

Федеральное государственное бюджетное учреждение науки
Институт вычислительной математики и математической геофизики
Сибирского отделения Российской академии наук

На правах рукописи



Каргаполова Нина Александровна

**АЛГОРИТМЫ ЧИСЛЕННОГО СТОХАСТИЧЕСКОГО
МОДЕЛИРОВАНИЯ НЕСТАЦИОНАРНЫХ МЕТЕОРОЛОГИЧЕСКИХ И
БИОКЛИМАТИЧЕСКИХ ПРОЦЕССОВ**

01.01.07 – вычислительная математика

ДИССЕРТАЦИЯ

на соискание ученой степени
доктора физико-математических наук

Научный консультант
доктор физико-математических наук
Огородников Василий Александрович

Новосибирск – 2022

ОГЛАВЛЕНИЕ

Введение	5
Глава 1. Некоторые вопросы моделирования случайных процессов и полей	19
1.1 Моделирование двумерных однородных случайных полей с экспоненциальной корреляционной функцией	19
1.1.1 Модифицированный нерандомизированный алгоритм «по строкам и столбцам»	22
1.1.2 Рандомизированный алгоритм «по строкам и столбцам»	30
1.2 Стохастическая интерполяция случайного процесса	37
1.3 Алгоритмы моделирования условных негауссовских случайных процессов с точечными и интервальными условиями	46
Глава 2. Стохастические «генераторы погоды»	55
2.1 Стохастическое моделирование рядов метеорологических параметров и их комплексов.....	55
2.1.1 Стохастическая модель совместных рядов индикатора выпадения осадков, минимальной и максимальной за сутки приземной температуры воздуха	55
2.1.2 Моделирование совместных нестационарных временных рядов различных метеорологических параметров	70
2.2 Моделирование неоднородного нестационарного пространственно-временного поля комплекса метеорологических параметров на сети метеостанций	86
Глава 3. Численные стохастические модели биоклиматических процессов...	103
3.1 Биоклиматические индексы. Определения и обозначения	103
3.2 Применение стохастических «генераторов погоды» для моделирования временных рядов биоклиматических индексов	108
3.2.1 Моделирование временных рядов биоклиматических индексов с учётом их суточного хода	109

3.2.2.	Стохастическое моделирование временных рядов среднесуточного индекса жары	123
3.3	Моделирование временных рядов биоклиматических индексов на основе определяющих формул	126
3.3.1	Стохастические модели безусловных рядов биоклиматических индексов	126
3.3.2	Стохастические модели условных рядов биоклиматических индексов	137
3.4	Моделирование пространственных и пространственно-временных полей биоклиматических индексов.....	143
3.4.1	Моделирование пространственных и пространственно-временных полей индекса жары и индекса холодового стресса на сети метеостанций	143
3.4.2	Об интерполяции пространственного поля индекса холодового стресса со станций в узлы регулярной сетки	159
Глава 4.	Исследование некоторых классов случайных процессов с периодическими характеристиками	168
4.1	Модели периодически коррелированных кусочно-постоянных и кусочно-линейных процессов	168
4.1.1	Модель периодически коррелированного кусочно-постоянного процесса на основе стационарного случайного процесса с дискретным временем	168
4.1.2	Модель периодически коррелированного кусочно-постоянного процесса на основе стационарного случайного процесса и регулярного потока	174
4.1.3	Модель периодически коррелированного кусочно-линейного процесса на основе стационарного случайного процесса и регулярного потока	180
4.2	Асимптотически периодически коррелированные процессы	191

4.3	Неоднородные марковские последовательности с матрицей переходных вероятностей, периодически зависящей от времени	195
4.4	Модели асимптотически периодически коррелированных кусочно-постоянных процессов	203
4.4.1	Модель асимптотически периодически коррелированного кусочно-постоянного процесса на основе бинарной неоднородной цепи Маркова	204
4.4.2	Модель асимптотически периодически коррелированного кусочно-постоянного процесса на основе нерегулярного точечного потока	207
	Заключение	211
	Список литературы	213
	Приложение 1. Известные методы моделирования гауссовских и негауссовских случайных процессов	235
	Приложение 2. Сведения о метеорологических станциях	239
	Приложение 3. Некоторые результаты верификации моделей, предложенных в параграфах 3.3 и 3.4 Главы 3	243

ВВЕДЕНИЕ

Актуальность и степень разработанности темы исследований. Теоретическая и практическая значимость работы.

При решении различных задач физики, климатологии и метеорологии, популяционной биологии и эпидемиологии, а также задач в других областях науки часто используются численные стохастические модели исследуемых процессов. Аппарат численного стохастического моделирования различных классов случайных процессов к настоящему времени достаточно хорошо развит [23, 41, 43, 48, 50-53, 59, 62, 66, 81, 85, 171, 175, 209]. Методы моделирования выбирают в зависимости от решаемой задачи и существующих требований на точность и время счёта. Но, несмотря на обилие разработанных методов и алгоритмов, при решении некоторых конкретных задач существующие методы не всегда дают приемлемые (по тем или иным критериям) результаты. Поэтому разработка гибких и легко адаптируемых под различные требования алгоритмов моделирования случайных процессов имеет очевидную теоретическую и практическую значимость.

Стохастический подход к моделированию и исследованию метеорологических и климатических процессов начал активно развиваться в 60 – 70-ые годы XX века [27, 44, 45, 58, 65, 107, 120, 155, 204]. В этот период стохастический подход чаще всего применялся для моделирования стационарных временных рядов суточного разрешения отдельных метеорологических параметров (в основном – суточного количества осадков). Позднее появились модели нестационарных временных рядов, пространственных и пространственно-временных полей отдельных метеопараметров и их комплексов [106, 117, 133, 157, 170, 182, 183, 187, 188, 200]. На основе входных данных различных типов (данных срочных метеонаблюдений, осредненных данных наблюдений и т.д.) построенные стохастические модели позволяют моделировать метеорологические процессы с заданным временным и/или пространственным разрешением. В литературе

стохастические модели метеопроцессов часто называют стохастическими «генераторами погоды». По своей сути, «генераторы погоды» являются стохастическими моделями и соответствующими им пакетами программ, позволяющими численно моделировать длинные ряды случайных чисел, обладающих статистическими свойствами, повторяющими основные свойства реальных метеорологических рядов и полей [92, 165].

Об актуальности разработки стохастических «генераторов погоды», и их теоретической и практической значимости свидетельствует то, что в настоящее время «генераторы погоды» активно используются для решения задач в различных областях науки, в том числе

- в климатологии и метеорологии при исследованиях, связанных с изменением климата и с анализом экстремальных метеособытий [118, 185, 172, 192, 206],

- в гляциологии при исследовании процессов формирования снежного покрова (изменения во времени толщины снега, содержания льда и талой воды в нём, процессов снеготаяния, сублимации и замерзания талой воды в толще снега) [13],

- в гидрологии при исследовании процессов массопереноса и биогенной нагрузки в системе водосбор – водоём [12, 34, 35, 89],

- в ветроэнергетике для проведения оценки эффективности действия ветрогенераторов [159, 199],

- при решении задач агрометеорологии, связанных с исследованием продукционного процесса сельскохозяйственных растений [9, 61, 86, 181],

- при анализе транспортных и транспортно-технологических систем [24, 38].

В зависимости от решаемой задачи, моделируются либо значения отдельных метеопараметров и их комплексов (приземной температуры воздуха, количества осадков, относительной влажности воздуха, модуля и направления скорости ветра и др.), либо их некоторые специальные характеристики (индикаторы наличия / отсутствия осадков, минимальная и максимальная за сутки температура и т.п.). Кроме это, при решении задач используются стохастические модели с различным

шагом по времени (например, месячным, суточным, трёхчасовым, часовым или минутным).

Следует отметить, что численное стохастическое моделирование метеорологических процессов часто является чрезвычайно трудоёмким (особенно, если моделируются комплексы метеопараметров как функции временных и пространственных координат). В связи с этим, разработка и реализация эффективных стохастических «генераторов погоды» является актуальной и практически значимой задачей.

Окружающая среда чрезвычайно сильно влияет на здоровье человека и его самочувствие. Загрязнение воздуха и воды, высокий уровень естественного и техногенного радиационного излучения, шумовое и световое загрязнение оказывают негативное воздействие на здоровье. Ещё одним природным фактором, который может оказать отрицательное влияние на состояние человека, является тепловое воздействие окружающей среды.

Первые попытки описать зависимость теплового состояния человека от погодных условий были сделаны в 20 – 30-х годах прошлого века [130, 168, 207]. С тех пор было разработано большое количество биометеорологических показателей (индексов), описывающих тепловой эффект (или, другими словами, характеризующих уровень биоклиматического комфорта / дискомфорта) различных комбинаций метеорологических параметров (температуры и относительной влажности воздуха, скорости ветра, атмосферного давления, солнечного излучения и т. д.) на человека. В настоящее время в литературе можно найти описание более 100 различных биоклиматических индексов [119], обзоры и классификация которых представлены, например, в [16, 26, 72, 83, 105, 110]. Некоторые из этих индексов, например, биоклиматический индекс суровости метеорежима и эквивалентно-эффективная температура воздуха, используются для описания теплового воздействия температуры воздуха и других метеорологических параметров в течение всего года [74, 96, 103]. Другие (например, индекс жары и Humidex) применяются исключительно для описания тепловых эффектов в летнее время [97, 161, 163, 191]. Отдельный класс индексов

составляют биоклиматические показатели, описывающие влияние холодного воздуха на тепловой баланс человека, например, приведенная температура и индекс холодового стресса [1, 179].

Биоклиматические индексы (БИ) нашли широкое применение. Значения некоторых из них можно увидеть в ежедневном прогнозе погоды (соответствующие значения обозначаются, например, «Ощущается как», «Real feel») [90, 91]. В ряде стран (в том числе, в США, Китае, Канаде, Франции) многие из этих индексов активно используются в работе оперативных систем предупреждения о наступлении некомфортных погодных условий [160]. Биоклиматические индексы используют также при исследовании связей между погодными условиями и обострениями различных заболеваний человека [22, 189, 194], для медико-климатической классификации территорий и их климато-рекреационного районирования [17, 25, 68]. Для решения таких задач необходимо, среди прочего, исследовать характеристики различных редких неблагоприятных биоклиматических явлений. Провести такое исследование по реальным данным не всегда представляется возможным в силу ограниченного (и, часто, малого) объёма выборки. Однако такие исследования можно провести, используя модельные значения БИ. Соответственно, разработка численных моделей, адекватно воспроизводящих различные свойства реальных биоклиматических процессов, является актуальной и значимой задачей.

Традиционно для исследования свойств биоклиматических процессов, характеризующихся изменением значений БИ во времени и по пространству, используются 2 подхода. Первый из них – статистический. В рамках этого подхода проводится статистический анализ реальных значений биоклиматических показателей в определённой местности в заданный период времени [63, 208]. В рамках второго, динамического, подхода для исследования применяются динамические модели метеорологических процессов [84, 115, 121]. В данной диссертационной работе рассмотрен новый подход к исследованию биоклиматических процессов. Суть этого подхода заключается в том, что биоклиматические процессы рассматриваются как случайные процессы и поля, и,

следовательно, могут быть исследованы с использованием аппарата стохастического моделирования. Впервые этот подход был предложен при непосредственном участии автора диссертационной работы в статье [148]. Позднее стохастический подход был использован и другими авторами [123, 124, 197].

При построении стохастических «генераторов погоды» и стохастических моделей БИ с временным шагом меньшим 1 суток необходимо учитывать суточный ход реальных процессов. Одним из способов такого учёта является применение предположения о периодической коррелированности рассматриваемых временных рядов и полей. Напомним, что случайный процесс $\xi(t)$, $t \in \mathbb{R}$ называется периодически коррелированным, если существует такое число $T > 0$, что математическое ожидание $E\xi(t)$, дисперсия $D\xi(t)$ и корреляционная функция $corr(\xi(t_1), \xi(t_2))$ процесса $\xi(t)$ удовлетворяют условиям

$$\begin{aligned} \forall t \quad E\xi(t+T) &= E\xi(t), \quad D\xi(t+T) = D\xi(t), \\ \forall t_1, t_2 \quad corr(\xi(t_1+T), \xi(t_2+T)) &= corr(\xi(t_1), \xi(t_2)). \end{aligned}$$

Процессы такого типа были использованы в работах [21, 69-71], связанных с изучением ритмики океанологических процессов. В вышеуказанных работах была изучена спектральная структура таких процессов. Задача моделирования периодически коррелированных процессов решается различными способами. Разработаны, например, алгоритмы моделирования с использованием векторных процессов авторегрессии [20, 70, 73, 175], модели с использованием некоторых типов точечных потоков [176], спектральные модели [164]. Эти алгоритмы позволяют моделировать случайные процессы с заданной корреляционной структурой. Однако при решении конкретных задач возникают некоторые трудности, связанные, например, с выбором матричных коэффициентов, гарантирующих стационарность процесса авторегрессии. В связи с этим, представляется значимым разработать новые алгоритмы моделирования некоторых классов периодически коррелированных рядов.

Цели и задачи

Целью работы является разработка и исследование алгоритмов численного моделирования случайных процессов с заданными вероятностными свойствами, разработка численных стохастических моделей метеорологических и биоклиматических процессов, учитывающих их временную нестационарность и пространственную неоднородность.

Для достижения указанной цели необходимо решить следующие задачи:

- 1) разработать алгоритмы моделирования двумерных однородных случайных полей с корреляционными функциями специального вида, алгоритмы стохастической интерполяции стационарных, нестационарных процессов и неоднородных полей, алгоритмы моделирования условных негауссовских процессов;
- 2) разработать стохастические «генераторы погоды» для численного моделирования комплексов нестационарных метеорологических процессов и неоднородных пространственно-временных полей;
- 3) разработать алгоритмы стохастического моделирования временных рядов, пространственных и пространственно-временных полей биоклиматических индексов, учитывающие нестационарность по времени и неоднородность по пространству реальных процессов;
- 4) разработать алгоритмы моделирования периодически коррелированных кусочно-постоянных и кусочно-линейных случайных процессов;
- 5) разработать алгоритмы моделирования асимптотически периодически коррелированных кусочно-постоянных случайных процессов.

Научная новизна

Научная новизна работы определяется совокупностью полученных результатов:

- 1) впервые предложены и разработаны стохастические модели временных рядов, пространственных и пространственно-временных полей биоклиматических индексов (биоклиматического индекса суровости метеорежима, индекса

холодового стресса, индекса жары и энтальпии влажного воздуха), учитывающие временную нестационарность и пространственную неоднородность реальных процессов. Модели основаны на специально разработанных численных стохастических моделях метеорологических процессов и на применении определяющих биоклиматические индексы формул;

- 2) разработаны новые стохастические модели нестационарных неоднородных негауссовских временных рядов и пространственно-временных полей комплексов метеорологических параметров и их специальных характеристик, и на их основе рассчитаны различные характеристики аномальных метеособытий;
- 3) предложены новые алгоритмы моделирования двумерных однородных случайных полей и условных негауссовских процессов с точечными и интервальными условиями;
- 4) разработано семейство новых алгоритмов моделирования периодически коррелированных и асимптотически периодически коррелированных кусочно-постоянных и кусочно-линейных случайных процессов, основанных на использовании различных типов точечных потоков.

Методология и методы исследования

В диссертационной работе для достижения поставленной цели и решения сформулированных выше задач использовались:

– аппарат теории методов Монте-Карло, включая методы численного моделирования различных классов случайных процессов (в том числе, гауссовских и негауссовских, стационарных, периодически коррелированных и нестационарных);

– аппарат теории вероятностей и математической статистики, в частности, теория случайных процессов, методы оценивания параметров стохастических моделей по реальным данным, критерии проверки статистических гипотез;

– аппарат вычислительной математики, например, численные методы решения нелинейных уравнений и проблемы собственных значений, методы численного интегрирования;

– аппарат математического анализа и линейной алгебры (включая теорию пределов, теорию интегрирования, теорию матриц);

– методы статистической метеорологии и климатологии;

– языки программирования C++ и FORTRAN для написания вычислительных программ.

Основные положения, выносимые на защиту

На защиту выносятся следующие результаты исследований:

- 1) алгоритмы стохастического моделирования временных рядов, пространственных и пространственно-временных полей биоклиматических индексов, позволяющие учитывать их нестационарность по времени и неоднородность по пространству;
- 2) алгоритмы стохастического моделирования временных рядов и пространственно-временных полей комплексов метеорологических процессов, позволяющие учитывать их нестационарность и/или неоднородность;
- 3) алгоритмы моделирования периодически коррелированных и асимптотически периодически коррелированных кусочно-постоянных и кусочно-линейных случайных процессов, алгоритмы моделирования условных случайных негауссовских процессов, специального класса однородных двумерных случайных полей и метод стохастической интерполяции случайных процессов.

Соответствие паспорту специальности

Данное диссертационное исследование выполнено согласно паспорту специальности 01.01.07 «Вычислительная математика». Результаты диссертации удовлетворяют формуле специальности «вычислительная математика – область

науки, к которой относятся разработка и теория методов численного решения математических задач, возникающих при моделировании естественнонаучных и прикладных проблем, а также реализация методов в практическом решении задач с применением современных ЭВМ» и соответствуют 1-ому, 2-ому и 4-ому пунктам из основных направлений специальности:

- Создание алгоритмов численного решения задач алгебры, анализа, дифференциальных и интегральных уравнений, математической физики, теории вероятностей и статистики, типичных для приложений математики к различным областям науки и техники.
- Разработка теории численных методов, анализ и обоснование алгоритмов, вопросы повышения их эффективности.
- Реализация численных методов в решении прикладных задач, возникающих при математическом моделировании естественнонаучных и научно-технических проблем, соответствие выбранных алгоритмов специфике рассматриваемых задач.

Степень достоверности и апробация результатов

В диссертационной работе использованы научные методы обоснования полученных результатов и выводов. При получении теоретических результатов корректно использовался необходимый математический аппарат, проведены доказательства сформулированных утверждений. При численной реализации алгоритмов проводились исследования корректности их работы (тестирование). Разработанные модели прошли детальную верификацию.

Представленные результаты прошли научное рецензирование в процессе публикации в ведущих отечественных и зарубежных журналах. Материалы работы докладывались и обсуждались со специалистами в области моделирования случайных процессов и со специалистами-климатологами на представительных российских и международных конференциях.

Основные результаты диссертационного исследования обсуждались на объединенном семинаре Института вычислительной математики и математической

геофизики СО РАН и кафедры вычислительной математики Механико-математического факультета Новосибирского государственного университета «Методы Монте-Карло в вычислительной математике и математической физике», а также были представлены на 37 российских и международных конференциях: IX Международной научно-технической конференции «Аналитические и численные методы моделирования естественнонаучных и социальных проблем» (Россия, Пенза, 28-31 октября 2014), Конференции молодых ученых ИВМиМГ СО РАН (Россия, Новосибирск, 7-9 апреля 2014, 06-08 апреля 2015, 11-13 апреля 2016), Международной конференции «Актуальные проблемы вычислительной и прикладной математики» (Россия, Новосибирск, 8-11 июня 2014, 19-23 октября 2015), Рабочей группе «Аэрозоли Сибири» (Россия, Томск, 27-30 ноября 2018, 25-29 ноября 2019), Международной конференции «Дистанционные методы зондирования Земли и фотограмметрия, мониторинг окружающей среды, геоэкология» в рамках Научного конгресса «Интерэкспо ГЕО-Сибирь» (Россия, Новосибирск, 14-15 апреля 2015, 17-20 апреля 2017, 24-27 апреля 2018, 17-26 апреля 2019), International Workshop «Applied Methods of Statistical Analysis», (Россия, Новосибирск и Белокуриха 14-19 сентября 2015; Красноярск, 18-22 сентября 2017; Новосибирск, 18-20 сентября 2019), Марчуковских научных чтениях (Россия, Новосибирск, 25-30 июня 2017, 8-10 октября 2018, 1-6 июля 2019, 19-23 октября 2020, 4-8 октября 2021), 3rd Workshop on Stochastic Weather Generators (Франция, Ванн, 17-20 мая 2016), 4th Stochastic Modeling Techniques and Data Analysis International Conference (Мальта, Валетта, 1-4 июня 2016), 5th Data Analysis and Modeling of the Earth System Conference (Германия, Гамбург, 26-28 сентября 2016), International Workshop on Data Science & Environment (Франция, Брест, 03-07 июля 2017), International Workshop on Simulation (Австрия, Вена, 21-25 сентября 2015; Испания, Барселона, 25-29 июня 2018), 9th International Workshop on Applied Probability (Венгрия, Будапешт, 18-21 июня 2018), International Conference on Extreme Value Analysis (Нидерланды, Делфт, 25-30 июня 2017; Хорватия, Загреб, 1-5 июля 2019), 6th International Conference on Time Series and Forecasting (Испания, Гранада, 25-27 сентября 2019), International Conference on

Simulation and Modeling Methodologies, Technologies and Applications – SIMULTECH (Испания, Мадрид, 26-29 июля 2017; Португалия, Порту, 29-31 июля 2018; Чехия, Прага, 29-31 июля 2019), European Simulation and Modelling Conference (Испания, Лас-Пальмас, 26-28 октября 2016; Португалия, Лиссабон, 24-28 октября 2017; Бельгия, Гент, 24-26 октября 2018; Испания, Пальма-де-Майорка, 28-30 октября 2019).

Представленные в диссертационной работе результаты были получены в ходе выполнения исследований по государственным заданиям (0251-2021-0002, 0315-2019-0002, 0315-2016-0002, 0315-2014-0002), гранту Президента РФ по государственной поддержке ведущих научных школ НШ-5111.2014.1, гранту Благотворительного фонда В. Потанина ГПК-50/16, гранту Президента РФ МК-659.2017.1, грантам РФФИ № 11-01-00641-а, 12-01-00727-а, 15-01-01458_а, 15-01-08988_а, 16-31-00123_мол_а, 16-31-00038_мол_а, 18-01-00149_а, 19-41-543001_р_мол_а, крупному научному проекту МинОбрНауки 075-15-2020-787, гранту РНФ № 21-71-00007.

Личный вклад автора

Все основные научные результаты диссертационной работы получены лично автором или при его непосредственном участии. В частности, вклад автора диссертационного исследования был определяющим при разработке, теоретическом и численном исследовании предложенных алгоритмов моделирования случайных процессов, при построении, обосновании и верификации моделей метеорологических и биоклиматических процессов, реализации комплекса вычислительных программ для проведения численных экспериментов. Конфликт интересов с соавторами отсутствует.

Публикации

По теме диссертационной работы Каргаполовой Н.А. опубликовано 33 статьи [5, 6, 28-31, 33, 94, 101, 134-154, 172, 173, 178], в том числе 22 работы [101, 135-138, 140-149, 151-154, 172, 173, 178] в изданиях, зарегистрированных в базах

данных Web of Science и/или Scopus, из которых 12 статей [101, 138, 145-149, 152-154, 172, 178] – в журналах из перечня ВАК. В опубликованных работах отражено основное содержание, результаты и выводы диссертационного исследования.

Объем и структура диссертации

Диссертация состоит из введения, 4 глав, заключения, списка литературы и 3 приложений. Диссертация изложена на 247 страницах, включает библиографический список из 210 наименований, содержит 74 рисунка и 46 таблиц.

Краткое содержание работы

В **Главе 1** предложены алгоритмы моделирования двумерных однородных случайных полей с корреляционными функциями специального вида. Кроме этого, в **Главе 1** рассмотрены вопросы, связанные со стохастической интерполяцией стационарных и нестационарных случайных процессов, в частном случае представляющей собой рандомизированный вариант метода интерполяции обратных взвешенных расстояний. Изучена корреляционная структура процесса непрерывного аргумента, построенного с помощью стохастической интерполяции процесса дискретного аргумента. В **Главе 1** предложены также алгоритмы моделирования условных негауссовских процессов с точечными и интервальными условиями. Разработанный алгоритм моделирования условных негауссовских процессов с точечными условиями основан на методе нормализации, а алгоритмы моделирования процессов с интервальными условиями – на комбинации метода обратных функций распределения либо с методом исключения, либо с методом моделирования условных гауссовских процессов со специальным образом выбранными точечными условиями. Основные результаты, представленные в **Главе 1**, опубликованы в работах [5, 6, 101, 150, 153, 172].

Глава 2 посвящена разработке, исследованию и верификации стохастических «генераторов погоды». В этой главе предложены алгоритмы стохастического моделирования временных рядов и пространственно-временных

полей комплексов метеорологических параметров и их специальных характеристик. В частности, в Главе 2 предложен стохастический «генератор погоды», предназначенный для моделирования совместного пространственно-временного поля приземной температуры воздуха, вектора скорости ветра и количества осадков на сети метеорологических станций, расположенных в южной части Байкальской природной территории. Поле моделируется в предположении о его пространственной неоднородности и временной периодической коррелированности с периодом равным 1 суткам. Для моделирования негауссовского поля был применён метод нормализации и итерационный алгоритм, предложенный в работе [210]. Результаты, представленные в Главе 2, опубликованы в работах [28, 29, 33, 135, 136, 140, 143, 144, 147, 149, 151, 173].

В Главе 3 приведены результаты разработки, исследования, верификации и сравнения различных моделей временных рядов, пространственных и пространственно-временных полей биоклиматических индексов (индекса холодного стресса, биоклиматического индекса суровости метеорежима, индекса жары, энтальпии влажного воздуха). При построении моделей использовались 2 подхода, первый из которых основан на применении стохастических «генераторов погоды», а второй – на непосредственном применении определяющих формул для каждого из рассматриваемых биоклиматических индексов. На основе второго подхода, были построены и верифицированы, в том числе, модели пространственно-временного поля среднесуточного индекса жары на сети метеостанций, расположенных на юге России, пространственного и пространственно-временного полей ИХС на юге Западной Сибири на сети метеостанций и на регулярной сетке. Результаты, представленные в Главе 3, опубликованы в работах [94, 138, 139, 141, 142, 145, 146, 148, 153, 154].

В Главе 4 рассмотрено семейство конструктивно определенных периодически коррелированных кусочно-постоянных и кусочно-линейных случайных процессов непрерывного аргумента, построенных на основе стационарных случайных процессов дискретного аргумента и случайных потоков точек различного типа. Кроме этого, исследованы некоторые свойства

асимптотически периодически коррелированных случайных процессов, в том числе, кусочно-постоянных процессов, построенных с использованием точечных потоков. Основные результаты, представленные в Главе 4, опубликованы в работах [30, 31, 134, 137, 152, 178].

В **Заключении** приведены основные результаты диссертационной работы и представлены перспективы дальнейшей разработки темы исследования.

В **Приложении 1** описаны некоторые известные методы моделирования гауссовских и негауссовских случайных процессов. Эти методы были использованы в Главах 2 и 3.

Приложение 2 содержит сведения о метеорологических станциях, данные наблюдений с которых были использованы для определения параметров и верификации стохастических моделей метеорологических и биоклиматических процессов в Главах 2 и 3.

В **Приложении 3** представлены некоторые результаты верификации моделей, предложенных в параграфах 3.3 и 3.4 Главы 3.

Благодарности

Выражаю искреннюю признательность и благодарность своему научному консультанту д.ф.-м.н. Огородникову Василию Александровичу за многолетнюю поддержку, постоянное внимание к работе и ценные рекомендации. Мне приятно выразить глубокую благодарность к.ф.-м.н. Хлебниковой Елене Ивановне (ФГБУ ГГО им. А.И. Воейкова, Санкт-Петербург) за полезные консультации и предоставленные материалы. Хочу сказать огромное спасибо члену-корреспонденту РАН Михайлову Геннадию Алексеевичу, д.ф.-м.н. Каргину Борису Александровичу и всем коллегам в ИВМиМГ СО РАН за профессиональные советы и создание доброжелательной атмосферы, способствующей проведению научных исследований. Отдельно хочу поблагодарить всех соавторов публикаций за совместную работу и многочисленные полезные дискуссии.

Глава 1. Некоторые вопросы моделирования случайных процессов и полей

1.1. Моделирование двумерных однородных случайных полей с экспоненциальной корреляционной функцией

При решении различных прикладных задач с использованием методов стохастического моделирования, в частности задач статистической метеорологии, океанологии, гидрологии и популяционной биологии, необходимо учитывать корреляционную структуру реальных процессов и полей. Во многих случаях соответствующие корреляционные функции, необходимые для моделирования, получают аппроксимацией выборочных корреляционных функций, оцененных по данным наблюдений, специальными положительно определенными функциями непрерывного аргумента. Например, для описания корреляционной структуры пространственных двумерных полей приземной температуры воздуха и геопотенциала на заданной территории, если предположение об изотропности поля согласуется с реальными данными, используют аппроксимирующие функции вида

$$r(\rho) = \exp(-\alpha\rho^2),$$

где $\rho = \sqrt{x^2 + y^2}$ – расстояние между двумя точками поля. В случае, когда с данными согласуется предположение об однородности поля, используют функции более общего вида [11, 60, 82]

$$r(x, y) = \exp\left(-\left(\alpha x^2 + \beta xy + \gamma y^2\right)^\theta\right).$$

Существующие алгоритмы моделирования случайных полей, например, векторные алгоритмы, основанные на методе условных распределений, на векторных моделях авторегрессии, скользящего среднего, смешанных моделях авторегрессии и

скользящего среднего и др., позволяют строить реализации гауссовских полей дискретных аргументов с заданной корреляционной структурой [175]. Однако, они обладают общим недостатком: в общем случае моделирование большого количества реализаций случайного поля оказывается достаточно трудоёмкой задачей в смысле затрат времени на вычисления и затрат машинной памяти. Кроме того, для корреляционных функций указанного выше вида стандартные алгоритмы моделирования вычислительно неустойчивы, поскольку соответствующие корреляционные матрицы обладают большим числом обусловленности и собственными числами, близкими к нулю. Так, например, в [5] приведены результаты численных экспериментов, показывающих, что алгоритм, основанный на методе условных распределений вероятностей неустойчив при $\theta > 0.85$. В качестве иллюстрации, на рисунке 1.1 приведены теоретические изолинии двумерного гауссовского поля с корреляционной функцией

$$r(x, y) = \exp\left(-\left(0.00098x^2 - 0.00016xy - 0.00147y^2\right)^{0.9}\right)$$

и изолинии поля с такой корреляционной функцией, смоделированного с использованием метода условных распределений (метод условных распределений описан в Приложении 1) Оценка корреляционной функции проводилась по 10^5 реализаций поля размерности 25×30 узлов.

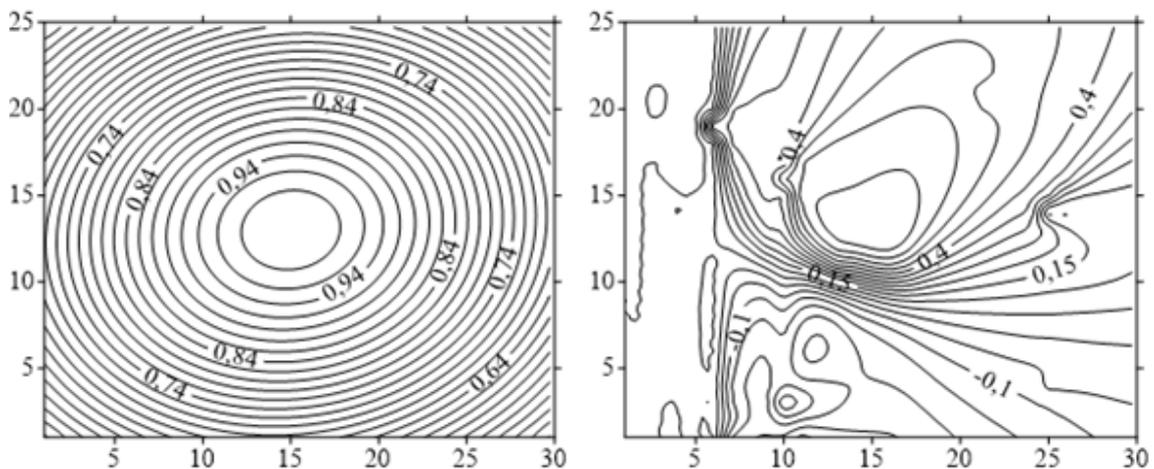


Рис 1.1. Теоретические (слева) и эмпирические (справа) изолинии.

В связи с вышеуказанными недостатками, для различных классов корреляционных функций разрабатывают специальные более экономичные и устойчивые алгоритмы моделирования [57, 66, 95, 175].

Для корреляционных функций вида $r(\rho) = \exp(-\alpha\rho^2)$ и $r(x, y) = \exp(-(\alpha x^2 + \gamma y^2))$ существует эффективный алгоритм моделирования гауссовских полей на регулярных сетках. В литературе его называют алгоритмом «по строкам и столбцам» [23, 27, 175]. При его реализации корреляционные матрицы задаются отдельно на горизонтальных и вертикальных сечениях поля. В основе этого алгоритма лежит тот факт, что корреляционная матрица поля представима в виде прямого произведения корреляционных матриц его горизонтального и вертикального сечений. Преимуществом данного метода является то, что он сводится к последовательному использованию алгоритмов моделирования гауссовских векторов с теплицевыми корреляционными матрицами, заданными функциями вида $r(\rho) = \exp(-\alpha\rho^2)$. Для них алгоритмы, основанные на методе условных распределений, реализуются с высокой точностью.

Алгоритм моделирования «по строкам и столбцам» имеет специфику, которую можно проиллюстрировать следующим примером. Предположим, что мы моделируем однородное гауссовское поле на равномерной прямоугольной сетке размера $m \times n$ узлов (для простоты используем единичный шаг сетки). Элементы корреляционных матриц горизонтального $R^{(1)} = (r_{ij}^{(1)})$, $i, j = 1, 2, \dots, n$ и вертикального $R^{(2)} = (r_{kp}^{(2)})$, $k, p = 1, 2, \dots, m$ сечений поля задаются функциями $r_{ij}^{(1)} = \exp(-\alpha_1|i-j|^2)$ и $r_{kp}^{(2)} = \exp(-\alpha_2|k-p|^2)$, соответственно. После реализации алгоритма «по строкам и столбцам» корреляционная функция построенного однородного случайного поля имеет вид

$$r(i, j, k, p) = \exp(-\alpha_1(i-j)^2 - \alpha_2(k-p)^2).$$

При этом, сечение горизонтальной плоскостью поверхности корреляционной функции между значением поля в произвольном фиксированном узле и значениями во всех остальных узлах будет представлять собой эллипс с главными осями, ориентированными вдоль осей Ox, Oy .

Однако, в прикладных задачах достаточно часто встречаются корреляционные функции случайных полей, изолинии которых являются эллипсами с главными осями, повернутыми относительно положительных направлений осей Ox, Oy на некоторый угол φ . Например, для корреляционной функции поля геопотенциала на уровне 500 мбар угол поворота $\varphi \approx 45^\circ$ [11].

В этом параграфе приведены 2 модификации алгоритма «по строкам и столбцам», предложенные в работах [6, 101, 150]. Первая их них позволяет моделировать поля, для которых угол поворота $\varphi \neq 0^\circ$. Вторая модификация метода «по строкам и столбцам» значительно расширяет класс корреляционных функций, для которых он применим.

1.1.1. Модифицированный нерандомизированный алгоритм «по строкам и столбцам»

Алгоритм моделирования

Сформулируем алгоритм моделирования двумерного гауссовского однородного случайного поля $\{\eta_j^i\}$, $i = 1, \dots, m$, $j = 1, \dots, n$ на равномерной прямоугольной сетке размерности $m \times n$ узлов (для удобства шаг сетки равен 1).

Алгоритм 1.1.

Шаг 1. Моделируем двумерное случайное поле $\{\varphi_j^i\}$, как массив стандартных гауссовских независимых случайных величин:

$$\begin{pmatrix} \varphi_1^1 & \cdots & \varphi_n^1 \\ \vdots & \ddots & \vdots \\ \varphi_1^m & \cdots & \varphi_n^m \end{pmatrix},$$

где $\vec{\varphi}^{(i)T} \sim N(0, I)$, $\vec{\varphi}^{(i)T} = (\varphi_1^i, \varphi_2^i, \dots, \varphi_n^i)$ – вектор-строка, $i = 1, \dots, m$.

Шаг 2. Преобразованием

$$\vec{\xi}^{(i)T} = \vec{\varphi}^{(i)T} L^{(1)T}$$

для каждого $i = 1, \dots, m$ получаем m независимых векторов-строк

$$\vec{\xi}^{(i)T} = (\xi_1^i, \xi_2^i, \dots, \xi_n^i).$$

Здесь

$$L^{(1)} = \begin{pmatrix} l_{1,1}^{(1)} & 0 & \cdots & 0 \\ l_{2,1}^{(1)} & l_{2,2}^{(1)} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ l_{n,1}^{(1)} & l_{n,2}^{(1)} & \cdots & l_{n,n}^{(1)} \end{pmatrix}$$

есть нижнетреугольная матрица такая, что $L^{(1)} L^{(1)T} = R^{(1)}$, а матрица

$$R^{(1)} = \begin{pmatrix} r_{1,1}^{(1)} & \cdots & r_{1,n}^{(1)} \\ \vdots & \ddots & \vdots \\ r_{n,1}^{(1)} & \cdots & r_{n,n}^{(1)} \end{pmatrix} -$$

это заданная корреляционная матрица горизонтального сечения поля. Из векторов-строк $\vec{\xi}^{(i)T}$ формируем поле $\{\xi_j^i\}$, $i = 1, \dots, m$, $j = 1, \dots, n$

$$\{\xi_j^i\} = \begin{pmatrix} \xi_1^1 & \cdots & \xi_n^1 \\ \vdots & \ddots & \vdots \\ \xi_1^m & \cdots & \xi_n^m \end{pmatrix},$$

векторы-строки $\vec{\xi}^{(i)T} \sim N(0, R^{(1)})$ которого независимы между собой.

Шаг 3. Составим p векторов $\vec{\zeta}^{(k)}$ из случайных величин ξ_j^i , «лежащих» на прямых определенного наклона относительно горизонтального направления сетки, т.е.

$$\vec{\zeta}_{(k)} = \left(\xi_k^1, \xi_{k+d}^2, \xi_{k+2d}^3, \dots, \xi_{k+(m-1)d}^m \right)^T,$$

где $k = 1, \dots, n - (m-1)d$, $d = 0, 1, 2, \dots$ – параметр сдвига на каждом горизонтальном слое сетки. Для иллюстрации этого способа построения векторов $\vec{\zeta}_{(k)}$ приведен рисунок 1.2. Так как векторы-строки $\vec{\xi}^{(i)T} = (\xi_1^i, \xi_2^i, \dots, \xi_n^i)$ поля $\{\xi_j^i\}$ независимы между собой, то векторы $\vec{\zeta}_{(k)} \sim N(0, I)$.

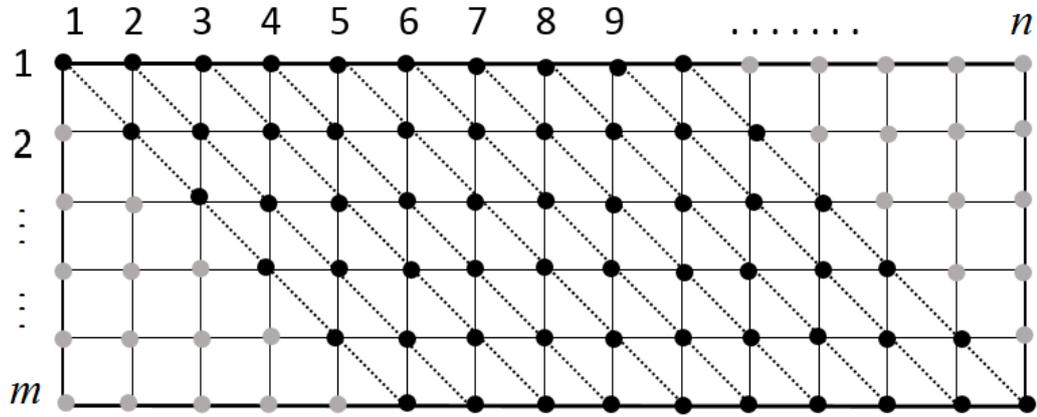


Рис. 1.2. Схема построения векторов $\vec{\zeta}_{(k)}$ на Шаге 3 модифицированного алгоритма «по строкам и столбцам» при $d = 1$. Черным цветом обозначены случайные величины, из которых составляются векторы $\vec{\zeta}_{(k)}$, $k = 1, \dots, n - (m-1)d$.

Отличие Шага 3 в модифицированном алгоритме от стандартного алгоритма «по строкам и столбцам» состоит в том, что при реализации последнего после преобразований на втором шаге случайные величины связываются между собой по узлам сетки, находящимся на её вертикальных слоях.

Шаг 4. Заключительным преобразованием

$$\vec{v}_{(k)} = L^{(2)} \vec{\zeta}_{(k)},$$

получаем векторы $\vec{v}_{(k)} = (v_k^1, v_{k+d}^2, v_{k+2d}^3, \dots, v_{k+(m-1)d}^m)^T$, $k = 1, \dots, n - (m-1)d$. Здесь матрица

$$L^{(2)} = \begin{pmatrix} l_{1,1}^{(2)} & 0 & \dots & 0 \\ l_{2,1}^{(2)} & l_{2,2}^{(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ l_{n,1}^{(2)} & l_{n,2}^{(2)} & \dots & l_{n,n}^{(2)} \end{pmatrix} -$$

нижнетреугольная матрица такая, что $L^{(2)}L^{(2)T} = R^{(2)}$, а матрица

$$R^{(2)} = \begin{pmatrix} r_{1,1}^{(1)} & \dots & r_{1,m}^{(1)} \\ \vdots & \ddots & \vdots \\ r_{m,1}^{(1)} & \dots & r_{m,m}^{(1)} \end{pmatrix} -$$

заданная корреляционная матрица вертикального сечения поля. Из построенных векторов формируем случайное поле $\{v_j^i\}$, $i = 1, \dots, m$, $j = 1, \dots, n$. При $d = 1$ поле

имеет вид:

$$\begin{pmatrix} v_1^1 & v_2^1 & \dots & v_{m-1}^1 & v_m^1 & \dots & v_{n-(m-1)}^1 & \xi_{n-(m-1)+1}^1 & \dots & \xi_n^1 \\ \xi_1^2 & v_2^2 & \dots & v_{m-1}^2 & v_m^2 & \dots & v_{n-(m-1)}^2 & v_{n-(m-1)+1}^2 & \dots & \xi_n^2 \\ \dots & \dots \\ \xi_1^{m-1} & \xi_2^{m-1} & \dots & v_{m-1}^{m-1} & v_m^{m-1} & \dots & v_{n-(m-1)}^{m-1} & v_{n-(m-1)+1}^{m-1} & \dots & \xi_n^{m-1} \\ \xi_1^m & \xi_2^m & \dots & \xi_{m-1}^m & v_m^m & \dots & v_{n-(m-1)}^m & v_{n-(m-1)+1}^m & \dots & v_n^m \end{pmatrix}.$$

Так как в рассматриваемом случае для моделирования используется прямоугольная сетка, то после «отбрасывания» по $d(m-1)$ векторов-столбцов с боковых сторон данного массива случайных величин получаем итоговое случайное поле $\{\eta_j^i\}$,

$i = 1, \dots, m$, $j = 1, \dots, n - 2d(m-1)$:

$$\{\eta_j^i\} = \begin{pmatrix} v_m^1 & v_{m+1}^1 & \dots & v_{n-(m-1)}^1 \\ v_m^2 & v_{m+1}^2 & \dots & v_{n-(m-1)}^2 \\ \dots & \dots & \dots & \dots \\ v_m^m & v_{m+1}^m & \dots & v_{n-(m-1)}^m \end{pmatrix} = \begin{pmatrix} \eta_1^1 & \eta_2^1 & \dots & \eta_{n-2(m-1)}^1 \\ \eta_1^2 & \eta_2^2 & \dots & \eta_{n-2(m-1)}^2 \\ \dots & \dots & \dots & \dots \\ \eta_1^m & \eta_2^m & \dots & \eta_{n-2(m-1)}^m \end{pmatrix}.$$

Следует отметить, что в данном случае был рассмотрен угол θ наклона прямых, по которым связываются случайные величины, принадлежащий

интервалу $[-\pi/4, 0)$. В общем случае $\theta \in [-\pi/4, \pi/4] / \{0\}$, и в силу дискретности сетки, угол наклона прямых θ изменяется тоже дискретно. При этом

$$\operatorname{tg}\theta = \begin{cases} -1/d, & \theta \in [-\pi/4, 0), \\ 1/d, & \theta \in (0, \pi/4], \end{cases}$$

где d – целочисленный параметр сдвига. В случае, когда $\theta \in (0, \pi/4]$, на Шаге 3 векторы $\vec{\zeta}_{(k)}$, $k = 1, \dots, n - (m-1)d$ формируются из случайных величин ξ_j^i , $i = 1, \dots, m$, $j = 1, \dots, n$ следующим образом:

$$\vec{\zeta}_{(k)} = \left(\xi_{k+(m-1)d}^1, \dots, \xi_{k+2d}^{m-2}, \xi_{k+d}^{m-1}, \xi_k^m \right)^T.$$

Все дальнейшие рассуждения приведены для случая $\theta \in [-\pi/4, 0)$, если не оговорено иное.

Отметим, что от значения параметра сдвига d при фиксированной размерности сетки ($m \times n$ узлов) зависит количество p векторов $\vec{\zeta}_{(k)}$, $k = 1, \dots, p$, а именно, $p = n - (m-1)d$. Это, в свою очередь, приводит к уменьшению размерности итогового поля $\{\eta_j^i\}$, $i = 1, \dots, m$, $j = 1, \dots, n - 2d(m-1)$. Если необходимо построить ровно p векторов $\vec{\zeta}_{(k)}$ при заданном сдвиге d и фиксированном числе узлов m по вертикальному сечению сетки, то необходимое число узлов сетки по горизонтали вычисляется по формуле $n = p + d(m-1)$.

Замечание 1.1. В приложениях бывает удобно моделировать случайные поля не на прямоугольной сетке, а, например, на сетке в форме параллелограмма. В этих случаях можно не «отбрасывать» по $d(m-1)$ векторов-столбцов массива, а использовать получившееся поле нужной формы.

Корреляционная структура поля

Исследуем корреляционную структуру построенного с помощью вышеописанного модифицированного алгоритма «по строкам и столбцам»

гауссовского случайного поля $\{\eta_j^i\}$, $i = 1, \dots, m$, $j = 1, \dots, n - 2d(m - 1)$. Элементы корреляционных матриц $R^{(1)} = (r^{(1)}(i, j))$ и $R^{(2)} = (r^{(2)}(k, p))$ вычисляются следующим образом:

$$r^{(1)}(i, j) = \exp(-\alpha_1 |i - j|^2), \quad i, j = 1, \dots, n,$$

$$r^{(2)}(k, p) = \exp(-\alpha_2 |k - p|^2), \quad k, p = 1, \dots, m.$$

Здесь α_1 и α_2 – произвольные параметры. Поскольку случайные величины η_j^i имеют нулевое среднее и единичную дисперсию, то коэффициенты корреляции между ними имеют вид

$$E\eta_i^k \eta_j^p = r(i, j, k, p) = r^{(2)}(k, p)r^{(1)}(i - d(k - 1), j - d(p - 1)),$$

где $i, j = 1, \dots, n - 2d(m - 1)$, $k, p = 1, \dots, m$. С учетом того, что шаг сетки равен 1, перепишем формулу, определяющую корреляционную структуру поля $\{\eta_j^i\}$ в декартовых координатах Oxy :

$$E\eta_i^k \eta_j^p = \exp(-\alpha_2 |k - p|^2) \exp(-\alpha_1 |i - d(k - 1) - j + d(p - 1)|^2) =$$

$$= \exp(-\alpha_2 y^2) \exp(-\alpha_1 |x - d y|^2) = \exp(-\alpha_1 x^2 + 2d\alpha_1 xy - (d^2\alpha_1 + \alpha_2)y^2).$$

Таким образом, нами доказано следующее утверждение.

