых антенных решетках (АР).

Результаты. Синтезированы два варианта ПВО, формирующие оптимальную предпороговую статистику обнаружителя ММО РЛС с M -элементной передающей и N -элементной приемной АР. Показано, что структурные схемы этих вариантов различаются последовательностями пространственной и временной обработки. Доказано, что за счет увеличенного размера «эквивалентной» апертуры приемной АР разрешающая способность ММО РЛС по угловым координатам существенно лучше, чем у традиционной РЛС при одинаковом отношении сигнал/шум.

Выводы. Показано, что структурные схемы синтезированных обнаружителей ММО РЛС обеспечивают статистически оптимальные показатели качества обнаружения.

Ключевые слова: ММО РЛС, многоканальная пространственно-временная обработка, алгоритм быстрого преобразования Фурье, многолучевая антенная решетка, предпороговая статистика обнаружителя, разрешающая способность по угловым координатам

Для цитирования: Вовшин Б.М., Пушков А.А. Синтез согласованной многоканальной обработки ортогональных сигналов ММО радиолокационных станций. *Russian Technological Journal*. 2026;14(4):68–81. <https://doi.org/10.32362/2500-316X-2026-14-4-68-81>, <https://elibrary.ru/IZWLXU>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

We consider a multiple-input, multiple-output (MIMO) radar system based on an equidistant antenna array (AA) configuration with M -element transmitters and N -element receivers. The distances between the antennas are denoted as d_t and d_r (index t for transmit and r for receive), respectively. According to MIMO principles, the AA transmits M component vector of orthogonal coherent signals, $\mathbf{U}(t) = \{U_m(t)\}_{m=1}^M$, with

partial complex amplitudes, $U_m(t)$, where t is time [1, 2]. As demonstrated in [3, 4], this approach defines a unique, parallel space survey in the $\Delta\theta_{svy}$ sector, which is consistent with the directional pattern (DP) beamwidth of the AA. The system operates in target acquisition mode (Fig. 1).

The special case of using a single weakly directional antenna (WDA) with $N = 1$ for reception, which is transformed by aperture synthesis into a virtual subarray (VSA), is considered in [5]. By synthesizing a separable STP, this study demonstrates the equivalence of target detection results against uncorrelated white Gaussian noise (WGN) backgrounds in MIMO radar with parallel space

surveys in WDA reception and AA transmission, as well as in traditional radar with WDA transmission and AA reception.

The results obtained in [5] are generalized in the present work to a multichannel ($N > 1$) MIMO structure. This involves radar reception without altering the transmission AA structure or the parameters of the ranging signals. Here, while the use of MIMO radar can deliver a significant improvement in performance, no equivalence should be expected in the results of observing targets in the compared variants; rather, the improvement primarily applies to the angular coordinate resolution at a fixed energy potential. In the compared radars, STP synthesis is employed to detect a point target with an a priori unknown three-dimensional vector of informative parameters: $\mathbf{a} = \{R_{TRG}, V_{rTRG}, \theta_{TRG}\}$. Here, $R_{TRG} = c\tau_{TRG}/2$ is the range capacity; V_{rTRG} is the radial velocity; θ_{TRG} is the angular coordinate of the target; c is the speed of light; τ_{TRG} is the time delay of the echo signal.

The predicted advantages of the MIMO radar systems over conventional radars are primarily due to the additional degrees of freedom, by which means the dimensionality of the spatial coordinate space is increased to improve the accuracy and precision of the system. In conventional radars, the transmitting WDA (or AA) has

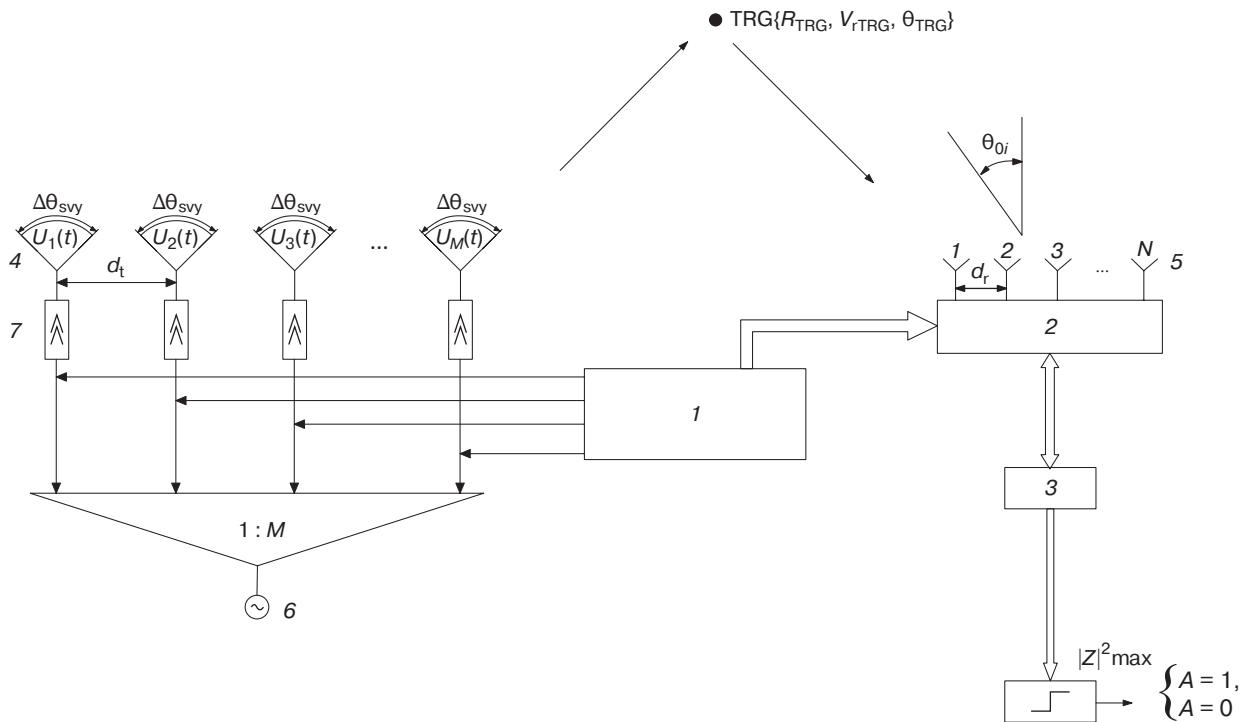


Fig. 1. Structure of MIMO radar with parallel space survey:

(1) orthogonal ranging signal generator; (2) space-time processing (STP); (3) pre-threshold statistics generator; (4) transmitting antenna array (AA); (5) receiving AA; (6) driving oscillator; and (7) power amplifier.

$\text{TRG}\{R_{TRG}, V_{rTRG}, \theta_{TRG}\}$ is the target; A is the threshold value used to make a decision regarding the presence or absence of a signal; $|Z|^2_{\max}$ represents the maximum pre-threshold statistics value

a single phase center for emitting a coherent ranging signal, $\mathbf{u}_0(t)$. In MIMO radars, conversely, each partial signal with a complex amplitude $\dot{U}_m(t)$, $m \in 1, M$ has a precisely known phase center with coordinates (x_m, y_m) , resulting in a phase “coloring” of the signals, which can lead to improved performance. When combined with an N -element reception AA, this allows for a significant improvement in angular resolution as compared with an N -element conventional radar system having WDA transmission and AA reception, which is inversely proportional to the product $M \times N$.

In order to determine the resolution of the radar in terms of angular coordinates, the present work sets out to synthesize a consistent multichannel signal processing system for detecting and analyzing spatiotemporal signals reflected from targets and received by MIMO radars. The system is compared with conventional radar detectors against the WGN background.

When solving this problem, we kept the framework of the narrow-band system assumption for ranging signals to ensure orthogonality in either the time or frequency domain [6, 7]. In the first case, all partial signals may be generated using orthogonal codes for the initial phases of discrete signals τ_{dscr} at the carrier frequency f_0 ($1/\tau_{\text{dscr}} \ll f_0$) analogous to a phase code modulated within transmitting AA channels. In the second case, orthogonality may be ensured by separating carrier frequencies by Δf between adjacent ranging signals within the transmitting AA. This may be considered as analogous to a stepped-frequency linear frequency modulated (SF-LFM) signal for which $(M-1)\Delta f \ll f_0$. Assuming a narrow-band probe signal system simplifies the process, as it allows for the following:

- Separating the processing into spatial and temporal components (non-separable STP for broadband signals is considered in [7]).
- Assuming that the MIMO radar operates at a specific wavelength of $\lambda_0 = c/f_0$, and thus, the Doppler frequency shift $F_D = 2V_{\text{rTRG}}/\lambda_0$ is approximately the same for all components of the vector signal.
- Calculating the delay time of the ranging signal at the aperture of the transmitting AA:

$$T_{\text{ant}} = \frac{(M-1)d_t}{c} \ll \tau_{\text{eq}},$$

where τ_{eq} is the equivalent pulse duration corresponding to the bandwidth $\Delta f = 1/\tau_{\text{eq}}$ (for the receiving AA of the MIMO radar).

In addition, we assume that there are no reflections or sources of external radiation within the survey sector $\pm\theta_{\text{svy}}/2$. Under these conditions, we can consider the radiators of the receiving AA to be excited by an M -dimensional vector signal $\mathbf{U}_{\text{TRG}}$, which is reflected from a point target:

$$\mathbf{U}_{\text{TRG}}(R_{\text{TRG}}, V_{\text{rTRG}}, \theta_{\text{TRG}}) = \left\{ U_m(\tau_{\text{TRG}}, V_{\text{rTRG}}) \exp \left(j \frac{2\pi d_t}{\lambda_0} (m-1) \sin \theta_{\text{TRG}} \right) \right\}_{m=1}^M. \quad (1)$$

It is further assumed that the components of the vector $\mathbf{a}$ are a priori unknown. Following excitation of the N -element receiving AA, the M -dimensional vector (1) transforms into an $N \times M$ vector of the echo signal. This can be expressed as a Kronecker product, as follows:

$$\mathbf{U}_r(\tau_{\text{TRG}}, t', \theta_{\text{TRG}}) = \mathbf{\beta}^H(\theta_{\text{TRG}}) \otimes \mathbf{U}_{\text{TRG}}(t, \mathbf{a}), \quad (2)$$

$$\text{where } \mathbf{\beta}^H(\theta_{\text{TRG}}) = \left\{ \exp \left[j \frac{2\pi d_t}{\lambda_0} (m-1) \sin \theta_{\text{TRG}} \right] \right\}_{n=1}^N$$

is the spatial phase progression vector between the receiving AA radiators, while H is the sign of Hermitian conjugacy (transposition, T , and complex conjugation, $*$).

Based on the accepted conditions and assumptions, we consider the conventional statistical approach to the STP synthesis, which is optimal according to the Neumann–Pearson criterion.

STATISTICAL APPROACH TO DETERMINING THE WEIGHT VECTOR OF OPTIMAL PROCESSING IN MIMO RADAR

The problem of synthesizing multichannel STP at the output of the MIMO radar’s receiving AA is considered against the background of the δ -correlated WGN using the conventional formulation based on the likelihood ratio [8, 9]. The sampled input $\mathbf{y}$ is represented by the N -dimensional vector $\mathbf{u}$ of the interference (noise) mixture $\mathbf{p}$ and possibly useful signal ($A = 1$) as follows:

$$\mathbf{y} = A\mathbf{u} + \mathbf{p}, \quad \mathbf{y} = \{y_i\}_{i=1}^N, \quad A = [1, 0]. \quad (3)$$

According to the well-established results from the theory of optimal detection based on the Neumann–Pearson criterion [10], the multichannel processing weight vector can be expressed as follows:

$$\mathbf{r} = \mathbf{\Phi}^{-1} \mathbf{u}^*(\alpha_0), \quad \mathbf{\Phi} = \left\{ \varphi_{ij} \right\}_{i,j}^N = \overline{\mathbf{p}\mathbf{p}^H}, \quad (4)$$

where $\mathbf{\Phi}^{-1}$ is an $N \times N$ matrix that is the inverse correlation matrix (CM) of the interference $\mathbf{\Phi}$, while the overbar denotes statistical averaging.

In Eq. (4), the N -dimensional vector $\mathbf{u}(\alpha_0)$ is a signal with an expected three-dimensional vector of informative parameters, $\alpha_0 = \{\tau_0, V_{r0}, \theta_0\}$. The CM $\mathbf{\Phi}$ with $\varphi_{ij}(t, s) = \overline{\varphi_i(t) \varphi_j^*(s)}$ elements exhibits cross-correlation between the i th and j th reception channels at times t and s . For the existence of an inverse matrix, $\mathbf{\Phi}^{-1}$, the CM should be Hermitian and positive-definite [11].

The STP synthesis for the given problem involves searching for the vector $\mathbf{r}$ in (4) that provides sufficient pre-threshold statistics in the form of a weighted scalar sum [12] as follows:

$$\xi = \mathbf{y}^T \boldsymbol{\Phi}^{-1} \mathbf{u}(\mathbf{a}_0) = \sum_{i=1}^N y_i r_i, \quad (5)$$

where the energy detection parameter is the following:

$$q = [\mathbf{u}^H(\mathbf{a}_0) \boldsymbol{\Phi}^{-1} \mathbf{u}(\mathbf{a}_0)]^{1/2}. \quad (6)$$

Provided the processing is linear, such pre-threshold statistics can be used to optimize the detection quality indicators by enabling the use of known detection characteristics [8, 9].

The method for calculating the random variable ξ in Eq. (5) determines the structural design of optimal processing consistent against the WGN background.

Due to the narrow bandwidth, the parameters of the signal reflected from a target with coordinates $\mathbf{a}_{\text{TRG}} = \{\tau_0, V_{\text{TRG}}, \theta_{\text{TRG}}\}$ can be separated into time-frequency and spatial components:

$$\mathbf{u}(t, \mathbf{a}_{\text{TRG}}) = \left\{ u_m(t, \tau_{\text{TRG}}, V_{\text{TRG}}) e^{j\varphi_m(\theta_{\text{TRG}})} \right\}_{m=1}^M, \quad (7)$$

where $\varphi_m(\theta_{\text{TRG}}) = \frac{2\pi}{\lambda} d_t(m-1) \sin \theta_{\text{TRG}}$ is the spatial phase progression of the m th component of the signal.

Thus, the processing of the components can be divided into time-frequency and spatial sequences.

We define the types of CM $\boldsymbol{\Phi}$ included in Eqs. (4) and (5) and note its characteristics for STP against WGN in the MIMO radar of the considered design (Fig. 1). It would be reasonable to assume that the noise in the N -element receiving AA channels is independent $\mathbf{p}\mathbf{p}_j^H = 0$ at $i \neq j$, $i, j = 1, N$, while it is δ -correlated in time. For an arbitrary i th channel having a spectral noise power density of P_{0i} , CM has the following form:

$$\boldsymbol{\Phi}_m(t, s) = P_{0i} \mathbf{I} \delta(t - s), \quad (8)$$

where $\mathbf{I} = \text{diag}\{\varphi_{ii}\}_{i=1}^N$ is the identity matrix.

Provided that the orthogonal components $\mathbf{u}(\mathbf{a}_{\text{TRG}})$ pass through the corresponding matched filters (MF), it is additionally possible to assume the noise in time-frequency channels is independent and δ -correlated [13].

To facilitate the CM description, we assign a double index $y_{(t)} = \{y_n^m\}_{n,m \neq 1}^{N,M}$ to the input samples. Here, n represents the channel index of the receiving AA, while m represents the index of signal components corresponding to the channel index in the transmitting AA.

Given the aforementioned assumptions and considering the filtering properties of the δ -function, the CM for the stationary WGN will be as follows:

$$\begin{aligned} \boldsymbol{\Phi}(t, s) &= \delta(t - s) \overline{\mathbf{y}(t) \mathbf{y}^*(s)} = \left\{ y_n^m(t) y_p^q(t) \right\} = \\ &= \text{diag} \left\{ y_{n,n}^{m,m} \right\}_{n,m=1}^{N,M} = \boldsymbol{\Phi}. \end{aligned} \quad (9)$$

In (9) it is assumed that the CM elements are zero for $n \neq p$ and $m \neq q$.

The sequence of stages in the STP process is determined by the order in which the $N \times M$ diagonal elements are grouped by indices, resulting in two variants of the $\boldsymbol{\Phi}$ matrix [14].

Variant 1. CM $\boldsymbol{\Phi}$ is represented by an $N \times M$ diagonal block matrix

$$\boldsymbol{\Phi}_1 = \begin{pmatrix} \mathbf{B}_{01}^{(m)} & & & \\ & \mathbf{B}_{02}^{(m)} & & \\ & & \ddots & \\ & & & \mathbf{B}_{0n}^{(m)} \\ & & & & \ddots \\ & & & & & \mathbf{B}_{0N}^{(m)} \end{pmatrix}. \quad (10)$$

Here, each of the N blocks is also represented by the diagonal CM noise between the time-frequency channels generated by the M MFs for the m th signal component in the n th channel of the receiving AA

$$\mathbf{B}_{0n}^{(m)} = \text{diag} \left\{ p_{0n}^{(m)} \right\}_{m=1}^M, \quad \mathbf{B}_{0n} = p_{0n} \mathbf{I},$$

$$\text{at } p_{0n}^{(1)} = p_{0n}^{(2)} = \dots = p_{0n}^{(M)} = p_{0n}. \quad (11)$$

Variant 2. CM $\boldsymbol{\Phi}$ is represented by an $M \times N$ -diagonal block matrix

$$\boldsymbol{\Phi}_2 = \begin{pmatrix} \mathbf{B}_{01}^{(n)} & & & \\ & \mathbf{B}_{02}^{(n)} & & \\ & & \ddots & \\ & & & \mathbf{B}_{0m}^{(n)} \\ & & & & \ddots \\ & & & & & \mathbf{B}_{0M}^{(n)} \end{pmatrix}. \quad (12)$$

In this representation, each of the M CM blocks is the diagonal CM of noise between the N spatial AA channels for one m th signal component

$$\mathbf{B}_{0m}^{(n)} = \text{diag} \left\{ p_{0m}^{(n)} \right\}_{n=1}^N, \quad \mathbf{B}_{0m} = p_{0m} \mathbf{I},$$

$$\text{at } p_{0m}^{(1)} = p_{0m}^{(2)} = \dots = p_{0m}^{(M)} = p_{0m}. \quad (13)$$

As follows from (10) and (12), the statistical approach to synthesizing STP for MIMO radars differs from the conventional methods involving CM dimensionality of interference $M \times N$ (instead of an N -dimensional vector), as well as the ability to combine time-frequency and spatial processing in any order.

ANALYSING SUFFICIENT STATISTICS

For a sufficient detector of pre-threshold statistics, Eq. (5) can be represented in analogue form using the correlation integral of the complex envelopes of the received mixture and the expected signal $Z = \int \mathbf{Y}^H(t, \mathbf{a}_{\text{TRG}}) \dot{\mathbf{U}}(t, \mathbf{a}_0) dt$. To ensure reliable detection characteristics, it is assumed that the echo signal has a random Rayleigh amplitude and a uniform initial phase ranging from 0 to 2π . Taking into account (11) and (12), the sufficient pre-threshold statistics against WGN background can then be written in terms of the squared modulus of the correlation integral [10, 12] as follows:

$$\xi = |Z|^2 = \sum_{n=1}^N \sum_{m=1}^M \frac{1}{P_{0n,m}^{(m)(n)}} \left| \int_{-\infty}^{\infty} \dot{U}_m(t, \tau_{\text{TRG}}, \theta_{\text{TRG}}) \dot{U}_m(t, \tau_0, \theta_0) \exp[j2\pi(f_0 + \Delta f_{Dm})(t - t_{0m})] dt \right|^2 \times \exp \left[j \frac{2\pi d_t(m-1)}{\lambda} (\sin \theta_{\text{TRG}} - \sin \theta_0) \right] \exp \left[j \frac{2\pi d_r(n-1)}{\lambda} (\sin \theta_{\text{TRG}} - \sin \theta_0) \right], \quad (14)$$

where Δf_{Dm} is the Doppler frequency shift for the m th signal component; t_{0m} is the constant delay in the temporal accumulation filter for the m th signal component; $\tau_{\text{TRG}} = 2R_{\text{TRG}}/c$ is the time delay for the expected echo signal.

Equation (14) can be considered as a four-dimensional ambiguity function in range–speed–angle coordinates. In this context, the double summation can be reduced to the product of the transmitting and receiving AA multipliers, which is a characteristic feature of MIMO radars [15]. Here, the DP refers to the cross section of the ambiguity function when the time-frequency parameters τ_0 and V_{r0} are fixed.

For simplicity, it is assumed that the Doppler frequency shift will be approximately the same for all signal components due to their narrow bandwidth ($\Delta f_{Dm} \approx \Delta f_D$), and that the time delay in all accumulation filters is also equal ($t_{0m} = t_0$). In this context, the MIMO radar is considered to be in target detection mode, meaning that the only value of interest is the maximum value ξ applied to the threshold. It is also assumed that temporal accumulation is achieved through M MFs whose maximum response is independent of the signal delay, $\tau_{\text{TRG}} = 2R_{\text{TRG}}/c$. At time t_0 , the maximum value of the integral in Eq. (14) is reached to obtain the optimal SNR at the MF output.

In this case, the maximum MF response value for the m th component of the signal is proportional to its energy, E_m :

$$\dot{U}_{m \text{ mf}}(0) = \int_{-\infty}^{\infty} |U_m(t, 0, 0)|^2 dt = kE_m, \quad (15)$$

where k is the proportionality coefficient.

Regarding Eq. (14), the concept of a DP is meaningful in terms of the instantaneous cross-section of the ambiguity function at the instantaneous time $t_0 = 0$, i.e., when a maximum appears at the MF output.

Equation (14) shows that the calculation method of pre-threshold statistics determines the structure of the detector, which depends on the order of summation by the “ m ” and “ n ” indices, i.e., the sequence of temporal and spatial accumulation. In the first case, temporal accumulation occurs before spatial accumulation, while in the second case, the order is reversed. In this case, the CM of noise corresponds to Eqs. (11) and (12). Taking the above comments and simplifications into account, we can now consider variants for constructing STP structural schemes.

STRUCTURAL DIAGRAMS OF MIMO RADAR DETECTORS

Variant 1 (temporal + spatial accumulation)

In this STP variant, formula (14) first performs coherent summation by the “ m ” index, followed by summation by the “ n ” index.

Rewriting formula (14) in terms of the condition $p_{0n}^{(m)} = p_{0n}$ gives the following form:

$$\xi = \left| \left\{ \sum_{m=1}^M \dot{U}_{m \text{ mf}}(0) \exp \left[j \frac{2\pi d_t(m-1)}{\lambda} (\sin \theta_{\text{TRG}} - \sin \theta_0) \right] \right\} \sum_{n=1}^N \frac{A_n}{P_{0n}} \exp \left[j \frac{2\pi d_r(n-1)}{\lambda} (\sin \theta_{\text{TRG}} - \sin \theta_0) \right] \right|^2. \quad (16)$$

The processing can then be divided into two consecutive stages.

Stage 1.1

In Eq. (16), the expression within the curly brackets represents the summation of the results from the temporal accumulation of M signal components that compensates for their spatial phase in the expected direction θ_0 . The second sum is the product of the N -element receiving AA with a d_t step, where the amplitude distribution, $\{A_n\}_{n=1}^N$, is normalized to the noise power in the radiator channels.

It should be noted that the sum over the index “ m ” corresponds formally to the multiplier of an M -element AA with a d_t step, where the maxima of MF responses to the signal components $A_m = U_{m\text{ mf}}(0)$ act as the amplitude distribution. The N weakly directional elements of the receiving AA are transformed into a VSA [5, 16]. The DP $F_v(\theta_n, \theta_0)$ for this VSA is determined by the parameters of the transmitting AA. For $V_{m\text{ mf}}(0) = \text{const}$, we obtain the following:

$$F_v(\theta_{\text{TRG}}, \theta_0) = \text{sinc} \left[\frac{\pi L_{\text{an.t}}}{\lambda_0} (\sin \theta_{\text{TRG}} - \sin \theta_0) \right], \quad (17)$$

where $\text{sinc } x = \sin x/x$, $L_{\text{an.t}} = (M-1)d_t$ is the length of the transmitting AA.

If the transmitting AA step is chosen from the unambiguous scanning condition within the angular survey sector, $\Delta\theta_{\text{svy}}$, then it is limited by the $d_t \leq \lambda_0[1 + |\sin \theta_0|]^{-1}$ ratio. As a result, diffraction lobes will not occur within the VSA DP $F_v(\theta_n, \theta_0)$.

Stage 1.2

The spatial accumulation stage in the receiving AA virtually consists of N subarrays, each having a DP in the form of Eq. (17).

If the receiving AA step is the same length as the transmitting AA step, $d_r = L_{\text{an.t}}$, then its equivalent length can be assumed to be $L_{\text{eq.r}} = (N-1)L_{\text{an.t}} = (N-1)(M-1)d_t$. In this case, the “equivalent” receiving AA is completely filled with $N \times M$ elements with a d_t step. Taking into account (17), the DP will take the following form:

$$\begin{aligned} F_r(\theta_{\text{TRG}}, \theta_0) &= F_{\text{el}}(\theta_{\text{TRG}}) F_v(\theta_{\text{TRG}}, \theta_0) \sum_{n=1}^N \frac{1}{P_{0n}} \left[j2\pi \frac{d_r}{\lambda_0} (n-1)(\sin \theta_{\text{TRG}} - \sin \theta_0) \right] \approx \\ &\approx F_v(\theta_{\text{TRG}}, \theta_0) \sum_{n=1}^N \frac{1}{P_{0n}} \left[j2\pi \frac{d_t}{\lambda} (M-1)(n-1)(\sin \theta_{\text{TRG}} - \sin \theta_0) \right]. \end{aligned} \quad (18)$$

Here no diffraction maxima are present. The second equality takes into account the isotropic nature of the elementary radiator DPs, $F_{\text{el}}(\theta) = \text{const}$.

For a MIMO radar in target search mode, it is essential to perform a parallel space survey, since the components of the ranging signal simultaneously “illuminate” the entire survey area. Here, due to the lack of a priori information about the expected target direction, multipath beamforming is required in order to prevent energy loss in the parallel survey.

To achieve this, the “equivalent” receiving AA should form $M \times N$ orthogonal beams (reception channels) in the following sequence. First, a fan of M beams in all N VSAs should be generated to create “secondary” reception channels (“primary” means N radiator channels here). Then, “tertiary” channels corresponding to a fan of N sharply directed beams should be generated within each of the M secondary beams.

The secondary and tertiary beams can be generated using beamforming circuit in the form of analogue Butler matrix [17] or M - and N -point fast Fourier transform (FFT) according to the assumption that $M = 2^p$ and $N = 2^q$ ($p, q = 2, 3, 4, \dots$) [18].

To generate the secondary channels at the VSA output for the i th DP, which are oriented in the direction $\theta_{0i} = \arcsin \left[\left(i - \frac{1}{2} \right) \frac{\lambda_0}{(M-1)d_t} \right]$, the transformation matrix $\mathbf{W}(\theta_{0i}) = \left\{ \exp \left(-j \frac{2\pi i m}{M} \right) \right\}_{i=-M/2}^{M/2}$ can be used. If vector $\mathbf{V}_{0\text{ mf}}(0)$ represents digitized samples at time $t = 0$, then the following response is obtained for an arbitrary direction towards the target in the i th secondary channel:

$$F_v(0, \theta_{\text{TRG}}, \theta_{0i}) = \sum_{m=1}^{M/2} U_{m\text{ mf}}(0) \cos \left[\frac{2\pi d_t}{\lambda_0} \left(m - \frac{1}{2} \right) \left(\sin \theta_{\text{TRG}} - \frac{\left(i - \frac{1}{2} \right) \lambda_0}{L_{\text{an.t}}} \right) \right]. \quad (19)$$

As a result, the same M -beam amplitude DPs are generated at the output of the N VSAs. However, the beams with the same index will differ in terms of phase progression, which is determined by the d_r step of the receiving AA.

In order to generate tertiary beams, the secondary beam signals having the same index (i) must be re-indexed according to the necessary compensation of phase progression. Since $L_{eq,r} = (N - 1)L_{an,r}$, the width of the tertiary beams is N times narrower than that of secondary beams. The relationship between the primary, secondary, and tertiary beams is clearly illustrated in Fig. 2.

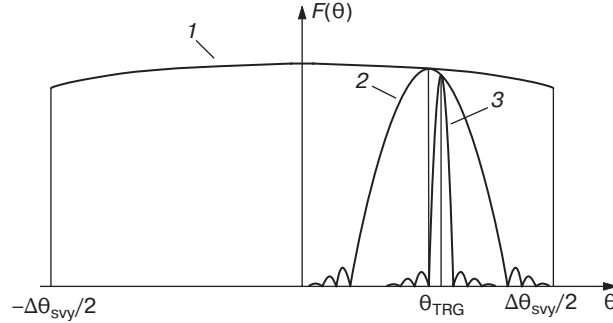


Fig. 2. DP of receiving channels:

(1) DP of the primary receiving radiator; (2) DP of the i th secondary channel at the VSA output; and (3) DP of the s th tertiary channel at the output of the i th secondary channel

The generation of tertiary beams is also facilitated by the FFT matrix $\mathbf{W}(\theta_{0s}) = \left\{ \exp\left(-j \frac{2\pi s n}{N}\right) \right\}_{n=-N/2}^{N/2}$. Beams with orientation $\theta_{0s} = \arcsin\left[\left(s - \frac{1}{2}\right)\lambda_0 / (N-1)d_r\right]$ are implemented in each i th secondary reception channel ($i \in \overline{1, M}$). For this reason, it is convenient to assign a double indexing of θ_{0is} to the phasing angles of the tertiary beams, amounting to $M \times N$, i.e., the s th ray in the i th direction in the multipath VSA DP. Consequently, a fan of tertiary beams is generated from the instantaneous (at $t = 0$) values of the envelope signals sampled from the $N \times M$ outputs of the VSA.

Consequently, the receiving AA has an $N \times M$ beam DP for the receiving AA consisting of N VSAs with M -ray partial DPs (19). The dependence of the response of the tertiary channel “ is ” for a signal from an arbitrary direction θ_{TRG} takes the following form:

$$F_r(0, \theta_{TRG}, \theta_{0is}) = F_{el}(\theta_{TRG}) F_v(0, \theta_{TRG}, \theta_{0is}) \sum_{n=1}^{N/2} \frac{A_n}{P_{0n}} \cos \left\{ \frac{\pi(M-1)d_t}{\lambda_0} \left(n - \frac{1}{2} \right) \left(\sin \theta_{TRG} - \frac{\left(s - \frac{1}{2}\right)\lambda_0}{L_{an,r}} \right) \right\}. \quad (20)$$

This formula can be interpreted as the tertiary DP of the spatial channel “ is ” or as the cross-section of the ambiguity function at time $t = 0$ in the “zero” speed channel.

Variant 2 (spatial + temporal accumulation)

In this STP variant, the coherent summation in Eq. (14) is first implemented by the “ n ” index and then subsequently by the “ m ” index.

We rewrite Eq. (14) in a different order using the $p_{0n}^{(m)} = p_0^{(m)}$ condition:

$$\xi = \left\| \sum_{n=1}^N A_n \exp \left[j \frac{2\pi d_r (n-1)}{\lambda} (\sin \theta_{TRG} - \sin \theta_0) \right] \right\| \sum_{m=1}^M \frac{1}{P_0^{(m)}} U_{m \text{ mf}} \exp \left[j \frac{2\pi d_t (m-1)}{\lambda} (\sin \theta_{TRG} - \sin \theta_0) \right]. \quad (21)$$

As in the first case, we split the processing into two stages.

Stage 2.1

For this variant, since θ_{TRG} in (21) is unknown a priori, spatial coherent integration must first be performed for the N expected directions of signal arrival. Using the FFT algorithm with the transformation matrix

$\mathbf{W}(\theta_{0s}) = \left\{ \exp\left(-j \frac{2\pi s n}{N}\right) \right\}_{n=-N/2}^{N/2}$, the first sum in (21) for an arbitrary s th phasing angle can be expressed as follows:

$$F_r(0, \theta_{\text{TRG}}, \theta_{0s}) = \sum_{n=0}^{N/2} A_n \cos \left[\frac{\pi d_r}{\lambda_0} \left(n - \frac{1}{2} \right) \left(\sin \theta_{\text{TRG}} \frac{\left(s - \frac{1}{2} \right) \lambda_0}{(N-1)d_r} \right) \right]. \quad (22)$$

For the maximum angle of phasing $\sin \theta_{0N/2}$, it is important that the condition $d_r \leq \lambda_0(1 + |\sin \theta_{0N/2}|)^{-1}$ be imposed on a step of the reception array in order to prevent diffraction maxima in DP. Then, at the first stage, an N -beam DP can be generated in the receiving AA that overlaps a given survey sector. This corresponds to the implementation of secondary receiving channels.

Unlike the previous variant, it should be noted that the DP in (22) does not have the index “v,” since the channels of the receiving AA physically exist. In particular, during digital charting, analogue-to-digital converters (ADCs) must be installed in each of the N primary channels. The DP $F_r(0, \theta_{\text{TRG}}, \theta_{0s/2})$ can be interpreted as “virtual” if, in formula (21), the summation over the “ m ” index is considered an array multiplier and the summation over the “ n ” index is considered a partial DP of the subarray.

Within each s th secondary channel, M tertiary channels should now be generated. As in the first variant, $N \times M$ receiving channels will also be implemented in the second variant.

Stage 2.2

The use of MF in each s th secondary channel enables the temporary accumulation of signal components and compensates for their spatial phase progression. Following the temporary accumulation at time $t = 0$, the second sum in (21) can be interpreted as an M element AA multiplier, where the maximum signal component values at the SF output act as amplitude distribution.

In this variant, the equivalent MIMO radar receiving AA is considered as an array consisting of M subarrays, each having N elements and a single d_r step. For the equivalent $N \times M$ -element AA to be completely filled with radiators of the same step size and of the same length as in variant 1, the transmitting AA can be sparse with an equidistant step size equal to $d_t = (N-1)d_r$.

Next, a multipath diagram is implemented in the equivalent receiving AA to ensure that the sharply directed tertiary beam falls within the same i th direction as the s th secondary beam. Therefore, in this variant, the beams are double-indexed “ si ”, rather than “ is ,” as in variant 1. We apply the FFT algorithm to the samples of the output signals

of the secondary channels using the transformation matrix $\mathbf{W}(\theta_{0i}) = \left\{ \exp\left(-j \frac{2\pi i m}{M}\right) \right\}_{i=1}^M$. Thus, the DP of the si -tertiary channel, taking into account Eq. (22), acquires the following form:

$$F_r(0, \theta_{\text{TRG}}, \theta_{0si}) = F_{\text{el}}(\theta_{\text{TRG}}) F_v(0, \theta_{\text{TRG}}, \theta_{0s}) \sum_{n=0}^{M/2} \frac{A_n}{P_0^{(m)}} \cos \left[\frac{\pi d_t}{\lambda_0} \left(m - \frac{1}{2} \right) \left(\sin \theta_{\text{TRG}} \frac{\left(i - \frac{1}{2} \right) \lambda_0}{(N-1)d_r} \right) \right], \quad (23)$$

where $F_v(0, \theta_{\text{TRG}}, \theta_{0s})$ is the DP defined by Eq. (22).

Thus, sufficient statistics for ξ (16) are maximized in the phasing directions.

COMPARATIVE ANALYSIS OF STP VARIANTS

In the synthesized STP variants, the sufficient pre-threshold statistics ξ , which is the same for all variants, is equal to the squared modulus of the correlation integral (14). This corresponds to the classical detector of a quasi-deterministic signal having a random uniform phase and random Rayleigh

amplitude against the WGN background [9, 12]. Therefore, the same well-known formulas and curves can be used for detection characteristics that depend only on the power SNR (q^2) in both cases. According to the Neumann–Pearson criterion, the detection quality indicators D and F , as well as the detection threshold η , are related by the following approximate ratios:

$$D = \frac{1}{F^{1+q^2}}, \quad \eta = \sqrt{2 \ln \frac{1}{F}}, \quad (24)$$

where D and F are the conditional probabilities of correct detection and false alarm, respectively.

It should also be noted that any parallel survey radar requires a device for selecting the spatial channel with the maximum SNR. Since the detection threshold may be exceeded in multiple reception channels simultaneously, the direction to the target is determined by the index of this channel.

The STP variants under comparison differ in the sequence of temporal and spatial signal accumulation. While this sequence does not affect the final result, it can significantly impact the practical implementation of MIMO radar. In particular, it affects the construction of the transmitting and receiving ARs.

The sparse array used for reception in variant 1 has a step size of $d_r = (M-1)d_t$, which is equal to the length of the transmitting AA. In variant 2, the transmitting AA with a step $d_t = (N-1)d_r$ is considered as sparse. Furthermore, the d_t step in variant 1 and the d_r step in variant 2 can be considered to be equivalent since they are both selected based on the absence of diffraction maxima in the angular survey areas for all scanning angles.

As shown in Fig. 3, the block diagram of the transmitting M -element AA is the same for both STP variants.

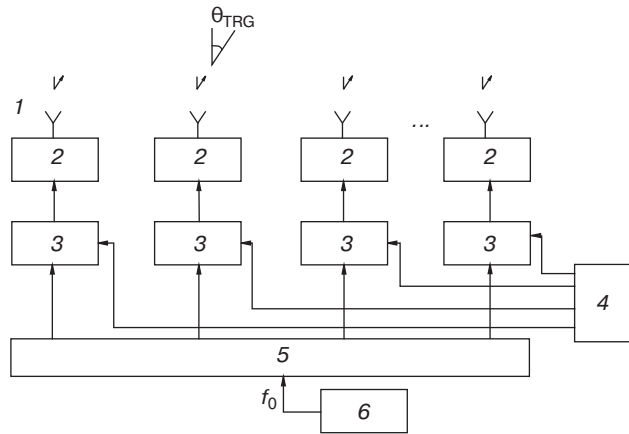


Fig. 3. Block diagram of the transmitting AA:

- (1) radiators, (2) power amplifier,
(3) modulators (mixers), (4) complex signal generator,
(5) divider $1:M$, and (6) driving oscillator

In order to consider the practicality and suitability of their implementation, we will now move on to a qualitative analysis of STP variants at the receiving AA output. The block diagram of variant 1 with initial temporary accumulation is shown in Fig. 4. This assumes $M > N$, which simplifies the construction of the receiving AA. Multibeam VSA DP is provided by an M -point FFT based on the signals of its primary

channels. The signals in all secondary channels are processed by an N -point FFT to generate $M \times N$ tertiary channels with sharply directional DPs. These channels correspond to the equivalent length of the receiving AA, which is equal to $L_{an.eq} = (M-1)(N-1)d_t$.

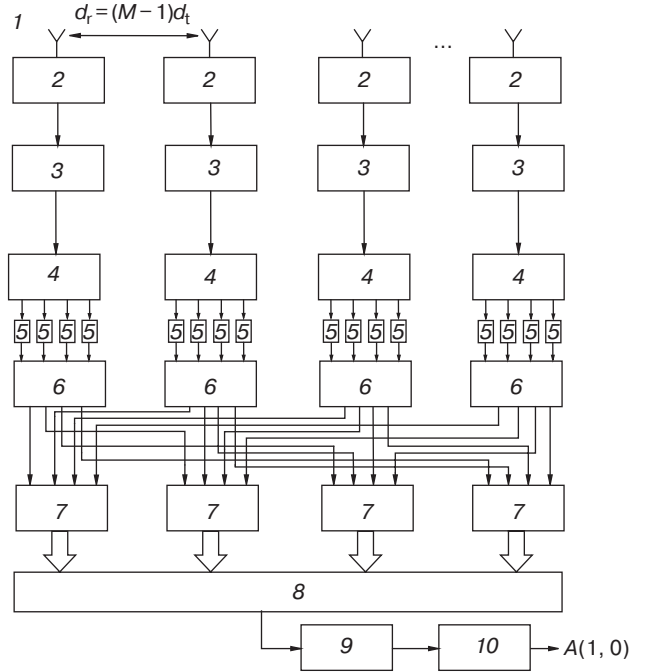


Fig. 4. Block diagram of the first STP variant

- (wide arrows represent M -dimensional vectors):
(1) primary emitters; (2) amplifier-converter unit;
(3) ADC; (4) divider $1:M$; (5) MF component of
a complex signal; (6) FFT-temporary accumulation
(generation of secondary channels); (7) FFT-spatial
accumulation (generation of tertiary channels);
(8) formation of pre-threshold statistics; (9) threshold
device (decision-making); (10) maximum selection

Figure 5 shows the STP block diagram for variant 2 with initial spatial accumulation. The assumption of $M < N$ in this variant simplifies the construction of the transmitting AA. Spatial processing is provided here by an N -point FFT of samples in the primary channels, which is reduced to generating N secondary channels (beams). Following the temporary accumulation of secondary channel samples in the MF due to the M -point FFT, as in variant 1, $M \times N$ tertiary channels are generated, whose DP is determined by an AA length that is equal to $L_{an.eq} = (M-1)(N-1)d_r$.

Assuming a fixed SNR (since the d_t step in variant 1 is equal to the d_r step in variant 2 for a given scan sector), it can be inferred that the MIMO radar's angular resolution is the same for the compared STP variants, as determined by $L_{an.eq}$. Here, the blocks for formation of pre-threshold statistics (block 8) involve selection of a tertiary channel having a maximum value of $\xi = \max|Z|^2$ (block 10), while the threshold device (block 9) in the compared STP variants are identical to those in conventional radars and have no specific features.

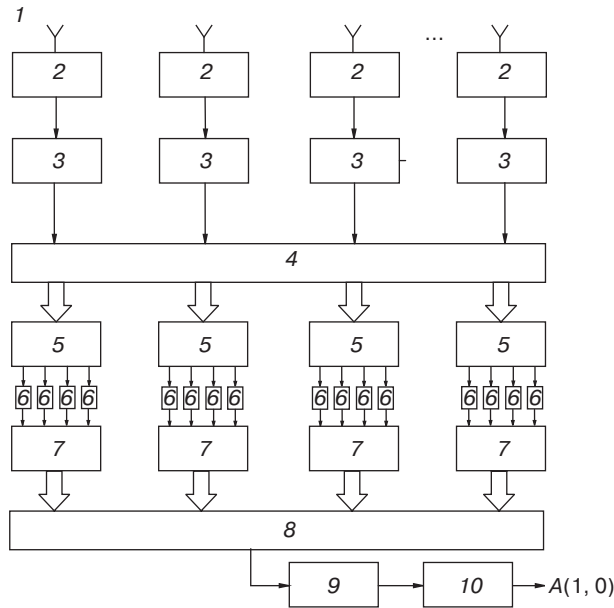


Fig. 5. Block diagram of the second STP variant (wide arrows represent N - and M -dimensional vectors): (1) primary radiators; (2) an amplifier-converter unit; (3) an ADC; (4) FFT spatial accumulation (generation of secondary channels); (5) divider 1 : M ; (6) MF components of a complex signal; (7) FFT temporal accumulation (generation of tertiary channels); (8) formation of pre-threshold statistics; (9) a threshold device (decision-making); (10) maximum selection

An analysis of the STP schemes in Figs. 4 and 5 shows that variant 1 may be preferable when implementing a compact, multielement transmitting AA based on relatively low-power amplifier modules. A compact AA allows for easier synchronization of the modulators' operation to achieve the desired coherence of the orthogonal signal components. For a sparse receiving AA, this variant reduces the number of amplifier-converter units and ADCs. However, implementing their synchronous operation may present issues due to the spatial separation of the ADCs.

It should be noted that several sources regard STP variant 1 as the only viable ground-based radar, for instance [19, 20]. For medium- and long-range radars, recall that the transmitting AA of the MIMO radar serves as the VSA of the "equivalent" receiving AA.

The second STP variant is preferable when implementing a compact multibeam receiving AA with either analogue (Butler matrix) or digital (FFT) beamforming. In this variant, the transmitting AA is sparse; however, due to the increased power of the channel amplifiers, it should provide the same energy potential as variant 1, albeit having fewer M channels. Due to the spatial separation of the channels, maintaining signal component coherence may be challenging. Variant 2 has recently been used in airborne MIMO aircraft radars [21].

