

УТВЕРЖДАЮ:

И.о. директора Государственного
учреждения «Институт прикладной
математики и механики»,



доктор физ.-мат. наук,
старший научный сотрудник
С.Н. Судаков

16 февраля 2022 г.

ОТЗЫВ

ведущей организации на диссертацию Рычки Ольги Валентиновны на тему:
«Совершенствование методов выявления и корректировки аномальных
измерений для повышения качества линейных регрессионных моделей»,
представленной на соискание ученой степени кандидата технических наук по
специальности 05.13.18 – Математическое моделирование, численные
методы и комплексы программ (технические науки)

1. Актуальность для науки и практики

Диссертационная работа Рычки О. В. направлена на решение проблем, возникающих при построении прогнозов на основе парных линейных регрессионных моделей, которые вызваны появлением в исходных данных аномальных значений. Такие значения оказывают существенное влияние на точность и дальность прогноза. Поэтому для построения надёжных прогнозов необходим предварительный анализ исходных статистических данных, с целью обнаружения и обработки аномальных значений. Аномальные данные могут возникать в результате технических и человеческих ошибок, влияния факторов, имеющих объективный характер, но проявляющихся эпизодически.

Существующие на сегодняшний день методы выявления и обработки аномальных значений недостаточно эффективны при использовании их для данных, имеющих линейную регрессионную зависимость. В связи с этим, задача разработки и совершенствования методов, алгоритмов и программных

средств поиска и корректировки аномалий, решению которой посвящена диссертация Рычки О.В., является актуальной.

2. Значимость полученных результатов

Основные результаты диссертационной работы заключаются в следующем:

- разработан усовершенствованный метод поиска аномальных данных, преимуществом которого является независимость от исходного объёма данных и возможность учитывать при поиске угол наклона графика функции парной линейной регрессии;

- на основе разработанного метода предложены две модификации, позволяющие сократить время анализа;

- предложен метод корректировки аномальных данных, заключающийся в замене аномальных значений на значения, которые будут расположены на границах области надёжности;

- доказана возможность использования предложенных в работе методов поиска аномалий и обработки данных для нелинейных регрессионных моделей с внутренней линейностью;

- на основе данных вычислительных экспериментов, проведено сравнение, демонстрирующее эффективность предложенных автором методов над наиболее известными методами поиска аномалий.

Практическая значимость работы заключается в:

- реализации предложенных методов в виде алгоритмов и разработке на их основе комплекса программ;

- обосновании рекомендуемых параметров использования предложенных методов поиска и обработки данных на основе которых даны практические рекомендации по выбору конкретного метода.

Практическая значимость результатов работы подтверждается их успешным внедрением в несколько научно-исследовательских работ, выполненных в Донецком национальном техническом университете, в ООО

НПО «Интермет», в учебный процесс. Все результаты внедрения подтверждены соответствующими справками, имеющимися в диссертации.

3. Обоснованность и достоверность научных положений, выводов и заключений

Все представленные в диссертации научные положения, выводы и рекомендации являются полностью обоснованными и достоверными.

Обоснованность определяется широким обзором существующих методов выявления аномалий, анализом их достоинств и недостатков.

Основные результаты работы базируются на научно обоснованных методах исследования. Достоверность подтверждается результатами успешной апробации предложенных в работе методов на тестовых наборах данных, результатами сравнения предложенного метода поиска аномалий с существующими методами на тех же наборах данных. Эффективность предложенного в работе метода сравнивалась с такими методами выявления выбросов как: метод Кука, «ящик с усами», Эктона, Титьена-Мура-Бэкмана, Прескотта-Лунда. Сравнение показало преимущество предложенного в диссертационной работе метода над существующими, заключающееся в обнаружении всех аномальных данных.

Помимо этого, в работе осуществлялось тестирование на реальных статистических данных, которые представляют собой зависимость оборота розничной торговли непродовольственными товарами, млн. руб. от среднедушевого денежного дохода населения в Российской Федерации, руб./месяц. Данные были взяты поквартально с 2013 по 2019 гг. Также после поиска аномальных данных и их корректировки был построен прогноз на следующие периоды. Полученные по найденной линейной регрессионной модели данные незначительно отличаются от реальных данных в прогнозируемых периодах. Это говорит, об эффективности, предлагаемых в работе методов повышения качества линейных регрессионных моделей.