Утверждение 1.1. Корреляционная функция построенного поля $\{\eta_j^i\}$ имеет вид

$$r(x, y) = \exp(-\alpha_1 x^2 + 2d\alpha_1 xy - (d^2\alpha_1 + \alpha_2)y^2).$$

Изолинии данной корреляционной функции описываются уравнениями

$$f(x, y) = -\alpha_1 x^2 + 2d\alpha_1 xy - (d^2\alpha_1 + \alpha_2)y^2 + C = 0$$

и представляют собой кривые второго порядка, а именно, эллипсы, главные оси которых повернуты на некоторый угол φ относительно координатных осей Ox, Oy . Соответствующая иллюстрация приведена на рисунке 1.3.

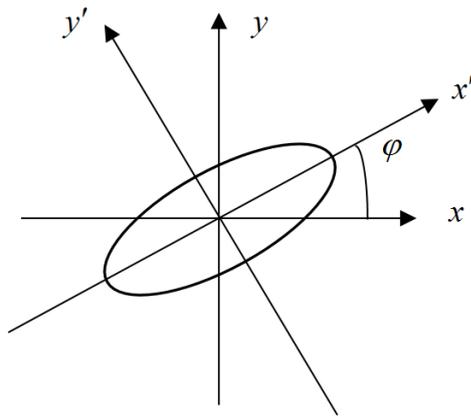


Рис. 1.3. Пример изолинии $f(x, y)$.

Угол поворота φ вычисляется по формуле

$$\operatorname{tg} 2\varphi = \frac{2d\alpha_1}{\alpha_2 + \alpha_1(d^2 - 1)}.$$

Из последней формулы следует, что изменяя коэффициенты α_1 и α_2 при задании корреляционных матриц $R^{(1)}$ и $R^{(2)}$, соответственно, можно контролировать угол поворота φ изолиний корреляционной функции поля.

Замечание 1.2. Если величина $\theta \in (0, \pi/4]$, то коэффициенты корреляции между случайными величинами поля $\{\eta_j^i\}$ будут равны

$$E\eta_i^k \eta_j^p = r(i, j, k, p) = r^{(2)}(k, p)r^{(1)}(i + d(k - 1), j + d(p - 1)),$$

а уравнение изолиний примет вид

$$f(x, y) = -\alpha_1 x^2 - 2d\alpha_1 xy - y^2(\alpha_2 + d^2\alpha_1) + C = 0.$$

При этом угол поворота φ может быть найден из уравнения

$$\operatorname{tg} 2\varphi = \frac{-2d\alpha_1}{\alpha_2 + \alpha_1(d^2 - 1)}.$$

Замечание 1.3. Следует отметить, что существуют некоторые ограничения на выбор параметров α_1 , α_2 , и угол поворота φ . Во-первых, $\alpha_1, \alpha_2 > 0$. Во-вторых, должно выполняться условие

$$\operatorname{tg} 2\varphi \in \left(0, \frac{2d}{d^2 - 1}\right),$$

т. е. пределы изменения угла поворота φ связаны с параметром сдвига d . Кроме того, существует связь между скоростями убывания корреляционной функции поля по координатным осям. Действительно, при $y = 0$ корреляционная функция поля $\{\eta_j^i\}$ имеет вид: $r(x) = \exp(-\alpha_1 x^2)$, то есть скорость убывания функции по оси Ox зависит только от параметра α_1 . При $x = 0$ корреляционная функция описывается выражением $r(y) = \exp(-(d^2 \alpha_1 + \alpha_2) y^2)$, и скорость убывания корреляционной функции по оси Oy зависит от суммы $(d^2 \alpha_1 + \alpha_2)$. Таким образом, при изменении параметра α_1 будет изменяться скорость убывания корреляционной функции не только по оси Ox , но и по оси Oy .

Очевидно, что справедлива следующая лемма.

Лемма 1.1. В главных осях $Ox'y'$ уравнение изолиний корреляционной функции поля имеет вид

$$f(x', y') = Ax'^2 + By'^2 + C = 0,$$

где

$$A = -\alpha_1 (d^2 \sin^2 \varphi - d \sin 2\varphi + \cos^2 \varphi) - \alpha_2 \sin^2 \varphi,$$

$$B = -\alpha_1 (d^2 \cos^2 \varphi + d \sin 2\varphi + \sin^2 \varphi) - \alpha_2 \cos^2 \varphi.$$

Корреляционную функцию поля в новой системе координат $Ox'y'$ можно переписать в виде

$$r(x', y') = \exp(Ax'^2 + By'^2).$$

При $y' = 0$ корреляционная функция описывается уравнением $r(x') = \exp(Ax'^2)$, а при $x' = 0$ – уравнением $r(y') = \exp(By'^2)$ соответственно. Кроме этого, при $d = 1$ на параметры A и B накладываются ограничения в виде неравенств

$$0 < A < \alpha_1 + \frac{\alpha_2}{2}, \quad 1 < B < 2\alpha_1 + \frac{\alpha_2}{2}.$$

Очевидным является тот факт, что график функции, получающийся при сечении поверхности $z = r(x, y)$ произвольной плоскостью вида $ax + by = 0$, обладает точками перегиба, причём все точки перегиба (независимо от выбора секущей плоскости $ax + by = 0$) принадлежат плоскости $z = \exp(-0.5)$.

На рисунке 1.4 приведены примеры изолиний функции $z = r(x, y)$ при 2 наборах параметров $\alpha_1, \alpha_2, \varphi, d$.

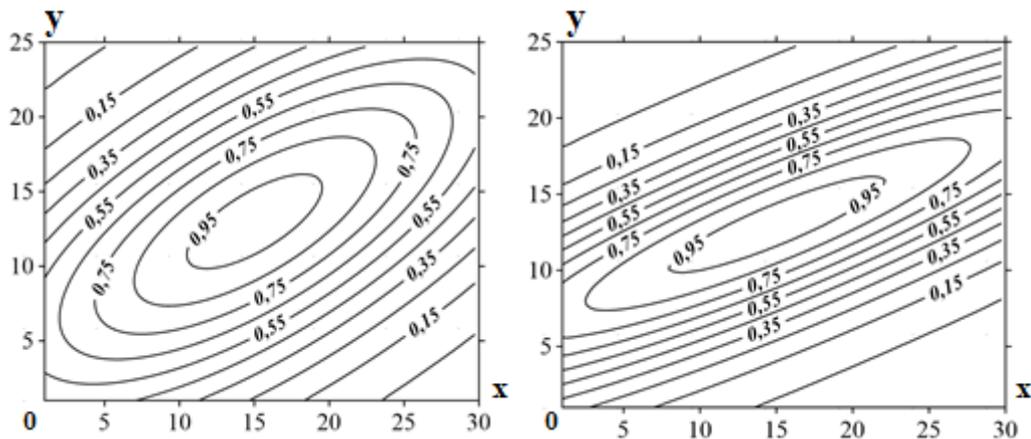


Рис. 1.4. Изолинии корреляционной функции $z = r(x, y)$ Слева – при $\alpha_1 = \alpha_2 = 0.005, \varphi = 31.7, d = 1$, справа – при $\alpha_1 = \alpha_2 = 0.005, \varphi = 22.5, d = 2$.

1.1.2. Рандомизированный алгоритм «по строкам и столбцам»

Рассмотрим ещё одну модификацию алгоритма «по строкам и столбцам». По аналогии с [52, 57], рандомизируя значения одного из параметров α_1 или α_2 , можно получить достаточно широкий класс корреляционных функций, для которых применим приведенный выше алгоритм моделирования.

Пусть угол поворота главных осей эллипса зафиксирован, т.е.

$$\operatorname{tg} 2\varphi = c,$$

где $c \in \left(0, \frac{2d}{d^2 - 1}\right)$ – некоторая константа. В таком случае можно выразить параметр α_2 через α_1 :

$$\alpha_2 = \left(\frac{2d}{c} - d^2 + 1\right)\alpha_1.$$

Подставляя полученное выражение для α_2 в формулу, описывающую корреляционную функцию поля, имеем:

$$r(x, y) = \exp\left(-\alpha_1\left((x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2\right)\right) = \exp(-D\alpha_1),$$

где $D = D(x, y) = (x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2 \geq 0$ (равенство $D = 0$ выполняется только если $x = y = 0$).

Пусть теперь α_1 – случайная величина с плотностью одномерного распределения $f(\lambda)$, $\lambda \in (a, b)$, $0 < a < b$. Будем считать, что $f(\lambda) > 0$, $\lambda \in (a, b)$. Рассмотрим несколько примеров корреляционных функций вида

$$R(x, y) = \int_a^b f(\lambda) \exp(-D\lambda) d\lambda.$$

Все примеры будут приведены для сдвига $d = 1$.

Пример 1.1. Пусть $f(\lambda)$ – плотность равномерного распределения на интервале (a, b) . Тогда

$$\begin{aligned} R(x, y) &= \int_a^b \frac{1}{b-a} \exp(-D\lambda) d\lambda = \frac{\exp(-aD) - \exp(-bD)}{(b-a)D} = \\ &= \frac{\exp\left(-a\left((x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2\right)\right) - \exp\left(-b\left((x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2\right)\right)}{(b-a)\left((x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2\right)}, \\ &\quad (x, y) \in \mathbb{R}^2 \setminus (0, 0), \\ &\quad R(0, 0) = 1. \end{aligned}$$

Если $a = 0$, $b = 1$, то

$$R(x, y) = \frac{1 - \exp\left(-\left((x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2\right)\right)}{\left((x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2\right)}, \quad (x, y) \in \mathbb{R}^2 \setminus (0, 0).$$

Пример 1.2. Если $f(\lambda)$ – плотность экспоненциального распределения с параметром $\gamma > 0$, то

$$R(x, y) = \frac{\gamma}{\gamma + D} = \frac{\gamma}{\gamma + (x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2}, \quad (x, y) \in \mathbb{R}^2.$$

Пример 1.3. Если $f(\lambda)$ – плотность треугольного распределения (распределения Симпсона), т.е.

$$f(\lambda) = \begin{cases} \frac{4(\lambda - a)}{(b - a)^2}, & \lambda \in \left(a, \frac{a + b}{2}\right], \\ \frac{4(b - \lambda)}{(b - a)^2}, & \lambda \in \left(\frac{a + b}{2}, b\right), \end{cases}$$

то

$$R(x, y) = \frac{4}{(b - a)^2} \frac{\left(\exp(-aD) - 2\exp\left(-\frac{a + b}{2}D\right) + \exp(-bD)\right)}{D^2},$$

$$(x, y) \in \mathbb{R}^2 \setminus (0, 0),$$

$$R(0, 0) = 1.$$

На рисунке 1.5 приведены графики сечений поверхности $z = R(x, y)$ плоскостью $y = x$ при $x > 0$ при различных плотностях $f(\lambda)$. Следует отметить, что, изменяя плотности $f(\lambda)$, можно получать корреляционные функции случайного поля с различным характером их убывания. Приведенные на рисунке 1.5 функции обладают общим свойством: у всех них есть точки перегиба.

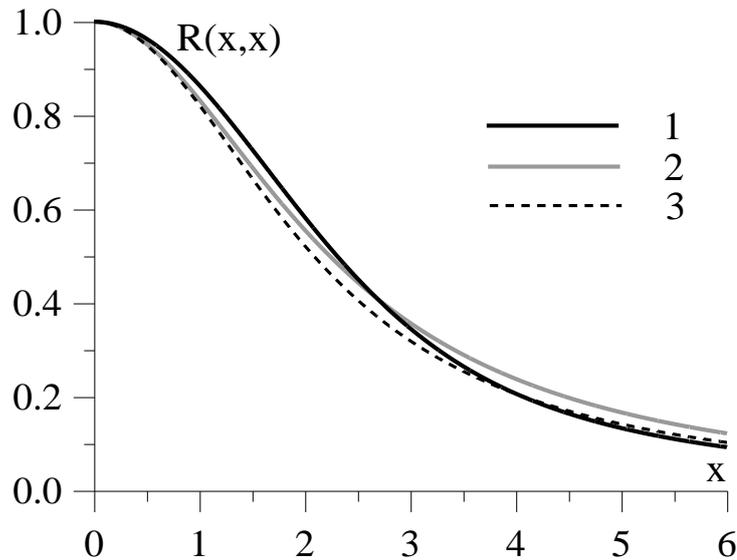


Рис. 1.5. Графики сечений поверхности $z = R(x, y)$ плоскостью $y = x$ при $x > 0$.

Кривая 1: при $f(\lambda)$ – плотности равномерного распределения на интервале $(0, 0.3)$, кривая 2: при $f(\lambda)$ – плотности экспоненциального распределения с параметром $\gamma = 5$, кривая 3: при $f(\lambda)$ – плотности экспоненциального распределения с параметром $\gamma = 3.7$, усеченного на отрезок $(0, 0.7)$.

Утверждение 1.2. Если $0 < a < b < +\infty$, плотность распределения $f(\lambda)$ непрерывна на $[a, b]$ и $f(\lambda) > 0$, $\lambda \in [a, b]$, то график функции, являющейся сечением плоскостью $Ax + By = 0$ поверхности

$$z = R(x, y) = \int_a^b f(\lambda) \exp(-D\lambda) d\lambda,$$

где $D = D(x, y) = (x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2$, имеет точки перегиба.

Доказательство. Рассмотрим сечение поверхности $z = R(x, y)$ плоскостью $x = 0$. В плоскости Ozy получим кривую

$$z = \int_a^b f(\lambda) e^{-my^2\lambda} d\lambda = r(y),$$

где $m = \frac{2d}{c} + 1$ и $m > 0$, т.к. $d > 0, c > 0$. Для того чтобы доказать, что кривая

$z = r(y)$ имеет точки перегиба, необходимо показать, что существуют $\frac{dz}{dy}, \frac{d^2z}{dy^2}$ и

функция $\frac{d^2z}{dy^2}$ меняет знак.

Вычисление частных производных

$$\frac{dz}{dy} = \frac{d \left(\int_a^b f(\lambda) e^{-my^2 \lambda} d\lambda \right)}{dy} = -2m \int_a^b \lambda f(\lambda) y e^{-my^2 \lambda} d\lambda,$$

$$\frac{d^2z}{dy^2} = 2m \int_a^b (2m\lambda^2 y^2 - \lambda) f(\lambda) e^{-my^2 \lambda} d\lambda.$$

осуществляется по правилу Лейбница дифференцирования интеграла по параметру [88].

Осталось доказать, что функция $\frac{d^2z}{dy^2}$ непрерывна в $[h, q]$ и меняет знак.

Согласно [88], функция $\frac{d^2z}{dy^2}$ непрерывна по y , если подынтегральная функция

$(2m\lambda^2 y^2 - \lambda) f(\lambda) e^{-my^2 \lambda}$ определена и непрерывна как функция двух переменных λ и y в прямоугольнике $[a, b] \times [h, q]$.

Очевидно, что если $2m\lambda^2 y^2 - \lambda < 0$ для всех $\lambda \in [a, b]$, то $\frac{d^2z}{dy^2} < 0$, и если

$2m\lambda^2 y^2 - \lambda > 0$ для всех $\lambda \in [a, b]$, то $\frac{d^2z}{dy^2} > 0$. Поскольку $\lambda > 0, m > 0$, верна

следующая цепочка:

$$2m\lambda^2 y^2 - \lambda < 0 \Leftrightarrow 2m\lambda y^2 - 1 < 0 \Leftrightarrow y^2 < \frac{1}{\sqrt{2m\lambda}}.$$

Пусть $y \geq 0$. При $\lambda \in [a, b]$ выполняются неравенства

$$\frac{1}{\sqrt{2mb}} < \frac{1}{\sqrt{2m\lambda}} < \frac{1}{\sqrt{2ma}}$$

и очевидно, что

$$\forall b > 0 \exists M_1 < \frac{1}{\sqrt{2mb}} : 2m\lambda^2 y^2 - \lambda|_{y=M_1} < 0 \quad \forall \lambda \in [a, b],$$

$$\forall a > 0 \exists M_2 > \frac{1}{\sqrt{2ma}} : 2m\lambda^2 y^2 - \lambda|_{y=M_2} > 0 \quad \forall \lambda \in [a, b].$$

Все рассматриваемые функции определены для $y \in [h, q]$, где $h < \frac{1}{\sqrt{2mb}}$,

$q > \frac{1}{\sqrt{2ma}}$, следовательно, $\frac{d^2z}{dy^2}$ меняет знак, и у функции $z = r(y)$ существует

точка перегиба при $y > 0$.

Аналогичным образом можно показать, что у функции $z = r(y)$ существует точка перегиба при $y < 0$.

Доказательство существования точки перегиба функции $z = \tilde{r}(x)$ при сечении поверхности $z = R(x, y)$ плоскостью $y = 0$ получается из предыдущего заменой y на x при $m = 1$.

Осталось рассмотреть случай сечения поверхности $z = R(x, y)$ плоскостью $y = kx$, $k \neq 0$. Введём новую систему координат $Ox'y'z'$ такую, что ось Oz' совпадает с исходной осью Oz и

$$x' = x \cos \beta + y \sin \beta,$$

$$y' = -x \sin \beta + y \cos \beta,$$

где $\cos \beta = \frac{1}{\sqrt{1+k^2}}$, $\sin \beta = \frac{k}{\sqrt{1+k^2}}$. Фактически, мы осуществили поворот осей

координат Oxy на угол β против часовой стрелки. При этом,

$$x = \frac{1}{\sqrt{1+k^2}} x' - \frac{k}{\sqrt{1+k^2}} y', \quad y = \frac{k}{\sqrt{1+k^2}} x' + \frac{1}{\sqrt{1+k^2}} y'.$$

В новой системе координат плоскость $y = kx$ описывается уравнением $y' = 0$, а поверхность $z = R(x, y)$ – уравнением

$$z = R'(x', y') = \int_a^b f(\lambda) e^{-D'\lambda} d\lambda,$$

где

$$\begin{aligned} D'(x', y') &= D(x, y) \Big|_{\substack{x=x' \\ y=y'}} = \left[(x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1 \right) y^2 \right]_{\substack{x=x' \\ y=y'}} = \\ &= \left[\frac{1}{\sqrt{1+k^2}} x' - \frac{k}{\sqrt{1+k^2}} y' - \frac{dk}{\sqrt{1+k^2}} x' - \frac{d}{\sqrt{1+k^2}} y' \right]^2 + \\ &\quad + \left(\frac{2d}{c} - d^2 + 1 \right) \left[\frac{k}{\sqrt{1+k^2}} x' + \frac{1}{\sqrt{1+k^2}} y' \right]^2, \\ D'(x', 0) &= (x')^2 \left[\left(\frac{1-dk}{\sqrt{1+k^2}} \right)^2 + \left(\frac{2d}{c} - d^2 + 1 \right) \frac{k^2}{1+k^2} \right] = M'(x')^2 \end{aligned}$$

и $M' \neq 0$ т.к. $c \in \left(0, \frac{2d}{d^2 - 1} \right)$. Доказательство существования точки перегиба проводится аналогично приведённому выше.

Утверждение доказано.

Следует отметить, что рандомизированный алгоритм, в отличие от стандартного алгоритма «по строкам и столбцам», является приближенным, в том смысле, что одномерные распределения построенного поля не являются стандартными нормальными. Одномерные распределения смоделированного поля зависят от выбора плотности распределения рандомизированного параметра и от координат точки поля, в которой определяется его распределение. В качестве иллюстрации приведем рисунке 1.6, на котором изображены плотность $\phi_{0,1}(x)$ стандартного нормального распределения, а также оценки плотностей $\psi(x)$ одномерных распределений в двух точках смоделированного поля. Оценки плотностей проводились по 2×10^7 реализаций.

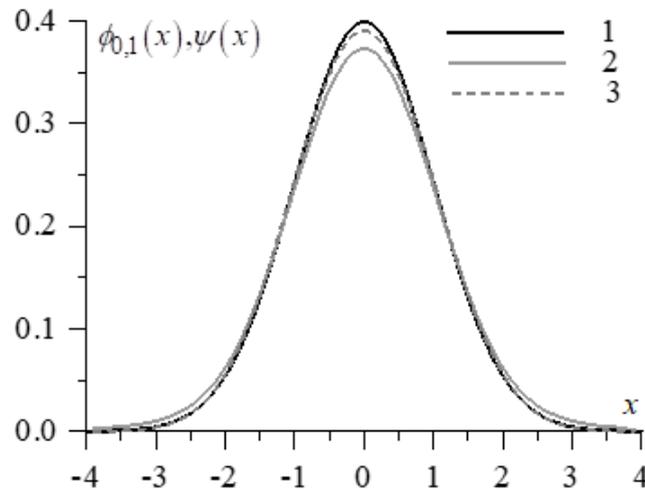


Рис. 1.6. Графики одномерных распределений. Кривая 1 – плотность $\phi_{0,1}(x)$ стандартного нормального распределения, кривые 2, 3 – эмпирические плотности $\psi(x)$, оцененные в 2 различных точках поля, при $f(\lambda) = 2e^{-2\lambda}$.

1.2. Стохастическая интерполяция случайного процесса

Интерполяция случайного процесса

Пусть $\dots, \xi_{-2}, \xi_{-1}, \xi_0, \xi_1, \xi_2, \xi_3 \dots$ – стационарный в широком смысле случайный процесс с плотностью одномерного распределения $f(x)$, математическим ожиданием μ , дисперсией $0 < \sigma^2 < \infty$ и корреляционной функцией $r_{|i-j|}$, $i, j \in \mathbb{Z}$.

Рассмотрим случайный процесс η_t , $t \in \mathbb{R}$, определяемый по правилу:

$$\eta_t = \xi_j \quad (1.1)$$

с вероятностью ω_j^t при $t = i + h$, $i \in \mathbb{Z}$, $0 < h < 1$, где j – одна из m ближайших

к t целочисленных точек, $0 < \omega_j^t < 1$, $\sum_{j=1}^m \omega_j^t = 1$ и

$$\eta_t = \xi_i \quad (1.2)$$

при $t = i$, $i \in \mathbb{Z}$. Очевидно, что справедливо следующая лемма

Лемма 1.2. Процесс $\eta_t, t \in \mathbb{R}$ имеет плотность одномерного распределения $f(x)$, математическое ожидание μ и дисперсию σ^2 .

Корреляционная функция процесса

Определим корреляционную функцию $\text{corr}(\eta_t, \eta_p)$ рассматриваемого процесса.

Рассмотрим случай $m = 2k, k \in \mathbb{Z}^+$ (т.е. случай, когда стохастическая интерполяция в точку t производится из чётного числа целочисленных точек). В этом случае, m ближайшими к t и p целочисленными точками будут точки $i - k + 1, \dots, i - 1, i, i + 1, \dots, i + k$ и $n - k + 1, \dots, n - 1, n, n + 1, \dots, n + k$, соответственно.

Лемма 1.3. В условиях данного параграфа справедливы равенства

1. при $t = i, p = n, i, n \in \mathbb{Z}$

$$\text{corr}(\eta_t, \eta_p) = r_{|i-n|};$$

2. при $t = i + h, p = n, i, n \in \mathbb{Z}, 0 < h < 1$

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k} \omega_j^t r_{|j-n|};$$

3. при $t = i, p = n + g, i, n \in \mathbb{Z}, 0 < g < 1$

$$\text{corr}(\eta_t, \eta_p) = \sum_{l=n-k+1}^{n+k} \omega_l^p r_{|i-l|};$$

4. при $t = i + h, p = n + g, i, n \in \mathbb{Z}, 0 < h, g < 1$

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k} \sum_{l=n-k+1}^{n+k} \omega_j^t \omega_l^p r_{|j-l|}.$$

Доказательство. Действительно,

1. если $t = i$, $p = n$, $i, n \in \mathbb{Z}$, то $\text{corr}(\eta_t, \eta_p) = r_{|i-n|}$ по построению процесса $\eta_t, t \in \mathbb{R}$;

2. если $t = i + h$, $p = n$, $i, n \in \mathbb{Z}$, $0 < h < 1$, то

$$\begin{aligned} E\eta_t\eta_p &= \sum_{j=i-k+1}^{i+k} \omega_j^t E\xi_j\xi_n = \sum_{j=i-k+1}^{i+k} \omega_j^t \left(r_{|j-n|} \sigma^2 + \mu^2 \right) = \\ &= \sigma^2 \sum_{j=i-k+1}^{i+k} \omega_j^t r_{|j-n|} + \mu^2, \\ \text{corr}(\eta_t, \eta_p) &= \frac{E\eta_t\eta_p - E\eta_t E\eta_p}{\sqrt{D\eta_t} \sqrt{D\eta_p}} = \sum_{j=i-k+1}^{i+k} \omega_j^t r_{|j-n|}; \end{aligned}$$

3. если $t = i$, $p = n + g$, $i, n \in \mathbb{Z}$, $0 < g < 1$, то верны равенства

$$\begin{aligned} E\eta_t\eta_p &= \sum_{l=n-k+1}^{n+k} \omega_l^p E\xi_i\xi_l = \sum_{l=n-k+1}^{n+k} \omega_l^p \left(r_{|i-l|} \sigma^2 + \mu^2 \right) = \\ &= \sigma^2 \sum_{l=n-k+1}^{n+k} \omega_l^p r_{|i-l|} + \mu^2, \\ \text{corr}(\eta_t, \eta_p) &= \frac{E\eta_t\eta_p - E\eta_t E\eta_p}{\sqrt{D\eta_t} \sqrt{D\eta_p}} = \sum_{l=n-k+1}^{n+k} \omega_l^p r_{|i-l|}; \end{aligned}$$

4. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $0 < h, g < 1$

$$\begin{aligned} E\eta_t\eta_p &= \sum_{j=i-k+1}^{i+k} \sum_{l=n-k+1}^{n+k} \omega_j^t \omega_l^p E\xi_j\xi_l = \sum_{j=i-k+1}^{i+k} \sum_{l=n-k+1}^{n+k} \omega_j^t \omega_l^p \left(r_{|j-l|} \sigma^2 + \mu^2 \right) = \\ &= \sigma^2 \sum_{j=i-k+1}^{i+k} \sum_{l=n-k+1}^{n+k} \omega_j^t \omega_l^p r_{|j-l|} + \mu^2 \sum_{j=i-k+1}^{i+k} \sum_{l=n-k+1}^{n+k} \omega_j^t \omega_l^p = \\ &= \sigma^2 \sum_{j=i-k+1}^{i+k} \sum_{l=n-k+1}^{n+k} \omega_j^t \omega_l^p r_{|j-l|} + \mu^2, \\ \text{corr}(\eta_t, \eta_p) &= \frac{E\eta_t\eta_p - E\eta_t E\eta_p}{\sqrt{D\eta_t} \sqrt{D\eta_p}} = \sum_{j=i-k+1}^{i+k} \sum_{l=n-k+1}^{n+k} \omega_j^t \omega_l^p r_{|j-l|}. \end{aligned}$$

Лемма доказана.

Если стохастическая интерполяция в точку $t = i + h$, $i \in \mathbb{Z}$, $0 < h < 1$ производится из нечётного числа $m = 2k + 1$, $k \in \mathbb{Z}^+$ целочисленных точек, то набор ближайших точек зависит от h :

- если $h < 0.5$, то ближайшими точками являются $i - k, \dots, i - 1, i, i + 1, \dots, i + k$,
- если $h > 0.5$, то ближайшими точками являются $i - k + 1, \dots, i - 1, i, i + 1, \dots, i + k + 1$,
- при $h = 0.5$ точки $i - k$ и $i + k + 1$ равноудалены от t , поэтому можно рассмотреть два варианта:
 - интерполяция «слева», когда в качестве ближайших точек берутся точки $i - k, \dots, i - 1, i, i + 1, \dots, i + k$,
 - интерполяция «справа», при которой используются точки $i - k + 1, \dots, i - 1, i, i + 1, \dots, i + k + 1$.

Лемма 1.4. Функция $\text{corr}(\eta_t, \eta_p)$ вычисляется по формулам:

1. при $t = i$, $p = n$, $i, n \in \mathbb{Z}$

$$\text{corr}(\eta_t, \eta_p) = \eta_{|i-n|};$$

2. при $t = i$, $p = n + g$, $i, n \in \mathbb{Z}$, $0 < g < 0.5$,

$$\text{corr}(\eta_t, \eta_p) = \sum_{l=n-k}^{n+k} \omega_l^p \eta_{|i-l|};$$

3. при $t = i$, $p = n + g$, $i, n \in \mathbb{Z}$, $g = 0.5$, если используется интерполяция «слева»,

$$\text{corr}(\eta_t, \eta_p) = \sum_{l=n-k}^{n+k} \omega_l^p \eta_{|i-l|}$$

и

$$\text{corr}(\eta_t, \eta_p) = \sum_{l=n-k+1}^{n+k+1} \omega_l^p \eta_{|i-l|}$$

при интерполяции «справа»;

4. при $t = i$, $p = n + g$, $i, n \in \mathbb{Z}$, $0.5 < g < 1$,

$$\text{corr}(\eta_t, \eta_p) = \sum_{l=n-k+1}^{n+k+1} \omega_l^p r_{|i-l|};$$

5. при $t = i + h$, $p = n$, $i, n \in \mathbb{Z}$, $0 < h < 0.5$

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \omega_j^t r_{|j-n|};$$

6. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $0 < h < 0.5$, $0 < g < 0.5$,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \sum_{l=n-k}^{n+k} \omega_j^t \omega_l^p r_{|j-l|};$$

7. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $0 < h < 0.5$, $g = 0.5$ если используется интерполяция «слева»,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \sum_{l=n-k}^{n+k} \omega_j^t \omega_l^p r_{|j-l|}$$

и

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \sum_{l=n-k+1}^{n+k+1} \omega_j^t \omega_l^p r_{|j-l|}$$

при интерполяции «справа»;

8. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $0 < h < 0.5$, $0.5 < g < 1$,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \sum_{l=n-k+1}^{n+k+1} \omega_j^t \omega_l^p r_{|j-l|};$$

9. при $t = i + h$, $p = n$, $i, n \in \mathbb{Z}$, $h = 0.5$ если используется интерполяция «слева»,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \omega_j^t r_{|j-n|}$$

и

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \omega_j^t r_{|j-n|}$$

при интерполяции «справа»;

10. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $h = 0.5$, $0 < g < 0.5$, если используется интерполяция «слева»,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \sum_{l=n-k}^{n+k} \omega_j^t \omega_l^p r_{|j-l|}$$

и

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \sum_{l=n-k}^{n+k} \omega_j^t \omega_l^p r_{|j-l|}$$

при интерполяции «справа»;

11. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $h = 0.5$, $g = 0.5$ если используется интерполяция «слева»,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \sum_{l=n-k}^{n+k} \omega_j^t \omega_l^p r_{|j-l|}$$

и

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \sum_{l=n-k+1}^{n+k+1} \omega_j^t \omega_l^p r_{|j-l|}$$

при интерполяции «справа»;

12. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $h = 0.5$, $0.5 < g < 1$ если используется интерполяция «слева»,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k}^{i+k} \sum_{l=n-k+1}^{n+k+1} \omega_j^t \omega_l^p r_{|j-l|}$$

и

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \sum_{l=n-k+1}^{n+k+1} \omega_j^t \omega_l^p r_{|j-l|}$$

при интерполяции «справа»;

13. при $t = i + h$, $p = n$, $i, n \in \mathbb{Z}$, $0.5 < h < 1$,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \omega_j^t r_{|j-n|};$$

14. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $0.5 < h < 1$, $0 < g < 0.5$,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \sum_{l=n-k}^{n+k} \omega_j^t \omega_l^p r_{|j-l|};$$

15. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $0.5 < h < 1$, $g = 0.5$ если используется интерполяция «слева»,

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \sum_{l=n-k}^{n+k} \omega_j^t \omega_l^p r_{|j-l|}$$

и

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \sum_{l=n-k+1}^{n+k+1} \omega_j^t \omega_l^p r_{|j-l|}$$

при интерполяции «справа»;

16. при $t = i + h$, $p = n + g$, $i, n \in \mathbb{Z}$, $0.5 < h < 1$, $0.5 < g < 1$

$$\text{corr}(\eta_t, \eta_p) = \sum_{j=i-k+1}^{i+k+1} \sum_{l=n-k+1}^{n+k+1} \omega_j^t \omega_l^p r_{|j-l|}.$$

Доказательство леммы 1.4 аналогично доказательству леммы 1.3.

Утверждение 1.3. При фиксированном значении h , $0 < h < 1$ процесс $\zeta_i = \eta_{i+h}$, $i \in \mathbb{Z}$ является стационарным в широком смысле.

Доказательство. Доказательство этого утверждения следует из лемм 1.2-1.4.

Сравнение с нерандомизированной линейной интерполяцией

Рассмотренная стохастическая интерполяция является рандомизированной модификацией линейной интерполяции, при которой значение случайного процесса ψ_t , $t \in \mathbb{R}$ определяется по правилу:

$$\psi_t = \sum_{j=i-k+1}^{i+k} \omega_j^t \xi_j$$

при $t = i + h$, $i \in \mathbb{Z}$, $0 < h < 1$,

$$\psi_t = \xi_i$$

при $t = i$, $i \in \mathbb{Z}$. Следует отметить, что хотя математические ожидания процессов η_t и ψ_t совпадают и равны μ , ряд вероятностных характеристик этих процессов существенно различаются. В частности, справедливо Утверждение 1.4.

Утверждение 1.4. В условиях данного параграфа верно неравенство

$$D\psi_t \leq D\eta_t.$$

Равенство дисперсий процессов η_t и ψ_t достигается тогда и только тогда, когда

$$\forall i, j \in \mathbb{Z} \quad r_{|i-j|} = 1.$$

Доказательство. Очевидно, что $D\eta_t = \sigma^2$. Вычислим $D\psi_t$ при $m = 2k$, $k \in \mathbb{Z}^+$. Пусть $t = i + h$, $i \in \mathbb{Z}$, $0 < h < 1$. Тогда

$$\begin{aligned} D\psi_t &= E\psi_t^2 - \mu^2 = E\left(\sum_{j=i-k+1}^{i+k} \omega_j^t \xi_j\right)^2 - \mu^2 = \\ &= \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t E\xi_j \xi_l - \mu^2 = \\ &= \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t (\sigma^2 r_{|j-l|} + \mu^2) - \mu^2 = \\ &= \sigma^2 \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t r_{|j-l|} + \mu^2 \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t - \mu^2 = \\ &= \sigma^2 \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t r_{|j-l|}. \end{aligned}$$

Поскольку $\sum_{j=1}^m \omega_j^t = 1$, справедливы равенства

$$\sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t = \sum_{j=i-k+1}^{i+k} \left(\omega_j^t \sum_{l=i-k+1}^{i+k} \omega_l^t \right) = \sum_{j=i-k+1}^{i+k} \omega_j^t = 1.$$

Следовательно,

$$D\psi_t = \sigma^2 \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t r_{|j-l|} \leq \sigma^2 \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t = \sigma^2 = D\eta_t.$$

Кроме этого,

$$\begin{aligned}
D\psi_t &= D\eta_t \Leftrightarrow \\
\Leftrightarrow \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t r_{|j-l|} &= \sigma^2 \sum_{j=i-k+1}^{i+k} \sum_{l=i-k+1}^{i+k} \omega_j^t \omega_l^t \Leftrightarrow \\
&\Leftrightarrow \forall i, j \in \mathbb{Z} \quad r_{|i-j|} = 1.
\end{aligned}$$

Доказательство утверждения при $m = 2k + 1, k \in \mathbb{Z}^+$ проводится аналогичным образом.

Замечание 1.4. Очевидно, что

$$D\eta_t - D\psi_t = \sigma^2 \sum_{j=i-k+1}^{i+k} \sum_{\substack{l=i-k+1 \\ l \neq j}}^{i+k} \omega_j^t \omega_l^t (1 - r_{|j-l|}).$$

Стохастическая интерполяция нестационарного процесса

Пусть процесс $\dots, \xi_{-2}, \xi_{-1}, \xi_0, \xi_1, \xi_2, \xi_3 \dots$ нестационарный с плотностями одномерных распределений $f_i(x)$, математическими ожиданиями и дисперсиями μ_i и $0 < \sigma_i^2 < \infty$, соответственно, корреляционной функцией $\text{corr}(\xi_i, \xi_j) = r_{ij}, i, j \in \mathbb{Z}$. Построим процесс η_t согласно формулам (1.1)-(1.2). Очевидно, что справедливы леммы 1.5 и 1.6.

Лемма 1.5. При $t = i, i \in \mathbb{Z}$ плотность одномерного распределения процесса η_t суть $f_i(x)$. При $t = i + h, i \in \mathbb{Z}, 0 < h < 1$ плотность одномерного распределения процесса η_t – смесь плотностей $f_j(x)$ с весами ω_j^t , где

- $j = i - k + 1, i - k + 2, \dots, i + k$ при $m = 2k, k \in \mathbb{Z}^+$,
- $j = i - k, i - k + 1, \dots, i + k$ при $m = 2k + 1, k \in \mathbb{Z}^+, h < 0.5$ или $h = 0.5$ и интерполяция «слева»,
- $j = i - k + 1, i - k + 2, \dots, i + k + 1$ при $m = 2k + 1, k \in \mathbb{Z}^+, h > 0.5$ или $h = 0.5$ и интерполяция «справа».

Лемма 1.6. При $t = i + h, i \in \mathbb{Z}, 0 < h < 1$

$$E\eta_t = \sum_j \omega_j^t \mu_j, \quad D\eta_t = \sum_j \omega_j^t \sigma_j^2 + \sum_j \omega_j^t \mu_j^2 - (E\eta_t)^2.$$

Здесь суммирование по j осуществляется в пределах, которые определяются значениями m и h вышеуказанным образом.

В пункте 2 параграфа 3.4 Главы 3 для интерполяции поля индекса холодного стресса используется стохастическая интерполяция неоднородного двумерного случайного поля. Математическое ожидание и дисперсия интерполированного поля вычисляются аналогичным образом.

1.3. Алгоритмы моделирования условных негауссовских случайных процессов с точечными и интервальными условиями

Моделирование условных гауссовских и негауссовских процессов применяется при решении различных задач метеорологии, океанологии и геофизики [49, 177, 186]. Для моделирования условных гауссовских процессов, удовлетворяющих точечным условиям, существует эффективный алгоритм, представленный в [175]. Моделирование условных гауссовских процессов, являющихся решением недоопределенной системы линейных уравнений можно осуществлять по алгоритму, описанному в [67].

Перейдём к рассмотрению вопроса, связанного с моделированием условных негауссовских временных рядов $\vec{\zeta} = (\zeta_1, \zeta_2, \dots, \zeta_N)$ на основе реальных данных. Моделируемый ряд должен удовлетворять одному из 2 типов условий:

1. точечные условия:

$$\zeta_i = c_i, \quad i \in \Omega, \quad \Omega \subset \{1, 2, \dots, N\}, \quad c_i \in \mathbb{R},$$

2. интервальные условия:

$$\zeta_i \in (a_i, b_i), i \in \Omega, \Omega \subset \{1, 2, \dots, N\}, -\infty \leq a_i, b_i \leq +\infty.$$

Интервалы (a_i, b_i) могут быть открытыми, замкнутыми или полужамкнутыми. Для удобства будем считать, что $\Omega = \{1, 2, \dots, K\}$, $K < N$.

Моделирование условных рядов с точечными условиями

Пусть, $\zeta_i^j, i = 1, 2, \dots, N, j = 1, 2, \dots, NY$ – выборка реальных данных (например, выборка значений какого-либо метеопараметра). Моделирование условного негауссовского вектора $\vec{\zeta} = (\zeta_1, \zeta_2, \dots, \zeta_N)$, удовлетворяющего точечным условиям $\zeta_i = c_i, i \in \Omega, \Omega \subset \{1, 2, \dots, N\}, c_i \in \mathbb{R}$, можно осуществлять с использованием метода нормализации [46, 77, 112] по приближенному Алгоритму 1.2, описанному в [172].

Алгоритм 1.2.

1. Преобразуем реальные данные

$$\eta_i^j = \Phi^{-1}\left(G_i\left(\zeta_i^j\right)\right), i = 1, 2, \dots, N, j = 1, 2, \dots, NY.$$

Здесь $\Phi(x)$ – функция одномерного стандартного нормального распределения, $G_i(x)$ – функция одномерного распределения i -ой компоненты вектора $\vec{\zeta}$. При этом преобразовании условия $\zeta_i = c_i, i \in \Omega$ превращаются в условия $\eta_i = c_i, i \in \Omega$.

2. Оцениваем по $\eta_i^j, i = 1, 2, \dots, N, j = 1, 2, \dots, NY$ корреляционную матрицу R размерности $N \times N$:

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},$$

где R_{11} – матрица, составленная из коэффициентов корреляции между $\eta_i, i \in \Omega$, R_{22} – матрица, составленная из коэффициентов корреляции между свободными компонентами $\eta_i, i \in \{1, 2, \dots, N\} / \Omega$, R_{12}, R_{21} – соответствующие взаимные матрицы коэффициентов корреляции.

3. Моделируем гауссовский вектор $\vec{\eta} = \begin{pmatrix} \vec{\eta}^{(1)} \\ \vec{\eta}^{(2)} \end{pmatrix}$, удовлетворяющий условиям $\eta_i = c_i, i \in \Omega$ [23, 175]. Компоненты вектора $\vec{\eta}^{(1)}$ есть $c_i, i \in \Omega$, а подвектор $\vec{\eta}^{(2)}$ вычисляется по формуле

$$\vec{\eta}^{(2)} = R_{21}R_{11}^{-1}\vec{\eta}^{(1)} + A_1\vec{\phi},$$

где $\vec{\phi}$ – вектор, составленный из независимых случайных величин со стандартным нормальным распределением, а нижнетреугольная матрица A_1 такова, что

$$A_1 A_1^T = R_{22} - R_{21}R_{11}^{-1}R_{12}.$$

4. Производим преобразование

$$\zeta_i = G_i^{-1}(\Phi(\eta_i)), i = 1, 2, \dots, N.$$

Замечание 1.5. На 4-ом шаге Алгоритма 1.2 используется приближение, которое состоит в том, что функции условных одномерных распределений $G_i^{cond}(x)$ свободных компонент $\zeta_i, i = K + 1, \dots, N$ вектора $\vec{\zeta}$ в общем случае отличны от $G_i(x), i = K + 1, \dots, N$, а используемое преобразование даёт величины с функциями распределения $G_i(x), i = K + 1, \dots, N$. Если бы функции условных распределений $G_i^{cond}(x)$ были известны, то на 4-ом шаге Алгоритма 1.2 следовало бы использовать преобразование

$$\zeta_i = \left(G_i^{cond}\right)^{-1}(\Phi(\eta_i)), i = 1, 2, \dots, N.$$

Однако, в реальных задачах функции $G_i^{cond}(x)$ чаще неизвестны, и последнее преобразование применить невозможно.

Моделирование условных рядов с интервальными условиями

Для моделирования вектора $\vec{\zeta}$, удовлетворяющего интервальным условиям $\zeta_i \in (a_i, b_i), i \in \Omega, \Omega \subset \{1, 2, \dots, N\}, -\infty \leq a_i, b_i \leq +\infty$, можно использовать Алгоритм

1.3, основанный на методе обратных функций распределения (см. Приложение 1) и методе исключения (переборе траекторий).

Алгоритм 1.3.

1. По данным $\zeta_i^j, i=1,2,\dots,N, j=1,2,\dots,NY$ оцениваем выборочную корреляционную матрицу вектора $\vec{\zeta}$ и в рамках метода обратных функций распределения вычисляем корреляционную матрицу R'_ζ вспомогательного безусловного гауссовского процесса.

2. Моделируем реализацию безусловного гауссовского вектора $(\xi_1, \xi_2, \dots, \xi_N)$ с нулевым средним и корреляционной матрицей R'_ζ .

3. Проверяем, удовлетворяет ли реализация интервальным условиям $\xi_i \in (a'_i, b'_i), i \in \Omega$, где $a'_i = \Phi^{-1}(G_i(a_i)), b'_i = \Phi^{-1}(G_i(b_i)), i \in \Omega$. Если хотя бы одно из условий не выполнено, траекторию исключаем («забываем») и повторяем шаг 2.

4. Трансформируем реализацию условного вектора

$$\zeta_i = G_i^{-1}(\Phi(\xi_i)), i = 1, \dots, N.$$

Замечание 1.6. В данной работе на 2-ом шаге Алгоритма 1.3 моделирование безусловного гауссовского вектора $\vec{\xi}$ с заданной корреляционной матрицей R'_ζ осуществлялось с помощью спектрального разложения матрицы R'_ζ :

$$\vec{\xi} = VD^{1/2}\vec{\phi},$$

где $\vec{\phi}$ – вектор, составленный из независимых случайных величин со стандартным нормальным распределением, диагональная матрица D составлена из собственных чисел $d_1 > d_2 > \dots > d_N$ матрицы R'_ζ , а матрица $V = (\vec{v}_1, \dots, \vec{v}_N)$ – из нормированных собственных векторов \vec{v}_i , соответствующих d_i .

Алгоритм 1.3 основан на методе исключения, соответственно, согласно [53], его трудоёмкость s пропорциональна математическому ожиданию целочисленной

случайной величины, имеющей геометрическое распределение с параметром $p = P(\xi_i \in (a'_i, b'_i), i \in \Omega)$, т.е.

$$s \sim \frac{1}{p} = \frac{1}{P(\xi_i \in (a'_i, b'_i), i \in \Omega)}.$$

Очевидно, что $s \geq 1$ и что при жёстких условиях (например, при $a_i \gg E\xi_i$ или при $b_i - a_i \ll 1$) на 3-ем шаге Алгоритма 1.3 исключается очень большое число траекторий, поэтому Алгоритм 1.3 оказывается весьма трудоёмким. В качестве примеров в таблицах 1.1, 1.2 приведено время моделирования 10^5 реализаций гауссовского вектора $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_{11})$ с нулевым средним и корреляционной матрицей

$$R = (r_{ij})_{i,j=1}^{11}, r_{ij} = e^{-0.05|i-j|},$$

на который наложены интервальные условия $\xi_i \in (a, b), i \in \Omega$.

Естественным образом возникает вопрос о том, как уменьшить трудоёмкость моделирования. При моделировании процесса с точечными условиями на 3-шаге Алгоритма 1.3 использовался метод условных распределений. При этом учитывался тот факт, что при фиксированных значениях ряда компонент гауссовского вектора распределение свободных компонент также является гауссов-

Таб. 1.1. Время моделирования как функция интервала (a, b) при $\Omega = \{1, 2, \dots, 10\}$.

(a, b)	Время, сек.
$(-1, 1)$	4.95
$(-0.5, 1)$	10.83
$(0, 1)$	126.06
$(0.1, 1)$	288.45
$(0.5, 1)$	$5 \cdot 10^4$

Таб. 1.2. Время моделирования как функция Ω при $(a,b)=(0,3,1)$.

Ω	Время, сек.
$\Omega = \{1\}$	4.06
$\Omega = \{1,2,3\}$	10.08
$\Omega = \{1,2,\dots,5\}$	36.97
$\Omega = \{1,2,\dots,7\}$	194.42
$\Omega = \{1,2,\dots,10\}$	2080.70

ским [2]. При замене точечных условий на интервальные, это свойство не сохраняется что можно продемонстрировать на следующем примере. Пусть $\vec{\xi} = (\xi_1, \xi_2)$ – стандартный гауссовский вектор с коэффициентом корреляции ρ между компонентами. Плотность $f(x)$ условного распределения компоненты ξ_2 при условии $a_1 < \xi_1 < b_1$ имеет вид

$$f(x|a_1 < \xi_1 < b_1) = \frac{1}{\Phi(b_1) - \Phi(a_1)} \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{a_1}^{b_1} e^{-\frac{u^2 - 2\rho ux + x^2}{2(1-\rho^2)}} du,$$

где $\Phi(x)$ – функция стандартного нормального распределения. Эта плотность не является гауссовской, что показано на рисунке 1.7. Для сравнения на этом же рисунке приведён график гауссовской плотности распределения $f_{gau}(x)$ с математическим ожиданием и дисперсией как у $f(x)$.