The qualitative comparative analysis reveals that the viability of implementing one variant over another depends significantly on several subjective technical factors. These include the frequency band, the ultrahigh frequency and digital element bases used, and the permissible design dimensions of the MIMO radar.

In conclusion, we explain the benefits of angular resolution provided by MIMO radar with any STP variant as compared to the fixed SNR of a conventional radar design. In radar theory, it is assumed that the angular resolution of the radar at the operating wavelength λ_0

has antenna aperture length L equal to $\Delta\theta_0 = \frac{\lambda_0}{L \cos \theta}$,

where θ is the beam deviation angle from the normal. Often, the resolution is determined by the equivalent Rayleigh criterion through the antenna beam width at half power $\Delta\theta_{0.5}$.

The above relation for $\Delta\theta_0$ is derived from the following conditional criterion. Consider two radiation sources in space with amplitudes A_1 and A_2 , separated by an angle $\delta\theta$. The total area at the opening of a linear antenna aligned along the z -axis is then given by:

$$E(x) = A_1 \exp(j2\pi f_0 t) \times \exp\left(j \frac{2\pi}{\lambda_0} z \sin \theta\right) \left[1 + \frac{A_2}{A_1} \exp\left(j \frac{2\pi}{\lambda_0} z \delta\theta \cos \theta\right) \right]. \quad (25)$$

In Eq. (25), the multiplier in front of the square brackets corresponds to the linear phase progression along the aperture (along the z -axis) from the A_1 first source. The presence of a second source may be determined by the modulating multiplier inside the square brackets, which varies slowly depending on z . Here it is assumed that this modulation has been established if at least one complete period of it is observed. This implies a lower limit on the aperture length of $2\pi L \delta\theta \cos \theta / \lambda \geq 2\pi$. The condition of equality implies the expression for $\Delta\theta_0 = \delta\theta_{\min}$.

For STP variants 1 and 2 considered in Eqs. (16) and (21), the sufficient statistics involving double summations by indices $n \in (1 - M)$ and $m \in (1 - M)$ correspond to obtaining DP of equivalent receiving ARs consisting of $M \times N$ elements, with unambiguous steps for scanning d_t and d_r , respectively. Since this determines the elongation of the aperture resolution in MIMO radars, a proportional improvement in angular resolution is achieved as compared to N -element AA in conventional radars.

For example, in variant 1, the physical step of the receiving AA is equal to the length of the transmitting AA. This is due to the temporary processing at the radiators' outputs; these can be interpreted as the VSAs with a length equal to $L_{\text{an.t}}$. In this scenario, the equivalent length of the receiving AA aperture is $(M - 1)$ times greater than that of a conventional N -element array

with a step size of d_t . Therefore, the modulating phase multiplier in Eq. (25) achieves a value of 2π with a separation of the sources $\delta\theta$ that is $(M-1)$ times lower.

CONCLUSIONS

The results of the study on STP synthesis against the WGN background in MIMO radar with transmitting M -element and receiving N -element AAs support the following conclusions:

1. According to the Neumann–Pearson criterion, the optimal STP against the WGN background is a combination of the temporal and spatial accumulation of M orthogonal signal components ranged by the transmitting AA. The accumulation sequence depends on the method used to represent the interference CM in Eqs. (10) and (11), resulting in two variants for constructing a consistent STP (Figs. 4 and 5). In both variants, the length of the equivalent (virtual) aperture $L_{\text{an.eq}}$ of the receiving AA is proportional to the product $M \times N$, whose single-element step is determined by the scanning sector.
2. The first variant of the agreed STP consistently implements the following:
 - Temporary accumulation of M signal components in the MF from the N outputs of the primary radiators of the receiving AA.
 - Spatial accumulation during M -beam diagramming in N VSAs due to the M -point FFT (generation of secondary receiving channels).
 - Spatial accumulation due to the N -point FFT at the output of all secondary channels, resulting in the generation of $M \times N$ orthogonal beams of equivalent AA (generation of tertiary receiving channels).
3. The second variant of the agreed STP consistently implements the following:
 - Spatial accumulation of the signal due to the N -point FFT at the outputs of the primary radiators of the receiving AA (generation of secondary reception channels).
 - Temporary accumulation of M signal components in the MF from the N outputs of the secondary channels.

- Spatial accumulation due to the M -point FFT at the MF output of all secondary channels, resulting in the generation of $M \times N$ orthogonal beams of equivalent AA (generation of tertiary receiving channels).
4. The same detection quality indicators can be achieved in both STP variants if they provide the same radiated power and product $M \times N$, as well as having an unambiguous scanning step. Therefore, the processing steps relating to the formation of the pre-threshold statistics and the detection threshold for decision-making in Figs. 4 and 5 are identical.
 5. The size of the equivalent aperture of receiving AA is the same for the variants under comparison. Consequently, they implement the same angular resolution, which is inversely proportional to $L_{\text{an.eq}}$. In the first variant, the equivalent AA is interpreted as an $M \times N$ -element array consisting of $N \times M$ -element VSAs ($M > N$). In the second variant, it is interpreted as an $M \times N$ -element array consisting of $M \times N$ -element VSAs ($M < N$).

Abbreviations

AA—antenna array;
ADC—analog-to-digital converter;
CM—correlation matrix;
DP—directional pattern;
FFT—Fast Fourier transform;
MF—matched filter AA;
MIMO—Multiple-Input, Multiple-Output;
SNR—signal-to-noise ratio;
STP—space-time processing;
VSA—virtual subarray;
WDA—weakly directional antenna;
WGN—white Gaussian noise.

Authors' contributions

B.M. Vovshin—general statement of the problem, development of principles for constructing two options of STP, development of methods for their comparison.

A.A. Pushkov—development of structural design schemes, derivation of analytical ratios, and analysis of the results of comparison two options of STP.

REFERENCES

1. Bergin J., Guerci J.R. *MIMO Radar. Theory and Application*. Artech House; 2018, 229 p.
2. Slukin G.P., Chapurskii V.V. (Eds.). *Prikladnye zadachi teorii MIMO sistem radiolokatsii i radiologografii (Applied Problems of the Theory of MIMO Radar and Radiologography Systems)*. Moscow: Bauman Press; 2024, 275 p. (In Russ.). ISBN 978-5-7038-6393-0
3. Vovshin B.M. Parallel Surveillance Ultra-wide-band Radars with Orthogonal Ranging Signals. In: *Proceedings on International Radar Symposium (IRS-2007)*. Cologne, Germany. P. 461–466.
4. Li J., Stoica P. *MIMO Radar Signal Processing*. Wiley; 2009, 472 p. ISBN 978-0-470-17898-0

5. Vovshin B.M., Pushkov A.A., Khalturina E.M. On the equivalence of characteristics and specularity of the construction of the traditional and MIMO radars with parallel view of space based on antenna arrays. *Russian Technological Journal*. 2025;13(3):73–83. <https://doi.org/10.32362/2500-316X-2025-13-3-73-83>
6. Ermolaev V.T., Flaksman A.G., Elokhin A.V., Sorokin I.S. *Prostranstvennaya obrabotka signalov v MIMO sistemakh sotovoy svyazi (Spatial Signal Processing in MIMO Cellular Communication Systems)*. Nizhny Novgorod; 2020, 134 p. (In Russ.).
7. Vovshin B.M., Kulikov K.V. Spatial multi-channel reception in MIMO radar with parallel survey of space with orthogonal multi-frequency signals. *Uspekhi sovremennoi radioelektroniki = J. Achievements of Modern Radioelectronics*. 2016;8: 67–79 (in Russ.). <https://elibrary.ru/wjilgx>
8. Chernyak V.S. Signal detection in MIMO radars. *Uspekhi sovremennoi radioelektroniki = J. Achievements of Modern Radioelectronics*. 2014;7:35–48 (in Russ.). <https://elibrary.ru/sjlqgt>
9. Shirman Ya.D. (Ed.). *Radioelektronnye sistemy. Osnovy postroeniya i teoriya (Radioelectronic systems. Fundamentals of Construction and Theory)*. Moscow: Radiotekhnika; 2007, 512 p. (In Russ.).
10. Shirman Ya.D., Manzhos V.N. *Teoriya i tekhnika obrabotki radiolokatsionnoi informatsii na fone pomekh (Theory and Technology of Processing Radar Information Against a Background of Interference)*. Moscow: Radio i svyaz'; 1981, 416 p. (In Russ.).
11. Strang G. *Lineinaya algebra i ee primeneniya (Linear Algebra and its Applications)*: transl. from Engl. Moscow: Mir; 1980, 454 p. (In Russ.).
[Strang G. *Linear Algebra and its Applications*. Elsevier Sci.; 1976, 428 p.]
12. Vovshin B.M. *Teoreticheskie osnovy radiolokatsionnogo nablyudeniya tselei (Theoretical Foundations of Radar Surveillance of Targets)*. Moscow: Almaz; 2022, 312 p. (In Russ.). ISBN 978-5-6042294-5-3
13. Tahcfulloh S., Hendratoro G. FPMIMO: A General MIMO Structure with Overlapping Subarrays for Various Radar Applications. *IEEE Access*. 2020;8:11248–11267. <https://doi.org/10.1109/ACCESS.2020.2965192>
14. Vovshin B.M. *Sluchainye protsessy v radiotekhnike i radiolokatsii (Random Processes in Radio Engineering and Radar)*. Moscow: Almaz; 2024, 296 p. (In Russ.). ISBN 978-5-6052360-0-9
15. Chapurskii V.V., Slukin G.P., Filatov A.A., Koroteev D.E. Virtual MIMO radars and their comparison based on generalized uncertainty functions. *Uspekhi sovremennoi radioelektroniki = J. Achievements of Modern Radioelectronics*. 2023;77(5):5–19 (in Russ.).
16. Wang H., Yu J., Yu W., De Ben. Adaptive waveform design for distributed OFDM MIMO radar system in multi-target scenario. *Chinese J. Aeronautics*. 2018;31(3):567–574. <https://doi.org/10.1016/j.cja.2018.01.016>
17. Kosogorov A.A. (Ed.). *Antenny, SVCh ustroystva i ikh tekhnologii (Antennas, Microwave Devices and Their Technologies)*. Moscow: Fizmatlit; 2023, 400 p. (In Russ.). ISBN 978-5-9221-1966-5
18. Tan L., Jiang J. *Digital Signal Processing. Fundamentals and Applications*. Academic Press; 2019, 838 p.
19. Chernyak V.S. About new and old ideas in radar: MIMO radars. *Uspekhi sovremennoi radioelektroniki = J. Achievements of Modern Radioelectronics*. 2011;2:5–20 (in Russ.). <https://elibrary.ru/ntmkap>
20. Dorey J., Blanchard Y., Christophe F. The RIAS project, a new approach to air surveillance radar. *Radiant. J. Dutch. Meteor. Soc*. 1984;20:21–24.
21. Klare J., Saalman O. MIRA-CLE X: A new Imaging MIMO Radar for Multi-Purpose Applications. In: *Proceedings of the 7th European Radar Conference (EuRAD)*, Paris, France. IEEE. 2010. P. 129–132.

СПИСОК ЛИТЕРАТУРЫ

1. Bergin J., Guerci J.R. *MIMO Radar. Theory and Application*. Artech House; 2018, 229 p.
2. *Прикладные задачи теории ММО систем радиолокации и радиолокации*; под ред. Г.П. Слукина, В.В. Чапурского. М.: Изд-во МГТУ им. Н.Э. Баумана; 2024, 276 с. ISBN 978-5-7038-6393-0
3. Vovshin B.M. Parallel Surveillance Ultra-wide-band Radars with Orthogonal Ranging Signals. In: *Proceedings on International Radar Symposium (IRS-2007)*. Cologne, Germany. P. 461–466.
4. Li J., Stoica P. *MIMO Radar Signal Processing*. Wiley; 2009, 472 p. ISBN 978-0-470-17898-0
5. Вовшин Б.М., Пушков А.А., Халтурина Е.М. Об эквивалентности характеристик и «зеркальности» построения традиционных и ММО РЛС при параллельном обзоре пространства на основе антенных решеток. *Russian Technological Journal*. 2025;13(3):73–83. <https://doi.org/10.32362/2500-316X-2025-13-3-73-83>
6. Ермолаев В.Т., Флакман А.Г., Елохин А.В., Сорокин И.С. *Пространственная обработка сигналов в ММО системах сотовой связи*. Нижний Новгород: Изд-во ННГУ; 2020, 134 с.
7. Вовшин Б.М., Куликов К.В. Пространственно-многоканальный прием в ММО РЛС параллельного обзора пространства с ортогональными многочастотными сигналами. *Успехи современной радиоэлектроники*. 2016;8:67–79. <https://elibrary.ru/wjilgx>
8. Черняк В.С. Обнаружение сигналов в ММО РЛС. *Успехи современной радиоэлектроники*. 2014;7:35–48. <https://elibrary.ru/sjlqgt>
9. *Радиоэлектронные системы. Основы построения и теория*; под ред. Я.Д. Ширмана. М.: Радиотехника; 2007, 512 с.
10. Ширман Я.Д., Манжос В.Н. *Теория и техника обработки радиолокационной информации на фоне помех*. М.: Радио и связь; 1981, 416 с.
11. Стренг Г. *Линейная алгебра и ее применения*: пер. с англ. М.: Мир; 1980, 454 с.

12. Вовшин Б.М. *Теоретические основы радиолокационного наблюдения целей*. М.: ПАО «НПО «Алмаз»; 2022, 312 с. ISBN 978-5-6042294-5-3
13. Tahcfulloh S., Hendranto G. FPMIMO: A General MIMO Structure with Overlapping Subarrays for Various Radar Applications. *IEEE Access*. 2020;8:11248–11267. <https://doi.org/10.1109/ACCESS.2020.2965192>
14. Вовшин Б.М. *Случайные процессы в радиотехнике и радиолокации*. М.: ПАО «НПО «Алмаз»; 2024, 296 с. ISBN 978-5-6052360-0-9
15. Чапурский В.В., Слукин Г.П., Филатов А.А., Коротеев Д.Е. Виртуальное МИМО РЛС и их сравнение на основе обобщенных функций неопределенности. *Успехи современной радиоэлектроники*. 2023;77(5):5–19. <https://doi.org/10.18127/j20700784-202305-01>
16. Wang H., Yu J., Yu W., De Ben. Adaptive waveform design for distributed OFDM MIMO radar system in multi-target scenario. *Chinese J. Aeronautics*. 2018;31(3):567–574. <https://doi.org/10.1016/j.cja.2018.01.016>
17. *Антенны, СВЧ устройства и их технологии*; под ред. А.А Косогорова. М.: Физматлит; 2023, 400 с. ISBN 978-5-9221-1966-5
18. Tan L., Jiang J. *Digital Signal Processing. Fundamentals and Applications*. Academic Press; 2019, 838 p.
19. Черняк В.С. О новых и старых идеях в радиолокации: МИМО РЛС. *Успехи современной радиоэлектроники*. 2011;2:5–20. <https://elibrary.ru/ntmkap>
20. Dorey J., Blanchard Y., Christophe F. The RIAS project, a new approach to air surveillance radar. *Radiant. J. Dutch. Meteor. Soc.* 1984;20:21–24.
21. Klare J., Saalmann O. MIRA-CLE X: A new Imaging MIMO Radar for Multi-Purpose Applications. In: *Proceedings of the 7th European Radar Conference (EuRAD)*, Paris, France. IEEE. 2010. P. 129–132.

About the Authors

Boris M. Vovshin, Dr. Sci. (Eng.), Professor, Department of Radio Wave Processes and Technology, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia); Chief Researcher, ALMAZ R&P Corp. (110, Dmitrovskoe sh., Moscow, 127411 Russia). E-mail: boris@eleron.net. Scopus Author ID 7801355022, <https://orcid.org/0009-0003-1357-5866>

Alexander A. Pushkov, Cand. Sci. (Eng.), Chief Researcher, ALMAZ R&P Corp. (110, Dmitrovskoe sh., Moscow, 127411 Russia). E-mail: aapushkov@mail.ru. <https://orcid.org/0009-0005-3831-4528>

Об авторах

Вовшин Борис Михайлович, д.т.н., профессор, кафедра радиоволновых процессов и технологий, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78); главный научный сотрудник, ПАО «Научно-производственное объединение «Алмаз» имени академика А.А. Расплетина (ПАО «НПО «Алмаз») (127411, Россия, Москва, Дмитровское шоссе, д. 110). E-mail: boris@eleron.net. Scopus Author ID 7801355022, <https://orcid.org/0009-0003-1357-5866>

Пушков Александр Александрович, к.т.н., начальник НТЦ, ПАО «Научно-производственное объединение «Алмаз» имени академика А.А. Расплетина (ПАО «НПО «Алмаз») (127411, Россия, Москва, Дмитровское шоссе, д. 110). E-mail: aapushkov@mail.ru. Scopus Author ID 54407412700, <https://orcid.org/0009-0005-3831-4528>

Translated from Russian into English by K. Nazarov

Edited for English language and spelling by Thomas A. Beavitt

Modern radio engineering and telecommunication systems
Современные радиотехнические и телекоммуникационные системы

UDC 621.391.8

<https://doi.org/10.32362/2500-316X-2026-14-4-82-95>

EDN VLPBLV



RESEARCH ARTICLE

Hybrid neural network architecture ResBiLSTM-BiGRU for efficient noise suppression in radio channels with digitally modulated signals

Nguyen Van Dung[@],
Phan Trong Hanh

Le Quy Don Technical University, Hanoi, Vietnam

[@] Corresponding author, e-mail: nguyenvandungvtdt@lqdtu.edu.vn

• Submitted: 24.11.2025 • Revised: 24.01.2026 • Accepted: 25.05.2026

Abstract

Objectives. The work sets out to develop an effective method for suppressing additive Gaussian noise in radio channels to ensure the high-quality reconstruction of digitally modulated signals at low signal-to-noise ratio (SNR) levels.

Methods. To address this problem, a hybrid architecture is proposed that combines Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) recurrent neural networks. The ResBiLSTM-BiGRU model enables more stable training and improved reconstruction quality, as well as more accurate modeling of the temporal structure of signals. The performance of the model was evaluated using a multi-modulation dataset with various input SNR levels.

Results. As confirmed by the experimental results, the proposed model achieves an average improvement of 2–3 dB in output SNR compared with the Relativistic Average Generative Adversarial Network (RaGAN) under identical simulation conditions, as well as a more than a 10-dB improvement as compared with the wavelet-transform-based method. The model is characterized by lower mean squared error values and higher correlation coefficients across the entire input SNR range from –10 to 10 dB. The peak SNR values also indicate high signal reconstruction quality, especially at low input SNR levels. Analysis of architectural parameters shows that the best performance is achieved with 128 hidden units and a deep ResBiLSTM-BiGRU architecture.

Conclusions. The obtained results confirm the effectiveness and practical applicability of the proposed ResBiLSTM-BiGRU hybrid architecture for suppressing additive Gaussian noise in radio channels and reconstructing digitally modulated signals, particularly under low-SNR conditions.

Keywords: cognitive radio networks, automatic modulation classification, digitally modulated signals, deep learning, hybrid ResBiLSTM-BiGRU network, noise suppression, signal-to-noise ratio, mean squared error, correlation coefficient, peak signal-to-noise ratio

For citation: Nguyen V.D., Phan T.H. Hybrid neural network architecture ResBiLSTM-BiGRU for efficient noise suppression in radio channels with digitally modulated signals. *Russian Technological Journal*. 2026;14(4):82–95. <https://doi.org/10.32362/2500-316X-2026-14-4-82-95>, <https://elibrary.ru/VLPBLV>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Гибридная архитектура нейронной сети ResBiLSTM-BiGRU для эффективного подавления шума в радиоканале с цифровыми модулированными сигналами

В.З. Нгуен[@],
Ч.Х. Фан

Вьетнамский государственный технический университет им. Ле Куй Дона, Ханой, Вьетнам
[@] Автор для переписки, e-mail: nguyenvandungvtdt@lqdtu.edu.vn

• Поступила: 24.11.2025 • Доработана: 24.01.2026 • Принята к опубликованию: 25.05.2026

Резюме

Цели. Целью данной работы является разработка эффективного метода подавления аддитивного гауссовского шума в радиоканале, обеспечивающего высокое качество восстановления цифровых модулированных сигналов при низких значениях отношения сигнал/шум (signal-to-noise ratio, SNR).

Методы. Для решения поставленной задачи предложена гибридная архитектура, объединяющая рекуррентные нейронные сети типов долгой краткосрочной памяти (long short-term memory, LSTM) и управляемых рекуррентных блоков (gated recurrent unit, GRU). Модель ResBiLSTM-BiGRU обеспечивает более стабильное обучение, повышенное качество восстановления и более точное моделирование временной структуры сигналов. Эффективность модели оценивалась на многомодуляционном наборе данных при различных уровнях входного SNR.

Результаты. Экспериментальные исследования показали, что предложенная модель в среднем обеспечивает увеличение выходного SNR на 2–3 дБ по сравнению с релятивистской усредненной генеративно-сопоставительной сетью (relativistic average generative adversarial networks, RaGAN) и более чем на 10 дБ по сравнению с методом на основе вейвлет-преобразования при одинаковых условиях моделирования. Модель характеризуется меньшими значениями среднеквадратической ошибки и более высокими коэффициентами корреляции во всем диапазоне входных SNR от –10 до 10 дБ. Значения пикового SNR также свидетельствуют о высоком качестве восстановления сигналов, особенно при низких уровнях входного SNR. Анализ архитектурных параметров показал, что наилучшие результаты достигаются при использовании 128 скрытых узлов и глубокой архитектуры ResBiLSTM-BiGRU.

Выводы. Полученные результаты подтверждают эффективность и практическую применимость гибридной архитектуры ResBiLSTM-BiGRU для подавления аддитивного гауссовского шума в радиоканале и восстановления цифровых модулированных сигналов, особенно в условиях низкого SNR.

Ключевые слова: когнитивные радиосети, автоматическая классификация модуляции, сигналы цифровой модуляции, глубокое обучение, гибридная сеть ResBiLSTM-BiGRU, подавление шума, отношение сигнал/шум, среднеквадратическая ошибка, коэффициент корреляции, пиковое отношение сигнал/шум

Для цитирования: Нгуен В.З., Фан Ч.Х. Гибридная архитектура нейронной сети ResBiLSTM-BiGRU для эффективного подавления шума в радиоканале с цифровыми модулированными сигналами. *Russian Technological Journal*. 2026;14(4):82–95. <https://doi.org/10.32362/2500-316X-2026-14-4-82-95>, <https://elibrary.ru/VLPBLV>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

The increased use of modern communication technologies such as 5G/6G networks and the Internet of Things (IoT) has led to greater demands for high-quality signal transmission through complex, noisy communication channels. This has given rise to a significant challenge in the form of additive white Gaussian noise (AWGN). By significantly reducing signal quality, AWGN complicates various applications, including automatic modulation classification, spectral sensing and resource management in cognitive radio networks.

Traditional noise suppression methods, such as linear filtering and signal transformations in the time and frequency domains, have been used for several decades. While these methods demonstrate satisfactory performance in linear channels with medium to high signal-to-noise ratios (SNRs), they become less effective in the presence of nonlinear signals and low SNR conditions. This limits their practical applicability. However, they are ineffective in the presence of nonlinear signals and low SNR conditions.

In recent years, deep learning has opened up new possibilities for radio signal processing. Models such as denoising autoencoders (DAEs) [9, 10], generative adversarial networks (GANs) [11–13], diffusion models [14, 15], and transformers have been shown to be effective at suppressing noise. However, each of these models has its limitations. For instance, DAEs become less efficient at low SNRs. GANs and their variations are difficult to train and require a lot of computational resources. Although diffusion models and transformers perform well in high-noise scenarios, they also have high computational requirements and long output delays. Long short-term memory (LSTM) and gated recurrent unit (GRU) networks do not exploit their complementary capabilities fully when used separately.

Based on these findings, the present work proposes a ResBiLSTM-BiGRU hybrid architecture that combines the strengths of LSTMs and GRUs. While LSTMs facilitate the learning of long-term dependencies, GRUs offer compactness and rapid convergence. The study aims to develop an efficient and practical approach to

AWGN suppression in digital radio channels for next-generation communication systems.

Specifically, the study:

- develops a specialized ResBiLSTM-BiGRU hybrid architecture for noise reduction in digital radio signals;
- conducts a thorough analysis of the results across a wide SNR range of from -10 dB to 10 dB and various modulation techniques; and
- demonstrates the efficacy of the proposed approach compared to traditional noise reduction methods using metrics such as SNR, mean square error (MSE), root mean square error (RMSE) and peak signal-to-noise ratio (PSNR).

1. RELATED RESEARCH

Over several decades, both conventional techniques and modern deep learning-based approaches have been used to address AWGN. While conventional techniques remain popular due to their simplicity and reliability, they are less effective under more challenging scenarios, such as those involving low SNR or nonlinear channels. Table 1 summarizes the key features, advantages, and limitations of some of the most commonly used conventional noise reduction methods.

Contemporary approaches based on deep learning offer high flexibility and efficacy in addressing AWGN, especially in low SNR settings. However, there are still several challenges to overcome when implementing deep learning models for noise reduction in digital communications. Table 2 outlines some recent research efforts to address AWGN in digital modulation systems.

The experimental phase of the research involves a comparative analysis of two noise reduction methods. One of these techniques is based on wavelet transformations (WT), while the other uses relativistic average generative adversarial networks (RaGAN). The results suggest that a hybrid architecture combining ResBiLSTM and BiGRU networks achieves more consistent signal recovery accuracy when used in challenging communication channels with low SNR for AWGN suppression across various modulation techniques.

Table 1. Characteristics of conventional noise reduction methods

Method group	Primary methods	Key features	Limitations
Linear filtering [1, 2]	Wiener filtering, Kalman filtering, adaptive least mean squares (LMS) algorithm, and recursive least squares (RLS) algorithm	• Wiener filter optimizes MSE in a linear environment; • Kalman filtering is effective for linear dynamical systems; • LMS/RLS provides adaptive filtering for a changing channel	• linear properties of signal and noise are required; • reduced efficiency with low SNR; • limitations when working in non-linear channels or with non-stationary noise
Noise reduction in the transformed area [3–6]	WT, Fourier transform, and discrete cosine transform (DCT)	• WT provides signal analysis in both time and frequency domains simultaneously; • the Fourier transform and DCT represent a signal exclusively in the frequency domain; • these methods are most effective for signals with a pronounced and stable frequency structure	• the effectiveness of WT depends on the choice of the basic function and the threshold; • the Fourier transform and DCT are ineffective with AWGN; • partial loss of local information is possible
Statistical methods [7, 8]	Minimum mean square error (MMSE), maximum <i>a posteriori</i> probability (MAP) method, principal component analysis (PCA), singular value decomposition (SVD)	• MMSE/MAP uses probabilistic signal and noise models; • PCA/SVD allows you to separate signal and noise components; • effective with a complete statistical description of the environment	• precise knowledge of signal and noise distributions is required; • high computational complexity; • efficiency is significantly reduced with an inaccurate model

Table 2. Characteristics of deep learning methods

Author(s), Year, Study	Architecture	Data / Channel	Key concepts and benefits	Limitations
Ruifeng Duan, Ziyu Chen, Haiyan Zhang, Xu Wang, Wei Meng, Guodong Sun, 2023 [10]	Autoencoder with skip connections and channel-wise attention mechanism	Digital modulated signals; AWGN	Combination of two residual encoder-decoder network connections and an advanced SE-ResNet ¹ module to enhance feature extraction and reduce computational complexity	Comparison only with RLS, LMS, and PCA
Shoushuai He, Lei Zhu, Changhua Yao, Lei Wang, Zhen Qin, 2022 [11]	Modified GAN architecture	Digital modulated signals; AWGN + interference	Restoring parts is better than traditional filters	The difficulty of learning, the difficulty of setting up
Liang Peng, Shengliang Fang, Youchen Fan, Mengtao Wang, Zhao Ma, 2023 [12]	RaGAN	Digital modulated signals; AWGN	Using GAN to generate a clean signal from a noisy one, improving the quality of recovery compared to traditional filters	Learning complexity, high computational load
Tong Wu, Zhiyong Chen, Dazhi He, Liang Qian, Yin Xu, Meixia Tao, Wenjun Zhang, 2023 [14]	The diffusion model is used to predict and suppress noise after channel alignment	Relay and AWGN communication channel	Determines the diffusion process according to the conditional distribution of the received signal, and can be integrated as a noise reduction module to improve signal quality	High computational complexity
Zelin Ji, Shuo Wang, Kuojun Yang, Qinchuan Zhang, Peng Ye, 2024 [16]	GRU-noise reduction module with classifier	Cyclic signal; digital modulated signal; AWGN channel	Pre-training of the noise canceller on a cyclic signal followed by fine-tuning on a digitally modulated signal; increased accuracy by more than 20% at low SNR values	Limited set of modulations; no evaluation using clean noise reduction metrics (MSE/SNR)

¹ The Squeeze-and-Excitation Residual Network (SE-ResNet) is a convolutional neural network architecture combining the skip connections of ResNet with a channel-wise attention mechanism based on Squeeze-and-Excitation blocks. This enables the adaptive recalibration of feature channels by modeling their mutual dependencies.

2. PROPOSED METHOD

2.1. ResBiLSTM-BiGRU network architecture for suppressing additive Gaussian noise and compensating for an imperfect transmission channel of basic I/Q² data

The paper assumes that the received data is basic I/Q data transmitted over a communication channel and corrupted by AWGN. In order to provide an effective noise suppression model for these signals, a ResNet³-based architecture called ResBiLSTM-BiGRU is proposed. The learning branch of this architecture is a hybrid network combining LSTM and GRU in cascade.

Before discussing this hybrid network architecture, a comparison of the key characteristics of LSTMs and GRUs is presented [17]. Table 3 summarizes their main differences.

Figure 1 depicts the overall architecture of the ResBiLSTM-BiGRU network. The network follows a similar structure to ResNet, comprising two parallel branches. The right branch serves as the primary data transmission path to ensure stable information flow and prevent the occurrence of gradient attenuation or explosion within the network. In contrast, the trainable left branch incorporates a hybrid LSTM-GRU structure alongside auxiliary layers, which is referred to as a “Hybrid LSTM-GRU.”

The model receives as input a basic I/Q signal contaminated by AWGN, represented as a 3D tensor with dimensions of $[B, T, 2]$. Here, B represents the batch size, T denotes the length of the time sequence, and 2 is the number of channels (I and Q). Due to the direct connection provided by the right branch, we will focus on data processing within the Hybrid LSTM-GRU branch for further analysis.

At the initial stage of processing, the data is passed through bidirectional long short-term memory (BiLSTM) layers. These layers extract time dependencies in both the forward and backward directions to capture long-term correlations within the data. The output of these layers is then fed into bidirectional gated recurrent units (BiGRU), which refine the features and capture local dependencies with higher frequency and lower computational cost due to their simplified gating structure.

Following the BiGRU processing, the output is normalized using layer normalization for stability in learning. The normalized data is then fed into a linear layer to reduce the dimension of the input. Finally, the outputs the direct and indirect paths are combined to generate the final output of the model.

This structure is intended to suppress AWGN and compensate for distortions caused by imperfections in the transmission channel that affect the received signal.

Table 3. Comparative characteristics of LSTM and GRU architectures

Characteristic	LSTM	GRU
Number of gates	Three: entry, forgetting, and exit	Two: update, reset
Number of states	Two: cell state and hidden state	One: hidden state
Memory mechanism	Linear transfer of information through the cell state resistant to long-term dependencies	Merging new and previous information in a hidden state, effective for medium and short dependencies
Model complexity	Higher	Lower
Number of parameters	Significantly more	Less (by 25–30% compared to LSTM)
Learning rate	Lower	Higher
Modeling long-term dependencies	Very high	High, but inferior to LSTM
Modeling short-term dynamics	Less sensitive	More sensitive
Overfitting tendency	Higher	Lower
Memory costs	Higher	Lower
Convergence behavior	Slower, hyperparameter-sensitive	Faster and more stable
Typical applications	Long sequences, signals with trend or drift	Noisy signals, real-time processing, limited resources

² I stands for the in-phase component, Q stands for the quadrature component, and the quadrature component refers to two orthogonal components of a complex signal.

³ A residual network is an architecture for convolutional neural networks that uses skip connections.

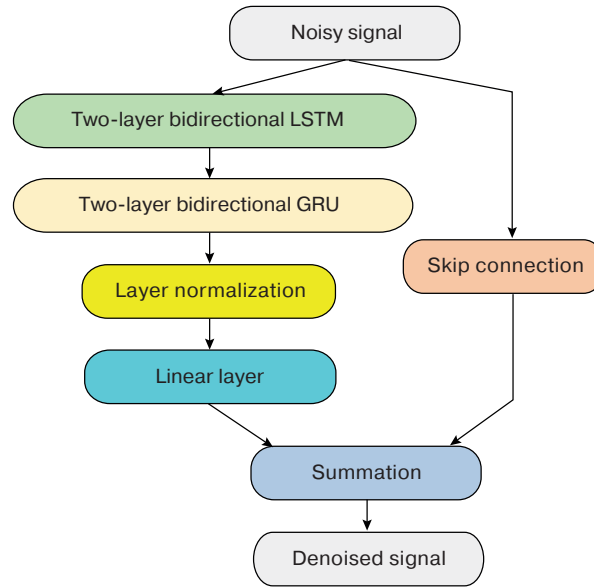


Fig. 1. Architecture of the ResBiLSTM-BiGRU hybrid network for noise suppression mixed with a digital modulation signal

2.2. Justification of the feasibility of the Hybrid LSTM-GRU structure

As illustrated in Fig. 1, a key feature of the proposed model is its hybrid architecture, which combines LSTM and GRU (Hybrid LSTM-GRU). Figure 2 provides a more detailed representation of this architecture.

Figure 2 depicts a network that operates on basic I/Q data. This modulated signal, which has been reduced to the base frequency band, contains information about amplitude, frequency and phase modulation. When transmitting over a channel, this signal is subject to additive noise, and can be modelled as follows:

$$x(n) = A(n)e^{j\phi} + \eta(n), \quad (1)$$

where $A(n)$ and $\phi(n)$ are slowly varying components carrying information about modulation, the relationship

between symbols and details of the transmission medium (including multipath effects and fading), and where $\eta(n)$ represents additive noise typically modelled as a white Gaussian process.

In this context, processing I/Q data involves suppressing additive noise, $\eta(n)$, and identifying signal patterns, as well as accounting for the effects of multipath/fading in $A(n)$ and $\phi(n)$. In the time domain, AWGN $\eta(n)$ varies rapidly and carries short-term information, whereas the signal patterns and multipath/fading effects are considered to be more slowly varying and to carry long-term information. In other words, the signal $x(n)$ in equation (1) exhibits multiscale temporal behavior.

The mathematical model of the I/Q data presented in equation (1) and the above analysis determine the network architecture shown in Fig. 2 directly.

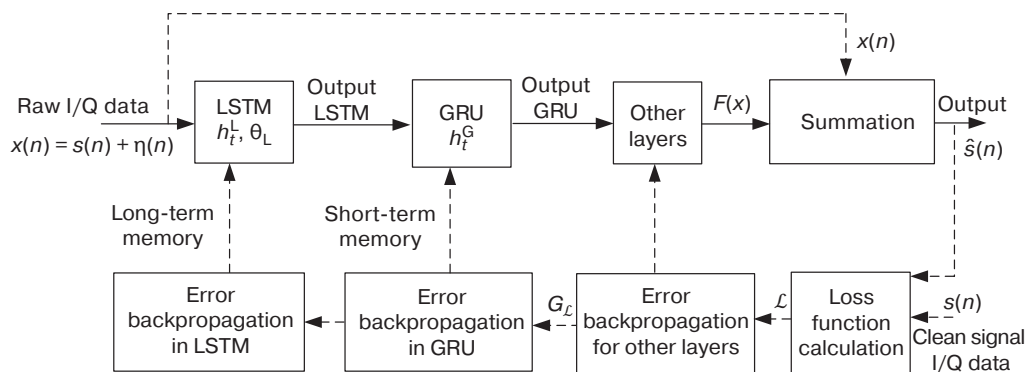


Fig. 2. Sequential Hybrid LSTM-GRU architecture:

$x(n)$ is the noisy signal, $s(n)$ is the clean signal, $\hat{s}(n)$ is the reconstructed signal, h_t^L, h_t^G are the hidden states of the LSTM and GRU networks, respectively, θ_L is the set of LSTM parameters, and $F(x)$ represents the transformation of the input signal through the network. The final output is produced using skip connection: $\hat{s}(n) = x(n) + F(x)$; $\mathcal{L}$ is the loss function, $G_{\mathcal{L}}$ is the loss gradient, n is the count, and t is the time step

The underlying I/Q signals exhibit multiscale temporal patterns and can also be considered statistically independent, since the white Gaussian noise component is uncorrelated with the others. During the training process, the network architecture (Fig. 2) enables the LSTM unit to capture primarily long-term signal components, while the GRU specializes in modeling short-term dynamics. From a physical perspective, the LSTM unit focuses on extracting characteristic signal patterns and multipath or fading effects, whereas the GRU unit mainly extracts information relating to the additive white noise component.

Therefore, the multiscale time structure and statistical independence of the I/Q signal components are necessary for the hybrid neural network to automatically allocate roles to the LSTM and GRU components without imposing any prior constraints.

2.3. Dataset

The study is based on the RadioML 2018.01A open DeepSig⁴ dataset. This dataset, which is widely used for radio signal modulation recognition tasks, includes 24 different modulation types, presented at 26 SNR levels ranging from -20 to $+30$ dB in 2 dB increments. A total of 4096 individual frames are included for each “modulation–SNR” pair. Each frame contains 1024 complex signal samples, which are represented as I/Q components with dimensions of (1024, 2). The total dataset comprises 2555904 frames.

To tailor the dataset for noise suppression, signals with an SNR of 30 dB are isolated and used as a reference for “clean” signals with minimal noise. Based on these, a noisy set is formed in which AWGN is added to each signal at SNR levels ranging from -10 to $+10$ dB in 2 dB increments according to a Gaussian distribution. Thus, two subsets are formed: clean modulated signals and noisy modulated signals.

The main parameters of the dataset are given in Table 4.

The dataset is divided into three parts for training, validation, and testing according to the proportions 80%, 10%, and 10%, respectively. Two models are trained using the training and validation sets: a Hybrid LSTM-GRU and a RaGAN. The test set is then used to compare the efficacy of these two models and evaluate the performance of the WT-based noise suppression methods.

Table 4. Dataset main parameters

Parameter	Value
Modulation types	8 types: BPSK ⁵ , QPSK ⁶ , OQPSK ⁷ , 8PSK ⁸ , 16QAM ⁹ , 64QAM ¹⁰ , GMSK ¹¹ , and OOK ¹²
SNR ranges	11 values ranging from -10 to 10 dB in 2 dB increments
Number of samples in one frame	1024
Number of frames per “modulation–SNR” pair	4096
Data format in the frame	1024 complex I/Q samples (in-phase and quadrature components), size (1024, 2)
Total number of frames	360448 frames ($8 \times 11 \times 4096$)
Communication channel	AWGN

3. EXPERIMENTAL RESULTS

To evaluate the effectiveness of the proposed ResBiLSTM-BiGRU hybrid architecture, experiments have been conducted using a test dataset. The performance of the model is compared to two representative methods, namely a deep RaGAN model and a conventional WT-based noise suppression method.

The RaGAN model uses a generative adversarial network with a relativistic discriminator based on bidirectional LSTM (BiLSTM) networks. We denote the original (real) signal as s_r , while the synthesized (fake) signal is denoted as s_f . The discriminator calculates the probability that the original data is more realistic than the synthesized data. The loss functions for the discriminator (D) and the generator (G) can be expressed as follows:

$$L_D^{\text{Ra}} = -E \left[\log \left(D_{\text{RaGAN}}(s_r, s_f) \right) \right] - E \left[\log \left(1 - D_{\text{RaGAN}}(s_f, s_r) \right) \right],$$

$$L_G^{\text{Ra}} = -E \left[\log \left(1 - D_{\text{RaGAN}}(s_r, s_f) \right) \right] - E \left[\log \left(D_{\text{RaGAN}}(s_f, s_r) \right) \right],$$

where $E[\cdot]$ denotes the expectation operator, and D_{RaGAN} is the relativistic discriminator used in RaGAN.

⁴ <https://www.deepsig.ai/datasets/>. Accessed November 11, 2025.

⁵ Binary Phase Shift Keying.

⁶ Quadrature Phase Shift Keying.

⁷ Offset Quadrature Phase Shift Keying.

⁸ 8-Phase Shift Keying.

⁹ 16-Quadrature Amplitude Modulation.

¹⁰ 64-Quadrature Amplitude Modulation.

¹¹ Gaussian Minimum Shift Keying.

¹² On-Off Keying.

To improve the quality of clean signal recovery, the overall loss function of the generator is formulated as a weighted sum of content loss and adversarial loss as follows:

$$L_G = L_{\text{con}} + \lambda L_G^{\text{Ra}},$$

where λ is the balancing coefficient. Content loss, which is based on MSE and absolute difference, is defined as:

$$L_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n \left(s_{r_i} - G(s_f)_i \right)^2, \quad L_1 = \frac{1}{n} \sum_{i=1}^n \left| s_{r_i} - G(s_f)_i \right|,$$

$$L_{\text{con}} = \frac{L_{\text{MSE}} + L_1}{2}.$$

This combination of loss functions allows the generator to accurately reconstruct the noisy signal while considering both the quality of content and competitive component.

The classical WT method operates in the frequency domain. The noisy signal $y(t)$ is decomposed into a set of wavelet coefficients:

$$c_{j,k} = \int y(t) \psi_{j,k}(t) dt,$$

where $\psi_{j,k}(t)$ are the basic wavelet functions at scale j and position k . Threshold filtering of coefficients is then applied:

$$\tilde{c}_{j,k} = \begin{cases} \text{sign}(c_{j,k}) \left(|c_{j,k}| - \lambda \right), & |c_{j,k}| > \lambda, \\ 0, & |c_{j,k}| \leq \lambda. \end{cases}$$

Then, the signal is reconstructed using the inverse wavelet transform, as follows:

$$\hat{x}(t) = \sum_{j,k} \tilde{c}_{j,k} \psi_{j,k}(t).$$

This approach enables the removal of high-frequency noise components by setting the corresponding coefficients to zero.

The efficacy of noise suppression was evaluated using standard quality recovery metrics, such as MSE, SNR, correlation coefficient (CC) and PSNR. The formulas for these metrics can be found in Table 5.

The model is trained using MSE loss function and the Adam¹³ optimizer for 150 epochs. During training, the adaptive reduction mechanism called Reduce learning rate on plateau or ReduceLROnPlateau, is employed, whereby the learning rate decreases if there is no improvement in the metrics over five consecutive epochs. To prevent overfitting and reduce calculation time, the Early stopping¹⁴ strategy is also employed,

Table 5. Metrics for evaluating the quality of noise reduction

Metric	Formula
Signal-to-noise ratio (SNR)	$\text{SNR} = 10 \cdot \lg \left(\frac{\sum_{i=1}^N x_i^2}{\sum_{i=1}^N (x_i - \hat{x}_i)^2} \right) \text{ dB},$ where x is the original clean signal, $\hat{x}$ is reconstructed signal, N is the number of samples
Mean squared error (MSE)	$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2$
Correlation coefficient (CC)	$\text{CC} = \frac{\sum_{i=1}^N (x_i - \bar{x})(\hat{x}_i - \bar{\hat{x}})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (\hat{x}_i - \bar{\hat{x}})^2}},$ where $\bar{x}$ are the average values of the original clean signal, $\bar{\hat{x}}$ are average values of the reconstructed signal
Peak signal-to-noise ratio (PSNR)	$\text{PSNR} = 10 \cdot \lg \frac{x_{\max}^2}{\text{MSE}}.$ The $x_{\max}$ parameter is defined as the maximum amplitude value of a clean one-dimensional signal, calculated from its discrete time realization in a test sample

which is triggered when the loss function stops decreasing. As shown in Fig. 3, the learning outcomes demonstrate a stable decrease in the loss function for both the training and test samples, as well as the model achieving stable and optimal error values.

