

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертацию Горшенина Андрея Константиновича

*«Полупараметрические методы анализа неоднородных данных и их применение
в задачах математического моделирования»,*

*представленную на соискание ученой степени доктора физико-
математических наук по специальности 05.13.18 – «Математическое
моделирование, численные методы и комплексы программ»*

Диссертация А.К. Горшенина посвящена развитию подходов на основе смешанных вероятностных моделей и выборок случайного объема для описания распределений характеристик процессов в различных прикладных областях. Кроме того, значительное внимание уделяется развитию статистических методов оценивания параметров смешанных распределений, объединенных в диссертации термином «полупараметрические». **Актуальность темы диссертационной работы** основывается на известных проблемах при использовании классических законов, в частности гауссовских, в качестве моделей эмпирических распределений в физике турбулентной плазмы и других областях. В подобных условиях может нарушаться корректность выводов, которые делаются на основе недостаточно обоснованных предположений о данных. Выбор в качестве основного исследовательского инструмента в диссертации аппарата теории вероятностей и математической статистики обосновывается стохастической природой многих реальных процессов.

Работа А.К. Горшенина ориентирована на строгое теоретическое обоснование вида вероятностных распределений, которые могут появиться в рамках рассматриваемых моделей для данных (например, сумм или максимального наблюдения). Это позволяет избежать эмпирического подхода при выборе соответствующей модели, например, для распределений размеров частиц лунного реголита или статистического определения величины экстремальных осадков.

Диссертация имеет следующую структуру: введение, 7 глав, заключение, список литературы (458 позиций), 28 таблиц, 175 рисунков и 30 вычислительных алгоритмов, описание которых выделено в отдельные фрагменты псевдокода. Объем диссертации составляет 355 страниц.

Во введении описана актуальность исследовательского направления, сформулированы цели, задачи, а также основные полученные соискателем результаты.

В главах 1-3 доказаны теоремы для выборок, объем которых является случайной величиной, аналитически исследованы свойства получаемых распределений, а также описаны и изучены основные методы и статистические алгоритмы, которые в дальнейших главах применены для анализа неоднородных данных в ряде различных прикладных областей.

В частности, результат теоремы 1.2 из раздела 1.3 использован в главе 6 для определения критического порогового уровня для объемов осадков. Новый вариант центральной предельной теоремы (теорема 1.10 из раздела 4) использован в главе 4 для обоснования использования моделей на основе логнормальных смесей для распределений размеров частиц лунного реголита. Описанный в главе 2 метод скользящего разделения смесей и теоретические результаты данного раздела, ориентированные на обоснование корректности вычислительных методов оценивания параметров смесей, существенным образом используются в главах 5-6 для анализа статистической структуры процессов в плазме и оценивания распределений неизвестных коэффициентов стохастического дифференциального уравнения, описывающего турбулентные потоки передачи тепла между океаном и атмосферой.

Весьма интересным представляется содержание главы 4, посвященной анализу распределений размеров частиц лунного реголита. Важным здесь представляется продемонстрированное соискателем отсутствие линейности в зависимости математического ожидания и дисперсии. Кроме того, сформулирована идея соотнесения параметров полученных смесей на основе кластеризации их параметров с химическим составом лунных реголитов. Работы по моделированию внешних воздействий на грунт на внеземных объектах имеют большое значение для космических исследований, поэтому эти методы, в совокупности с изучением плазмохимических реакций, могут быть весьма полезны в прикладном использовании.

Стоит отметить подходы на основе бутстреп-симуляции выборок, применяемые в этом разделе, а также в главе 5 при аппроксимации спектров: разнообразные подходы к имитационному моделированию весьма популярны в современных физических исследованиях. Также в главе 5 описаны результаты применения метода скользящего разделения смесей для детального изучения физических эффектов, возникающих в процессе электронно-циклотронного резонансного нагрева плазмы в стеллараторе Л-2М.

В главе 6 с использованием результатов главы 1 и на основе развития статистических процедур и применения методов машинного обучения решена задача выбора критериев экстремальности осадков и интенсивностей.

В главе 7 описаны разработанные соискателем программные комплексы. Весьма интересным представляется идея динамической визуализации результатов анализа методом скользящего разделения смесей (раздел 7.1), которая позволяет наглядно продемонстрировать изменение параметров во времени.

В заключении подводятся краткие итоги диссертационной работы.

Основные результаты диссертации, выносимые на защиту, сформулированы в виде семи пунктов.

Первый из них (ему посвящена глава 1 диссертации) ориентирован на развитие смешанных вероятностных моделей для выборок со случайным объемом с аналитическим исследованием свойств таких моделей и относится к направлению *«Разработка новых математических методов моделирования объектов и явлений»* паспорта специальности 05.13.18 – «Математическое моделирование, численные методы и комплексы программ».

Второй и третий результаты (глава 2) – полученные соискателем доказательства устойчивости в метрике Леви дисперсионно-сдвиговых и конечных сдвиговых смесей нормальных распределений относительно возмущений параметров смешивающего распределения, а также комплекс полупараметрических методов анализа неоднородных данных и результаты аналитического исследования некоторых их свойств в моделях аддитивного зашумления конечными смесями и округления наблюдений – соответствуют пункту *«Развитие качественных и приближенных аналитических методов исследования математических моделей»* паспорта специальности 05.13.18.