Для уменьшения трудоёмкости моделирования условных векторов с интервальными условиями можно использовать приближенный Алгоритм 1.4.

Алгоритм 1.4.

Пусть n – требуемое число реализаций случайного вектора $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_N)$, удовлетворяющего условиям $\xi_i \in (a'_i, b'_i)$, $i = 1, 2, \dots, K$. На каждом интервале

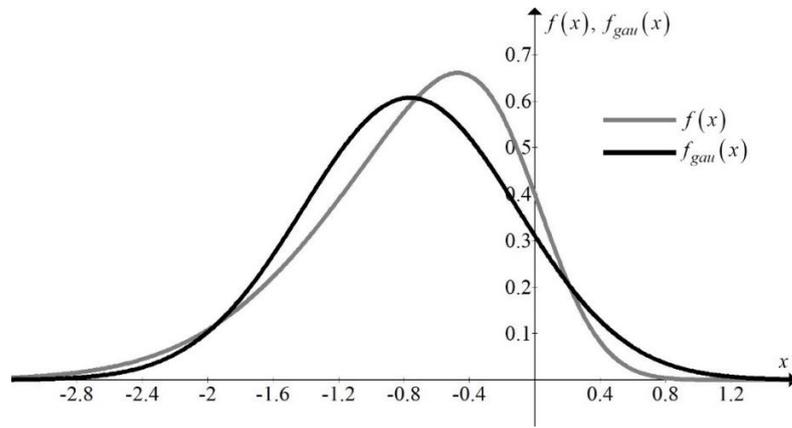


Рис. 1.7. Плотности одномерных распределений $f(x)$, $f_{gau}(x)$.

(a'_i, b'_i) определим равномерную сетку (см. рисунок 1.8)

$$c_i^j = a'_i + (j-1) \frac{(b'_i - a'_i)}{n-1}, \quad j=1, 2, \dots, n, \quad i=1, 2, \dots, K.$$

В качестве j -ой реализации вектора $\vec{\xi}$ рассматриваем реализацию условного вектора $\vec{\xi}' = (\xi'_1, c, \dots, \xi'_N)$, удовлетворяющего точечным условиям $\xi'_i = c_i^j$, $j=1, 2, \dots, n$, $i=1, 2, \dots, K$.

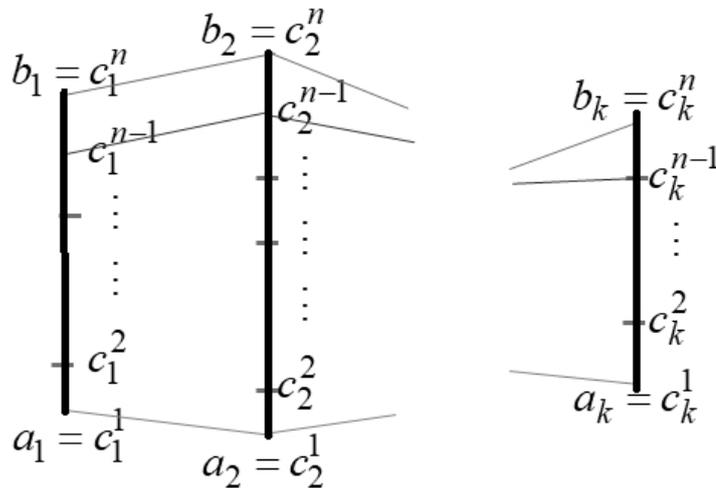


Рис. 1.8. Сетка c_i^j , $j=1, 2, \dots, n$, $i=1, 2, \dots, K$.

По времени моделирования Алгоритм 1.4 выигрывает у Алгоритма 1.3. Так, например, для вышерассмотренного вектора $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_{11})$ время

моделирования при $\Omega = \{1, 2, \dots, 10\}$ и $(a, b) = (0.5, 1)$ оставляет всего 156.42 сек., против $5 \cdot 10^4$ сек. при моделировании с помощью Алгоритма 1.3. Однако, предложенный Алгоритм 1.4 обладает недостатком: математическое ожидание и дисперсия смоделированного процесса отличается от требуемых. На рисунках 1.9, 1.10 приведены графики оценок условного математического ожидания и условной дисперсии гауссовского вектора $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_{30})$ с нулевым средним и корреляционной матрицей

$$R = (r_{ij})_{i,j=1}^{30}, r_{ij} = e^{-0.05|i-j|},$$

на который наложены интервальные условия $\xi_i \in (a, b)$, $i = 1, 2, \dots, 10$, полученных по 10^7 реализаций Алгоритма 1.3 и Алгоритма 1.4. Численные эксперименты показывают, то при широких интервалах (a_i, b_i) отклонение математического ожидания модельного процесса от требуемого может быть весьма существенным. Однако при узких интервалах (a_i, b_i) это отклонение незначительно.

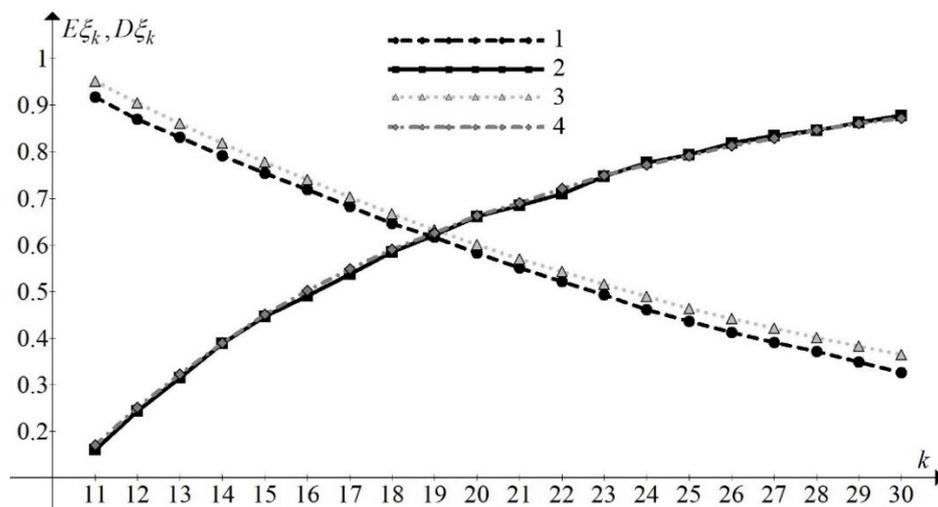


Рис. 1.9. Кривые 1, 2 – условные математическое ожидание $E\xi_k$ и дисперсия $D\xi_k$, оцененные по реализациям Алгоритма 1.3, кривые 3, 4 – $E\xi_k$ и $D\xi_k$, оцененные по реализациям Алгоритма 1.4. $(a, b) = (0.5, 1.5)$.

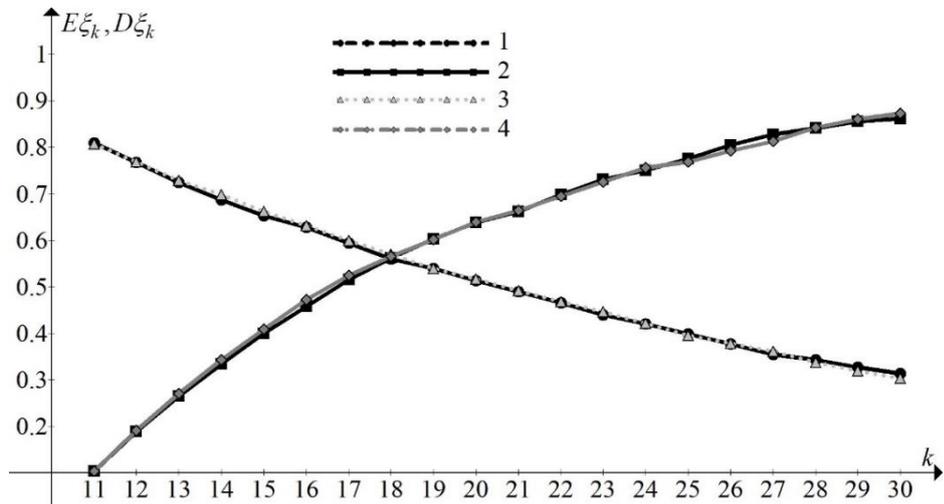


Рис. 1.10. Кривые 1, 2 – условные математическое ожидание $E\xi_k$ и дисперсия $D\xi_k$, оцененные по реализациям Алгоритма 1.3, кривые 3, 4 – $E\xi_k$ и $D\xi_k$, оцененные по реализациям Алгоритма 1.4. $(a, b) = (0.7, 1.0)$.

Отметим, что для интервальных условий специального вида (например, $(a'_i, b'_i) = (-\infty, b_i)$, $i \in \Omega$) существуют и другие алгоритмы моделирования условного гауссовского процесса, которые могут быть успешно применены на 2 и 3 шагах Алгоритма 1.3. Обзор этих алгоритмов приведён в [162, 203].

Глава 2. Стохастические «генераторы погоды»

В параграфе 2.1 этой главы представлены численные стохастические модели временных рядов комплексов метеорологических параметров и их специальных характеристик. Предложенные модели позволяют учитывать суточный и годовой ход реальных метеопроцессов. В параграфе 2.2 описана стохастическая модель пространственно-временного поля комплекса метеопараметров, учитывающая пространственную неоднородность рассматриваемого поля и присущий ему суточный ход. Некоторые из предложенных моделей использованы в Главе 3 для решения задач биоклиматологии.

2.1. Стохастическое моделирование рядов метеорологических параметров и их комплексов

2.1.1. Стохастическая модель совместных рядов индикатора выпадения осадков, минимальной и максимальной за сутки приземной температуры воздуха

В этом параграфе рассматривается стохастическая модель совместных временных рядов трех метеорологических характеристик: индикатора наличия/отсутствия осадков в течение суток, минимальной и максимальной за сутки приземной температуры воздуха. Моделирование такого набора характеристик метеорологических процессов встречается в различных стохастических «генераторах погоды» (см., например, [193]), однако, чаще всего,

осуществляется в предположении стационарности метеопроцессов на временных интервалах длительностью до 1 месяца.

В данной работе для учёта годового хода реальных метеорологических процессов использовано предположение о нестационарности рядов рассматриваемых параметров. Для моделирования индикаторного ряда используется латентный нестационарный гауссовский процесс, а моделирование рядов температуры осуществляется в предположении гауссовости их одномерных распределений. Ниже также приведены специальные преобразования уравнений, связывающих корреляционные функции латентного гауссовского и индикаторного процессов. Эти преобразования позволяют существенно сократить время вычислений значений корреляционной функции латентного процесса при заданных значениях корреляционной функции индикаторного.

Основная моделирующая формула и входные параметры модели

Опишем формально правила и алгоритмы, по которым проводилось моделирование. Моделируется процесс $(\vec{I}^T, \vec{A}^T, \vec{E}^T)$, где $\vec{I}^T = (I_1, I_2, \dots, I_N)^T$ – вектор-столбец минимальных суточных температур, $\vec{A}^T = (A_1, A_2, \dots, A_N)^T$ – вектор максимальных суточных температур, $\vec{E}^T = (E_1, E_2, \dots, E_N)^T$ – вектор индикаторов наличия / отсутствия осадков. Значение $E_j = 1$ или $E_j = 0$ показывает, что в сутки номер j осадки наблюдались или отсутствовали, соответственно. Здесь N – число дней в рассматриваемом временном интервале, $N \in \{1, 2, \dots, 365\}$. Основная моделирующая формула имеет вид

$$\begin{aligned} I_j &= \sigma_j^I \zeta_j^I + \mu_j^I, \\ A_j &= \sigma_j^A \zeta_j^A + \mu_j^A, \\ E_j &= \begin{cases} 1, & \zeta_j^E \leq c_j, \\ 0, & \zeta_j^E > c_j, \end{cases} \end{aligned} \quad (2.1)$$

где

$$\begin{pmatrix} \mu_1^I \\ \mu_2^I \\ \dots \\ \mu_N^I \end{pmatrix}, \begin{pmatrix} \mu_1^A \\ \mu_2^A \\ \dots \\ \mu_N^A \end{pmatrix}, \begin{pmatrix} \sigma_1^I \\ \sigma_2^I \\ \dots \\ \sigma_N^I \end{pmatrix}, \begin{pmatrix} \sigma_1^A \\ \sigma_2^A \\ \dots \\ \sigma_N^A \end{pmatrix}$$

суть известные векторы средних значений и среднеквадратических отклонений минимальной и максимальной температуры воздуха, соответственно. Далее значения компонент этих векторов будут оцениваться по реальным данным.

Пороговые значения c_j для каждого $j = 1, 2, \dots, N$ определяются из уравнения

$$P(E_j = 1) = P(\xi_j^E \leq c_j) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c_j} \exp\left(-\frac{t^2}{2}\right) dt = p_j,$$

где p_j – заданная вероятность наличия осадков в j -е сутки (например, оцененная по реальным данным) [42]. Отметим, что равенство $c_j = 0$ выполняется тогда и

только тогда, когда $p_j = \frac{1}{2}$, неравенство $c_j > 0$ выполняется только при $p_j > \frac{1}{2}$, а

неравенство $c_j < 0$ – только при $p_j < \frac{1}{2}$. Все дальнейшие рассуждения приведены

для случая $p_j \neq 0$ и $p_j \neq 1$ при $j = 1, 2, \dots, N$. Величины $\xi_j^I, \xi_j^A, \xi_j^E$ есть компоненты

совместного гауссовского ряда $\left(\left(\bar{\xi}^I \right)^T, \left(\bar{\xi}^A \right)^T, \left(\bar{\xi}^E \right)^T \right)$ с нулевым средним и такой

специально выбранной корреляционной матрицей

$$G = \begin{pmatrix} G_{II} & G_{IA} & G_{IE} \\ G_{AI} & G_{AA} & G_{AE} \\ G_{EI} & G_{EA} & G_{EE} \end{pmatrix},$$

что процесс $(\vec{I}^T, \vec{A}^T, \vec{E}^T)$ обладает заданной корреляционной матрицей

$$R = \begin{pmatrix} R_{II} & R_{IA} & R_{IE} \\ R_{AI} & R_{AA} & R_{AE} \\ R_{EI} & R_{EA} & R_{EE} \end{pmatrix}.$$

Индекс XU у блоков G_{XU}, R_{XU} означает, что эти блоки составлены из взаимных коэффициентов корреляции между значениями метеорологических характеристик X и U во все возможные сутки i и j , где $X, U \in \{I, A, E\}$, $i, j = 1, 2, \dots, N$. Отметим, что матрицы R и G имеют размерность $3N \times 3N$.

Корреляционная матрица вспомогательного гауссовского процесса

Возникает вопрос о том, как определить такую корреляционную матрицу G вспомогательного гауссовского процесса $\left(\left(\bar{\xi}^I \right)^T, \left(\bar{\xi}^A \right)^T, \left(\bar{\xi}^E \right)^T \right)$, чтобы построенный согласно преобразованию (2.1) процесс обладал необходимой корреляционной матрицей R . Отметим, что в работе [174] был рассмотрен аналог преобразования (2.1) в более простом случае. Для того чтобы определить все блоки матрицы G через R и p_j воспользуемся леммами 2.1-2.3. Далее элементы подматриц G_{XU}, R_{XU} , стоящие на пересечении i -ой строки и j -го столбца, обозначены $g_{XU}(i, j)$ и $r_{XU}(i, j)$, соответственно. Полагаем, что $|r_{XU}(i, j)| \neq 1$ и $|g_{XU}(i, j)| \neq 1$ при $i \neq j$,

Лемма 2.1.

$$\begin{aligned} G_{II} &= R_{II}, \\ G_{AA} &= R_{AA}, \\ G_{IA} &= R_{IA}, \\ G_{AI} &= R_{AI}. \end{aligned}$$

Доказательство этой леммы очевидно.

Лемма 2.2. Элементы $g_{IE}(i, j)$, $i \neq j$ матрицы G_{IE} вычисляются по правилу

$$g_{IE}(i, j) = -r_{IE}(i, j) \sqrt{2\pi p_j (1 - p_j)} \exp\left(\frac{c_j^2}{2}\right), \quad i, j = 1, 2, \dots, N.$$

Доказательство. Справедливость леммы следует из следующей цепочки равенств

$$r_{IE}(i, j) = \frac{E(I_i E_j) - EI_i EE_j}{\sqrt{DI_i} \sqrt{DE_j}} = \frac{E\left(\left(\sigma_i^I \xi_i^I + \mu_i^I\right) E_j\right) - \mu_i^I p_j}{\sigma_i^I \sqrt{p_j(1-p_j)}} =$$

$$= -\frac{g_{IE}(i, j)}{\sqrt{2\pi p_j(1-p_j)}} \exp\left(-\frac{c_j^2}{2}\right).$$

Элементы матриц G_{EI}, G_{AE}, G_{EA} определяются аналогичным образом.

Лемма доказана.

Очевидно, что $r_{EE}(i, i) = g_{EE}(i, i) = 1, i = 1, 2, \dots, N$.

Лемма 2.3. Если $i \neq j$, то $g_{EE}(i, j)$ есть решение уравнения

$$r_{EE}(i, j) = \begin{cases} \frac{\frac{1}{2} p_i + \frac{1}{2} p_j - p_i p_j}{\sqrt{p_i(1-p_i)} \sqrt{p_j(1-p_j)}} - \frac{T(c_i, a_1) + T(c_j, a_2)}{\sqrt{p_i(1-p_i)} \sqrt{p_j(1-p_j)}}, & \text{при } c_i c_j > 0, \\ \frac{\frac{1}{2} p_i + \frac{1}{2} p_j - p_i p_j - \frac{1}{2}}{\sqrt{p_i(1-p_i)} \sqrt{p_j(1-p_j)}} - \frac{T(c_i, a_1) + T(c_j, a_2)}{\sqrt{p_i(1-p_i)} \sqrt{p_j(1-p_j)}}, & \text{при } c_i c_j < 0, \\ \frac{2 \arcsin g_{EE}(i, j)}{\pi}, & \text{при } c_i = 0, c_j = 0, \\ \frac{2T\left(c_i, \frac{g_{EE}(i, j)}{\sqrt{1-g_{EE}^2(i, j)}}\right)}{\sqrt{p_i(1-p_i)}}, & \text{при } c_i \neq 0, c_j = 0, \end{cases} \quad (2.2)$$

где $T(c, a)$ – функция Оуэна [80, 180]

$$T(c, a) = \frac{1}{2\pi} \int_0^a \exp\left(-\frac{c^2(1+t^2)}{2}\right) \frac{dt}{1+t^2},$$

$$a_1 = \frac{c_j - c_i g_{EE}(i, j)}{c_i \sqrt{1-g_{EE}^2(i, j)}}, \quad a_2 = \frac{c_i - c_j g_{EE}(i, j)}{c_j \sqrt{1-g_{EE}^2(i, j)}}.$$

Доказательство. Из определения коэффициента корреляции и определения процесса \bar{E}^T следует выполнение равенств

$$r_{EE}(i, j) = \frac{P(E_i = 1, E_j = 1) - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} \quad (2.3)$$

для $i, j = 1, 2, \dots, N$. Здесь

$$P(E_i = 1, E_j = 1) = P(\xi_i^E \leq c_i, \xi_j^E \leq c_j) = F(c_i, c_j, g_{EE}(i, j)), \quad (2.4)$$

где

$$F(h, k, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^h \int_{-\infty}^k \exp\left(-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right) dx dy.$$

Равенства (2.3) и (2.4) позволяют численно определить значение всех элементов матрицы G_{EE} . Справедливы формулы Оуэна [80, 180]

$$F(c_i, c_j, g_{EE}(i, j)) = \begin{cases} \frac{1}{2}\Phi_{Lap}(c_i) + \frac{1}{2}\Phi_{Lap}(c_j) - T(c_i, a_1) - T(c_j, a_2) + \frac{1}{2}, & \text{если } c_i \geq 0, c_j \geq 0 \text{ или } c_i < 0, c_j < 0, \\ \frac{1}{2}\Phi_{Lap}(c_i) + \frac{1}{2}\Phi_{Lap}(c_j) - T(c_i, a_1) - T(c_j, a_2), & \text{если } c_i < 0, c_j \geq 0 \text{ или } c_i \geq 0, c_j < 0, \end{cases} \quad (2.5)$$

где

$$\Phi_{Lap}(c_i) = \frac{1}{\sqrt{2\pi}} \int_0^{c_i} \exp\left(-\frac{t^2}{2}\right) dt.$$

Соответственно, уравнение (2.3) при $c_i c_j > 0$ принимает вид

$$r_{EE}(i, j) = \frac{\frac{1}{2}\Phi_{Lap}(c_i) + \frac{1}{2}\Phi_{Lap}(c_j) - T(c_i, a_1) - T(c_j, a_2) + \frac{1}{2} - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}},$$

а при $c_i c_j < 0$ – вид

$$r_{EE}(i, j) = \frac{\frac{1}{2}\Phi_{Lap}(c_i) + \frac{1}{2}\Phi_{Lap}(c_j) - T(c_i, a_1) - T(c_j, a_2) - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}}.$$

Верны также равенства

$$p_i = \frac{1}{2} + \Phi_{Lap}(c_i), i = 1, 2, \dots, N.$$

Следовательно,

$$r_{EE}(i, j) = \frac{\frac{1}{2}p_i + \frac{1}{2}p_j - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} - \frac{T(c_i, a_1) + T(c_j, a_2)}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} \text{ при } c_i c_j > 0,$$

$$r_{EE}(i, j) = \frac{\frac{1}{2}p_i + \frac{1}{2}p_j - p_i p_j - \frac{1}{2}}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} - \frac{T(c_i, a_1) + T(c_j, a_2)}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} \text{ при } c_i c_j < 0.$$

Если $c_i = c_j = 0$, то

$$p_i p_j = \frac{1}{4}, \sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)} = \frac{1}{4}.$$

Кроме того, согласно [80] справедливо равенство

$$F(0, 0, g_{EE}(i, j)) = \frac{1}{4} + \frac{\arcsin g_{EE}(i, j)}{2\pi}.$$

Следовательно, уравнение (2.3) принимает вид

$$r_{EE}(i, j) = \frac{2\arcsin g_{EE}(i, j)}{\pi}.$$

Указанные в формулировке леммы равенства

$$r_{EE}(i, j) = \frac{2T\left(c_i, \frac{g_{EE}(i, j)}{\sqrt{1-g_{EE}^2(i, j)}}\right)}{\sqrt{p_i(1-p_i)}}$$

при $c_i \neq 0, c_j = 0$ следуют из равенств (2.3)-(2.5) и того, что

$$\Phi(0) = 0, \lim_{a \rightarrow \infty} T(h, a) = \frac{1 - 2\Phi(|h|)}{4}, T(h, a) = -T(h, -a).$$

Лемма доказана.

Следует отметить, что уравнение (2.3) разрешимо относительно $g_{EE}(i, j)$ не при всех $r_{EE}(i, j), p_i, p_j$, и справедливо следующее утверждение.

Утверждение 2.1. При заданных p_i, p_j уравнение (2.3) разрешимо относительно $g_{EE}(i, j)$ только при $r_{EE}(i, j) \in (a, b) \subseteq (-1, 1)$, где

$$a = \frac{F\left(\Phi_{Lap}^{-1}(p_i), \Phi_{Lap}^{-1}(p_j), -1\right) - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}},$$

$$b = \frac{F\left(\Phi_{Lap}^{-1}(p_i), \Phi_{Lap}^{-1}(p_j), 1\right) - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}}.$$

Доказательство этого утверждения следует из монотонности функции $F(\cdot, \cdot, \cdot)$ как функции $g_{EE}(i, j)$ при $g_{EE}(i, j) \in (-1, 1)$.

Замечание 2.1. Из монотонности правой части уравнения (2.3) как функции $g_{EE}(i, j)$ следует также единственность решения уравнения (2.3).

Замечание 2.2. В работе [4] приведены значения a, b при $p_i, p_j \geq \frac{1}{2}$ или $p_i, p_j < \frac{1}{2}$, выраженные без использования функции $F(\cdot, \cdot, \cdot)$:

$$a = \frac{\frac{1}{2}p_i + \frac{1}{2}p_j - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} - \frac{1 - \frac{1}{2}\left(\Phi_{Lap}\left(\left|\Phi_{Lap}^{-1}(p_i)\right|\right) + \Phi_{Lap}\left(\left|\Phi_{Lap}^{-1}(p_j)\right|\right)\right)}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}},$$

$$b = \frac{\frac{1}{2}p_i + \frac{1}{2}p_j - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} - \frac{\frac{1}{2}\left|\Phi_{Lap}\left(\left|\Phi_{Lap}^{-1}(p_i)\right|\right) - \Phi_{Lap}\left(\left|\Phi_{Lap}^{-1}(p_j)\right|\right)\right|}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}}.$$

Замечание 2.3. Матрица G , составленная их коэффициентов корреляции, вычисленных вышеуказанным способом (леммы 2.1-2.3), может оказаться не положительно определенной, т.е. матрица G может оказаться не корреляционной матрицей случайного процесса. В этом случае перед моделированием гауссовского процесса необходимо регуляризовать матрицу G , и моделирование процесса осуществлять уже с регуляризованной корреляционной матрицей. Далее будем считать, что матрица G регуляризована.

Моделирование процесса $(\vec{I}^T, \vec{A}^T, \vec{E}^T)$

После того как матрица G полностью определена, можно приступить к моделированию гауссовского процесса $\left(\left(\vec{\xi}^I \right)^T, \left(\vec{\xi}^A \right)^T, \left(\vec{\xi}^E \right)^T \right)$ и последующему построению процесса $(\vec{I}^T, \vec{A}^T, \vec{E}^T)$ согласно основной моделирующей формуле (2.1). Моделирование гауссовского процесса с заданной корреляционной матрицей можно осуществлять с помощью любого известного алгоритма (см., например, [175]). В численных экспериментах, результаты которых приведены ниже, для моделирования использовался метод, основанный на разложении Холецкого корреляционной матрицы (см. Приложение 1). Здесь и всюду далее, в качестве датчика псевдослучайных чисел α_k , равномерно распределенных на отрезке $(0,1)$, используется широко известный мультипликативный генератор [53, 66]

$$\alpha_k = 2^{-40} n_k,$$

$$n_0 = 1, n_{k+1} \equiv n_k \cdot 5^{17} \pmod{2^{40}}.$$

По смыслу минимальной и максимальной температуры в каждой моделируемой реализации во все моменты времени $j = 1, 2, \dots, N$ должно выполняться условие $I_j \leq A_j$. Однако, использование моделирующей формулы (2.1) выполнение этого неравенства не гарантирует. Соответственно, при численной реализации рассматриваемой модели необходимо проверять выполнение условия $I_j \leq A_j$. Реализации вектора $(\vec{I}^T, \vec{A}^T, \vec{E}^T)$, в которых требуемое неравенство не выполняется, требуется исключать, т.е. не включать в выборку, по которой производятся проверка адекватности модели и дальнейшие исследования.

Верификация модели

Верификация предложенной модели проводилась на основе данных многолетних метеорологических наблюдений на ряде метеостанций,

расположенных на территории Российской Федерации. Полученные на метеостанциях данные о суточном количестве осадков были преобразованы в индикаторные ряды, характеризующие чередование суток с осадками и без них. Для оценки параметров модели использовались данные, собранные за $NY = 34$ года в период с 1976 по 2009 гг., а также дополнительные данные за декабрь 1975 г. и январь 2010 г. Данные предоставлены Всероссийским научно-исследовательским институтом гидрометеорологической информации – Мировым центром данных, Обнинск, Россия (<http://meteo.ru/data>). Описание массива данных дано в [8]. Отметим, что в Главах 2 и 3 верификация каждой модели временных рядов проводилась по данным минимум с 16 метеостанций, расположенных в различных климатических зонах (полный список метеостанций, данные с которых использовались в диссертационной работе, указан в Приложении 2).

Временные ряды рассматриваемых метеорологических характеристик (минимальной и максимальной температуры воздуха, индикаторов осадков) являются нестационарными, что наглядно проиллюстрировано на рисунках 2.1-2.3, где приведены графики выборочных средних значений минимальной и максимальной температуры, и индикатора осадков на годовом интервале (т.е. при $N = 365$). В связи с тем, что оценки средних значений проводятся по выборке относительно малого объёма, присутствует значительная статистическая погрешность. Кроме этого, при $N > 11$ размерность корреляционной матрицы R превышает объём выборки, по которой оцениваются элементы матрицы, в следствие чего матрица R оказывается вырожденной. Для уменьшения статистической погрешности оценки вероятностных характеристик процесса $(\vec{I}^T, \vec{A}^T, \vec{E}^T)$ при оценивании было использовано симметричное скользящее окно осреднения шириной $2L+1$ суток. В этом случае для оценки параметра в j -ый день использовались данные не только за этот день, но и за L предыдущих, и L последующих дней, а объём выборки возрос до $NY(2L+1)$ элементов. Формально, если X_j^i, Y_k^i – реальные значения 2 метеохарактеристик соответствующие j -ому и

k -ому дням ($j, k = 1, 2, \dots, N$) в i -ый год наблюдений ($i = 1, 2, \dots, NY$), то элемент $r_{XY}(j, k)$ матрицы R вычисляется по правилу

$$r_{XY}(j, k) = \frac{EX_j Y_k - EX_j \times EY_k}{\sqrt{DX_j} \sqrt{DY_k}},$$

$$EX_j Y_k = \frac{1}{NY(2L+1)} \sum_{i=1}^{NY} \sum_{m=-L}^L X_{j+m}^i Y_{k+m}^i,$$

$$EX_j = \frac{1}{NY(2L+1)} \sum_{i=1}^{NY} \sum_{m=-L}^L X_{j+m}^i,$$

$$DX_j = \frac{1}{NY(2L+1)} \sum_{i=1}^{NY} \sum_{m=-L}^L (X_{j+m}^i)^2.$$

Подробная информация о свойствах такой процедуры сглаживания приведена, например, в [3].

Естественным образом возникает вопрос о выборе ширины окна осреднения. Если выбрать окно малой ширины, то объем выборки возрастет недостаточно для того, чтобы существенно уменьшить статистическую погрешность оценки. Если же число L выбрать очень большим, то, осредняя данные по $2L+1$ дню, можно потерять временные изменения значения оцениваемого параметра, присущие реальному процессу. В данной работе ширина окна осреднения выбиралась по правилу

$$L = \underset{l \in Z^+}{\operatorname{argmin}} (NY(2l+1) > 3N). \quad (2.6)$$

Так, например, $L = 16$ при $N = 365$. На Рис. 2.1, 2.2 приведены оценки μ_j^I, μ_j^A и μ_j^E при использовании окна осреднения такой ширины. На Рис. 2.3 приведены оценки коэффициентов корреляции $r_{II}(1, 1+h), r_{AA}(1, 1+h), r_{EE}(1, 1+h)$ при таком же окне осреднения.

Приведём несколько примеров характеристик, использованных для верификации предложенной модели совместных временных рядов $(\vec{I}^T, \vec{A}^T, \vec{E}^T)$.

Рассмотрим такую характеристику как среднее число AN_0 дней в месяце, в которые минимальная за сутки температура ниже 0°C , а максимальная – выше. Эта

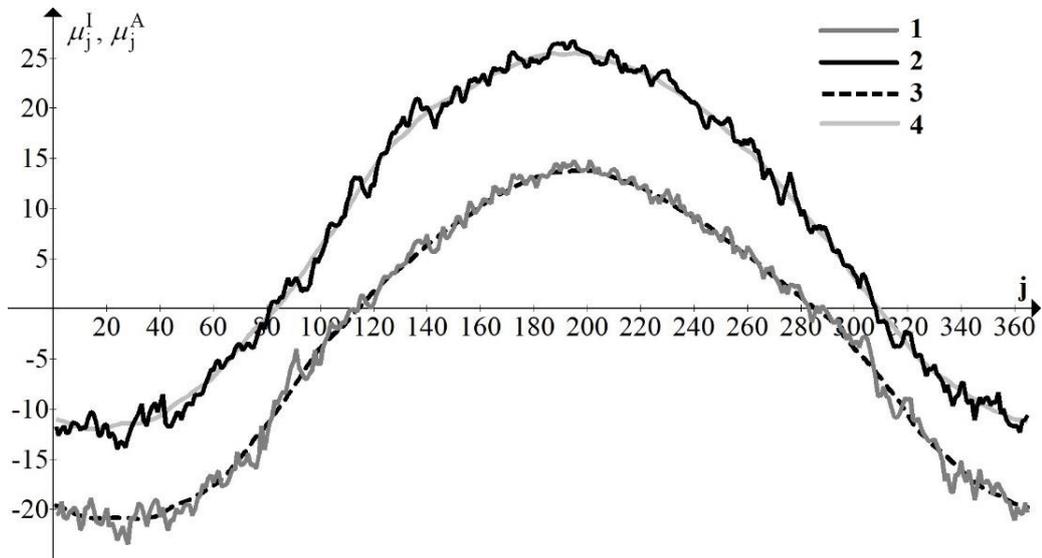


Рис. 2.1. Оценки математического ожидания минимальной и максимальной температуры. Кривые 1 и 2 – оценки μ_j^I и μ_j^A , построенные без использования окна осреднения. Кривые 3 и 4 – оценки μ_j^I и μ_j^A , построенные с использованием окна осреднения. Огурцово (Новосибирск).

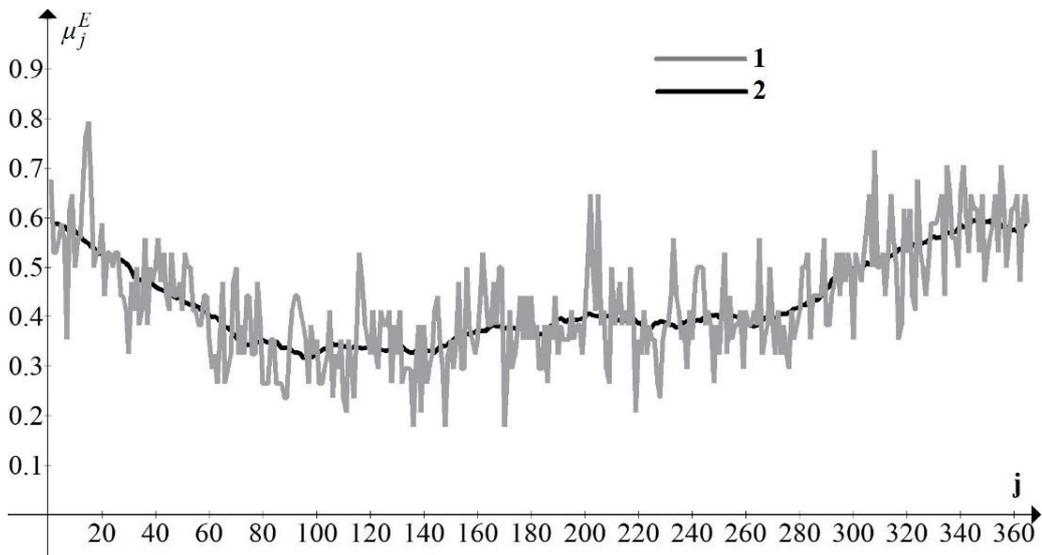


Рис. 2.2. Оценки математического ожидания μ_j^E индикаторов осадков. Кривая 1 – без использования окна осреднения. Кривая 2 – при использовании окна осреднения. Огурцово (Новосибирск).

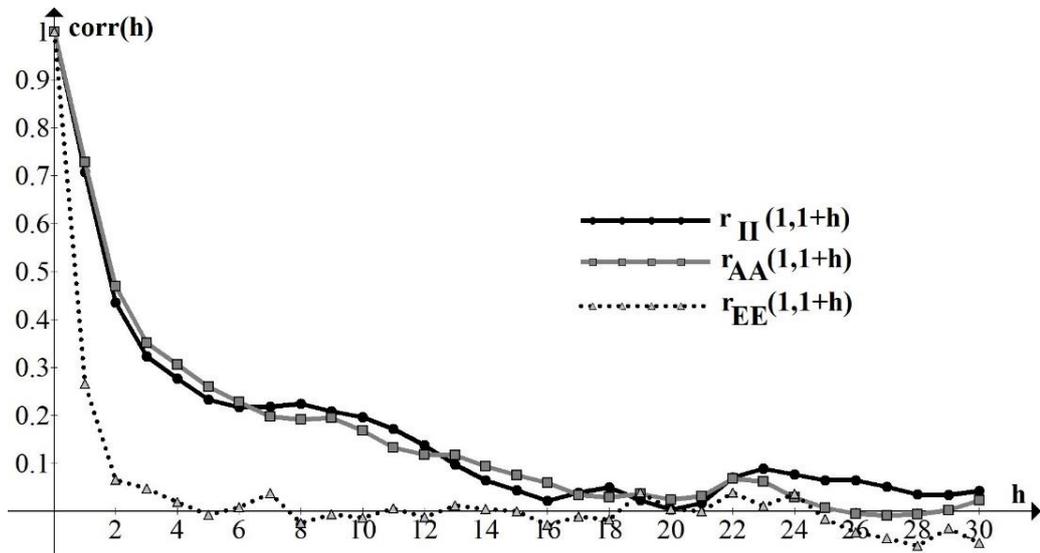


Рис. 2.3. Оценки коэффициентов корреляции. $r_{II}(1,1+h)$, $r_{AA}(1,1+h)$, $r_{EE}(1,1+h)$. Огурцово (Новосибирск), 1-30 января.

характеристика не является входным параметром модели, поэтому может быть использована для верификации. Моделирование проводилось при N равном числу дней в месяце и L выбранном из условия (2.6). В таблице 2.1 приведены значения оценок этой характеристики для различных месяцев, полученные по реальным данным и по ансамблю 10^6 модельных траекторий. Как видно из таблицы 2.1, модель воспроизводит данную характеристику с достаточной степенью точности. Здесь и всюду далее в Главах 2 и 3 оценки по реальным данным обозначены РД, а оценки по модельным траекториям – МД, σ – среднеквадратическое отклонение оценки рассматриваемой характеристики по выборке реальных данных имеющегося объёма. Значение σ находилось численно.

В таблице 2.2 приведены результаты оценивания ещё одной характеристики, использованной для верификации модели. Сравнение оценок по реальным и модельным данным среднего числа AN_k за летние месяцы временных интервалов длительностью ровно k суток, характеризующихся минимальной температурой выше 20°C и отсутствием осадков, показывает, что и эта характеристика реального процесса хорошо воспроизводится моделью. Моделирование проводилось при $N = 92$ (число дней в летние месяцы).

Таб. 2.1. Оценки среднего числа AN_0 дней в месяце с $I_j < 0, A_j > 0$. Санкт-Петербург.

Месяц	Среднее число дней AN_0	
	РД, $AN_0 \pm \sigma$	МД
Октябрь	4.7 ± 0.465	4.9
Ноябрь	8.9 ± 0.881	9.1
Декабрь	5.7 ± 0.564	5.3
Январь	9.3 ± 0.921	9.4
Февраль	7.7 ± 0.762	7.4
Март	16.3 ± 1.614	16.8

Таб. 2.2. Оценки среднего числа AN_k . Астрахань.

Длительность k , сутки	РД, $AN_k \pm 3\sigma$	МД
$k=1$	5.3 ± 0.909	5.72
$k=2$	2.0 ± 0.343	2.20
$k=4$	0.7 ± 0.120	0.76
$k=5$	0.6 ± 0.103	0.59
$k=8$	3.7 ± 0.635	4.18
$k=11$	4.5 ± 0.772	4.06

В общей сложности, для верификации этой модели были использованы более 15 различных характеристик. В целом, модель хорошо воспроизводит свойства реального процесса – более чем в 90% проведённых численных экспериментах оценки характеристик по модельным данным принадлежат доверительному интервалу шириной 2σ для соответствующих оценок по реальным данным.

Расширенная модель

Рассмотренную модель несложно расширить, заменив моделирование индикатора осадков \vec{E}^T моделированием количества $\vec{D}^T = (D_1, D_2, \dots, D_N)^T$ выпавших осадков:

$$D_j = E_j C_j, \quad j = 1, \dots, N,$$

где $\vec{C}^T = (C_1, C_2, \dots, C_N)^T$ – негауссовский процесс, описывающий количество выпавших за сутки осадков при условии их наличия. В этом случае моделируется процесс $(\vec{I}^T, \vec{A}^T, \vec{E}^T, \vec{C}^T)$ с корреляционной матрицей

$$\tilde{R} = \begin{pmatrix} R_{II} & R_{IA} & R_{IE} & R_{IC} \\ R_{AI} & R_{AA} & R_{AE} & R_{AC} \\ R_{EI} & R_{EA} & R_{EE} & O \\ R_{CI} & R_{CA} & O & R_{CC} \end{pmatrix},$$

где O – нулевая матрица размерности $N \times N$. Элементы $\tilde{g}_{IC}(i, j)$, $\tilde{g}_{AC}(i, j)$, $\tilde{g}_{CI}(i, j)$, $\tilde{g}_{CA}(i, j)$, $\tilde{g}_{CC}(i, j)$, $i, j = 1, 2, \dots, N$ соответствующей корреляционной матрицы \tilde{G} вспомогательного гауссовского процесса вычисляются в рамках метода обратных функций распределений.

Отметим, что мультипликативный подход к моделированию количества осадков, т.е. представление D_j в виде произведения индикатора наличия / отсутствия осадков и их количества при условии их наличия

$$D_j = E_j C_j, \quad j = 1, \dots, N$$

является одним из наиболее часто применяемых подходов при разработке стохастических «генераторов погоды» [60, 157, 188, 204].

Продемонстрируем на одном примере, что расширенная модель комплекса характеристик $(\vec{I}^T, \vec{A}^T, \vec{E}^T, \vec{C}^T)$ адекватно описывает реальные метеопроцессы. В качестве функции $F_i(x)$ одномерного распределения количества осадков при условии их наличия была взята аппроксимация выборочной функции распределения смесью двух гамма-распределений с параметрами, определенными

по алгоритму, предложенному в [47]. В таблице 2.3 приведены оценки вероятностей $P(lev)$ возникновения 5 последовательных дней с осадками больше lev мм. Оценки проводились по реальным данным и 10^8 модельных реализаций при $N = 31$.

Следует отметить, что, в отличие от известных «генераторов погоды» WGEN и LARS-WG [193], предложенная модель позволяет воспроизводить корреляционную структуру рядов рассматриваемых метеорологических характеристик, а также их изменчивость, связанную с годовым ходом реальных процессов.

Таб. 2.3. Оценки вероятности $P(lev)$ возникновения 5 последовательных дней с осадками больше lev мм. Огурцово (Новосибирск), июль.

lev	РД, $P(lev) \pm 3\sigma$	МД
10	0.18 ± 0.013	0.192
15	0.10 ± 0.015	0.098
20	0.06 ± 0.018	0.058
25	0.04 ± 0.022	0.042
30	0.03 ± 0.025	0.029

2.1.2. Моделирование совместных нестационарных временных рядов различных метеорологических параметров

В этом параграфе рассмотрены некоторые подходы к моделированию совместных нестационарных негауссовских временных рядов различных комбинаций следующих метеорологических параметров: приземной температуры воздуха, его относительной влажности, модуля скорости ветра и атмосферного давления. Предложенные подходы позволяют учесть при моделировании суточный ход реальных метеорологических процессов.

Перед тем как перейти к детальному описанию моделей, отметим, что для моделирования временных рядов температуры воздуха с различным временным разрешением используют как детерминированные гидродинамические, так и стохастические модели (см., например, [98, 113, 126, 131, 156]). Изучение относительной влажности атмосферного воздуха, особенно в условиях глобального потепления, чаще всего проводится на основе уравнений гидродинамики и статистического анализа реальных наблюдений, стохастический подход практически не используется (обзор работ с описанием используемых подходов представлен в [108]). Заметим, что временные ряды относительной влажности часто моделируются с шагом 24 часа. Для получения значений относительной влажности с меньшим временным шагом обычно используются различные методы интерполяции, однако применение этих методов часто приводит к существенным ошибкам [114]. Для моделирования временных рядов скорости ветра различного разрешения разработано большое количество стохастических моделей [79, 93, 102, 104, 167]. В ряде работ под скоростью ветра понимают вектор скорости ветра, и моделируют либо его ортогональные компоненты, либо модуль скорости ветра и его направление. В других работах под скоростью ветра понимают лишь модуль скорости ветра.

Описание модели периодически коррелированного случайного процесса

Опишем детально принцип построения модели совместных временных рядов температуры воздуха (измеряемой в градусах Цельсия), его относительной влажности (в долях единицы) и атмосферного давления (единица измерения – гПа). Модели совместных рядов других комбинаций метеопараметров (например, температуры воздуха и модуля скорости ветра) строятся по такому же принципу.

Необходимо построить модель совместного временного ряда

$$\begin{pmatrix} \vec{T} \\ \vec{H} \\ \vec{P} \end{pmatrix} = \left(\vec{T}^T, \vec{H}^T, \vec{P}^T \right)^T = (T_1, T_2, \dots, T_{8d}, H_1, H_2, \dots, H_{8d}, P_1, P_2, \dots, P_{8d})^T$$

на временном интервале длительностью d суток с 3-часовым разрешением (при таком разрешении каждый метеопараметр характеризуется 8 значениями в сутки). Для того чтобы учесть суточный ход реальных метеопроцессов (наиболее сильно выраженный у рядов температуры воздуха) воспользуемся предположением о том, что каждый из векторов-столбцов $\vec{T}, \vec{H}, \vec{P}$ является периодически коррелированным случайным процессом с периодом равным 1 суткам. Для того чтобы полностью определить вектор $(\vec{T}^T, \vec{H}^T, \vec{P}^T)^T$ необходимо задать одномерные плотности распределения $t_i(x), h_i(x), p_i(x), i = 1, 2, \dots, 8d$ температуры, относительной влажности и атмосферного давления и корреляционную матрицу

$$R = \begin{pmatrix} R_T & R_{TH} & R_{TP} \\ R_{HT} & R_H & R_{HP} \\ R_{PT} & R_{PH} & R_P \end{pmatrix}.$$

размерности $(3 \cdot 8d) \times (3 \cdot 8d)$. Здесь $R_X, X \in \{T, H, P\}$ – автокорреляционная матрица метеопараметра X , а $R_{XY}, X, Y \in \{T, H, P\}$ – взаимные корреляционные матрицы метеопараметров X, Y .

Определим плотности распределения $t_k(x), h_k(x), p_k(x), k = 1, 2, \dots, 8$ первых восьми компонент векторов $\vec{T}, \vec{H}, \vec{P}$. Плотности распределения остальных компонент определяются по правилу

$$\begin{aligned} t_i(x) &= t_{8j+k}(x) = t_k(x), \\ h_i(x) &= h_{8j+k}(x) = h_k(x), \\ p_i(x) &= p_{8j+k}(x) = p_k(x), \\ i &= 9, \dots, 8d, \quad k = 1, 2, \dots, 8, \quad j = 1, \dots, d-1. \end{aligned}$$

При таком определении плотностей распределения будет выполнено условие периодичности функций математического ожидания и дисперсии, выполнение которого необходимо при использовании предположения о периодической коррелированности векторов $\vec{T}, \vec{H}, \vec{P}$.