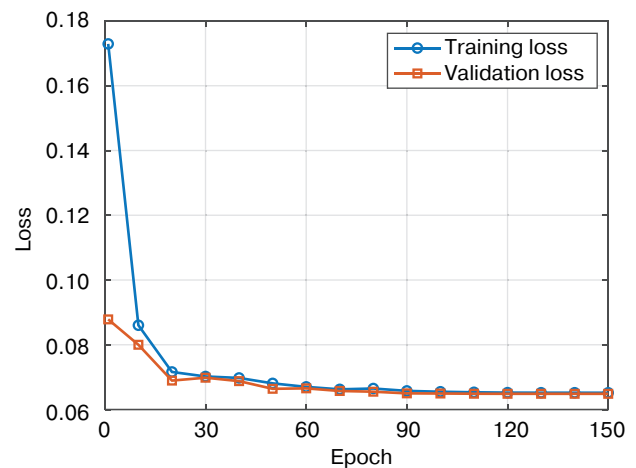


Fig. 3. Values of training and validation loss during the training process

¹³ Adaptive moment estimation is an optimization algorithm used for training neural networks that combines the concepts of the momentum method and adaptive gradient scaling.

¹⁴ Early stopping is a regularization method in which the learning process of a neural network is interrupted when the error on the validation data stops decreasing (and usually begins to increase), which prevents overfitting.

Figure 4a shows the results of comparing different noise suppression methods using the SNR metric. The proposed model delivers optimal performance across the entire input SNR range, spanning from -10 to $+10$ dB. Compared to RaGAN, the average improvement is approximately 2.27 dB. At an extremely low SNR level of -10 dB, the improvement is just 0.41 dB; however, it then gradually increases to stabilize within the range of 3.0–3.5 dB for high SNR values. Compared to WT, the advantage is even more pronounced, providing an average increase of 10.5 dB. In the low SNR range of -10 to 0 dB, the increase ranges from 6.1 to 12.3 dB. In the high SNR range (2 to 10 dB), the noise-suppression improvement remains between 11 and 13.5 dB.

Figure 4b shows the results of comparing the average MSE values. The proposed model consistently provides the lowest MSE, indicating more accurate signal reconstruction than RaGAN and WT. A particularly significant improvement can be observed compared to the WT; here, the MSE is 0.2556 at an SNR of -10 dB, whereas for WT it is 1.077. At an SNR of 10 dB, the MSE decreases to 0.0024, whereas WT remains at 0.0233 (almost ten times higher). Compared to RaGAN, the difference is less pronounced but still evident across the entire SNR range. For instance, at an SNR of -10 dB, the model's MSE is 0.255 vs 0.279 for RaGAN. At an SNR of 10 dB, the

respective values are 0.0024 and 0.0042, indicating approximately 1.7 times greater accuracy.

Figure 4c shows a comparison of the correlation coefficient between the original and reconstructed signals. The proposed model achieves the highest values across the entire SNR range, confirming its superior ability to preserve the temporal structure of the signal. At an SNR of -10 dB, the correlation coefficient is 0.726, which is higher than that of RaGAN (0.718) and almost double that of WT (0.366). As the SNR increases, the performance of all models improves; however, the proposed method consistently outperforms the others. At an SNR of 0 dB, the correlation coefficient reaches 0.980, surpassing the values obtained for RaGAN (0.971) and WT (0.822). At high SNR levels (>6 dB), all techniques approach the saturation region (>0.99 for the proposed method and RaGAN). Nevertheless, the proposed method continues to outperform WT (0.979 at 10 dB vs 0.998 for the former).

Figure 4d presents the PSNR comparison results. The proposed model exhibits the highest PSNR values over the entire SNR range, confirming its superior signal reconstruction capabilities. At an SNR of -10 dB, the PSNR is 11.72 dB, which exceeds the corresponding values for RaGAN (11.31 dB) and WT (3.47 dB). As the SNR increases, the gap between the proposed model and

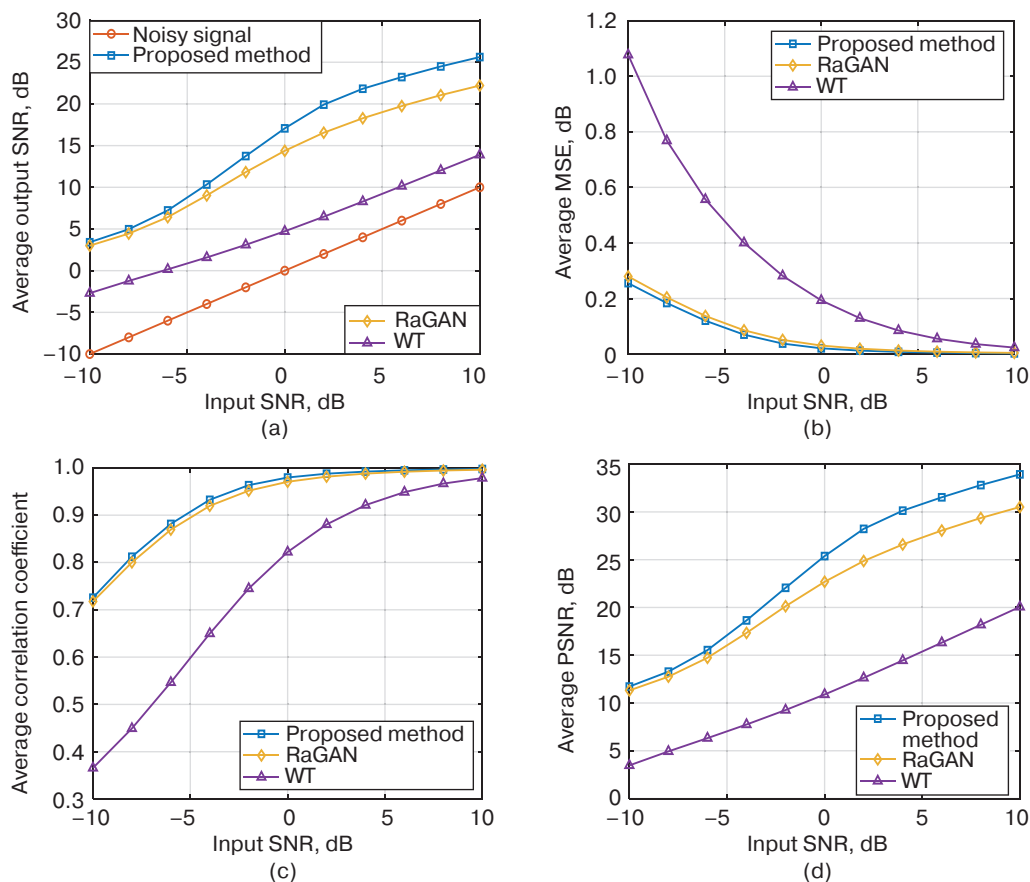


Fig. 4. Results of comparing noise suppression methods: (a) output SNR, (b) MSE, (c) correlation coefficient, and (d) PSNR

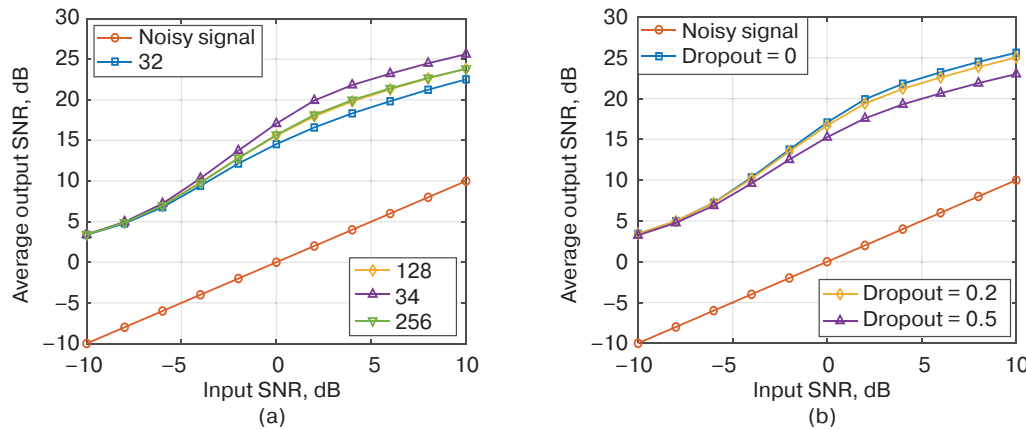


Fig. 5. Results of comparing noise suppression methods:
(a) by the number of hidden nodes and (b) by the dropout value

RaGAN remains consistent at around 2–3 dB. The benefit is particularly notable in comparison to WT: at 0 dB, the PSNR is 25.38 dB for the model vs 10.89 dB for WT; at 10 dB, the model's PSNR reaches 33.93 dB, which is nearly 14 dB greater than that of WT (20.06 dB).

The influence of the number of hidden layers in the ResBiLSTM-BiGRU hybrid model is illustrated in Fig. 5a. A significant improvement in signal reconstruction accuracy is achieved by increasing the number of layers from 32 to 64. For example, at an SNR of 0 dB, the performance metric increases from 14.54 dB to 15.59 dB. Further increasing the number of layers to 128 results in optimal performance, with an SNR of 0 dB yielding a value of 17.07 dB. This is 1.5 dB higher than the 64-layer configuration and 2.5 dB higher than the 32-layer configuration. However, no further

improvement is observed with 256 layers, at which configuration the performance remains comparable to that of the 64-layer model, and is lower than that of the 128-layer variant over the entire SNR range. These results suggest that 128 layers provide the best balance between reconstruction accuracy and model complexity.

Figure 5b shows the impact of the dropout rate¹⁵ on model performance. Without dropout, the model exhibits maximum efficiency. Setting the dropout value to 0.2 results in only a minor decrease in reconstruction accuracy while maintaining overall stability. However, increasing the dropout rate to 0.5 results in a significant reduction in performance across the entire SNR range.

Figure 6 depicts the results of comparing various combinations of the LSTM and GRU blocks. Using LSTM alone provides limited performance, for

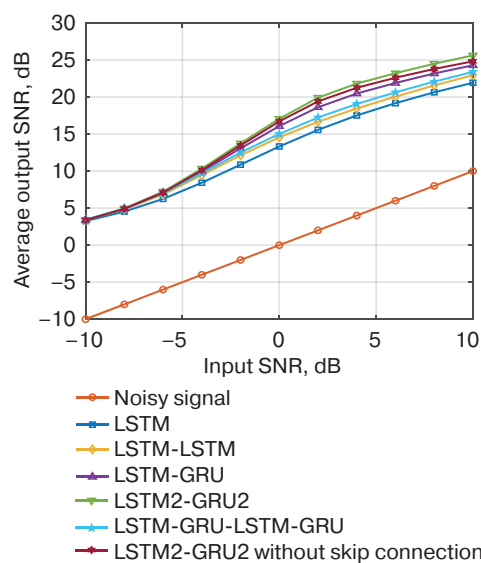


Fig. 6. Results of comparing performance for different versions of the hybrid LSTM-GRU architecture

¹⁵ A hyperparameter of the neural network regularization method that determines the probability of unintentional pruning of neurons during the training process in order to prevent overfitting.

example, 13.33 dB at an SNR of 0 dB. Adding the GRU component (the LSTM-GRU configuration) significantly improves signal reconstruction accuracy, achieving 16.08 dB at 0 dB. The best performance is demonstrated by deeper architectures, such as LSTM2-GRU2 with a skip connection (ResBiLSTM-BiGRU), which achieve maximum values of 17.07 dB at 0 dB and 25.62 dB, respectively, for all configurations considered. Although variants of the LSTM-GRU-LSTM-GRU and LSTM2-GRU2 configurations without a skip connection also perform well, they are inferior to the LSTM2-GRU2-Res architecture in terms of stability and absolute metric values across the entire SNR range.

Figure 7 shows the results of the proposed method for the following eight types of modulation: BPSK, QPSK, OQPSK, 8PSK, 16QAM, 64QAM, GMSK, and OOK. At low input SNRs, 64QAM, 16QAM, and 8PSK demonstrate the lowest efficiency, exhibiting a lower output SNR, PSNR, and correlation coefficient, while the MSE is higher. BPSK, GMSK, and OOK show the best results under the same

conditions. As the SNR increases, the difference between the modulations decreases; however, higher orders of QAM and 8PSK continue to demonstrate slightly lower efficiency.

Figure 8 shows some examples of time domain signal comparisons at different SNR levels after noise suppression methods have been applied.

CONCLUSIONS

A novel hybrid ResBiLSTM-BiGRU neural network architecture for noise suppression in digital modulation signals transmitted over an AWGN channel is proposed. Experimental results confirm that the proposed model achieves an average SNR reconstruction improvement of 2–3 dB compared to RaGAN, and an improvement of over 10 dB as compared to WT. The model also demonstrates the lowest MSE values and highest correlation coefficients across the entire input SNR range from –10 to 10 dB. The PSNR values additionally confirm the model's high signal reconstruction accuracy, particularly at low SNR levels.

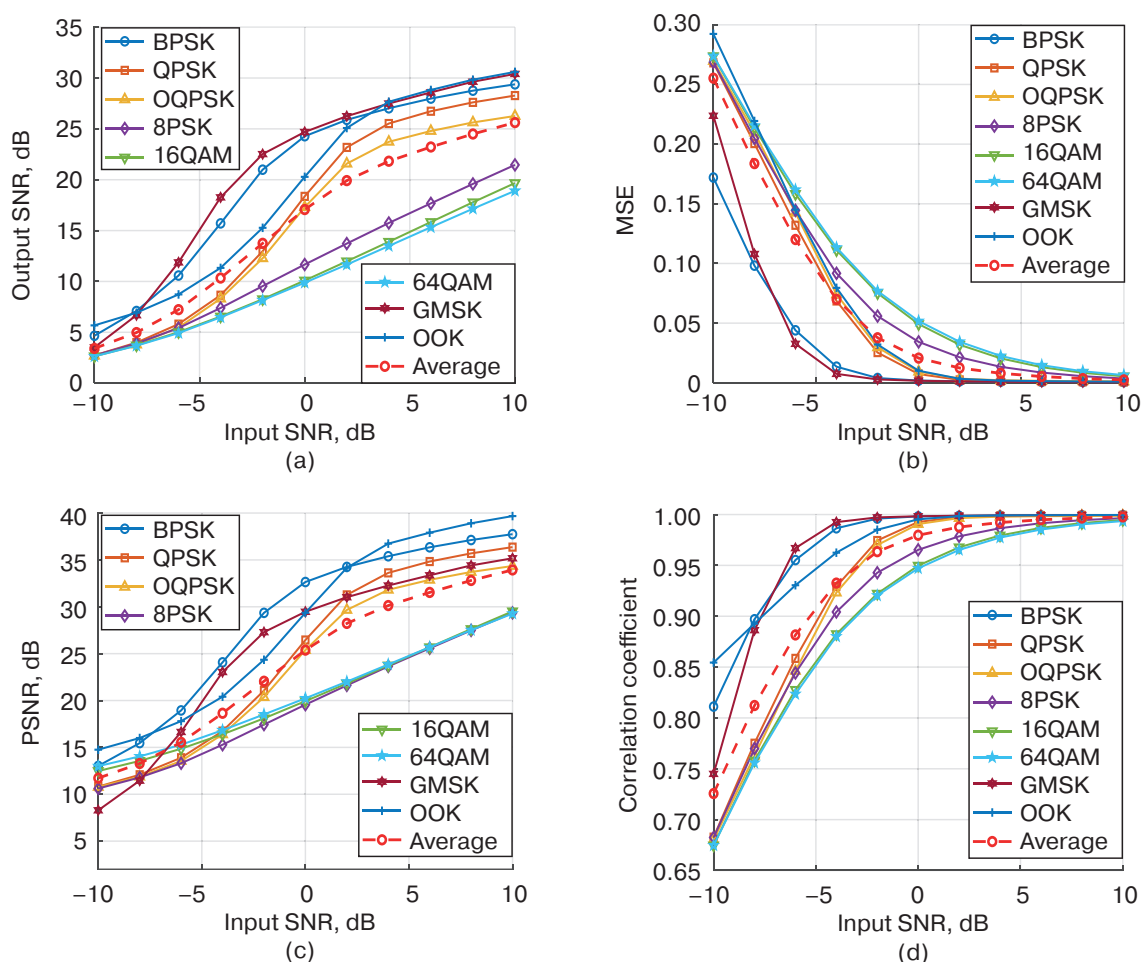


Fig. 7. Results of evaluating the accuracy of the proposed method for eight types of modulation: BPSK, QPSK, OQPSK, 8PSK, 16QAM, 64QAM, GMSK, and OOK for different SNR input levels: (a) output SNR, (b) MSE, (c) PSNR, and (d) correlation coefficient

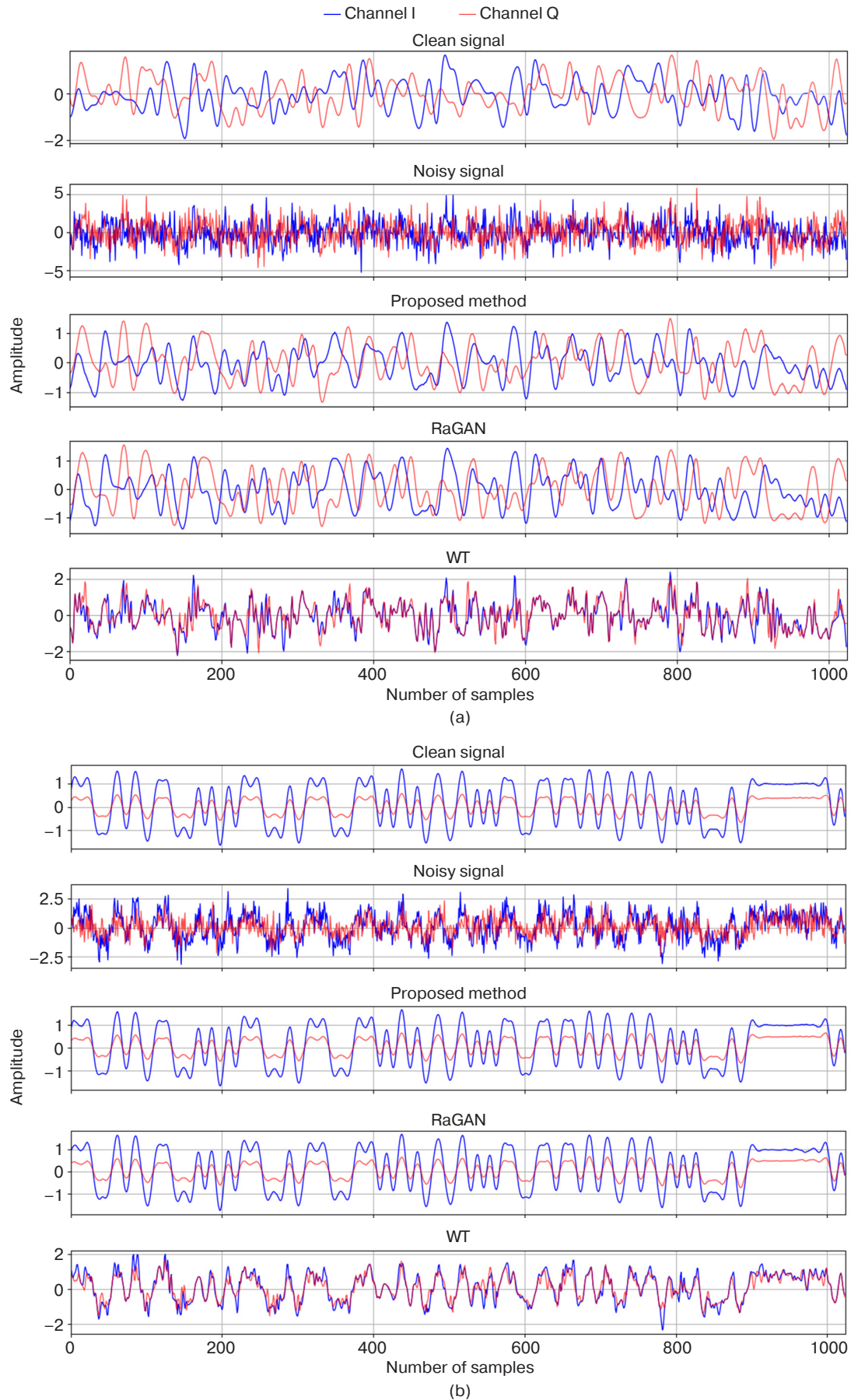


Fig. 8. Signals after noise suppression by various models for the modulated signal:
(a) 16QAM at SNR = -6 dB and (b) BPSK at SNR = 0 dB

Analysis of the parameters shows that the optimal configuration has 128 hidden nodes, while additional nodes increase the dropout rate leading to a significant reduction in accuracy. Among the studied architectures, the deep ResBiLSTM-BiGRU model demonstrates the highest overall efficiency. These results confirm the effectiveness and promising potential of the proposed ResBiLSTM-BiGRU hybrid model for suppressing additive Gaussian noise.

Future research will focus on integrating the noise reduction module with the automatic modulation classification system in order to improve the overall performance of a radio communication system.

Authors' contribution

All authors contributed equally to the research work.

REFERENCES

1. Wang W. Evaluation of Adaptive Filtering Algorithms in Signal Denoising: Insights from Simulation and Algorithmic Enhancement. *Theoretical and Natural Science*. 2025;101(1):61–69. <https://doi.org/10.54254/2753-8818/2025.CH22293>
2. Lei J., Zhou J., Jian Z., Liu H., Zhang L., Feng Y., Wu Z. A Wiener Filter Denoising Based Intelligent Modulation Recognition. In: *2022 IEEE/CIC International Conference on Communications in China (ICCC)*. IEEE. 2022. P. 88–93. <https://doi.org/10.1109/ICCC55456.2022.9880789>
3. Guo T., Zhang T., Lim E., López-Benítez M., Ma F., Yu L. A Review of Wavelet Analysis and Its Applications: Challenges and Opportunities. *IEEE Access*. 2022;10:58869–58903. <https://doi.org/10.1109/ACCESS.2022.3179517>
4. Akkaya S. Wavelet-Based Denoising Strategies for Non-Stationary Signals in Electrical Power Systems: An Optimization Perspective. *Electronics*. 2025;14(16):3190. <https://doi.org/10.3390/electronics14163190>
5. Facciolo G., Pierazzo N., Morel J.-M. Conservative Scale Recomposition for Multiscale Denoising. (The Devil is in the High Frequency Detail). *SIAM J. Imaging Sci.* 2017;10(3):1603–1626. <https://doi.org/10.1137/17M1111826>
6. Wahab M.F., Gritti F., C. O'Haver T. Discrete Fourier transform techniques for noise reduction and digital enhancement of analytical signals. *TrAC Trends in Analytical Chemistry*. 2021;143(4):116354. <https://doi.org/10.1016/j.trac.2021.116354>
7. Zhao S., Ma L., Xu L., Liu M., Chen X. A Study of Fault Signal Noise Reduction Based on Improved CEEMDAN-SVD. *Appl. Sci.* 2023;13(19):10713. <https://doi.org/10.3390/app131910713>
8. Wang H., Yao Y., Zhang G., Pan J., Gao L., Jin H., Wang C. Enhanced Noise Suppression in Partial Discharge Signals via SVD and VMD with Wavelet Thresholding. *Modelling and Simulation in Engineering*. 2024;2024:5676986. <https://doi.org/10.1155/2024/5676986>
9. Almazrouei E., Gianini G., Almoosa N., Damiani E. A Deep Learning Approach to Radio Signal Denoising. In: *2019 IEEE Wireless Communications and Networking Conference Workshop (WCNCW)*. IEEE. 2019. <https://doi.org/10.1109/WCNCW.2019.8902756>
10. Duan R., Chen Z., Zhang H., Wang X., Meng W., Sun G. Dual Residual Denoising Autoencoder with Channel Attention Mechanism for Modulation of Signals. *Sensors*. 2023;23(2):1023. <https://doi.org/10.3390/s23021023>
11. He S., Zhu L., Yao C., Wang L., Qin Z. A Novel Approach Based on Generative Adversarial Network for Signal Enhancement in Wireless Communications. *Wireless Communications and Mobile Computing*. 2022. <https://doi.org/10.1155/2022/7050573>
12. Peng L., Fang S., Fan Y., Wang M., Ma Z. A Method of Noise Reduction for Radio Communication Signal Based on RaGAN. *Sensors*. 2023;23(1):475. <https://doi.org/10.3390/s23010475>
13. Yao H., Zhou Q., Chen Z., Huang L., Xu D., Xuan Q. Denoising of Radio Modulation Signal Based on Deep Learning. In: Xuan Q., Xiang Y., Xu D. (Eds.). *Deep Learning Applications*. 2023. Ch. 7. P. 125–143. https://doi.org/10.1142/9789811266911_0007
14. Wu T., Chen Z., He D., Qian L., Xu Y., Tao M. CDDM: Channel Denoising Diffusion Models for Wireless Semantic Communications. *IEEE Trans. on Wireless Commun.* 2024;23(9):11168–11183. <https://doi.org/10.1109/TWC.2024.3379244>
15. Letafati M., Ali S., Latva-Aho M. Conditional Denoising Diffusion Probabilistic Models for Data Reconstruction Enhancement in Wireless Communications. *IEEE Trans. on Machine Learning in Communications and Networking*. 2025;3:133–146. <https://doi.org/10.1109/TMLCN.2024.3522872>
16. Ji Z., Wang S., Yang K., Zhang Q., Ye P. Transfer Learning Guided Noise Reduction for Automatic Modulation Classification. In: *2025 IEEE 102nd Vehicular Technology Conference (VTC2025-Fall)*. IEEE. 2025. <https://doi.org/10.1109/VTC2025-Fall65116.2025.11310099>
17. Chung J., Gulcehre C., Cho K.H., Bengio Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv preprint. arXiv:1412.3555*. 2014. <http://arxiv.org/abs/1412.3555>
18. O'Shea T.J., Roy T., Clancy T.C. Over-the-Air Deep Learning Based Radio Signal Classification. *IEEE J. Selected Topics in Signal Process.* 2018;12(1):168–179. <https://doi.org/10.1109/JSTSP.2018.2797022>

About the Authors

Nguyen Van Dung, Cand. Sci. (Eng.), Lecturer, Department of Circuit Theory – Measurement, Faculty of Radio-Electronic Engineering, Le Quy Don Technical University (236, Hoang Quoc Viet, Ha Noi, Viet Nam). E-mail: nguyenvandungvtdt@lqdtu.edu.vn. <https://orcid.org/0000-0002-2810-1204>

Phan Trong Hanh, Cand. Sci. (Eng.), Lecturer, Department of Circuit Theory – Measurement, Faculty of Radio-Electronic Engineering, Le Quy Don Technical University (236, Hoang Quoc Viet, Ha Noi, Viet Nam). E-mail: maitrongphan@gmail.com. <https://orcid.org/0009-0006-2093-7150>

Об авторах

Нгуен Ван Зунг, к.т.н., преподаватель, кафедра теории цепей – измерения, Факультет радиотехники и электроники, Вьетнамский государственный технический университет им. Ле Куй Дона (Вьетнам, Ханой, ул. Хоанг Куок Вьет, д. 236). E-mail: nguyenvandungvtdt@lqdtu.edu.vn. <https://orcid.org/0000-0002-2810-1204>

Фан Чонг Хань, к.т.н., преподаватель, кафедра теории цепей – измерения, Факультет радиотехники и электроники, Вьетнамский государственный технический университет им. Ле Куй Дона (Вьетнам, Ханой, ул. Хоанг Куок Вьет, д. 236). E-mail: maitrongphan@gmail.com. <https://orcid.org/0009-0006-2093-7150>

Translated from Russian into English by K. Nazarov

Edited for English language and spelling by Thomas A. Beavitt

Modern radio engineering and telecommunication systems
Современные радиотехнические и телекоммуникационные системы

UDC 537.874+537.877+621.37

<https://doi.org/10.32362/2500-316X-2026-14-4-96-105>

EDN TAFVVJ



RESEARCH ARTICLE

Tracking-based formation of vector radio images

Sergey S. Shadinov[@],
Mihail S. Kostin

MIREA – Russian Technological University, Moscow, 119454 Russia[@] Corresponding author, e-mail: shadinov@mirea.ru

• Submitted: 26.12.2025 • Revised: 06.02.2026 • Accepted: 26.05.2026

Abstract

Objectives. The aim of this article is to investigate the possibility of forming vector radio images of small-size objects (SSO) using a digital vector network analyzer (VNA) with a controllable bandwidth and frequency step of the tracking generator.

Methods. The work is based on the comprehensive application of the following methods: analysis of radio wave processes in the context of radio vision, the apparatus of N -pole theory, numerical methods of statistical radio engineering, and computer modeling (including both electromagnetic and functional-block simulation).

Results. A new method for vector tracking-formation of signal radio images of SSO has been developed. Its fundamental difference from traditional pulsed methods of near-field radio vision lies in the use of an S -parametric model of a spatial N -pole to describe the inhomogeneous medium, as opposed to approaches based on scalar signal transformation. To analyze the accuracy of the method, a model of operator ψ -transformation has been proposed, allowing for the evaluation of projective distortions in vector radio images of SSO signatures. Verification of the method was carried out on a custom-developed software-hardware complex based on the R&S ZNLE6 VNA. The complex provides coherent processing and verification of signal radio images of SSOs recorded under the influence of non-fluctuating interference.

Conclusions. During the research, a method for vector tracking-based formation of radio images with spectro-temporal unfolding of the power spectral density function was proposed. This method is free from the limitations of pulsed analogs and can be used as one of the solutions for verifying SSO in near-range radio vision. The foundation of the proposed method can be a model of an inhomogeneous medium as a spatially-oriented system of coupled N -poles, described by S -parameters. The practical effectiveness has been experimentally confirmed: the probability of verifying a SSO reaches at least 0.94 (Neumann–Pearson criterion) with a signal-to-noise ratio ≥ 15 dB and a signal-to-interference ratio ≥ 9 dB. These results are of practical value for the development of near-field radio vision technologies.

Keywords: vector radio image, radio identification, small-size object, object signature, radiogenome

For citation: Shadinov S.S., Kostin M.S. Tracking-based formation of vector radio images. *Russian Technological Journal*. 2026;14(4):96–105. <https://doi.org/10.32362/2500-316X-2026-14-4-96-105>, <https://elibrary.ru/TAFVVJ>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Трекинг-формирование векторных радиоизображений

С.С. Шадинов[@],
М.С. Костин

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[@] Автор для переписки, e-mail: shadinov@mirea.ru

• Поступила: 26.12.2025 • Доработана: 06.02.2026 • Принята к опубликованию: 26.05.2026

Резюме

Цели. Целью статьи является исследование особенностей формирования векторных радиоизображений малоразмерных объектов (МРО) при помощи цифрового векторного анализатора цепей (ВАЦ) с управляемыми полосой и шагом перестройки частоты трекинг-генератора.

Методы. Работа основана на комплексном применении следующих методов: анализ радиоволновых процессов в контексте радиовидения, теория N -полюсников, численные методы статистической радиотехники и компьютерное моделирование (включая электродинамическое и функционально-блочное).

Результаты. Разработан новый метод векторного трекинг-формирования сигнальных радиоизображений МРО. Его принципиальное отличие от традиционных импульсных методов ближнего радиовидения заключается в использовании S -параметрической модели пространственного N -полюсника для описания неоднородной среды вместо подходов, основанных на скалярном сигнальном преобразовании. Для анализа точности метода предложено операторное Ψ -преобразование, позволяющее оценить проекционные искажения векторных радиоизображений сигнатур МРО. Верификация метода осуществлена на созданном программно-аппаратном комплексе на базе ВАЦ R&S ZNLE6 (производитель – Rohde & Schwarz, Германия). Комплекс обеспечивает согласованную обработку и верификацию сигнальных радиоизображений МРО, регистрируемых в условиях воздействия нефлуктуационных помех, а также аддитивного белого гауссовского шума.

Выводы. В ходе исследования предложен метод векторного трекинг-формирования радиоизображений со спектрально-временной разверткой функции спектральной плотности мощности, свободный от ограничений импульсных аналогов, который может быть использован в качестве одного из решений верификации МРО в ближнем радиовидении. Основой предложенного метода может служить модель неоднородной среды как пространственно-ориентированной системы связанных N -полюсников, описываемой S -параметрами. Практическая эффективность подтверждена экспериментально: вероятность верификации МРО достигает не менее 0.95 (критерий Неймана – Пирсона) при отношении сигнал/шум ≥ 15 дБ и отношении сигнал/помеха ≥ 9 дБ. Результаты представляют практическую ценность для развития технологий ближнего радиовидения.

Ключевые слова: векторное радиоизображение, радиоидентификация, малоразмерный объект, сигнатура объекта, радиогеном

Для цитирования: Шадинов С.С., Костин М.С. Трекинг-формирование векторных радиоизображений. *Russian Technological Journal*. 2026;14(4):96–105. <https://doi.org/10.32362/2500-316X-2026-14-4-96-105>, <https://elibrary.ru/TAFVVJ>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

The signal verification process in short-range radio vision systems designed for small-sized objects (SSOs) is influenced by several factors. Key aspects include the selection of authenticators, system configuration, and post-processing algorithms. Since the authenticators that define the object's properties manifest as specific deformations in its scattered field spectrum, the primary step in creating a radio image involves analyzing the reflected field structure and identifying features within the radio signal image. For probing purposes, either a pulse signal or a signal with a linearly varying power spectral density (PSD) function can be employed [1].

The use of signals having a gradual (linear) frequency-time sweep has advantages over the classical pulse method. These include high stability of signal phase synchronization, flexible vector and scalar scaling, the ability to selectively analyze the object's response to different frequencies while suppressing interference, high matched filtering efficiency, and the ability to adapt the shape of the probing signal to the object's specified impulse response (IR) [2].

The relevance of the study consists in the potential use of the described method to construct SSO radio images in complex, dispersive, reactive, and nonlinear environments.

Since the material environment and the radio vision object that generate complex radiophysical distortions can be considered as an N -pole with an S -matrix, the proposed technique takes into account the contribution of each harmonic to the synthesized radio image in accordance with a given feature cluster. Consequently, vector tracking-based formation of radio images with spectral-temporal resolution becomes an efficient solution for noise-resistant short-range radio vision, ensuring the effective detection and verification of SSO signal radio images in the presence of non-fluctuating interference and noise [3].

While traditional radio vision systems utilize scalar signals, such as ultrashort or linearly modulated radio pulses, vector signals, which are represented by an array of amplitudes and phases, allow the temporal profile of a radio image to be reconstructed using a polynomial. By controlling the frequency tuning step of the tracking generator, compressed spectral sensing techniques can be applied to reduce the number of measurements per sample while maintaining information. This presents new opportunities for addressing radio vision challenges, including those related to priority radio guidance systems [4]. This paper therefore proposes generating vector radio vision signals in the form of S -parameters from a heterogeneous environment. The proposed S -parameter method offers several benefits, including the ability to facilitate the radio identification

of small objects in the presence of noise and non-fluctuating interference based on the coherent reception of the radio vision signal. Several studies conducted by foreign researchers have demonstrated the need for in-phase reception using MIMO SAR¹ technology when processing deflated radio images [5]. Most research in the field of radio vision focuses on 2D visualization methods for radio images of objects [6], based on multi-position scanning [7]. The method proposed in the present paper enables the generation of signal radio images in a single-position configuration without raster scanning by obtaining an S_{21} vector profile [8].

1. S -PARAMETRIC N -POLE SYSTEM CORRESPONDING TO SSO

The proposed method involves representing SSO and its surrounding environment (the scene) as an N -pole system. This system can be analyzed using the theory of N -poles and S -parameters. The example of a quadcopter-type SSO, whose radio signatures are analyzed further in this paper, are shown in Fig. 1.



Fig. 1. Quadcopter-type SSO when capturing a radio image

In this case, the non-uniform system consists in a four-pole network comprising two antennas connected to a vector network analyzer (VNA) and the SSO. This network is defined by a radio image derived from the parameter S_{21} , which can be described by the following analytical expression:

$$\dot{S}_{21} = \frac{\dot{U}_{\text{Ref}}(n)}{\dot{U}_{\text{Tx}}(n)}, \quad (1)$$

where $\dot{U}_{\text{Ref}}(n) = \sum_{n=1}^N \sum_{m=1}^M k_n A_{\text{Tx}} \exp(j(\omega_n t + \varphi_{nm})) * h_a * h_{\text{Env}}$ is the signal received after being scattered by the

¹ Multiple Input Multiple Output Synthetic Aperture Radar is a combination of two technologies: MIMO (Multiple Input Multiple Output), which uses multiple input and output transmission channels; and SAR (Synthetic Aperture Radar), which employs a synthetic aperture antenna.

SSO signature; $\dot{U}_{Tx}(n) = \sum_{n=1}^N A_{Tx} \exp(j(\omega_n t + \varphi_n)) * h_a$

is the transmitted signal; A_{Tx} is the signal amplitude; $\omega_n = 2\pi f_n$ is an angular frequency of the n th harmonic of the synthesized radio vision signal; φ_n is the phase of the n th harmonic; h_a is the antenna IR; h_{Env} is the environment IR; $N = 200$ is the number of harmonics in the synthesized signal; M is the number of elementary reflecting planes forming the instantaneous signature of the object; $*$ is mathematical convolution; and k_n is the attenuation decrement specified for each discrete value of the formant as follows:

$$k_n = \sqrt{\frac{P_{Tx} D_1 D_2 \lambda_n^2}{(4\pi R)^2}}, \quad (2)$$

where P_{Tx} is the power of the tracking signal formant; λ_n is the wavelength of the n th harmonic of the synthesized signal; D_1 and D_2 are directivity coefficients of the transmitting and receiving antennas, respectively; and R is the distance from the transmitting antenna to the signal reception point.

The discrete variation in the frequency of the synthesized signal within a given range is described by summation over n in (1), while its interaction with the SSO is described by summation over m . However, despite the presence of phase information in (1), fundamental ambiguity arises in the form of phase uncertainty. In the absence of attenuation, adding a phase component that is a multiple of 2π to the sum does not alter the value of the entire expression.

Therefore, to verify the SSO, it is necessary to capture and process radio images of SSO signatures (that is, radio images of the SSO at a specified angular position) in order to create a database of reference radio images. In this case, the size of the sample and the accuracy of the database determine the probability of recognition according to a given SSO verification criterion.

2. OPERATOR-BASED Ψ -TRANSFORM OF VECTOR RADIOIMAGES

Several studies by foreign researchers have proposed a method for compensating for amplitude attenuation of the radio vision signal in order to correct distortions in the resulting radio images [9]. The amplitude-phase operator presented in the present paper compensates for distortions in radio images caused by the linear displacement of an object perpendicular to the radio vision antenna. Since slight fluctuations on the order of the wavelength lead to significant distortion of radio images, the position of the SSO relative to the antenna array is an important parameter when recording radio images of SSO signatures [10]. Due to the labor-intensive nature

of recording SSOs at a multitude of distances from the antenna array, a Ψ -transformation operator is proposed, which uses amplitude and phase operators to significantly reduce the required number of SSO reference radio images. The amplitude operator, $A\{\cdot\}$, which models natural radio wave attenuation, is defined by a set of coefficients calculated using attenuation decrement (2). The phase operator $\Psi\{\cdot\}$ determines the phase shift for each n th harmonic component of the tracking signal. The initial distance from the SSO to the antenna's phase center is F , while the wavelength λ_1 for the harmonic with frequency f_1 is defined as c/f_1 , where c is the speed of light. Each m th elementary plane of the object contributes its own phase contribution, φ_m , upon signal reflection. Thus, the resulting phase φ_n for the n th harmonic with frequency f_n consist in the superimposition (sum)

$$\varphi_n = \varphi_{m=1} + \varphi_{m=2} + \dots + \varphi_{m=M} = \sum_{m=1}^M \varphi_m \text{ of all these}$$

contributions. The general geometry of the problem for an arbitrary distance F is shown in Fig. 2.

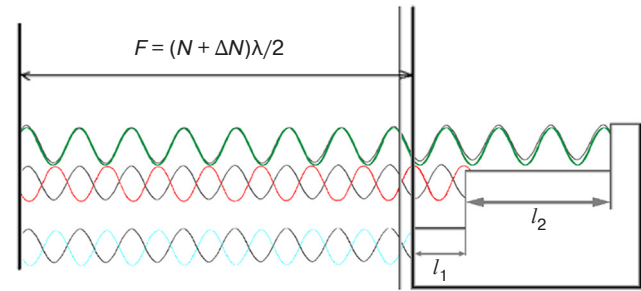


Fig. 2. Harmonics reflected from the SSO elementary planes; l_1 and l_2 are distances between planes in the SSO signature

For a harmonic with frequency f_n , the distance F is a multiple of the wavelength ($F = N\lambda_n$). Since the phase factor $2\pi N$ is an integer that does not alter the value of the total phase φ_n , it can be omitted. Consequently, the phase φ_n , which is determined by the object's geometry (i.e., the arrangement of the elementary planes), is invariant to the term $2\pi N$ and thus provides information about the SSO shape.

Consider the displacement and corresponding distortion of the signature of a radio vision object located at a distance of $F + \Delta F$ (Fig. 3).

In this case, the displacement is defined as $\frac{\lambda}{2}(N + \Delta N) + \frac{\lambda}{2}(N_1 + \Delta N_1)$, where ΔN_1 is the residual displacement for the distance ΔF , $\Delta N = \frac{\varphi}{2\pi} \leq 1$,

while $\varphi = \Delta N 2\pi$ is the phase at the initial distance F . The additional phase at the new distance $F + \Delta F$ for each

elementary plane of the object is then defined as $\Delta\varphi = \Delta N_1 2\pi$. The resulting phase is defined as

$$\varphi_n = \sum_{m=1}^M \varphi_m + M\Delta\varphi. \quad \text{The number of elementary}$$

planes M into which the topology of the SSO surface is divided is not a fixed parameter. Rather, its value depends on the signal frequency range: as the wavelength decreases (i.e., as the frequency increases), smaller surfaces are analyzed, leading to an increase in M . Figure 4 illustrates this dependence.

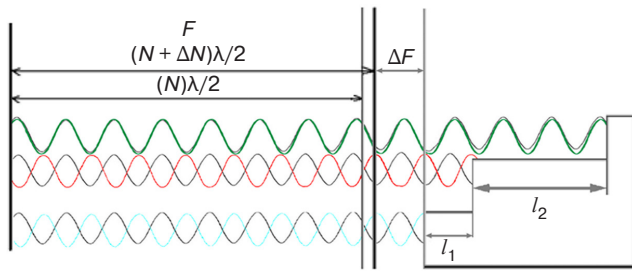


Fig. 3. Reflection from the SSO topological inhomogeneities for the harmonic when the distance changes by ΔF

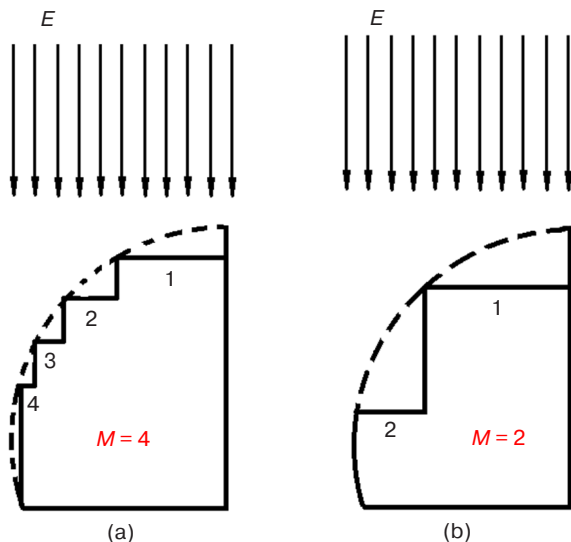


Fig. 4. Discretization of the MRO surface at different wavelengths ($\lambda_1 > \lambda_2$); (a) $\lambda_1 = c/f_1$; (b) $\lambda_2 = c/f_2$; E is intensity of the incident electric field

Since the synthesized probe signal consists of a sum of harmonics that are separated in time, the number of elementary planes M is a function of frequency. The value of M is calculated for each spectral component f_n according to its wavelength [11]. Thus, for the n th harmonic of frequency f_n , we denote the number of planes considered as M_n .