Предложенные методы реализованы в виде программных модулей, объединенных в единый комплекс программ. В диссертационной работе

описаны этапы проектирования и программные модули системы. С помощью разработанного комплекса программ было проведено ряд вычислительных экспериментов, подтверждающих его функциональность и эффективность.

Таким образом, полученные Рычкой О.В. результаты научно обоснованы, а достоверность подтверждена результатами тестирования и опытной эксплуатации разработанного программного комплекса.

4. Общая характеристика работы

Диссертационная работа Рычки О.В. представляет собой содержательную, целостную работу. Полученные автором результаты изложены последовательно и корректно. Работа состоит из введения, четырёх глав, заключения, списка использованной литературы и 4 приложений.

Во введении обоснована актуальность темы диссертации, сформулированы цели и задачи, указаны объект и предмет исследования, сформулирована научная новизна, перечислены методы исследования, выделены основные результаты, которые выносятся на защиту, отмечаются практическая и теоретическая значимость полученных результатов, приводится информация об апробации работы, указано количество публикаций.

В первом разделе кратко рассматриваются основные понятия и предпосылки применения регрессионного анализа для построения моделей и осуществления на их основе прогнозов. Проведён анализ наиболее известных методов обнаружения аномальных данных. Выделены их достоинства и недостатки. На основании проведённого анализа определены цели и задачи работы.

Во втором разделе предложены и подробно описаны метод поиска аномалий, две его модификации, метод корректировки найденных аномальных данных. Отобраны критерии оценки эффективности предложенных методов.

В третьем разделе представлена программная реализация предложенных во втором разделе методов, приведено описание и выполнена

проверка работоспособности программного комплекса. Тестирование программного комплекса осуществлялось на ряде экспериментальных выборок.

В четвертом разделе доказывається эффективность предложенных методов с использованием вычислительных экспериментов. Проведён сравнительный анализ методов между собой и даны рекомендации по их использованию.

В заключении подтверждаются все научные положения и выводы, сделанные автором, а также завершенность диссертационной работы в целом.

5. Замечания:

1) в первом разделе достаточно подробно исследуются существующие методы поиска аномалий, однако, данный раздел можно было сократить, представив описания методов в сравнительной таблице;

2) в работе представлено описание для 9 экспериментальных наборов, однако не указано общее количество проведённых экспериментов;

3) необходимо было дать более подробную рекомендацию относительно количества отбрасываемых данных;

4) для разработанного комплекса программ не приводится полное руководство пользователя;

5) листинг кода программных модулей прокомментирован недостаточно подробно.

В целом указанные замечания не снижают ценности и практической значимости диссертационной работы.

6. Заключение по работе

Диссертационная работа Рычки О.В. «Совершенствование методов выявления и корректировки аномальных измерений для повышения качества линейных регрессионных моделей» является завершенной научно-квалификационной работой и содержит новые научные результаты и положения.

Результаты, полученные Рычкой О.В. в ходе выполнения диссертационной работы, обоснованы и имеют практическую значимость. Предложенные методы прошли апробацию на различных наборах данных и показали свою эффективность в сравнении с иными существующими методами.

В автореферате в достаточном объёме изложены основные результаты исследований, приведённых в диссертации.

Тематика и содержание работы соответствуют паспорту специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ (технические науки).

Таким образом, диссертационная работа соответствует требованиям п. 2.2. «Положения о присуждении ученых степеней» Донецкой Народной Республики, а ее автор, Рычка Ольга Валентиновна, заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ (технические науки).

Отзыв обсуждён и одобрен на заседании научного семинара отдела теории управляющих систем «Современные проблемы в машинном обучении и анализе данных» Государственного учреждения «Институт прикладной математики и механики» «25» ноября 2021 г., протокол № 2

Учёный секретарь ГУ «Институт прикладной математики и механики»,
кандидат технических наук
283048, ДНР, г. Донецк, ул. Розы Люксембург, д. 74
Тел.: +38 (062) 311-03-91; +38 (062) 311-01-75
E-mail: math.iamm@mail.ru

А.Ю. Максимова

Я, Александра Юрьевна Максимова, даю согласие на автоматизированную обработку моих персональных данных

А.Ю. Максимова

Подпись к.т.н. Максимовой А.Ю. удостоверяю:
начальник отдела кадров