Каждая из подматриц $R_X, R_{XY}, X, Y \in \{T, H, P\}$ матрицы R является блочно-тёплицевой с блоками размерности 8×8 . Ниже будет продемонстрировано, что такая структура матрицы R позволяет существенно уменьшить трудоёмкость моделирования рассматриваемого вектора по сравнению со случаем, когда предположение о периодической коррелированности не используется. Далее модель вектора $\left(\bar{T}^T, \bar{H}^T, \bar{P}^T\right)^T$ с корреляционной матрицей R и плотностями одномерных распределений $t_i(x), h_i(x), p_i(x), i = 1, 2, \dots, 8d$ будет обозначаться как модель ПКСП.

Плотности одномерных распределений метеопараметров и корреляционная матрица процесса

Для построения стохастической модели использование выборочных одномерных распределений метеопараметров нецелесообразно, поскольку выборочные распределения не имеют хвостов, и, следовательно, не позволяют оценивать вероятности возникновения экстремальных значений метеозлементов. В связи с этим, следует аппроксимировать выборочные распределения какими-либо аналитическими плотностями, которые с одной стороны, не сильно искажают вид плотности распределения и его моменты, а с другой – обладают хвостами.

Численные эксперименты показали, что на всех рассмотренных метеостанциях для аппроксимации плотностей одномерных распределений температуры в k -ый срок измерений ($k = 1, 2, \dots, 8$) лучше всего подходят смеси плотностей двух нормальных распределений

$$t_k(x) = \theta_k \frac{1}{b_{k1} \sqrt{2\pi}} \exp\left(-\frac{(x - a_{k1})^2}{2b_{k1}^2}\right) + (1 - \theta_k) \frac{1}{b_{k2} \sqrt{2\pi}} \exp\left(-\frac{(x - a_{k2})^2}{2b_{k2}^2}\right),$$

$$0 < \theta_k < 1, k = 1, 2, \dots, 8.$$

Параметры $\theta_k, a_{k1}, b_{k1}^2, a_{k2}, b_{k2}^2$ аппроксимирующих смесей для каждого срока измерений выбирались по алгоритму, приведённому в [40]. Алгоритм позволяет выбирать параметры смеси $t_k(x)$ таким образом, чтобы среднее значение,

дисперсия и коэффициент асимметрии случайной величины с распределением $t_k(x)$ совпадали с соответствующими выборочными значениями, оцененными по реальным данным, и при этом значение функционала Пирсона было минимальным среди всех аппроксимаций заданного вида с заданными первыми 3 моментами. Отметим, что для определения 5 параметров плотности $t_k(x)$ не следует использовать метод моментов (по реальным данным моменты высоких порядков оцениваются с большой статистической погрешностью), а параметры, полученные при применении метода максимального правдоподобия, отличаются от $\theta_k, a_{k1}, b_{k1}^2, a_{k2}, b_{k2}^2$, определённых по алгоритму из [40], менее, чем на 10^{-4} . На рисунке 2.4 приведены выборочные гистограммы плотностей одномерных распределений температуры для 2 сроков и соответствующие им аппроксимирующие плотности $t_k(x)$. Построение гистограмм проводилось по данным метеонаблюдений, собранных на метеостанции в период с 1993 по 2015 (т.е. за $NY = 23$ года).

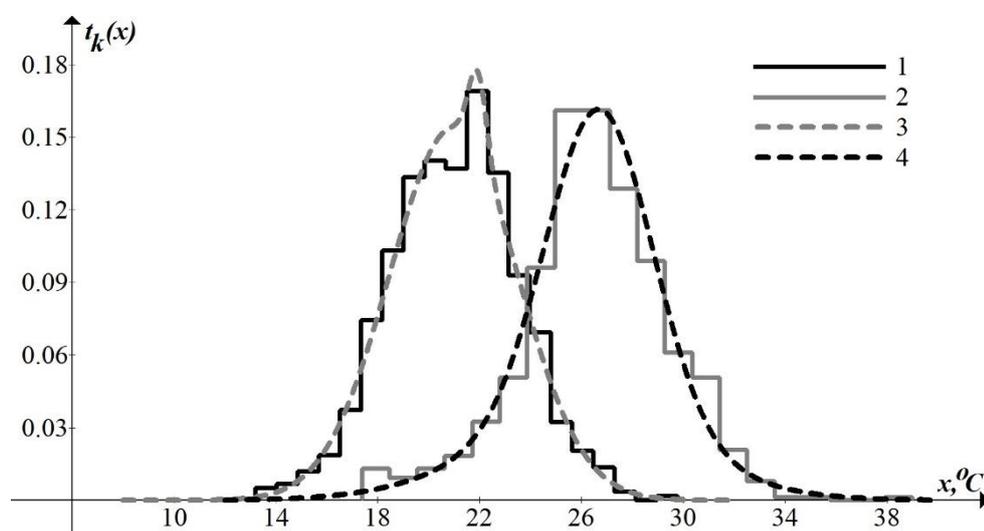


Рис. 2.4. Выборочные гистограммы приземной температуры воздуха в Сочи, 1-10 июля. Кривые 1 и 2 – гистограммы в сроки измерений $k = 1$ и $k = 6$, соответственно. Кривые 3 и 4 – аппроксимирующие плотности распределений $t_1(x)$ и $t_6(x)$.

Аппроксимировать одномерные распределения относительной влажности значительно сложнее. Дело в том, что если вид распределения температуры не существенно зависит от местоположения метеостанции, то вид распределений влажности сильно меняется от станции к станции (на рисунке 2.5 приведены графики гистограмм относительной влажности на 3 метеостанциях). Соответственно, для каждой станции необходимо подбирать аппроксимирующие плотности $h_k(x)$ подходящего вида. Так, например, бета-распределение достаточно хорошо аппроксимирует распределение относительной влажности в Сочи, но при этом оно не подходит для описания распределения этого метеоэлемента в Томске. В зависимости от станции, выборочные распределения относительной влажности в рамках рассматриваемой в диссертационной работе модели, аппроксимируются бета-распределением, смесью бета-распределений или смесью усеченных на $[0,1]$ нормальных распределений.

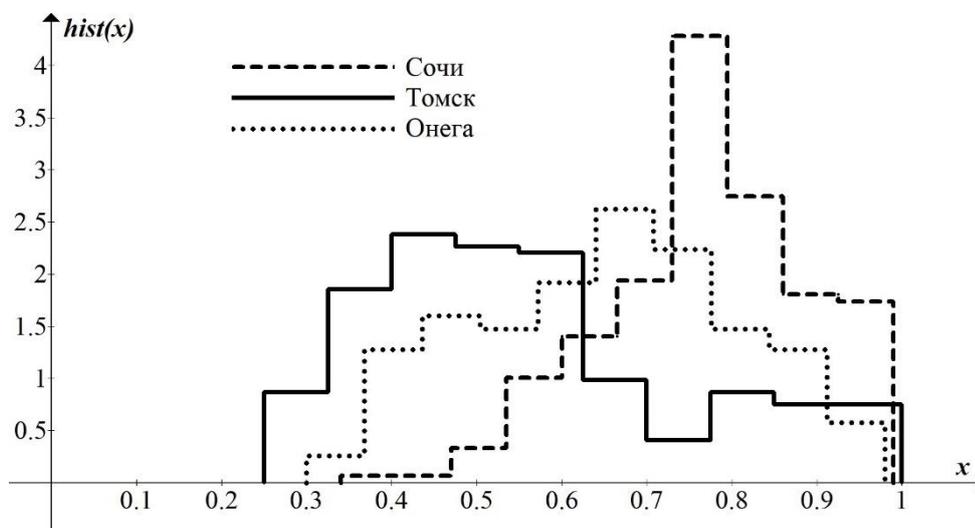


Рис. 2.5. Гистограммы одномерных распределений относительной влажности, $k = 3$. 1-10 июня.

Для аппроксимации одномерных распределений атмосферного давления в данной работе была использована смесь $p_k(x)$ двух усеченных на $(0, +\infty)$ нормальных распределений. Следует отметить, что для ряда метеостанций в

качестве аппроксимирующих плотностей распределения хорошо подходят смеси гамма-распределений

$$p_k(x) = \theta_k \frac{x^{\alpha_{k1}-1} \exp\left(-\frac{x}{\beta_{k1}}\right)}{\Gamma(\alpha_{k1}) \beta_{k1}^{\alpha_{k1}}} + (1-\theta_k) \frac{x^{\alpha_{k2}-1} \exp\left(-\frac{x}{\beta_{k2}}\right)}{\Gamma(\alpha_{k2}) \beta_{k2}^{\alpha_{k2}}}, \quad x > 0$$

$$0 < \theta_k < 1, \quad k = 1, 2, \dots, 8.$$

Однако, параметры гамма-распределений таковы, что при дальнейшем моделировании рядов с требуемыми распределениями методом обратных функций распределения возникают вычислительные проблемы, связанные с вычислением значений соответствующих функций распределения.

Замечание 2.4. В тех случаях, когда наряду с другими метеопараметрами моделировался модуль скорости ветра $\vec{V}^T = (V_1, V_2, \dots, V_{8d})^T$, в качестве плотности его одномерного распределения $w_k(x)$, как и в работах [102, 122], использовалось плотности распределения Вейбулла

$$w_k(x) = \frac{\nu_k}{\lambda_k} \left(\frac{x}{\lambda_k}\right)^{\nu_k-1} \exp\left(-\left(\frac{x}{\lambda_k}\right)^{\nu_k}\right), \quad x \geq 0, \quad k = 1, 2, \dots, 8.$$

В качестве матрицы R для построения стохастической модели используем выборочную матрицу, оцененную по реальным данным наблюдений за NY лет. Как и ранее, должно выполняться условие, что объём выборки превышает размерность матрицы R (напомним, что матрица R имеет размерность $24d \times 24d$). Отметим, что использование предположения о периодической коррелированности метеорологических процессов как при оценивании матрицы R , так и при построении выборочных гистограмм и оценки моментов одномерных распределений, увеличивает объём выборки до размера $d \times NY$. В случае, если данные наблюдений собраны за период длительностью более 24 лет, требуемое условие выполняется даже если никакое дополнительное окно осреднения не используется. При меньших значениях NY (на всех рассмотренных метеостанциях $NY > 15$) достаточно использовать окно шириной $2L+1=3$. На рисунках 2.6, 2.7 в

качестве примеров приведены графики оценок элементов $r_T(i, i+h)$ корреляционной матрицы R_T и элементов $r_T(1, 1+h)$, $r_H(1, 1+h)$ матриц R_T , R_H , соответственно.

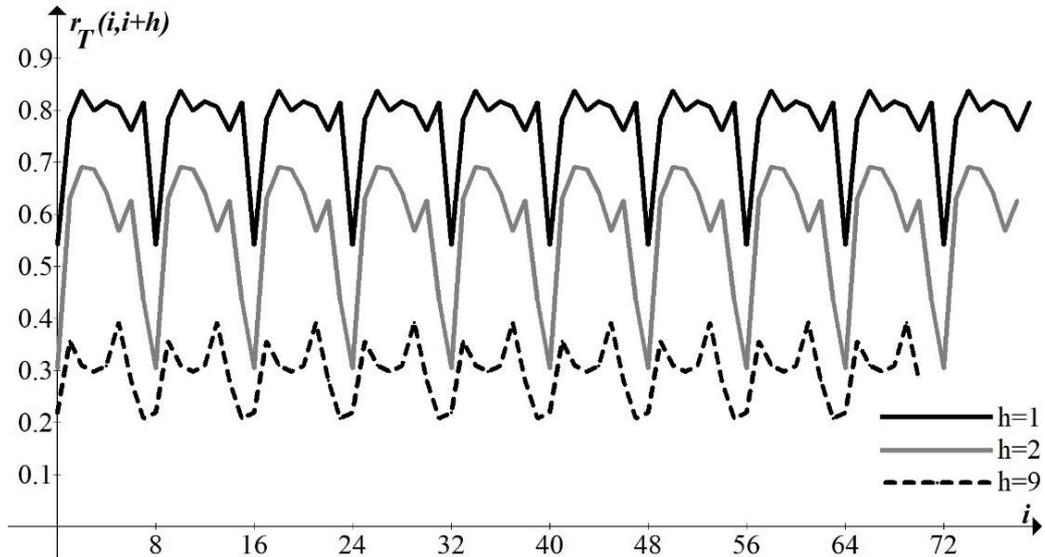


Рис. 2.6. Оценки элементов $r_T(i, i+h)$ корреляционной матрицы R_T . Томск, 1-10 июня.

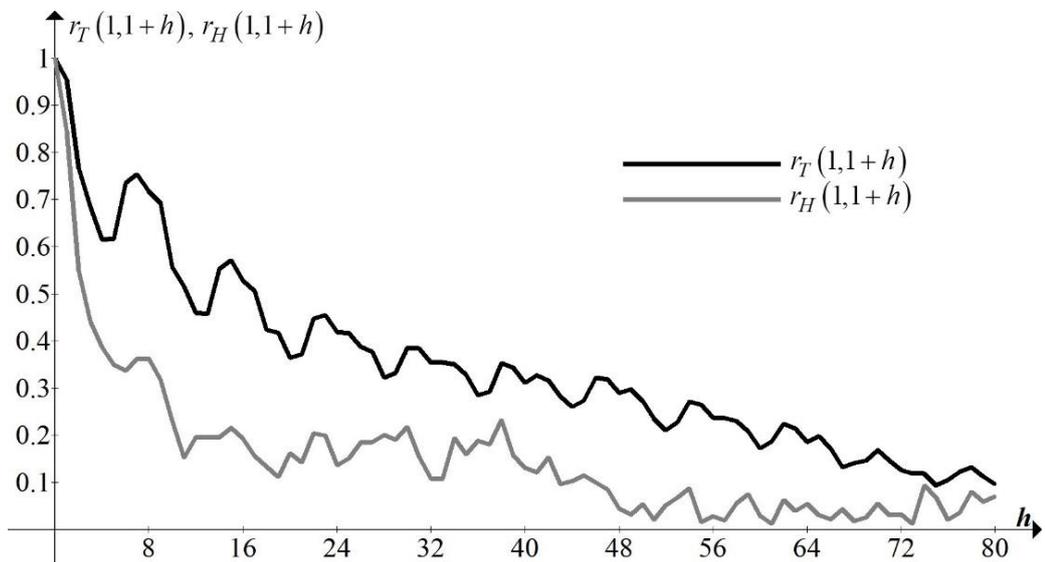


Рис. 2.7. Оценки элементов $r_T(1, 1+h)$, $r_H(1, 1+h)$ матриц R_T , R_H , соответственно. Сочи, 1-10 июля.

Алгоритм моделирования

Моделирование временного ряда $\left(\bar{T}^T, \bar{H}^T, \bar{P}^T\right)^T$ с одномерными плотностями распределений $t_k(x), h_k(x), p_k(x), k=1,2,\dots,8$ и корреляционной матрицей R осуществлялось методом обратных функций распределения [53, 64, 66] по Алгоритму 2.1.

Алгоритм 2.1.

1. С помощью функций одномерных распределений $FT_k(x), FH_k(x), FP_k(x), k=1,2,\dots,8$ соответствующих вышеописанным плотностям $t_k(x), h_k(x), p_k(x), k=1,2,\dots,8$ вычисляется корреляционная матрица R_G вспомогательного гауссовского процесса

$$\left(\bar{\xi}_T^T, \bar{\xi}_H^T, \bar{\xi}_P^T\right)^T = \left(\xi_{T1}, \dots, \xi_{T8d}, \xi_{H1}, \dots, \xi_{H8d}, \xi_{P1}, \dots, \xi_{P8d}\right)^T.$$

Элементы $r_{G,X}(i,l), r_{G,XY}(i,l)$ подматриц $R_{G,X}, R_{G,XY}, X, Y \in \{T, H, P\}$ матрицы R_G , соответствующих подматрицам $R_X, R_{XY}, X, Y \in \{T, H, P\}$, суть решения уравнений

$$r_X(i,l) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} FX_k^{-1}(\Phi(y)) FX_n^{-1}(\Phi(z)) \phi(y, z, r_{G,X}(i,l)) dy dz,$$

$$r_{XY}(i,l) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} FX_k^{-1}(\Phi(y)) FY_n^{-1}(\Phi(z)) \phi(y, z, r_{G,XY}(i,l)) dy dz,$$

$$i = 8j + k, l = 8m + n,$$

$$j, m = 0, 1, \dots, d-1, k, n = 1, \dots, 8.$$

Здесь $\Phi(x)$ – функция стандартного нормального распределения, $\phi(y, z, \rho)$ – плотность двумерного гауссовского вектора с нулевым математическим ожиданием и единичной дисперсией у каждой из компонент и коэффициентом корреляции между компонентами равным ρ .

2. Моделируются траектории гауссовского процесса с корреляционной матрицей R_G . В данной работе моделирование производилось методом, основанным на спектральном разложении $R_G = VDV^T$ симметричной матрицы R_G , где диагональная матрица D составлена из собственных чисел $d_i, i=1,2,\dots,24d$ матрицы R_G , а матрица $V = (\vec{v}_1, \dots, \vec{v}_{24d})$ – из нормированных собственных векторов \vec{v}_i этой матрицы, соответствующих d_i .

3. Проводится преобразование каждой траектории гауссовского процесса в траектории негауссовского процесса по правилу

$$T_i = FT_k^{-1}(\Phi(\xi_{Ti})), H_i = FH_k^{-1}(\Phi(\xi_{Hi})), P_i = FP_k^{-1}(\Phi(\xi_{Pi})), \\ i = 8j + k, j = 0, 1, \dots, d-1, k = 1, \dots, 8.$$

При численной реализации Алгоритма 2.1 в тех случаях, когда матрица R_G оказывалась не положительно определённой, проводилась регуляризация матрицы R_G . Регуляризация осуществлялась с помощью алгоритма, описанного в [52]. В рамках этого алгоритма, матрица R_G представлялась в виде $R_G = VDV^T$. Отрицательные элементы d_i диагональной матрицы D заменялись на $|d_i|$, а $d_j = 0$ – на $d_j = \varepsilon > 0$, при этом формировалась диагональная матрица D' . Дальнейшее моделирование осуществлялось с корреляционной матрицей $R_G' = VD'V^T$. Применение регуляризации привело к тому, что выборочная корреляционная матрица, оцененная по траекториям, полученным на последнем шаге Алгоритма 2.1 отличается от матрицы R . В качестве иллюстрации, на рисунке 2.8 приведены оценки коэффициентов корреляции $corr(T_1, T_{1+h})$ and $corr(T_h, T_{h+4})$ по реальным данным (т.е. первая строка и четвёртая диагональ матрицы R_T) и оценки соответствующих коэффициентов корреляции по 10^8 модельных траекторий.

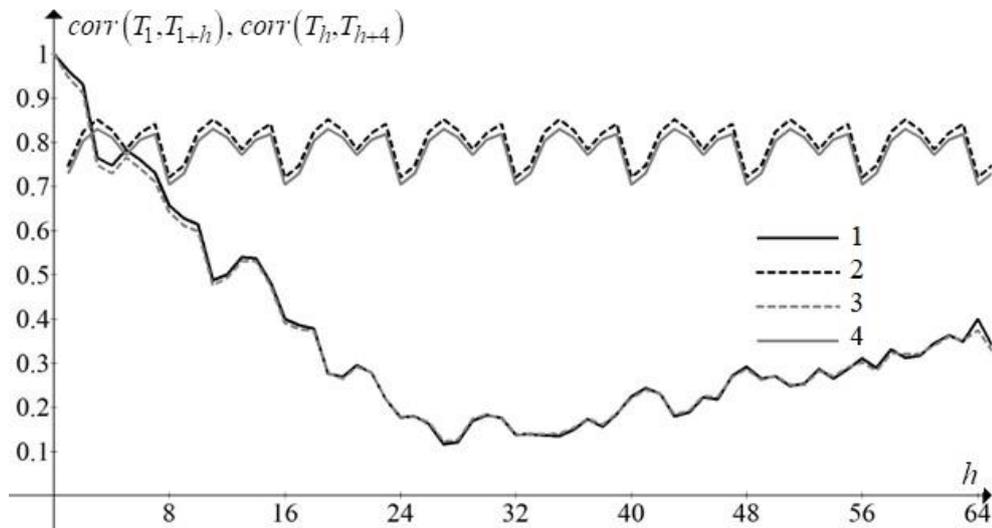


Рис. 2.8. Оценки коэффициентов корреляции $\text{corr}(T_1, T_{1+h})$ и $\text{corr}(T_h, T_{h+4})$. Кривые 1 и 2 – по реальным данным, кривые 3 и 4 – по модельным траекториям. Сочи, 1-8 февраля.

Описание модели нестационарного процесса

Наряду с рассмотренной моделью временных рядов $(\vec{T}^T, \vec{H}^T, \vec{P}^T)^T$ как периодически коррелированного процесса, была построена модель рядов $(\vec{T}^T, \vec{H}^T, \vec{P}^T)^T$ как нестационарного процесса с плотностями одномерных распределений $t_i^n(x), h_i^n(x), p_i^n(x), i=1, 2, \dots, 8d$ рассматриваемых метеопараметров и корреляционной матрицей R^n , не обладающей блочно-тёплицевой структурой. Принципы определения параметров указанных плотностей и построения траекторий такого процесса совпадают с принципами, использованными при разработке модели ПКСП. Назовём эту модель моделью НСП. При оценки корреляционной матрицы R^n и построении гистограмм распределения по реальным данным было использовано симметричное скользящее окно осреднения минимальной ширины $2L+1$, удовлетворяющей условию $24d < (2L+1)NY$.

Сравнение трудоёмкости моделирования

Численная реализация модели НСП существенно более трудоёмка, чем реализация модели ПКСП.

Действительно, при численной реализации модели НСП на первом шаге алгоритма моделирования на основе метода обратных функций необходимо вычислить $num_1 = 12d(24d - 1)$ элементов симметричной матрицы R_G^n , лежащих выше её главной диагонали (здесь R_G^n – корреляционная матрица вспомогательного гауссовского процесса). При этом, при реализации Алгоритма 2.1, как уже упоминалось выше, блочно-тёплицева структура подматриц матрицы R (и, соответственно, такая же структура матрицы R_G) позволяет сократить количество num_2 вычисляемых элементов. В силу того, что

$$R_{XY} = (R_{YX})^T, X, Y \in \{T, H, P\}$$

и блочно-тёплицевости $R_X, R_{XY}, X, Y \in \{T, H, P\}$ достаточно определить элементы первой блочной строки матриц $R_T, R_H, R_P, R_{TH}, R_{TP}, R_{HP}$. Размерность каждой из блочных строк есть $8 \times 8d$, т.е. всего надо определить $6 \times 8 \times 8d = 384d$ элементов. Поскольку на главных диагоналях матриц R и R_G стоят единицы (их $24d$ штук), число вычисляемых элементов находится по формуле

$$num_2 = 384d - 24d = 360d.$$

Следовательно, справедлива следующая лемма.

Лемма 2.4.

$$\frac{num_1}{num_2} = \frac{12d(24d - 1)}{360d} \approx 0.8d.$$

Из вышеприведённого утверждения следует, что, например, при $d = 10$ суткам, при реализации модели ПКСП необходимо вычислить примерно в 8 раз меньше элементов матрицы R_G , чем при реализации модели НСП.

Аналогичным образом доказывается лемма 2.5.

Лемма 2.5. При моделировании вектора 2 метеопараметров верно соотношение

$$\frac{num_1}{num_2} = \frac{8d(16d-1)}{3 \times 8 \times 8d - 16d} \approx 0.73d,$$

а при моделировании вектора 4 метеопараметров –

$$\frac{num_1}{num_2} = \frac{16d(32d-1)}{10 \times 8 \times 8d - 32d} \approx 0.84d.$$

Верификация моделей

Была проведена верификация моделей ПКСП и НСП. Отметим, что для верификации необходимо сравнивать оценки по реальным и модельным данным тех характеристик рассматриваемого процесса, которые, с одной стороны, не являются входными параметрами модели, а с другой – надёжно оцениваются по выборке реальных данных ограниченного объёма. Ниже приведены несколько примеров таких характеристик.

Одной из использованных для верификации характеристик процесса была вероятность p_j того, что хотя бы 1 раз за рассматриваемый временной интервал длительностью d дней температура воздуха и его относительная влажность были ниже / выше заданных уровней $lt^{\circ}C$ и lh , соответственно. В таблице 2.4 для некоторых уровней $lt^{\circ}C$ и lh приведены оценки этой характеристики, полученные по реальным метеоданным и модельным траекториям. Проведенные численные эксперименты показали, что обе модели хорошо воспроизводят данную характеристику реального процесса (более чем в 90% экспериментов оценки по модельным траекториям попали в доверительный интервал $(p_j - 3\sigma, p_j + 3\sigma)$ оценок по реальным данным). Если рассматривать интервал $(p_j - \sigma, p_j + \sigma)$, то модель ПКСП оказывается предпочтительнее – оценки по траекториям этой модели попадают в интервал $(p_j - \sigma, p_j + \sigma)$ почти вдвое чаще, чем оценки по траекториям модели НСП.

Таб. 2.4. Оценки вероятностей p_j того, что хотя бы 1 раз за рассматриваемый временной интервал длительностью d дней температура была ниже уровня lt °C, а относительная влажность – выше lh . Сочи, 1-10 февраля.

lt °C	lh	РД, $p_j \pm \sigma$	Модель ПКСП	Модель НСП
14	0.2	0.952 ± 0.020	0.954	0.947
14	0.5	0.842 ± 0.028	0.817	0.809
10	0.2	0.822 ± 0.043	0.824	0.796
10	0.5	0.769 ± 0.041	0.733	0.711
6	0.8	0.267 ± 0.038	0.344	0.310
2	0.2	0.164 ± 0.040	0.172	0.179
-2	0.5	0.017 ± 0.012	0.016	0.025
-2	0.8	0.005 ± 0.011	0.015	0.022

Ещё одной характеристикой, которая была применена для верификации моделей было среднее число AN дней на рассматриваемом временном интервале длительностью d суток, в которые минимальное за сутки давление превышало уровень lp гПа. В таблице 2.5 приведены оценки этой характеристики по реальным данным и модельным траекториям. Эта характеристика также хорошо воспроизводится обеими моделями.

При верификации моделей также проводилось сравнение оценок по реальным и модельным данным вероятности pc изменения температуры воздуха более чем на ct °C не более чем за m измерений, т.е.

$$pc = P\left(\exists i, j : |T_i - T_j| > ct, |i - j| \leq m\right).$$

В таблице 2.6 приведены оценки pc для $m = 8$. В доверительный интервал $(pc - \sigma, pc + \sigma)$ оценок по реальным данным оценки по модельным траекториям, полученным по обеим моделям, попали менее, чем в 50% экспериментов, а в

интервал $(pc - 3\sigma, pc + 3\sigma)$ – в 82 % экспериментов. Можно сказать, что эта характеристика удовлетворительно воспроизводится обеими моделями.

Таб. 2.5. Оценки среднего числа AN дней на рассматриваемом временном интервале длительностью d суток, в которые минимальное за сутки давление превышало lp гПа. Сочи, 1-10 февраля.

lp гПа	РД, $AN \pm \sigma$	Модель ПКСП	Модель НСП
990	9.565 ± 0.248	9.644	9.620
996	8.261 ± 0.562	8.425	8.177
1002	6.609 ± 0.749	6.272	5.948
1008	4.783 ± 0.763	4.160	4.030
1014	2.478 ± 0.631	2.279	2.194
1020	0.870 ± 0.365	0.814	0.731

Таб. 2.6. Оценки вероятности pc . Сочи, 1-10 февраля.

ct °C	РД, $pc \pm 3\sigma$	Модель ПКСП	Модель НСП
1	0.996 ± 0.003	1.000	0.999
5	0.507 ± 0.102	0.574	0.492
7	0.176 ± 0.075	0.242	0.184
11	0.004 ± 0.015	0.017	0.008
13	0.001 ± 0.003	0.003	0.001

Результаты сравнения оценок других характеристик временных рядов $(\bar{T}^T, \bar{H}^T, \bar{P}^T)^T$, а также рядов $(\bar{T}^T, \bar{V}^T)^T$, $(\bar{T}^T, \bar{P}^T)^T$, $(\bar{T}^T, \bar{H}^T)^T$, \bar{T}^T приведены в работах автора [28, 136, 140, 147, 149]. Для всех этих рядов обе модели прошли верификацию и могут быть использованы для численного исследования тех характеристик метеопроцессов, которые не могут быть исследованы по реальным данным. Кроме этого, построенные модели могут быть использованы при

исследовании временных рядов различных биоклиматических индексов (что будет продемонстрировано в параграфе 3.2 Главы 3).

Следует отметить, что модель ПКСП, в отличие от модели НСП, применима только на относительно коротких временных интервалах длительностью d суток (в пределах 11-15 суток). На более длинных временных интервалах у рядов реальных метеопроцессов проявляется сезонный ход, существенно влияющих на характеристики процессов, а модель ПКСП этот ход не учитывает. В качестве иллюстрации приведём следующий пример. В таблице 2.7 приведены оценки среднесуточной температуры воздуха AT_i в i -ый день рассматриваемого временного интервала длительностью d дней. Из таблицы 2.7 отчётливо видно, что модель ПКСП не смогла воспроизвести характерный для рассматриваемого временного интервала с 1 по 31 мая тренд на повышение среднесуточной температуры.

Таб 2.7. Оценки среднесуточной температуры. Сочи, 1-31 мая.

i	РД, $AT_i \pm 3\sigma$	Модель ПКСП	Модель НСП
1	13.87 ± 0.86	16.01	13.96
11	15.99 ± 0.93	16.06	15.83
21	18.47 ± 0.92	15.99	18.01
31	19.45 ± 0.87	16.04	19.52

Заметим, что некоторые характеристики реальных процессов хорошо воспроизводятся моделью ПКСП даже на длительных временных интервалах. В качестве примера в таблице 2.8 приведены оценки среднего числа D дней в месяце, в которые минимальная температура за сутки превысила заданный уровень $lev^{\circ}C$. Наряду с оценками D по реальным данным и по модельным траекториям в этой таблице приведены результаты оценки D по траекториям широко используемого стохастического «генератора погоды» WGEN [188, 193]. Видно, что предложенные в данной работе модели не уступают WGEN по точности воспроизведения

рассматриваемой характеристики. Численная реализация WGEN осуществлялась согласно исходному коду, приведённому в [188].

Таб. 2.8. Оценки среднего числа D . Сочи, 1-31 июля.

$lev^{\circ}C$	РД, $D \pm \sigma$	Модель ПКСП	Модель НСП	WGEN
14	30.78 ± 0.14	30.81	30.76	30.80
16	29.83 ± 0.19	29.63	29.64	29.60
18	25.70 ± 0.24	25.31	25.98	25.30
20	16.74 ± 0.28	16.70	16.79	16.70
22	7.09 ± 0.32	6.96	7.03	7.07
24	1.70 ± 0.39	1.63	1.62	1.60

2.2. Моделирование неоднородного нестационарного пространственно-временного поля комплекса метеорологических параметров на сети метеостанций

В этом параграфе рассматривается стохастический «генератор погоды», предназначенный для моделирования совместного пространственно-временного поля приземной температуры воздуха, вектора скорости ветра и количества осадков на сети метеорологических станций, расположенных в южной части Байкальской природной территории. Предложенный «генератор погоды» позволяет моделировать поля температуры воздуха и компонент скорости ветра с трёхчасовым разрешением, а поле осадков – с 12-часовым шагом по времени. Для того чтобы учесть суточный ход реальных метеопроцессов, при построении модели использовано предположение о периодической коррелированности поля по

времени с периодом равным 1 суткам. Кроме этого, использовано предположение о пространственной неоднородности поля.

Описание модели

Рассматривается сеть из $NS = 8$ метеостанций, местоположение которых указано на рисунке 2.9. Отметим, что в метеоданных содержатся сведения о модуле скорости ветра W м/с и его направлении φ , указываемого в румбах, а моделируются значения ортогональных компонент вектора \vec{V} скорости ветра

$$\begin{aligned} u &= W \cos \tilde{\varphi}^o, \\ v &= W \sin \tilde{\varphi}^o, \end{aligned}$$

где $\tilde{\varphi}^o = 0.08889\varphi$.

Введём следующие обозначения. Пусть $k, k = 1, 2, \dots, NS$ – номер метеостанции, $i, i = 1, 2, \dots, ND$ – номер дня на интервале длительностью ND дней. Обозначим

$$\vec{\xi}_i^k = \left(T_{1,i}^k, \dots, T_{8,i}^k, u_{1,i}^k, v_{1,i}^k, \dots, u_{8,i}^k, v_{8,i}^k, P_{1,i}^k, P_{2,i}^k \right)^T$$

вектор-столбец, компоненты $T_{j,i}^k, j = 1, 2, \dots, 8$ которого есть значения температуры воздуха в j -ый срок измерений в i -ый день на k -ой метеостанции, $u_{j,i}^k, v_{j,i}^k$ – компоненты вектора скорости ветра в этот же момент времени на k -ой метеостанции, а $P_{1,i}^k$ и $P_{2,i}^k$ – количество выпавших осадков за первую и вторую половину i -ых суток на k -ой метеостанции, соответственно. Вектор $\vec{\xi}_i^k$ состоит из $NC = 26$ компонент. Обозначим также

$$\vec{\xi}_i = \left(\vec{\xi}_i^1, \dots, \vec{\xi}_i^{NS} \right)^T$$

вектор-столбец значений рассматриваемых метеопараметров на всех метеостанциях во все сроки измерений в i -ый день и

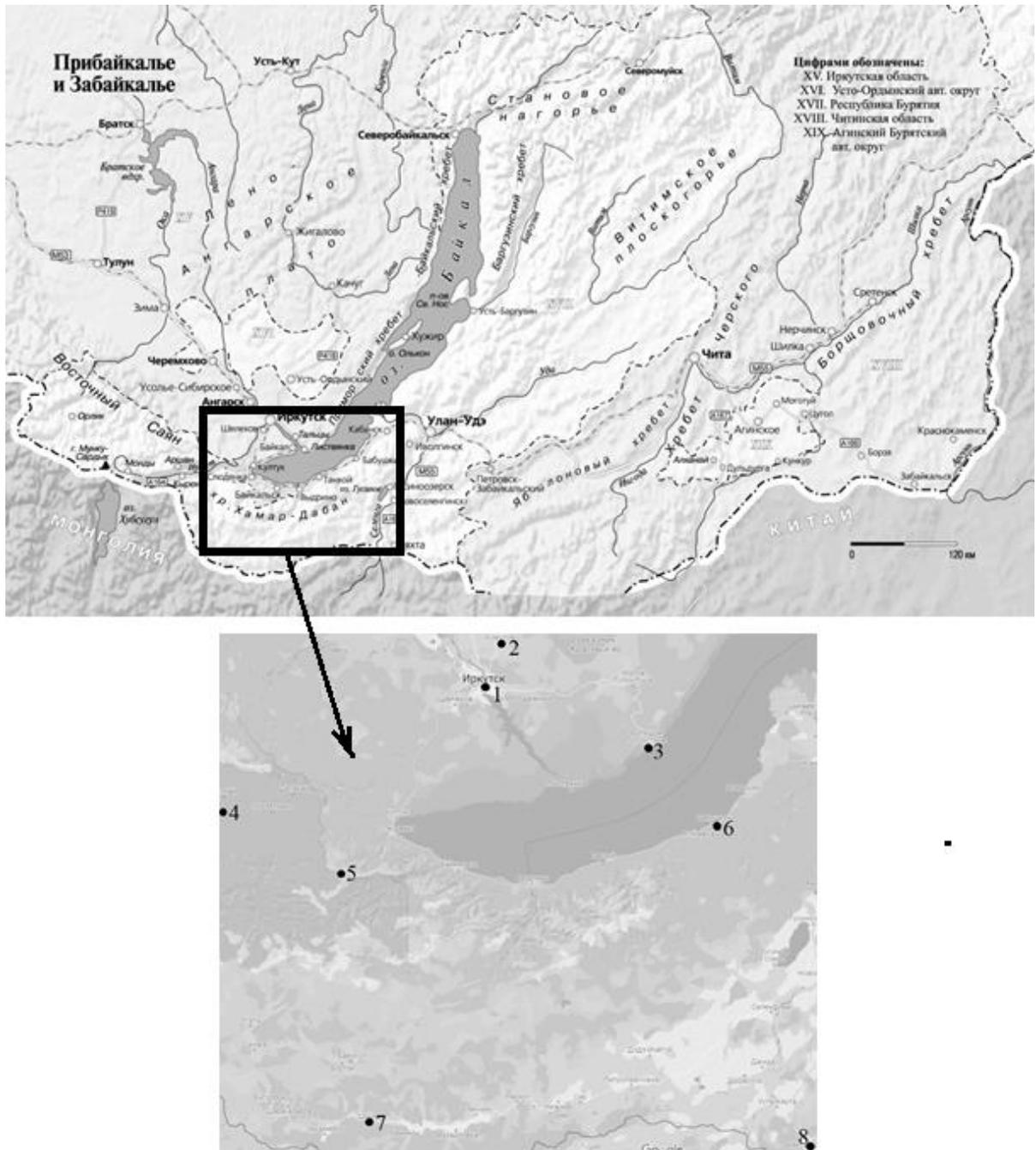


Рис. 2.9. Расположение метеостанций в южной части Байкальской природной территории. 1 – Иркутск, 2 – Хомутово, 3 – Большое Голоустное, 4 – Тунка, 5 – Хамар-Дабан, 6 – Бабушкин, 7 – Цакир, 8 – Кяхта.

$$\vec{\xi} = \left(\overline{\xi_1}^T, \dots, \overline{\xi_{ND}}^T \right)^T$$

вектор-столбец значений рассматриваемых метеопараметров на всех метеостанциях во все сроки измерений во все дни. Если предположить, что

последовательность векторов $\vec{\xi}_i, i=1,2,\dots,ND$ является стационарной, то корреляционная матрица R вектора $\vec{\xi}$ является блочно-тёплицевой и имеет вид

$$R = \begin{pmatrix} R_0 & R_1 & R_2 & \dots & R_{ND-1} \\ R_1^T & R_0 & R_1 & \dots & R_{ND-2} \\ \dots & \dots & \dots & \dots & \dots \\ R_{ND-2}^T & & R_1^T & R_0 & R_1 \\ R_{ND-1}^T & R_{ND-2}^T & \dots & R_1^T & R_0 \end{pmatrix},$$

где $R_h, h=0,\dots,ND-1$ – несимметричные нетёплицевы блоки размерности $(NS \cdot NC) \times (NS \cdot NC)$. Отметим, что матрица R имеет размерность $(ND \cdot NS \cdot NC) \times (ND \cdot NS \cdot NC)$ (например, при $ND=10$ размерность матрицы R – 2080×2080). Последовательности

$$\begin{aligned} &T_{1,1}^k, \dots, T_{8,1}^k, T_{1,2}^k, \dots, T_{8,2}^k, \dots, T_{1,ND}^k, \dots, T_{8,ND}^k, \\ &u_{1,1}^k, \dots, u_{8,1}^k, u_{1,2}^k, \dots, u_{8,2}^k, \dots, u_{1,ND}^k, \dots, u_{8,ND}^k, \\ &v_{1,1}^k, \dots, v_{8,1}^k, v_{1,2}^k, \dots, v_{8,2}^k, \dots, v_{1,ND}^k, \dots, v_{8,ND}^k \end{aligned}$$

температуры воздуха и компонент скорости ветра k -ой станции являются периодически коррелированными с периодом $p=8$ измерениям (что соответствует периоду равному 24 часам), а последовательность $P_{1,1}^k, P_{2,1}^k, P_{1,2}^k, P_{2,2}^k, \dots, P_{1,ND}^k, P_{2,ND}^k$ количества осадков – периодически коррелированной с периодом $p=2$.

В данной работе элементы матрицы R оценивались по данным метеонаблюдений, собранным за $NY=36$ лет. Использование предположения о стационарности последовательности векторов $\vec{\xi}_i, i=1,2,\dots,ND$ приводит к тому, что элементы матрицы R оцениваются по выборке объёма $NY \cdot ND$. То, что объём выборки существенно меньше размерности $(ND \cdot NS \cdot NC) \times (ND \cdot NS \cdot NC)$ матрицы R , приводит к вырожденности матрицы R . То есть, строго говоря, матрица R , составленная из оценок коэффициентов корреляции по реальным данным, не является положительно определённой. Дальнейшее моделирование

будет осуществляться с матрицей R' , представляющей собой регуляризованную матрицу R :

$$R' = (1 - \varepsilon)R + \varepsilon I,$$

где I – единичная матрица, $\varepsilon = 10^{-5}$ [43]. Разница между элементами матриц R, R' не превышает статистическую погрешность оценки элементов матрицы R по реальным данным.

Для аппроксимации эмпирической функции одномерного распределения температуры, как и в параграфе 2.1 Главы 2, используем смесь 2 нормальных распределений с параметрами выбранному по алгоритму, предложенному в [40]. Обозначим $FT_j^k(x)$, $j = 1, 2, \dots, 8$, $k = 1, 2, \dots, NS$ соответствующие функции распределений температуры в j -ый срок измерений на k -ой метеостанции.

При моделировании пространственно-временного поля компонент скорости ветра в работе [104] использовано предположение о гауссовости одномерных распределений исследуемых величин. Из рисунка 2.10 видно, что для рассматриваемой в этой работе территории использование такого предположения нецелесообразно – выборочные одномерные распределения компонент существенно отличаются от гауссовских. В статье [129] указывается, что одномерные распределения являются негауссовскими, и при моделировании используются выборочные функции распределения компонент скорости ветра. В данной работе считаем, что функции распределения $FU_j^k(x), FV_j^k(x)$, $j = 1, 2, \dots, 8$, $k = 1, 2, \dots, NS$ компонент скорости ветра суть смесь 2 нормальных распределений, аппроксимирующих выборочные (графики соответствующих плотностей распределения $fu_j^k(x), fv_j^k(x)$ приведены на рисунке 2.10). Параметры смесей определяются по тому же алгоритму, что использован при аппроксимации распределений температуры воздуха.

В параграфе 2.1 Главы 2 для моделирования временных рядов количества выпавших осадков использовался мультипликативный подход. Этот же подход был использован для моделирования пространственно-временных полей количества

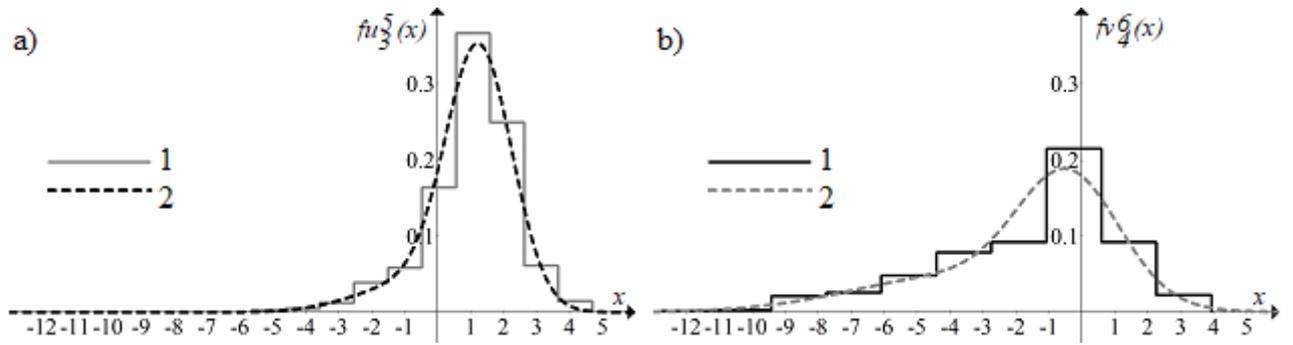


Рис. 2.10. Выборочные (кривые 1) и аппроксимирующие (кривые 2) плотности распределений компонент скорости ветра. График а) – u -компонента, Хамар-Дабан ($k = 5$), $j = 3$. График б) – v -компонента, Бабушкин ($k = 6$), $j = 4$. 10-20 июля.

осадков в работах [135, 151, 172]. Продемонстрируем в этом параграфе иной подход к моделированию. Пусть $x_0 = 0, x_1 = 0.1, x_2 < \dots < x_N$ – последовательность точек. По реальным данным оценивается эмпирическая функция одномерного распределения $\overline{FP}_j^k(x)$, $j = 1, 2, k = 1, 2, \dots, NS$ полусуточного количества осадков. Отметим, что $\overline{FP}_j^k(x_1)$ есть оценка вероятности отсутствия осадков в j -ую половину суток на k -ой метеостанции. Определим функцию $FP_j^k(x)$ одномерного распределения осадков согласно формуле

$$FP_j^k(x) = \begin{cases} a_j^k(i)x + b_j^k(i), & x_{i-1} \leq x < x_i, \quad i = 1, \dots, N, \\ \left(1 - \exp \left(- \left(\frac{x}{\lambda_j^k} \right)^{\gamma_j^k} \right) \right)^{\alpha_j^k}, & x > x_N. \end{cases}$$

Соответствующая плотность распределения имеет вид

$$fp_j^k(x) = \begin{cases} a_j^k(i), & x_{i-1} \leq x < x_i, \quad i = 1, \dots, N, \\ \frac{\alpha_j^k \gamma_j^k}{\lambda_j^k} \left(\frac{x}{\lambda_j^k} \right)^{\gamma_j^k - 1} \left(1 - \exp \left(- \left(\frac{x}{\lambda_j^k} \right)^{\gamma_j^k} \right) \right)^{\alpha_j^k - 1} \exp \left(- \left(\frac{x}{\lambda_j^k} \right)^{\gamma_j^k} \right), & x > x_N. \end{cases}$$

Параметры $a_j^k(i), b_j^k(i), i=1, \dots, N$ выбираются из условия

$$\begin{cases} a_j^k(i)x_{i-1} + b_j^k(i) = \overline{FP}_j^k(x_{i-1}), \\ a_j^k(i)x_i + b_j^k(i) = \overline{FP}_j^k(x_i). \end{cases}$$

Параметры $\alpha_j^k, \gamma_j^k, \lambda_j^k$ «хвоста» распределения выбираются из условия непрерывности функции $FP_j^k(x)$ в точке x_N и близости первых трёх моментов распределения случайной величины с функцией распределения $FP_j^k(x)$ к оценкам этих моментов по реальным данным. В таблице 2.9 приведены оценки математического ожидания \overline{m}_j^{-k} , стандартного отклонения $\overline{\sigma}_j^{-k}$ и коэффициента асимметрии \overline{A}_j^{-k} по реальным данным и значения m_j^k, σ_j^k, A_j^k этих статистических характеристик, вычисленные по функции распределения $FP_j^k(x)$. На рисунке 2.11 изображены графики $\overline{FP}_j^k(x)$ и $FP_j^k(x)$. Отметим, что функция

$$g(x) = \left(1 - \exp \left(- \left(\frac{x}{\lambda_j^k} \right)^{\gamma_j^k} \right) \right)^{\alpha_j^k}$$

использованная в качестве «хвоста» функции распределения $FP_j^k(x)$ при $\alpha_j^k = 1$ есть функция распределения Вейбулла. Использование «хвостов» такого типа было предложено в работе [39].

Два алгоритма моделирования пространственно-временного поля

Моделирование негауссовского вектора $\vec{\xi} = \left(\overline{\xi}_i^T, \dots, \overline{\xi}_{ND}^T \right)^T$ с

корреляционной матрицей R' и функциями

$$FT_j^k(x), FU_j^k(x), FV_j^k(x), FP_i^k(x), j=1, 2, \dots, 8, i=1, 2, k=1, 2, \dots, NS$$

одномерных распределений его компонент можно осуществлять, например, методом обратных функций распределения. В данной работе будут использованы

Таб. 2.9. Сравнение реальных и модельных моментов распределения
полусуточных сумм осадков. 10-20 июля.