Since the value of $\Delta\varphi M_n$ usually exceeds 2π , the remainder from the division $\Delta\varphi_R = \frac{\Delta\varphi M_n}{2\pi R}$, where R is

an integer, is important. The resulting phase when the object is located at a distance $F + \Delta F$ from the antenna

$$\text{array is therefore } \varphi_n = \sum_{m=1}^M \varphi_m + \Delta\varphi_R.$$

The value of the elementary plane M_n for the n th harmonic is directly proportional to its wavelength, while the number of M_n increases as the frequency rises. Therefore, the frequency grid $\{f_n\}$ specified by the algorithm uniquely determines the corresponding set of values $\{M_n\}$. To perform the calculation, it is sufficient to determine the base value M_1 from the low-dimensionality condition at frequency f_n .

Although the implemented method does not allow for the direct measurement of the parameter M_n , it is possible to record the number of phase jumps corresponding to the desired value of M_n when recording the time-domain signal.

Thus, to obtain the transformed radio image from the reference image $\langle \hat{S}_{21} \rangle$ from the reference $\hat{S}_{21}$ when the antenna is displaced by a distance F along its axis of symmetry, the operator composition $\langle \hat{S}_{21} \rangle = A\{\Psi\{\hat{S}_{21}\}\}$ is applied.

3. HARDWARE-SOFTWARE SYSTEM FOR REAL-TIME RADAR VISION BASED ON VNA

The radar vision system (Fig. 5) consists of two HE400UW antennas having a range of 0.03–6.0 GHz, as well as a ZNLE6 VNA that has a range of 0.001–6.0 GHz (manufactured by Rohde & Schwarz, Germany).

The peaks in the antenna radiation patterns are mutually aligned at an angle of 60° with a distance of 1 m between the phase centers of the antennas. The VNA is connected to the control computer via Ethernet. The computer runs custom software, whose block diagram is shown in Fig. 6 in the *Simulink* environment.

The generated reference radio images are stored in a database (Block 1). The SSO radio image under study is received via Block 3 from VNA and fed into Block 2 to calculate the correlation coefficient (CC). Next, the decision-making module (Block 4) outputs a reference radio image corresponding to the SSO under study. If, based on estimates of the signal-to-noise ratio (SNR) and the signal-to-interference ratio (SIR), the CC value exceeds the 0.9 threshold, the system identifies the SSO.

Figure 7 shows the examples of radio images of SSO (quadcopter) signatures taken from three different angles.

The decrease in amplitude observed in the images is due to frequency-dependent attenuation: high-frequency radio waves are attenuated more rapidly by the environment than low-frequency ones.

Figure 8 shows the VNA screen during real-time recording of an SSO signature radio image.

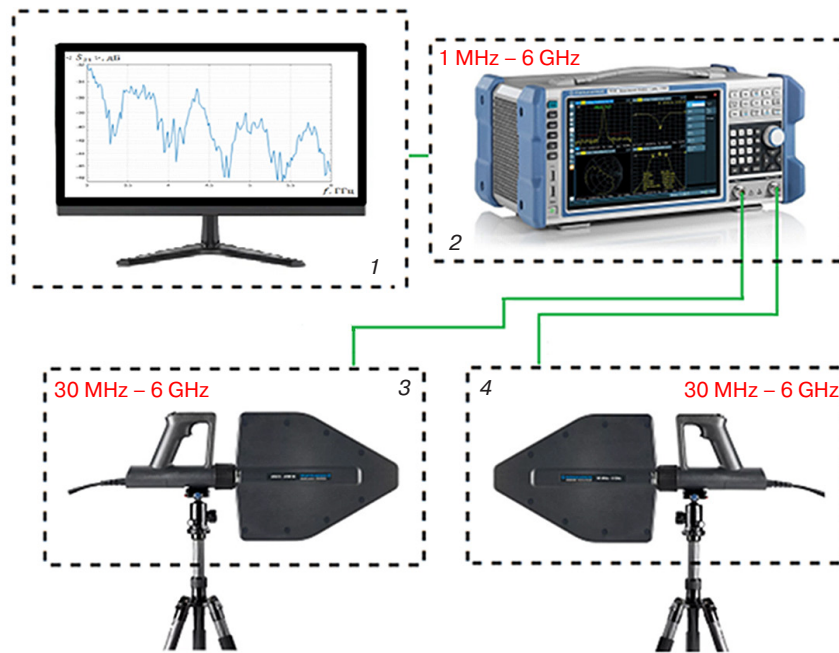


Fig. 5. Block diagram of the real-time radio vision system:
(1) software implementation in the *Simulink*² environment, (2) VNA,
(3) receiving antenna, and (4) transmitting antenna

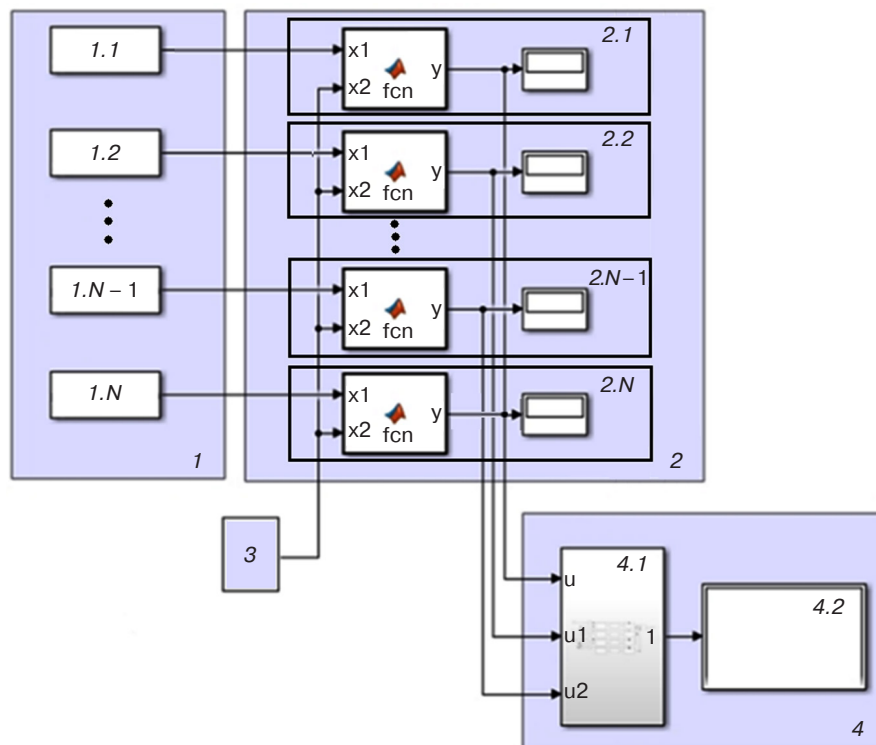


Fig. 6. Block diagram of the radio vision system software suite:
(1) reference radio image database block (1.1, ..., 1.N are reference radio images);
(2) correlation processing block (2.1, ..., 2.N are correlators with graphical output of the correlation coefficient);
(3) VNA vector data transmission block; (4) identification block (4.1 is the block for finding maximum CC;
4.2 is the module for outputting the name of the corresponding reference radio image); x_1 and x_2 are inputs to the FCN block, in which the correlation of two radio images is compared; u , u_1 , and u_2 are inputs of the block 4.1

² <https://www.mathworks.com/products/simulink.html>. Accessed May 22, 2026.

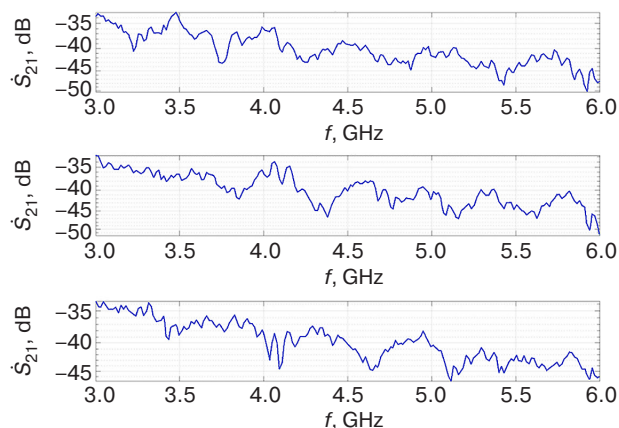


Fig. 7. Radio images of quadcopter-type SSO signatures at three scanning angles: (1) 0°, (2) 45°, and (3) 90°

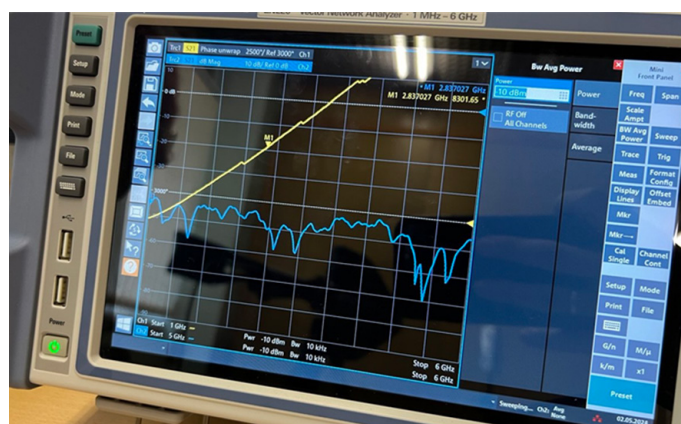


Fig. 8. VNA screen during real-time recording of SSO signature radio images

The radio image of the SSO signature is displayed on the VNA screen as the lower curve. The upper curve shows the phase run-up with the characteristic phase jumps that correspond to the reflection of the radio vision signal from the elementary planes of the SSO signature.

4. EVALUATION OF THE REPRODUCIBILITY OF SSO VECTOR RADIO IMAGES

A series of experiments involving 30 consecutive radio image acquisitions for selected signatures from three SSOs was conducted to evaluate the effectiveness of the method. Imaging performed at three different angles for each object resulted in a sample set of 270 radio images. The proportion of successful identifications was then calculated for each “SSO–angle” combination in the sample. A radio image is considered successful if its CC with the reference image exceeded the 0.9 threshold on the Chaddock scale. Thus, the ratio of successful identifications to the total number of trials was determined for each angular position of each SSO.

According to the Neyman–Pearson criterion, the method’s effectiveness, as averaged over the

experimental sample for the three SSO signatures and satisfying the $CC \geq 0.9$ condition, is 95%.

The method’s immunity to interference is evaluated by analyzing the performance of coherent reception when applied to a tracking signal. This is its advantage over pulse-based radio vision [12]. Due to the harmonic nature of the signal, for which the frequency and phase are known at every instant, synthesizing a matched filter becomes a standard problem. The filter’s frequency response corresponds to the signal spectrum (up to the coefficient), while its inverse frequency response is the mirror image having a delay equal to the signal duration [13]. Figure 9 shows the spectral representation $|U(f)|$ of the signal accumulated PSD ($N = 200$ harmonics) and interference.

A comparative assessment of the noise immunity of two PSD formation methods is conducted (Fig. 10): impulse (curve 1) and time-frequency sweep of the PSD (curve 2) [14]. To this end, the graphs depict the dependence of SNR and SIR on CC of random signatures in the presence of additive white Gaussian noise and non-fluctuating band interference [15].

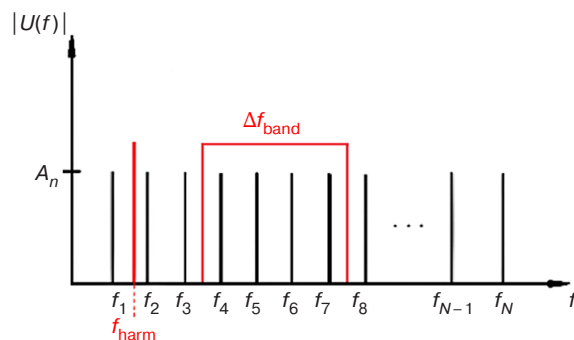


Fig. 9. Synthesized PSD of the signal and PSD of interference: harmonic f_{harm} and band Δf_{band}

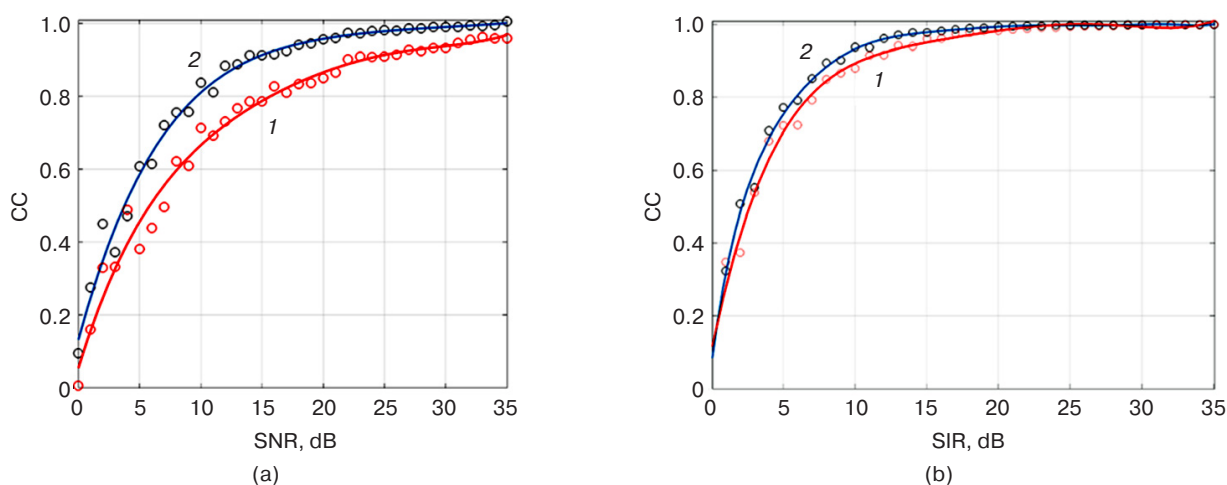


Fig. 10. Distribution of the radio image signature CC as a function of SNR (a) and SIR (b)

The graphs show that, due to coordinated reception, the method for vector tracking-based formation of signal radio image has an advantage of 8 and 5 dB, respectively, compared to pulse radio imaging at $CC = 0.9$. This ensures SSO verification at $SNR \geq 15$ dB and $SIR \geq 9$ dB for the time-frequency sweep method, and at $SNR \geq 23$ dB and $SIR \geq 14$ dB for the pulse-generated PSD method.

CONCLUSIONS

Scientific and practical results regarding the applicability of the tracking-based formation method for verifying vector radio images of SSO signatures have been presented. This was achieved using a radio visualizer laboratory prototype based on a digital VNA (R&S ZNLE6) and two HE400UWB ultra-wideband antennas. The prototype implements a semi-automatic processing cycle in *Simulink* with real-time correlation analysis. The scientific and practical results obtained are as follows:

- The developed method for vector tracking-based formation of signal radio images overcomes the

limitations inherent of scalar approaches based on impulse spectrum formation.

- The representation of a heterogeneous radio wave propagation environment has been validated as an S -parameter system of coupled N -poles, taking into account their spatial orientation.
- The proposed algorithm for the operator Ψ -transform establishes the dependence of projection distortions in vector radio images of signatures on the spatial position of the object.
- It is confirmed that the recognition of SSO signatures based on vector radio images ensures SSO verification at $SNR \geq 15$ dB and $SIR \geq 9$ dB with a defined $CC = 0.9$.
- Experimental data confirms reproducible SSO recognition at the 95% level when using a sample of $N \geq 200$ harmonics in coherent radio vision reception at $SNR \geq 15$ dB and $SIR \geq 9$ dB.

Authors' contribution

The authors' contributions to the preparation of the theoretical section and the conduct of the experiments were distributed equally.

REFERENCES

1. Shadinov S.S. Method of scanning spectral-temporal sweep for obtaining radio images. *Oboronnyi kompleks – nauchno-tekhnicheskomu progressu Rossii = Defense Industry Achievements – Russian Scientific and Technical Progress*. 2024;4(164):57–60 (in Russ.). <https://elibrary.ru/bmhwsc>
2. Shadinov S.S., Kostin S.S., Konyashkin G.V., Korchagin A.S., Romanovskii M.Y., Gusein-zade N.G. Vector S-parametric analysis of signal phase dynamic radio images. *Doklady Rossiiskoi akademii nauk. Fizika, tekhnicheskie nauki = Doklady Physics*. 2023;512(1):78–86 (in Russ.). <https://doi.org/10.31857/S2686740023050115>
3. Shadinov S.S., Konyashkin G.V. *Program for Capturing and Recognizing Vector Radio Images*: Certificate of State Registration of Computer Program 2025614644 RF. Publ. 25.02.2025 (in Russ.).
4. Kostin M.S., Boikov K.A. Digital technologies of signal radio vision and radio monitoring. *Russian Technological Journal*. 2024;12(4):59–69. <https://doi.org/10.32362/2500-316X-2024-12-4-59-69>
5. Song S., Lu J., Xing S., et al. Near Field 3D Millimeter-Wave SAR Image Enhancement and Detection with Application of Antenna Pattern Compensation. *Sensors*. 2022;22(1):4509. <https://doi.org/10.3390/s22124509>
6. Xie Q., Quan X., He B. A Remote Sensing and Meteorological Data-Based Methodology for Wildfire Danger Assessment for China. In: *IGARSS 2020 – 2020 IEEE International Geoscience and Remote Sensing Symposium*. 2020. P. 6798–6801. <https://doi.org/10.1109/IGARSS39084.2020.9324063>
7. Li X., Li D., Liu H., Wan J., Chen Z., Liu Q. A-BFPN: An Attention-Guided Balanced Feature Pyramid Network for SAR Ship Detection. *Remote Sens*. 2022;14(15):3829–3847. <https://doi.org/10.3390/rs14153829>
8. Papadopoulos K., Jelali M. A Comparative Study on Recent Progress of Machine Learning-Based Human Activity Recognition with Radar. *Appl. Sci*. 2023;13(23):12728–12761. <https://doi.org/10.3390/app132312728>
9. Kulikov G.V., Dang X.Kh. Influence of quadrature transformation imbalance on the noise immunity of signal reception with amplitude-phase shift keying. *Russian Technological Journal*. 2024;12(1):59–68. <https://doi.org/10.32362/2500-316X-2024-12-1-59-68>
10. Yanik M.E., Wang D., Torlak M. Development and Demonstration of MIMO-SAR mmWave Imaging Testbeds. *IEEE Access*. 2020;8:126019–126038. <https://doi.org/10.1109/ACCESS.2020.3007877>
11. Chen H., Long Z., Niu L., et al. Millimeter-wave SFCW SAR imaging system based on in-phase signal measurement with simplified transceiver. *Opt. Express*. 2020;28(2):1526–1538. <https://doi.org/10.1364/oe.380266>
12. Kim S., Krieger G., Villano M. Volume Structure Retrieval Using Drone-Based SAR Interferometry with Wide Fractional Bandwidth. *Remote Sens*. 2024;16(8):1352. <https://doi.org/10.3390/rs16081352>
13. Manzoni M., Tebaldini S., Monti-Guarnieri A.V., et al. Automotive SAR imaging: potentials, challenges, and performances. *Int. J. Microwave Wireless Technol*. 2023;16(1):3–12. <https://doi.org/10.1017/S1759078723000326>
14. Bolkhovskaya O., Maltsev A., Sergeev V. The wavefront estimation and signal detection in multi-element antenna arrays at low SNR. In: *Proceedings 2018 2nd European Conference on Electrical Engineering and Computer Science (EECS)*. 2018. P. 497–501. <https://doi.org/10.1109/EECS.2018.00097>
15. Bolkhovskaya O.V., Maltsev A.A., Sergeev V.A. A passive system for source detection and distance measurement based on signal wavefront estimation. *Radiotekhnika = J. Radioengineering*. 2022;86(9):98–112 (in Russ.). <https://doi.org/10.18127/j00338486-202209-11>

СПИСОК ЛИТЕРАТУРЫ

1. Шадинов С.С. Метод сканирующей спектрально-временной развертки для получения радиоизображений. *Оборонный комплекс – научно-техническому прогрессу России*. 2024;4(164):57–60. <https://elibrary.ru/bmhwsc>
2. Шадинов С.С., Костин С.С., Коняшкин Г.В., Корчагин А.С., Романовский М.Ю., Гусейн-заде Н.Г. Векторный S-параметрический анализ сигнальных фазодинамических радиоизображений. *Доклады Российской академии наук. Физика, технические науки*. 2023;512(1):78–86. <https://doi.org/10.31857/S2686740023050115>
3. Шадинов С.С., Коняшкин Г.В. *Программа захвата и распознавания векторных радиоизображений*: Свидетельство о государственной регистрации программы для ЭВМ 2025614644 РФ; опубл. 25.02.2025.
4. Костин М.С., Бойков К.А. Цифровые технологии сигнального радиовидения и радиомониторинга. *Russian Technological Journal*. 2024;12(4):59–69. <https://doi.org/10.32362/2500-316X-2024-12-4-59-69>
5. Song S., Lu J., Xing S., et al. Near Field 3D Millimeter-Wave SAR Image Enhancement and Detection with Application of Antenna Pattern Compensation. *Sensors*. 2022;22(1):4509. <https://doi.org/10.3390/s22124509>
6. Xie Q., Quan X., He B. A Remote Sensing and Meteorological Data-Based Methodology for Wildfire Danger Assessment for China. In: *IGARSS 2020 – 2020 IEEE International Geoscience and Remote Sensing Symposium*. 2020. P. 6798–6801. <https://doi.org/10.1109/IGARSS39084.2020.9324063>
7. Li X., Li D., Liu H., Wan J., Chen Z., Liu Q. A-BFPN: An Attention-Guided Balanced Feature Pyramid Network for SAR Ship Detection. *Remote Sens*. 2022;14(15):3829–3847. <https://doi.org/10.3390/rs14153829>
8. Papadopoulos K., Jelali M. A Comparative Study on Recent Progress of Machine Learning-Based Human Activity Recognition with Radar. *Appl. Sci*. 2023;13(23):12728–12761. <https://doi.org/10.3390/app132312728>
9. Куликов Г.В., Данг С.Х. Влияние разбаланса квадратурного преобразования на помехоустойчивость приема сигналов с амплитудно-фазовой манипуляцией. *Russian Technological Journal*. 2024;12(1):59–68. <https://doi.org/10.32362/2500-316X-2024-12-1-59-68>

10. Yanik M.E., Wang D., Torlak M. Development and Demonstration of MIMO-SAR mmWave Imaging Testbeds. *IEEE Access*. 2020;8:126019–126038. <https://doi.org/10.1109/ACCESS.2020.3007877>
11. Chen H., Long Z., Niu L., et al. Millimeter-wave SFCW SAR imaging system based on in-phase signal measurement with simplified transceiver. *Opt. Express*. 2020;28(2):1526–1538. <https://doi.org/10.1364/oe.380266>
12. Kim S., Krieger G., Villano M. Volume Structure Retrieval Using Drone-Based SAR Interferometry with Wide Fractional Bandwidth. *Remote Sens*. 2024;16(8):1352. <https://doi.org/10.3390/rs16081352>
13. Manzoni M., Tebaldini S., Monti-Guarnieri A.V., et al. Automotive SAR imaging: potentials, challenges, and performances. *Int. J. Microwave Wireless Technol*. 2023;16(1):3–12. <https://doi.org/10.1017/S1759078723000326>
14. Bolkhovskaya O., Maltsev A., Sergeev V. The wavefront estimation and signal detection in multi-element antenna arrays at low SNR. In: *Proceedings 2018 2nd European Conference on Electrical Engineering and Computer Science (EECS)*. 2018. P. 497–501. <https://doi.org/10.1109/EECS.2018.00097>
15. Болховская О.В., Мальцев А.А., Сергеев В.А. Система пассивного обнаружения и измерения расстояния до источника на основе оценки волнового фронта сигнала. *Радиотехника*. 2022;86(9):98–112. <https://doi.org/10.18127/j00338486-202209-11>

About the Authors

Sergey S. Shadinov, Cand. Sci. (Eng.), Associate Professor, Department of Radio Wave Processes and Technologies, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: shadinov@mirea.ru. <https://orcid.org/0009-0002-3798-5723>

Mihail S. Kostin, Dr. Sci. (Eng.), Professor, Head of the Department of Radio Wave Processes and Technologies, Deputy Director, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: kostin_m@mirea.ru. Scopus Author ID 57208434671, RSCI SPIN-code 5819-2178, <http://orcid.org/0000-0002-5232-5478>

Об авторах

Шадинов Сергей Семенович, к.т.н., доцент, кафедра радиоволновых процессов и технологий, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: shadinov@mirea.ru. <https://orcid.org/0009-0002-3798-5723>

Костин Михаил Сергеевич, д.т.н., профессор, заведующий кафедрой радиоволновых процессов и технологий, заместитель директора Института радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: kostin_m@mirea.ru. Scopus Author ID 57208434671, SPIN-код РИНЦ 5819-2178, <http://orcid.org/0000-0002-5232-5478>

Translated from Russian into English by K. Nazarov

Edited for English language and spelling by Thomas A. Beavitt

Micro- and nanoelectronics. Condensed matter physics
Микро- и нанoeлектроника. Физика конденсированного состояния

UDC 538.945

<https://doi.org/10.32362/2500-316X-2026-14-4-106-115>

EDN TTKFOC



RESEARCH ARTICLE

Photoemission modeling of a two-liquid charge-carrier system formed at strong electron–phonon coupling

Stanislava V. Doronkina ^{1, @}, Anastasia S. Babayants ²,
Ivan V. Lisitsa ³, Anna E. Myasnikova ²

¹ MIREA – Russian Technological University, Moscow, 119454 Russia

² Southern Federal University, Rostov-on-Don, 344090 Russia

³ Rostov State Medical University, Rostov-on-Don, 344022 Russia

@ Corresponding author, e-mail: doronkina@mirea.ru

• Submitted: 11.01.2026 • Revised: 28.01.2026 • Accepted: 03.06.2026

Abstract

Objectives. The nature of the “waterfalls” observed in the photoemission spectra of cuprate high-temperature superconductors is explained by calculating the angle-resolved photoemission spectra of systems with significant electron–phonon interaction and dispersion law characteristic of cuprates. The resulting spectrum is compared with the results of experiments on superconducting cuprates.

Methods. Photoemission spectra obtained within a two-component model, comprising a liquid of large-radius bipolarons and a Fermi liquid of delocalized charge carriers, are calculated based on Fermi’s golden rule using the bipolaron and polaron binding energies, polarization field energy, and wave functions of charge carriers in them. The approach is based on previous findings that such a two-liquid system forms the ground and weakly excited states of systems with strong Fröhlich electron–phonon interaction.

Results. Angle-resolved photoemission spectra of systems with strong electron–phonon interaction and cuprate-like dispersion were calculated using the parameters of a two-fluid charge-carrier system, which were obtained by minimizing the free energy of the system. Features of the resulting spectra observed in the experimental spectra of cuprate superconductors with the corresponding doping level provide a basis to analyze their physical causes.

Conclusions. The model used for the calculations is characterized by different energies of relaxation following the photoemission from the self-trapped state and delocalized state, which coexist and divide the momentum space in accordance with the Pauli exclusion principle under long-range strong electron–phonon interaction. This energy difference, which leads to the formation of so-called waterfalls in the angle-resolved photoemission spectra (ARPES spectra) in the nodal direction, also determines their heights. The wave vector of the waterfall is associated with the equilibrium size of bipolarons at given doping level and temperature by the uncertainty principle, as well as the change in this wave vector with varying doping level that agrees with the phenomena observed in the experimental spectra.

Keywords: cuprate superconductors, strong electron–phonon interaction, large-radius bipolaron, photoemission spectroscopy

For citation: Doronkina S.V., Babayants A.S., Lisitsa I.V., Myasnikova A.E. Photoemission modeling of a two-liquid charge-carrier system formed at strong electron–phonon coupling. *Russian Technological Journal*. 2026;14(4):106–115. <https://doi.org/10.32362/2500-316X-2026-14-4-106-115>, <https://elibrary.ru/TTKFOC>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Моделирование фотоэмиссии двухжидкостной системы носителей заряда, формирующейся при сильном электрон-фононном взаимодействии

С.В. Доронкина ^{1, @}, А.С. Бабаянц ²,
И.В. Лисица ³, А.Э. Мясникова ²

¹ МИРЭА – Российский технологический университет, Москва, 119454 Россия

² Южный федеральный университет, Ростов-на-Дону, 344090 Россия

³ Ростовский государственный медицинский университет, Ростов-на-Дону, 344022 Россия

@ Автор для переписки, e-mail: doronkina@mirea.ru

• Поступила: 11.01.2026 • Доработана: 28.01.2026 • Принята к опубликованию: 03.06.2026

Резюме

Цели. С целью объяснения природы «водопадов», наблюдаемых в фотоэмиссионных спектрах купратных высокотемпературных сверхпроводников, в работе выполнен расчет фотоэмиссионных спектров с угловым разрешением систем с сильным электрон-фононным взаимодействием (ЭФВ) и законом дисперсии, характерным для купратов, и произведено сравнение полученного спектра с результатами экспериментов на сверхпроводящих купратах.

Методы. Фотоэмиссионные спектры рассчитаны на основе золотого правила Ферми при использовании энергии связи биполярона и полярона, энергии поля поляризации и волновых функции носителей заряда в них, полученных в рамках двухкомпонентной модели, включающей жидкость биполяронов большого радиуса и ферми-жидкость делокализованных носителей заряда. Ранее было показано, что такая двухжидкостная система образует основное и слабо возбужденные состояния систем с сильным фрэлиховским ЭФВ.

Результаты. Рассчитаны спектры фотоэмиссионной спектроскопии с угловым разрешением от систем с сильным ЭФВ и купратоподобной дисперсией при использовании параметров двухжидкостной системы носителей заряда, полученных минимизацией свободной энергии системы. Полученные спектры демонстрируют особенности, наблюдаемые в экспериментальных спектрах купратных сверхпроводников с соответствующим уровнем допирования, что позволяет разобраться в физических причинах этих особенностей.

Выводы. В использованной для расчета модели различная энергетическая цена релаксации системы после фотоэмиссии из автолокализованного и делокализованного состояний, которые при дальнедействующем сильном ЭФВ сосуществуют и делят импульсное пространство в соответствии с принципом Паули, приводит к возникновению «водопадов» в спектрах ARPES¹ в нодальном направлении. Высота «водопада» определяется различием в энергетической стоимости релаксации. Волновой вектор «водопада» связан с равновесным размером биполяронов при данном уровне допирования и температуре соотношением неопределенностей, изменение его величины с допированием согласуется с наблюдаемым в экспериментальных спектрах.

¹ Angle-resolved photoemission spectroscopy – фотоэлектронная спектроскопия с угловым разрешением.

Ключевые слова: купратные сверхпроводники, сильное электрон-фононное взаимодействие, биполярон большого радиуса, фотоэмиссионная спектроскопия

Для цитирования: Доронкина С.В., Бабаянц А.С., Лисица И.В., Мясникова А.Э. Моделирование фотоэмиссии двухжидкостной системы носителей заряда, формирующейся при сильном электрон-фононном взаимодействии. *Russian Technological Journal*. 2026;14(4):106–115. <https://doi.org/10.32362/2500-316X-2026-14-4-106-115>, <https://elibrary.ru/TTKFOC>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Since the discovery of high-temperature superconductivity, both experimental and theoretical approaches for studying cuprate high-temperature superconductors (HTSCs) have been developed. However, the mechanism of charge-carrier pair formation, which is the fundamental cause of superconductivity in cuprates, remains unclear. Currently, there is no single, generally accepted explanation for the origin of multiple phases in cuprates. In particular, in hole-doped systems, up to seven different phases are observed, whose formation mechanisms (for example, the pseudogap and “strange metal” phases) do not all yet have a clear interpretation [1]. An explanation is also required for the properties of the normal state of cuprate HTSCs, including the appearance of an energy gap in the charge-carrier spectrum at temperatures exceeding the superconducting transition temperature T_c (critical temperature), different dependencies of the charge ordering period on the doping level in the charge density wave phase in various cuprate systems, a sharp change in the shape of the Fermi surface from small electron pockets to a large hole surface at a sixfold increase in its volume with an insignificant increase in the doping level, and the sign reversal of the Hall coefficient.

The explicit and strong dependencies of the properties of the system on the composition of the material are also yet to be satisfactorily explained: for example, T_c differs greatly among different cuprate families (for single-layer cuprate based on lanthanum, the critical temperature is $T_c = 40$ K; for HTSC based on yttrium, $T_c = 93$ K; and the maximum critical temperature ($T_c = 133$ K) is currently exhibited by a mercury-containing superconductor). Another unsolved problem is the existence of superconductivity in systems that differ in terms of their properties and structure. For example, while pnictides (iron-based high-temperature superconductors) contain a superconducting phase, their parent compound is a metal, rather than a dielectric. Meanwhile, the symmetry of the order parameter (i.e., the wave function of the Bose condensate) has s' -symmetry

in contrast to d -symmetry in cuprates. However, the existence of electron correlations in pnictides has not been unambiguously established [2].

Since comprising systems having strong electron correlations, cuprates have been conventionally studied using tight-binding models. However, researchers have noted [1] that these models cannot explain the full set of experimental data on cuprates [3]. In addition to electron correlations, an important factor accelerating progress in this direction concerns the influence of the crystal lattice on the states of carriers in HTSC materials [1]. This influence supplements the theory of superconductivity with the dependence of the properties of the system on the structure of the material [4], which, as described above, significantly affects the properties of cuprates. Thus, when taking into account the dependence of the properties of high-temperature superconductors on their structure, the electron–phonon interaction (EPI) should be considered.

1. CUPRATE MODEL USED IN CALCULATING ARPES SPECTRA

It is necessary to note the existence of two types of EPI: strong and weak. The strong interaction can arise either due to deformation of the crystal lattice, which changes the electric field and creates a deformation potential, or due to Coulomb repulsion between electrons. In the first case, localized states of carriers are formed, which are smaller in size than the unit cell. Such states are called small-radius polarons and bipolarons. In the second case, the Coulomb interaction leads to the formation of large-radius polarons and bipolarons, having sizes several times larger than the unit cell. This second type of interaction, which is referred to as long-range, is more suitable for descriptions of the properties of cuprate HTSCs due to their ionic crystal lattice, which is known to have high polarizability.

The presence of the strong EPI in cuprates is evidenced by a pronounced decrease in the dispersion of phonons near the wave vector of charge ordering [5], which is called “softening” of phonon modes. Another

of the most obvious manifestations of the strong EPI in cuprates is the giant broadening of bands in the angle-resolved photoemission spectra (ARPES spectra) [4]. The discrepancy between the short-range EPI model and the observed properties of cuprates lies in the different modes of motion of small and large bipolarons. Cuprate HTSCs are known to exhibit a relatively high conductivity even at low doping levels. However, small-radius bipolarons cannot move coherently; their motion along the crystal is only possible via thermally activated hops, during which a carrier leaves the polarization potential well at one site and forms it at an adjacent one. Unlike the case of large-radius bipolarons, small-radius bipolarons do not coexist with delocalized carriers due to the low probability of such hops and the small size of small-radius bipolarons. Therefore, systems with short-range interactions cannot exhibit the high conductivity that occurs in cuprates. For this reason, the most suitable model for describing the properties of cuprate HTSCs is the strong long-range EPI model.

It was previously shown [6] that the ground and weakly excited states of such a system constitute a two-component system of coexisting bipolarons of size much larger than the lattice constant and delocalized carriers. To calculate the parameters of this state in the model of a compressible bipolaronic liquid, we generalized the variational method to systems of interacting compressible self-trapped carriers coexisting with delocalized ones. The resulting doping–temperature phase diagram, charge ordering period, and transition temperature, as well as its dependence on the number of conducting layers in the unit cell, are in qualitative and quantitative agreement with experimental data on cuprates. The previously calculated ARPES spectra [7] are also in not only qualitative but also quantitative agreement with experimental results. However, since they were calculated before obtaining the parameters of the ground and weakly excited states of the two-fluid system of charge carriers [6], their calculations were based on the polaron and bipolaron radii, binding energies, and lattice deformation energies calculated for an isolated polaron and bipolaron. For this reason, it becomes necessary to check whether the agreement between the calculation results and the experimental data is maintained when using these quantities found by minimizing the free energy of the system [6]. It turns out that the radius of the polaron and bipolaron in the conducting plane and in the direction perpendicular to it, as well as the binding energy of the polaron and bipolaron and the energy of deformation of the crystal lattice in the polaron and bipolaron, differ quite strongly in a bipolaron liquid from the case of an isolated polaron and bipolaron in a medium with the same polarizability of the crystal lattice. This is due both to the strong anisotropy (quasi-two-dimensionality) of

the cuprates and the compression of the self-localized states compared to their size in the case of a single polaron or bipolaron.

It is important to note that the photoemission spectra of cuprates recorded in different directions in momentum space have different features [8]. In the antinodal direction, i.e., along the faces of the Brillouin zone, the experimental spectra exhibit the absence of spectral weight near the Fermi surface, or a pseudogap, whose width can reach 0.06 eV. Unlike a conventional superconducting gap, the pseudogap remains open in both the superconducting and normal states. However, the nature of the so-called pseudogap phase in the phase diagram of cuprates in the doping level–temperature coordinates remains a subject of discussion. Within the framework of the approach under consideration, a method for solving the Schrödinger equation, similar to the finite element method, was developed to find stationary states of delocalized charge carriers located in the additional bipolaron potential. This method was used to show that the pseudogap arises due to the absence of stationary states of charge carriers with average momenta near the antinodal directions [9].

In addition to the peak broadening discussed above, two characteristic features are observed in the nodal spectra of cuprates recorded in the direction of the Brillouin zone diagonal: a kink [10] in the low-energy part of the spectrum and a region of vertical dispersion, which is referred to both as a “waterfall” and as a high-energy anomaly [11–14]. The second feature was explained within the framework of the strong long-range EPI model as due the fact that the same system contains carriers in delocalized and localized states. These carriers, which divide the momentum space in accordance with the Pauli exclusion rule, give different energy contributions to the photoemission manifested in the experimental spectra as regions of vertical dispersion [7]. It was possible to calculate the wave vector of vertical dispersion, k_0 , by determining the half-width and position of the bands in the photoemission spectrum.

The magnitude of the wave vector k_0 is determined by the bipolaron R through the uncertainty principle as

$$k_0 = \frac{\sqrt{\pi}}{R} \quad [6], \text{ while the bipolaron radius is determined}$$

by the minimization of the free energy of the system. Minimization of the free energy of the system at low temperatures and low doping levels characteristic of superconducting cuprates showed that the system is virtually free of polarons [6]. However, since in such a system the compression of self-localized states with increasing carrier density is energetically favorable, a hole produced by photoemission of an electron will

form a hole polaron following relaxation to the minimum hole energy region if the doping is below the optimal value. At optimal and higher doping, further compression of self-localized states becomes energetically unfavorable, so the photohole remains delocalized.

The calculations used the permittivity values (static and high-frequency), which were taken from the optical spectra, and the effective mass of carriers, which was taken from the ARPES experiment data [15].

2. RESEARCH METHODS

Due to the complexity of the system under consideration, it was necessary to use a combination of several existing models and methods, as well as to further develop them. The carrier dispersion in cuprates can be determined directly from the experimental ARPES spectra based on the correspondence of the dispersion of the ARPES band in systems with strong EPI with the carrier dispersion without EPI. According to ARPES data [16], the dispersion of electrons and holes near the extrema in hole-doped cuprates is expressed by a parabolic function, while the Fermi surface consists of four circular arcs centered at the point $(\pm\pi/a, \pm\pi/a)$, where a is the lattice constant. The bipolaron binding energy and the energy of the deformed phonon vacuum in sparse and dense systems are calculated using the variational method that takes electron correlations into account in the effective mass approximation of the mean energy functional [15]. The state vector of the system is written in the adiabatic approximation using the coherent state representation for the phonon field [17]. The application of this representation allows for an analytical calculation of the high-energy part of the ARPES spectrum according to Fermi's golden rule [15]. The original theory of bipolarons considered the polarization field as a classical field with an energy equal to its average value (doubled binding energy [15]). As a result, the width of the predicted band in the optical absorption spectrum due to the photodissociation of polarons could be determined only by the wave function of the electron in the polaron [18]. The analytical calculation of broad Gaussian bands arising as a result of the generation of different numbers of phonons in different occurrences of photodecay of polarons with Poisson probability in optical conduction [19] and ARPES spectra took into account quantum fluctuations of the phonon field in a polaron on the basis of quantum coherent states [17, 20].

The ground and weakly excited states of the system are determined by minimizing its free energy over the radius of self-localized states [6]. The free energy is obtained using the charge-carrier distribution function

in systems with a strong Fröhlich EPI. This charge-carrier distribution function is constructed using the Gibbs method [21]. As noted above, its main feature consists in the possible coexistence of self-localized and delocalized carriers occupying different regions of the momentum space. As a result, photoemission from these different regions of the momentum space is characterized by different energies, which are calculated according to energy conservation equations, including those involving a change in the energy of the electron and phonon fields due to relaxation after photoemission.

The law of conservation of energy (the argument of the δ -function in the expression for the probability of transition from the initial to the final state of the system) relates the photoelectron energy (usually represented through the so-called binding energy) with the number of emitted phonons ν . For example, in the case where photoemission occurs from a delocalized state and the photohole remaining after the photoemission forms a hole polaron, $\hbar\Omega + E(k) = E_{\text{kin}} + \Phi + E_p + \nu\hbar\omega$, where $\hbar\Omega$ is the energy of a photon with frequency Ω , $E(k)$ is the dispersion law of the charge carrier, E_{kin} is the kinetic energy of the photoelectron, Φ is the work function, E_p is the binding energy of the resulting hole polaron, and $\nu\hbar\omega$ is the energy of phonons with frequency ω , respectively. In photoemission spectroscopy, the photoelectron binding energy is usually introduced as $\varepsilon = E_{\text{kin}} - (\hbar\Omega - \Phi)$. Then the law of conservation of energy takes the form $E(k) = \varepsilon + E_p + \nu\hbar\omega$. Along with the momentum of the charge carrier in the initial state, it contains only two variables, ε and ν , which allows the second to be expressed through the first and substituted into the probability P_ν of emitting ν phonons. Then, neglecting the dispersion of phonons, we can express the probability $A(k, \varepsilon)$ to detect a photoelectron with momentum k and binding energy ε in the form [19, 20]

$$A(k, \varepsilon) \propto |\psi_k|^2 P_{\nu(\varepsilon)}, \quad (1)$$

$$P_{\nu(\varepsilon)} = \sum_{\{v_{\vec{q}}\}=\nu} \prod_{\vec{q}} \left| \left\langle v_{\vec{q}} \mid d_{\vec{q}} \right\rangle \right|^2 = \frac{\bar{\nu}^{\nu(\varepsilon)-1}}{(\nu(\varepsilon)-1)!} e^{-\bar{\nu}}, \quad (2)$$

$$\nu(\varepsilon) = \frac{E_p - \varepsilon}{\hbar\omega}, \quad \bar{\nu} = \sum_{\vec{q}} |d_{\vec{q}}|^2 = \frac{\Delta E_{\text{pvd}}}{\hbar\omega}, \quad (3)$$

where ψ_k is the Fourier transform of the electron wave function in the initial state; $\bar{\nu}$ is the average number of phonons emitted as a result of relaxation of the system; ΔE_{pvd} is the of phonon vacuum deformation due to strong EPI; summation in expression (2) is over sets of $v_{\vec{q}}$, in which the sum of $v_{\vec{q}}$ over all harmonics of the phonon field is equal to ν ; and $d_{\vec{q}}$ is the average

value of the creation operator of a longitudinal optical phonon with energy $\hbar\omega$ and wave vector $\vec{q}$. In the case of photoemission of delocalized electron with the formation of a hole polaron, ΔE_{pvd} is equal to E_{pvd} minus the lattice deformation energy in the polaron.