К направлению *«Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий»* относятся:

- полупараметрический подход к статистическому оцениванию распределений случайных коэффициентов стохастических дифференциальных уравнений Ланжевена (глава 2 с демонстрацией применения в главе 6);
- статистическая методология построения моделей сгруппированных скрытых наблюдений при заданных характерных точках их эмпирической функции распределения (глава 4);

– комплекс методов и алгоритмов статистической идентификации и классификации экстремальных наблюдений, на основе обобщенных отрицательных биномиальных распределений числа наблюдений и обобщенных гамма-моделей для данных (глава 6).

Наконец, созданные соискателем программные комплексы для автоматизации обработки массивов неоднородных данных на высокопроизводительных вычислительных ресурсах, реализующие разработанные полупараметрические методы, решение с их помощью некоторых задач математического моделирования в физике плазмы, селенологии, метеорологии, океанологии относятся к пунктам *«Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента»* и *«Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента»*.

Таким образом, полученные А.К. Горшениным **результаты являются новыми и полностью соответствуют областям специальности 05.13.18 – «Математическое моделирование, численные методы и комплексы программ»**. При их получении использованы классические и развиваемые соискателем математические методы и подходы к анализу данных. Доказательства теорем корректны, вычислительные процедуры детально описаны и исследованы в том числе аналитически, продемонстрированы содержательные примеры применения разработанных методов для анализа прикладных данных. Таким образом, полученные в работе **результаты представляются полностью научно обоснованными и достоверными**. Их **теоретическая значимость** заключается в существенном развитии подходов в области вероятностно-статистического моделирования, а **прикладная** – в решении с использованием развитых соискателем методов задач сразу из нескольких весьма различающихся прикладных областей.

К достоинствам работы можно также отнести проведение расчетов с использованием ресурсов центра коллективного пользования «Информатика» ФИЦ «Информатика и управление» РАН. Это открывает новые **возможности для дальнейшего прикладного использования результатов** в области анализа данных. В частности, разработанные программные комплексы могут эффективно использоваться для анализа частиц, моделирующих лунный грунт и получаемых в экспериментах с использованием гиротрона в качестве источника

СВЧ излучения. Эти исследования ориентированы на решение важной задачи тестирования свойств материалов, используемых при конструировании космических исследовательских аппаратов для функционирования в сложных внеземных условиях.

Соискателем опубликованы 82 печатные работы по тематике диссертации, включая 31 статью в журналах, рекомендованных ВАК РФ, и 51 работу в изданиях, индексируемых в Web of Science и (или) Scopus. В Роспатенте зарегистрированы 39 свидетельств о государственной регистрации программ для ЭВМ. **Личный вклад соискателя, в том числе в работах с соавторами, не вызывает сомнений.** Результаты представлялись на ведущих международных и российских научных мероприятиях по тематике диссертации, а исследования соискателя в разное время были поддержаны грантами и стипендиями Президента РФ для молодых ученых, Российским научным фондом, Российским фондом фундаментальных исследований Президента РФ.

По изложению материала в работе имеются некоторые **замечания.**

1. В разделах 2.2 и 2.3 рассматриваются результаты для двух моделей – добавления и расщепления компоненты. Следовало бы более подробно обосновать выбор моделей такого вида с точки зрения приложений.

2. Для прогнозирования физических экспериментальных данных в разделе 5.2.4 использованы нейронные сети прямого распространения, в то время как далее в 5.3 для прогнозирования моментных характеристик используются и LSTM-сети. Чем обоснован выбор архитектур в разделе 5.2?

3. В работе в разделах 5.2.2 и 6.5.2 не обсуждается скорость работы метода определения связности компонент для физических и океанологических данных.

Указанные замечания не снижают общее положительное впечатление от диссертации. **Существенные недочеты в работе не обнаружены.** Автореферат корректно отражает суть проведенных исследований и полученных А.К. Горшениным результатов.

Диссертация А.К. Горшенина является научно-квалификационной работой высокого научного уровня. Результаты, представленные в диссертации, являются **научным достижением** в развитии методов математического моделирования на основе выборок случайного объема, предельных теорем теории вероятностей и полупараметрических статистических методов. Кроме того, разработан широкий набор программных решений, ориентированных на автоматизацию прикладного анализа неоднородных данных.

Докторская диссертация А.К. Горшенина соответствует всем требованиям Положения о присуждении ученых степеней, утвержденного Постановлением Правительства РФ от 24 сентября 2013 г. №842. Горшенин Андрей Константинович безусловно заслуживает присуждения степени доктора физико-математических наук по специальности 05.13.18 – «Математическое моделирование, численные методы и комплексы программ».

Официальный оппонент

Майоров Сергей Алексеевич,
доктор физико-математических наук по
специальностям 01.04.08 – «Физика и химия плазмы»
и 05.13.18 – «Теоретические основы
математического моделирования, численные методы
и комплексы программ», доцент, ведущий научный
сотрудник теоретического отдела Федерального
государственного бюджетного учреждения науки
Федерального исследовательского центра «Институт
общей физики им. А.М. Прохорова Российской
академии наук»

Адрес: 119991 ГСП-1, Москва, ул. Вавилова, д.38

Тел.: +7-905-784-50-58, 499-135-02-47

e-mail: mayorov_sa@mail.ru

«_29_» апреля 2021 г.

С.А. Майоров

Подпись доктора физико-математических наук, доцента, ведущего научного
сотрудника теоретического отдела ИОФ РАН Майорова Сергея Алексеевича
заверяю.

И.о. Ученого секретаря ИОФ РАН д.ф.-м.н.



В.В. Глушков