Метеостанция	Половина суток j	Математическое ожидание		Дисперсия		Коэффициент асимметрии	
		\bar{m}_j^k	m_j^k	$\bar{\sigma}_j^k$	σ_j^k	\bar{A}_j^k	A_j^k
Иркутск	$j = 1$	1.67	1.64	4.99	5.06	5.47	5.48
Иркутск	$j = 2$	1.41	1.38	3.70	3.74	4.43	4.05
Большое Голоустное	$j = 1$	0.85	0.82	2.87	3.00	6.54	6.60
Большое Голоустное	$j = 2$	0.85	0.86	2.40	2.61	5.02	4.77
Цакир	$j = 1$	1.68	1.58	4.49	4.36	5.89	5.98
Цакир	$j = 2$	1.58	1.58	3.40	3.59	4.32	4.22

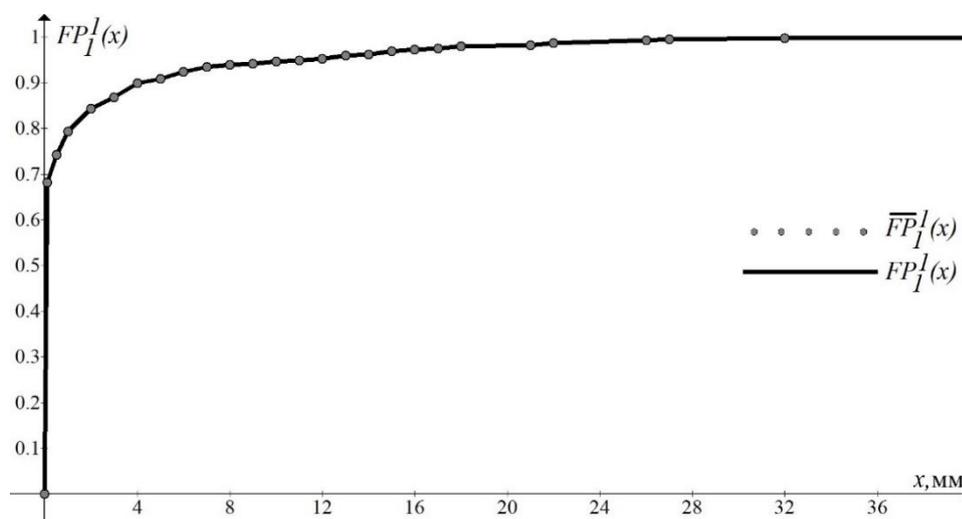


Рис 2.11. Выборочное $\overline{FP}_j^k(x)$ и модельное $FP_j^k(x)$ одномерные распределения
полусуточных осадков. Иркутск ($k = 1$), $j = 1$. 10-20 июля

2 других подхода к моделированию.

Первый подход представляет собой специальный вариант метода обратных функций распределения, основанный на нормализации исходных данных [46, 77].

Алгоритм 2.2.

1. Нормализуем реальные данные

$$T_{j,i}^k(m), u_{j,i}^k(m), v_{j,i}^k(m), P_{l,i}^k(m), \\ j = 1, \dots, 8, l = 1, 2, i = 1, \dots, ND, k = 1, \dots, NS, m = 1, \dots, NY,$$

собранные в m -ый год наблюдений на k -ой метеостанции в j -ый срок наблюдений или в l -ую половину i -ых суток по правилу

$$nT_{j,i}^k(m) = \Phi^{-1}\left(FT_j^k\left(T_{j,i}^k(m)\right)\right), nu_{j,i}^k(m) = \Phi^{-1}\left(FU_j^k\left(u_{j,i}^k(m)\right)\right), \\ nv_{j,i}^k(m) = \Phi^{-1}\left(FV_j^k\left(v_{j,i}^k(m)\right)\right), nP_{l,i}^k(m) = \Phi^{-1}\left(FP_l^k\left(P_{l,i}^k(m)\right)\right), \\ j = 1, \dots, 8, l = 1, 2, i = 1, \dots, ND, k = 1, \dots, NS, m = 1, \dots, NY$$

и сформируем из нормализованных данных векторы

$$\vec{\eta}(m) = \left(\vec{\eta}_i(m)^T, \dots, \vec{\eta}_{ND}(m)^T\right)^T, m = 1, \dots, NY \text{ с такой же структурой, как у векторов}$$

$$\vec{\xi} = \left(\vec{\xi}_i^T, \dots, \vec{\xi}_{ND}^T\right)^T. \text{ По полученной выборке } \vec{\eta}(m), m = 1, \dots, NY \text{ оценим}$$

коэффициенты корреляции, составим из них блочно-тёплицеву матрицу R_η и регуляризируем её по правилу

$$R_\eta' = (1 - \varepsilon)R_\eta + \varepsilon I, \varepsilon = 10^{-5}.$$

2. Моделируем гауссовский вектор $\vec{\eta} = \left(\vec{\eta}_1^T, \dots, \vec{\eta}_{ND}^T\right)^T$ с корреляционной матрицей R_η' и структурой как у вектора $\vec{\xi}$ (экономичный алгоритм моделирования гауссовского вектора с блочно-тёплицевой корреляционной матрицей приведен, например, в [175]). Обозначим $\eta_{j,i}^{T,k}, \eta_{j,i}^{u,k}, \eta_{j,i}^{v,k}, \eta_{j,i}^{P,k}$ компоненты вектора $\vec{\eta}$.

3. Вычисляем вектор $\vec{\xi} = \left(\vec{\xi}_i^T, \dots, \vec{\xi}_{ND}^T\right)^T$ с помощью преобразования

$$T_{j,i}^k = \left(FT_j^k\right)^{-1}\left(\Phi\left(\eta_{j,i}^{T,k}\right)\right), u_{j,i}^k = \left(FU_j^k\right)^{-1}\left(\Phi\left(\eta_{j,i}^{u,k}\right)\right),$$

$$v_{j,i}^k = (FV_j^k)^{-1} \left(\Phi \left(\eta_{j,i}^{v,k} \right) \right), \quad P_{l,i}^k = (FP_l^k)^{-1} \left(\Phi \left(\eta_{l,i}^{P,k} \right) \right),$$

$$j = 1, \dots, 8, \quad l = 1, 2, \quad i = 1, \dots, ND, \quad k = 1, \dots, NS$$

и последующей замены

$$P_{l,i}^k = \begin{cases} 0, & \text{если } P_{l,i}^k < 0.1 \text{ мм,} \\ P_{l,i}^k, & \text{иначе.} \end{cases}$$

Второй подход к моделированию негауссовского вектора основан на итерационном алгоритме, предложенном в работе [210]. Этот алгоритм позволяет моделировать негауссовский вектор $\vec{\zeta} = (\zeta_1, \dots, \zeta_N)^T$ с ковариационной матрицей R_ζ и заданными функциями $F_k(x)$, $k = 1, \dots, N$ одномерных распределений. Опишем этот алгоритм в общем виде. Переход к моделированию вектора $\vec{\xi} = \left(\vec{\xi}_i^T, \dots, \vec{\xi}_{ND}^T \right)^T$ с корреляционной матрицей R' и функциями $FT_j^k(x), FU_j^k(x), FV_j^k(x), FP_i^k(x)$, $j = 1, 2, \dots, 8$, $i = 1, 2$, $k = 1, 2, \dots, NS$ одномерных распределений осуществляется очевидным образом.

Алгоритм 2.3.

1. Моделируется m независимых реализаций каждой случайной величины ζ_k , $k = 1, \dots, N$ с функцией распределения $F_k(x)$. Из полученных реализаций формируется матрица

$$Y = \left(\bar{y}^{(1)}, \dots, \bar{y}^{(N)} \right)$$

размерности $m \times N$. Здесь $\bar{y}^{(k)}$ – k -ый столбец матрицы Y , состоящий из m реализаций случайной величины ζ_k , $k = 1, \dots, N$. Строки матрицы Y можно рассматривать как реализации случайного вектора, компоненты которого независимы и имеют функции одномерных распределений $F_1(x), \dots, F_N(x)$.

2. По строкам матрицы Y оценивается ковариационная матрица \tilde{R} размерности $N \times N$.

3. Осуществляется разложение Холецкого для матриц R_ζ и \tilde{R} :

$$R_\zeta = CC^T, \quad \tilde{R} = \tilde{C}\tilde{C}^T.$$

4. Вычисляется матрица \tilde{Y} по правилу:

$$\tilde{Y} = Y(\tilde{C}^T)^{-1}C^T.$$

5. Элементы столбца $\bar{y}^{(k)}$ сортируются в соответствии с элементами столбца $\bar{y}^{(k)}$ следующим образом. Для каждого $j=1, \dots, m$ в столбце $\bar{y}^{(k)}$ находится элемент, который является j -ым по величине. Обозначим его $\tilde{y}_j^{(k)}$. Пусть $\tilde{y}_j^{(k)}$ стоит в строке с номером q . Затем, находится j -ый по величине элемент $y_j^{(k)}$ в столбце $\bar{y}^{(k)}$. Пусть он стоит в строке с номером l . Осуществляется перестановка элемента $y_j^{(k)}$ в q -ую строку. Стоящий на этом месте элемент переносится в l -ую строку. Эта процедура повторяется для каждого столбца $\bar{y}^{(1)}, \dots, \bar{y}^{(N)}$.

6. Шаги 2-5 повторяются до тех пор, пока не будет выполнен какой-либо критерий останова (например, пока не будет выполнено условие $\|R_\zeta - \tilde{R}\| < \varepsilon_\zeta$).

Также как и при реализации метода нормализации, при использовании итерационного алгоритма на последнем шаге моделирования вектора

$\bar{\xi} = \left(\bar{\xi}_i^T, \dots, \bar{\xi}_{ND}^T \right)^T$ осуществляется замена

$$P_{l,i}^k = \begin{cases} 0, & \text{если } P_{l,i}^k < 0.1 \text{ мм.} \\ P_{l,i}^k, & \text{иначе.} \end{cases}$$

Алгоритм 2.3 был численно реализован коллегой автора диссертационной работы Акентьевой Мариной Сергеевной.

В этой работе, в силу большой размерности матрицы R' , и, как следствия, вычислительной трудоёмкости моделирования, Алгоритм 2.3 был принудительно остановлен после 5 итераций. Оказалось, что при этом

$\|R_T - \tilde{R}_T\| < 10^{-3}$, $\|R_u - \tilde{R}_u\| < 10^{-3}$, $\|R_v - \tilde{R}_v\| < 10^{-3}$, $\|R_P - \tilde{R}_P\| < 10^{-2}$, где R_X и \tilde{R}_X , $X \in \{T, u, v, P\}$ – матрицы, составленные из элементов матриц R' и \tilde{R} , соответственно, отвечающих за автокорреляции элемента X . Кроме этого, численные эксперименты показали, что $\|R_P - \tilde{R}_P\|$ не уменьшается при увеличении числа итераций.

Верификация модели

С помощью Алгоритмов 2.3 и 2.3 было построено $3 \cdot 10^5$ модельных реализаций вектора $\vec{\xi} = \left(\overline{\xi}_i^T, \dots, \overline{\xi}_{ND}^T \right)^T$. Ниже на рисунках 2.12, 2.13 приведены результаты сравнения корреляционных коэффициентов температуры воздуха, и компоненты скорости ветра и количества осадков, оцененных по реальным данным и по модельным реализациям. Разница между значениями элементов матриц R_T, R_u, R_v и элементов соответствующих матриц, оцененных по модельным данным при использовании итерационного Алгоритма 2.3, столь мала, что визуально графики на рисунке 2.12 неразличимы. Разница между элементами R_T, R_u, R_v и соответствующими коэффициентами корреляции, оцененными по модельным данным при использовании метода нормализации существенно больше и в ряде численных экспериментов приближалась к 3σ , где σ – среднеквадратическое отклонение оценки коэффициента корреляции по выборке реальных данных. В то же время, как видно из рисунка 2.13, оба метода воспроизводят значения элементов матрицы R_P несколько хуже (хотя отклонения оценок корреляционных коэффициентов по модельным выборкам от оценок по реальным данным не превышают 3σ). Во всех проведённых численных экспериментах итерационный метод показал более высокую точность воспроизведения корреляционной структуры полей осадков.

Так же, как и в параграфе 2.1 этой Главы, верификация предложенной модели проводилась путём сравнения оценок различных характеристик пространственно-

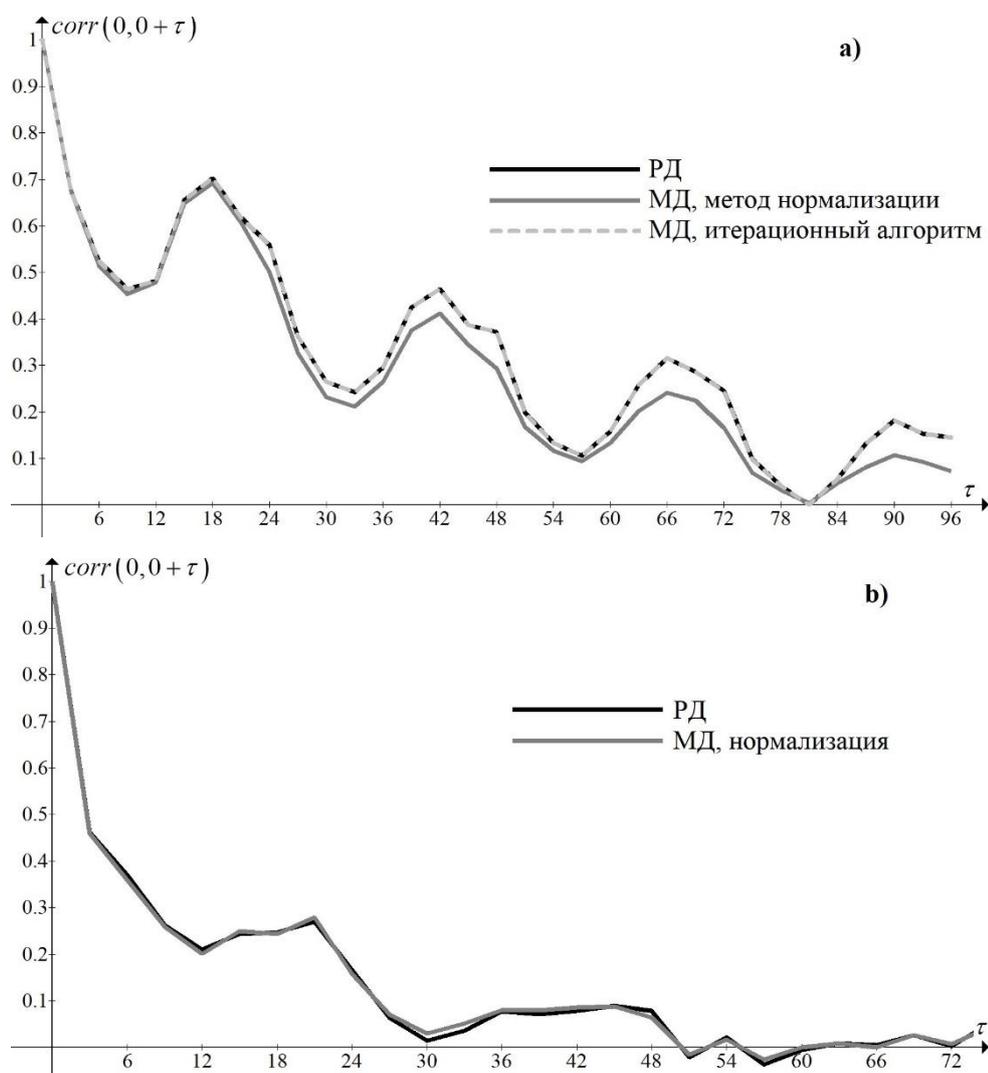


Рис. 2.12. Оценки коэффициентов корреляции между значениям метеопараметра в моменты времени $t = 0$ и $t + \tau$, $\tau = 0, 3, 6, \dots$ часов, вычисленные по реальным и модельным данным при использовании Алгоритмов 2.2 и 2.3. График а) – температура воздуха, график б) – u-компонента скорости ветра. Иркутск, 10-20 июля.

временного поля по модельным и реальным данным. В качестве примеров использованных для верификации характеристик на рисунках 2.14-2.18 приведены оценки

- вероятности

$$p_a = P(T_{cp}^1 > lt, T_{cp}^2 > lt, T_{cp}^4 > lt)$$

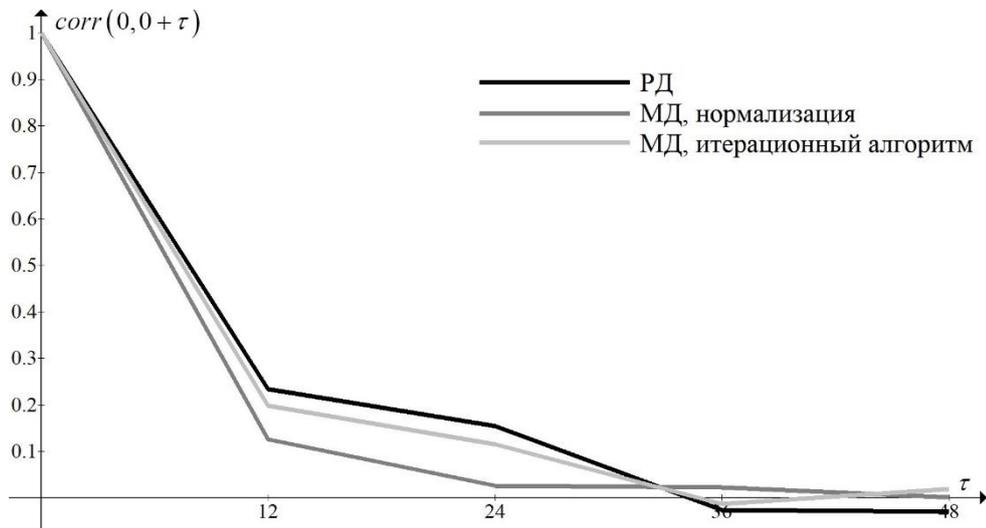


Рис. 2.13. Оценки коэффициентов корреляции количества осадков в моменты времени $t = 0$ и $t + \tau$, $\tau = 0, 12, 24, \dots$ часов, вычисленные по реальным и модельным данным при использовании Алгоритмов 2.2 и 2.3. Иркутск, 10-20 июля.

того, что средняя полусуточная температура воздуха T_{cp}^k превышает заданный уровень lt °C одновременно в Иркутске ($k = 1$), Хомутово ($k = 2$) и Тунке ($k = 4$),

- вероятности

$$p_b = P\left(T_{cp}^6 > lt, T_{cp}^7 > lt, T_{cp}^8 > lt\right)$$

того, что средняя полусуточная температура воздуха превышает заданный уровень lt °C одновременно в Бабушкине, Цакире и Кяхте,

- вероятности

$$p_c = P\left(T_{cp}^8 > lt, wind_{cp}^8 > lw\right)$$

того, что средняя полусуточная температура воздуха T_{cp}^8 превышает заданный уровень lt °C и средний за эти же полусутки модуль скорости ветра

$$wind_{cp}^8 = \sqrt{\left(u_{cp}^8\right)^2 + \left(v_{cp}^8\right)^2} \text{ больше } lw = 2 \text{ м/с в Кяхте,}$$

- вероятности

$$p_d = P\left(T_{cp}^3 > lt, wind_{cp}^3 > lw\right)$$

того, что средняя полусуточная температура воздуха T_{cp}^3 превышает заданный уровень lt °C и средний за эти же полусутки модуль скорости ветра $wind_{cp}^3$ больше $lw = 2$ м/с в Большом Голоульном,

- вероятности

$$p_e = P\left(T_{cp}^8 > lt, P_{cp}^8 > lp\right)$$

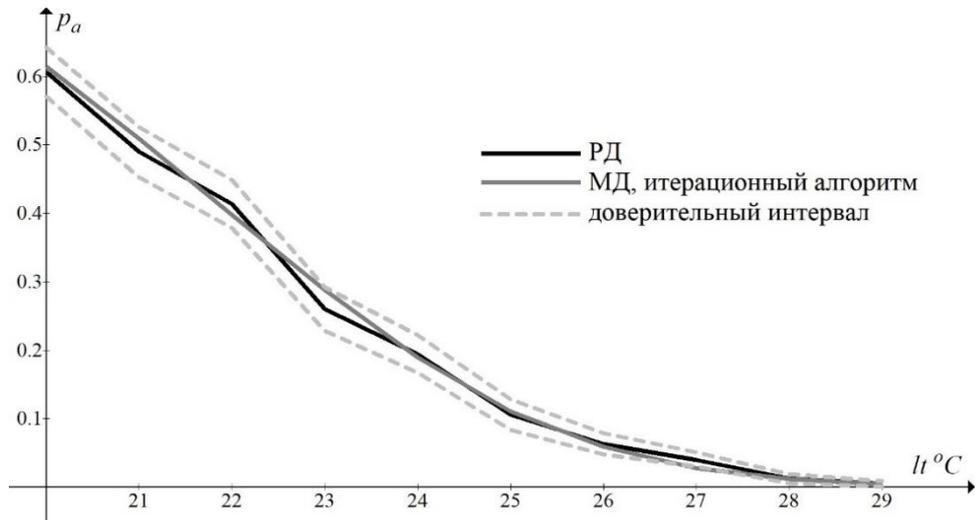
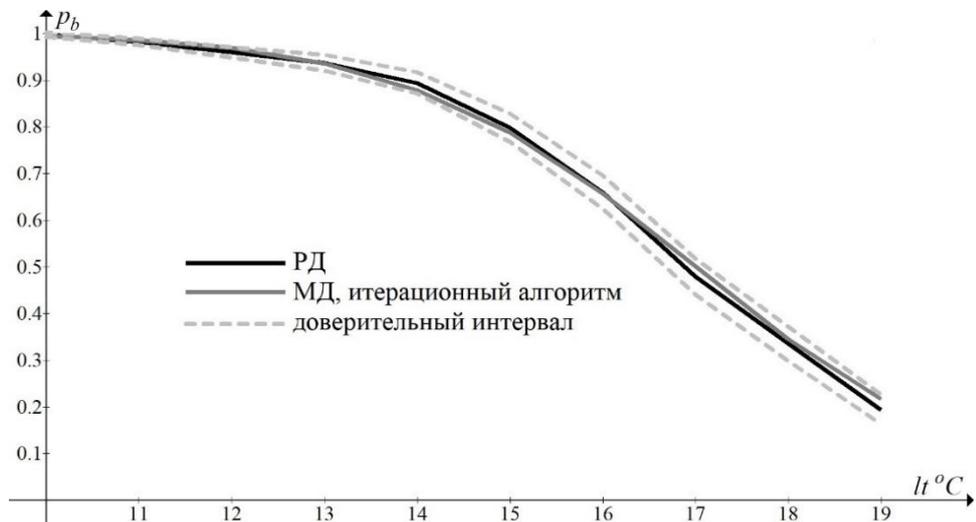
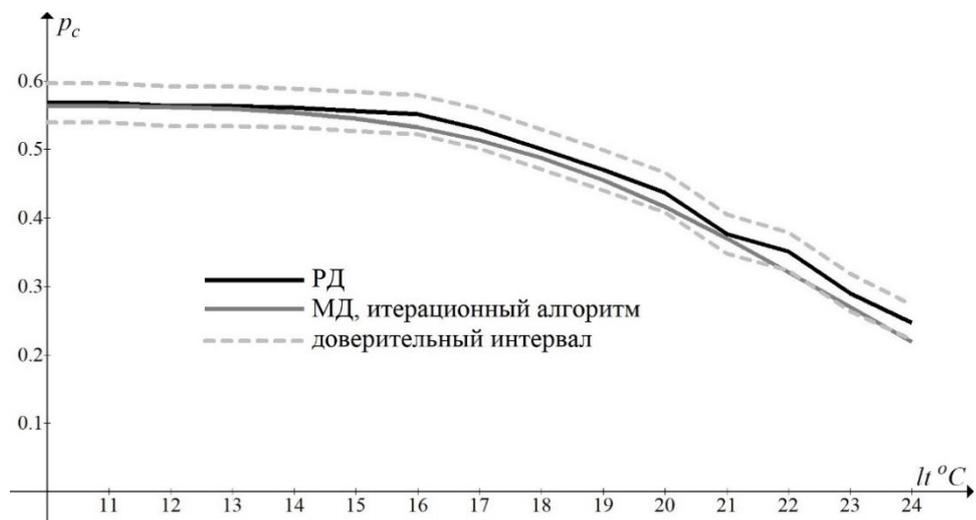
того, что средняя полусуточная температура воздуха T_{cp}^8 превышает заданный уровень lt °C и полусуточное количество P_{cp}^8 осадков больше $lp = 1$ мм в Кяхте.

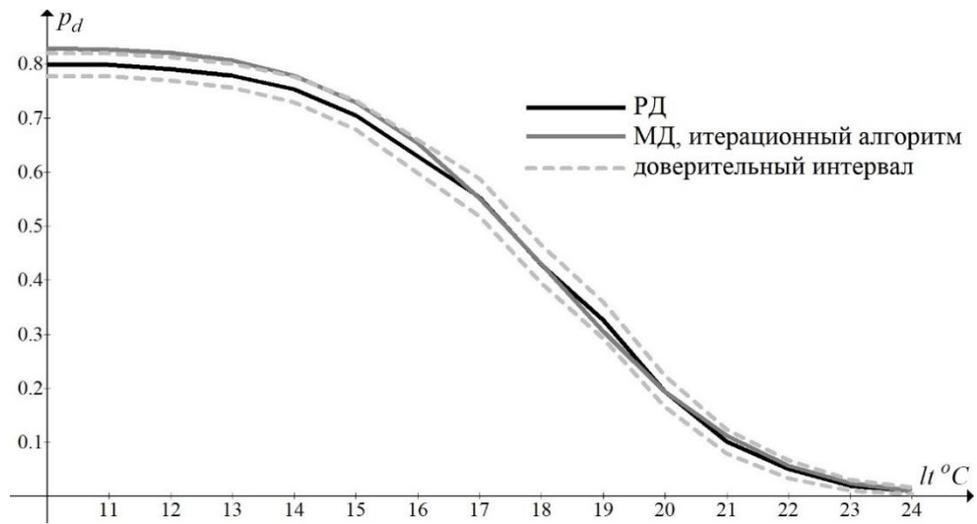
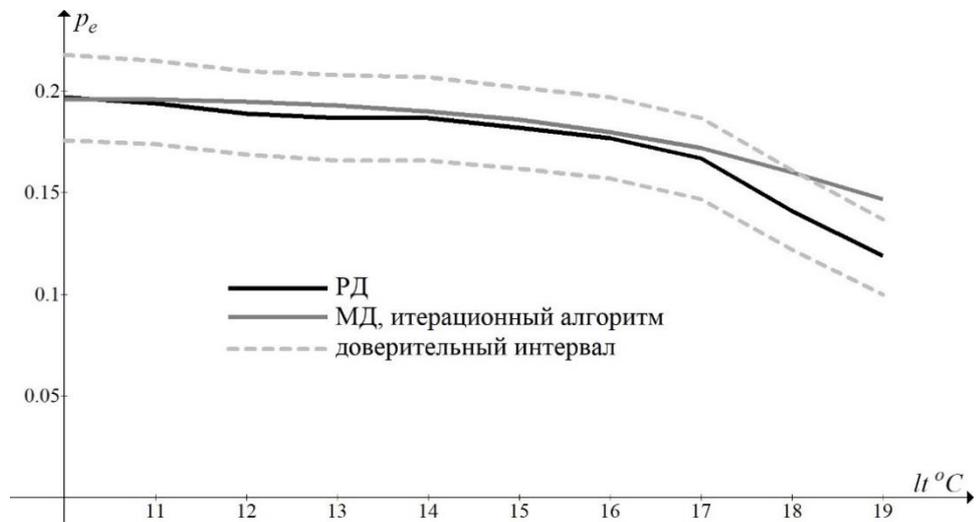
Средняя полусуточная температура воздуха T_{cp}^k вычислялась как

$$T_{cp}^k = \frac{1}{4ND} \sum_{j=1}^{ND} \sum_{i=1}^4 T_{i,j}^k$$

и соответствует средней температуре воздуха в период с 8 утра до 17 часов дня местного времени. Оценки вышеуказанных вероятностей проводились по реальным данным и модельным реализациям, полученным с помощью Алгоритма 2.3. Для оценок по реальным данным приведены границы доверительных интервалов $p_X \pm \sigma$, $X \in \{a, b, c, d, e\}$.

Верификация показала, что предложенная модель хорошо воспроизводит свойства рассматриваемого пространственно-временного поля. В параграфе 3.4 Главы 3 будет приведён пример применения этой модели для исследования пространственно-временного поля одного из биоклиматических индексов.

Рис. 2.14. Оценки вероятностей p_a . 10-20 июля.Рис. 2.15. Оценки вероятностей p_b . 10-20 июля.Рис. 2.16. Оценки вероятностей p_c . 10-20 июля.

Рис. 2.17. Оценки вероятностей p_d . 10-20 июля.Рис. 2.18. Оценки вероятностей p_e . 10-20 июля.

3.1. Биоклиматические индексы. Определения и обозначения

Индекс холодного стресса

Индекс холодного стресса (ИХС) впервые был предложен в работе [196]. Этот сравнительно простой индекс был использован для описания воздействия на человека низких температур воздуха и ветра в холодное время года. Согласно [166, 179], с 2001 г. для вычисления ИХС используется формула:

$$W = \begin{cases} 13.12 + 0.6215T - 11.37V^{0.16} + 0.3965TV^{0.16}, & \text{если } V \geq 5 \text{ км / ч,} \\ T + \frac{-1.59 + 0.1345T}{5}V, & \text{если } V < 5 \text{ км / ч,} \end{cases} \quad (3.1)$$

где T °С – приземная температура воздуха, V км / ч – скорость ветра на высоте 10 м над землёй. Единицей измерения индекса холодного стресса, по определению, считается градус Цельсия (°С). ИХС W °С представляет собой температуру воздуха, при которой в тени в безветренных условиях скорость охлаждения открытых участков кожи совпадает со скоростью их охлаждения при температуре T °С и скорости ветра V км / ч. ИХС вычисляют только при $T \leq 10$ °С.

Биоклиматический индекс суровости метеорежима

Биоклиматический индекс суровости метеорежима (БИСМ) является интегральным показателем степени биоклиматического дискомфорта при различных видах жизнедеятельности. Этот биоклиматический индекс является функцией целого ряда метеорологических и физико-географических параметров:

приземной температуры воздуха T °C, атмосферного давления P гПа, модуля скорости ветра V м/с, относительной влажности воздуха H %, а также высоты L м местности над уровнем моря [68, 103]. Значение БИСМ B вычисляется следующим образом:

$$B = \frac{\tilde{T}(P - 266)(1 - 0.02V)}{75\tilde{H}\tilde{S}}. \quad (3.2)$$

Здесь $\tilde{T}, \tilde{H}, \tilde{S}$ – температурный, влажностный и радиационный коэффициенты, определяемые равенствами

$$\tilde{T} = \begin{cases} 1 - 0.0089(22 - T), & T < 22^\circ \text{C}, \\ 1 - 0.0263(T - 22), & T \geq 22^\circ \text{C}, \end{cases}$$

$$\tilde{H} = \begin{cases} 1 + 0.6 \frac{50 - H}{100}, & H < 50\%, \\ 1 + 0.6 \frac{H - 50}{100}, & H \geq 50\%, \end{cases}$$

$$\tilde{S} = \begin{cases} 1, & H < 2000 \text{ м}, \\ 1 + 0.45 \frac{L - 2000}{1000}, & H \geq 2000 \text{ м}. \end{cases}$$

По определению, БИСМ – безразмерная величина, принимающая значения в интервале $[0,10]$. Считается, что погодные условия являются комфортными при значениях $B \in [8,10]$. Значения B , лежащие в интервалах $[7,8)$ и $[6,7)$ характеризуют относительную комфортность метеоусловий 1 и 2 уровня. Если $B \in [5,6)$ или $B \in [4,5)$, то полагают, что погода характеризуется 2 различными степенями компенсируемого дискомфорта, а при $B \in [0,4)$ – некомпенсируемым дискомфортом. В [83] указано, что БИСМ – это эмпирическая мера комфорта, отражающего суровость климатического влияния на организм человека. Этот индекс является относительно универсальным, так как позволяет учитывать влияние на организм человека не только низких температур, но и жарких условий полупустынных и пустынных районов, а также избыток ультрафиолета в горах и его дефицит в арктических широтах.

На рисунке 3.1 приведены графики измеряемой 8 раз в сутки температуры воздуха и соответствующие ей значения индекса холодного стресса и биоклиматического индекса суровости метеорежима.

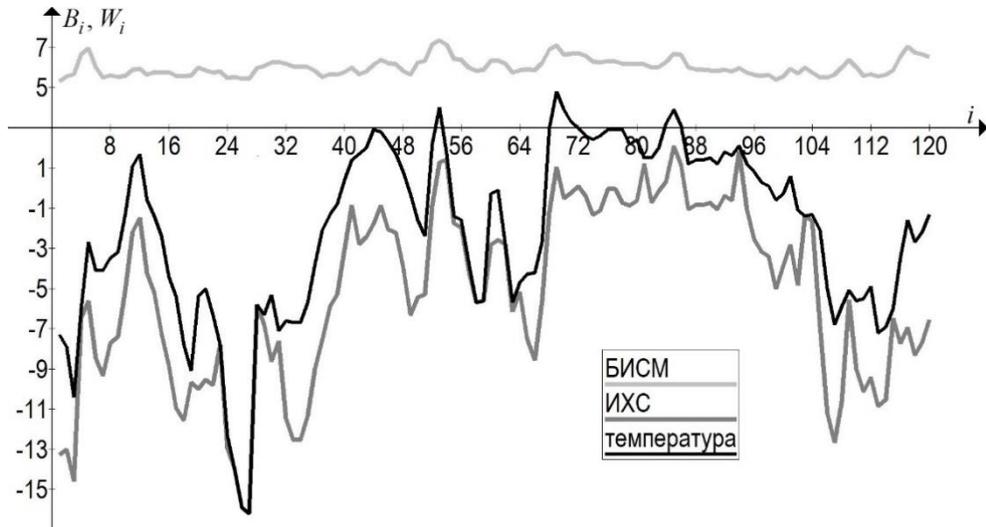


Рис. 3.1. Реальные значения температуры воздуха, ИХС и БИСМ. Онега, 1-15 марта 2020.

Индекс жары

Для оценки влияния тёплого воздуха на состояние человека в летний период используют такой биоклиматический показатель как индекс жары [97, 184, 190, 191, 201, 202]. В работах [201, 202] индекс жары (ИЖ) определён таблично как функция температуры воздуха и его относительной влажности. Для вычисления ИЖ при значениях температуры и влажности, отсутствующих в таблицах, используют различные интерполяционные формулы [97]. Одной из наиболее часто используемых является формула [191]

$$HI = T - 1.0799e^{0.03755T} \left(1 - e^{0.0801(D-14)} \right), \quad (3.3)$$

где HI °C – индекс жары, T °C – приземная температура воздуха, D °C – температура точки росы

$$D = \frac{237.3\alpha}{17.27 - \alpha}, \quad \alpha = \frac{17.27T}{237.3 + T} + \ln HI,$$

H – относительная влажность воздуха, измеряемая в долях единицы. Характер теплового воздействия при различных значениях ИЖ представлен в таблице 3.1 [72]. При применении ИЖ необходимо учитывать, что он определялся для людей среднего возраста и комплекции, находящихся в состоянии относительного покоя.

Таб. 3.1. Возможное влияние ИЖ на физическое состояние человека.

Категория	Индекс жары, °C	Возможные расстройства здоровья для населения в группах высокого риска
Экстремальная опасность	54 и выше	Тепловой или солнечный удар
Опасность	41–54	Солнечный удар, мышечные судороги, спазм сосудов от перегрева при длительном воздействии и / или физической активности
Экстремальная осторожность	32–41	Возможны солнечный удар, мышечные судороги, спазм сосудов от перегрева при длительном воздействии и / или физической активности
Осторожность	27–32	Переутомление при длительном воздействии и / или физической активности

Энтальпия влажного воздуха

Наряду с моделями трёх вышеуказанных БИ, в этой главе будет рассмотрена стохастическая модель временных рядов энтальпии влажного воздуха (ЭВВ). Формально, ЭВВ не является биоклиматическим показателем, однако, значения ЭВВ тесно связаны с ощущением «комфортности» метеорологических условий. По определению [72, 55], значение J кДж/кг ЭВВ рассчитывается по следующей формуле:

$$J = c_c T + d(L + c_v T), \quad (3.4)$$

где T °C – приземная температура воздуха, d кг/кг – влагосодержание воздуха при температуре T °C и давлении P гПа, $L = 2501$ кДж/кг – удельная теплота парообразования [78], $c_c = 1.005$ кДж/(кг·°C) – удельная теплоемкость сухого воздуха [54], $c_v = 1.86$ кДж/(кг·°C) – удельная теплоемкость водяного пара [78]. Для вычисления значений J энтальпии влажного воздуха по значениям метеорологических параметров, формулу (3.4) можно переписать в виде

$$J = c_c T + 0.622 H \frac{W(T)}{P} (L + c_v T), \quad (3.5)$$

где H – относительная влажность воздуха (в долях единицы), а давление $W(T)$ насыщенного водяного пара при температуре T °C вычисляется по формуле [125]:

$$W(T) = 6.112 \left(1.0016 + 3.15 \cdot 10^{-6} P - \frac{0.074}{P} \right) \exp \left(\frac{17.62T}{243.12 + T} \right). \quad (3.6)$$

На рисунке 3.2 приведены графики измеряемой 8 раз в сутки температуры воздуха и соответствующие ей значения индекса и энтальпии влажного воздуха.

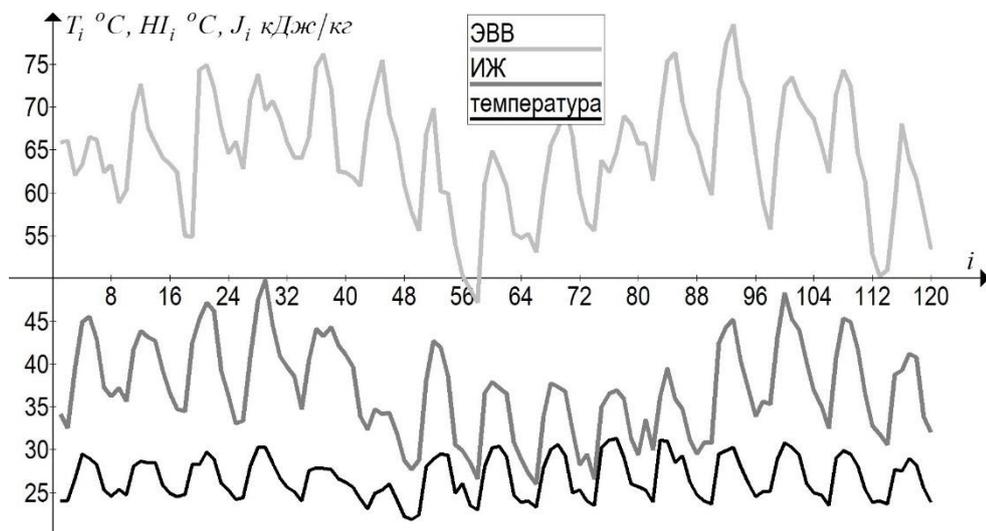


Рис. 3.2. Реальные значения температуры воздуха, ИЖ и ЭВВ. Сочи, 1-15 августа 2015.

3.2. Применение стохастических «генераторов погоды» для моделирования временных рядов биоклиматических индексов

Формально, существует 2 подхода к стохастическому моделированию временных рядов БИ, схематично описанные на рисунке 3.3.

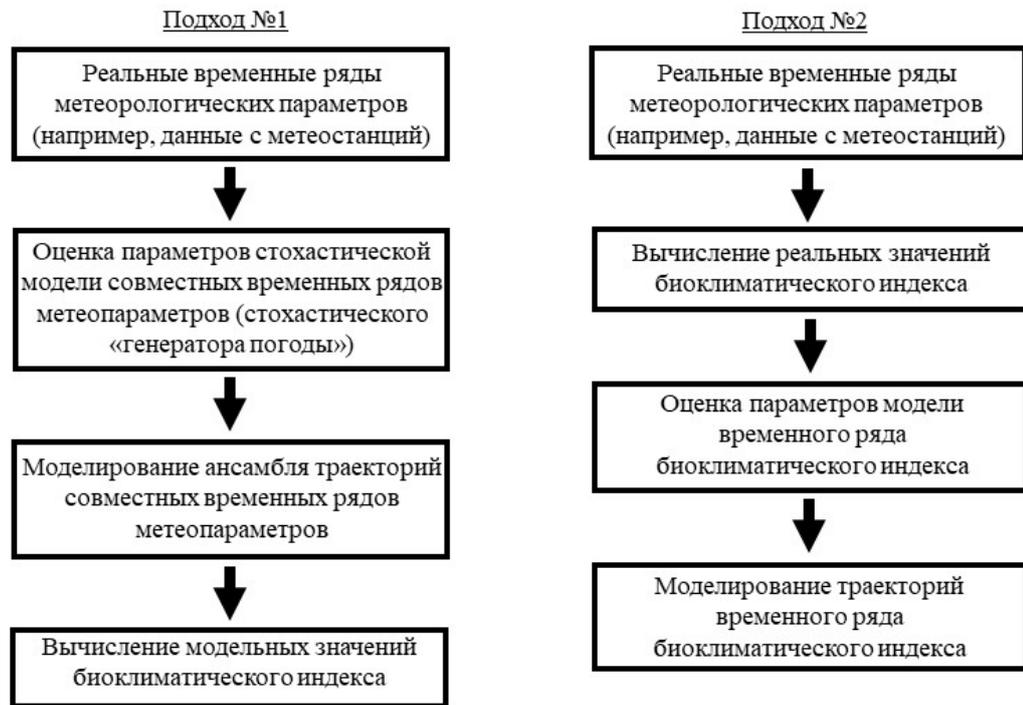


Рис. 3.3. Блок-схемы 2 подходов к моделированию рядов биоклиматических индексов.

Первый подход основан на том, что БИ являются функциями нескольких метеорологических параметров. В связи с этим, представляется естественным сначала смоделировать временные ряды совместных рядов метеорологических параметров (т.е. воспользоваться «генератором погоды»), а затем по полученным модельным траекториям вычислить модельные значения БИ. Этот подход является достаточно универсальным, в том смысле, что по одному ансамблю траекторий, полученных при реализации «генератора погоды», можно вычислить значения различных БИ. В рамках второго подхода, для каждого конкретного БИ строится своя стохастическая модель, при реализации которой моделируются именно

значения этого БИ (далее будем называть этот подход подходом на основе определяющих формул).

В этом параграфе будут приведены примеры применения первого подхода к моделированию временных рядов БИ.

3.2.1. Моделирование временных рядов биоклиматических индексов с учётом их суточного хода

Для моделирования временных рядов биоклиматических индексов с учётом их суточного хода в рамках 1-го подхода в качестве «генератора погоды» будет использована стохастическая модель совместных периодически коррелированных рядов метеорологических параметров, описанная в параграфе 2.1 Главы 2.

Моделирование временных рядов энтальпии влажного воздуха

Вначале рассмотрим возможное применение описанного стохастического «генератора погоды» для моделирования временных рядов энтальпии влажного воздуха.

Необходимо смоделировать реализации временных рядов ЭВВ $\vec{J}^T = (J_1, J_2, \dots, J_{8d})^T$ с 3-часовым разрешением на временном интервале длительностью d дней (при таком временном разрешении каждые сутки характеризуются 8 значениями ЭВВ, и вектор-столбец \vec{J}^T состоит $8d$ компонент). Поскольку, согласно формулам (3.5) и (3.6), ЭВВ есть функция температуры воздуха, его относительной влажности и атмосферного давления, для моделирования реализаций вектора \vec{J} можно воспользоваться Алгоритмом 3.1.

Алгоритм 3.1.

1) Моделируется реализация вектора

$$\left(\vec{T}^T, \vec{H}^T, \vec{P}^T \right)^T = (T_1, T_2, \dots, T_{8d}, H_1, H_2, \dots, H_{8d}, P_1, P_2, \dots, P_{8d})^T$$

по Алгоритму 2.1 (параграф 2.1 Главы 2).

2) Построенная реализация вектора $\left(\bar{T}^T, \bar{H}^T, \bar{P}^T\right)^T$ используется для вычисления значений вектора \vec{J}^T по формулам (3.5) и (3.6).

Предположение о периодической коррелированности процесса $\left(\bar{T}^T, \bar{H}^T, \bar{P}^T\right)^T$ позволяет нам учесть суточный ход, присущий реальным рядам ЭВВ.

Следует отметить, что на коротких временных интервалах, на которых возможно использование предложения о периодической коррелированности процесса $\left(\bar{T}^T, \bar{H}^T, \bar{P}^T\right)^T$, суточный и сезонный ход давления слабо выражены, и изменчивость давления слабо влияет на значение ЭВВ. В связи с этим, при вычислении ЭВВ \vec{J}^T вместо значений давления $P_i, i=1, \dots, 8d$ можно использовать среднее на рассматриваемом временном интервале давление

$$P = \frac{1}{8d} \sum_{i=1}^{8d} P_i.$$

В этом случае формулы (3.5) и (3.6) принимают вид

$$J_i = c_c T_i + 0.622 H_i \frac{W(T_i)}{P} (L + c_v T_i),$$

$$W(T_i) = 6.112 \left(1.0016 + 3.15 \cdot 10^{-6} P - \frac{0.074}{P} \right) \exp \left(\frac{17.62 T_i}{243.12 + T_i} \right), \quad (3.7)$$

$$i = 1, 2, \dots, 8d.$$

В качестве примера, на рисунке 3.4 приведены графики ЭВВ, вычисленной по формулам (3.5), (3.6) и по формулам (3.7).

Использование формулы (3.7) позволяет упростить Алгоритм 3.1 моделирования вектора \vec{J}^T ЭВВ:

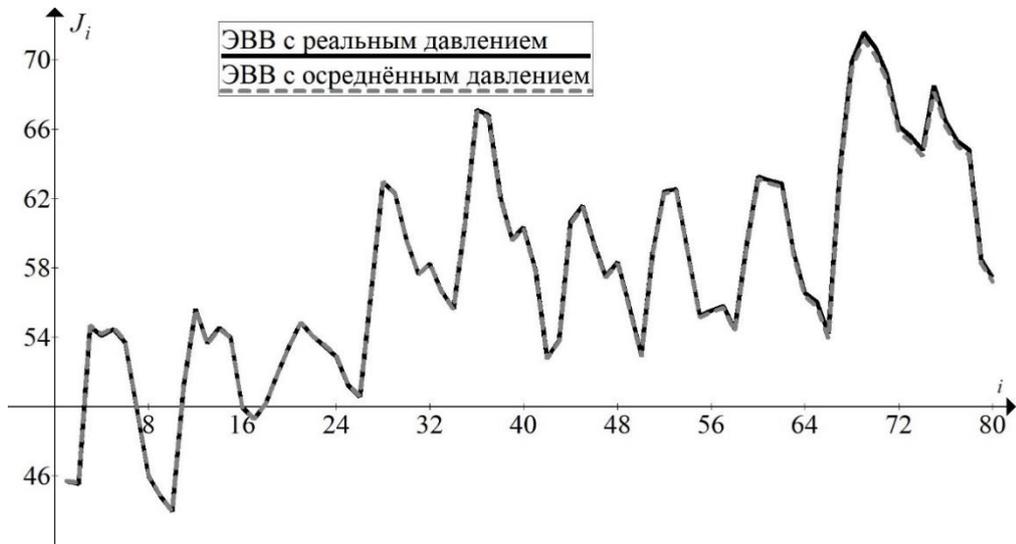


Рис 3.4. Энтальпия влажного воздуха. Сочи, 1-10 июля 2015

Алгоритм 3.2.