In the case of photoemission of a carrier from an electron bipolaron, the relaxation of the system after the photoemission includes the transition of an electron from the Fermi surface to a self-localized state. Here, the excess kinetic energy of this electron is transferred to the phonon subsystem, leading to additional emission of an average $\bar{\nu} = \hbar^2 k_F^2 / (2m^* \hbar\omega)$ phonons, where k_F is the Fermi wave vector of the electron and m^* is the effective mass of the electron. The recovery of a bipolaron from a polaron is additionally accompanied by the emission of phonons, whose average number is $\bar{\nu}_2 \approx \bar{\nu}_1$, where $\bar{\nu}_1$ is the number of phonons emitted during the decay of the bipolaron polarization “coat” to a polaron one, since the initial and final states of the phonon vacuum are reversed for the decay and recovery cases. The average numbers of emitted phonons are summed up in accordance with the properties of the Gaussian distribution as follows:

$$\begin{aligned} \bar{\nu} &= \bar{\nu}_1 + \bar{\nu}_2 + \bar{\nu}' \cong 2\bar{\nu}_1 + \bar{\nu}' \cong \\ &\cong 2(E_{\text{pvd}}^{\text{b}} - E_{\text{pvd}}^{\text{p}}) / \hbar\omega + \hbar^2 k_F^2 / (2m^* \hbar\omega), \end{aligned} \quad (4)$$

where $E_{\text{pvd}}^{\text{b}}$ and $E_{\text{pvd}}^{\text{p}}$ are the phonon vacuum deformation energies of the bipolaron and polaron, respectively. The law of conservation of energy at doping above the optimal level and the photoelectron momentum $k < k_0$ has the form (of course, at low temperature):

$$E_{\text{b}} + 0 = E_{\text{b}} + \varepsilon + \nu \hbar\omega, \quad (5)$$

where E_{b} is the binding energy of the bipolaron. In the case of hole doping below the optimal value, the binding energy of the hole polaron formed by the hole emerging from the Fermi surface is added to the right-hand side of expression (5), while $E_{\text{pvd}}^{\text{p}} / \hbar\omega$ is added to the right-hand side of expression (4).

The coefficient 2 in expression (4) at the difference of the phonon vacuum deformation energies in the bipolaron and polaron corresponds to the case where part of the phonon “coat” of the bipolaron has completely decayed to a “coat” in the polaron. However, since both the decay of part of the polarization “coat” of the bipolaron and the relaxation of the electron from the Fermi surface to the momentum region $k < k_0$ corresponding to self-localized electrons are associated with phonon processes, their times are of the same order. In this case, the coefficient will be smaller than 2. Since calculating its value is quite difficult, we used the first of the two cases considered (systems with doping below and above the optimal) to determine the coefficient from the

experimental spectrum, and the second case as a test (the coefficient is the same in both cases). The obtained coefficient value 0.8 for doping above the optimal value demonstrated good agreement between the low doping case calculated with this coefficient and the experimental data. The low value of the coefficient is also consistent with the value measured recently in a dynamic study of X-ray spectra following resonant (through an optical slit) photoexcitation of the complex oxide BaBiO_3 , representing the parent compound of HTSC bismuthates. The observed rapid recovery of the state of the system after photoexcitation, while unexpected, should be accompanied by deformation of the crystal lattice [22].

3. RESULTS AND DISCUSSION

Figure shows the calculated angle-resolved photoemission spectra of strongly interacting electron–phonon systems with cuprate-like dispersion. Figures (a) and (b) illustrate the case of low doping ($p = 0.05$) and doping above the optimal value ($p = 0.17$), respectively. The figures were constructed using the results of minimization of the free energy of a system with the inverse effective permittivity

$$\frac{1}{\varepsilon^*} = \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} = 0.27 \quad (\text{where } \varepsilon_\infty = 3.45 \text{ is the high-}$$

frequency permittivity, and $\varepsilon_0 = 50$ is the static permittivity), the effective mass $m^* = 1.2m_e$ of electrons and holes near the minimum of their energies (where m_e is the mass of an electron), two conducting layers in the unit cell with a distance between the layers of $d = 4.1 \text{ \AA}$, and the “thickness” of a polaron and bipolaron of $h = 6 \text{ \AA}$. The values of $1/\varepsilon^*$ and ε_0 correspond to those observed in cuprates ($\varepsilon_0 = 30\text{--}50$ [23, 24], $\varepsilon_\infty = 3.5\text{--}4.0$ [25]). At the used parameters of the system, the calculated radius of the bipolaron and polaron at low $p = 0.05$ is $13 \text{ \AA}$, while at $p = 0.17$, it is $7.1 \text{ \AA}$.

Figure (a) is in good agreement with the experimental spectra in the nodal direction in the undoped parent compound $\text{Ca}_2\text{CuO}_2\text{Cl}_2$ [14], as well as in the lightly doped compound $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ [26] having the superconducting transition temperature $T_c = 5 \text{ K}$. Figure (b) is consistent with the experimental spectrum of the Bi2201 system in the nodal direction [11]. The difference of the bipolaron radii in the lightly and heavily doped system is manifested in a noticeable shift in the momentum of the “vertical dispersion” k_0 toward an increase with increasing doping (due to a decrease in the bipolaron radius). A similar shift is also clearly expressed in the experimental spectra. The position of the high-energy intensity maximum also undergoes a shift, here to the region of lower binding energy with increasing doping. The physical reason for this shift is the lack of the possibility of self-trapping of the photohole at doping

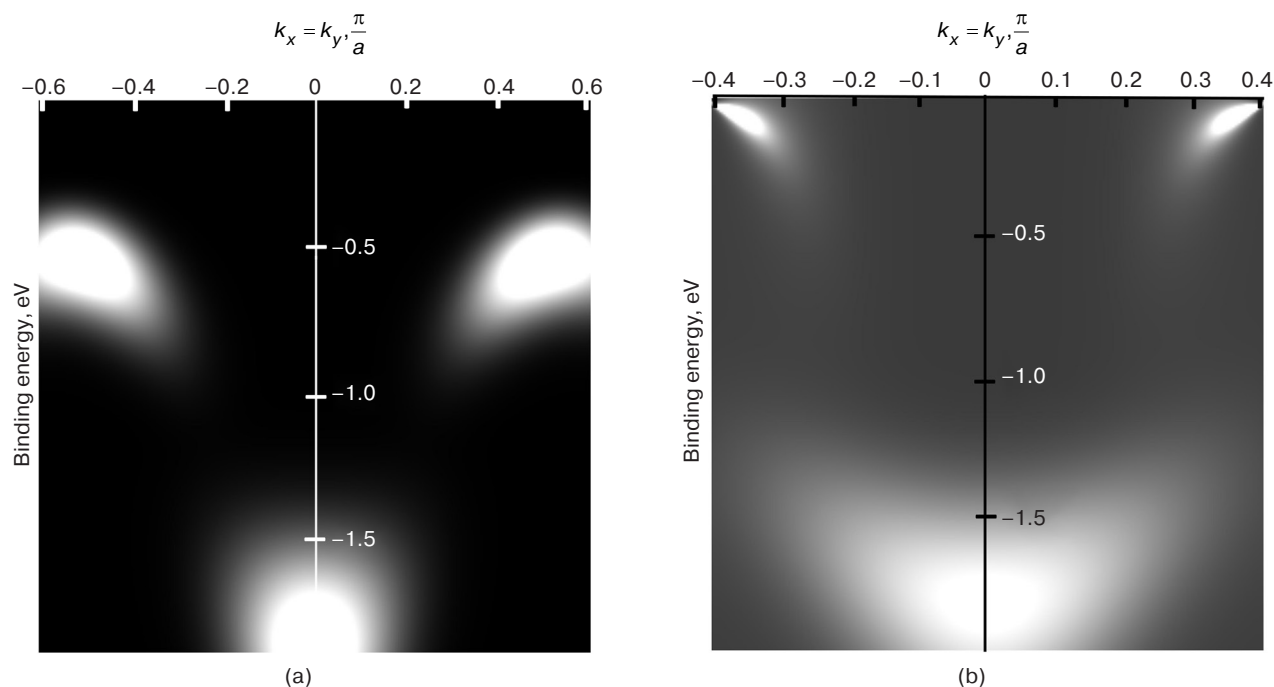


Figure. Calculated maps of the intensity of angle-resolved photoemission spectra of strongly interacting electron–phonon systems with cuprate-like dispersion at (a) low and (b) high doping levels. The horizontal axis shows the projections of the charge-carrier wave vector k_x and k_y in the copper–oxygen plane of cuprates along the nodal direction (the diagonal of the first Brillouin zone): a is the size of the unit cell of the crystal, $a = 3.86 \text{ \AA}$

levels higher than optimal, which is reflected in the absence of terms describing the hole polaron formation in expressions (4) and (5). The resulting coefficient value, which takes into account the degree of decay of the polarization “coat” during photoemission from the electron bipolaron state before the bipolaron is recovered due to relaxation of the system to an energetically more favorable state, indicates a comparable rate of decay of the polarization “coat” and electron relaxation from the Fermi surface to the momentum level of carriers in self-trapped states.

The model used for the calculation is characterized by the difference in the energies of relaxation of the system after photoemission from the self-trapped and delocalized states, which coexist under long-range strong EPI and divide the momentum space in accordance with the Pauli principle. This difference in the energies of relaxation leads to the appearance of “waterfalls” in the ARPES spectra in the nodal direction. It is this energy difference that determines the “waterfall” height. The wave vector k_0 , which characterizes the position of the “waterfall,” is determined by the equilibrium size of the bipolarons at a given doping level and temperature through the uncertainty principle. The change in k_0 with increasing doping level is consistent with that observed in the experimental spectra [11, 14, 26]. The width of the bands in the ARPES spectrum that arise with the

participation of self-trapped states is very large due to the production of a different number of phonons during the decay or change in the coherent state of the phonon field, which is also in agreement with the experimental data. The absence of intersection of the spectral density with the Fermi surface at low doping $p < 0.1$ [8] is due to the fact that a delocalized hole is not a stationary state at such doping; the stationary state is a hole polaron (at $p < 0.04$) or bipolaron (at $0.04 < p$) [6].

In the last two years, interest in theoretical modeling of the occurrence of high-energy features in the ARPES spectra of cuprates has been renewed due to the discovery of similar structures in the ARPES spectra of nickelates [27, 28], which relatively new high-temperature superconductors were discovered 7 years ago. A general model for the formation of “waterfalls” may be useful for developing a general theory of HTSC cuprates and nickelates. A recently proposed model [29] has a significant drawback: within this framework, vertical dispersion occurs only within a narrow range of model parameters, when the energy U of repulsion of two electrons at a copper (nickel) site is from $7t$ to $8t$, where t is the energy of an electron hopping to the nearest site. The narrow range of parameters that leads to the emergence of “vertical dispersion” is a consequence of its occurrence in the model [29] based on the local coincidence of the slope of the two dispersion curves: the curve for the dispersion of quasiparticles,

which is calculated in the Hubbard model using the quantum Monte Carlo method, and that associated with the dispersion of photoelectrons. However, the model considered in this paper, which is not limited by such narrow applicability boundaries, can be consequently applied to the calculation of photoemission of nickelates.

CONCLUSIONS

Thus, in this study, angle-resolved photoemission spectra were calculated for systems with strong EPI and cuprate-like dispersion using the parameters of a two-fluid charge-carrier system, calculated by minimizing the free energy of the system. The observed spectral features at low and high doping levels are consistent with experimental data on cuprates, thus providing a basis for

discussing the possible physical causes of these spectral features.

Authors' contributions

S.V. Doronkina—developing the programs and calculating the spectra, discussing results, and writing the text of the article.

A.S. Babayants—developing the programs and calculating the spectra, discussing results.

I.V. Lisitsa—developing the programs and calculating the spectra, discussing results.

A.E. Myasnikova—problem statement, discussing results.

REFERENCES

1. Keimer B., Kivelson S.A., Norman M.N., Uchida S., Zaanen J. From quantum matter to high-temperature superconductivity in copper oxides. *Nature*. 2015;518:179–186. <https://doi.org/10.1038/nature14165>
2. Stewart G.R. Superconductivity in iron compounds. *Rev. Mod. Phys.* 2011;83:1589–1652. <https://doi.org/10.1103/RevModPhys.83.1589>
3. Bozovic I., Ahn Ch. A new frontier for superconductivity. *Nature Phys.* 2014;10:892–895. <https://doi.org/10.1038/nphys3177>
4. Shen K.M., Ronning F., Meevasana W., Lu D.H., Ingle N.J.C., Baumberger F., et al. Angle-resolved photoemission studies of lattice polaron formation in the cuprate $\text{Ca}_2\text{CuO}_2\text{Cl}_2$. *Phys. Rev. B*. 2007;75:075115. <https://doi.org/10.1103/PhysRevB.75.075115>
5. Le Tacon M., Bosak A., Souliou S.M., Dellea G., Loew T., et al. Inelastic X-ray scattering in $\text{YBa}_2\text{Cu}_3\text{O}_{6.6}$ reveals giant phonon anomalies and elastic central peak due to charge-density-wave formation. *Nature Phys.* 2014;10:52–58. <https://doi.org/10.1038/nphys2805>
6. Myasnikova A.E., Doronkina S.V., Arutyunyan R.R., Dzhantemirov A.H. Free energy of a two-liquid system of charge carriers in strongly coupled electron and phonon fields and common nature of three phases in hole-doped cuprates. *J. Phys.: Condens. Matter*. 2024;36:325601. <https://doi.org/10.1088/1361-648X/ad459d>
7. Myasnikova A.E., Zhileeva E.A., Moseykin D.V. Relaxation of strongly coupled electron and phonon fields after photoemission and high-energy part of ARPES spectra of cuprates. *J. Phys.: Condens. Matter*. 2018;30(12):125601. <https://doi.org/10.1088/1361-648X/aaad3e>
8. Damascelli A., Hussain Z., Shen Z.-X. Angle-Resolved Photoemission Studies of the Cuprate Superconductors. *Rev. Mod. Phys.* 2003;75(2):473–541. <https://doi.org/10.1103/RevModPhys.75.473>
9. Doronkina S.V., Myasnikova A.E., Dzhantemirov A.H., Lutsenko A.V. Topological pseudogap in highly polarizable layered systems with 2D hole-like dispersion. *Physica E: Low-Dimens. Syst. Nanostruct.* 2022;136:115052. <https://doi.org/10.1016/j.physe.2021.115052>
10. Lanzara A., Bogdanov P.V., Zhou X.J., et al. Evidence for ubiquitous strong electron–phonon coupling in high-temperature superconductors. *Nature*. 2001;412:510–514. <https://doi.org/10.1038/35087518>
11. Xie B.P., Yang K., Shen D.W., et al. High-Energy Scale Revival and Giant Kink in the Dispersion of a Cuprate Superconductor. *Phys. Rev. Lett.* 2007;98:147001. <https://doi.org/10.1103/PhysRevLett.98.147001>
12. Graf J., Gweon G.-H., McElroy K., et al. Universal High Energy Anomaly in the Angle-Resolved Photoemission Spectra of High Temperature Superconductors: Possible Evidence of Spinon and Holon Branches. *Phys. Rev. Lett.* 2007;98:067004. <https://doi.org/10.1103/PhysRevLett.98.067004>
13. Moritz B., Schmitt F., Meevasana W., et al. Effect of strong correlations on the high energy anomaly in hole- and electron-doped high- T_c superconductors. *New J. Phys.* 2009;11:093020. <https://doi.org/10.1088/1367-2630/11/9/093020>
14. Ronning F., Shen K.M., Armitage N.P., Damascelli A., Lu D.H., et al. Anomalous high-energy dispersion in angle-resolved photoemission spectra from the insulating cuprate $\text{Ca}_2\text{CuO}_2\text{Cl}_2$. *Phys. Rev. B*. 2005;71:094518. <https://doi.org/10.1103/PhysRevB.71.094518>
15. Pekar S.I. *Untersuchungen über die Electronentheorie der Kristalle*. Berlin: Akademie-Verlag; 1954, 184 p. (In Germ.).
16. Rösch O., Gunnarsson O., Zhou X.J., et al. Polaronic behavior of undoped high- T_c cuprate superconductors from angle-resolved photoemission spectra. *Phys. Rev. Lett.* 2005;95:227002. <https://doi.org/10.1103/PhysRevLett.95.227002>
17. Myasnikov E.N., Myasnikova A.E., Kusmartsev F.V. Coherence of the lattice polarization in large-polaron motion. *Phys. Rev. B*. 2005;72:224303. <https://doi.org/10.1103/PhysRevB.72.224303>
18. Emin D. Phonon-mediated attraction between large bipolarons: Condensation to a liquid. *Phys. Rev. Lett.* 1994;72(7):1052–1055. <https://doi.org/10.1103/PhysRevLett.72.1052>

19. Myasnikov E.N., Myasnikova A.E., Mastropas Z.P. Multiphonon generation during photodissociation of slow Landau–Pekar polarons. *J. Exp. Theor. Phys.* 2006;102(3):480–496. <https://doi.org/10.1134/S1063776106030113>
[Original Russian Text: Myasnikov E.N., Myasnikova A.E., Mastropas Z.P. Multiphonon generation during photodissociation of slow Landau–Pekar polarons. *Zhurnal eksperimental'noi i teoreticheskoi fiziki.* 2006;129(3):548–565 (in Russ.). <https://www.elibrary.ru/hsxhvd>]
20. Myasnikova A.E., Myasnikov E.N. Correlation of optical conductivity and angle-resolved photoemission spectra of strong-coupling large polarons and its display in cuprates. *Phys. Rev. B.* 2008;77(16):165136. <https://doi.org/10.1103/PhysRevB.77.165136>
21. Myasnikova A.E., Myasnikov E.N., Moseykin D.V., Zuev I.S. Distribution of charge carriers at strong electron–phonon interaction and “vertical dispersion” in ARPES spectra of cuprates. *Phys. Lett. A.* 2015;379(5):458–465. <https://doi.org/10.1016/j.physleta.2014.10.051>
22. Menushenkov A.P., Ivanov A., Neverov V., et al. Direct evidence of real-space pairing in BaBiO₃. *Phys. Rev. Res.* 2024;6:023307. <https://doi.org/10.1103/PhysRevResearch.6.023307>
23. Meevasana W., Baumberger E., Tanaka K., et al. Extracting the spectral function of the cuprates by a full two-dimensional analysis: Angle-resolved photoemission spectra of Bi₂Sr₂CuO₆. *Phys. Rev. B.* 2008;77:104506. <https://doi.org/10.1103/PhysRevB.77.104506>
24. Mishchenko A.S., Nagaosa N., Shen K.M., Shen Z.-X., Zhou X.J., Devereaux T.P. Polaronic metal in lightly doped high-T_c cuprates. *Europhys. Lett.* 2011;95(5):57007. <https://doi.org/10.1209/0295-5075/95/57007>
25. Zhang W., Liu G., Meng J., et al. High Energy Dispersion Relations for the High Temperature Bi₂Sr₂CaCu₂O₈ Superconductor from Laser-Based Angle-Resolved Photoemission Spectroscopy. *Phys. Rev. Lett.* 2008;101:017002. <https://doi.org/10.1103/PhysRevLett.101.017002>
26. Valla T., Kidd T.E., Yin W.-G., Gu G.D., et al. High-Energy Kink Observed in the Electron Dispersion of High-Temperature Cuprate Superconductors. *Phys. Rev. Lett.* 2007;98:167003. <https://doi.org/10.1103/PhysRevLett.98.167003>
27. Sun W., Jiang Z., Xia C., et al. Electronic structure of superconducting infinite-layer lanthanum nickelates. *Sci. Adv.* 2025;11(4):eadr5116. <https://doi.org/10.1126/sciadv.adr5116>
28. Ding X., Fan Y., Wang X., et al. Cuprate-like electronic structures in infinite-layer nickelates with substantial hole dopings. *Natl. Sci. Rev.* 2024;11(8):nwae194. <https://doi.org/10.1093/nsr/nwae194>
29. Krsnik J., Held K. Local correlations necessitate waterfalls as a connection between quasiparticle band and developing Hubbard bands. *Nat. Comm.* 2025;16(1):255. <https://doi.org/10.1038/s41467-024-55465-7>

About the Authors

Stanislava V. Doronkina, Cand. Sci. (Phys.-Math.), Senior Lecturer, Higher Mathematics Department – 3, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: st.doronkina@gmail.com. Scopus Author ID 57219795495, RSCI SPIN-code 9086-8195, <https://orcid.org/0000-0002-0546-8740>

Anastasia S. Babayants, Student, Physics Faculty, Southern Federal University (105/42, Bolshaya Sadovaya ul., Rostov-on-Don, 344006 Russia). E-mail: 9aababayanz@gmail.com. <https://orcid.org/0009-0004-9940-3842>

Ivan V. Lisitsa, Assistant, Department of Medical Physics, Mathematics, and Information Technology, Rostov State Medical University (29, Nakhichevanskii per., Rostov-on-Don, 344022 Russia). E-mail: nikkola2014@bk.ru. <https://orcid.org/0009-0006-7927-2157>

Anna E. Myasnikova, Dr. Sci. (Phys.-Math.), Professor, Department of Theoretical and Computational Physics, Southern Federal University (105/42, Bolshaya Sadovaya ul., Rostov-on-Don, 344006 Russia). E-mail: myasnikova67@yandex.ru. Scopus Author ID 6602715354, <https://orcid.org/0000-0001-7262-8292>

Об авторах

Доронкина Станислава Валерьевна, к.ф.-м.н., старший преподаватель, кафедра высшей математики – 3, Институт перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: st.doronkina@gmail.com. Scopus Author ID 57219795495, SPIN-код РИНЦ 9086-8195, <https://orcid.org/0000-0002-0546-8740>

Бабаянц Анастасия Станиславовна, студент, физический факультет, ФГАОУ ВО «Южный федеральный университет» (344006, Россия, Ростов-на-Дону, ул. Большая Садовая, д. 105/42). E-mail: 9aababayanz@gmail.com. <https://orcid.org/0009-0004-9940-3842>

Лисица Иван Васильевич, ассистент, кафедра медицинской физики, математики и информационных технологий, ФГБОУ ВО «Ростовский государственный медицинский университет» (344022, Россия, Ростов-на-Дону, пер. Нахичеванский, д. 29). E-mail: nikkola2014@bk.ru. <https://orcid.org/0009-0006-7927-2157>

Мясникова Анна Эдуардовна, д.ф.-м.н., доцент, профессор кафедры теоретической и вычислительной физики, ФГАОУ ВО «Южный федеральный университет» (344006, Россия, Ростов-на-Дону, ул. Большая Садовая, д. 105/42). E-mail: myasnikova67@yandex.ru. Scopus Author ID 6602715354, <https://orcid.org/0000-0001-7262-8292>

Translated from Russian into English by V. Glyanchenko

Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling
Математическое моделирование

UDC 532.5;519.6

<https://doi.org/10.32362/2500-316X-2026-14-4-116-124>

EDN QPOWCM



RESEARCH ARTICLE

Numerical simulation of Couette flow in a micropolar fluid using the projection method

Kristina V. Gubareva ^{1, @}, Evgenii Yu. Prosviryakov ^{2, 3}, Anton V. Eremin ¹

¹ Samara State Technical University, Samara, 620062 Russia

² Ural Federal University, Yekaterinburg, 620062 Russia

³ Institute of Engineering Science, Ural Branch of the Russian Academy of Sciences, Yekaterinburg, 620049 Russia

@ Corresponding author, e-mail: r.kristina2017@mail.ru

• Submitted: 28.12.2025 • Revised: 07.02.2026 • Accepted: 24.05.2026

Abstract

Objectives. The study develops an efficient numerical algorithm for simulating a two-dimensional flow of a micropolar fluid in a plane channel with a moving upper wall (Couette flow) in order to investigate the influence of micropolarity parameters on the flow structure.

Methods. The equations governing the dynamics of a micropolar fluid are solved using the projection method with explicit time integration. Spatial discretization is performed by the finite difference method on a uniform 51×51 grid. The convective terms are approximated using a first-order upwind scheme to ensure stability at moderate Reynolds numbers. The microrotation and momentum conservation equations are solved separately, followed by a pressure correction to satisfy the incompressibility condition.

Results. Steady-state velocity and microrotation fields were obtained for a Couette flow at Reynolds number $Re = 100$ with micropolarity parameters $N = 0.3$ and $m = 0.015$. The numerical method demonstrated stable convergence within 5000 iterations to achieve residuals of $3 \cdot 10^{-6}$ for velocity components and $1.5 \cdot 10^{-6}$ for microrotation. Visualized results include the vector velocity field and streamlines, as well as distributions of microrotation, vorticity, and energy dissipation. A nonuniform microrotation field was formed having maximum values (~ 0.45 rad/s) localized in the corner regions of the channel.

Conclusions. The developed algorithm effectively simulates a micropolar fluid flow in a channel. The numerical method exhibits stable convergence to yield physically meaningful results. Micropolarity is confirmed to significantly alter the flow structure in comparison with the Newtonian case, leading to the development of a transverse velocity component and a nonlinear microrotation distribution. The obtained distributions of field characteristics can serve as a basis for further research into the rheological properties of micropolar media and for the verification of experimental data.

Keywords: micropolar fluid, Couette flow, projection method, numerical simulation, Cosserat continuum, microrotation, finite difference method

For citation: Gubareva K.V., Prosviryakov E.Yu., Eremin A.V. Numerical simulation of Couette flow in a micropolar fluid using the projection method. *Russian Technological Journal*. 2026;14(4):116–124. <https://doi.org/10.32362/2500-316X-2026-14-4-116-124>, <https://elibrary.ru/QPOWCM>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Численное моделирование течения Куэтта микрополярной жидкости методом проекции

К.В. Губарева ^{1, @}, Е.Ю. Просвирыков ^{2, 3}, А.В. Еремин ¹

¹ Самарский государственный технический университет, Самара, 443100 Россия

² Уральский федеральный университет имени первого Президента России Б.Н. Ельцина, Екатеринбург, 620062 Россия

³ Институт машиноведения имени Э.С. Горкунова Уральского отделения Российской академии наук, Екатеринбург, 620049 Россия

@ Автор для переписки, e-mail: r.kristina2017@mail.ru

• Поступила: 28.12.2025 • Доработана: 07.02.2026 • Принята к опубликованию: 24.05.2026

Резюме

Цели. Целью работы является разработка эффективного численного алгоритма для моделирования двумерного течения микрополярной жидкости в плоском канале с движущейся верхней стенкой (течение Куэтта) и исследование влияния параметров микрополярности на структуру течения.

Методы. Для решения уравнений микрополярной жидкости использован метод проекции с явной схемой интегрирования по времени. Пространственная дискретизация выполнена методом конечных разностей на равномерной сетке 51×51 . Конвективные члены аппроксимированы противопоточной схемой первого порядка для обеспечения устойчивости при умеренных числах Рейнольдса. Для уравнения микровращения и уравнений сохранения импульса применена раздельная схема решения с последующей коррекцией давления для удовлетворения условия несжимаемости.

Результаты. Получены стационарные поля скорости и микровращения для течения Куэтта при числе Рейнольдса $Re = 100$ и параметрах микрополярности $N = 0.3$, $m = 0.015$. Продемонстрирована устойчивая сходимость численного метода за 5000 итераций с точностью до $3 \cdot 10^{-6}$ для компонент скорости и $1.5 \cdot 10^{-6}$ для микровращения. Визуализированы векторное поле скорости, линии тока, распределение микровращения, завихренности и диссипации. Наблюдается неоднородное поле микровращения с максимальными значениями (~ 0.45 рад/с) в угловых областях канала.

Выводы. Разработанный алгоритм позволяет эффективно моделировать течение микрополярной жидкости в канале. Численный метод демонстрирует устойчивую сходимость и физическую осмысленность результатов. Установлено, что микрополярность существенно изменяет структуру течения по сравнению с ньютоновским случаем, приводя к формированию поперечной компоненты скорости и нелинейного распределения микровращения. Полученные распределения полевых характеристик могут служить основой для дальнейших исследований реологических свойств микрополярных сред и верификации экспериментальных данных.

Ключевые слова: микрополярная жидкость, течение Куэтта, метод проекции, численное моделирование, континуум Коссера, микровращение, метод конечных разностей

Для цитирования: Губарева К.В., Просвирыков Е.Ю., Еремин А.В. Численное моделирование течения Куэтта микрополярной жидкости методом проекции. *Russian Technological Journal*. 2026;14(4):116–124. <https://doi.org/10.32362/2500-316X-2026-14-4-116-124>, <https://elibrary.ru/QPOWCM>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Micropolar fluids are a class of non-Newtonian media that possess internal rotational degrees of freedom. Unlike classical Newtonian fluids described by a symmetric stress tensor, micropolar media are characterized by the presence of asymmetric stresses and an additional microrotational field [1, 2]. This class of fluid adequately models the behavior of suspensions, polymer solutions, liquids with impurities, and some biological media [3, 4]. Contemporary research actively uses such models to analyze flows in microfluidic devices [5] and biomechanical systems [6]. Promising areas also include modeling technical devices such as journal bearings [7] and analyzing complex flows taking into account additional physical factors, such as magnetohydrodynamic effects in hybrid nanofluids [8]. To solve coupled problems of heat and mass transfer in such media, effective numerical methods are being developed, including methods for splitting integral transforms [9].

A theoretical description of micropolar fluids was first proposed by Eringen within the framework of the Cosserat continuum theory [1, 10]. A micropolar fluid is a special case of the nonlinear elastic Cosserat continuum, in which the local frame is reduced to a single rotation angle to significantly simplify the mathematical description while preserving key physical features. However, the theory introduces an additional rotational degree of freedom, which leads to a significant complication of the system of motion equations. The mathematical model includes equations of conservation of momentum by a symmetric stress tensor, an equation for the microrotational field, and a continuity equation [2, 11]. An important direction in the theory of micropolar fluids is the search for exact solutions of the equations of motion, which serves as the basis for the verification of numerical methods [12, 13].

Couette flow, representing one of the fundamental flows in fluid dynamics, is a shear flow between two parallel plates, one of which moves at a constant velocity [14]. For Newtonian fluids, this flow has a simple analytical solution with a linear velocity profile. However, for micropolar fluids, the presence of microrotation leads to a nonlinear velocity profile and the possible occurrence of secondary flows, which

is of significant interest for the study of rheological properties [15, 16]. Recent works have demonstrated the relevance of studying exact solutions for Couette flows taking into account various physical factors [17].

The numerical simulation of micropolar fluid flows involves significant computational difficulties due to the complexity of the system of equations and the need to simultaneously solve equations for velocity, pressure, and microrotation. Conventional methods for solving the Navier–Stokes equations require modification to account for micropolar effects. Currently, such problems are solved by finite difference methods [18], finite element methods, and splitting approaches [19].

The projection—or splitting—method is one of the most effective approaches for solving incompressible fluid equations [20]. Here, the basic idea consists in dividing the problem into stages: first solving the equations of motion without taking pressure into account, followed by correction to satisfy the incompressibility condition by solving the Poisson equation for pressure. For micropolar fluid equations, this method requires an additional stage of solving the equation for microrotation [21].

Despite the significant number of works on micropolar fluids, issues of efficient numerical simulation of Couette flow for such media remain insufficiently studied. Specifically, investigation is required on the stability of numerical schemes, the influence of discretization parameters on solution accuracy, and the microrotation field distribution at various micropolarity parameter values [16, 22].

In the present work, an algorithm for numerical simulation of a two-dimensional Couette flow of a micropolar fluid according to the projection method is presented. The developed method is implemented using an upwind scheme for convective terms to ensure stability at moderate Reynolds numbers. The influence of micropolarity parameters on the flow structure and microrotation field distribution is studied. The main novelty of the work consists in the adaptation of the projection method for the simultaneous solution of the equations of motion and microrotation using a stabilizing upwind discretization. This ensures the stability of the calculations while maintaining acceptable accuracy. The results of the work can be used to verify modern exact solutions [13, 17, 23].

MATHEMATICAL MODEL

The original system of equations for an incompressible micropolar fluid in dimensional form was given in the fundamental works of A. Eringen [1] and G. Lukaszewicz [2]. According to these works, the equations of conservation of momentum, angular momentum, and continuity are written as follows:

$$\begin{aligned}\rho \left(\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right) &= -\nabla p + (\mu + \kappa) \nabla^2 \mathbf{u} + \kappa \nabla \times \boldsymbol{\omega}, \\ j \rho \left(\frac{\partial \boldsymbol{\omega}}{\partial t} + (\mathbf{u} \cdot \nabla) \boldsymbol{\omega} \right) &= -2\kappa \boldsymbol{\omega} + \kappa \nabla \times \mathbf{u} + \gamma \nabla^2 \boldsymbol{\omega}, \\ \nabla \cdot \mathbf{u} &= 0,\end{aligned}$$

where $\mathbf{u}$ is the velocity vector, p is the pressure, $\boldsymbol{\omega}$ is the microrotation vector, ρ is the fluid density, μ is the classical dynamic viscosity, κ is the rotational viscosity coefficient, j is the microinertia, and γ is the microviscosity coefficient.

For the considered problem of Couette flow in a rectangular channel, the following assumptions are made:

1. The fluid is incompressible ($\nabla \cdot \mathbf{u} = 0$).
2. There are no body forces or external moments.
3. The flow is isothermal (temperature effects are not taken into account).
4. Physical parameters μ , κ , and γ are constant.
5. The microviscosity coefficient γ is related to the classical viscosity μ and the characteristic size l of the microstructure by the expression $\gamma = \mu l^2$, while the microinertia is of the order of $j \sim \rho l^2$. This is a standard assumption in the theory of micropolar fluids [2, 16].

For numerical implementation, the equations were nondimensionalized. The characteristic scales chosen were length L (channel height), velocity U (velocity of the moving upper wall), time L/U , pressure ρU^2 , and microrotation U/L . Introducing dimensionless variables

$$\tilde{x} = \frac{x}{L}, \quad \tilde{\mathbf{u}} = \frac{\mathbf{u}}{U}, \quad \tilde{t} = \frac{tU}{L}, \quad \tilde{p} = \frac{p}{\rho U^2}, \quad \tilde{\boldsymbol{\omega}} = \frac{\boldsymbol{\omega} L}{U},$$

and substituting them into the dimensional equations taking the assumptions into account after discarding the tildes to simplify the notation, we obtain a system of equations in dimensionless form:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \frac{1}{\text{Re}} \nabla^2 \mathbf{u} + \frac{N}{\text{Re}} \nabla \times \boldsymbol{\omega}, \quad (1)$$

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + (\mathbf{u} \cdot \nabla) \boldsymbol{\omega} = \frac{1}{m^2 \text{Re}} \nabla^2 \boldsymbol{\omega} + \frac{2N}{m^2 \text{Re}} (\boldsymbol{\Omega} - \boldsymbol{\omega}), \quad (2)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (3)$$

with the dimensionless parameters $\text{Re} = \frac{\rho UL}{\mu + \kappa}$ (Reynolds number), $N = \frac{\kappa}{\mu + \kappa}$ (micropolarity parameter, $0 \leq N \leq 1$),

and $m = \frac{l}{L}$ (dimensionless characteristic length

of micropolarity), as well as with the vorticity vector $\boldsymbol{\Omega} = \nabla \times \mathbf{u}$.

In a two-dimensional formulation of the problem in the (x, y) plane, the only nonzero component of microrotation is ω_z , while the vorticity is defined as

$$\Omega_z = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}.$$

To simplify the notation, we will below omit the subscripts z , implying scalar quantities ω and Ω .

Physically, the parameter N characterizes the relative contribution of rotational degrees of freedom to the rheological behavior of the fluid, varying from 0 (Newtonian fluid) to 1 (strongly expressed micropolarity) [2].

For the Couette flow in the two-dimensional formulation, a rectangular region of size $L_x \times L_y$ is considered. The upper wall ($y = L_y$) moves at a constant velocity U_0 along the x axis, while the lower wall ($y = 0$) is stationary. The side walls ($x = 0$ and $x = L_x$) are also considered to be stationary. The boundary conditions for the velocity are

$$\begin{aligned}\mathbf{u}(x, 0) &= 0, \quad \mathbf{u}(x, L_y) = (U_0, 0), \\ \mathbf{u}(0, y) &= (L_x, y) = 0.\end{aligned} \quad (4)$$

For the microrotation field, homogeneous Neumann conditions are applied on all boundaries:

$$\frac{\partial \boldsymbol{\omega}}{\partial n} = 0. \quad (5)$$

These boundary conditions, which are selected under the assumption of the absence of microrotational flow through solid boundaries in the first approximation, are widely used in the literature on numerical simulation of micropolar flows [16, 21].

NUMERICAL METHOD AND VERIFICATION

The system of equations (1)–(3) was solved by the projection method with an explicit time integration scheme. The projection method for the system of equations of a micropolar fluid was implemented in the form of the following sequence of stages at each time step [20]:

1. The microrotation equation (2) is solved.
2. The intermediate velocity field (predictor) is calculated using the equations of motion (1) without taking into account the pressure gradient.
3. The Poisson equation for pressure (7) is solved.
4. The velocity is corrected taking into account the found pressure (equation (8)).

In the first stage, microrotation equation (2) is solved using an upwind scheme for the convective term. Discretization is performed using the finite difference method on a uniform grid with steps Δx and Δy in spatial coordinates. For stability at moderate Reynolds numbers, the convective terms are approximated by a first-order upwind scheme [18]. Since ω is a scalar quantity in the two-dimensional formulation, the convective operator is expanded as

$$(\mathbf{u} \cdot \nabla)\omega = u \frac{\partial \omega}{\partial x} + v \frac{\partial \omega}{\partial y},$$

where $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$. Each component is approximated independently, taking into account the direction of the local flow: the value is taken from the cell located upstream (opposite to the direction of velocity). With the numbering of grid nodes adopted in the work (j and i correspond to the coordinates x and y , respectively), the component along the x axis is approximated as

$$(u \cdot \nabla)\omega \approx \begin{cases} u_{i,j} \frac{\omega_{i,j} - \omega_{i,j-1}}{\Delta x}, & \text{if } u_{i,j} > 0, \\ u_{i,j} \frac{\omega_{i,j+1} - \omega_{i,j}}{\Delta x}, & \text{if } u_{i,j} < 0, \end{cases} \quad (6)$$

while the component along the y axis is approximated similarly with replacing subscripts ($j \rightarrow i$) and $\Delta x \rightarrow \Delta y$. Complete approximation of $(\mathbf{u} \cdot \nabla)\omega$ is obtained by summing both components. The Laplacian in equation (2) is calculated using the standard five-point scheme. Similar splitting methods, which have demonstrated utility for solving the Navier–Stokes equations, have been previously studied for incompressible flows [19, 24].

In the second stage, momentum conservation equation (1) are solved without taking into account the pressure gradient (velocity predictor). For this purpose, an upwind scheme for convective terms is also used. The terms containing microrotation are calculated using central differences.

The third stage involves solving the Poisson equation for pressure:

$$\nabla^2 p = \frac{1}{\Delta t} \nabla \cdot \mathbf{u}^*, \quad (7)$$

where $\mathbf{u}^*$ is the intermediate velocity field obtained in the second stage, and Δt is the time step. Equation (7) is solved by the successive overrelaxation method under the Neumann boundary conditions on pressure.

In the final stage, the velocity is corrected taking into account the pressure:

$$\mathbf{u}^{n+1} = \mathbf{u}^* - \Delta t \nabla p, \quad (8)$$

where $\mathbf{u}^{n+1}$ is the velocity in the new time layer.

To ensure the stability of the numerical scheme, the time step is selected from the Courant–Friedrichs–Lewy conditions:

$$\Delta t \leq \min \left(\frac{\Delta x}{|u|_{\max}}, \frac{\Delta y}{|v|_{\max}} \right), \quad (9)$$

as well as additional limitations related to viscosity and micropolarity parameters.

The energy dissipation is calculated from the formula [2]

$$\Phi = (1 + N)D_{ij}D_{ij} + \frac{1}{m^2} \nabla \omega \cdot \nabla \omega + 2N(\Omega - \omega)^2, \quad (10)$$

where $D_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$ is the strain rate tensor.

The vorticity is calculated as the velocity curl:

$$\Omega = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \quad (11)$$

The criterion for the convergence of the iterative process is the achievement of a given accuracy for the residuals of all field variables:

$$\max(|\mathbf{u}^{n+1} - \mathbf{u}^n|, |\omega^{n+1} - \omega^n|) < \varepsilon, \quad (12)$$

where ε is the specified accuracy; in the calculations, it was taken as equal to 10^{-7} .

The numerical method was verified by comparison with the classical solution of the Couette problem for a Newtonian fluid (at $N = 0$). In this case, the resulting velocity profile coincided to within 0.5% with the analytical linear profile $u(y) = U_0 y / L_y$. All calculations and visualizations of results are performed in the *MATLAB* environment, version R2023b [25].¹

RESULTS AND DISCUSSION

The numerical simulation was carried out for a channel with dimensions $L_x = 0.1$ m and $L_y = 0.1$ m. The flow parameters were as follows: Reynolds number $Re = 100$; micropolarity parameter $N = 0.3$; characteristic length $m = 0.015$; upper wall velocity $U_0 = 0.3$ m/s. The parameters were selected in accordance with characteristic values for technical micropolar media, where m is in the range of 0.01–0.1 and N corresponds to moderately expressed micropolar properties [2, 21]. The computational grid consisted of 51×51 nodes, which provided sufficient resolution for flow structure analysis. Similar grid parameters and approaches to modeling have been used in modern research of flows in cavities and channels [16, 22].

¹ https://www.mathworks.com/products/new_products/release2023b.html. Accessed December 28, 2025.

Figure 1 presents the velocity vector field for the steady-state flow. The characteristic profile of the Couette flow with the predominant flow along the x axis is observed. The maximum velocity, which is achieved at the upper moving wall and amounts to 0.3 m/s, corresponds to the boundary condition. At the lower stationary wall, the velocity tends to zero. The side walls create stagnation zones in which the velocity vectors have a minimum value. The ratio of the maximum transverse velocity to the maximum longitudinal velocity is 0.3, which indicates a noticeable transverse component, characteristic of two-dimensional models of micropolar flows [16].

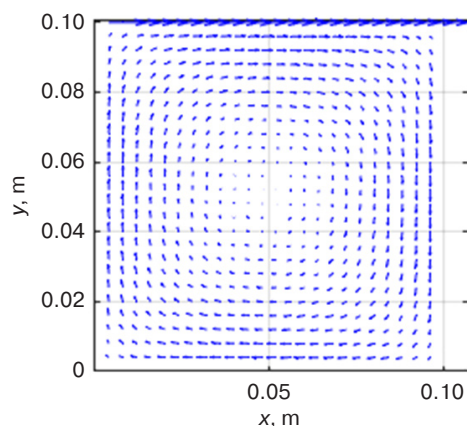


Fig. 1. Vector field of velocity u

The streamlines shown in Fig. 2 demonstrate predominantly parallel flow with slight deviations in the corner regions. At the given flow parameters, pronounced secondary vortices do not form, as consistent with theoretical expectations at moderate values of the micropolarity parameter N and agreeing with results of flow simulations in cavities at similar parameter values [22].

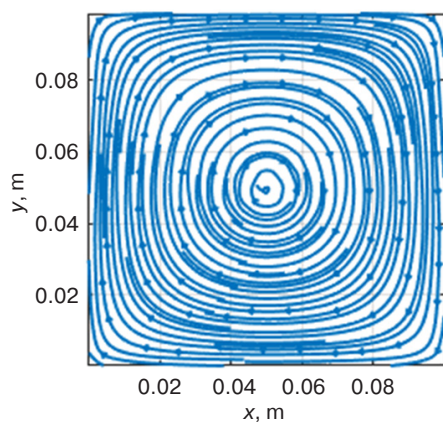


Fig. 2. Streamlines

Figure 3 illustrates the distribution of the microrotation module $|\omega|$. The most intense microrotation is observed

in the corner regions of the channel, where the interaction of the stationary side walls with the moving upper wall creates maximum velocity gradients. The maximum value of the microrotation modulus reaches 0.45 rad/s. The minimal magnitude of microrotation observed in the center of the channel is consistent with theoretical concepts about the attenuation of rotational degrees of freedom far from the boundaries. The resulting microrotation field is significantly nonuniform, having clearly defined maxima in the near-wall regions, primarily along the upper and lower walls at $x \approx 0.025$ and $x \approx 0.075$. The obtained distribution is qualitatively consistent with the results of other studies of micropolar fluid flows in channels [16] and with conclusions of works on exact solutions for unidirectional flows [13, 17].

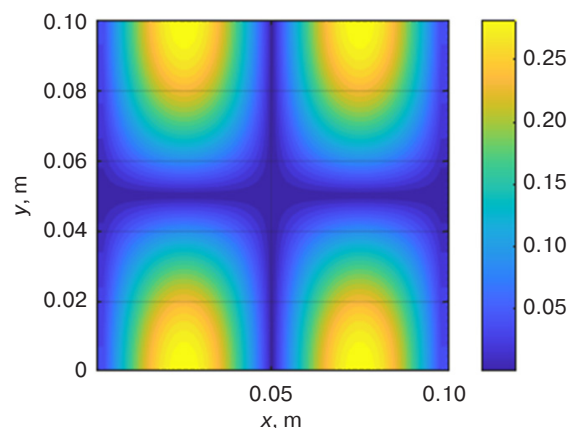


Fig. 3. Distribution of microrotation modulus $|\omega|$

The convergence of the numerical method is demonstrated in Fig. 4 on a semilogarithmic scale. The residuals for all three field variables show a monotonic decrease. After 2000 iterations, the residuals for the velocity components stabilize at $3 \cdot 10^{-6}$, while for microrotation, stability occurs at $1.5 \cdot 10^{-6}$, indicating that a steady state has been reached. The calculation was stopped at 5000 iterations, since further reduction of the residuals occurred extremely slowly with stabilized field distributions. The total calculation time was 55.0 s. This convergence rate, which is acceptable for the class of problems under consideration, confirms the effectiveness of the applied algorithm, which is consistent with studies of the convergence of projection methods [19].

Figure 5 presents the distribution of vorticity Ω calculated using equation (11). The maximum values of vorticity are observed in the corner regions of the channel, where the flow experiences the greatest deformations due to interaction with fixed walls. The correlation of the vorticity distribution with the microrotation field confirms the physical consistency of the solution. The maximum vorticity value is 1.6 s^{-1} , which corresponds to the characteristic velocity gradient in the flow.

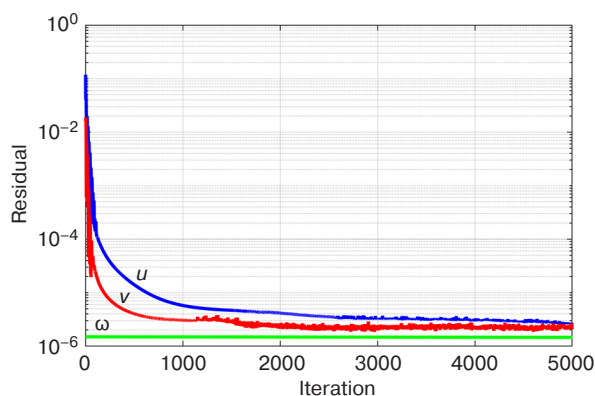


Fig. 4. Convergence of the numerical method: residuals versus iteration count

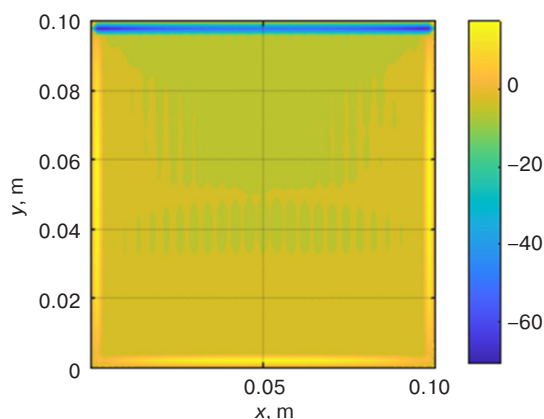


Fig. 5. Distribution of vorticity Ω

Figure 6 demonstrates the distribution of energy dissipation Φ calculated using formula (10). The maximum dissipation is observed near the walls, especially in the corners, where velocity gradients are greatest. The overall dissipation level corresponds to a moderate Reynolds number and micropolarity parameters. Analysis of the dissipation distribution demonstrates a significant contribution of micropolar effects to the overall energy balance, especially in the near-wall regions. The study of dissipative processes is an important aspect of the analysis of flows of structured fluids [23].

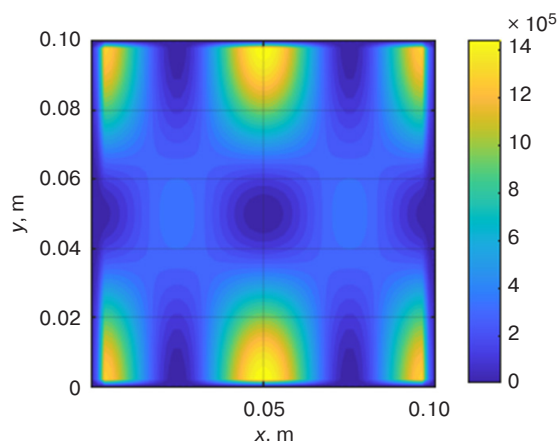


Fig. 6. Distribution of energy dissipation Φ

The quality of the numerical solution was verified by analyzing the incompressibility condition (3). The maximum velocity divergence in the computational domain did not exceed 10^{-4} s^{-1} , thus confirming the adequacy of the projection method for solving micropolar fluid equations.

CONCLUSIONS

In this work, a numerical algorithm for simulating a two-dimensional Couette flow of a micropolar fluid has been developed and implemented. The projection method with upwind approximation of convective terms was used to ensure the stability of the calculations at moderate Reynolds numbers. The described approach, which is based on contemporary studies of the convergence of splitting methods, is applicable to simulating complex flows.

Physically meaningful distributions of velocity, microrotation, vorticity, and energy dissipation were obtained. Micropolarity is found to significantly alter the flow structure in comparison with the Newtonian case: a noticeable transverse velocity component ($v/u \approx 0.3$) is observed, while the microrotation reaches its maximum values ($\sim 0.45 \text{ rad/s}$) in the corner regions. The numerical method demonstrates stable convergence with a monotonic decrease in residuals over 5000 iterations to $3 \cdot 10^{-6}$ for velocity components and $1.5 \cdot 10^{-6}$ for microrotation.

The developed algorithm can be used for the study of rheological properties of micropolar fluids at various values of the parameters N and m , as well as for flow analysis in more complex geometries. The practical significance of the results lies in the applicability of the developed approach for modeling biofluids, processing polymeric materials, and designing microfluidic devices, where it is necessary to account for micropolar effects. The obtained numerical solutions can serve as a test for verifying new exact solutions to the equations of micropolar hydrodynamics.

A promising direction for further research would consist in an expansion of the method to describe three-dimensional flows, taking into account transient effects, heat and mass transfer, and magnetohydrodynamic forces. Here it will also be of interest to compare the calculation results with data obtained by the method for splitting integral transforms for related problems. The results of this work contribute to the development of numerical methods for modeling micropolar media, as well having relevance for the verification of experimental data and analytical solutions.

Authors' contributions

K.V. Gubareva—development and implementation of the numerical algorithm, performing computations, processing and visualization of results.

E.Yu. Prosviryakov—problem statement, formalization of the mathematical model, scientific supervision.

A.V. Eremin—verification of the method, analysis of the obtained data, interpretation of results and formulation of conclusions.

REFERENCES

1. Eringen A.C. Simple Microfluids. *Int. J. Eng. Sci.* 1964;2(2):205–217. [https://doi.org/10.1016/0020-7225\(64\)90005-9](https://doi.org/10.1016/0020-7225(64)90005-9)
2. Lukaszewicz G. *Micropolar Fluids: Theory and Applications*. Basel: Birkhäuser; 1999, 320 p. <https://doi.org/10.1007/978-1-4612-0641-5>
3. Ariman T., Turk M.A., Sylvester N.D. Microcontinuum Fluid Mechanics – A Review. *Int. J. Eng. Sci.* 1973;11(8):905–930. [https://doi.org/10.1016/0020-7225\(73\)90038-4](https://doi.org/10.1016/0020-7225(73)90038-4)
4. Cowin S.C. The Theory of Polar Fluids. *Adv. Appl. Mech.* 1974;14:279–347. [https://doi.org/10.1016/S0065-2156\(08\)70034-6](https://doi.org/10.1016/S0065-2156(08)70034-6)
5. Squires T.M., Quake S.R. Microfluidics: Fluid Physics at the Nanoliter Scale. *Rev. Mod. Phys.* 2005;77(3):977–1026. <https://doi.org/10.1103/REVMODPHYS.77.977>
6. Carvalho V., Pinho D., Lima R.A., Teixeira J.C., Teixeira S. Blood Flow Modeling in Coronary Arteries: A Review. *Fluids*. 2021;6(2):53. <https://doi.org/10.3390/fluids6020053>
7. Mathur P., Soni S. Transient nonlinear modeling of lemon bore journal bearings with micropolar fluid and eccentricity effects. *Multiscale and Multidiscip. Model. Exp. and Des.* 2025;8:467. <https://doi.org/10.1007/s41939-025-01065-x>
8. Tassaddiq A. Impact of Cattaneo-Christov heat flux model on MHD hybrid nano-micropolar fluid flow and heat transfer with viscous and joule dissipation effects. *Sci. Rep.* 2021;11(1):67. <https://doi.org/10.1038/s41598-020-77419-x>
9. Kartashov E.M. Development of model representations of thermal reaction viscoelastic bodies on the temperature field. *Russian Technological Journal*. 2024;12(6):80–90. <https://doi.org/10.32362/2500-316X-2024-12-6-80-90>
10. Eringen A.C. Theory of Micropolar Fluids. *J. Math. Mech.* 1966;16(1):1–18. <https://doi.org/10.1512/IUMJ.1967.16.16001>
11. Ahmadi G. Self-Similar Solution of Incompressible Micropolar Boundary Layer Flow Over a Semi-Infinite Plate. *Int. J. Eng. Sci.* 1976;14(7):639–646. [https://doi.org/10.1016/0020-7225\(76\)90006-9](https://doi.org/10.1016/0020-7225(76)90006-9)
12. Burmasheva N.V., Prosviryakov E.Y. Exact Solutions to the Navier–Stokes equations for unidirectional flows of micropolar fluids in a mass force field. *Diagnostics, Resource and Mechanics of Materials and Structures (DReaM)*. 2024;3:41–63. <https://doi.org/10.17804/2410-9908.2024.3.041-063>
13. Baranovskii E.S., Ershkov S.V., Prosviryakov E.Y., Yudin A.V. Exact Solutions to the Oberbeck–Boussinesq Equations for Describing Three-Dimensional Flows of Micropolar Liquids. *Symmetry*. 2024;16(12):1669. <https://doi.org/10.3390/sym16121669>
14. Batchelor G.K. *An Introduction to Fluid Dynamics*. Cambridge: Cambridge University Press; 1967, 615 p.
15. Alwawi F., Sulaiman I.M., Swalmeh M.Z., Yaseen N. Energy transport boosters of magneto micropolar fluid flowing past a cylinder: A case of laminar combined convection. *Proceedings of the Institution of Mechanical Engineers. Part C: Journal of Mechanical Engineering Science*. 2022;236(22):10902–10913.
16. Ali N., Nazeer M., Javed T., Abbas F. A numerical study of micropolar flow inside a lid-driven triangular enclosure. *Meccanica*. 2018;53(13):3279–3299. <https://doi.org/10.1007/s11012-018-0884-5>
17. Burmasheva N.V., Prosviryakov E.Y. The Inhomogeneous Couette Flow of a Micropolar Fluid. *Russ. J. Nonlin. Dyn.* 2025;21(3):345–358. <https://doi.org/10.20537/nd250601>
18. Patankar S.V. *Numerical Heat Transfer and Fluid Flow*. New York: Hemisphere Publishing Corporation; 1980, 197 p.
19. Maeda M., Soga K. More on convergence of Chorin’s projection method for incompressible Navier–Stokes equations. *J. Math. Fluid Mech.* 2022;24(2):41. <https://doi.org/10.1007/s00021-021-00652-5>
20. Chorin A.J. Numerical solution of the Navier-Stokes equations. *Mathematics of Computation*. 1968;22(104):745–762. <https://doi.org/10.1090/S0025-5718-1968-0242392-2>
21. Venkatadri K., Maheswari S., Lakshmi C.V., Prasad V.R. Numerical simulation of lid-driven cavity flow of micropolar fluid. *IOP Conf. Ser.: Mater. Sci. Eng.* 2018;402:012168. <https://doi.org/10.1088/1757-899X/402/1/012168>
22. Batool S., Rasool G., Alshammari N., Khan I., Kaneez H., Hamadneh N. Numerical analysis of heat and mass transfer in micropolar nanofluids flow through lid driven cavity: Finite volume approach. *Case Studies in Thermal Engineering*. 2022;37:102233. <https://doi.org/10.1016/j.csite.2022.102233>
23. Baranovskii E.S., Prosviryakov E.Y., Ershkov S.V. Mathematical analysis of steady non-isothermal flows of a micropolar fluid. *Nonlin. Anal.: Real World Appl.* 2025;84:104294. <https://doi.org/10.1016/j.nonrwa.2024.104294>
24. Kartashov E.M. Method for splitting integral transformation in problems of complex heat transfer. *Russian Technological Journal*. 2025;13(6):104–115. <https://doi.org/10.32362/2500-316X-2025-13-6-104-115>
25. Gubareva K.V., Prosviryakov E.Yu. MATLAB Implementation of a Projection Method for 2D Micropolar Couette Flow. *Mendeley Data*. V1. 2026. <https://doi.org/10.17632/mpjtfj7krc.1>

About the Authors

Kristina V. Gubareva, Cand. Sci. (Eng.) Associate Professor, Department of Industrial Thermal Power Engineering, Samara State Technical University (244, Molodogvardeyskaya ul., Samara, 443100 Russia). E-mail: r.kristina2017@mail.ru. Scopus Author ID 57216361463, RSCI SPIN-code 4171-9816, <https://orcid.org/0000-0002-9845-8372>

Evgenii Yu. Prosviryakov, Dr. Sci. (Phys.-Math.), Professor, Department of Information Technology and Automation, Institute of Radioelectronics and Information Technologies, Ural Federal University (32, Mira ul., Yekaterinburg, 620062 Russia); Head of the Sector of Nonlinear Vortex Hydrodynamics, Institute of Engineering Science, Ural Branch of the Russian Academy of Sciences (34, Komsomolskaya ul., Yekaterinburg, 620049 Russia). E-mail: evgen_pros@mail.ru. Scopus Author ID 57189461740, ResearcherID E-6254-2016, RSCI SPIN-code 3880-5690, <https://orcid.org/0000-0002-2349-7801>

Anton V. Eremin, Dr. Sci. (Eng.), Associate Professor, Vice-Rector for Scientific Work, Head of the Department of Industrial Thermal Power Engineering, Samara State Technical University (244, Molodogvardeyskaya ul., Samara, 443100 Russia). E-mail: a.v.eremin@list.ru. Scopus Author ID 56395547000, ResearcherID D-6936-2014, RSCI SPIN-code 3892-0775, <https://orcid.org/0000-0002-2614-6329>

Об авторах

Губарева Кристина Владимировна, к.т.н., доцент, кафедра «Промышленная теплоэнергетика», ФГБОУ ВО «Самарский государственный технический университет» (443100, Россия, Самара, ул. Молодогвардейская, д. 244). E-mail: r.kristina2017@mail.ru. Scopus Author ID 57216361463, SPIN-код РИНЦ 4171-9816, <https://orcid.org/0000-0002-9845-8372>

Просви́ряков Евге́ний Ю́рьевич, д.ф.-м.н., доцент, профессор, кафедра информационных технологий и систем управления, Институт радиоэлектроники и информационных технологий – РтФ, Уральский федеральный университет имени первого Президента России Б.Н. Ельцина (620062, Россия, Екатеринбург, ул. Мира, д. 32); заведующий сектором нелинейной вихревой гидродинамики, ФГБУН «Институт машиноведения имени Э.С. Горкунова Уральского отделения Российской академии наук» (620049, Россия, Екатеринбург, ул. Комсомольская, д. 34). E-mail: evgen_pros@mail.ru. Scopus Author ID 57189461740, ResearcherID E-6254-2016, SPIN-код РИНЦ 3880-5690, <https://orcid.org/0000-0002-2349-7801>

Еремин Антон Владимирович, д.т.н., доцент, проректор по научной работе, заведующий кафедрой «Промышленная теплоэнергетика», ФГБОУ ВО «Самарский государственный технический университет» (443100, Россия, Самара, ул. Молодогвардейская, д. 244). E-mail: a.v.eremin@list.ru. Scopus Author ID 56395547000, ResearcherID D-6936-2014, SPIN-код РИНЦ 3892-0775, <https://orcid.org/0000-0002-2614-6329>

Translated from Russian into English by V. Glyanchenko

Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling
Математическое моделирование

UDC 537.86:519.25

<https://doi.org/10.32362/2500-316X-2026-14-4-125-136>

EDN RCOVQR



RESEARCH ARTICLE

Modernization of a statistical data processing algorithm for integrated background monitoring

Vladimir A. Ivanov ^{1, 2}, Sergey G. Paramonov ²,
Ekaterina S. Zhigacheva ^{2, @}, Ekaterina A. Pozdnyakova ²

¹ MIREA – Russian Technological University, Moscow, 119454 Russia

² Yu.A. Izrael Institute of Global Climate and Ecology, Moscow, 107258 Russia

@ Corresponding author, e-mail: kosjatko@gmail.com

• Submitted: 02.10.2025 • Revised: 19.01.2026 • Accepted: 01.06.2026

Abstract

Objectives. Since the 1980s, the Integrated Background Monitoring Network (IBMoN) operated by RosHydroMet¹ in the Russian Federation has been collecting experimental data on background levels of pollutants in environmental media, including atmospheric air, precipitation, surface water, soil, and vegetation. As part of its operations, IBMoN regularly updates recommendations and regulatory documents regarding chemical analysis methods and requirements for monitoring site placement every 5–10 years. However, the statistical framework of the network has not been revised since 1995 (Algorithm 1995). This paper sets out to modernize selected steps of the statistical algorithm for processing data on air quality obtained from IBMoN.

Methods. In the study, correlation analysis, trend analysis, descriptive mathematical statistics, and structural algorithmization methods were used.

Results. Based on a review of Algorithm 1995 steps requiring modernization, a structure for the updated data processing algorithm version, Algorithm 2025, was developed. This new algorithm will enable the joint analysis of RosHydroMet data to mitigate the impact of noise and fluctuations in the time series of background concentrations of air pollutants. The effectiveness of the proposed methods for modernizing selected steps of Algorithm 1995 was substantiated.

Conclusions. Since it is not possible to obtain data in the format required by RosHydroMet without modifying some steps of the Algorithm 1995, this algorithm requires modernization. As a result of modernizing the selected steps, noise can be reduced by up to 70% to eliminate the main types of random and systematic errors caused by the human factor. However, future work is required to identify suitable approaches for modernizing the remaining steps that are not covered in this study.

Keywords: algorithm, time series, background monitoring, environmental pollution, statistical processing, trends

¹ RosHydroMet is the Federal Service for Hydrometeorology and Environmental Monitoring of the Russian Federation. <https://www.meteorf.gov.ru/>. Accessed May 22, 2026. (In Russ.).

For citation: Ivanov V.A., Paramonov S.G., Zhigacheva E.S., Pozdnyakova E.A. Modernization of a statistical data processing algorithm for integrated background monitoring. *Russian Technological Journal*. 2026;14(4):125–136. <https://doi.org/10.32362/2500-316X-2026-14-4-125-136>, <https://elibrary.ru/RCOVQR>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Модернизация алгоритма статистической обработки данных сети комплексного фоновый мониторинга

В.А. Иванов^{1, 2}, С.Г. Парамонов²,
Е.С. Жигачева^{2, @}, Е.А. Позднякова²

¹ МИРЭА – Российский технологический университет, Москва, 119454 Россия

² Институт глобального климата и экологии имени академика Ю.А. Израэля, Москва, 107258 Россия

@ Автор для переписки, e-mail: kosjatko@gmail.com

• Поступила: 02.10.2025 • Доработана: 19.01.2026 • Принята к опубликованию: 01.06.2026

Резюме

Цели. На территории Российской Федерации с 1980-х гг. функционирует сеть комплексного фоновый мониторинга (КФМ) Росгидромета². В задачи программы КФМ входит получение экспериментальных данных о фоновом содержании загрязняющих веществ в различных природных средах (атмосферный воздух, атмосферные осадки, поверхностные воды, почва, растительность). В рамках работы сети регулярно (1 раз в 5–10 лет) обновляются рекомендации и нормативные документы о методах химического анализа, требованиях к размещению станций наблюдения и пр. Однако статистический аппарат сети не подвергался пересмотру с 1995 г. (Алгоритм 1995). Цель работы – модернизация выбранных шагов алгоритма статистической обработки данных о состоянии и загрязнении атмосферного воздуха, получаемых на сети КФМ.

Методы. Использованы методы корреляционного, трендового анализа, описательной математической статистики и структурной алгоритмизации.

Результаты. Рассмотрены шаги Алгоритма 1995, требующие модернизации. Создана структура обновленного алгоритма обработки данных – Алгоритм 2025, который предположительно позволит проводить совместный анализ данных Росгидромета, а также приведет к уменьшению влияния шумов и флуктуаций во временных рядах данных о фоновых концентрациях загрязняющих веществ в атмосферном воздухе. Предложены подходы к модернизации выбранных шагов Алгоритма 1995 и показана эффективность их модернизации.

Выводы. Невозможно получить данные в формате, рекомендуемом Росгидрометом, без изменения некоторых шагов устаревшего Алгоритма 1995, и необходима его модернизация. Модернизация выбранных шагов алгоритма позволит сократить шумы посредством устранения до 70% основных типов случайных и систематических ошибок, обусловленных человеческим фактором. Необходимо разработать корректные подходы к модернизации оставшихся шагов алгоритма, которые не были рассмотрены в этой работе.

Ключевые слова: алгоритм, временные ряды, фоновый мониторинг, загрязнение окружающей среды, статистическая обработка данных, тренды

² Росгидромет – федеральная служба по гидрометеорологии и мониторингу окружающей среды. <https://www.meteorf.gov.ru/>. Дата обращения 22.05.2026. [RosHydroMet is the Federal Service for Hydrometeorology and Environmental Monitoring of the Russian Federation. <https://www.meteorf.gov.ru/>. Accessed May 22, 2026. (In Russ.).]

Для цитирования: Иванов В.А., Парамонов С.Г., Жигачева Е.С., Позднякова Е.А. Модернизация алгоритма статистической обработки данных сети комплексного фоновое мониторинга. *Russian Technological Journal*. 2026;14(4):125–136. <https://doi.org/10.32362/2500-316X-2026-14-4-125-136>, <https://elibrary.ru/RCOVQR>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

Glossary

Filter exposure is a process in which the air passes through a filter mounted in a special holder.

Blank sample (field blank sample) is a sample that undergoes all the usual manipulations (unpacking, transportation to the sampling site, etc.), except for exposure.

Limit of detection (LOD) in chemical analysis is the minimum concentration of a substance in a sample that can be reliably distinguished from the background concentration.

Outliers (extreme values) are values in the input data that are far beyond the range of other observations. Outliers do not fit into the general model of the analyzed process due to differing significantly from the surrounding values. Outliers can be caused by measurement errors, incorrect data input, or result from high variability.

Geophysical year in Russia corresponds to the period of the hydrological year, which is assumed to begin on October 1 with the cold season, which lasts until March 31, while the warm season extends from April 1 to September 30.

INTRODUCTION

Environmental pollution is a major concern facing humanity. During the 1960s and 1970s, when systemic observations of pollution levels and environmental conditions started to be carried out, international and national monitoring initiatives were undertaken through the state monitoring network of the Soviet Union. The concept of integrated background monitoring (IBMoN) for environmental pollution emerged during the late 1970s. The network of IBMoN monitoring stations established in the early 1980s continues to collect data on pollutant concentrations in different natural environments [1, 2]. Yu.A. Izrael Institute of Global Climate and Ecology (IGCE)³ serves as the scientific and methodological center for the IBMoN network. Its main functions include providing methodological guidance for the observation network, including the publication of technical and regulatory documents concerning chemical analysis methods, as well as recommendations for equipping observation stations and quality control of observation results. As well as carrying out some measurements, the network collects, stores and summarizes information. The IBMoN's dataset comprises time series of pollutant concentrations in various media, consisting in the processed results of laboratory measurements obtained using standardized methods and calculated according to the 1995 algorithm for processing experimental data on ambient air pollutant concentrations. These observation results are published in RosHydroMet and IGCE periodicals [3, 4].

In accordance with regulatory documents⁴ and publications from RosHydroMet and IGCE, long-term observation results should be published in annual sequence, with an averaging interval of one calendar year (from January to December). However, historically, the averaging interval for IBMoN observations was the geophysical year (from October to September). For this reason, up until 2023, the main results from the IBMoN network were published by RosHydroMet and IGCE for the geophysical year rather than the calendar year. This made it difficult to compare these data with the publications of other RosHydroMet organizations. In 2023, a decision was made to switch to a standard averaging interval of a calendar year. Meanwhile, improvements in techniques for working with noisy signals and technical capabilities for processing statistical data have led to a need to revise the algorithm used for processing IBMoN observation data.

The present work sets out to update selected stages of the algorithm for processing statistical data on atmospheric air quality and pollution collected from the IBMoN network. These updates will allow the data to be compared with observations from other RosHydroMet networks, thereby reducing the impact of noise and variations on background air pollutant levels.

³ <http://www.igce.ru/>. Accessed May 22, 2026. (In Russ.).

⁴ Order of the Federal Service for Hydrometeorology and Environmental Monitoring of the Russian Federation No. 798 of December 30, 2015, "On Approval of Regulations for the Annual Review of State and Environmental Pollution in the Russian Federation". <https://www.meteorf.gov.ru/upload/iblock/952/Prikaz-RHM-20151230-798.pdf>. Accessed May 22, 2026. (In Russ.).

OBJECTS AND METHODS

The study examines time series data from 1988 to 2019 on pollutant concentrations in the ambient air of the Astrakhan State Natural Biosphere Reserve⁵. The pollutants analyzed include total suspended particulate (TSP), sulfur dioxide (SO₂), sulfate (SO₄²⁻), and lead (Pb), which were measured according to standardized methods such as RD 52.04.186-89⁶ and RD 55.44.593-2015 [2]. Since the measurement error for these methods is up to 25%, field blank samples are subtracted from the measurements to eliminate noise due to storage, transportation, and sample preparation.⁷ TSP, SO₂, and SO₄²⁻ are measured at the IBMoN monitoring station, while Pb is measured in a scientific and methodological institutional laboratory.

The time series of pollutant concentration data, which is based on intervals, is not equidistant. Due to variations in observation programs over the study period, observation results have characterized the concentrations of sampled substances over various time periods over different years (one value per week, one every three days, one per day, etc.) [5]. Since, gaps appear in the data series due to missing data, the number of measurements varies. For example, the measurements concerning TSP have an annual range of 22 measurements in 2004 and 351 measurements in 1992.

STATISTICAL PROCESSING ALGORITHM

We consider the operation of the Algorithm 1995 (Fig 1), which has been in use since 1995 for processing time series data on atmospheric pollutant concentrations obtained from observation stations and the IBMoN network's institutional scientific and methodological laboratory.

Step 1: Collection of experimental data. This step involves the gathering and recording of observational results.

Step 2: Data entry into paper and electronic logbooks. Instrument operators record data immediately following analysis in a paper laboratory logbook, which records the causes of potential distortions in results that reduce data reliability. On a monthly basis, data is transferred

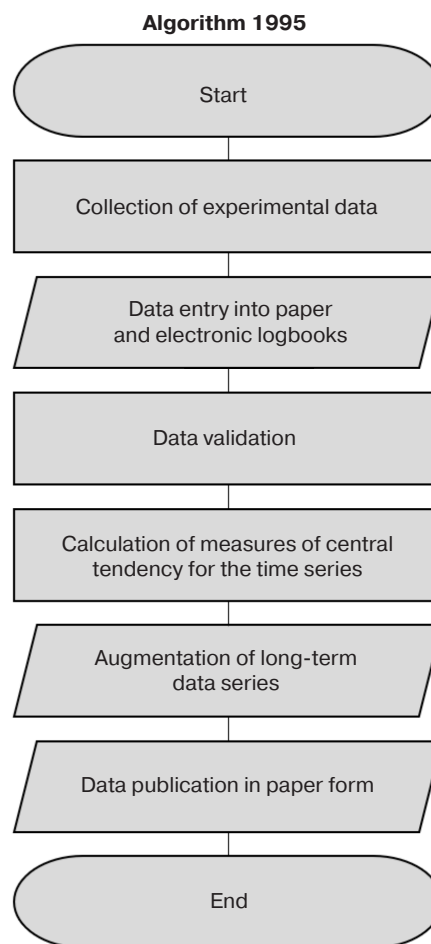


Fig. 1. Existing algorithm for processing time series of data on ambient air pollutant concentrations

to electronic spreadsheets using *DBase*⁸ and *FoxPro*⁹ software on a PC. Two operators or a single operator double-enter the data to eliminate misprints.

Step 3: Data validation. Validation is performed through assessment (by one expert). This includes decisions regarding removal of outliers (extreme values), known false values (negative values), and values below the instrument's LOD.

Step 4: Calculation of measures of central tendency for the time series. Averages are calculated using the Algorithm 1995 in tandem with the geometric mean method for both the cold and warm seasons, as well as for the entire year. For the warm season, the time series from April to September are used to calculate the geometric mean. For the cold season, the time series from October of the previous year through to March of the current year are used. The annual geometric mean is then calculated based on geometric means of the warm and cold seasons. These geometric means are used

⁵ <https://astrakhanzapoved.ru/>. Accessed May 22, 2026. (In Russ.).

⁶ RD 52.04.186-89. Guidance Document for the Control of Atmospheric Pollution (approved by the State Committee for Hydrometeorology of the USSR on January 6, 1989 and the Chief State Sanitary Inspector of the USSR on May 16, 1989). (In Russ.).

⁷ The EMEP Manual for Sampling and Chemical Analysis. EMEP/CCC – Report 1/95.Revision 2001. Norwegian Institute for Air Research; 2001.

⁸ <https://www.dbase.com/>. Accessed May 22, 2026.

⁹ <https://www.microsoft.com/>, official support has been terminated.

according to the assumption of a multimodal log-normal distribution of the data series under consideration.¹⁰

Step 5: Augmentation of long-term data series. The calculated seasonal and annual averages are then added to the existing long-term series.

Step 6: Data publication in paper form. This involves saving the data in a spreadsheet and creating graphs and charts. The materials are then prepared for publication by RosHydroMet and IGCE.

ASSESSMENT OF THE STATISTICAL SIGNIFICANCE OF TRENDS

The presence, significance, and direction of trends in atmospheric pollutant concentrations within the Astrakhan Nature Reserve were assessed using the Mann–Kendall trend test [6] and a nonparametric linear smoothing method based on the Theil–Sen estimator [4, 7].

In accordance with the requirements for applying the Mann–Kendall trend test to datasets with 10 or fewer observations, trends can be assessed using the S statistic. For datasets with more than 10 observations, the Z statistic can be calculated from S . All time series used in this study contain more than 10 data points.

The S statistic parameter is calculated using the following formula:

$$S = \sum_{b=1}^{n-1} \sum_{a=b+1}^n \text{sgn}(x_a - x_b), \quad (1)$$

where x_a and x_b are the annual values for years a and b , respectively, and $\text{sgn}(x_a - x_b)$ can be calculated using the following formula:

$$\text{sgn}(x_a - x_b) = \begin{cases} 1, & x_a - x_b > 0, \\ 0, & x_a - x_b = 0, \\ -1, & x_a - x_b < 0, \end{cases} \quad (2)$$

where sgn is the sign function, which takes the values -1 , 0 , and $+1$.

The variance of the S statistic parameter ($\text{VAR}(S)$) is calculated using a formula that takes into account possible correlations, as follows:

$$\begin{aligned} \text{VAR}(S) = \\ = \frac{n(n-1)(2n+5) - \sum_{p=1}^q t_p(t_p-1)(2t_p+5)}{18}. \end{aligned} \quad (3)$$

The Z statistic is calculated using the following formula:

$$Z = \begin{cases} \frac{S-1}{\sqrt{\text{VAR}(S)}}, & S > 0 \\ 0, & S = 0 \\ \frac{S+1}{\sqrt{\text{VAR}(S)}}, & S < 0 \end{cases}, \quad (4)$$

where q is the number of tied groups, n is the number of values in the data series, and t_p is the number of values in p th group.

A positive value of Z indicates an upward trend, whereas a negative value suggests a downward trend.

The Sen method can be used to estimate the linear trend slope.

In order to estimate the trend slope, we need to calculate the slopes Q_i for all pairs of values, as follows:

$$Q_i = \frac{x_a - x_b}{a - b}, \quad (5)$$

where $a > b$, while x_a and x_b are the data values at time points a and b , respectively.

If the time series contains n values of x_a , then the number of slope estimates Q_i is as follows:

$$N = \frac{n(n-1)}{2}. \quad (6)$$

The Sen slope estimate is the median value of these N values of Q_i , ranked in order from smallest to largest. The Sen slope estimate is calculated using the following formula:

$$Q = Q_{[(N+1)/2]}, \text{ if } N \text{ is odd,}$$

or

$$Q = \frac{1}{2} \left(Q_{[N/2]} + Q_{[(N+2)/2]} \right), \text{ if } N \text{ is even.} \quad (7)$$

To study the relationship and the possibility of combining different data series, a Spearman's rank correlation analysis can be conducted [8], since the sample has a non-normal (non-Gaussian) data distribution.

CRITICAL ANALYSIS OF DATA PROCESSING USING ALGORITHM 1995

Up until 2023, the primary findings from the IBMoN network of measurements were published in the RosHydroMet and IGCE logbooks for geophysical years rather than calendar years. This made it difficult to compare these data with those released by other RosHydroMet organizations. Since joint analysis of information from various observational networks must

¹⁰ Paramonov S.G. *Background Pollution of the Atmosphere in the European Territory of Russia*: Dis. Sci. Thesis (Geogr.). Moscow: 1994, 20 p. (In Russ.).

be carried out, it is crucial to ensure the possibility of converting data from the IBMoN network into calendar-year average values.

However, due to the annual geometric mean being calculated in step 4 of Algorithm 1995 based on the geometric means of the cold and warm seasons rather than the results of individual measurements (i.e., to account for seasonal variations), it is not possible to use this algorithm version to perform the transition from averaging over a geophysical year to a calendar year. The reason why this approach is not possible when calculating for a calendar year is due to the need to shift the averaging periods.

Since 1995, advancements in digital technologies have provided for the implementation of additional constraints and validation during data digitization. This has helped to prevent certain types of error, such as accidental and systematic mistakes, including the entry of random characters (e.g., alphabetic), incorrect formatting (e.g., dates in American or British formats), and others. The use of additional validation measures also facilitates the detection of erroneous data. For instance, errors resulting from incorrect decimal placements (e.g., significant increases or decreases in values) can be identified visually through the use of graphs. Once errors have been identified, operators can correct the relevant datasets accordingly.

Advances in analytical techniques and instrumentation have reduced the LOD to identify increasingly lower concentrations of pollutants. However, each technique has an intrinsic LOD consisting in the lowest amount or concentration of a target substance in a test sample that can be reliably distinguished from zero by a particular measurement system. A significant feature of data collected from an extensive background monitoring network is the typically low level of measured pollutant concentrations, which may fall below the detection limit of the technique. However, Algorithm 1995 does not provide a standard procedure for handling below LOD values. Previously, human error from data entry operators entering data into a logbook or expert reviewers validating the data may have resulted in these values being deleted, assigned to zero, or assigned to half of the instrument's LOD. This can lead to discrepancies in calculated averages, even if the original data is identical.

On this basis, we consider which steps in Algorithm 1995 require modernization. By “modernization,” we mean the replacement, addition, or alteration of steps, as well as modifications to the content of those steps (e.g., approaches, mechanisms, data structures, input and output data, and parameters).

Step 1 “Collection of experimental data” does not require any changes, due to the values being obtained according to approved chemical analysis methods.

Step 2 “Data entry into paper and electronic logbooks” requires some updating. The primary goal of this step is to ensure data integrity by creating paper and electronic copies of the dataset that are completely consistent with one another. Simultaneously, measures should be taken to prevent random and systematic errors in the data recording and entry process.

At present, the storage of paper logbooks is regulated by GOST ISO/IEC 17025-2019¹¹. However, the use of electronic logs eliminates the risk of accidental mistakes when filling out spreadsheets due to eliminating the entry of duplicate data.

Step 3 “Data validation” requires revision. Due to the change in the sampling time period from a geophysical year to a calendar year, all of the time series will need to be re-formatted for the entire observation period.

Algorithm 1995 lacks a data imputation mechanism, i.e., a method of replacing missing values. Consequently, the amount of experimental data can vary significantly (from 5–10 to 20–31 values per month). This may be due to a lack of data for a particular period, or to limitations in the analytical methods. Since the concentration of pollutants in ambient air often exhibits pronounced seasonal variability [9–11], significant differences in the number of measurements obtained at different times of the year can lead to inaccurate calculation results. Limitations of analytical methods are often related to their sensitivity or the LOD for the target substance. Therefore, a unified approach is needed to replace below LOD values with corresponding established values for all measured substances [12, 13].

In Algorithm 1995, outliers were eliminated through the data validation process. However, the procedure for assigning values to emissions was not specified, resulting in additional possible errors in the calculations. While the need to eliminate outliers is due to the application of the geometric mean, which is sensitive to extreme values [14, 15], concentrations that are two, three, or more magnitudes higher than the main dataset, e.g., those associated with the emergence of local pollution sources or meteorological phenomena, may occur the field of atmospheric air pollution. Therefore, this step should be modernized by incorporating a data imputation stage to replace missing and below LOD values, along with a unified approach to handling outliers.

Step 4 “Calculation of average values” requires modification. In this step, the geophysical annual average values must be converted into values representing the calendar year in accordance with RosHydroMet requirements. Since the Algorithm 1995 approach is not feasible for this purpose, it is proposed that geometric

¹¹ GOST ISO/IEC 17025-2019. Interstate Standard. *General requirements for the competence of testing and calibration laboratories*. Moscow: Standartinform; 2021. (In Russ.).

means be replaced with medians [16, 17], which are more resistant to outliers [18, 19].

Step 5 “Augmentation of long-term data series” should be updated due to changes in the source data calculations. Thus, according to new the recommendations, long-term data series, which were also previously based on geometric means, should consist of median values when processing measurement results.

Step 6 “Data publication in paper form” should be adjusted in the light of technological developments in online and electronic media.

Step 2.1 “Updating the long-term raw data series” should also be incorporated into Algorithm 2025 due to the fact that the development of methods and technologies, along with the accumulation of data, often necessitates revisiting discarded values or re-calculating based on the raw data. For this reason, it is necessary to update the original algorithm with the one illustrated in Fig. 2.

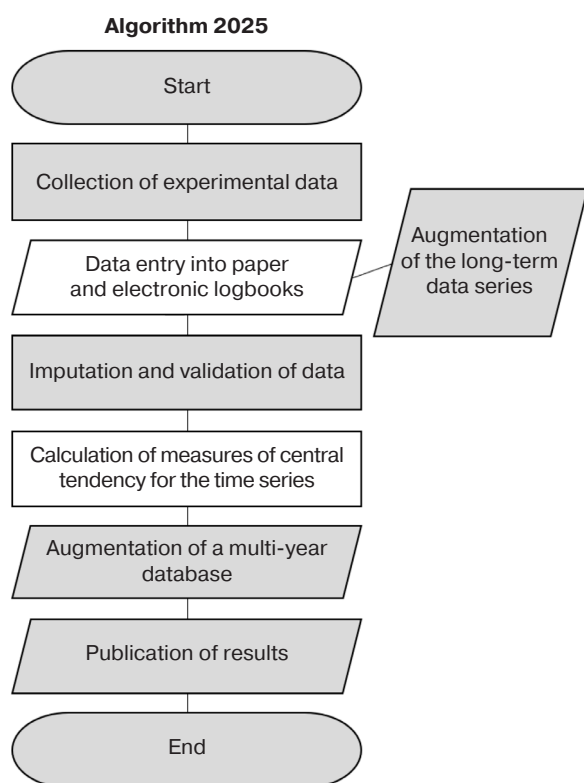


Fig. 2. Modernized algorithm for processing time series data on the pollutant concentrations in atmospheric air, with the steps highlighted in white and detailed further in the study

ANALYSIS AND DISCUSSION OF RESULTS

The paper examines the possibility of modifying Steps 2 and 4 (“Data entry into paper and electronic logbooks” and “Calculation of measures of central tendency for the time series,” respectively) and its impact

on the final outcomes. The efficacy of modernizing the algorithm’s steps is assessed using data collected from the Astrakhan Nature Reserve. The process of data imputation has not yet been considered due to the need to use the same source data for comparative analysis.

Impact of modernization of Step 2 in Algorithm 1995 on the quality of data collected on atmospheric air pollution

Modern electronic spreadsheet editors such as *Microsoft Excel*, *LibreOffice Calc*, and *Gnumeric*¹² can be used to eliminate common errors that occur during manual data entry such as typos, incorrect formatting, and re-entry. For instance, *Microsoft Excel* offers several methods for restricting the use of certain characters, such as setting the allowed data type or using a formula to check if a cell contains a specific character. If the character is present, the formula generates an error and prevents further input. Additionally, numerical values can be visually represented or conditionally formatted to identify outliers, intentionally false values, or data requiring verification.

Therefore, using modern information processing and storage capabilities eliminates the need for re-filling tables at Step 2 involving a second operator, without compromising the accuracy of table filling. In certain cases, eliminating the “human element” in the modernization of Step 2 will reduce accidental mistakes and consequently increase the accuracy of the tables.

Impact of modernization of Step 4 in Algorithm 1995 on the quality of data collected on atmospheric air pollution

In some cases, extremely high values may occur in data series on the concentration of pollutants in the atmosphere. In such cases, it is not always possible to determine whether such values are erroneous based on experimental data. For example, an exceptionally high value could be a valid consequence of the presence of an additional local pollution source or the long-range transport of pollutants from a major emission source. As such, in the case of seasonal emission sources, an assessment based solely on multi-year datasets may lead to an incorrect conclusion that the data are erroneous. Therefore, the unconditional removal of exceptionally high values from initial data sets would be inappropriate due to the possibility of information loss and consequent distortion of observation results [20].

¹² <https://gnome.pages.gitlab.gnome.org/gnumeric-web/>. Accessed May 22, 2026.

Outliers are primarily excluded from calculations as a means of deriving an integral characteristic that represents the level of a particular pollutant characteristic of the year as a whole. In Algorithm 1995, the geometric mean is used as this characteristic. However, due to the is sensitivity of this indicator to extreme values, calculating the geometric mean requires the removal of outliers. In contrast, the median is a more robust characteristic of a data series due to being less affected by extreme values [15].

In 2023, RosHydroMet's IBMoN determined the necessity of switching from averaging over a geophysical year to the period a calendar year. It was in connection with this decision that the possibility of using the annual median instead of the annual geometric mean was also explored.

A review of the datasets on pollutant concentrations in the atmosphere revealed that the geometric mean and median values generally fall within the same range, with only slight differences. However, an exception was noted in the figures for 2004, where the median value calculated for the calendar year ($23.13 \mu\text{g}/\text{m}^3$) was significantly higher than the geometric mean for the geophysical year ($8.39 \mu\text{g}/\text{m}^3$). This is likely due to an increase in sulfur dioxide concentrations during October–December of 2004 and the corresponding shift in the peaks of sulfur dioxide concentration. For other pollutants and time periods, there have also been shifts in peak concentrations and intensities, although they continue fall within the same ranges.

The feasibility of changing the averaging period—that is, transitioning from calculating averages based on a geophysical year to calculating them based on a calendar year—has been assessed using Spearman's correlation analysis. Following a ranking of the dataset values, correlation analysis was performed on the resulting ranks to derive the correlation coefficient r using the following formula:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (8)$$

where n is the number of values in the data series; x_i and y_i are the individual sample points numbered i ; $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ and $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ are the mean values of the samples.

The results of the correlation analysis are presented in Table 1 below.

The findings from the correlation analysis demonstrate a strong correspondence between the different averaging methods and calculation types. Based on preliminary estimates, the use of median values is likely to maintain data accuracy while also allowing for the processing of measurement results without the need to exclude outliers. This is a significant benefit of the proposed approach.