- 1) Вычисляется значение P .
- 2) Моделируется реализация вектора

$$\left(\vec{T}^T, \vec{H}^T \right)^T = (T_1, T_2, \dots, T_{8d}, H_1, H_2, \dots, H_{8d})^T$$

по Алгоритму 2.1 (параграф 2.1 Главы 2).

- 3) Построенная реализация вектора $\left(\vec{T}^T, \vec{H}^T \right)^T$ используется для вычисления значений вектора \vec{J}^T по формуле (3.7).

Моделирование временных рядов индекса жары

Если требуется смоделировать вектор $\vec{HI}^T = (HI_1, HI_2, \dots, HI_{8d})^T$ значений индекса жары на том же временном интервале длительностью d суток, то, поскольку ИЖ есть функция температуры воздуха и его относительно влажности, это может быть сделано по Алгоритму 3.3:

Алгоритм 3.3.

- 1) Моделируется реализация вектора

$$\left(\bar{T}^T, \bar{H}^T\right)^T = (T_1, T_2, \dots, T_{8d}, H_1, H_2, \dots, H_{8d})^T$$

по Алгоритму 2.1 (параграф 2.1 Главы 2).

2) Построенная реализация вектора $\left(\bar{T}^T, \bar{H}^T\right)^T$ используется для вычисления значений вектора \bar{HI}^T по формуле (3.3).

Верификация моделей временных рядов энтальпии влажного воздуха и индекса жары

Автором диссертационного исследования в [148] было показано, что построенные таким образом реализации временных рядов ЭВВ и ИЖ близки по различным статистическим свойствам к реальным рядам ЭВВ и ИЖ. Ниже приведены несколько характеристик, использованных для сравнения. Оценка параметров стохастического «генератора погоды» для моделирования вектора $\left(\bar{T}^T, \bar{H}^T\right)^T$ и оценка характеристик реальных рядов ЭВВ и ИЖ проводились на основе данных наблюдений, собранных в период с 1993 по 2015 гг. на 18 метеостанциях, расположенных на территории РФ в различных климатических зонах. Для оценки тех же вероятностных характеристик по модельным данным использовались 10^6 траекторий комплекса $\left(\bar{T}^T, \bar{H}^T\right)^T$.

Одной из характеристик, по которой проводилось сравнение реальных и модельных рядов ЭВВ, является их корреляционная матрица. В качестве примера, на рисунке 3.5 приведены оценки коэффициентов корреляции $corr(J_1, J_{1+h})$ и $corr(J_i, J_{i+1})$. Для всех рассмотренных метеостанций и интервалов времени оценки коэффициентов корреляции по модельным траекториям отклоняются от оценок по реальным данным менее, чем на 3σ , а в большинстве случаев – менее, чем на σ .

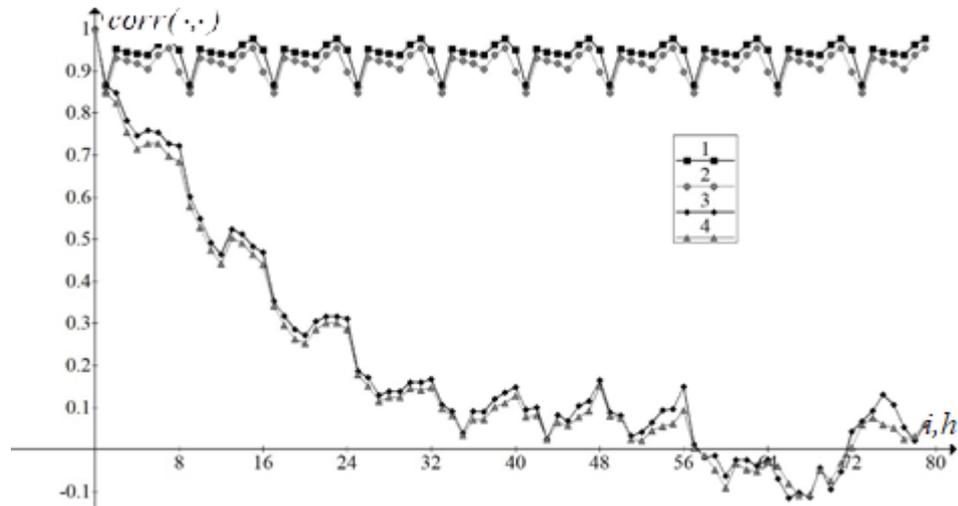


Рис. 3.5. Оценки коэффициентов корреляции. Кривые 1 и 2 – оценки $corr(J_i, J_{i+1})$ по реальным и модельным данным, соответственно; кривые 3, 4 – оценки $corr(J_1, J_{1+h})$ по реальным и модельным данным. Томск, 1-10 июля.

Ещё одной характеристикой, которая была использована для исследования, насколько стохастическая модель комплекса $(\vec{T}^T, \vec{H}^T)^T$ пригодна для моделирования рядов ЭВВ и изучения их свойств, являлось среднее число $N(lev)$ суток на рассматриваемом временном интервале длительностью d дней, в которые минимальная за сутки энтальпия выше уровня lev кДж/кг. В таблице 3.2 приведены оценки $N(lev)$, полученные по реальным и модельным данным. Для всех рассмотренных метеостанций, интервалов времени и уровней lev оценка данной характеристики по модельным траекториям попадает в соответствующие интервалы $(N(lev) - 3\sigma, N(lev) + 3\sigma)$, и, более того, в большинстве случаев – в интервалы $(N(lev) - \sigma, N(lev) + \sigma)$. Поэтому можно считать, что данная характеристика хорошо воспроизводится моделью.

Как и для временных рядов ЭВВ, для рядов ИЖ было проведено сравнение оценок различных характеристик процесса по реальным и модельным данным. Одной из характеристик, которая была использована для верификации модели, построенной по Алгоритму 3.3, являлась вероятность

Таб. 3.2. Оценки среднего числа $N(lev)$ дней с минимальной энтальпией за сутки выше lev кДж/кг. Томск.

lev	1-10 июня		1-10 июля		1-10 августа	
	РД, $N(lev) \pm \sigma$	МД	РД, $N(lev) \pm \sigma$	МД	РД, $N(lev) \pm \sigma$	МД
20	6.435 ± 0.604	6.385	10.000 ± 0.062	9.948	10.000 ± 0.087	9.904
24	4.739 ± 0.626	4.470	9.609 ± 0.228	9.479	9.522 ± 0.220	9.472
28	3.348 ± 0.554	2.839	8.348 ± 0.422	8.234	8.348 ± 0.402	8.282
32	2.130 ± 0.427	1.599	7.000 ± 0.535	6.860	6.261 ± 0.543	6.262
36	0.826 ± 0.284	0.751	5.217 ± 0.578	5.156	4.044 ± 0.550	3.978
40	0.087 ± 0.154	0.250	3.522 ± 0.530	3.185	2.348 ± 0.434	2.067
44	0.000 ± 0.075	0.067	2.000 ± 0.389	1.556	0.913 ± 0.272	0.822
48	0.000 ± 0.035	0.018	0.693 ± 0.234	0.592	0.478 ± 0.148	0.278
52	0.000 ± 0.016	0.005	0.130 ± 0.121	0.177	0.174 ± 0.071	0.072
56	0.000 ± 0.001	0.001	0.000 ± 0.055	0.043	0.000 ± 0.019	0.015

$$p(k) = P(HI_i > k, i = 1, \dots, 8d).$$

Из таблицы 3.3 видно, что оценки по реальным и модельным данным близки. Для всех рассмотренных метеостанций и интервалов времени оценки вероятностей $p(k)$ по модельным траекториям отклоняются от оценок по реальным данным менее, чем на 3σ , а в 84% случаев – менее, чем на σ .

Ещё одной характеристикой, применённой для сравнения реальных и модельных рядов ИЖ, была доля $s(l, k)$ периодов длительностью k измерений с индексом жары выше $l^\circ C$ среди всех периодов длительностью k измерений. На рисунке 3.6 приведены примеры оценок $s(l, k)$ по реальным и модельным данным. Численные эксперименты показали, что оценки по модельным данным для всех рассмотренных метеостанций, временных интервалов, уровней l и длительностей

k не отклоняются от соответствующих оценок по реальным данным более чем на 2σ , а в большинстве случаев отклонение не превышает σ .

Таб. 3.3. Оценки вероятности $p(k)$.

k	Сочи, 1-10 мая		Томск, 1-10 июля	
	РД, $p(k) \pm \sigma$	МД	РД, $p(k) \pm \sigma$	МД
12	0.707 ± 0.037	0.709	0.910 ± 0.021	0.910
16	0.342 ± 0.042	0.326	0.730 ± 0.037	0.737
20	0.107 ± 0.026	0.111	0.511 ± 0.040	0.499
24	0.021 ± 0.011	0.028	0.303 ± 0.034	0.298
28	0.002 ± 0.003	0.004	0.144 ± 0.025	0.149
32	0.000 ± 0.001	0.000	0.054 ± 0.014	0.052
36	0.000 ± 0.000	0.000	0.008 ± 0.006	0.014
40	0.000 ± 0.000	0.000	0.000 ± 0.003	0.004
44	0.000 ± 0.000	0.000	0.000 ± 0.001	0.001

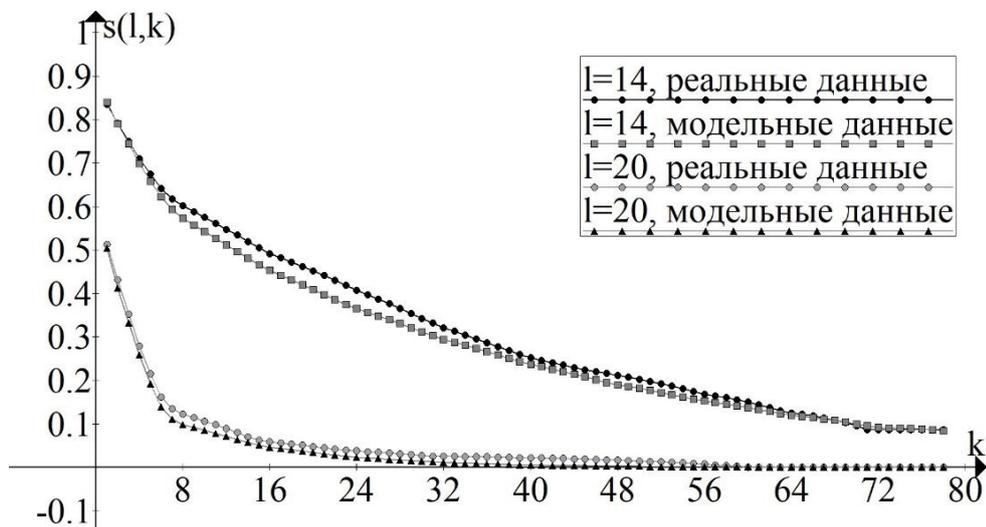


Рис. 3.6. Оценки $s(l, k)$. Томск, 1-10 июля.

В таблице 3.4 приведены оценки среднесуточного индекса жары $HI_{24}(i)$ в i -ые сутки, полученные по реальным и модельным данным. Для всех

рассмотренных метеостанций и временных интервалов различия оценками среднесуточного индекса жары по реальным и модельным данным не превышают 3σ .

Таб. 3.4. Оценки среднесуточного индекса жары $HI_{24}(i)$ в i -ые сутки.

i	Сочи, 1-10 мая		Томск, 1-10 июня	
	РД, $HI_{24}(i) \pm \sigma$	МД	РД, $HI_{24}(i) \pm \sigma$	МД
1	13.998 ± 0.702	14.638	13.890 ± 1.179	13.979
3	13.731 ± 0.714	14.616	13.041 ± 1.202	13.995
5	13.960 ± 0.697	14.611	12.069 ± 1.190	14.002
7	15.227 ± 0.705	14.604	15.378 ± 1.193	13.975
9	15.237 ± 0.695	14.611	15.748 ± 1.182	13.966

Следует отметить, что не все характеристики временных рядов ЭВВ и ИЖ хорошо воспроизводятся рассматриваемой моделью совместных рядов температуры и относительной влажности воздуха. Так, например, вероятность $P_{24}(\Delta)$ изменения энтальпии влажного воздуха более чем на Δ кДж/кг внутри 24-часового интервала воспроизводится плохо, и лишь в 26% рассмотренных случаев оценка $P_{24}(\Delta)$ по модельным данным попадает в интервал $(P_{24}(\Delta) - 3\sigma, P_{24}(\Delta) + 3\sigma)$ оценки этой характеристики по реальным значениям. Отметим, что при этом вероятность $P_3(\Delta)$ изменения энтальпии влажного воздуха более чем на Δ кДж/кг за 3 часа (т.е. вероятность события $|J_i - J_{i+1}| > \Delta$) для $\Delta < 10$ кДж/кг воспроизводится с высокой точностью (в 73% экспериментов оценка по модельным данным отклоняется от оценки по реальным значениям менее, чем на σ).

Моделирование временных рядов биоклиматического индекса суровости метеорежима

Перейдём к обсуждению подхода, основанного на применении стохастического «генератора погоды», к моделированию временных рядов биоклиматического индекса суровости метеорежима. Напомним, что, согласно определению по формуле (3.2), БИСМ является функцией приземной температуры воздуха T °C, атмосферного давления P гПа, модуля скорости ветра V м/с, относительной влажности воздуха H %, а также высоты L м местности над уровнем моря. Очевидно, что параметр L зависит только от географического расположения точки, для которой проводится моделирование, и не зависит от времени. Соответственно, в рамках рассматриваемого подхода, для вычисления модельных значений временного ряда $\vec{B} = (B_1, B_2, \dots, B_{8d})^T$ БИСМ необходимо моделировать совместные временные ряды

$$\left(\vec{T}^T, \vec{H}^T, \vec{V}^T, \vec{P}^T \right)^T = (T_1, \dots, T_{8d}, H_1, \dots, H_{8d}, V_1, \dots, V_{8d}, P_1, \dots, P_{8d})^T$$

4 метеопараметров. Но, как и в случае с моделированием ЭВВ, можно считать, что суточный и сезонный ход давления слабо выражены, и при вычислении БИСМ вместо значений давления $P_i, i = 1, \dots, 8d$ использовать среднее на рассматриваемом временном интервале давление P , оцененное по реальным данным. В этом случае, моделирование временного ряда \vec{B} сводится к реализации Алгоритма 3.4:

Алгоритм 3.4.

- 1) Вычисляется значение P .
- 2) Согласно Алгоритму 2.1 моделируется реализация вектора

$$\left(\vec{T}^T, \vec{H}^T, \vec{V}^T \right)^T = (T_1, \dots, T_{8d}, H_1, \dots, H_{8d}, V_1, \dots, V_{8d})^T.$$

3) Построенная реализация вектора $\left(\vec{T}^T, \vec{H}^T, \vec{V}^T\right)^T$, вычисленное значение P и известная высота L используются для вычисления значений вектора \vec{B} по формуле (3.2).

Верификация модели временных рядов биоклиматического индекса суровости метеорежима

В работе [94] приведены результаты верификации предложенной модели временных рядов \vec{B} и показано, что модельные траектории близки по своим статистическим характеристикам к реальным рядам БИСМ. В качестве примера, приведём 2 характеристики, использованные для верификации модели. Оценки характеристик процесса по модельным данным проводились по 10^6 траекторий, а для оценки входных параметров модели и реальных характеристик процесса были использованы данные метеонаблюдений в период с 1966 по 2016 гг.

В таблице 3.5 приведены оценки по реальным и модельным значениям среднего числа $AN(b_1, b_2)$ дней на рассматриваемом d -дневном интервале, в которые среднесуточный БИСМ находился в заданном интервале (b_1, b_2) . Рассматривались интервалы (b_1, b_2) , характеризующие уровень комфортности / дискомфорта метеоусловий. Сравнение оценок показывает, оценки по реальным и модельным данным близки друг к другу.

Для верификации модели также было проведено сравнение оценок по реальным и модельным траекториям среднего числа $AD(lev)$ дней на рассматриваемом d -дневном интервале, в которые минимальный за сутки БИСМ ниже заданного уровня lev (см. таблицу 3.6). Численные эксперименты показали, что оценка $AD(lev)$ по модельным данным попадает в интервал $(AD(lev) - \sigma, AD(lev) + \sigma)$, построенный по реальным данным, для 86% для рассмотренных временных интервалов и уровней lev .

Таб. 3.5. Оценки среднего числа $AN(b_1, b_2)$ дней, в которые среднесуточный БИСМ находится в заданном интервале (b_1, b_2) .

(b_1, b_2)	Пограничный, 1-15 июля		Томск, 16-30 января	
	РД, $AN(b_1, b_2) \pm \sigma$	МД	РД, $AN(b_1, b_2) \pm \sigma$	МД
[8,10]	0.065 ± 0.048	0.107	0.000 ± 0.000	0.000
[7,8)	13.065 ± 0.241	13.245	0.000 ± 0.000	0.004
[6,7)	1.870 ± 0.234	1.645	0.900 ± 0.028	1.094
[5,6)	0.000 ± 0.000	0.000	8.460 ± 0.531	8.500
[4,5)	0.000 ± 0.000	0.000	5.440 ± 0.533	5.159
[0,4)	0.000 ± 0.000	0.000	0.200 ± 0.001	0.244

Таб. 3.6. Оценки среднего числа $AD(lev)$ дней, в которые минимальный за сутки БИСМ ниже заданного уровня lev .

lev	Пограничный, 1-15 июля		Томск, 16-30 января	
	РД, $AD(lev) \pm \sigma$	МД	РД, $AD(lev) \pm \sigma$	МД
10.0	15.000 ± 0.000	15.000	15.000 ± 0.000	15.000
9.0	15.000 ± 0.004	14.999	11.580 ± 0.281	11.425
8.5	15.000 ± 0.010	14.996	6.600 ± 0.357	5.707
8.0	14.978 ± 0.023	14.981	2.680 ± 0.255	2.170
7.0	13.783 ± 0.204	13.707	0.120 ± 0.042	0.073
6.0	3.957 ± 0.348	3.824	0.000 ± 0.002	0.000
5.0	0.022 ± 0.036	0.049	0.000 ± 0.002	0.000

Таким образом, численные эксперименты показывают, что подход к моделированию БИ, основанный на использовании стохастического «генератора

погоды», может быть применён для исследования тех свойств рядов БИ, которые не могут быть детально исследованы по реальным данным в силу малого объёма их выборки.

Исследование зависимости оценок характеристик рядов ЭВВ и ИЖ от входных параметров

В тех случаях, когда характеристика реального процесса хорошо воспроизводится моделью, модельные траектории можно использовать для исследования чувствительности рассматриваемой характеристики к погрешностям оценивания входных параметров стохастической модели.

В качестве иллюстрации приведём результаты следующего численного эксперимента. Моделирование вектора $\left(\bar{T}^T, \bar{H}^T\right)^T$ осуществлялось с корреляционной матрицей, оцененной по реальным данным, и одномерными распределениями относительной влажности, описанными в параграфе 2.1 Главы 2, а параметры одномерных распределений температуры варьировались:

Случай А. Параметры определены как в параграфе 2.1 Главы 2.

Случаи Б, В, Г. Дисперсия температуры и веса, определяющие смеси нормальных распределений, определены как в параграфе 2.1 Главы 2, а математическое ожидание увеличено на $0.1\text{ }^{\circ}\text{C}$, $0.5\text{ }^{\circ}\text{C}$ и $1\text{ }^{\circ}\text{C}$, соответственно.

На рисунке 3.7 и в таблицах 3.7-3.9 приведены оценки некоторых характеристик вектора \vec{J} энтальпии влажного воздуха и вектора \overline{HI} индекса жары, вычисленные по модельным реализациям комплекса $\left(\bar{T}^T, \bar{H}^T\right)^T$ для случаев А-Г. Оценки проводились по 10^6 реализациям. Из представленных таблиц видно, что среднее EJ_k , дисперсия $DJ_k, k=1, \dots, 8$ ЭВВ и среднее число $N(lev)$ существенно изменяются даже при малом изменении среднего значения температуры (Случай Б). Из этого следует, что эти характеристики чувствительны к изменению

рассматриваемого входного параметра модели вектора $(\bar{T}^T, \bar{H}^T)^T$. Рассмотренные характеристики индекса жары (вероятность $p(k)$ и доля $s(l, k)$) также оказались чувствительными к изменениям среднего значения температуры.

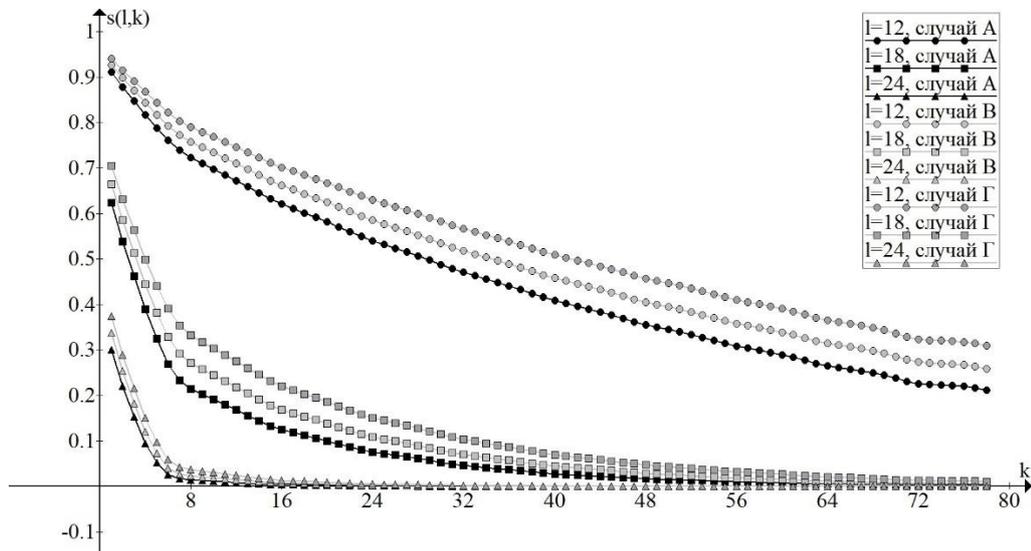


Рис. 3.7. Оценки $s(l, k)$. Томск, 1-10 июля.

Таб. 3.7. Оценки среднего EJ_k и дисперсии DJ_k ЭВВ, $k=1, \dots, 8$. Томск, 1-10 июля.

k	EJ_k				DJ_k			
	А	Б	В	Г	А	Б	В	Г
1	38.773	39.028	40.072	41.393	62.313	62.897	64.790	67.245
2	45.728	46.002	47.077	48.434	90.332	90.853	93.599	97.029
3	48.066	48.326	49.368	50.682	99.954	100.683	103.530	107.196
4	48.462	48.718	49.749	51.047	100.859	101.622	104.565	108.346
5	48.245	48.501	49.552	50.873	106.282	107.235	110.338	114.257
6	44.586	44.849	45.913	47.255	89.966	90.866	93.634	97.274
7	41.194	41.454	42.511	43.858	73.436	74.187	76.416	79.241
8	38.779	39.030	40.073	41.407	68.004	68.724	70.682	73.378

Таб. 3.8. Оценки среднего числа $N(lev)$ дней с минимальной энтальпией за сутки выше lev кДж/кг. Томск, 1-10 июля.

lev	А	Б	В	Г
20	9.948	9.955	9.974	9.987
24	9.479	9.523	9.674	9.812
28	8.234	8.301	8.610	8.972
32	6.860	6.935	7.261	7.641
36	5.156	5.264	5.713	6.233
40	3.185	3.304	3.805	4.434
44	1.556	1.646	2.031	2.586
48	0.592	0.643	0.865	1.203
52	0.177	0.198	0.292	0.449
56	0.043	0.050	0.079	0.135

Таб. 3.9. Оценки вероятности $p(k)$. Томск, 1-10 июля.

k	А	Б	В	Г
12	0.910	0.913	0.925	0.939
16	0.737	0.744	0.773	0.806
20	0.499	0.508	0.542	0.586
24	0.298	0.305	0.333	0.371
28	0.149	0.154	0.178	0.208
32	0.052	0.055	0.070	0.091
36	0.014	0.016	0.021	0.031
40	0.004	0.004	0.006	0.010
44	0.001	0.001	0.002	0.003

Замечание 3.1. С точки зрения статистической метеорологии, представленный выше эксперимент с изменением средних значений температуры может быть истолкован как реализация простейшего сценария изменения климата.

В такой трактовке, предложенная стохастическая модель позволила исследовать зависимость свойств временных рядов ЭВВ и ИЖ от изменений климата.

3.2.2. Стохастическое моделирование временных рядов среднесуточного индекса жары

В этом параграфе будет продемонстрировано применение стохастического «генератора погоды» для моделирования нестационарных временных рядов среднесуточного индекса жары (СИЖ).

В первую очередь необходимо определить, что такое среднесуточный индекс жары. К определению этого показателя есть 2 подхода, суть которых отображена на рисунке 3.8. В этом параграфе и всюду далее среднесуточные и среднеполусуточные значения биоклиматических индексов определяются с использованием 2-го подхода.

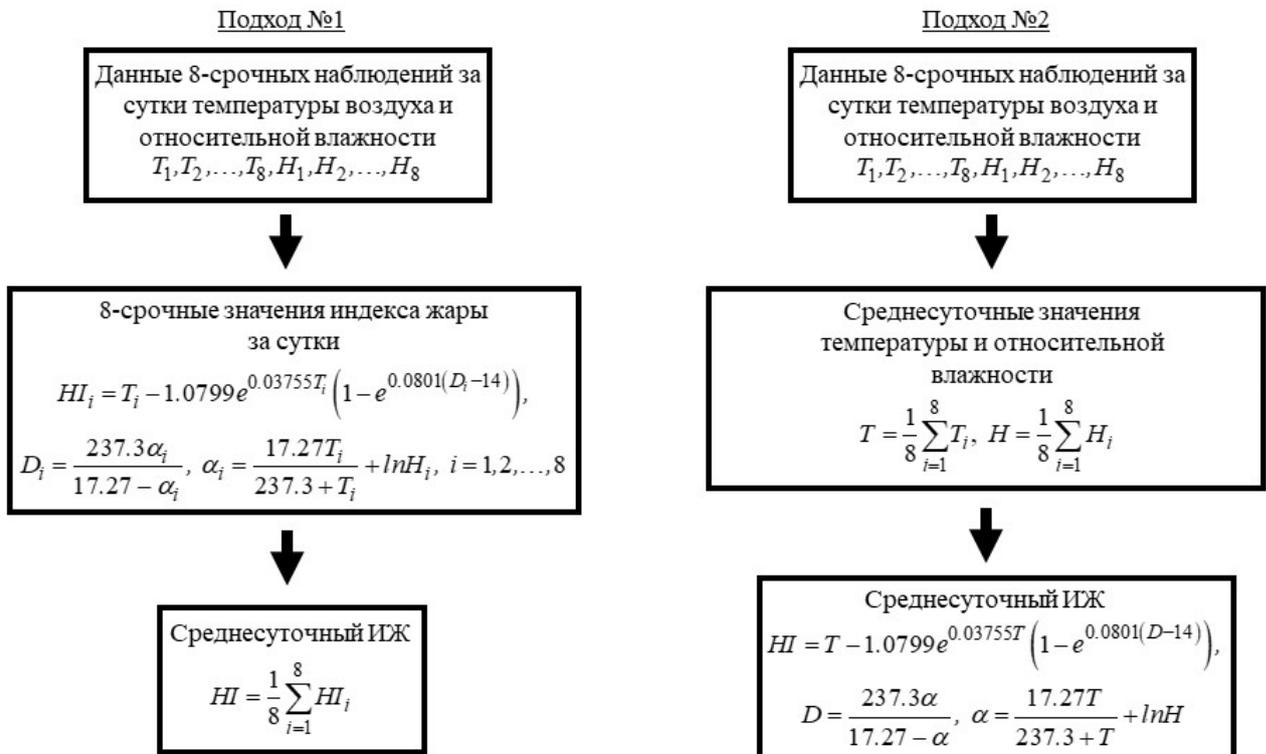


Рис. 3.8. Блок-схемы 2 подходов к определению среднесуточного ИЖ.

Описание модели и алгоритм моделирования

Пусть $T_i^j, H_i^j, i=1,2,\dots,N, j=1,2,\dots,NY$ – реальные значения среднесуточной температуры (СТ) и среднесуточной относительной влажности (СОВ) на интервале длительностью N дней, собранные за NY лет,

$$\tilde{R} = \begin{pmatrix} \tilde{R}_{TT} & \tilde{R}_{TH} \\ \tilde{R}_{HT} & \tilde{R}_{HH} \end{pmatrix}$$

есть выборочная корреляционная матрица, оцененная по $T_i^j, H_i^j, i=1,2,\dots,N, j=1,2,\dots,NY$, размерность которой $2N \times 2N$. Если $2N \geq NY$, то для обеспечения невырожденности матрицы \tilde{R} при оценке её элемента $\tilde{r}_{XY}(i,k), X,Y \in \{T,H\}, i,k=1,2,\dots,N$, как и ранее, используем скользящее симметричное окно осреднения шириной $2L+1$, где $L = \arg \min_{l \in \mathbb{Z}^+ \cup \{0\}} (2N < (2l+1)NY)$.

При построении модели будем считать, что плотности одномерных распределений СТ и СОВ есть смеси

$$t_i(x) = \theta_i \frac{1}{\sigma_{i1} \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_{i1})^2}{2\sigma_{i1}^2}\right) + (1 - \theta_i) \frac{1}{\sigma_{i2} \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_{i2})^2}{2\sigma_{i2}^2}\right), \quad x \in \mathbb{R},$$

$$h_i(x) = \nu_i \frac{x^{\alpha_{i1}-1} (1-x)^{\beta_{i1}-1}}{B(\alpha_{i1}, \beta_{i1})} + (1 - \nu_i) \frac{x^{\alpha_{i2}-1} (1-x)^{\beta_{i2}-1}}{B(\alpha_{i2}, \beta_{i2})}, \quad x \in [0,1]$$

2 гауссовских и 2 бета-распределений соответственно, где

$$\mu_{i1}, \mu_{i2} \in \mathbb{R}, \quad \sigma_{i1}, \sigma_{i2}, \alpha_{i1}, \beta_{i1}, \alpha_{i2}, \beta_{i2} > 0,$$

$$0 < \theta_i, \nu_i < 1, \quad i = 1, 2, \dots, N,$$

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx, \quad a, b > 0.$$

Параметры смесей $t_i(x), h_i(x)$ выбираем с помощью алгоритма, предложенного в [40] и EM-алгоритма [36]. Такой выбор плотностей $t_i(x), h_i(x)$ хорошо согласуется с выборочными плотностями одномерных распределений $\tilde{t}_i(x), \tilde{h}_i(x)$ СТ и СОВ в смысле критерия Пирсона. Моделирование траектории

$$HI^n = (HI_1^n, \dots, HI_N^n), n = 1, 2, \dots, NT$$

временного ряда $\overline{HI}^T = (HI_1, HI_2, \dots, HI_N)^T$ СИЖ на N -дневном временном интервале в рамках первого подхода осуществляется по Алгоритму 3.5 (здесь NT – общее число моделируемых траекторий) [142].

Алгоритм 3.5.

1) Моделирование траектории $(T^n, H^n) = (T_1^n, \dots, T_N^n, H_1^n, \dots, H_N^n)$ совместных временных рядов $(\overline{T}^T, \overline{H}^T)^T$ СТ и СОВ с плотностями одномерных распределений $t_i(x), h_i(x), i = 1, 2, \dots, N$ и корреляционной матрицей \tilde{R} .

2) Вычисление траектории $HI^n = (HI_1^n, \dots, HI_N^n)$ по формуле (3.3):

$$D_i^n = \frac{237.3\alpha_i^n}{17.27 - \alpha_i^n}, \alpha_i^n = \frac{17.27T_i^n}{237.3 + T_i^n} + \ln H_i^n,$$

$$HI_i^n = T_i^n - 1.0799e^{0.03755T_i^n} \left(1 - e^{0.0801(D_i^n - 14)} \right).$$

Замечание 3.2. В данном случае на 1-ом шаге Алгоритма 3.5 моделирование негауссовского процесса осуществлялось методом обратных функций распределения.

Замечание 3.3. В общем случае, временные ряды $(\overline{T}^T, \overline{H}^T)^T$ являются нестационарными, и корреляционная матрица \tilde{R} не является ни тёплицевой, ни блочно-тёплицевой. При относительно небольших значениях N (в пределах 8-12 дней), моделирование рядов $(\overline{T}^T, \overline{H}^T)^T$ можно проводить в предположении о их стационарности. В этом случае каждый из блоков $\tilde{R}_{XY}, X, Y \in \{T, H\}$ является тёплицевым, но вся матрица \tilde{R} тёплицевой не будет.

Для краткости изложения, результаты применения Алгоритма 3.5. приведены в следующем параграфе, где проведено сравнение рассматриваемого подхода к моделированию (моделирование с использованием «генератора погоды») с подходом, основанном на непосредственном (прямом) моделировании БИ.

3.3. Моделирование временных рядов биоклиматических индексов на основе определяющих формул

3.3.1. Стохастические модели безусловных рядов биоклиматических индексов

Перейдём к описанию второго подхода к моделированию временных рядов БИ. Напомним, что этот подход на основе определяющих формул не включает в себя шага, на котором моделируются совместные временные ряды метеорологических параметров.

Моделирование временных рядов СИЖ

Сначала применим этот подход для моделирования временных рядов СИЖ [142]. Пусть, по-прежнему, $T_i^j, H_i^j, i=1,2,\dots,N, j=1,2,\dots,NY$ – реальные значения СТ и СОВ на интервале длительностью N дней, собранные за NY лет. По формуле (3.3) вычислим реальные значения $HI_i^j, i=1,2,\dots,N, j=1,2,\dots,NY$ СИЖ, оценим по ним выборочную корреляционную матрицу \tilde{R}_{HI} размерности $N \times N$ и выборочные плотности одномерных распределений $\tilde{g}_i(x)$. Как и ранее, ширина $(2L+1)$ окна осреднения в том случае выбирается из условия $L = \arg \min_{l \in \mathbb{Z}^+ \cup \{0\}} (N < (2l+1)NY)$.

Выборочные плотности $\tilde{g}_i(x)$ аппроксимируем смесями $g_i(x)$ 2 гауссовских распределений с параметрами, определенными по алгоритму, описанному в [40].

На рисунке 3.9 приведены примеры гистограмм $\tilde{g}_i(x)$ и аппроксимирующих плотностей распределения $g_i(x)$, полученных по данным наблюдений 3 июля в период с 1977 по 2017 ($NY = 41, L = 1$). Отметим, что первоначально в численных экспериментах в качестве альтернативы смесям 2 гауссовских распределений были использованы также смеси 2 логистических распределений и 2 распределений Лапласа. Напомним, что плотности одномерных распределений случайных величин с логистическим распределением и распределением Лапласа имеют вид

$$\text{logist}(x) = \frac{e^{-(x-\mu)/s}}{s(1 + e^{-(x-\mu)/s})^2}, \mu \in \mathbb{R}, s > 0,$$

$$\text{lapl}(x) = \frac{\alpha}{2} e^{-\alpha|x-\beta|}, \alpha > 0, \beta \in \mathbb{R},$$

соответственно. Оказалось, что при аппроксимации смеси $\tilde{g}_i(x)$ смесями этих распределений уступают использованным в итоге смесям $g_i(x)$ в смысле критерия Пирсона.

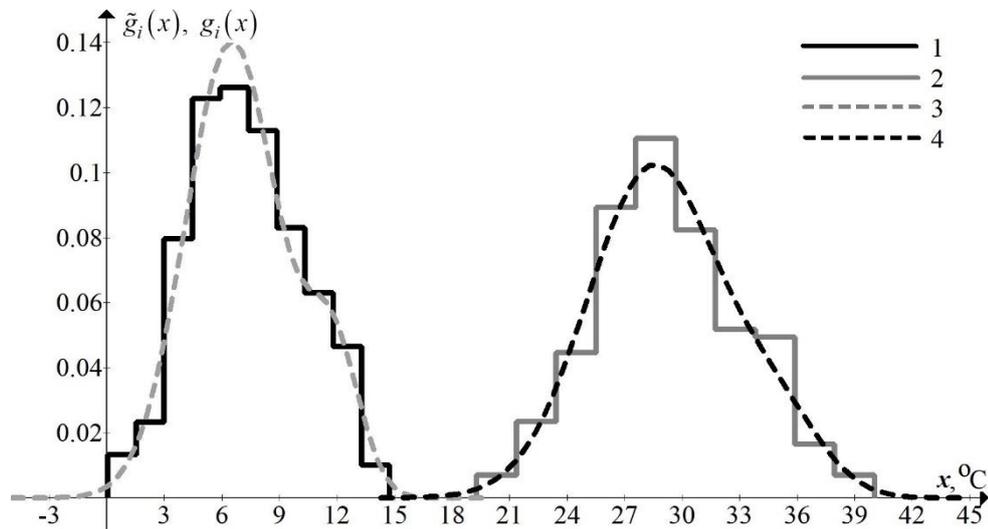


Рис. 3.9. Графики $\tilde{g}_i(x)$ и $g_i(x)$. Кривые 1 и 2 – гистограммы $\tilde{g}_i(x)$ на метеостанциях в Сулаке и Дербенте, кривые 3 и 4 – плотности $g_i(x)$ на метеостанциях в Сулаке и Дербенте.

Моделирование траектории $HI^n = (HI_1^n, \dots, HI_N^n)$ временного ряда СИЖ в рамках второго подхода (назовём это Алгоритмом 3.6) сводится к моделированию нестационарного негауссовского вектора с заданными одномерными распределениями $g_i(x), i = 1, 2, \dots, N$ и корреляционной матрицей \tilde{R}_{HI} . В данной работе моделирование проводилось методом обратных функций распределения.

Верификация и сравнение 2 моделей

Проведём сравнение рассмотренного подхода к моделированию временных рядов СИЖ с подходом, предложенным в параграфе 3.2 Главы 3. Сравнение будем проводить по точности и по трудоёмкости моделирования.

Для сравнения 2 подходов к моделированию по точности, было проведено сравнение оценок некоторых характеристик временных рядов \overline{HI}^T . В качестве примера, в таблицах 3.10 и 3.11 приведены оценки среднего числа $AN(lev)$ дней в рассматриваемом временном промежутке с СИЖ выше lev^oC и оценки вероятностей $p(l) = P(|HI_i - HI_{i+1}| > l)$. В целом, модель, построенная на основе определяющих формул (подход № 2), достаточно хорошо воспроизводит свойства реальных временных рядов СИЖ. По всем проведённым численным экспериментам (по данным с разных станций, в различные периоды времени, с различными значениями N и оценками различных характеристик), процент попадания оценок по модельным данным в доверительные интервалы шириной 2σ при использовании этого подхода равен 68%, а в доверительные интервалы шириной 6σ – 89%. Модель, предложенная в параграфе 3.2 Главы 3 (основанная на применении «генератора погоды»), воспроизводит ряд характеристик так же хорошо, однако некоторые характеристики (например, $AN(lev)$) воспроизводятся неудовлетворительно.

Очевидно, что реализация второго подхода существенно менее трудоёмка по сравнению с реализацией первого подхода как на стадии определения входных

Таб. 3.10. Оценки среднего числа $AN(lev)$. Астрахань, 1-31 августа ($N = 31$, период наблюдений 1966-2000, $NY = 35$, $L = 0$), $NT = 10^7$.

lev	Оценка по реальным данным, $AN(lev) \pm \sigma$	Подход №1	Подход №2
20	28.206 ± 0.451	27.193	28.120
24	21.088 ± 0.857	14.463	21.053
28	10.441 ± 0.946	1.032	10.323
32	2.853 ± 0.529	0.007	2.758
36	0.500 ± 0.213	0.000	0.566
40	0.118 ± 0.085	0.000	0.138
44	0.029 ± 0.048	0.000	0.056
48	0.000 ± 0.011	0.000	0.004

Таб. 3.11. Оценки вероятностей $p(l)$. Сочи, 1-30 июня ($N = 30$, период наблюдений 1993-2015, $NY = 23$, $L = 1$), $NT = 10^7$.

l	Оценка по реальным данным, $p(l) \pm \sigma$	Подход №1	Подход №2
1	0.585 ± 0.021	0.577	0.638
2	0.304 ± 0.022	0.275	0.360
3	0.150 ± 0.018	0.111	0.183
4	0.078 ± 0.013	0.040	0.084
5	0.039 ± 0.008	0.013	0.035

параметров модели, так и при численной реализации. Действительно, в первом случае необходимо определить $10N$ параметров плотностей $t_i(x)$, $h_i(x)$, $i = 1, 2, \dots, N$ и оценить корреляционную матрицу размерности $2N \times 2N$, а во втором – $5N$ параметров одномерных распределений $g_i(x)$, $i = 1, 2, \dots, N$ и матрицу размерности $N \times N$. При численной реализации метода обратных функций

в первом случае приходится пересчитывать $N(2N-1)$ коэффициентов корреляции вспомогательного гауссовского процесса, во втором – $N(N-1)/2$ коэффициентов. Количество обращений к датчику случайны чисел отличается в 2 раза.

Моделирование временных рядов БИСМ

Вышеописанный подход на основе определяющих формул несложно применить и для моделирования временных рядов других БИ. Так в работе [141] была предложена и верифицирована модель временных рядов $\vec{B} = (B_1, B_2, \dots, B_{365})^T$ среднесуточного БИСМ на интервале годовой длины.

Из определения БИСМ прямо следует, что временные ряды \vec{B} являются негауссовскими. Кроме этого, на годовом интервале эти ряды нестационарны, причём годовой ход БИСМ отчётливо выражен на большинстве рассмотренных метеостанций. На рисунке 3.10 приведены графики реальных значений БИСМ на годовом интервале, а на рисунке 3.11 – оценки математического ожидания EB_i этого БИ (оценки проводились по данным за 1966-2015 гг. в Томске и Сочи, 1992-2020 в Мурманске и Териберке).

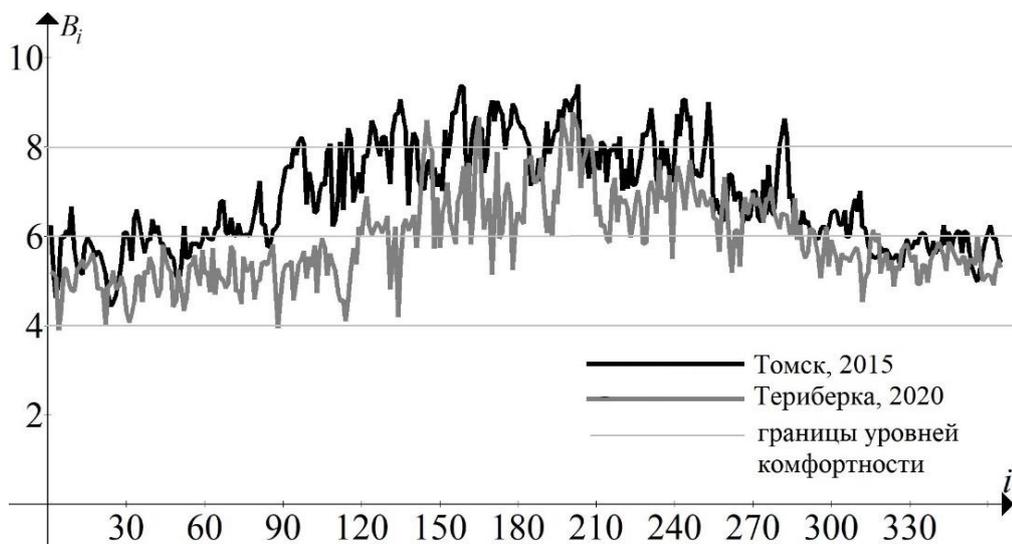


Рис. 3.10. Графики реальных значений БИСМ.

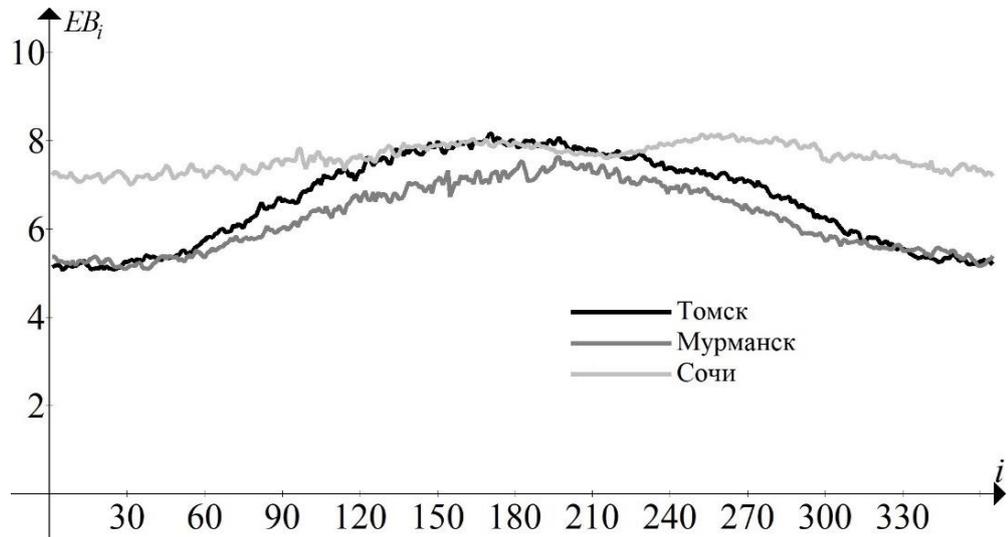


Рис. 3.11. Графики оценок среднего значения EB_i БИСМ.

Чрезвычайно важно отметить существенное отличие вычисления БИСМ в этом случае от случая, рассмотренного в параграфе 3.2 Главы 3. На годовом интервале атмосферное давление изменяется достаточно сильно (внутригодовые колебания на некоторых станциях достигают 90 гПа), и для расчёта БИСМ нельзя использовать усреднённое давление, как в ранее описанной модели (Алгоритм 3.4). В рассматриваемом случае БИСМ необходимо вычислять по формуле (3.2).

Моделирование временного ряда \vec{B} в предположении о его нестационарности и негауссовости осуществляется методом обратных функций распределения.

Для реализации этого подхода необходимо оценить по реальным данным B_i^j , $i = 1, 2, \dots, 365$, $j = 1, 2, \dots, NY$ корреляционную матрицу \tilde{R} процесса и определить плотности одномерных распределений $s_i(x)$, $i = 1, 2, \dots, 365$. Здесь, как и ранее, NY — количество лет, в которые проводились наблюдения на метеостанции. Численные эксперименты показали, что в качестве одномерных плотностей распределения $s_i(x)$, $i = 1, 2, \dots, 365$ среднесуточного БИСМ можно выбрать смеси 2 бета-распределений:

$$s_i(x) = \theta_i \frac{1}{10B(\alpha_{1i}, \beta_{1i})} \left(\frac{x}{10}\right)^{\alpha_{1i}-1} \left(1 - \left(\frac{x}{10}\right)\right)^{\beta_{1i}-1} +$$

$$+(1-\theta_i) \frac{1}{10B(\alpha_{2i}, \beta_{2i})} \left(\frac{x}{10}\right)^{\alpha_{2i}-1} \left(1-\left(\frac{x}{10}\right)\right)^{\beta_{2i}-1},$$

$$0 \leq \theta_i \leq 1, x \in [1, 10], i = 1, 2, \dots, 365,$$

которые в смысле критерия Пирсона хорошо аппроксимируют выборочные одномерные распределения $\tilde{s}_i(x), i = 1, 2, \dots, 365$. Параметры $\theta_i, \alpha_{1i}, \beta_{1i}, \alpha_{2i}, \beta_{2i}$ плотностей $s_i(x), i = 1, 2, \dots, 365$ были выбраны методом максимального правдоподобия. В выборочной корреляционной матрице $\tilde{R} = (\tilde{r}_{ij})_{i,j=1}^{365}$ большое количество элементов (в среднем, около 90 %) близки к 0. Это объясняется тем, что дальние корреляционные связи реального процесса очень слабы. Можно заменить в \tilde{R} малые значения коэффициентов корреляции нулями. Обозначим полученную при этом матрицу $R_0 = (r_{0ij})_{i,j=1}^{365}$. Как показали численные эксперименты, матрица R_0 оказывается отрицательно определенной. Для моделирования процесса с такой матрицей необходимо предварительно её регуляризовать. Однако регуляризации R_0 чрезвычайно сильно искажает значения ненулевых элементов матрицы: так, например, относительная погрешность $\frac{|\tilde{r}_{12} - r_{012}|}{|\tilde{r}_{12}|} \cdot 100\%$ достигает 50%. В связи с этим, при моделировании временных рядов БИСМ была использована выборочная корреляционная матрица \tilde{R} . Как и ранее, при построении выборочных одномерных распределений и оценивании выборочной корреляционной матрицы использовалось симметричное скользящее окно осреднения.