The impact of the modernization of step 4 in Algorithm 1995 on determining trends in air pollution data

Time series generated using Algorithm 1995 are used to further analyze trends and changes in pollutant levels in atmospheric air, the results of which are published in relevant publications [4]. To assess the effectiveness of the updated algorithm, the time series obtained should be evaluated for trends. As shown in Table 2, the use of data averaged over a calendar year instead of a geophysical year affects the ability to identify significant trends. The processing of empirical data on different pollutants (suspended particles, sulfur dioxide,

Table 1. Results of the Spearman rank correlation analysis for central trends in pollutant concentrations in ambient air

Correlation coefficient	TSP	SO ₂	SO ₄ ²⁻	Pb
Between geometric means when using different averaging periods (calendar and geophysical year)	0.83	0.94	0.89	0.74
Between geometric means and median values for a single averaging period (calendar year)	0.96	0.96	0.98	0.98

Table 2. Results of trend analysis of data obtained during processing in accordance with Algorithm 1995 and with the modified Algorithm 1995

Substance	Algorithm 1995	Modified Algorithm 1995
TSP	—	—
SO ₂	$\alpha < 0.01$, downward trend	$\alpha < 0.01$, downward trend
SO ₄ ²⁻	—	$\alpha < 0.05$, downward trend
Pb	—	$\alpha < 0.01$, downward trend

sulfates, and lead) using the updated algorithm revealed an increase in the number of statistically significant trends. However, the trend of previously identified trends remained unchanged.

A further advantage of switching to median calculations is to reduce the number of computational operations. More specifically, the calculation of annual geometric means based on Algorithm 1995 uses the geometric means obtained for the cold and warm seasons to perform the calculation. The entire data series can then be used to determine the median value.

CONCLUSIONS

An algorithm for processing experimental data on the concentration of pollutants in ambient air, which has been used since 1995 by the IBMoN of RosHydroMet, is discussed. The impossibility of obtaining data in the recommended format without modifying some steps of Algorithm 1995 is substantiated. A revised data processing algorithm, Algorithm 2025, has been developed to address this issue. The use of Algorithm 2025 should result in the production of data that meets the requirements of RosHydroMet while reducing the impact of noise and variability in time series on background concentrations of air pollutants.

Methods for modernizing selected steps of Algorithm 1995 connected with data entry into both paper and electronic logbooks have been discussed in detail and substantiated along with a determination of indicators of central tendencies. The effectiveness of a partially modified version of Algorithm 1995 has been tested on the example of a time series of data collected for the territory of the Astrakhan Nature Reserve. The modified algorithm can be used to convert data to

standard averaging periods (typical data format) used in RosHydroMet monitoring networks without losing accuracy. The modernization of these algorithm steps leads to a reduction in statistical noise to eliminate up to 70% of the main types of random and systematic errors that occur during manual data entry.

For the development of appropriate approaches to modifying the steps of the algorithm that are not covered in this study, as well as its adaptation for the Algorithm 2025 version, more work is required. The implementation of Algorithm 2025 for the processing of complex background monitoring data is required in order to preserve original data, further reduce random and systematic errors, standardize the process, and eliminate the influence of selective data usage.

ACKNOWLEDGMENTS

The research was partially implemented within the framework of the R&D project “Development and modernization of methods and technologies of complex background monitoring and comprehensive assessment of condition and pollution of the environment of the Russian Federation and its dynamics (based on the integrated results of RosHydroMet monitoring networks)” (Registration number 125031703848-8).

Authors' contributions

V.A. Ivanov—analysis and interpretation of data, development of the 2025 algorithm, and writing the text of the manuscript.

S.G. Paramonov—data collection, consultation on the 1995 algorithm, discussion of the results.

E.S. Zhigacheva—discussion of the results, editing the text of the manuscript.

E.A. Pozdnyakova—formulation of the task, discussion of the results.

REFERENCES

1. Gromov S.A., Paramonov S.G. Current status and prospects for the development of integrated background monitoring of environmental pollution. *Problemy ehkologicheskogo monitoringa i modelirovaniya ehkosistem = Problems of Ecological Monitoring and Ecosystem Modeling*. 2015;26(1):205–221 (in Russ.). <https://elibrary.ru/uybmod>
2. Burtseva L.V., Golubeva N.I., Konkova E.S. Development of quality assurance methods of heavy metal measurements in the atmospheric air and precipitations in the integrated background monitoring system of roshydromet. *Sistemy kontrolya okruzhayushchei sredy = Monitoring Systems of Environment (MSoE)*. 2019;4(38):20–26 (in Russ.). <https://doi.org/10.33075/2220-5861-2019-4-20-26>
3. Chernogaeva G.M. (Ed). *Obzor sostoyaniya i zagryazneniya okruzhayushchei sredy v Rossiiskoi Federatsii za 2021 god (Review of the State and Pollution of the Environment in the Russian Federation for 2021)*. Moscow: RosHydroMet; 2022, 219 p. (in Russ.). https://www.meteorf.gov.ru/upload/iblock/dc8/Obzor_2021.pdf
4. Chernogaeva G.M. (Ed). *Tendentsii i dinamika sostoyaniya i zagryazneniya okruzhayushchei sredy v Rossiiskoi Federatsii po dannym mnogoletnego monitoringa za poslednie 10 let: Analiticheskii obzor (Trends and Dynamics of the State and Pollution of the Environment in the Russian Federation According to Long-Term Monitoring Data over the Past 10 Years: Analytical Review)*. Moscow: RosHydroMet; 2017, 47 p. (in Russ.). ISBN 978-5-9500881-2-4
5. Burtseva L.V., Konkova E.S. Assessment of atmospheric pollution with copper in background areas of the Europe. *Problemy ehkologicheskogo monitoringa i modelirovaniya ehkosistem = Problems of Ecological Monitoring and Ecosystem Modeling*. 2020;31(1-2):14–33 (in Russ.). <https://doi.org/10.21513/0207-2564-2020-1-14-33>

6. Salmi T., Maata A., Antilla P., Ruoho-Airola T., Amnell T. *Detecting Trends of Annual Values of Atmospheric Pollutants by the Mann–Kendall Test and Sen's Slope estimates—the Excel Template Application Makesens*. Report code: FMI-AQ-31. Helsinki: Finnish Meteorological Institute; 2002, 36 p.
7. Perikhanyan Ye., Shahnazaryan G., Gabrielyan A. Long term trends of wet deposition and atmospheric concentrations of nitrogen and sulfur compounds at EMEP site in Armenia. *J. Atmos. Chem.* 2020;77(3):101–116. <https://doi.org/10.1007/s10874-020-09408-3>
8. Pavlova K.V. Application of Spearman's rank correlation coefficient in statistical research. *Bezopasnost'. Upravlenie. Iskusstvennyi intellekt.* 2023;3(3):28–32 (in Russ.). <https://elibrary.ru/ubhmwg>
9. Golobokova L.P., Berdashkinova O.I., Loskutova M.A., et al. Long-term Studies of Atmospheric Aerosol Chemical Composition at the Cape Baranov Ice Base Station. *Atmos. Ocean Opt.* 202437(1):14–23. <https://doi.org/10.1134/S1024856023700033> [Original Russian Text: Golobokova L.P., Berdashkinova O.I., Loskutova M.A., Rize D.D., Onishchuk N.A., Sakerin S.M., Turchinovich Yu.S. Long-term Studies of Atmospheric Aerosol Chemical Composition at the Cape Baranov Ice Base Station. *Optika atmosfery i okeana.* 2023;36(11):874–882 (in Russ.). <https://doi.org/10.15372/AOO20231102>]
10. Kukkonen J., Konttinen M., Bremer P., Salmi T., Saari H. The seasonal variation of urban air quality in northern European conditions. *Int. J. Environ. Pollut.* 2000;14(1-6):480–487. <https://doi.org/10.1504/IJEP.2000.000571>
11. Zareba M., Weglinska E., Danek T. Air pollution seasons in urban moderate climate areas through big data analytics. *Sci. Rep.* 2024;14(1):3058. <https://doi.org/10.1038/s41598-024-52733-w>
12. Sköld M. Trend detection with non-detects in long-term monitoring, a mixed model approach. *Environ. Monit. Assess.* 2023;195(6):663. <https://doi.org/10.1007/s10661-023-11285-8>
13. Ehksperiandova L.P., Belikov K.N., Khimchenko S.V., et al. Once again about determination and detection limits. *J. Anal. Chem.* 2010;65(3):223–228. <https://doi.org/10.1134/S1061934810030020> [Original Russian Text: Ehksperiandova L.P., Belikov K.N., Khimchenko S.V., Blank T.A. Once again about determination and detection limits. *Zhurnal analiticheskoi khimii.* 2010;65(3):229–234 (in Russ.). <https://elibrary.ru/kkilqd>]
14. Gilbert R.O. *Statistical Methods for Environmental Pollution Monitoring*. New York: Van Nostrand Reinhold Company Inc.; 1987, 334 p.
15. Rozenberg G.S., Shitikov V.K., Brusilovskii G.N. *Ehkologicheskoe prognozirovaniye (Funktsional'nye prediktory vremennykh ryadov) (Environmental Forecasting (Functional Predictors of Time Series))*. Tol'yatti; 1994, 182 p. (In Russ.). <https://elibrary.ru/zumjps>
16. Zheksenbaeva A.K. Statistical characteristics of rainfall in northern Kazakhstan during the vegetative period. *Nauka, novye tekhnologii i innovatsii Kyrgyzstana = Science, New Technologies and Innovations of Kyrgyzstan.* 2016;5:110–117 (in Russ.). <https://elibrary.ru/wdfhgb>
17. Bruce P., Bruce A. *Prakticheskaya statistika dlya spetsialistov Data Science (Practical Statistics for Data Scientists)*: transl. from Engl. BKHV-Petersburg; 2018, 302 p. (In Russ.). ISBN 978-5-9775-3974-6 [Bruce P., Bruce A. *Practical Statistics for Data Scientists*. O'Reilly; 2017, 368 p.]
18. Bokov V.N., Vorobyev V.N., Serebriyskiy I.A. Evaluation of the background pollution of atmosphere in St. Petersburg and suburbs. *Uchenye zapiski Rossiiskogo gosudarstvennogo gidrometeorologicheskogo universiteta = Proceedings of the Russian State Hydrometeorological University.* 2011;20:36–41 (in Russ.). <https://elibrary.ru/onoexv>
19. Lenskaya O.Yu., Kulyabova A.B. On the methods of analysis of irregular observations of the composition of atmospheric air pollution. *Vestnik Chelyabinskogo gosudarstvennogo univrsiteta = Bulletin of Chelyabinsk State University.* 2007;6:80–84 (in Russ.). <https://elibrary.ru/mvmwmp>
20. Van Zoest V.M., Stein A., Hoek G. Outlier Detection in Urban Air Quality Sensor Networks. *Water, Air, and Soil Pollution.* 2018;229(4):111. <https://doi.org/10.1007/s11270-018-3756-7>

СПИСОК ЛИТЕРАТУРЫ

1. Громов С.А., Парамонов С.Г. Современное состояние и перспективы развития комплексного фонового мониторинга загрязнения природной среды. *Проблемы экологического мониторинга и моделирования экосистем.* 2015;26(1): 205–221. <https://elibrary.ru/uybmod>
2. Бурцева Л.В., Голубева Н.И., Конькова Е.С. Разработка методов обеспечения качества измерений тяжелых металлов в атмосферном воздухе и осадках в системе комплексного фонового мониторинга Росгидромета. *Системы контроля окружающей среды.* 2019;4(38):20–26. <https://doi.org/10.33075/2220-5861-2019-4-20-26>
3. *Обзор состояния и загрязнения окружающей среды в Российской Федерации за 2021 год*; под ред. Г.М. Черногаевой. М.: Росгидромет; 2022, 219 с. https://www.meteorf.gov.ru/upload/iblock/dc8/Obzor_2021.pdf
4. *Тенденции и динамика состояния и загрязнения окружающей среды в Российской Федерации по данным многолетнего мониторинга за последние 10 лет: Аналитический обзор*; под ред. Г.М. Черногаевой. М.: Росгидромет; 2017, 47 с. ISBN 978-5-9500881-2-4
5. Бурцева Л.В., Конькова Е.С. Оценка загрязнения атмосферы медью в фоновых районах Европы. *Проблемы экологического мониторинга и моделирования экосистем.* 2020;31(1-2):14–33. <https://doi.org/10.21513/0207-2564-2020-1-14-33>
6. Salmi T., Maata A., Antilla P., Ruoho-Airola T., Amnell T. *Detecting Trends of Annual Values of Atmospheric Pollutants by the Mann–Kendall Test and Sen's Slope estimates—the Excel Template Application Makesens*. Report code: FMI-AQ-31. Helsinki: Finnish Meteorological Institute; 2002, 36 p.

7. Perikhanyan Ye., Shahnazaryan G., Gabrielyan A. Long term trends of wet deposition and atmospheric concentrations of nitrogen and sulfur compounds at EMEP site in Armenia. *J. Atmos. Chem.* 2020;77(3):101–116. <https://doi.org/10.1007/s10874-020-09408-3>
8. Павлова К.В. Применение коэффициента ранговой корреляции Спирмена в статистических исследованиях. *Безопасность. Управление. Искусственный интеллект.* 2023;3(3):28–32. <https://elibrary.ru/ubhwmwg>
9. Голобокова Л.П., Бердашкинова О.И., Лоскутова М.А., Ризе Д.Д., Онишук Н.А., Сакерин С.М., Турчинович Ю.С. Результаты многолетних исследований химического состава аэрозоля в атмосфере на стационаре «Ледовая база «Мыс Баранова». *Оптика атмосферы и океана.* 2023;36(11):874–882. <https://doi.org/10.15372/AOO20231102>
10. Kukkonen J., Konttinen M., Bremer P., Salmi T., Saari H. The seasonal variation of urban air quality in northern European conditions. *Int. J. Environ. Pollut.* 2000;14(1-6):480–487. <https://doi.org/10.1504/IJEP.2000.000571>
11. Zareba M., Weglinska E., Danek T. Air pollution seasons in urban moderate climate areas through big data analytics. *Sci. Rep.* 2024;14(1):3058. <https://doi.org/10.1038/s41598-024-52733-w>
12. Sköld M. Trend detection with non-detects in long-term monitoring, a mixed model approach. *Environ. Monit. Assess.* 2023;195(6):663. <https://doi.org/10.1007/s10661-023-11285-8>
13. Экспериандова Л.П., Беликов К.Н., Химченко С.В., Бланк Т.А. Еще раз о пределах обнаружения и определения. *Журнал аналитической химии.* 2010;65(3):229–234. <https://elibrary.ru/kkilqd>
14. Gilbert R.O. *Statistical Methods for Environmental Pollution Monitoring.* New York: Van Nostrand Reinhold Company Inc.; 1987, 334 p.
15. Розенберг Г.С., Шитиков В.К., Брусиловский Г.Н. *Экологическое прогнозирование (Функциональные предикторы временных рядов).* Тольятти; 1994, 182 с. <https://elibrary.ru/zumjps>
16. Жексенбаева А.К. Статистические характеристики осадков в северном Казахстане в вегетационный период. *Наука, новые технологии и инновации Кыргызстана.* 2016;5:110–117. <https://elibrary.ru/wdfhgb>
17. Брюс П., Брюс Э. *Практическая статистика для специалистов Data Science.* БХВ-Петербург; 2018, 302 с. ISBN 978-5-9775-3974-6
18. Боков В.Н., Воробьев В.Н., Серебрицкий И.А. Оценка фоновое загрязнения атмосферного воздуха в Санкт-Петербурге и его пригородах. *Ученые записки Российского государственного гидрометеорологического университета.* 2011;20:36–41. <https://elibrary.ru/onoexv>
19. Ленская О.Ю., Кулябова А.Б. О способах анализа нерегулярных наблюдений за составом загрязнения атмосферного воздуха. *Вестник Челябинского государственного университета.* 2007;6:80–84. <https://elibrary.ru/mvmwmp>
20. Van Zoest V.M., Stein A., Hoek G. Outlier Detection in Urban Air Quality Sensor Networks. *Water, Air, and Soil Pollution.* 2018;229(4):111. <https://doi.org/10.1007/s11270-018-3756-7>

About the Authors

Vladimir A. Ivanov, Postgraduate Student, Department of Modeling of Radiophysical Processes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia); Junior Researcher, Department of Measurements of Background Chemical Pollution of Continental Natural Systems, Yu.A. Izrael Institute of Global Climate and Ecology (20B, Glebovskaya ul., Moscow, 107258 Russia). E-mail: dunkelheit2015@yandex.ru. ResearcherID AFA-9383-2022, RSCI SPIN-code 7426-4230, <https://orcid.org/0009-0005-1001-8575>

Sergey G. Paramonov, Cand. Sci. (Geogr.), Leading Researcher, Environmental Pollution Assessment Department, Yu.A. Izrael Institute of Global Climate and Ecology (20B, Glebovskaya ul., Moscow, 107258 Russia). E-mail: sgpar@mail.ru. Scopus Author ID 35581383300, ResearcherID Z-5928-2019, RSCI SPIN-code 7005-0408, <https://orcid.org/0009-0004-6269-1262>

Ekaterina S. Zhigacheva, Cand. Sci. (Agricult.), Researcher, Department of Measurements of Background Chemical Pollution of Continental Natural Systems. Yu.A. Izrael Institute of Global Climate and Ecology (20B, Glebovskaya ul., Moscow, 107258 Russia). E-mail: kosjatko@gmail.com. Scopus Author ID 57323176200, ResearcherID U-4242-2018, RSCI SPIN-code 1675-8876, <https://orcid.org/0000-0003-2477-8653>

Ekaterina A. Pozdnyakova, Head of Department, Senior Researcher, Department of Measurements of Background Chemical Pollution of Continental Natural Systems, Yu.A. Izrael Institute of Global Climate and Ecology (20B, Glebovskaya ul., Moscow, 107258 Russia). E-mail: e.pozdnyakova@igce.ru. Scopus Author ID 59693836300, ResearcherID AAA-4825-2019, RSCI SPIN-code 3316-9451, <https://orcid.org/0009-0009-3169-9575>

Об авторах

Иванов Владимир Алексеевич, аспирант, кафедра «Моделирование радиофизических процессов», Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78); младший научный сотрудник, отдел измерений фонового химического загрязнения континентальных природных систем, ФГБУ «Институт глобального климата и экологии имени академика Ю.А. Израэля» (107258, Россия, Москва, ул. Глебовская, д. 20Б). E-mail: dunkelheit2015@yandex.ru. ResearcherID AFA-9383-2022, SPIN-код РИНЦ 7426-4230, <https://orcid.org/0009-0005-1001-8575>

Парамонов Сергей Геннадьевич, к.г.н., ведущий научный сотрудник, отдел оценки загрязнения окружающей среды, ФГБУ «Институт глобального климата и экологии имени академика Ю.А. Израэля» (107258, Россия, Москва, ул. Глебовская, д. 20Б). E-mail: sgpar@mail.ru. Scopus Author ID 35581383300, ResearcherID Z-5928-2019, SPIN-код РИНЦ 7005-0408, <https://orcid.org/0009-0004-6269-1262>

Жигачева Екатерина Сергеевна, к.с.-х.н., научный сотрудник, отдел измерений фонового химического загрязнения континентальных природных систем, ФГБУ «Институт глобального климата и экологии имени академика Ю.А. Израэля» (107258, Россия, Москва, ул. Глебовская, д. 20Б). E-mail: kosjatko@gmail.com. Scopus Author ID 57323176200, ResearcherID U-4242-2018, SPIN-код РИНЦ 1675-8876, <https://orcid.org/0000-0003-2477-8653>

Позднякова Екатерина Александровна, заведующий отделом, старший научный сотрудник, отдел измерений фонового химического загрязнения континентальных природных систем, ФГБУ «Институт глобального климата и экологии имени академика Ю.А. Израэля» (107258, Россия, Москва, ул. Глебовская, д. 20Б). E-mail: e.pozdnyakova@igce.ru. Scopus Author ID 59693836300, ResearcherID AAA-4825-2019, SPIN-код РИНЦ 3316-9451, <https://orcid.org/0009-0009-3169-9575>

Translated from Russian into English by K. Nazarov

Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling
Математическое моделирование

UDC 519.2

<https://doi.org/10.32362/2500-316X-2026-14-4-137-151>

EDN JNDMBQ



RESEARCH ARTICLE

Stochastic properties of the dynamics of the first derivative in Cauchy problems

Alexey N. Nazarov[@]

MIREA – Russian Technological University, Moscow, 119454 Russia

[@] Corresponding author, e-mail: a.nazarov06@bk.ru

• Submitted: 07.10.2025 • Revised: 23.01.2026 • Accepted: 23.03.2026

Abstract

Objectives. Advances in information communication technologies, coupled with the unlimited computational capabilities of modern automation systems, make it possible to reformulate and solve differential equations in the form of the Cauchy problem using a random process theory. The present work aims to develop a random process of a stepwise model of the dynamics of an object, whose characteristics and parameters are observed and processed through monitoring. This creates the prerequisites for the formation of new information communication approaches for automated analysis and synthesis of dynamic processes and systems, including Big Data.

Methods. The research took a geometric approach to estimating the probabilities of independent events along with the mean value theorem for integrals and a method for synthesizing stochastic models of dynamic processes for various subject areas based on graphical dependencies of the first derivative.

Results. Based on the methodology, results of the formation of new stochastic dependencies of the integral curve for solving the differential equation of dynamics of an arbitrary object are presented. A new random process of the dynamics of an arbitrary object was developed and investigated. Its parameters and characteristics are formed based on monitoring results in the info-communication space. This creates the prerequisites for automated analysis and synthesis of innovative research approaches to study dynamic processes and systems, including Big Data. A novel application area for the methodology of solving conditional optimization problems, previously utilized in the analysis of packet network teletraffic, was identified.

Keywords: differential equation, complete group of incompatible events, integral curve, model, link, chain, probability distribution density, random process

For citation: Nazarov A.N. Stochastic properties of the dynamics of the first derivative in Cauchy problems. *Russian Technological Journal*. 2026;14(4):137–151. <https://doi.org/10.32362/2500-316X-2026-14-4-137-151>, <https://elibrary.ru/JNDMBQ>

Financial disclosure: The author has no financial or proprietary interest in any material or method mentioned.

The author declares no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Стохастические свойства динамики первой производной в задачах Коши

А.Н. Назаров [®]

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[®] Автор для переписки, e-mail: a.nazarov06@bk.ru

• Поступила: 07.10.2025 • Доработана: 23.01.2026 • Принята к опубликованию: 23.03.2026

Резюме

Цели. Прогресс в области инфокоммуникаций и неограниченные вычислительные возможности современных систем автоматизации позволяют по-новому сформулировать и найти решение дифференциального уравнения в форме задачи Коши с позиций теории случайных процессов. Цель статьи – разработка случайного процесса ступенчатой модели динамики объекта, характеристики и параметры которой наблюдаются и обрабатываются в результате мониторинга.

Методы. Геометрический подход к оценке вероятностей независимых событий. Интегральная теорема о среднем. Метод синтеза стохастических моделей динамических процессов для различных предметных областей на основе графических зависимостей первой производной.

Результаты. Представлены новые методические результаты формирования новых стохастических зависимостей интегральной кривой решения дифференциального уравнения динамики произвольного объекта. Разработан и исследован новый случайный процесс динамики произвольного объекта, параметры и характеристики которого формируются по результатам мониторинга в инфокоммуникационном пространстве, что создает предпосылки автоматизированного анализа и синтеза новых подходов исследований динамических процессов и систем, включая большие данные. Найдена новая область применения методики решения задач условной оптимизации, ранее применявшаяся для исследования телетрафика пакетных сетей.

Выводы. Новые функциональные соотношения являются методической основой создания новых научно-обоснованных инфокоммуникационных систем и технологий и могут быть положены в основу синтеза алгоритмов, методов и моделей динамических систем и процессов в различных предметных областях. Проведенные теоретические исследования носят фундаментальный характер и позволяют считать обоснованным, что предлагаемый автором подход обладает общностью применения для автоматизированного анализа и синтеза динамических систем и средств на основе инфокоммуникационных технологий, включая большие данные.

Ключевые слова: дифференциальное уравнение, полная группа несовместных событий, интегральная кривая, модель, звено, цепь, плотность распределения вероятностей, случайный процесс

Для цитирования: Назаров А.Н. Стохастические свойства динамики первой производной в задачах Коши. *Russian Technological Journal*. 2026;14(4):137–151. <https://doi.org/10.32362/2500-316X-2026-14-4-137-151>, <https://elibrary.ru/JNDMBQ>

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

INTRODUCTION

Government-led initiatives to address dynamic challenges in support of the industrial and military development of the Soviet Union resulted in the development of random process theory [1]. This theory was highly effective due to the capabilities of hardware solutions based on analogue principles, which were relevant in the mid-20th century.

However, the emergence of digital approaches in the second half of the 20th century and the development of software systems and programming tools created conditions for the relatively rapid development of new hardware and software solutions. This has contributed to the development of new information and communication technologies, as well as geographically distributed large-scale systems based on them.

Cloud technologies provide researchers with practically unlimited computing resources that enable them to solve the most complex computational problems that could previously only be formulated in theoretical science. At the same time, there is a lack of development in fundamental science and engineering, especially in computer science, where heuristic methods dominate. Therefore, while such methods can produce results that may seem convincing, they are not necessarily scientifically sound, and could lead to future global problems.

It is becoming increasingly clear that the technological specifications, protocols, frameworks, and other methods for describing dynamic processes within the internet ecosystem do not provide sufficient confidence that implementation by engineers will lead to the creation of high-quality equipment and systems.

Here, the concept of emergence should be borne in mind: a phenomenon where new properties or behaviors emerge in a system that are not inherent in its individual components. Unexpected results may occur when simple elements interact within a larger system, especially when automation subsystems and new information technologies are introduced.

There are still gaps in the development of the industry related to dynamic processes and systems of an arbitrary nature, which can be described by differential equations. However, these systems are underestimated, inconsistent, and incompatible with the new technological capabilities and phenomena that have emerged as a result of the long-term paradigm shift in information and communication technology. Such gaps are associated with a lack of identified and explored interdisciplinary connections, and the impact of various theoretical and applied methodological advances in computer science, systems analysis and operations research.

Thus, it is crucial to develop formal procedures for analyzing and solving these differential equations, particularly those related to advancements in computer science.

FEATURES OF SOLVING DIFFERENTIAL EQUATIONS USING WELL-ESTABLISHED ANALYTICAL TECHNIQUES

Without going into the detailed methodological aspects of representing dynamic processes as differential equations^{1, 2, 3} and converting them to normal form as the Cauchy problem [2], it should be noted that the study by I.N. Pryadko and L.P. Petrova provides a scientifically rigorous methodological approach for solving the differential equations related to the Cauchy problem, including detailed proofs.

Let us consider the vector differential equation (or normal system, NS):

$$x' = f(x, t), \quad (1)$$

where $x = x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \dots \\ x_n(t) \end{pmatrix}$ is an unknown vector function

depending on a real independent variable, t , which may represent time and lie within a bounded set, and

$f(x, t) = \begin{pmatrix} f_1(x, t) \\ f_2(x, t) \\ \dots \\ f_n(x, t) \end{pmatrix}$ is a vector function.

Along with (1), we consider the initial condition (IC):

$$x_0 = x(t_0) = \begin{pmatrix} x_1(t_0) \\ x_2(t_0) \\ \dots \\ x_n(t_0) \end{pmatrix}. \quad (2)$$

The higher-order differential equation⁴

$$x^{(n)}(t) = F(t, x(t), x'(t), \dots, x^{(n-1)}(t)) \quad (3)$$

¹ Pryadko I.N., Petrova L.P. *Lecture notes, questions and tasks. Differential equations. Part 2. The Cauchy problem.* Voronezh: Voronezh State University; 2015, 287 p. (In Russ.).

² Zenkov A.V. *Systems of differential equations and elements of stability theory: Textbook for students.* Yekaterinburg: USTU – UPI; 2010, 54 p. (In Russ.).

³ Chostkovskaya O.P., Chernyakina I.V., Starinova O.L. *Solution and application of ordinary differential equations and their systems: Methodical manual.* Samara: Samara University; 2023, 126 p.

⁴ Zenkov A.V., 2010.

can be reduced to a system of equations (1) by introducing additional unknown functions

$$x(t_0) = x_0, x'(t_0) = x'_0, \dots, x^{(n-1)}(t_0) = x_0^{(n-1)}.$$

Therefore, unless otherwise specified, we consider the issue as it applies to solving systems NS (1) and IC (2).

The studies investigate the conditions for the existence and uniqueness of the solution to the Cauchy problem when the Lipschitz condition is satisfied on the interval of the domain t , as also studied in^{5, 6} [2].

If the function $F(t, x(t), x'(t), \dots, x^{(n-1)}(t))$ from (3) is holomorphic, i.e., if it can be expanded into a power series in the vicinity of the point $(t_0, x_0, x'_0, \dots, x_0^{(n-1)})$, then the above Cauchy problem has a unique solution, $x = x(t)$, represented by the series $x(t) = \sum_{p=0}^{\infty} B_p (t - t_0)^k$, which can be found using the method of undetermined coefficients [2].

If the point $(t_0, x_0, x'_0, \dots, x_0^{(n-1)})$ is special for the function F , then the solution to the Cauchy problem is found in the form of a generalized power series for a number of important cases [2].

The Cauchy problem for an n th-order ordinary differential equation (3) is equivalent to the problem of constructing a solution to the following sequence of equations:

$$\begin{aligned} y_0(t) &= x_0 + \int_{t_0}^t y_1(\tau) d\tau, \\ y_1(t) &= x'_0 + \int_{t_0}^t y_2(\tau) d\tau, \\ &\dots \\ y_{n-2}(t) &= x_0^{(n-2)} + \int_{t_0}^t y_{n-1}(\tau) d\tau, \\ y_{n-1}(t) &= x_0^{(n-1)} + \int_{t_0}^t F(\tau, y_0(\tau), \dots, y_{n-1}(\tau)) d\tau, \end{aligned}$$

where $x(t) = y_0(x)$, $x^{(k)}(t) = y_k(t)$, $1 \leq k \leq n$.

This sequence of equations can be solved using the iteration method, also known as the method of successive approximations or Picard's method [2].

Let an ordinary differential equation for two variables $x \in X$ and $t \in T$ solvable with respect to the first derivative be given:

$$x' = f(x, t). \quad (4)$$

Picard's method for this equation reduces to solving the following single equation:

$$y(t) = x_0 + \int_{t_0}^t f(y(\tau), \tau) d\tau,$$

for which a sequence of functions, $u_0(t), u_1(t), \dots, u_k(t)$, is constructed [2] according to the following rule:

$$\begin{aligned} u_0(t) &= x_0, \\ u_k(t) &= x_0 + \int_{t_0}^t f(u_{k-1}(\tau), \tau) d\tau. \end{aligned}$$

If the function $f(x, t)$ is continuous and has a continuous partial derivative $\frac{\partial f}{\partial x}$ in the neighborhood of

the point (x_0, t_0) , then a solution to the Cauchy problem

$$\frac{\partial f}{\partial x} = f(x, t), \quad x(t_0) = x_0$$

exists, is unique, and can be represented as a uniform limit

$$x(t) = \lim_{p \rightarrow \infty} u_p(t) \quad (5)$$

or as the following convergent series [2]:

$$x(t) = x_0 + \sum_{p=1}^{\infty} [u_p(t) - u_{p-1}(t)]. \quad (6)$$

The above considerations regarding Picard's method are applicable to NS (1) in multidimensional cases. Higher-order derivatives are derived from lower-order derivatives, and this determines the generality with which the methodological recommendations provided can be applied to the solution of differential equations with second- and higher-order derivatives.

Further research into solving the Cauchy problem may benefit from the application computational math techniques to approximate functions that describe the behavior of fast-moving processes or real-time functions. The transition to numerical methods for representing and manipulating real-time data has been made facilitated by new computational mathematical techniques that utilize powerful, high-precision signal processors capable of performing floating-point calculations, as well as high-capacity random-access memory (RAM) and read-only memory (ROM) devices. These advances in computational methods have made it possible to develop techniques for

⁵ Pryadko I.N., Petrova L.P. *Lecture notes, questions and tasks. Differential equations. Part 2. The Cauchy problem.* Voronezh: Voronezh State University; 2015, 287 p. (In Russ.).

⁶ Chostkovskaya O.P., Chernyakina I.V., Starinova O.L. *Solution and application of ordinary differential equations and their systems: Methodical manual.* Samara: Samara University; 2023, 126 p. (In Russ.).

representing and processing real-time information, which was previously hindered by the lack of suitable hardware capabilities.

A range of tools and techniques developed for automated measurement and analysis of the properties of modern dynamic systems enable the experimental evaluation of these limits for objects in diverse subject areas.

GRAPHICAL FEATURES OF THE FIRST DERIVATIVE DYNAMICS

The studies [3–8] explored aspects of solving Cauchy problems for differential equations through analytical and computational techniques at a new technological level. This has enabled the development of mathematical methods for determining values of the first-order derivative of a system's dynamics, which have been known since the Middle Ages.

Let us consider a differential equation formulated as a Cauchy problem with two variables, $x \in X$, $t \in [t_0, T]$, solved for the first derivative:

$$x' = f(x, t), \quad (7)$$

including multidimensional cases.

As demonstrated in [3–8], the solution to the Cauchy problem (7) may be achieved through the use of spline interpolation, as based on the findings of monitoring and analyzing experimental data. This technique has been successfully employed in the investigation of the features of packet teletraffic in contemporary information and communication networks [9, 10].

Note. To avoid methodological errors and ensure the accuracy of subsequent models, when plotting graphs in straight Cartesian coordinates with a metric system, the following mutually inverse method (bijection) will be used by default; here, the graph coordinates are generated by multiplying the actual coordinate value of the dynamic object with the dimension of the object domain. This product is then multiplied by an I (coordinate) function with a dimension that is opposite to that of the original coordinate value. Finally, the product is multiplied by the dimensionless scaling factor K (coordinate value), which is determined by the values of the axes of the graph. The scaling factor, which depends on the absolute value of the coordinates of the dynamic object, is chosen separately for each case to ensure clear visual representation of the graphical dependencies.

Example. Let $x' = x'(t)$ with a dimension of m/s. To plot the graph of $x' = x'(t)$ with axis markings: 1 mm $\leftrightarrow$ 2 m on the y -axis and 1 mm $\leftrightarrow$ 3 s on the x -axis, the formulas $x'_{\text{ord}} = x'I(x')K(x')$, $K(x') = 1/2$ and

$t_{\text{abs}} = tI(t)K(t)$, $K(t) = 1/3$ are obtained. These can then be used to plot the graph of $x'(t)$.

Geometrically, the general solution to the first-order differential equation (7) can be described as a family of curves in the xx_0t plane, depending on an arbitrary constant C . These curves are referred to as integral curves of the given differential equation.

A specific solution corresponds to a particular integral curve that passes through a given initial point. The trajectory of this curve in a Cartesian coordinate system (Fig. 1) may consist of geometrically complex segments due to the varying dynamics of the first derivative over the observation interval $[t_0, T]$. The differences between these segments arise from the behavior of the first derivative's values, which can either increase, remain constant, decrease down to zero, or exhibit negative (reversible) behavior. This property is relevant to the topic of the paper.

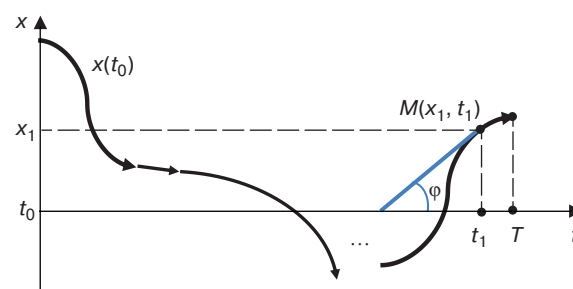


Fig. 1. Graph of a particular solution (integral curve) of a first-order differential equation. $M(x_1, t_1)$ is the point of tangency to the integral curve. The tangent intersects the horizontal axis at an angle of $\varphi = \arctg M(x_1, t_1)$. The initial condition is $x(t_0)$, and the observation interval is $[t_0, T]$

For each point tangent to the integral curve within the domain of definition of $f(x, t)$, equation (7) determines the value of the derivative x' . This specifies the slope of the tangent line to the integral curve that passes through this point and intersects the horizontal axis at an angle of $\varphi = \arctg M(x_1, t_1)$, as shown in Fig. 1.

Therefore, differential equation (7) allows us to graphically construct a set of directions—or, equivalently, determine a field of directions or a field of isoclines [2].

Advances in computer science form the basis for modern automated engineering solutions to estimate the dynamics of the first derivative (velocity) of a physical object. These solutions utilize the results of processing field measurements of isoclines [11–13].

Further investigation requires plotting the graph $G[t_0, T]$ of the dynamics of the first derivative over the observation interval $[t_0, T]$ in the $x'x_0t_0t$ plane. This can be accomplished in two ways. Firstly, if the

trajectory x' allows for the analytical solution^{7, 8} [2], then the graph $G[t_0, T)$ should be plotted based on these analytical dependencies. Secondly, if the curve x' is obtained from the results of processing measurement data and monitoring [3–13], then the graph $G[t_0, T)$ should be drawn based on those results.

Figure 2 shows a possible, highly curvilinear graph $G[t_0, T)$ of changes in x' over the observation interval $[t_0, T)$ for general object dynamics in the coordinates.

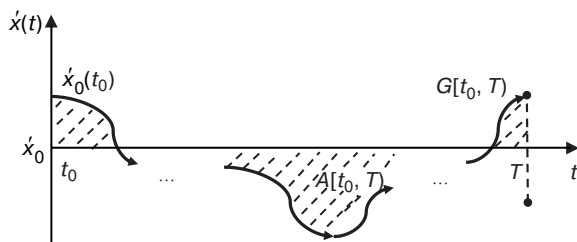


Fig. 2. The graph of the dynamics of the first derivative $G[t_0, T)$ within the observation interval $[t_0, T)$, where $x'(t_0)$ is the initial condition and $A[t_0, T)$ is the shaded flat figure swept by the integral curve $G[t_0, T)$

It is necessary to study how the numerical values of x' change over the observation interval $[t_0, T)$.

ALGORITHM Z FOR FORMING A CHAIN MODEL OF THE DYNAMICS OF THE FIRST DERIVATIVE AT TIME T

The graph of x' , or integral curve $G[t_0, T)$, in Fig. 2 consists of segments that represent the acceleration, deceleration, constant velocity, reverse deceleration, and reverse acceleration dynamics of x' . The physical significance of this behavior of the integral curve can be understood through the analysis of derivatives, with which engineers are familiar from their courses in mathematical analysis and advanced mathematics at technical universities.

1. The coordinates and duration of all local extrema of the integral curve $G[t_0, T)$ must be determined. Optimization techniques can be used for this purpose⁹. The multi-extremality of the dynamics

of the x' trajectory on the graph $G[t_0, T)$ should be explored using the tangent method outlined in the paper. Based on the outcomes of step 1 of the algorithm, a finite set of values for local extrema of the first derivative, as well as for $x'(t_0)$ and $x'(t)$, is generated. This set is arranged in ascending order of as follows:

$$\text{extr } x'[t_0, t) = \{(x'_j, t_j, t_j) : \ddot{x}(t_j) = 0\}_{j \in J(T)}.$$

Note. If the extremum x'_j is reached within the interval $[t_j, t_j + T_{pj}]$, which corresponds to $\ddot{x}(t_j) = 0$ for $\forall \tau \in [t_j, t_j + T_{pj}]$, then (x'_j, t_j, t_j) should be substituted by $(x'_j, t_j, t_j + T_{pj})$ in $\text{extr } x'[t_0, t)$

2. For $\forall j \in J(T)$, draw the j th secant $\text{Sec } x'_j$ from point x'_j . This is a straight line numbered j , which originates from the ordinate x'_j parallel to the horizontal axis until it intersects the vertical perpendicular restored from point T .
3. Fix the flat figure $A[t_0, T)$ swept by the integral curve $G[t_0, T)$ and set $j = \min\{j \in J(T)\}$.
4. Define a typical flat section $A[t_j, t_{j+1}) \subseteq A[t_0, T)$ by the following boundaries: part of the secant $\text{Sec } x'_j [t_j, t_{j+1}) \subseteq A_j[t_j, t_{j+1})$, while the fragment of the envelope (j th link), $G_j[t_j, t_{j+1}) \subseteq G[t_0, t)$ is at the opposite.

If $x'_j \leq x'_{j+1}$, then the section on the left is bounded by the vertical line restored from the point $g_j(x'_j, t_j) = G_j[t_j, t_{j+1}) \cap \text{Sec } x'_j \neq \Phi$. On the right, it approaches without a limit the vertical line restored from the point $G_j[t_j, t_{j+1}) \cap \text{Sec } x'_{j+1} \neq \Phi$, i.e., $g_j(x'_j < t_{j+1}) = G_j[t_j, t_{j+1}) \cap \text{Sec } x'_{j+1} \neq \Phi$.

If $x'_j > x'_{j+1}$, then the section on the left is bounded by the vertical line restored from the point $g_j(x'_j, t_j) = G_j[t_j, t_{j+1}) \cap \text{Sec } x'_{j+1} \neq \Phi$. On the right, it approaches without a limit the vertical line restored from the point $G_j[t_j, t_{j+1}) \cap \text{Sec } x'_j \neq \Phi$, i.e., $g_j(x'_j < t_{j+1}) = G_j[t_j, t_{j+1}) \cap \text{Sec } x'_j \neq \Phi$.

5. $j = j + 1$. If $j \in J(T)$, proceed to step 4.
6. Due to the above constructions and sequential enumeration of all j , the integral curve $G[t_0, T)$ can be represented as the following union (chain) of links:

$$G[t_0, T) = \bigcup_{j \in J(T)} G_j[t_j, t_{j+1}). \quad (8)$$

Stop. The chain model of the dynamics of the first derivative (8) is constructed, thus completing algorithm Z.

Note. When the dynamic behavior of the first derivative persists after time point T , the aforementioned algorithm Z is reiterated within a new observation interval and the chain model (8) is further refined.

Developing a method for evaluating the probabilities of the chain (8) ordinates, i.e., the first derivatives thereof, is of both scientific and practical significance.

⁷ Petrovskii I.G. *Lectures on the theory of ordinary differential equations*. Myshkis A.D., Oleinik O.A. (Eds.). Moscow: MSU; 1984, 296 p. (In Russ.).

⁸ Trukhan A.A., Ogorodnikova T.V. *Ordinary differential equations and methods of their solution*. Rows. *Elements of variational calculus*: Study guide for universities. 2nd ed. St. Petersburg: Lan; 2020, 268 p. ISBN 978-5-8114-6421-0. (In Russ.).

⁹ *Collection of math tasks for higher technical educational institutions*. In 4 parts. Part 3: *Study guide for higher technical educational institutions*. Efimov A.V., Pospelov A.S. (Eds.). 4th ed. rev. and add. Moscow: Physico-Mathematical Literature Publ.; 2002, 576 p. ISBN 5-94052-036-7. (In Russ.).

METHOD FOR GENERATING A COMPLETE SET OF VALUES FOR THE FIRST DERIVATIVE WITHIN AN OBSERVATION INTERVAL

Select the j th link from the chain (8):

$$G_j[t_j, t_{j+1}) \subseteq G[t_0, T], j \in J(T).$$

Choose an arbitrary point $g_{j_r}(x'_{j_r}, t_{j_r})$ in this link:

$$g_{j_r}(x'_{j_r}, t_{j_r}) \in G_j[t_j, t_{j+1}) \subseteq G[t_0, T], t_{j_r} \in [t_j, t_{j+1}), r = 1, 2, \dots \quad (9)$$

Note. The double index j_r in the notation for this point and its coordinates allows for other points in the same j th link and chain (8) to be denoted with the same ordinate value by varying the subscript. This will facilitate an understanding of the analyses process.

As defined by academician V.S. Pugachev [1], an arbitrary variable is a quantitative characteristic of an experiment that can take on one of several possible, albeit unpredictable values.

Denote the half-interval on the vertical ordinate axis in Fig. 3 as $OG_j[t_j, t_{j+1}) \subseteq OG[t_0, T]$. This includes the x'_{j_r} ordinate values of all points $g_{j_r}(x'_{j_r}, t_{j_r})$ (9) in the j th link of the chain (8), i.e.,

$$OG_j[t_j, t_{j+1}) = \left\{ \forall x'_{j_r} \in OG_j[t_j, t_{j+1}) : g_{j_r}(x'_{j_r}, t_{j_r}) \in G_j[t_j, t_{j+1}), r = 1, 2, \dots \right\}. \quad (10)$$

Fixing the value of an arbitrary point $g_{j_r}(x'_{j_r}, t_{j_r})$ in the j th link of the chain (8) in a graph is a result obtained through the application of either mathematical formulas or field measurements during observation. This is described as a stochastic event. The events involving the fixation of different points in the link are mutually exclusive in nature.