Верификация модели временных рядов БИСМ

Модель успешно прошла верификацию. Так, например, оценки по модельным данным среднего числа дней в году, когда среднесуточный БИСМ принадлежит заданному интервалу $[a, b]$ для всех рассмотренных метеостанций принадлежат интервалу $(AN - 2\sigma, AN + 2\sigma)$, где AN – оценка этой

характеристики по реальным данным. В качестве иллюстрации, оценки AN по реальным и модельным данным для метеостанций, расположенных в Томске и Сочи приведены в таблице 3.12. На рисунке 3.12 приведены значения среднего минимального $A_{min}(i)$ и среднего максимального $A_{max}(i)$ БИСМ за месяц. Отклонения оценок этих характеристик по реальным и модельным данным для 16 из 18 рассмотренных метеостанций не превышают σ . В таблицах 3.13, 3.14 приведены оценки доли $s(v, n)$ временных интервалов длительностью n дней с БИСМ не превышающим v среди всех n -дневных интервалов. Эта характеристика реального процесса также хорошо воспроизводится моделью.

Таб. 3.12. Оценки среднего числа дней в году, когда среднесуточный БИСМ принадлежит заданному интервалу $[a, b]$.

$[a, b]$	Томск		Сочи	
	РД, $AN \pm \sigma$	МД	РД, $AN \pm \sigma$	МД
$[8, 10]$	52.240 ± 2.042	53.197	105.440 ± 3.371	105.737
$[7, 8)$	99.360 ± 1.750	97.708	193.800 ± 2.918	193.401
$[6, 7)$	87.700 ± 1.963	88.122	64.000 ± 1.760	63.401
$[5, 6)$	97.400 ± 1.698	97.081	1.740 ± 0.315	2.407
$[4, 5)$	27.300 ± 1.919	27.886	0.020 ± 0.033	0.053
$[0, 4)$	1.000 ± 0.248	1.005	0.000 ± 0.003	0.000

Исследование изменчивости характеристик временных рядов БИСМ

Представляет интерес вопрос о том, как изменяются различные характеристики временных рядов БИСМ по времени (отражаются ли климатические изменения на характеристиках рядов БИСМ). Для исследования изменчивости простых характеристик (например, средних значений) можно провести следующий эксперимент. Для тех метеостанций, на которых имеется

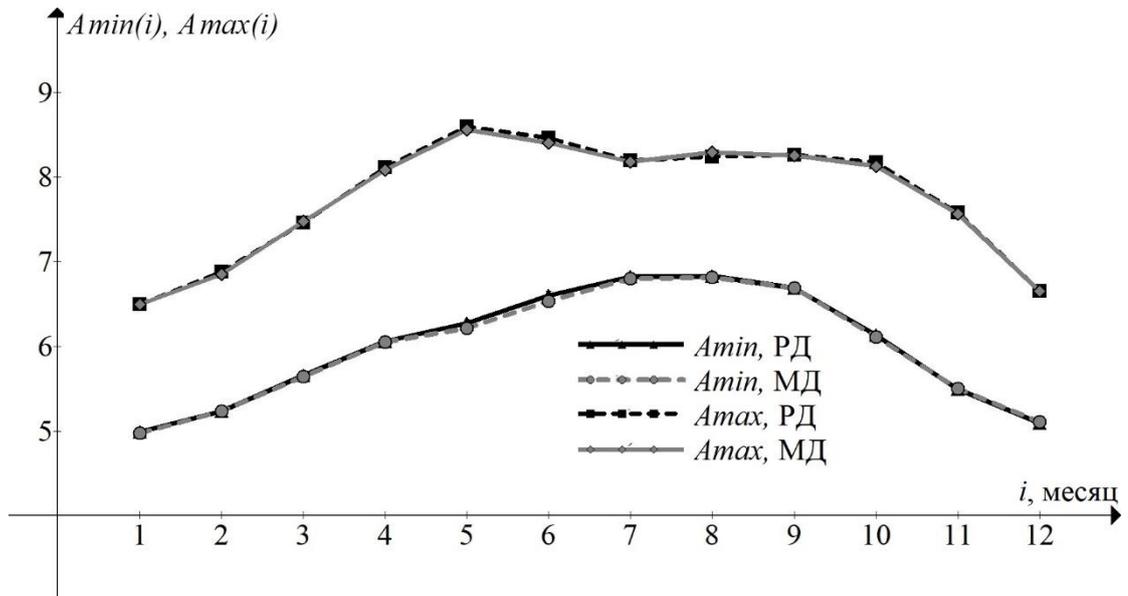


Рис. 3.12. Оценки $A_{min}(i)$ и $A_{max}(i)$. Пограничный (по данным за 1966-2015 гг.).

Таб. 3.13. Оценки доли $s(v, n)$. Томск.

n	$\nu = 8$		$\nu = 6$	
	РД, $s(v, n) \pm 3\sigma$	МД	РД, $s(v, n) \pm 3\sigma$	МД
2	0.803 ± 0.020	0.797	0.888 ± 0.021	0.887
4	0.733 ± 0.023	0.722	0.821 ± 0.030	0.818
6	0.685 ± 0.024	0.673	0.769 ± 0.036	0.763
8	0.649 ± 0.025	0.638	0.727 ± 0.039	0.719
10	0.621 ± 0.025	0.610	0.692 ± 0.042	0.682
12	0.598 ± 0.025	0.588	0.663 ± 0.044	0.650
14	0.579 ± 0.026	0.569	0.636 ± 0.044	0.622

достаточно большой объём данных наблюдений, разделим выборку реальных данных на 2 равные части. В таблице 3.15 приведены оценки математического ожидания μ БИСМ в Томске по 2 массивам данных: по данным за 1966-1990 гг. (РД-1) и 1991-2015 гг. (РД-2). В таблице 3.15 приведены также оценки выборочной

Таб. 3.14. Оценки доли $s(v, n)$. Пограничный.

n	$\nu = 8$		$\nu = 7$	
	РД, $s(v, n) \pm 3\sigma$	МД	РД, $s(v, n) \pm 3\sigma$	МД
2	0.888 ± 0.021	0.887	0.412 ± 0.017	0.411
4	0.821 ± 0.030	0.818	0.338 ± 0.015	0.338
6	0.769 ± 0.036	0.763	0.305 ± 0.015	0.304
8	0.727 ± 0.039	0.719	0.283 ± 0.016	0.281
10	0.692 ± 0.042	0.682	0.266 ± 0.016	0.264
12	0.663 ± 0.044	0.650	0.251 ± 0.017	0.249
14	0.636 ± 0.044	0.622	0.236 ± 0.017	0.235

дисперсии БИСМ по этим 2 выборкам. При использовании РД-1 и РД-2 для оценивания математического ожидания и дисперсии было применено окно осреднения шириной $2L+1=15$ дней, и объём выборок, по которым оценивались значения μ , составлял $n=375$ (25 лет наблюдений \times на 15 дней осреднения) элементов. Пусть μ_1 и μ_2 – оценки математического ожидания μ по РД-1 и РД-2, соответственно. Проверим при уровне значимости 0.05 нулевую гипотезу $\mu_1 = \mu_2$ при конкурирующей гипотезе $\mu_1 \neq \mu_2$. Поскольку объём выборок достаточно велик, для сравнения оценок μ_1 и μ_2 по 2 независимым выборкам из генеральной совокупности, распределенной не нормально, можно воспользоваться приближенным критерием, описанном в [15]. Согласно этому критерию, вычислим величину

$$Z' = \frac{|\mu_1 - \mu_2|}{\sqrt{s_1/n + s_2/n}},$$

где s_1, s_2 – выборочные дисперсии, оцененные по РД-1 и РД-2, соответственно. При заданном уровне значимости критическое значение $Z = 1.64$. При $|Z'| \leq Z$ нулевую гипотезу принимаем, при $|Z'| > Z$ – отвергаем. Как показали численные

эксперименты, примерно в половине случаев гипотеза $\mu_1 = \mu_2$ отвергается, и можно сказать, что климатические изменения повлияли на средние значения БИСМ.

Таб. 3.15. Оценки математического ожидания μ и дисперсии s^2 БИСМ. Томск.

день	Оценки математического ожидания		Оценки дисперсии		Z'
	РД-1, $\mu \pm 3\sigma$	РД-2, $\mu \pm 3\sigma$	РД-1, s_1	РД-2, s_2	
1 января	5.132±0.072	5.125±0.096	0.218	0.381	0.175
1 марта	5.623±0.093	5.861±0.072	0.362	0.214	-6.073
1 мая	7.293±0.134	7.283±0.097	0.751	0.388	0.182
1 июля	7.893±0.112	7.859±0.081	0.518	0.272	0.741
1 сентября	7.235±0.067	7.433±0.088	0.185	0.320	-5.396
1 ноября	5.985±0.085	6.183±0.060	0.326	0.149	-5.563

Для более сложных характеристик критерии, по которым можно сравнивать оценки по РД-1 и РД-2 неизвестны. В этом случае, можно воспользоваться предложенной и верифицированной моделью временных рядов БИСМ. Независимо друг от друга по РД-1 и РД-2 были оценены входные параметры вышеописанной модели, построены реализации временных рядов и была проведена верификация. В таблице 3.16 приведены оценки среднего минимального $A_{min}(i)$ и среднего максимално $A_{max}(i)$ БИСМ за месяц, полученные по 2 ансамблям из 10^8 траекторий. Видно, что значения $A_{min}(i)$ и $A_{max}(i)$, полученные при использовании данных из РД-2, несколько выше, чем при использовании данных из РД-1. Можно сказать, что во второй из рассматриваемых периодов времени климатические условия в смысле БИСМ были несколько мягче.

Таб. 3.16. Оценки $A_{min}(i)$ и $A_{max}(i)$. Томск.

i , месяц	$A_{min}(i)$		$A_{max}(i)$	
	РД-1	РД-2	РД-1	РД-2
Январь	4.22	4.35	5.95	6.12
Март	5.08	5.27	7.20	7.44
Май	5.99	6.40	8.80	9.04
Июль	6.84	6.93	8.79	8.75
Сентябрь	6.15	6.23	8.20	8.30
Ноябрь	4.78	4.88	5.95	6.12

3.3.2. Стохастические модели условных рядов биоклиматических индексов с точечными и интервальными условиями

Моделирование условных рядов среднесуточного ИЖ и БИСМ

В параграфе 1.3 Главы 1 были рассмотрены алгоритмы моделирования условных негауссовских процессов. Эти алгоритмы были использованы для моделирования и исследования временных рядов различных биоклиматических индексов. Так, например, вышеописанный подход был применен для моделирования условных рядов $\vec{HI} = (HI_1, HI_2, \dots, HI_N)$ среднесуточного индекса жары. В работе [139] приведены результаты верификации моделей условных рядов СИЖ с точечными и интервальными условиями, и показано, что модельные ряды близки по своим вероятностным свойствам к реальным при $N \leq 15$.

Рассмотрим два примера применения предложенных моделей условных негауссовских процессов для исследования тех характеристик процесса, которые не могут быть исследованы по реальным данным в силу малого объема реальной выборки.

Пример 3.1. Пусть $m(j)$ – условное среднее СИЖ HI_j , $j > i$ при условии $HI_k > a_k$, $k = \overline{1, i}$. На рисунке 3.13 приведены оценки $m(j)$, полученные по 10^7

реализациям модели условных временных рядов с интервальными условиями (Алгоритм 1.3) при условиях

$$a_k = EHI_k, a_k = EHI_k + \sqrt{DHI_k}, a_k = EHI_k + 2\sqrt{DHI_k},$$

где EHI_k, DHI_k – математическое ожидание и дисперсия безусловного процесса $\vec{HI} = (HI_1, HI_2, \dots, HI_N)$ с функциями одномерного распределения $G_k(x), k = 1, 2, \dots, N$. Для наглядности приведены также значения математического ожидания безусловного процесса.

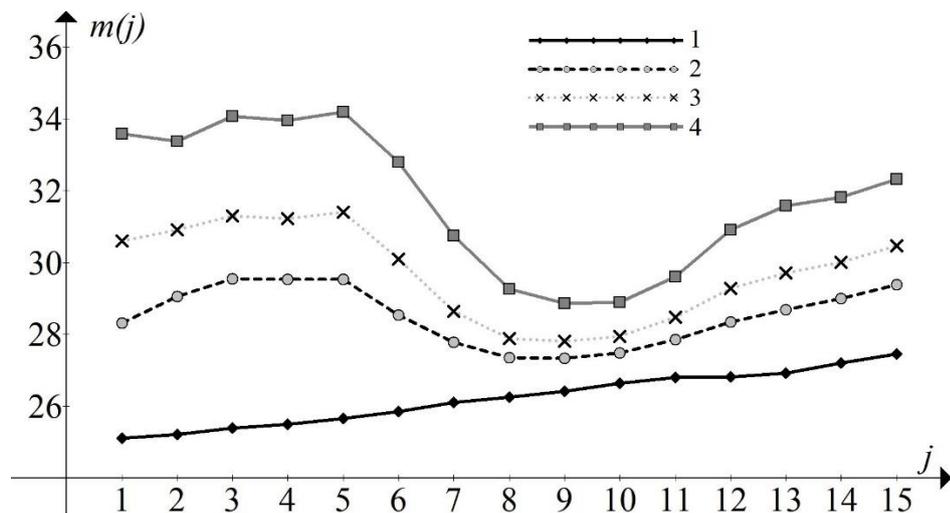


Рис. 3.13. Условное среднее СИЖ. Кривая 1 – средние значения безусловного процесса. Кривые 2, 3, 4 – условные средние при условиях

$a_k = EHI_k, a_k = EHI_k + \sqrt{DHI_k}, a_k = EHI_k + 2\sqrt{DHI_k}$, соответственного. Сочи, 1-15 июля ($N = 15$), $i = 5$.

Пример 3.2. С использованием Алгоритма 1.2 были построены 10^7 реализаций условных временных рядов $\vec{B} = (B_1, B_2, \dots, B_N)$ среднесуточного БИСМ с точечными условиями $B_k = a_k, k = \overline{1, i}$, и по ним были проведены оценки вероятности попадания свободных компонент $B_j, j > i$ в интервалы, характеризующие уровень комфорта / дискомфорта погодных условий. В таблице 3.17 приведены примеры оценок этих вероятностей. Фактически, в таблице 3.17

представлен один из вариантов вероятностного прогноза, при котором по данным $B_k = a_k$, $k = \overline{1, i}$ численно определяются вероятности наступления того или иного события (в данном случае – принадлежности значения среднесуточного БИСМ выбранному интервалу комфортности).

Таб. 3.17. Оценки вероятности попадания B_j в заданный интервал $[s, t)$. Томск, 1-30 января ($N = 30$), $i = 10$, a_k – реальные значения среднесуточного БИСМ 1-10 января 2015.

	$P(B_j \in [s, t))$			
	$B_j \in [0, 4)$	$B_j \in [4, 5)$	$B_j \in [5, 6)$	$B_j \in [6, 7)$
11.01 ($j = 11$)	0.000	0.691	0.309	0.000
12.01 ($j = 12$)	0.188	0.718	0.095	0.000
13.01 ($j = 13$)	0.024	0.678	0.298	0.000
14.01 ($j = 14$)	0.007	0.829	0.164	0.000
15.01 ($j = 15$)	0.290	0.681	0.029	0.000
16.01 ($j = 16$)	0.608	0.388	0.005	0.000
17.01 ($j = 17$)	0.010	0.820	0.170	0.000
18.01 ($j = 18$)	0.079	0.732	0.189	0.000
19.01 ($j = 19$)	0.024	0.331	0.621	0.024
20.01 ($j = 20$)	0.004	0.308	0.658	0.030

Моделирование условных рядов ИХС и вероятностный прогноз

По аналогии с вышеописанными моделями безусловных и условных рядов индекса жары, в работах [138, 154] были предложены и верифицированы модели безусловных и условных временных рядов $\vec{W} = (W_1, W_2, \dots, W_N)$ индекса холодного стресса с 3- и 12-часовым разрешением. Модели условных рядов, также как и для индекса жары, были построены для случаев точечных и интервальных условий.

Некоторые результаты верификации модели приведены в Приложении 3. Далее на примере модели условных временных рядов ИХС с 12-часовым разрешением будет исследована возможность применения стохастического подхода для прогнозирования значений индекса холодного стресса.

При численной реализации стохастической модели условного вектора мы получаем облако траекторий. Будем называть прогнозируемым значением полусуточного ИХС в момент времени i значение $E_{\text{cond}}W_i$ условного среднего, оцененного этому облаку. На рисунке 3.14 изображены условия, границы облака из 10^8 траекторий и оценённые значения условного среднего.

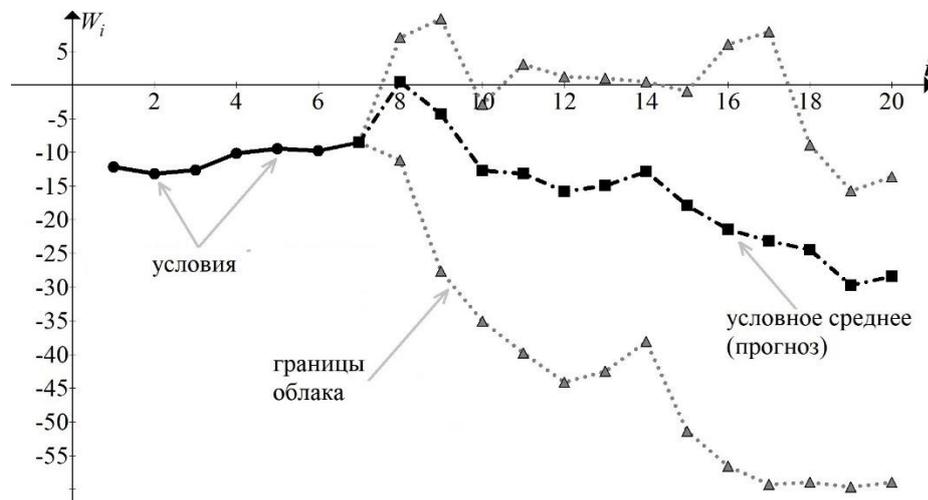


Рис. 3.14. Облако траекторий условного процесса.

Необходимо также определить критерий, согласно которому определяется оправдался прогноз или нет. Будем считать, что в момент времени i прогноз оправдался, если прогнозируемое значение $E_{\text{cond}}W_i$ и реальное значение \tilde{W}_i полусуточного ИХС принадлежат одному интервалу (rfl, rfr) . Границы интервалов определим согласно [18, 37]. Отметим, что границы этих интервалов связаны с оценкой рисков обморожения, и отличаются для жителей разных стран [18, 116]. Например, экстремальный холод, при котором обморожение незащищенных участков кожи происходит в течение 2 минут, характеризуется значениями ИХС ниже $-60\text{ }^{\circ}\text{C}$ и $-55\text{ }^{\circ}\text{C}$ для России и Канады, соответственно.

Обозначим

$$rate(i) = \frac{N_a(i)}{N_t} \times 100\%$$

долю оправдавшихся прогнозов для момента времени i , где N_t общее число сделанных прогнозов. а $N_a(i)$ – число оправдавшихся.

Следует отметить, что для оценки параметров модели и для исследования оправдываемости прогнозов следует использовать независимые массивы данных. Во всех численных экспериментах, результаты которых приведены ниже, для оценки параметров модели использовались данные, собранные на метеостанциях в период с 1967 по 2014 гг., а для исследования оправдываемости прогнозов – с 2015 по 2017 гг.

На рисунке 3.15 изображен график доли $rate(i)$ как функции i , оцененной по $N_t = 10^3$ прогнозов при числе точечных условий равным 13. Высокий уровень ($rate(i) > 80\%$) оправдываемости прогнозов на 1 сутки (первые 2 свободных от условий значения полусуточного ИХС) объясняется сильными корреляционными связями на таком временном промежутке. Далее, вместе с убыванием корреляций убывает и доля оправдавшихся прогнозов. Оправдываемость прогнозов на срок свыше 3 суток неудовлетворительно низкая ($rate(i) < 60\%$) при любом количестве условий.

На рисунке 3.16 изображен график доли $rate(i)$ как функции i , оцененной по $N_t = 10^3$ прогнозов при 13 интервальных условиях. Рассматривались 2 типа интервалов. Во-первых, интервалы $(a_j, b_j) = (rfl_j, rfr_j)$, $j \in \Omega$, где (rfl_j, rfr_j) – интервал, к которому принадлежит реальное значение \tilde{W}_j , $j \in \Omega$. Во-вторых, интервалы $(a_j, b_j) = (\tilde{W}_j - 1, \tilde{W}_j + 1)$, $j \in \Omega$ существенно меньшей длины, чем (rfl_j, rfr_j) . Как показали численные эксперименты, чем меньше длины интервалов

(a_j, b_j) , $j \in \Omega$, тем ближе доля оправдавшихся прогнозов к доле таких прогнозов при точечных условиях.

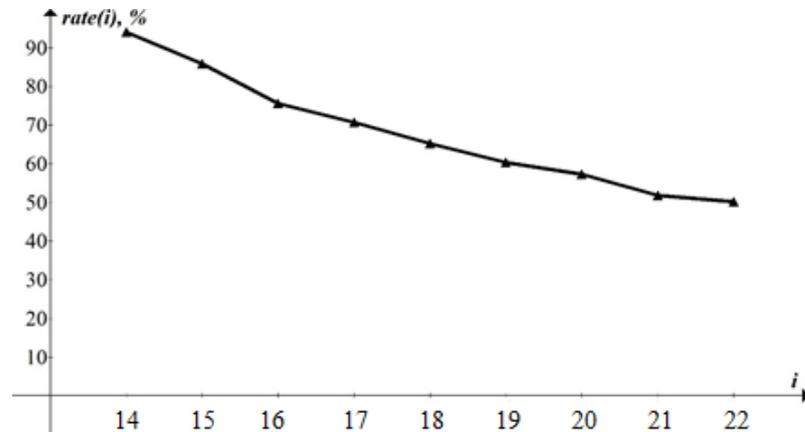


Рис. 3.15. Доля оправдавшихся прогнозов при точечных условиях.

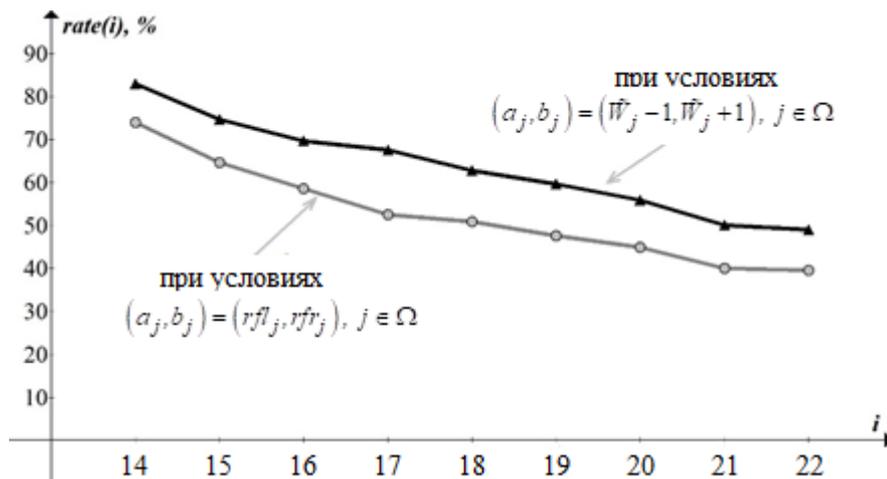


Рис. 3.16. Доля оправдавшихся прогнозов при интервальных условиях.

Необходимо отметить, что важно правильно выбрать количество условий. Численные эксперименты показали, что для модели с интервальными условиями оптимальное количество условий составляет от 10 до 15. При большем количестве условий, время моделирования значительно увеличивается, а качество прогноза не улучшается (значение $rate(i)$ не возрастает). Меньшее количество условий ухудшает точность прогноза. Следует отметить, что для большинства рассмотренных метеостанций и временных интервалов радиус корреляции временных рядов полусуточного ИХС составляет от 10 до 15 измерений, и это,

скорее всего, объясняет, почему указанное выше количество условий является оптимальным. То же самое и в случае точечных условий.

В целом, предложенные модели условных рядов полусуточного ИХС могут быть использованы для краткосрочного прогнозирования. Доля оправдавшихся долгосрочных прогнозов слишком мала.

3.4. Моделирование пространственных и пространственно-временных полей биоклиматических индексов

3.4.1. Моделирование пространственных и пространственно-временных полей индекса жары и индекса холодового стресса на сети метеостанций

Модель пространственно-временного поля СИЖ

Подход, примененный в параграфе 3.3 Главы 3 к моделированию безусловных рядов СИЖ можно применить не только для моделирования временных рядов, но и для моделирования пространственных и пространственно-временных полей этого биоклиматического индекса. Продемонстрируем это на подробно рассмотренном в [146] примере.

Построим стохастическую модель пространственно-временного поля СИЖ на 18 метеостанциях, расположенных между Чёрным, Азовским и Каспийским морями. На рисунке 3.17 представлена карта рассматриваемой территории и отмечены рассматриваемые метеостанции. Расстояние между станциями варьируется от 36 до 941 км.

Обозначим HI_s^t значение СИЖ в день t на метеостанции s . Рассмотрим пространственно-временное поле

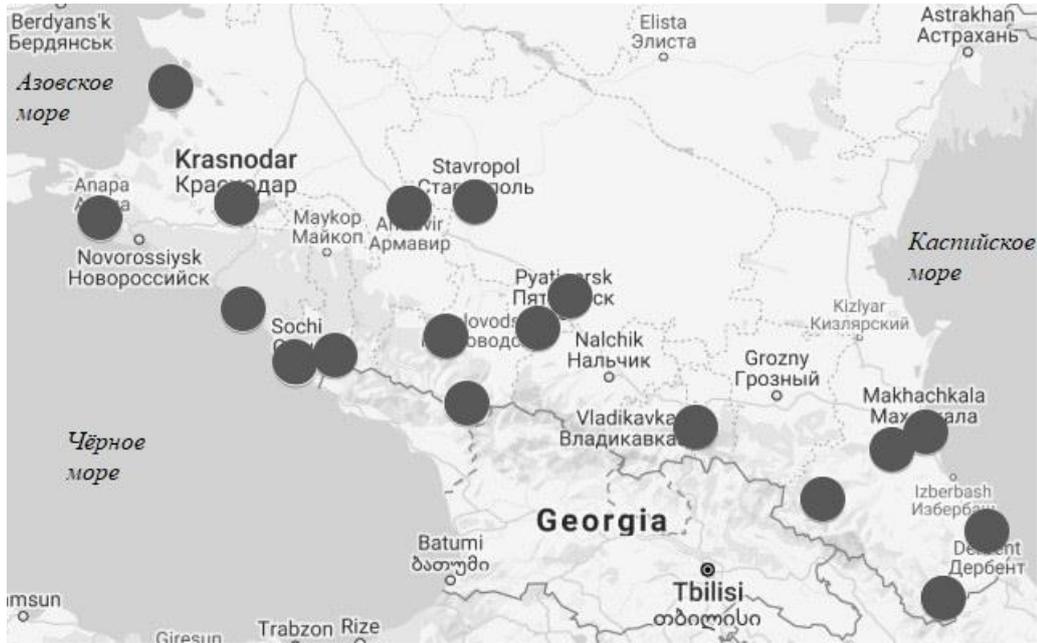


Рис. 3.17. Расположение метеорологических станций.

$$\overline{HI}^T = \left(HI_1^a, \dots, HI_{NS}^a, HI_1^{a+1}, \dots, HI_{NS}^{a+1}, \dots, HI_1^{a+N-1}, \dots, HI_{NS}^{a+N-1} \right)^T$$

СИЖ на сети $NS = 18$ станций на временном интервале длительностью N дней. Рассматриваемый временной интервал $[a, a + N - 1]$ принадлежит летнему периоду (с 1 июня по 31 августа).

В качестве модели реального поля СИЖ рассмотрим неоднородное по пространству и нестационарное по времени случайное поле с выборочной корреляционной матрицей \tilde{R}_{HI} размерности $18N \times 18N$, оцененной по значениям СИЖ, вычисленным по формуле (3.3) на основе данных метеонаблюдений, произведенных на станциях в период с 1977 по 2017, и плотностями одномерных распределений $g_{t,s}(x)$, представляющими собой смеси 2 гауссовских плотностей с параметрами, определенными как в параграфе 3.3 Главы 3 [146]. О неоднородности и нестационарности реального поля СИЖ свидетельствуют, в частности, оценки по реальным данным математического ожидания $E(HI_s^t)$ (см. рисунок 3.18).

Отметим две особенности выборочной корреляционной матрицы \tilde{R}_{HI} .

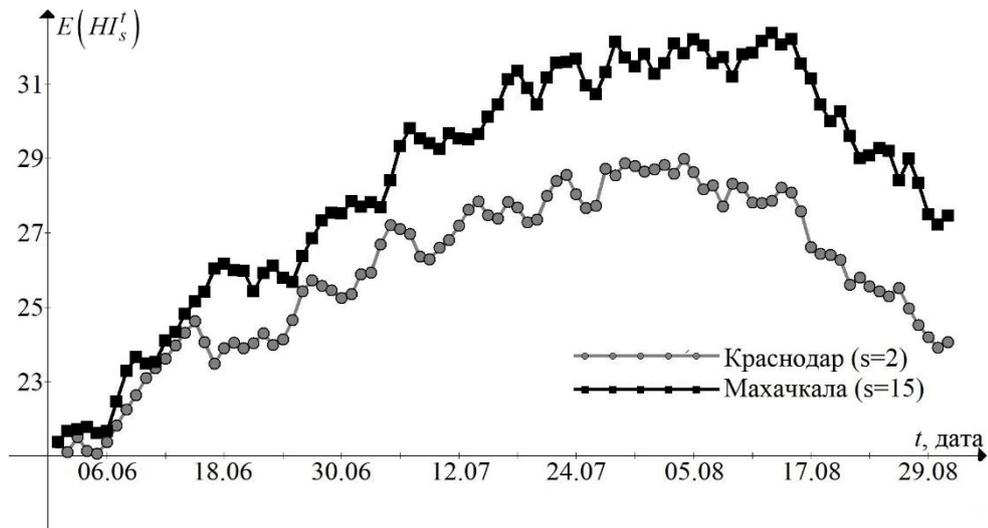


Рис. 3.18. Оценки математического ожидания $E(HI_s^t)$.

Во-первых, матрица \tilde{R}_{HI} не обладает какой-либо специальной структурой (например, матрица не является ни тёплицевой, ни разреженной), которая бы позволила применить экономичный по времени алгоритм моделирования. На рисунке 3.19 в качестве примера приведены оценки коэффициентов корреляции $corr(HI_s^t, HI_r^t)$ как функции расстояния между станциями. Видно, что даже при расстоянии между станциями s и r около 500 км значения СИЖ HI_s^t и HI_r^t сильно коррелированы.

Во-вторых, в большинстве случаев, матрица \tilde{R}_{HI} плохо обусловлена. В качестве примера, в таблице 3.18 приведены значения числа обусловленности $cond(\tilde{R}_{HI})$ матрицы \tilde{R}_{HI} при временном интервале длительностью N дней, начинающемся 1 июля.

Моделирование случайного поля \overline{HI}^T с корреляционной матрицей \tilde{R}_{HI} и плотностями одномерных распределений $g_{t,s}(x)$, $t = a, \dots, a + N - 1$, $s = 1, \dots, 18$ (и соответствующими им функциями одномерных распределений $G_{t,s}(x)$) осуществлялось 2 способами:

- 1) методом обратных функций распределения (МОФР),
- 2) методом нормализации (МН).

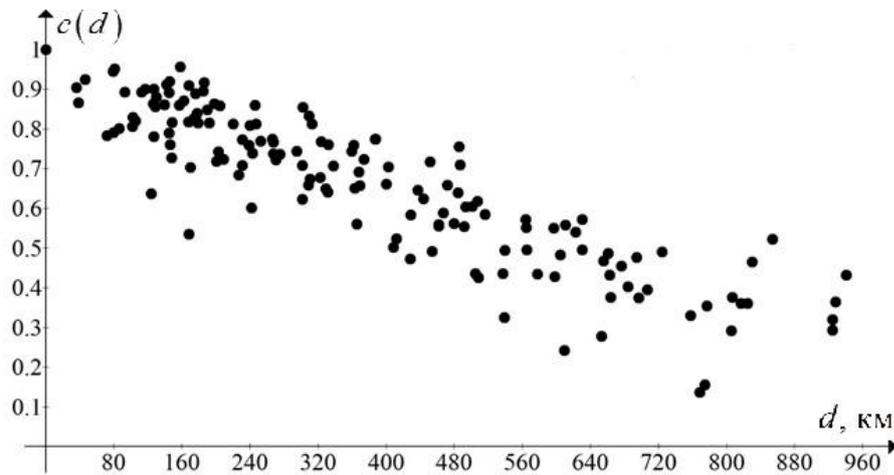


Рис.3.19. Выборочные коэффициенты корреляции $c(d) = \text{corr}(HI_s^t, HI_r^t)$ между всеми станциями как функция расстояния d между станциями, $t = 1$ июля.

Таб. 3.18. Число обусловленности $\text{cond}(\tilde{R}_{HI})$ матрицы \tilde{R}_{HI} .

N	Размерность матрицы \tilde{R}_{HI}	$\text{cond}(\tilde{R}_{HI})$
5	90×90	4.1×10^4
10	180×180	5.0×10^5
15	270×270	7.8×10^5
20	360×360	6.8×10^6

В работе [146] приведены результаты верификации предложенной модели поля \overline{HI}^T СИЖ при использовании этих 2 методов моделирования (часть результатов верификации представлена в Приложении 3). В целом, построенные траектории, обладают свойствами близкими, к свойствам реальных траекторий рассматриваемого поля. Ряд характеристик, например, вероятность $p_N(lev)$ того, что на протяжении N суток СИЖ будет выше заданного уровня lev °C на всех метеостанциях, воспроизводятся моделью, основанной на МОФР, несколько

точнее: оценки по траекториям, построенным с помощью МОФР, попадают в доверительные интервалы $(p_N(lev) - \sigma, p_N(lev) + \sigma)$ и $(p_N(lev) - 3\sigma, p_N(lev) + 3\sigma)$ оценок по реальным данным в 75% и 99% проведенных численных экспериментов, соответственно, а оценки по траекториям, полученным при использовании МН, – в 58% и 75 % экспериментов (при проведении численным экспериментов изменялись значения уровня lev °С, длительности N и начала временного интервала a).

В качестве иллюстрации характеристик, для исследования которых может быть применена предложенная модель, на рисунке 3.20 приведены графики оценки вероятности $pr(k, lev)$ того, что в заданный день хотя бы на k станциях СИЖ превысит уровень lev °С. Оценки получены по 10^6 траекторий, построенных с помощью МОФР. Из графиков видно, что вероятность того, что на большинстве станций (т.е. почти на всей рассматриваемой территории) СИЖ будет принимать значения, при которых людям необходимо соблюдать осторожность (СИЖ > 27°C), весьма значительна.

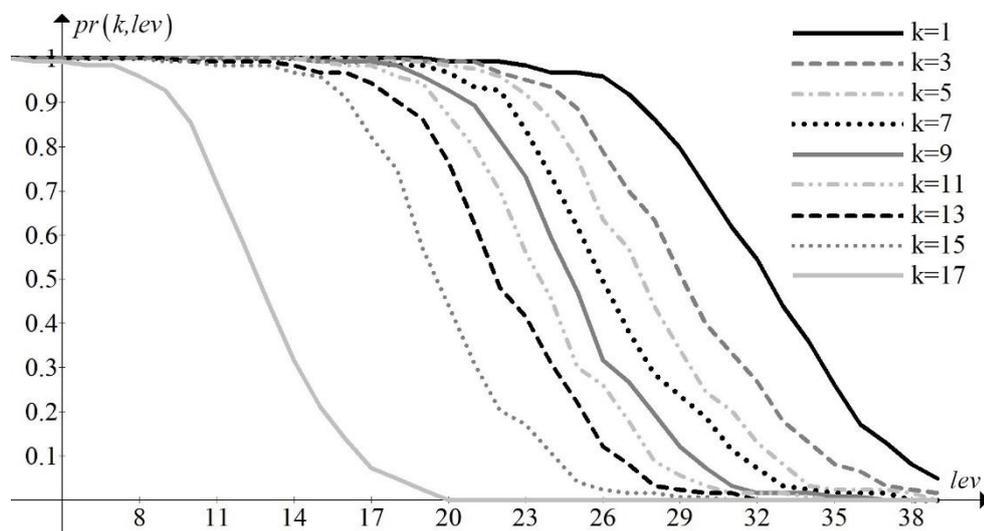


Рис. 3.20. Оценки вероятности $pr(k, lev)$. $N = 1$, $a = 10$ июля.

Сравнение двух методов моделирования

Перейдём к обсуждению вопросов, связанных с вычислительными особенностями, возникающими при реализации МОФР и МН.

На первом шаге реализации МОФР вычисляется корреляционная матрица R_{HI}^G вспомогательного гауссовского процесса. Как и матрица \tilde{R}_{HI} , матрица R_{HI}^G плохо обусловлена. В таблице 3.19 приведены примеры чисел обусловленности $cond(R_{HI}^G)$ матрицы R_{HI}^G при временном интервале длительностью N дней.

Таб. 3.19. Число обусловленности $cond(R_{HI}^G)$ матрицы R_{HI}^G .

N	Размерность матрицы R_{HI}^G	$cond(R_{HI}^G)$
5	90×90	2.2×10^5
10	180×180	6.1×10^5
15	270×270	3.2×10^6
20	360×360	7.1×10^6

При реализации 2-го шага МОФР моделируется гауссовский вектор с корреляционной матрицей R_{HI}^G . Для моделирования этого вектора необходимо представить матрицу R_{HI}^G в виде $R_{HI}^G = AA^T$, где A – нижнетреугольная матрица. Большое число обусловленности матрицы R_{HI}^G приводит к тому, что даже устойчивые алгоритмы разложения квадратной положительно определённой матрицы на произведение ниже- и верхнетреугольных матриц оказываются неприменимыми. Для решения этой проблемы проводилась описанная в [52] регуляризация матрицы R_{HI}^G , и гауссовский вектор моделировался с получившейся регуляризированной корреляционной матрицей G_{HI}^G . Численные эксперименты показали, что для того, чтобы матрица G_{HI}^G могла быть численно разложена в произведение $G_{HI}^G = A_G A_G^T$, достаточно брать $\varepsilon = 10^{-5}$, а искажение результатов

моделирования, вызванной заменой R_{HI}^G на G_{HI}^G , много меньше, чем искажения, вызванные статистической погрешностью оценивания матрицы \tilde{R}_{HI} по используемой выборке реальных данных.

При реализации метода нормализации возникает аналогичная проблема. Корреляционная матрица R_{HI}^N , оцененная по нормализованным значениям реального СИЖ, также оказывается плохо обусловленной (например, $cond(R_{HI}^N) \approx 6.2 \times 10^6$ при $N = 20$). Для регуляризации матрицы R_{HI}^N использовался тот же подход: матрица R_{HI}^N заменялась на R_{HI}^N '. Как и в случае с МОФР, погрешность, вносимая при регуляризации незначительна.

На рисунке 3.21 приведены графики первой строки корреляционной матрицы \tilde{R}_{HI} , оцененной по реальным данным, и первых строк корреляционных матриц, оцененных по 10^6 траекториям, полученным с помощью МОФР и МН.

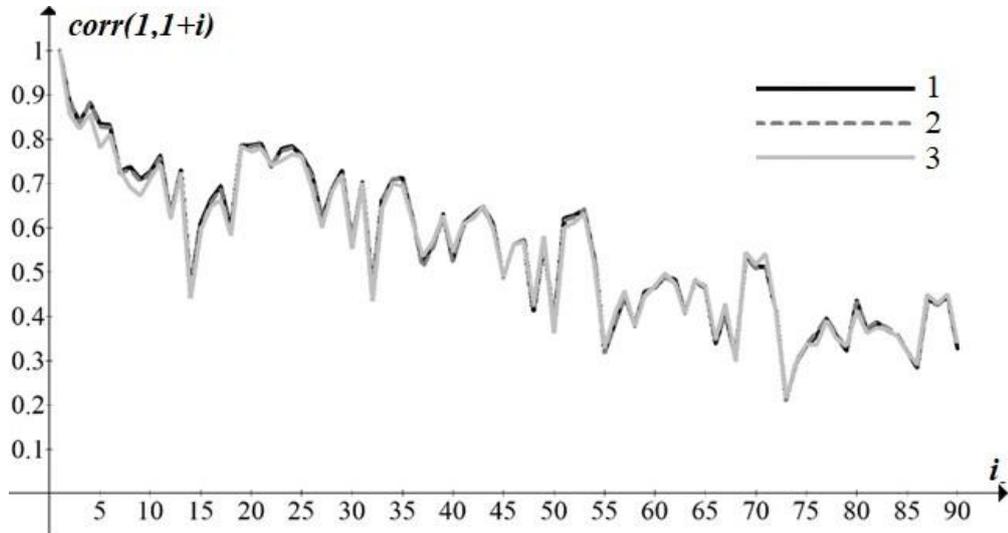


Рис. 3.21. Первые строки корреляционных матриц. Кривая 1 – оценка по реальным данным, кривая 2 – оценка по траекториям МОФР, кривая 3 – оценка по траекториям МН. $N = 5$, $a = 1$ августа.

Следует отметить, что моделирование методом нормализации гораздо менее затратно по времени, чем моделирование методом обратных функций

распределения. В таблице 3.20 приведено среднее время моделирования 10^4 траекторий поля в зависимости от числа N .

Таб. 3.20. Время моделирования, мин.

N	Размерность корреляционной матрицы	МОФР	МН
1	18×18	0.084	0.020
2	36×36	0.302	0.042
5	90×90	1.818	0.120
10	180×180	6.976	0.316
20	360×360	28.016	1.172
50	900×900	175.731	8.047

Модели пространственного и пространственно-временного полей среднесуточного ИХС

Наряду с моделью пространственно-временного поля СИЖ были построены стохастические модели пространственного поля $\vec{W}_s^T = (W_1, W_2, \dots, W_{NS})^T$ среднесуточного ИХС на сети NS станций и пространственно-временного поля $\vec{W}_{st}^T = (W_1^1, \dots, W_{NS}^1, W_1^2, \dots, W_{NS}^2, \dots, W_1^{ND}, \dots, W_{NS}^{ND})^T$ на сети NS станций на временном интервале длительностью ND дней [145, 153]. Рассматривались метеостанции, расположенные в 2 областях. Первая область (Область 1) расположена между $54^\circ 54'$ и $57^\circ 54'$ с.ш., $82^\circ 04'$ и $86^\circ 13'$ в.д. Приблизительные размеры этой области $260 \text{ км} \times 240 \text{ км}$. На этой территории расположено 7 метеостанций ($NS = 7$). Вторая область (Область 2) имеет существенно большую площадь. Область 2 расположена между $53^\circ 20'$ и $56^\circ 20'$ с.ш., $76^\circ 01'$ и $88^\circ 18'$ в.д. (размер Области 2 примерно $790 \text{ км} \times 330 \text{ км}$). В Области 2 расположено 11 метеостанций (что

соответствует $NS = 11$). На рисунке 3.22 показано расположение метеостанций в рассматриваемых областях.

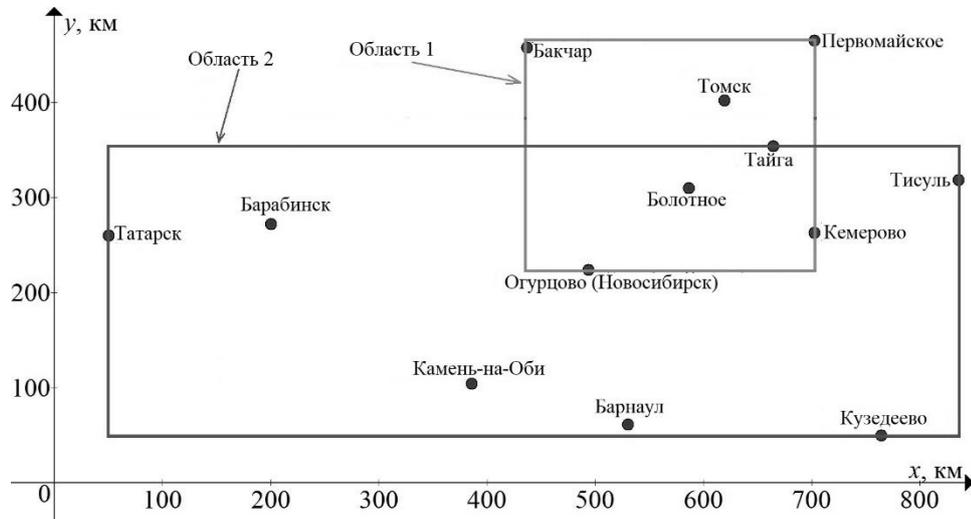


Рис. 3.22. Расположение метеостанций в рассматриваемых областях.

Необходимо отметить, что пространственные поля среднесуточного ИХС на рассматриваемых территориях являются неоднородными, а пространственно-временное поле является ещё и нестационарным. В качестве иллюстрации, на рисунке 3.23 приведены оценки математического ожидания EW_i^j среднесуточного ИХС на метеостанциях, расположенных в Области 2, в заданные моменты времени. На рисунке 3.24 представлены оценки математического ожидания EW_i^j и стандартного отклонения $\sqrt{DW_i^j}$ среднесуточного ИХС как функции времени на метеостанции Болотное. Здесь и далее для вычисления реальных значений ИХС по формуле (3.1) и для оценивания различных характеристик полей по реальным данным использованы метеоданные, собранные на метеостанциях в период с 1988 по 2017 гг.

Поля среднесуточного ИХС в Области 1 и 2 сильно коррелированы по пространству: наименьший из коэффициентов корреляции, оценённых по реальным синхронным данным, равен 0.55. На рисунке 3.25 изображены значения коэффициентов корреляции между ИХС на всех станциях в Области 1 и 2 как функции расстояния между станциями. Временные корреляционные связи между

значениями ИХС тоже достаточно сильны. На рисунке 3.26 изображены оценки коэффициентов корреляции $corr(W_i^j, W_l^k)$ между значениями среднесуточного ИХС на станции Тисуль 2 января и значениями среднесуточного ИХС на станциях Тисуль, Кемерово, Огурцово, Татарск в период со 2 по 18 января.