The space of possible outcomes is referred to as an arbitrary space, whose elements are called elementary outcomes. In practical terms, this space represents the set of all potential mutually exclusive experimental results.

Geometrical probability extends the conventional definition of probability by including cases having an infinite number of equiprobable outcomes.¹⁰

The probability of an event (j th link) occurring is determined using the geometrical concept of probability:

$$p_j = \frac{\text{len } G_j[t_j, t_{j+1})}{\text{len } G[t_0, T]}, \quad (11)$$

where $\text{len } G_j[t_j, t_{j+1})$ and $\text{len } G[t_0, T]$ are the lengths of the j th link and the entire chain, respectively. These values can be determined either from experimental monitoring data on the dynamics of the first derivative of the object, or by using formulas for calculating the length of a curve in a plane, which are taught in advanced mathematics courses at technical universities, for example [14]:

$$\begin{aligned} \text{len } G_j[t_j, t_{j+1}) &= \int_{t_j}^{t_{j+1}} \sqrt{1 + \ddot{x}(s)^2} ds, \\ \text{len } G[t_0, T] &= \int_{t_0}^T \sqrt{1 + \ddot{x}(s)^2} ds. \end{aligned}$$

The numerical value of the x'_{j_r} ordinate of point $g_{j_r}(x'_{j_r}, t_{j_r})$ (7) is an arbitrary variable.

Due to the construction of the graph of the first derivative, a bijection $\forall x'_{j_r} \leftrightarrow g_{j_r}(x'_{j_r}, t_{j_r})$ exists in each j th link of the chain (8).

As the presence of $\forall x'_{j_r} \in OG_j[t_j, t_{j+1})$ depends on the occurrence of the event $G_j[t_j, t_{j+1})$, and since there is no prior information about the preceding or subsequent derivative values, the occurrence of x'_{j_r} within the half-interval $OG_j[t_j, t_{j+1})$ is equally probable. This corresponds to the maximum uncertainty in the occurrence of this x'_{j_r} value, in terms of maximum entropy [15].

¹⁰ Melnikova I.N., Fastovets N.O. *Probability theory: Lecture notes*. Moscow: Gubkin University Publ.; 2017, 100 p. (In Russ.).

Figure 3 shows that due to the construction, each $OG_j[t_j, t_{j+1})$ has minimum and maximum values, $x'_{j-\min}$ and $x'_{j-\max}$, from which the following can be deduced:

$$OG_j[t_j, t_{j+1}) = [x'_{j-\min}, x'_{j-\max}) \text{ and} \\ \text{len } G_j[t_j, t_{j+1}) = x'_{j-\max} - x'_{j-\min}.$$

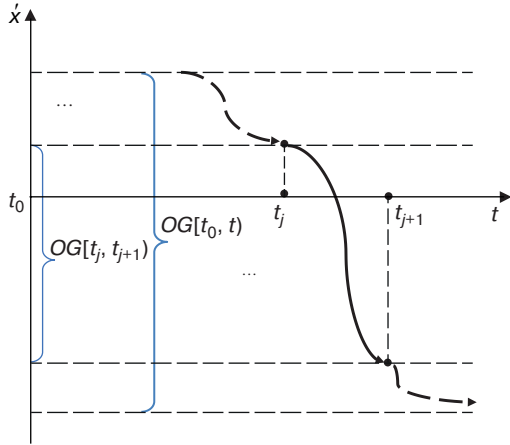


Fig. 3. Half-intervals of the ordinates:
 $OG[t_0, t]$ represents the entire chain, and $OG_j[t_j, t_{j+1})$
represents the j th link

Then, the probability of the x'_{j_r} ordinate value appearing within the half-interval $OG_j[t_j, t_{j+1})$ in the monitoring experiment is determined using (11), as follows:

$$px'_{j_r} = \frac{1}{x'_{j-\max} - x'_{j-\min}} p_j. \quad (12)$$

It follows from the above that the ordinate values of the points x' are stochastic.

Each arbitrary variable can be linked [1] to a specific event. For example, the occurrence of an event where the arbitrary variable x'_j takes on a value less than a specified value. Conversely, each event x'_j can be linked to an arbitrary variable, $P_j(x'_j)$. This represents the probability of a specific numerical value of the x'_{j_r} coordinate appearing on the vertical axis in a specific order in r different links of the chain (8).

Note: Due to the design depicted in Fig. 3, the different links in the chain (8) may correspond to different horizontal axis intervals, while having the same values on the vertical axis, as shown in Fig. 4.

If the following condition is met:

$$OG_j[t_j, t_{j+1}) \cap OG_p[t_p, t_{p+1}) = \\ = [x'_{j-\min}, x'_{j-\max}) \cap [x'_{p-\min}, x'_{p-\max}) = \quad (13) \\ = [x'_{j-\min}, x'_{j-\max}) \neq \Phi, j < p, j, \forall p \in J(T),$$

then all links $\forall p \in J(T)$ can be considered equivalent to the j th link.

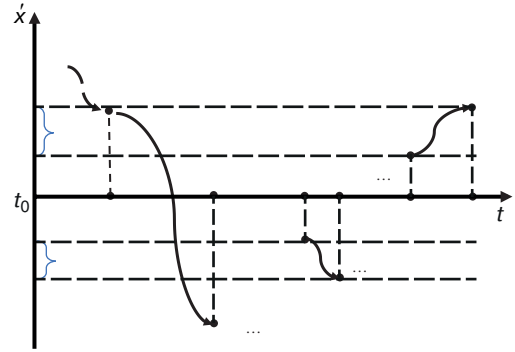


Fig. 4. Similar chain links. The solid line depicts three curved segments that are distinct yet similar chain links, which share common ordinates marked with vertical curly brackets

We form a supplementary set of link numbers, similar to the j th link in the chain (8)

$$Ax'_j = \{j, r_{j1}, r_{j2}, \dots, r_{jk}\}, \quad (14)$$

where r_{ji} is the number of links in chain (8) that are similar to the j th link.

Note that $\text{pow } Ax'_j$, the set cardinality or number of elements in set Ax'_j (14), is equal to the number of links in chain (8) that are similar to the j th link plus one.

We can then write the following formula:

$$P_j(x'_j) = \sum_{r \in Ax'_j} \frac{1}{x'_{j_r-\max} - x'_{j_r-\min}} p_r, \quad (15)$$

which coincides with formula (12) in the special case where there are no links similar to the j th link, with $x'_{j_r-\max} = x'_{j-\max}$, $x'_{j_r-\min} = x'_{j-\min}$.

By exhaustively enumerating all indices $j \in J(T)$ and applying the formulas (9)–(15) presented in this section, a complete set of events (i.e., the values of x'_j) can be derived: $\{P_j(x'_j)\}_{j \in J(T)}$, $\sum_{j \in J(T)} P_j(x'_j) = 1$ at time T within the observation interval.

Note. Based on the results of processing the experimental observation data, as t increases, the values of this complete set of events may change due to variations in $\text{len } G[t_0, t]$ (11) and $J(T)$ (15).

The existence of a complete set of such events is a defining feature of the first derivative.

It would be beneficial to investigate the impact of links similar to those in the chain model (8).

ALGORITHM W FOR GENERATING A CHAIN MODEL OF THE FIRST DERIVATIVE REFINED BY SIMILAR LINKS

1. Use algorithm Z to generate the chain model (8) of the first derivative at time t .

2. Using the method described above, determine the complete set of values of the first derivative within the observation interval for the chain model (8).
3. For $\forall j \in J(T)$, create a finite set, Bx' , of values for index j , arranged in ascending order. This set is formed from the values $x'_{j,-\max}^{\text{mul}}, x'_{j,-\max}, x'_{j,-\min}^{\text{mul}}, x'_{j,-\min}$, obtained in step 2, some of which may coincide. The elements of the set Bx' at time t are denoted as $b_1, b_2, \dots, b_{n(T)}$.
4. For $\forall i = 1, 2, \dots, n(T)$, draw the i th secant, $\text{Sec } b_i$, which is a straight line numbered i , originating from coordinate b_i on the vertical axis and running parallel to the horizontal axis until it intersects the vertical perpendicular restored from point T . Hold fix $i = 1$.
5. Define the new flat segment $B_i[t_i, t_{i+1}) \subseteq A[t_0, T)$ using the following boundaries. This is bounded below by the segment $\text{Sec } b_i \subset B_i[t_i, t_{i+1})$ and above by the new i th link $GNW_i[t_i, t_{i+1}) \subseteq G[t_0, T)$, which is a segment of the envelope.

If $b_i < b_{i+1}$, then the interval is bounded on the left by the vertical line reconstructed from the point $h_j(b_i, t_i) = GNEW_i[t_i, t_{i+1}) \cap \text{Sec } b_i \neq \Phi$. On the right, it approaches the vertical line reconstructed from the point $GNW_i[t_i, t_{i+1}) \cap \text{Sec } b_{i+1} \neq \Phi$, without bound. That is

$$h_i(b_i < t_i) = GNEW_i[t_i, t_{i+1}) \cap \text{Sec } b_{i+1} \neq \Phi.$$

If $b_i > b_{i+1}$, then the interval is bounded on the left by the vertical line reconstructed from the point $h_i(b_i, t_i) = GNEW_i[t_i, t_{i+1}) \cap \text{Sec } b_{i+1} \neq \Phi$, and on the right, it approaches the vertical line reconstructed from the point $GNW_i[t_i, t_{i+1}) \cap \text{Sec } b_i \neq \Phi$, without bound. That is

$$h_i(b_i < t_{i+1}) = GNEW_i[t_i, t_{i+1}) \cap \text{Sec } b_i \neq \Phi.$$

6. $i = i + 1$. If $i \leq n(T)$, proceed to step 5.
7. Based on the above constructions, and by sequentially iterating over all possible values of i , the envelope curve $G[t_0, T)$ can be represented as a new union (chain) that takes into account the characteristics of the similar segments, in the following way:

$$G[t_0, t) = \bigcup_{i=1, 2, \dots, n(T)} GNEW_i[t_i, t_{i+1}). \quad (16)$$

Stop. The end of algorithm W. A new chain model for the first derivative dynamics (16) has been developed and refined based on similar links.

These models and algorithms address the “heavy-tailed” problem in the probabilistic analysis of internet processes and modern information and communication systems. They provide new formalizations for problems related to research in areas such as Big Data analysis, data mining [16, 17], and artificial intellect [18].

Further investigation of the new chain model (16) for the dynamics of the first derivative, taking into

account its non-uniformity, is required. This involves considering the presence of links with sections of deceleration and acceleration in the dynamics of the derivative throughout the entire observation interval up to time T .

PROBABILITY CHARACTERISTICS OF TYPICAL LINKS

For circuit (16), we consider the typical links shown in Figs. 5–8.

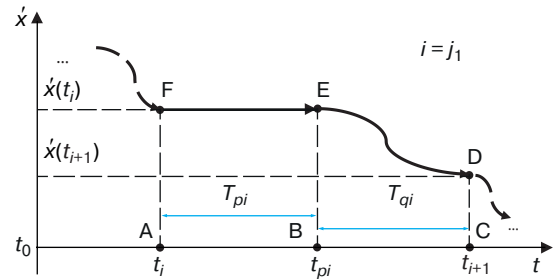


Fig. 5. Graph of the constancy and acceleration of the first derivative in the link $G_i[t_i, t_{i+1})$, $i = j_1$. $AB = T_{pi}$ and $BC = T_{qi}$ are the corresponding coordinate segments on the x -axis, representing the constancy and deceleration of the first derivative from $x'(t_i)$ to $x'(t_{i+1})$

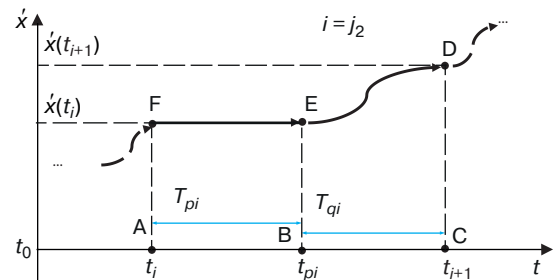


Fig. 6. Graph of the constancy and acceleration of the first derivative in the link $G_i[t_i, t_{i+1})$, $i = j_2$. $AB = T_{pi}$ and $BC = T_{qi}$ are the corresponding coordinate segments on the x -axis, representing the constancy and acceleration of the first derivative from $x'(t_i)$ to $x'(t_{i+1})$

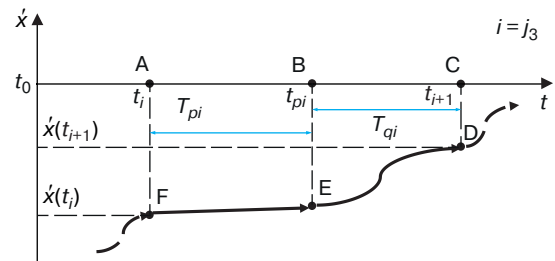


Fig. 7. Graph of the inverse constancy and deceleration of the first derivative in the link $G_i[t_i, t_{i+1})$, $i = j_3$. $AB = T_{pi}$ and $BC = T_{qi}$ are the corresponding coordinate segments on the x -axis, representing the constancy and deceleration of the first derivative from $x'(t_i)$ to $x'(t_{i+1})$

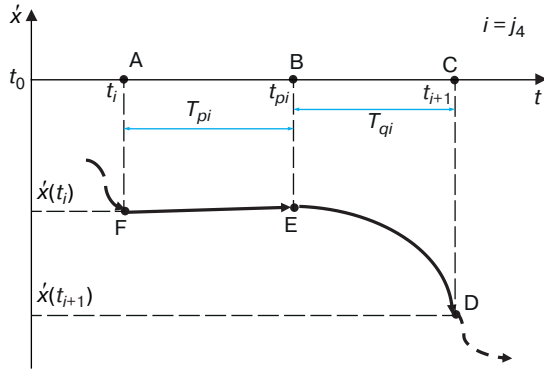


Fig. 8. Graph of the inverse constancy and acceleration of the first derivative in the link $G_i[t_i, t_{i+1}]$, $i = j_4$. $AB = T_{pi}$ and $DC = T_{qi}$ are the corresponding coordinate segments on the x -axis, representing the constancy and acceleration of the first derivative from $x'(t_i)$ to $x'(t_{i+1})$

In terms of probability theory, we examine the characteristics of typical links with irregularities in the dynamics of the first derivative (Figs. 5–8).

Analysis of these figures shows that two random events are possible: $x' = x'_{\text{const}_i}$ and $x' \neq x'_{\text{const}_i}$. The methodological aspects of calculating the probabilities of these events are of scientific and practical interest as they enable an evaluation of the stochastic moments of changes in the first derivative, allowing predictive conclusions to be drawn about the dynamics of these changes for purposes such as the control and monitoring of a moving objects.

Figures 5–8 show a plane figure ABCDEF, which illustrates the i th typical link of change in the dynamics of the first derivative, $i \in \{j_1, j_2, j_3, j_4\} \subset I$. Initially constant, and then becoming non-uniform from time t_{pi} (with deceleration in Fig. 5 and acceleration in Fig. 6, as well as similar changes in the reverse dynamics in Figs. 7 and 8), the change in the dynamics of the first derivative in the i th typical link occurs within the interval $[t_i, t_{i+1}) \subseteq [t_0, T)$.

The probability of the event where x' takes the x'_{const_i} value in the i th typical link within the interval $[t_i, t_{i+1})$ can be estimated using the concept of geometric probability [9, 10, 19, 20], as follows:

$$P(x' = x'_{\text{const}_i}) = \frac{\text{len } s_{1_i}[t_i, t_{pi}]}{\text{len } G[t_0, T]}, \quad (17)$$

$$P(x' \neq x'_{\text{const}_i}) = \frac{\text{len } s_{2_i}}{\text{len } G[t_0, T]} = q_i, \quad (18)$$

where

$$\text{len } s_{1_i}[t_i, t_{pi}] = \left| \int_{t_i}^{t_{pi}} \sqrt{1 + [\ddot{x}(s)|_{\ddot{x}(s)=0}]^2} ds \right| = t_{pi} - t_i = T_{pi}$$

is the length of the i th typical link with $x' = x'_{\text{const}_i}$,

$$\text{len } G[t_0, T] = \int_{t_0}^T \sqrt{1 + \ddot{x}(s)^2} ds \text{ is the length of the entire}$$

chain (8), $\text{len } s_{2_i}[t_{qi}, t_{i+1}] = \left| \int_{t_{qi}}^{t_{i+1}} \sqrt{1 + \ddot{x}(s)^2} ds \right|$ is the

length of the part of the i th typical link with $x' \neq x'_{\text{const}_i}$.

According to the mean value theorem for definite integrals [21], it follows that

$$\text{len } s_{2_i} = \left| \int_{t_{pi}}^{t_{i+1}} \ddot{x}(s) ds \right| = |\alpha_{\text{aver}}[t_{pi}, t_{i+1}]| T_{qi} \geq 0. \quad (19)$$

It follows from this at $\text{len } s_{2_i} \neq 0$ that

$$\left. \begin{aligned} \alpha_{\text{aver}}[t_{pi}, t_{i+1}] &\neq 0, \\ T_{qi} &\geq 0. \end{aligned} \right\} \quad (20)$$

Given the values of t_i , t_{pi} , and t_{i+1} obtained from measurements, it follows from the latter relations that the value of $\alpha_{\text{aver}}[t_{pi}, t_{i+1}]$ can be obtained by numerical calculation.

Note: The calculated values of $\alpha_{\text{aver}}[t_{pi}, t_{i+1}]$ may be negative due to the construction of the corresponding graphs for the inverse dynamics.

On the other hand, according to the mean value theorem for definite integrals [21], the following can be written:

$$\begin{aligned} \text{len } s_{1_i} + \text{len } s_{2_i} &= \left| \int_{t_i}^{t_{i+1}} \ddot{x}(s) ds \right| = \\ &= |\alpha_{\text{aver}}[t_i, t_{i+1}]| (t_{i+1} - t_i) \geq 0. \end{aligned} \quad (21)$$

From relations (19)–(21), we obtain

$$\begin{aligned} |\alpha_{\text{aver}}[t_{pi}, t_{i+1}]| T_{qi} &= \\ &= |\alpha_{\text{aver}}[t_i, t_{i+1}]| (t_{i+1} - t_i) - T_{pi} \geq 0. \end{aligned} \quad (22)$$

It follows from (22) that $|\alpha_{\text{aver}}[t_i, t_{i+1}]| (t_{i+1} - t_i) \geq T_{pi}$, while $|\alpha_{\text{aver}}[t_{pi}, t_{i+1}]| T_{qi} \leq |\alpha_{\text{aver}}[t_i, t_{i+1}]| (t_{i+1} - t_i)$ or $\frac{|\alpha_{\text{aver}}[t_{pi}, t_{i+1}]|}{|\alpha_{\text{aver}}[t_i, t_{i+1}]|} \leq \frac{t_{i+1} - t_i}{T_{qi}} = 1 + \frac{T_{pi}}{T_{qi}}$.

Applying (17), we obtain the following:

$$P(x' = x'_{\text{const}_i}) = p_i = \frac{T_{pi}}{\text{len } G[t_0, T]} \leq 1. \quad (23)$$

For the hypothetical case of $T_{qi} = 0$, the probability density function is

$$f(x' = x'_{\text{const}_i}) = \delta(x' - x'_{\text{const}_i}), \quad (24)$$

where δ is the Dirac delta function and coincides with the result obtained in [10].

The obtained relations require studying the distribution of the values of the first derivative in the i th typical link of the refined circuit model (16) within the interval $[t_i, t_{i+1}) \subseteq [t_0, T)$.

ANALYSIS OF THE DISPERSION OF THE FIRST DERIVATIVE DYNAMICS ON TYPICAL SECTIONS

Suppose that the chain model (16) consists of a single i th typical link. The shape of the probability density of the first derivative, $\mathfrak{D}(x', t \in [t_i, t_{i+1}])$, which yields the supremum¹¹ of the variance of the values, $D(x', t \in [t_i, t_{i+1}])$, in cases of deceleration and acceleration of the first derivative in the i th typical link (Figs. 5–8), is determined by Lemma 1.

Lemma 1. The supremum of the variance of the first derivative, $D(x', t \in [t_i, t_{i+1}])$ in the i th typical link, is attained at the probability density $\mathfrak{D}(x', t \in [t_i, t_{i+1}]) = p_i \delta(x' - x'_{\text{const}_i}) + q_i \delta(x' - x'_{\text{const}_{i+1}})$, which corresponds to the discrete type of the probability distribution function of the arbitrary variable x' .

Proof of Lemma 1. We analyze the variance $D(x', t \in [t_i, t_{i+1}])$ and the shape of the probability density $\mathfrak{D}(x', t \in [t_i, t_{i+1}])$ in Figs. 5–8. Let

$$x'_{\min_i} = \min(x'_{\text{const}_i}, x'_{\text{const}_{i+1}}), \quad x'_{\max_i} = \max(x'_{\text{const}_i}, x'_{\text{const}_{i+1}}).$$

$$\begin{aligned} D(x', t \in [t_i, t_{i+1}]) &= \\ &= \int_{-\infty}^{\infty} (s^2) f(s, t \in [t_i, t_{i+1}]) ds - \\ &- \left(\int_{-\infty}^{\infty} s f(s, t \in [t_i, t_{i+1}]) ds \right)^2 = \\ &= \left| \int_{x'_{\min_i}}^{x'_{\max_i}} (s^2) f(s, t \in [t_i, t_{i+1}]) ds \right| - \\ &- \left(\int_{x'_{\min_i}}^{x'_{\max_i}} s f(s, t \in [t_i, t_{i+1}]) ds \right)^2. \end{aligned}$$

Following the recommendations given in [10, 22], we perform an ε -partition of the segment $[x'_{\min_i}, x'_{\max_i}]$ of length $x'_{\max_i} - x'_{\min_i} > 0$, dividing it into non-intersecting half-intervals $0 < [x'_j, x'_{j+1}] < \varepsilon$, $j \in J(\varepsilon)$ of length $x'_{j+1} - x'_j$, each less than $\varepsilon > 0$.

We approximate $D(x', t \in [t_i, t_{i+1}])$ by the sum

$$\tilde{D}(x', t \in [t_i, t_{i+1}]) = \sum_{j \in J(\varepsilon)} z_j (e_j)^2 - \left(\sum_{j \in J(\varepsilon)} z_j e_j \right)^2,$$

where $z_j = \int_{x'_j}^{x'_{j+1}} f(s, t \in [t_i, t_{i+1}]) ds \geq 0$, e_j is defined

below,

$$\sum_{j \in J(\varepsilon)} z_j = 1. \quad (25)$$

The Lagrange multiplier method for $\tilde{D}(x', t \in [t_i, t_{i+1}]) \rightarrow \text{extr}$ subject to the constraints (25) yields the necessary condition for the optimum:

$$\begin{cases} e_j^2 + 2me_j - \lambda = 0, \quad j \in J(\varepsilon), \\ \sum_{j \in J(\varepsilon)} z_j = 1, \\ z_j \geq 0, \end{cases} \quad (26)$$

where $m = \sum_k e_k z_k$, $k \in J(\varepsilon)$, $k \neq j$; λ is the Lagrange multiplier.

Thus, for $\forall j \in J(\varepsilon)$, the value of e_j must be a solution to the same quadratic equation in (26). Consequently, for $\forall j \in J(\varepsilon)$, the values of e can take no more than two distinct values. Taking into account (17) and (18), which are characteristic of the half-interval $[t_i, t_{i+1}]$, these values are p_i (23) and q_i (18).

The construction of the refined chain model (16) and (24) makes the case of multiple roots of the quadratic equation in (26) impossible.

Therefore, the general form of the probability density function for the values of the first derivative of the object's dynamics for Figs. 5–8 is as follows:

$$\mathfrak{D}(x', t \in [t_i, t_{i+1}]) = p_i \delta(x' - x'_{\text{const}_i}) + q_i \delta(x' - x'_{\text{const}_{i+1}}).$$

Omitting the intermediate calculations from [10, 23], the following expression for the variance $D(x', t \in [t_i, t_{i+1}])$ can be written:

$$D(x', t \in [t_i, t_{i+1}]) = p_i q_i (x'_{\text{const}_i} - x'_{\text{const}_{i+1}})^2. \quad (27)$$

Thus, Lemma 1 is proved.

Corollary of Lemma 1. It follows from (27) that the supremum of $D(x', t \in [t_i, t_{i+1}])$ is attained at $p_i q_i = \max$ and $|x'_{\text{const}_i} - x'_{\text{const}_{i+1}}| = \max$.

We analyze the variance of the values of $D(x', t \in [t_i, t_{i+1}])$.

Omitting the intermediate calculations from [6, 9, and 10], the following expressions can be derived for the variance of x' :

$$\begin{aligned} D(x', t \in [t_i, t_{i+1}]) &= \\ &= \int_{-\infty}^{\infty} (s^2) f(s, t \in [t_i, t_{i+1}]) ds - \\ &- \left(\int_{-\infty}^{\infty} s f(s, t \in [t_i, t_{i+1}]) ds \right)^2 = \\ &= \left| \int_{x'_{\min_i}}^{x'_{\max_i}} (s^2) f(s, t \in [t_i, t_{i+1}]) ds \right| - \\ &- \left(\int_{x'_{\min_i}}^{x'_{\max_i}} s f(s, t \in [t_i, t_{i+1}]) ds \right)^2 = \\ &= p_i q_i (x'_{\text{const}_i} - x'_{\text{const}_{i+1}})^2, \end{aligned}$$

from which it follows that the supremum of $D(x', t \in [t_i, t_{i+1}])$ is attained when $x'_{\text{const}_i} = 0$ or $x'_{\text{const}_{i+1}} = 0$. This does not contradict the graphs in Figs. 5–8;

$$E(x', t \in [t_i, t_{i+1}]) =$$

$$= \int_{-\infty}^{\infty} s f(s, t \in [t_i, t_{i+1}]) ds = p_i x'_{\text{const}_i} + q_i x'_{\text{const}_{i+1}},$$

where p_i is taken from (23) and q_i from (18).

Of scientific interest is the study of the behavior of the dispersion dynamics of the random values of x' over the entire observation interval $[t_0, T)$.

¹¹ Least upper bound.

ANALYSIS OF THE DISPERSION OF THE FIRST DERIVATIVE DYNAMICS OVER THE ENTIRE OBSERVATION INTERVAL

As discussed previously, at time T , the full set of integral curve events of the x' dynamics is an inherent property of any first-order differential equation.

As in [8, 9, 23], we analyze the variance of the values $D(x', t \in [t_i, t_{i+1}))$ and the shape of the probability density of the derivative $f(x', t \in [t_i, t_{i+1}))$, which yields the supremum of $D(x', t \in [t_i, t_{i+1}))$. This problem is solved by Lemma 2.

Lemma 2. The supremum of the variance of the values of the first derivative, $D(x', t \in [t_i, t_{i+1}))$, over the entire observation area $[t_0, T]$, is reached based on the probability distribution density $\mathfrak{I}(x', t \in [t_i, t_{i+1})) = \sum_{j \in J(t)} (p_i \delta(x' - x'_{\text{const}_i}) + q_i \delta(x' - x'_{\text{const}_{i+1}}))$, which corresponds to a discrete probability distribution function of an arbitrary variable, x' .

Proof of Lemma 2. We analyze the variance value $D(x', t \in [t_i, t_{i+1}))$ and the probability distribution density type $\mathfrak{I}(x', t \in [t_0, T])$ for Fig. 2.

For the chain model (14), taking into account (27), we can write the following:

$$D(x', t \in [t_0, T]) = \sum_{i \in J(T)} D(x', t \in [t_i, t_{i+1})) = \\ = \sum_{i \in J(T)} p_i q_i (x'_{\text{const}_i} - x'_{\text{const}_{i+1}})^2.$$

The apparent feature is: the maximum of the sum of independent terms is the sum of their maxima.

Thus, Lemma 2 is proved.

The result of the stepwise approximation of the stochastic process of the first derivative at time T , which has broad applicability, is expressed as follows:

$$x'(t \in [t_0, T]) = \\ = \sum_{i \in J(T)} [x'_{\text{const}_i} p_i \Theta(t_{pi} - t_i) + x'_{\text{const}_{i+1}} q_i \Theta(t_{i+1} - t_i)], \quad (28)$$

where $\Theta(t) = \begin{cases} 0, & \text{when } t < 0, \\ 1, & \text{when } t \geq 0 \end{cases}$ is the Heaviside step

function.

Analysis of this model shows that it can be constructed using measured data on monitoring the object's dynamics, thus eliminating the need to solve the corresponding differential equation. The graph of this equation can be reconstructed from the parameters in (28).

Therefore, the step function model of first derivative dynamics can be used to study of dynamic systems based on their observed probabilistic characteristics.

CONCLUSIONS

The paper establishes and explores the stochastic nature of forming the integral curve of solutions to Cauchy's problem for first-order ordinary differential

equations that describe the dynamics of arbitrary objects. The integral curve can be obtained through calculation using analytical relationships or by monitoring the object's dynamics and acquiring and processing measurements of its characteristics.

As the actual dynamics of the object are highly heterogeneous, as demonstrated by the wide range of values of its first derivative, a chain model of the dynamics of the first derivative at a fixed observation time was developed, along with an associated algorithm Z for its formation. These are intended to provide primary analytical, numerical, and graphical interpretations of changes in the dynamics of the first derivative, suitable for subsequent stochastic analysis.

A geometric probabilistic approach has been used to discover new probabilistic properties of solutions to Cauchy's problems as stochastic processes of changes in the first derivative dynamics.

The concept of similar links in the chain model of the first derivative has been introduced, enabling the development of a more detailed model of the integral curve chain in the form of a chain of similar enveloping links. The algorithm W has also been devised for constructing a chain model of the first derivative that is refined by similar links.

It is demonstrated that the probability values of the first derivative constitute a complete set of disjoint events. For estimating these probabilities at any observation time, corresponding formulas are derived. The number of elements in such a complete group is a non-decreasing integer function of the abscissa axis (the horizontal coordinate). The existence of a complete group of events of this kind is a defining feature of the first derivative. Furthermore, it is demonstrated that this complete group of events is an inherent property of any first-order differential equation.

A method is proposed for estimating the probability of the values of the first derivative within an observation interval based on the concept of geometric probability and by using the ratio of link lengths in a refined chain model.

These models and algorithms address the "heavy-tailed" issue in probabilistic investigations of processes in contemporary information and communication systems and technologies.

Lemmas have been proved to justify the discrete form of the probability distribution of the values of the first derivative of an object's dynamics. Studies have also been conducted on the dispersion of this distribution, which justify its supremum. The derived formulas enable the dispersion of first derivative dynamics to be estimated based on experimental data, allowing a piecewise approximation of the random process for the dynamics of the first derivative to be formed over a fixed observation period.

The stepwise model of first-order dynamics is generally applicable, making it possible to study dynamic systems based on the observed probabilistic characteristics of its parameters. This allows the construction of a set of criteria and formulate dynamics-based problems in terms of stochastic processes, thus opening up new avenues for research.

In order to identify new dependencies in the dynamics of the first derivative, the methodological results obtained require further development in line with approaches to random processes. These results provide a foundation for novel formulations and statements of research problems in various subject areas, particularly those related to observations and monitoring in fields such as Big Data, data mining, and artificial intellect.

REFERENCES

1. Pugachev V.S. *Teoriya sluchainykh funktsii i ee primeneniye k zadacham avtomaticheskogo upravleniya (Theory of Random Functions and its Application to Automatic Control Problems)*. Moscow: Gostekhizdat; 1957, 659 p. (In Russ.).
2. Ponomarev K.K. *Sostavlenie differentsial'nykh uravnenii (Compilation of Differential Equations)*. Minsk: Vyshehishaya Shkola; 1973, 560 p. (In Russ.).
3. Bulychev Yu.G. Method of integral support curves for solving the Cauchy problem for ordinary differential equations. *U.S.S.R. Comput. Math. Math. Phys.* 1988;28(5):130–136. [https://doi.org/10.1016/0041-5553\(88\)90021-3](https://doi.org/10.1016/0041-5553(88)90021-3)
[Original Russian Text: Bulychev Yu.G. Method of integral support curves for solving the Cauchy problem for ordinary differential equations. *Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki*. 1988;28(10):1482–1490 (in Russ.).]
4. Bulychev Yu.G. Methods for the numerical-analytical integration of differential equations. *U.S.S.R. Comput. Math. Math. Phys.* 1991;31(9):28–39.
[Original Russian Text: Bulychev Yu.G. Methods for the numerical-analytical integration of differential equations. *Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki*. 1991;31(9):1305–1319 (in Russ.).]
5. Bulychev Yu.G., Burlai I.V. System theoretic approach to the modeling of complex dynamic systems in optimization problems with predictive model. *Autom. Remote Control*. 1996;57(3):329–338.
[Original Russian Text: Bulychev Yu.G., Burlai I.V. System theoretic approach to the modeling of complex dynamic systems in optimization problems with predictive model. *Avtomatika i telemekhanika*. 1996;3:34–46 (in Russ.). <https://www.elibrary.ru/movvbz>]
6. Bulychev Yu.G., Burlai I.V. A parametric identification method of control systems under inexact input data. *Avtomatika i telemekhanika*. 1997;11:56–65 (in Russ.).
7. Burlai I.V. The parametric identification of controlled systems based on the extended model of observation. *J. Comp. Syst. Sci. Inform.* 1977;36(4):524–529. <https://www.elibrary.ru/ledoyx>
[Original Russian Text: Burlai I.V. The parametric identification of controlled systems based on the extended model of observation. *Izvestiya Rossiiskoi Akademii nauk. Teoriya i sistemy upravleniya*. 1997;4:29–34 (in Russ.). <https://www.elibrary.ru/mozlgl>]
8. Bulychev Yu.G., Burlai I.V. Optimal estimation of normal regression parameters for the case of an extended observational model. *Problems Inform. Transmission*. 1993;29(3):31–41 (in Russ.).
9. Nazarov A.N. *Modeli i metody rascheta strukturno-setevykh parametrov ATM-setei (Models and Methods for Calculating the Structural and Network Parameters of ATM Networks)*. Moscow: Goryachaya liniya-Telekom; 2002, 256 p. (In Russ.). ISBN 5-93517-087-6
10. Nazarov A.N. Service Traffic Models with Bit Rate of Transmission in Wideband Integral Service Digital Networks. *Autom. Remote Control*. 1998;59(9):1245–1254.
[Original Russian Text: Nazarov A.N. Service Traffic Models with Bit Rate of Transmission in Wideband Integral Service Digital Networks. *Avtomatika i telemekhanika*. 1998;9:52–63 (in Russ.).]
11. Nazarov A.N. Two-Stage Method for Assessing the Adequacy of Local Polynomial Spline Interpolation of Bit Traffic in Broadband Digital Networks with Integrated Services. *Intellektual'nye sistemy = Intelligent Systems*. 1999;4(1-2):169–184 (in Russ.). <https://www.elibrary.ru/gpunwq>
12. Nazarov A.N. A method of parametric stable identification of dynamic bit traffic in broadband digital networks of integral service. *Intellektual'nye sistemy = Intelligent Systems*. 2000;5(1-4):111–125 (in Russ.). <https://www.elibrary.ru/ywaxjz>
13. Nazarov A.N. Models of Bit Traffic in Digital Trains and Nodes of Broadband Digital Networks of Integrated Service Based on ATM Technology, Taking into Account the Features of Statistical Compaction. In: *Models and Methods of Researching Information Networks: Collection of Scientific Papers*. Saint Petersburg; 1999. V. 2. P. 105–116 (in Russ.).

14. Fikhtengol'ts G.M. *Osnovy matematicheskogo analiza (Fundamentals of Mathematical Analysis)*. V. 1. Moscow: Nauka; 1968, 440 p. (In Russ.).
15. Prokaev A.N. The Maximum Entropy Principle in Decision Theory. *Vestnik Sankt-Peterburgskogo universiteta. Prikladnaya matematika. Informatika. Protsessy upravleniya = Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*. 2024;20(2):154–169 (in Russ.). <https://doi.org/10.21638/spbu10.2024.203>
16. Zinchenko A.P., Ukolova A.V., Demichev V.V., et al. *Tsifrovyye tekhnologii analiza dannykh v sel'skom khozyaistve (Digital Technologies for Data Analysis in Agriculture)*. Moscow: Nauchnyi konsul'tant; 2024, 260 p. (in Russ.). ISBN 978-5-907477-96-4
17. Gorelik A. *Korporativnoe ozero bol'shikh dannykh. Novyi podkhod k ispol'zovaniyu Big Data i Data Science v biznese (Corporate Big Data Lake. A New Approach to Using Big Data and Data Science in Business)*. Moscow: Eksmo; 2023, 272 p. (In Russ.). ISBN 978-5-04-107657-3
18. Katsov I. *Iskusstvennyi intellekt na predpriyatii: teoriya i praktika (Artificial Intelligence at the Enterprise: Theory and Practice)*; transl. from Engl. DMK Press; 2024, 710 p. (In Russ.). ISBN 978-5-93700-277-8
19. Gnedenko B.V. *Kurs teorii veroyatnostei (Course in Probability Theory)*. Moscow: Fizmatgiz; 1961, 406 p. (In Russ.).
20. Gmurman V.E. *Teoriya veroyatnostei i matematicheskaya statistika (Probability Theory and Mathematical Statistics)*. Moscow: Yurait; 2023, 480 p. (In Russ.). ISBN 978-5-534-00859-3
21. Kudryavtsev L.D. *Matematicheskii analiz (Mathematical Analysis)*. V. 1. Moscow: Vysshaya Shkola; 1973, 614 p. (In Russ.).
22. Opoitsev V.I. The nonlinear law of large numbers. *Autom. Remote Control*. 1994;55(4):511–519.
[Original Russian Text: Opoitsev V.I. The nonlinear law of large numbers. *Avtomatika i telemekhanika*. 1994;4:65–75 (in Russ.).]
23. Cooper G.R., McGilhem C.D. *Veroyatnostnye metody analiza signalov i sistem (Probabilistic Methods of Signal and System Analysis)*; transl. from Engl. Moscow: Mir; 1989, 376 p. (In Russ.).
[Cooper G.R., McGilhem C.D. *Probabilistic Methods of Signal and System Analysis*. Holt, Rinehart, and Winston; 1986, 408 p.]

СПИСОК ЛИТЕРАТУРЫ

1. Пугачёв В.С. *Теория случайных функций и ее применение к задачам автоматического управления*. М.: Гостехиздат; 1957, 659 с.
2. Пономарев К.К. *Составление дифференциальных уравнений*. Минск: Вышэйшая школа; 1973, 560 с.
3. Булычев Ю.Г. Метод опорных интегральных кривых решения задачи Коши для обыкновенных дифференциальных уравнений. *Журн. вычисл. матем. и матем. физ.* 1988;28(10):1482–1490.
4. Булычев Ю.Г. Методы численно-аналитического интегрирования дифференциальных уравнений. *Журн. вычисл. матем. и матем. физ.* 1991;9:1305–1319.
5. Булычев Ю.Г., Бурлай И.В. Системный подход к моделированию сложных динамических систем в задачах оптимизации с прогнозирующей моделью. *Автоматика и телемеханика*. 1996;3:34–46. <https://www.elibrary.ru/movvzb>
6. Булычев Ю.Г., Бурлай И.В. Метод параметрической идентификации систем управления при неточном задании исходных данных. *Автоматика и телемеханика*. 1997;11:56–65.
7. Бурлай И.В. Параметрическая идентификация управляемых систем на базе расширенной модели наблюдений. *Известия Российской Академии наук. Теория и системы управления*. 1997;4:29–34. <https://www.elibrary.ru/mozlgl>
8. Булычев Ю.Г., Бурлай И.В. Оптимальное оценивание параметров нормальной регрессии для случая расширенной модели наблюдений. *Проблемы передачи информации*. 1993;29(3):31–41.
9. Назаров А.Н. *Модели и методы расчета структурно-сетевых параметров АТМ-сетей*. М.: Горячая линия – Телеком; 2002, 256 с. ISBN 5-93517-087-6
10. Назаров А.Н. Модели трафика служб с битовой скоростью передачи информации в широкополосных цифровых сетях интегрального обслуживания. *Автоматика и телемеханика*. 1998;9:52–63.
11. Назаров А.Н. Двухэтапный метод оценки адекватности локальной полиномиальной сплайн-интерполяции битового трафика служб широкополосных цифровых сетей интегрального обслуживания. *Интеллектуальные системы*. 1999;4(1-2):169–184. <https://www.elibrary.ru/grunwq>
12. Назаров А.Н. Метод параметрической устойчивой идентификации динамического битового трафика в широкополосных цифровых сетях интегрального обслуживания. *Интеллектуальные системы*. 2000;5(1-4):111–125. <https://www.elibrary.ru/ywaxjz>
13. Назаров А.Н. Модели битового трафика в цифровых трактах и узлах широкополосных цифровых сетей интегрального обслуживания на технологии АТМ с учетом особенностей статистического уплотнения. В сб.: *Модели и методы исследования информационных сетей: сборник научных трудов*. СПб.: Тема; 1999. Вып. 2. С. 105–116.
14. Фиктengольц Г.М. *Основы математического анализа*. Т. 1. М.: Наука; 1968, 440 с.
15. Прокаев А.Н. Принцип максимума энтропии в теории принятия решений. *Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления*. 2024;20(2):154–169. <https://doi.org/10.21638/spbu10.2024.203>
16. Зинченко А.П., Уколова А.В., Демичев В.В., Тихонова А.В., Романцева Ю.Н., Харитонов А.Е., Быков Д.В., Ульянкин А.Е., Невзоров А.С. *Цифровые технологии анализа данных в сельском хозяйстве: монография*. М.: Изд-во «Научный консультант»; 2024, 260 с. ISBN 978-5-907477-96-4

17. Горелик А. *Корпоративное озеро больших данных. Новый подход к использованию Big Data и Data Science в бизнесе*. М.: Эксмо; 2023, 272 с. ISBN 978-5-04-107657-3
18. Кацов И. *Искусственный интеллект на предприятии: теория и практика*: пер. с англ. ДМК Пресс; 2024, 710 с. ISBN 978-5-93700-277-8
19. Гнеденко Б.В. *Курс теории вероятностей*. М.: Физматгиз; 1961, 406 с.
20. Гмурман В.Е. *Теория вероятностей и математическая статистика*. М.: Юрайт; 2023, 480 с. ISBN 978-5-534-00859-3
21. Кудрявцев Л.Д. *Математический анализ*. Т. 1. М.: Высшая школа; 1973, 614 с.
22. Опойцев В.И. Нелинейный закон больших чисел. *Автоматика и телемеханика*. 1994;4:65–75.
23. Купер Дж., Макгилом К. *Вероятностные методы анализа сигналов и систем*: пер. с англ. М.: Мир; 1989, 376 с.

About the Author

Alexey N. Nazarov, Dr. Sci. (Eng.), Professor, Department of Corporate information systems, Institute of Information Technology, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: a.nazarov06@bk.ru. Scopus Author ID 7201780424, ResearcherID G-3154-2013, RSCI SPIN-code 6032-5302, <https://orcid.org/0000-0002-0497-0296>

Об авторе

Назаров Алексей Николаевич, д.т.н., профессор, кафедра «Корпоративные информационные системы», Институт информационных технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: a.nazarov06@bk.ru. Scopus Author ID 7201780424, ResearcherID G-3154-2013, SPIN-код РИНЦ 6032-5302, <https://orcid.org/0000-0002-0497-0296>

Translated from Russian into English by K. Nazarov

Edited for English language and spelling by Thomas A. Beavitt

MIREA – Russian Technological University.
78, Vernadskogo pr., Moscow, 119454 Russian
Federation.
Publication date July 31, 2026.
Not for sale.

МИРЭА – Российский технологический
университет.
119454, РФ, г. Москва, пр-т Вернадского, д. 78.
Дата опубликования 31.07.2026 г.
Не для продажи.

<https://www.rtj-mirea.ru>