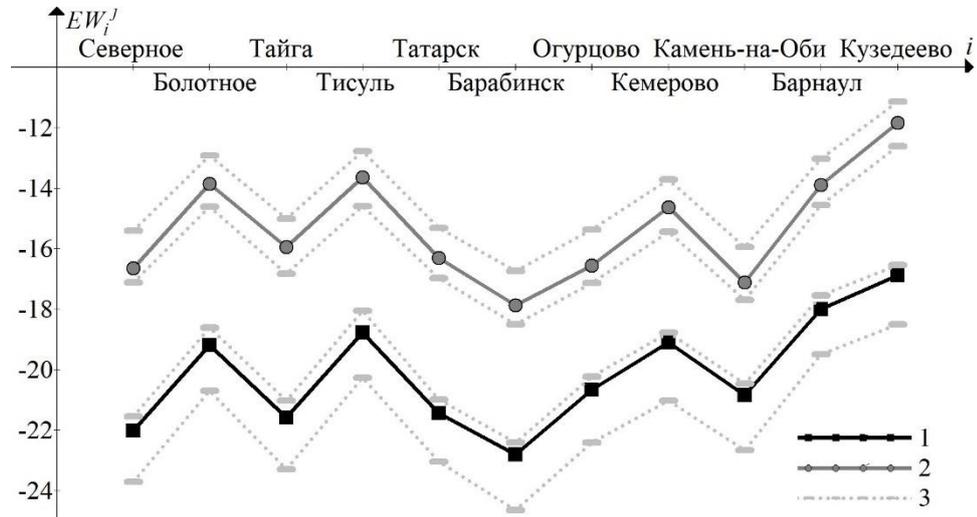


Рис. 3.23. Оценки математического ожидания EW_i^j среднесуточного ИХС. Кривая 1 – 15 января, кривая 2 – 1 марта, кривые 3 – границы соответствующих доверительных интервалов. Область 2.

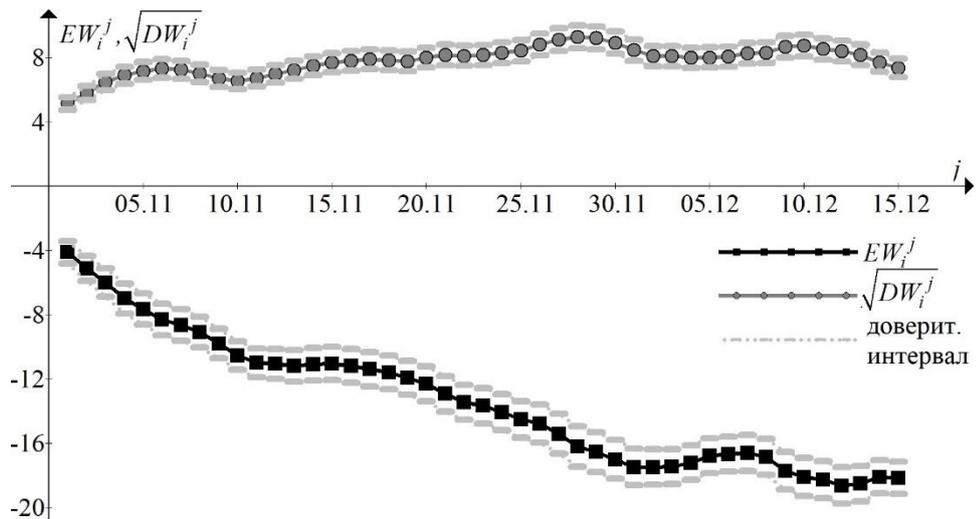


Рис. 3.24. Оценки математического ожидания EW_i^j и стандартного отклонения $\sqrt{DW_i^j}$ среднесуточного ИХС на метеостанции Болотное в период с 1 ноября по 15 декабря.

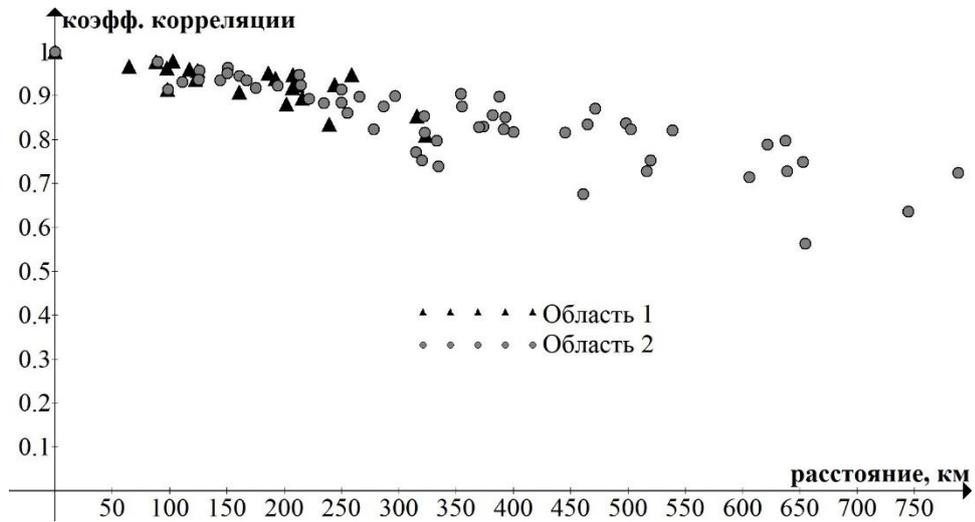


Рис. 3.25. Оценки коэффициентов корреляции среднесуточного ИХС как функция расстояния между метеостанциями. 16 января.

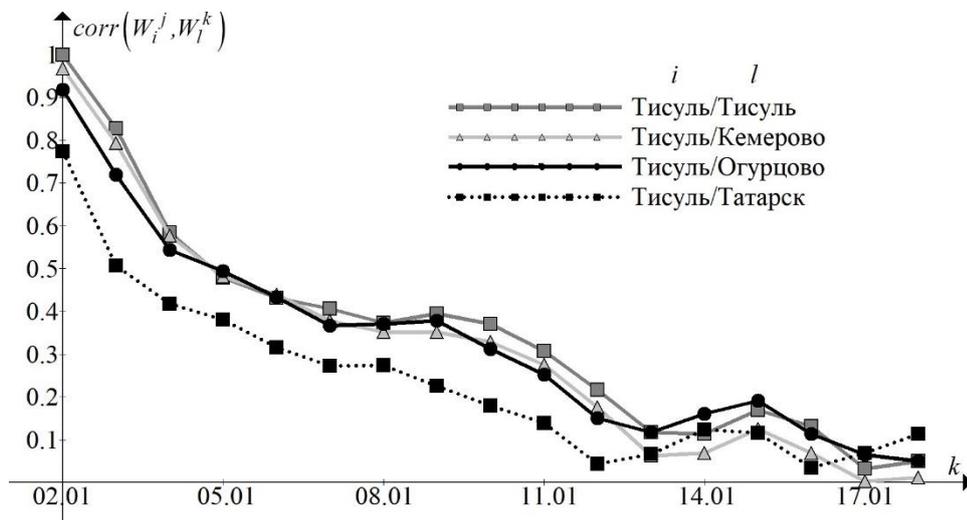


Рис. 3.26. Оценки коэффициентов корреляции между значениями среднесуточного ИХС на станции Тисуль 2 января и значениями среднесуточного ИХС на станциях Тисуль, Кемерово, Огурцово, Татарск в период со 2 по 18 января.

Отметим, что поскольку размерность пространственного поля \vec{W}_s^T среднесуточного ИХС меньше, чем объём выборки реальных данных, при оценивании выборочных коэффициентов корреляции, выборочных моментов распределения и при построении гистограммы необязательно использовать какое-либо окно осреднения, искусственно увеличивая объём выборки. Однако в данное

работе, было использовано скользящее окно осреднения шириной 5 суток. Выборочную корреляционную матрицу пространственно-временного поля \vec{W}_{st}^T без использования окна осреднения можно оценивать только при $ND = 1$. При $ND \geq 2$ размерность матрицы превысит объём выборки (равный 30, поскольку используются данные за 30 лет), и оцененная матрица не будет положительно определенной). При $ND \geq 2$ использовалось скользящее симметричное осреднение такой минимальной ширины $2L + 1$, чтобы выполнялось неравенство $30(2L + 1) > 18ND$. Максимальная ширина использованного окна осреднения бралась равной 11, что наложило ограничение на область применимости модели: для Области 1 максимальное значение ND равно 47, а для Области 2 – 30. При более широких окнах осреднения модель перестаёт воспроизводить временную нестационарность реального поля.

Реальные поля (и пространственное, и пространственно-временное) среднесуточного ИХС обладают выборочной корреляционной матрицей с гораздо меньшим числом обусловленности, чем поля СИЖ. Максимальное число обусловленности, которое наблюдалось в численных экспериментах, имело порядок 10^3 . При численной реализации модели поля ИХС регуляризация корреляционной матрицы либо не требовалась вовсе, либо могла быть проведена с параметром регуляризации $\varepsilon = 10^{-9}$.

Стохастическое моделирование пространственного поля \vec{W}_s^T (или пространственно-временного поля \vec{W}_{st}^T) осуществлялось методом обратных функций распределения в предположении, что одномерные распределения среднесуточного ИХС суть смеси $g_i(x), i = 1, \dots, NS$ (для случая пространственно-временного поля – смеси $g_i^j(x), i = 1, \dots, NS, j = 1, \dots, ND$) двух гауссовских распределений, а корреляционная структура поля задаётся выборочной корреляционной матрицей $R = \{r(i, j)\}_{i, j=1}^{NS}$ (для поля \vec{W}_s^T – $R = \{r(i, j)\}_{i, j=1}^{NS \times ND}$).

Верификация моделей пространственного и пространственно-временного полей среднесуточного ИХС

Результаты верификации моделей пространственного и пространственно-временного полей приведены в [145, 153]. В качестве иллюстрации того, что модельные и реальные траектории полей среднесуточного ИХС обладают схожими вероятностными свойствами, приведём 3 примера.

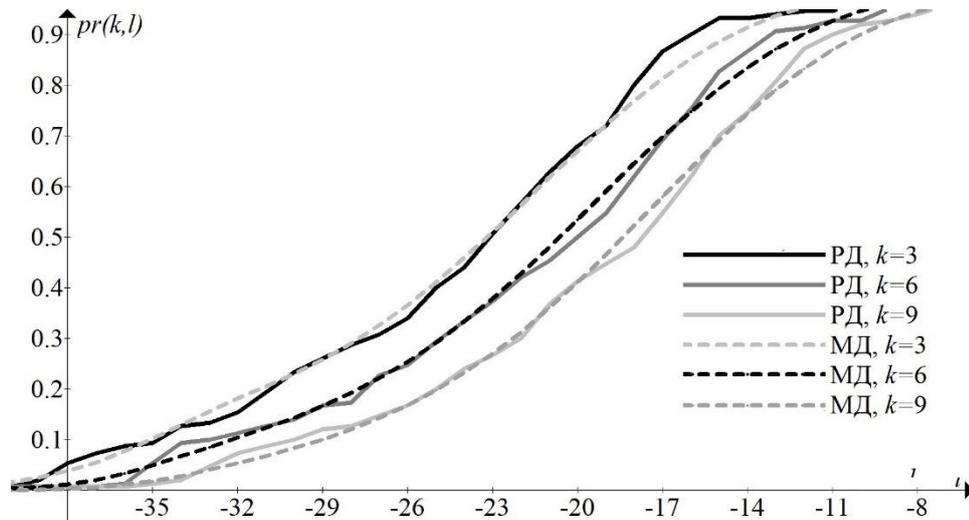
Для различных значений l °C по модельным и реальным данным оценивалась вероятность $p(l) = P(W_1 < l, W_2 < l, \dots, W_{NS} < l)$. Эта характеристика тесно связана с NS -мерным распределением поля \vec{W}_s^T . В таблице 3.21 приведены оценки $p(l)$ полученные по реальным данным и по 10^7 модельных траекторий. Для относительно высоких уровней l ($l \geq -15$ °C), примерно в 95% рассмотренных случаях оценки по модельным ансамблям попадают в доверительный интервал $(p(l) - \sigma, p(l) + \sigma)$ для оценки по реальным данным. Для более низких уровней l процент попадания в соответствующий доверительный интервал колеблется от 75% до 83% в зависимости от даты, для которой строится модель поля.

В качестве 2-го примера рассмотрим оценку вероятности $pr(k, l)$ того, что хотя бы на k метеостанциях одновременно среднесуточный ИХС ниже уровня l °C. На рисунке 3.27 приведены графики оценок $pr(k, l)$ для некоторых значений k , l и поля \vec{W}_s^T в Области 2. Видно, что при фиксированном k вероятность $pr(k, l)$ быстро убывает при убывании l . Для всех рассмотренных в численных экспериментах дат, оценки по модельным данным лежат в доверительных интервалах $(pr(k, l) - 3\sigma, pr(k, l) + 3\sigma)$ для оценок по реальным значениям ИХС.

Ещё одной из характеристик, которая применялась для верификации модели, была вероятность $P_{NS}^{ND}(lev)$ того, что среднесуточный ИХС ниже уровня lev °C на всех NS станциях в течение ND дней. В таблице 3.22

Таб. 3.21. Оценка вероятности $p(l)$. 5 декабря.

l °C	Область 1		Область 2	
	РД, $p(l) \pm \sigma$	МД	РД, $p(l) \pm \sigma$	МД
-3	0.980 ± 0.013	0.977	0.967 ± 0.017	0.958
-6	0.927 ± 0.023	0.913	0.873 ± 0.027	0.866
-9	0.707 ± 0.035	0.765	0.633 ± 0.039	0.684
-12	0.527 ± 0.039	0.566	0.440 ± 0.041	0.463
-15	0.367 ± 0.046	0.396	0.287 ± 0.057	0.288
-33	0.040 ± 0.436	0.027	0.000 ± 0.602	0.005
-36	0.013 ± 0.498	0.012	0.000 ± 0.699	0.001
-39	0.000 ± 0.572	0.004	0.000 ± 0.781	0.000
-42	0.000 ± 0.837	0.000	0.000 ± 0.813	0.000

Рис. 3.27. Оценка вероятности $pr(k,l)$. Область 2, 1 февраля.

приведены соответствующие оценки $P_{NS}^{ND}(lev)$ по реальным данным и по 10^9 модельных траекторий поля \vec{W}_{st}^T . В 96% проведенных численных экспериментов оценки по модельным данным лежат в доверительных интервалах $(P_{NS}^{ND}(lev) - 3\sigma, P_{NS}^{ND}(lev) + 3\sigma)$ оценок по реальным значениям ИХС.

Таб. 3.22. Оценки $P_{NS}^{ND}(lev)$. Область 2.

lev °C	5-10 января ($ND = 6$)		5-14 января ($ND = 10$)	
	РД, $P_{NS}^{ND}(lev) \pm \sigma$	МД	РД, $P_{NS}^{ND}(lev) \pm \sigma$	МД
0	0.871 ± 0.021	0.852	0.767 ± 0.024	0.789
-5	0.610 ± 0.036	0.589	0.482 ± 0.042	0.528
-10	0.276 ± 0.047	0.230	0.100 ± 0.053	0.084
-15	0.071 ± 0.085	0.059	0.018 ± 0.088	0.015
-20	0.019 ± 0.096	0.029	0.000 ± 0.101	0.007

В рассматриваемых Областях 1 и 2 сеть метеорологических станций очень редкая и не позволяет решать ряд задач биоклиматологии, связанных, например, с оценкой увеличения нагрузки на медицинские службы при наступлении холодов и с изучением распространения волн холода. Предложенные модели полей \overline{W}_s^T и \overline{W}_{st}^T не позволяют напрямую исследовать характеристики ИХС в тех точках пространства, где нет метеостанций. Однако, построенные модели можно расширить так, чтобы они позволили моделировать значения ИХС как на станциях, так и в узлах произвольной регулярной или нерегулярной сетки. Одним из наиболее простых способов такого расширения является пространственная интерполяция модельных траекторий со станций в узлы сетки. В пункте 3.4.2 параграфа 3.4 Главы 3 будет проведено сравнение нескольких методов интерполяции.

Модель пространственно-временного поля среднесуточного ИХС на основе стохастического «генератора погоды»

Предложенные в этом параграфе модели пространственно-временных полей биоклиматических индексов основаны на использовании определяющих формул. Однако, эти поля так же, как и временные ряды БИ, можно моделировать применяя стохастические «генераторы» погоды. В параграфе 2.2 Главы 2 была рассмотрена

стохастическая модель пространственно-временного поля комплекса метеоэлементов в южной части Байкальской природной территории. Продемонстрируем, что эта модель может быть применена для построения траекторий пространственно-временных полей индекса холодного стресса. Пусть, как и в параграфе 2.2 Главы 2, $\overline{\xi}_i^k = (T_{1,i}^k, \dots, T_{8,i}^k, u_{1,i}^k, v_{1,i}^k, \dots, u_{8,i}^k, v_{8,i}^k, P_{1,i}^k, P_{2,i}^k)^T$ – вектор, состоящий из значений температуры воздуха, u - и v -компонент скорости ветра с шагом по времени равным 3 часам, а также значений полусуточных сумм осадков в i -ые сутки на k -ой метеостанции. Очевидно, что можно вычислить значения модуля скорости ветра

$$wind_{j,i}^k = \sqrt{(u_{j,i}^k)^2 + (v_{j,i}^k)^2}$$

и, воспользовавшись формулой (3.1), значения индекса холодного стресса $W_{j,i}^k$ с трёхчасовым разрешением. Значения $W_{j,i}^k$, $j = 1, \dots, 8$, $i = 1, \dots, ND$, $k = 1, \dots, NS$ есть значения построенного на сети метеостанций пространственно-временного поля ИХС с трёхчасовым разрешением. Была проведена верификация такой модели поля. Она показала, что модель может быть использована для оценки вероятностей возникновения экстремальных значений ИХС. В качестве примера характеристики, использованной для верификации, в таблицах 3.23, 3.24 приведены оценки вероятности

$$p_{min} = P(\exists i \in \{1, 2, \dots, ND\} : \min W_i^k = \min \{W_{1,i}^k, W_{2,i}^k, W_{3,i}^k, W_{4,i}^k\} \in (a, b])$$

того, что на k -ой станции хотя бы единожды за ND дней минимальный за временной интервал с 8 утра до 17 дня местного времени ИХС $\min W_i^k$ принадлежит интервалу $(a, b]$. Приведённые в таблицах 3.23, 3.24 интервалы $(a, b]$ соответствуют 4 степеням риска обморожения кожи, установленным [18].

Так же как и в случае моделей на основе определяющих формул, построенные с помощью «генератора погоды» модели пространственно-временных полей БИ могут быть расширить так, чтобы они позволили моделировать значения БИ не только на станциях, но и в узлах произвольной сетки.

Таб. 3.23. Оценки вероятности p_{min} . 10-20 января.

$(a,b]$	Цакир		Хамар-Дабан	
	РД, $p_{min} \pm \sigma$	МД	РД, $p_{min} \pm \sigma$	МД
$(-24,-10]$	0.067 ± 0.044	0.060	0.720 ± 0.117	0.678
$(-34,-25]$	0.514 ± 0.100	0.475	0.189 ± 0.093	0.211
$(-59,-35]$	0.325 ± 0.106	0.372	0.014 ± 0.039	0.029
$(-\infty,-60]$	0.000 ± 0.001	0.000	0.000 ± 0.000	0.000

Таб. 3.24. Оценки вероятности p_{min} . 10-20 января.

$(a,b]$	Иркутск		Хомутово	
	РД, $p_{min} \pm \sigma$	МД	РД, $p_{min} \pm \sigma$	МД
$(-24,-10]$	0.450 ± 0.125	0.460	0.187 ± 0.079	0.169
$(-34,-25]$	0.361 ± 0.092	0.329	0.390 ± 0.090	0.402
$(-59,-35]$	0.089 ± 0.089	0.129	0.347 ± 0.111	0.348
$(-\infty,-60]$	0.000 ± 0.002	0.000	0.000 ± 0.007	0.001

3.4.2. Об интерполяции пространственного поля индекса холодного стресса со станций в узлы регулярной сетки

Методы пространственной интерполяции часто используются для определения значений различных метеорологических параметров в тех точках, где отсутствуют метеостанции (см., например, [99, 100, 111, 127, 169, 198]). Обзор литературы показывает, что качество интерполяции зависит не только от выбора метода интерполяции, но и от многих других факторов, включая степень пространственной неоднородности поля рассматриваемого метеопараметра, количество станций, где проводятся наблюдения, и их пространственное распределение, физико-географические особенности местности. Это означает, что

для каждой территории, метеопараметра и набора станций необходимо сравнить несколько методов интерполяции, чтобы выбрать тот, который дает наименьшую ошибку интерполяции. Аналогичная ситуация наблюдается и для биоклиматических индексов.

Сравним несколько способов пространственной интерполяции среднесуточного ИХС в Областях 1 и 2.

Методы интерполяции

Поскольку все рассматриваемые способы интерполяции хорошо известны и широко используются при решении задач климатологии и геостатистики [127, 158, 198, 205], приведём лишь их краткое описание. Всюду далее $W_i, i=1,2,\dots,NS$ – значение ИХС на i -ой станции с координатами (x_i, y_i) , W_{int} – интерполированное значение ИХС в точке с координатами (x_{int}, y_{int}) . Формально, необходимо рассматривать координаты (x_i, y_i, z_i) , однако, поскольку в рассматриваемых Областях 1 и 2 высоты станций над уровнем моря отличаются незначительно, третью координату мы учитывать не будем.

Первый из рассмотренных методов интерполяции – интерполяция из ближайшего соседа. В рамках этого подхода предполагается, что значение ИХС в заданной точке равно значению ИХС на ближайшей (в смысле декартова расстояния) метеостанции. Формально:

$$W_{int} = W_i,$$

$$i = \arg \min_{i=1,\dots,NS} \rho_i = \arg \min_{i=1,\dots,NS} \sqrt{(x_i - x_{int})^2 + (y_i - y_{int})^2}.$$

Отметим, что хотя метод интерполяции из ближайшего соседа является частным случаем рассмотренной далее интерполяции обратных взвешенных расстояний, его часто рассматривают как отдельный метод интерполяции. Далее мы будем обозначать интерполяцию с помощью этого метода как NN-интерполяцию.

Следующий метод – интерполяция обратных взвешенных расстояний (IDW-интерполяция). Значение W_{int} является линейной комбинацией значений ИХС W_i

на n ближайших к точке (x_{int}, y_{int}) станциях с весами, обратно пропорциональными расстоянию между точками (x_{int}, y_{int}) и (x_i, y_i) :

$$W = \sum_{i=1}^n \omega_i W_i, \quad \omega_i = \frac{1/\rho_i^\beta}{\sum_{j=1}^n (1/\rho_j^\beta)}, \quad \beta \geq 0, \quad 2 \leq n \leq NS.$$

Выбор параметров β , n осуществляется из условия минимума некоторого функционала (например, невязки или среднеквадратической ошибки интерполяции). В ряде работ отдельно выделяется IDW-интерполяция при $\beta = 2$. При таком значении β метод IDW-интерполяции называют методом обратных квадратов [19]. Будем обозначать этот способ как IDW-2.

Ещё один рассматриваемый в этой работе метод интерполяции – локальная полиномиальная интерполяция. Рассматривалась интерполяция ИХС полиномами двух переменных первой и второй степени (PN-1 и PN-2, соответственно). При использовании полиномов первой степени, значение W_{int} в точке (x_{int}, y_{int}) находилось из равенства

$$Ax_{int} + By_{int} + CW_{int} + D = 0.$$

Здесь $Ax + By + Cz + D = 0$ – уравнение плоскости, проходящей через точки (x_i, y_i, W_i) , $i = 1, 2, 3$, где (x_i, y_i) , $i = 1, 2, 3$ – координаты трёх ближайших к (x_{int}, y_{int}) метеостанций. В случае квадратичной интерполяции,

$$W_{int} = P_2(x_{int}, y_{int}),$$

где $P_2(x, y) = Ax^2 + By^2 + Cxy + Fx + Gy + H$ – многочлен второй степени, проходящий через точки (x_i, y_i, W_i) , $i = 1, 2, \dots, 6$, а (x_i, y_i) , $i = 1, 2, \dots, 6$ – координаты шести ближайших к (x_{int}, y_{int}) метеостанций.

Помимо вышеописанных детерминированных способов интерполяции рассмотрим недетерминированный метод: стохастическую интерполяцию (SI). Этот способ интерполяции детально описан в параграфе 1.2 Главы 1 для случая,

когда интерполируется случайный процесс как функция 1 переменной (времени). Интерполяция случайного поля проводится аналогичным образом [153, 177]:

$$W_{int} = W_i,$$

где i – реализация дискретной случайной величины X с распределением

$$P(X = i) = \frac{1/\rho_i^\beta}{\sum_{j=1}^{NS} (1/\rho_j^\beta)}, \quad \beta \geq 1, \quad i = 1, \dots, NS.$$

В отличие от ряда других недетерминированных интерполяционных методов (например, оптимальной интерполяции, ординарного или универсального кригинга [10, 19]), стохастическая интерполяция позволяет учитывать неоднородность поля среднесуточного ИХС и не требует знания корреляционных связей между значениями ИХС на станциях и в тех точках, куда проводится интерполяция.

Сравнение методов интерполяции

Выбор метода интерполяции будем осуществлять методом кросс-валидации [19, 132], в рамках которого для каждой из Областей 1 и 2 выбирается тот метод интерполяции, который обеспечивает минимум средней абсолютной ошибки интерполяции (MAE) или минимум среднеквадратической ошибки интерполяции (RMSE):

$$MAE(k) = \frac{1}{NS \cdot NY} \sum_{i=1}^{NS} \sum_{j=1}^{NY} |W_R(i, j, k) - W_I(i, j, k)|,$$

$$RMSE(k) = \sqrt{\frac{1}{NS \cdot NY} \sum_{i=1}^{NS} \sum_{j=1}^{NY} (W_R(i, j, k) - W_I(i, j, k))^2}.$$

Здесь NY – число лет, в которые проводились наблюдения на метеостанциях (в данном случае $NY = 30$), $W_R(i, j, k)$ и $W_I(i, j, k)$ – реальное и интерполированное значения среднесуточного ИХС на метеостанции i , $i = 1, \dots, NS$ в k -ый день $(1987 + j)$ -го года. Переменная k принимает значения от 1 до 90 включительно (что соответствует периоду с 1 января по 31 марта) и с 287 до 365 (15 октября – 31 декабря).

В таблицах 3.25, 3.26 приведены значения MAE и RMSE для некоторых значений k для Областей 1 и 2. Численные эксперименты показали, что для большинства значений k оптимальные значения n при IDW-интерполяции и стохастической интерполяции совпадают с NS , поэтому в таблицах 3.25, 3.26 приведены значения MAE и RMSE для этих методов для $n = NS$. Для тех k , для которых $n \neq NS$ в скобках указано оптимальное значение n . В скобках также указаны значения параметра β , минимизирующие $MAE(k)$.

Таб. 3.25. Результаты кросс-валидации. Область 1.

Метод интерполяции	10 декабря		10 января	
	MAE	RMSE	MAE	RMSE
NN	2.77	3.39	3.18	3.88
IDW-2	2.14	2.79	2.36	3.09
IDW	2.13 ($\beta = 1.7508$)	2.79	2.33 ($\beta = 1.2355$)	3.07
PN-1	4.29	5.35	4.52	5.76
PN-2	3.72	4.19	3.81	4.55
SI	1.90 ($\beta = 4.4031$)	2.72	2.10 ($\beta = 5.6859$)	2.98

Кросс-валидация была осуществлена для обеих рассматриваемых областей для каждого дня k с 1 января по 31 марта и с 15 октября по 31 декабря, т.е. всего 346 раз. Для всех дней наихудшие результаты (в смысле величины MAE и RMSE) были получены при интерполяции полиномами первой и второй степени (PN-1 и PN-2). Значения MAE и RMSE при NN-интерполяции выше, чем соответствующие значения при IDW-2, IDW и стохастической интерполяции в 84% экспериментов. Во всех проведенных экспериментах MAE и RMSE IDW-2 больше, чем соответствующие ошибки IDW-интерполяции. Наименьшие MAE и RMSE

получены при использовании IDW- интерполяции и стохастической интерполяции в 48% и 52% экспериментов соответственно.

Таб. 3.26. Результаты кросс-валидации. Область 2.

Метод интерполяции	10 декабря		10 января		17 октября	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
NN	2.38	3.18	2.67	3.44	1.52	1.92
IDW-2	2.05	2.81	2.32	3.06	1.28	1.70
IDW	2.01 ($\beta = 3.2640$)	2.72	2.27 ($\beta = 3.1215$)	2.96	1.18 ($\beta = 1.1690, n = 3$)	1.52
PN-1	5.83	14.18	5.54	15.21	1.93	2.83
PN-2	4.73	12.10	4.47	13.14	2.12	5.13
SI	2.02 ($\beta = 5.2188$)	2.81	2.22 ($\beta = 5.9361$)	3.03	1.26 ($\beta = 5.9860, n = 3$)	1.83

Ещё одним критерием, который был использован для выбора оптимального метода интерполяции среднесуточного ИХС в Областях 1 и 2, являлась относительная ошибка $re(i, l, k)$ оценки коэффициентов корреляции при использовании реальных и интерполированных данных:

$$re(i, l, k) = \frac{|c_{RR}(i, l, k) - c_{IR}(i, l, k)|}{|c_{RR}(i, l, k)|} \cdot 100\%,$$

$$i, l = 1, \dots, NS, k = 1, \dots, 90 \text{ или } k = 287, \dots, 365,$$

где $c_{RR}(i, l, k)$ – коэффициент корреляции ИХС на станциях i и l в k -ый день, оценённый по реальным данным

$$W_R(i, j, k), W_R(l, j, k), i, l = 1, \dots, NS, \\ j = 1, \dots, NY, k = 1, \dots, 90 \text{ или } k = 287, \dots, 365,$$

$c_{IR}(i, l, k)$ – коэффициент корреляции ИХС на станциях i и l в k -ый день, оценённый по интерполированным и реальным данным

$$W_I(i, j, k), W_R(l, j, k),$$

$$i, l = 1, \dots, NS, \quad j = 1, \dots, NY, \quad k = 1, \dots, 90 \text{ или } k = 287, \dots, 365,$$

соответственно. Этот критерий применялся для сравнения результатов стохастической и IDW-интерполяции при $n = NS$ и значениях β , полученных минимизацией RMSE. В качестве примера ниже приведены матрицы, составленные из $re(i, l, 41)$, $i, l = 1, \dots, NS$, вычисленных при использовании стохастической и IDW-интерполяции в Области 1:

$$re_{SI} = \begin{pmatrix} 0.0 & 2.3 & 4.5 & 0.9 & 0.9 & 10.2 & 9.4 \\ 5.8 & 0.0 & 1.6 & 2.2 & 2.5 & 9.3 & 2.3 \\ 4.5 & 1.1 & 0.0 & 2.3 & 2.3 & 7.7 & 2.4 \\ 5.7 & 1.2 & 1.1 & 0.0 & 4.0 & 5.8 & 0.4 \\ 6.5 & 0.8 & 2.4 & 5.5 & 0.0 & 6.4 & 0.6 \\ 7.1 & 2.0 & 2.1 & 3.6 & 1.7 & 0.0 & 0.7 \\ 9.0 & 3.8 & 5.1 & 0.8 & 1.9 & 8.1 & 0.0 \end{pmatrix},$$

$$re_{IDW} = \begin{pmatrix} 0.0 & 3.3 & 8.1 & 0.6 & 4.3 & 8.4 & 13.8 \\ 12.4 & 0.0 & 5.0 & 1.6 & 3.9 & 12.5 & 6.6 \\ 11.8 & 0.1 & 0.0 & 3.1 & 3.6 & 9.0 & 5.0 \\ 13.0 & 2.0 & 5.5 & 0.0 & 2.7 & 7.4 & 5.8 \\ 13.8 & 0.8 & 5.4 & 4.2 & 0.0 & 7.9 & 4.5 \\ 16.1 & 5.6 & 7.6 & 3.0 & 4.5 & 0.0 & 3.6 \\ 25.2 & 6.0 & 9.8 & 1.1 & 7.0 & 9.9 & 0.0 \end{pmatrix}.$$

В среднем, относительная ошибка $re(i, l, k)$, полученная при использовании стохастической интерполяции, меньше ошибки, возникающей при применении IDW-интерполяции. Это означает, что, согласно рассматриваемому критерию, стохастическая интерполяция предпочтительнее IDW-интерполяции.

Модель поля среднесуточного ИХС на сетке

Рассмотренные выше методы интерполяции позволяют преобразовать предложенную в предыдущем параграфе модель пространственного (или пространственно-временного) поля среднесуточного ИХС на метеостанциях в

модель поля в узлах произвольной стеки. Реализация случайного поля на сетке может быть получена по Алгоритму 3.7.

Алгоритм 3.7.

- 1) строится реализация поля \vec{W}_s^T (или \vec{W}_{st}^T) среднесуточного ИХС на станциях;
- 2) полученные значения ИХС на станциях стохастически интерполируются в узлы (x_j, y_k) , $j = 1, \dots, NX$, $k = 1, \dots, NY$ сетки с NX , NY узлами по осям Ox , Oy , соответственно.

Приведём 2 примера характеристик, исследование которых невозможно при использовании модели поля среднесуточного ИХС на метеостанциях, но может быть проведено при использовании модели поля на сетке.

В качестве первого примера, в таблице 3.27 приведены оценки вероятностей $P_{Bolot\ reg}(lev)$ того, что ИХС будет ниже заданного уровня $lev^{\circ}C$ на территории Болотнинского района. На территорию этого района попал 31 узел сетки с шагом 10 км в обоих направлениях. Поскольку территория Болотнинского района лежит и в Области 1, и в Области 2, в таблице 3.27 приведены оценки $P_{Bolot\ reg}(lev)$, полученные по модельным реализациям пространственного поля ИХС в каждой из областей. Для сравнения приведены вероятности $P_{Bolot}(lev)$ того, что ИХС ниже заданного уровня $lev^{\circ}C$ на метеостанции Болотное.

Второй пример применение модели поля на сетке – это оценка среднего числа $AN_{Nsk}(lev)$ узлов сетки, лежащих на территории Новосибирска, в которых в заданный день ИХС ниже $lev^{\circ}C$. Оценки $AN_{Nsk}(lev)$, приведенные в таблице 3.28, получены по 10^9 реализациям пространственно-временного поля.

Необходимо отметить, что все рассмотренные в этом параграфе методы интерполяции представляют собой методы пространственной интерполяции. При

построении модели пространственно-временного поля на сетке временные связи при интерполяции учитываются не напрямую, а косвенно – через временные связи между значениями ИХС на станциях.

Таб. 3.27. Оценки вероятностей $P_{Bolot\ reg}(lev)$. 20 декабря.

$lev\ ^\circ C$	$P_{Bolot\ reg}(lev) \pm \sigma$ (Область 1)	$P_{Bolot\ reg}(lev) \pm \sigma$ (Область 2)	$P_{Bolot}(lev)$
-5	0.955 ± 0.009	0.953 ± 0.010	0.964
-10	0.849 ± 0.024	0.845 ± 0.025	0.869
-15	0.587 ± 0.035	0.569 ± 0.039	0.676
-20	0.273 ± 0.056	0.269 ± 0.056	0.424
-25	0.115 ± 0.071	0.101 ± 0.078	0.214
-30	0.053 ± 0.089	0.050 ± 0.089	0.102
-35	0.008 ± 0.097	0.011 ± 0.101	0.060

Таб. 3.28. Оценки $AN_{Nsk}(lev)$. 7-23 декабря.

День	$lev\ ^\circ C$						
	0	-6	-12	-18	-24	-30	-36
07.12	10.980	10.493	7.360	4.320	1.993	0.973	0.337
09.19	10.980	10.507	7.687	4.973	2.720	1.113	0.387
11.12	10.983	10.600	8.013	5.320	3.080	1.207	0.407
13.12	10.985	10.648	8.333	5.773	3.100	1.233	0.426
15.12	10.987	10.693	8.900	5.807	2.487	1.273	0.437
17.12	10.990	10.787	9.200	5.893	2.588	1.331	0.487
19.12	10.990	10.813	9.313	5.920	2.747	1.413	0.488
21.12	10.992	10.887	9.587	6.233	3.353	1.507	0.512
23.12	10.993	10.983	9.887	6.613	3.833	1.693	0.696

Глава 4. Исследование некоторых классов случайных процессов с периодическими характеристиками

4.1. Модели периодически коррелированных кусочно-постоянных и кусочно-линейных процессов

4.1.1. Модель периодически коррелированного кусочно-постоянного процесса на основе стационарного случайного процесса с дискретным временем

Определение процесса

Пусть $\eta_i, i = 0, 1, 2, \dots$ – стационарный в широком смысле случайный процесс с функцией одномерного распределения $F_\eta(x)$, нулевым средним, единичной дисперсией и нормированной корреляционной функцией $r(|i - j|) = r_m, m = 0, 1, 2, \dots, r_0 = 1$. В [56] был предложен Алгоритм 4.1 построения кусочно-постоянного случайного процесса ξ_t непрерывного аргумента $t \geq 0$, основанного на кусочно-постоянном восполнении процесса $\eta_i, i = 0, 1, 2, \dots$ специального вида.

Алгоритм 4.1.

1. Независимо от $\eta_i, i = 0, 1, 2, \dots$ моделируется случайная величина α , распределенная на отрезке $[0, 1]$ с функцией одномерного распределения $F(x), x \in [0, 1]$, и для каждого $i = 0, 1, 2, \dots$ на оси t выбирается точка $i + \alpha$.
2. Значения процесса ξ_t определяются по правилу

$$\xi_t = \begin{cases} \eta_0, & t \in [0, \alpha] \\ \eta_i, & t \in (i-1 + \alpha, i + \alpha], \quad i = 1, 2, \dots \end{cases}$$

Отметим, что по построению значения процесса ξ_t в точках $t = i = 0, 1, 2, \dots$ совпадают со значениями последовательности η_i . На рисунке 4.1 приведён пример реализации кусочно-постоянного процесса ξ_t и опорного процесса $\eta_i, i = 0, 1, 2, \dots$

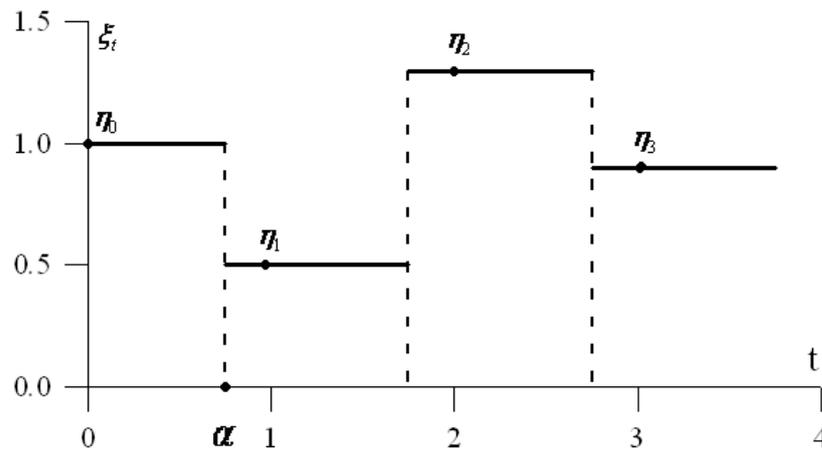


Рис. 4.1. Пример реализации кусочно-постоянного процесса ξ_t и опорного процесса $\eta_i, i = 0, 1, 2, \dots$

Свойства процесса

Свойства процесса ξ_t в случае равномерного распределения случайной величины α на отрезке $[0, 1]$, подробно изучены в [56, 176]. В этих работах приведено доказательство следующего утверждения.

Утверждение 4.1. Процесс ξ_t является стационарным в широком смысле.

В [176] рассмотрен также случай, когда распределение α на отрезке $[0, 1]$ отлично от равномерного. В этой статье показано, что справедливы утверждения 4.2 и 4.3.

Утверждение 4.2. Нормированная корреляционная функция процесса ξ_t имеет вид

$$\text{corr}(\xi_t, \xi_{t+\tau}) = \begin{cases} (1 - F(t_0 + h) + F(t_0))r_m + (F(t_0 + h) - F(t_0))r_{m+1}, & t_0 + h \leq 1, \\ (F(t_0) - F(t_0 + h))r_m + (1 + F(t_0 + h) - F(t_0))r_{m+1}, & t_0 + h > 1, \end{cases}$$

где $t_0 = \{t\}$, $h = \{\tau\}$, $m = [\tau]$ (здесь и далее $[x]$ и $\{x\}$ – целая и дробная части числа x , $x \in \mathbb{R}$, соответственно), и при целых неотрицательных значениях τ значения $\text{corr}(\xi_t, \xi_{t+\tau})$ совпадают с r_τ .

Утверждение 4.3. Построенный согласно Алгоритму 4.1 процесс ξ_t является периодически коррелированным с периодом $T = 1$.

Рассмотрим следующий пример.

Пример 4.1. Пусть случайная величина α распределена согласно усеченному экспоненциальному распределению, т.е.

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1 - e^{-\lambda x}}{1 - e^{-\lambda}}, & 0 < x \leq 1, \\ 1, & 1 < x \end{cases}$$

где $\lambda > 0$, а значения корреляционной функции стационарного процесса η_i определены как $r_i = e^{-\gamma i}$, $i = 0, 1, 2, \dots$. На рисунке 4.2 приведены графики функции $\text{corr}(\xi_t, \xi_{t+\tau})$ как функции переменной t при фиксированных значениях τ и $\lambda = 5$, $\gamma = 0.9$, $F_\eta(x) = \Phi_{0,1}$. На рисунке 4.3 изображены графики корреляционной функции $\text{corr}(\xi_t, \xi_{t+\tau})$ как функции переменной τ процесса ξ_t при фиксированных значениях t и значения корреляционной функции стационарного процесса η_i . Видно, что корреляционная функция $\text{corr}(\xi_t, \xi_{t+\tau})$ не имеет частной производной

$\frac{\partial \text{corr}(\xi_t, \xi_{t+\tau})}{\partial t}$ в точке $t = 1 - \{\tau\}$. Аналогичная ситуация возникает и при ряде

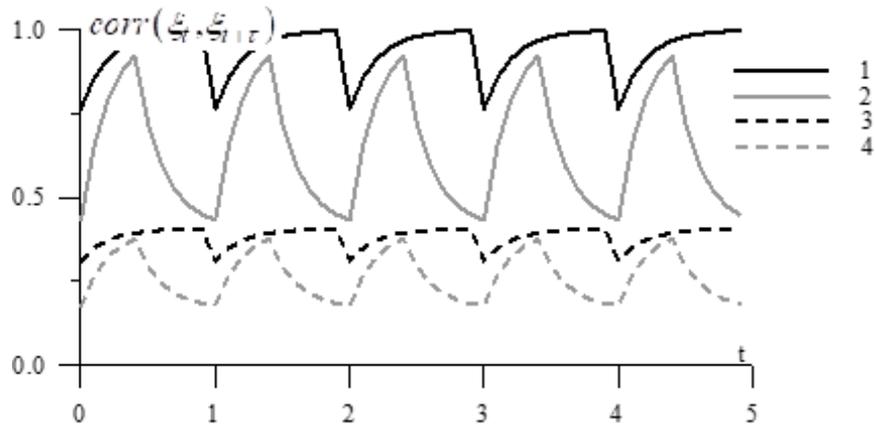


Рис. 4.2. Корреляционная функция $\text{corr}(\xi_t, \xi_{t+\tau})$ процесса ξ_t при фиксированных значениях τ . Кривая 1: $\tau = 0.1$, кривая 2: $\tau = 0.6$, кривая 3: $\tau = 1.1$, кривая 4: $\tau = 1.6$.

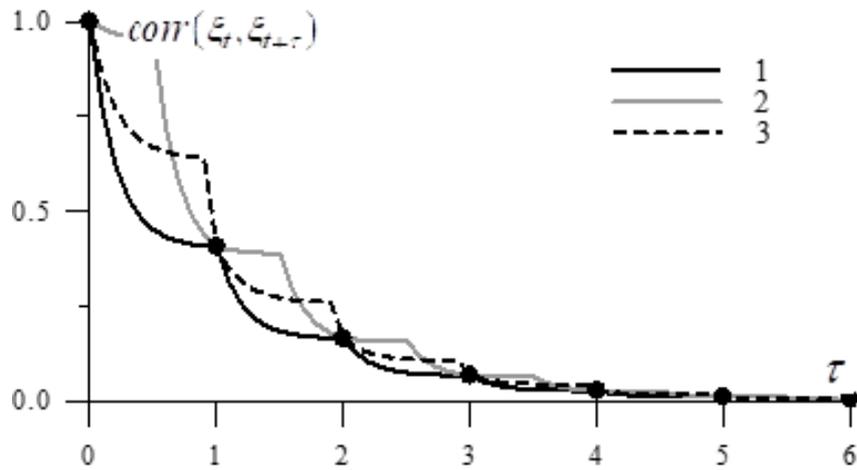


Рис. 4.3. Корреляционная функции $\text{corr}(\xi_t, \xi_{t+\tau})$ при фиксированных значениях t . Кривая 1: $t = 0$, кривая 2: $t = 0.5$, кривая 3: $t = 1.1$. Точками (\bullet) обозначены значения r_τ .

других функций $F_\eta(x)$, $F(x)$. Исследуем, при каких условиях эта производная существует [31, 134, 137].

Перепишем выражение для $\text{corr}(\xi_t, \xi_{t+\tau})$ в следующем виде:

$$\text{corr}(\xi_t, \xi_{t+\tau}) = \begin{cases} r_m + (r_m - r_{m+1})(F(\{t\}) - F(\{t\} + \{\tau\})), & \{t\} + \{\tau\} \leq 1, \\ r_{m+1} + (r_m - r_{m+1})(F(\{t\}) - F(\{t\} + \{\tau\} - 1)), & \{t\} + \{\tau\} > 1. \end{cases}$$

Исследуем $\text{corr}(\xi_t, \xi_{t+\tau})$ как функцию переменной t . Поскольку функция $\text{corr}(\xi_t, \xi_{t+\tau})$ является периодической по t , достаточно провести исследование только для $t \in [0, 1)$. Если $t \in [0, 1)$, то $\{t\} = t$, и выражение для корреляционной функции принимает вид

$$\text{corr}(\xi_t, \xi_{t+\tau}) = \begin{cases} r_m + (r_m - r_{m+1})(F(t) - F(t + \{\tau\})), & t + \{\tau\} \leq 1, \\ r_{m+1} + (r_m - r_{m+1})(F(t) - F(t + \{\tau\} - 1)), & t + \{\tau\} > 1. \end{cases}$$

Рассматриваемая частная производная $\frac{\partial \text{corr}(\xi_t, \xi_{t+\tau})}{\partial t}$ в точке $t = 1 - \{\tau\}$

существует тогда и только тогда, когда левая и правая частные производные в этой точке равны, т.е. должно выполняться равенство

$$\frac{\partial F(t + \{\tau\})}{\partial t} = \frac{\partial F(t + \{\tau\} - 1)}{\partial t}.$$

Если $t = 1 - \{\tau\}$, то $t + \{\tau\} = 1$, $t + \{\tau\} - 1 = 0$, и условие равенства левой и правой частных производных принимает вид $F'(0) = F'(1)$. Таким образом, доказано утверждение 4.4.

Утверждение 4.4. В условиях данного параграфа, если $F'(0) \neq F'(1)$ или не существует хотя бы одной из производных $F'(0)$, $F'(1)$, то функция $\text{corr}(\xi_t, \xi_{t+\tau})$ не будет иметь частной производной по t в точке $t = 1 - \{\tau\}$.

С учётом того, что на интервале $[n, n+1)$ для любого целого неотрицательного n выполняется равенство $\{t\} = t - n$, аналогичное

исследование можно провести для $corr(\xi_t, \xi_{t+\tau})$ как функции τ . Несложно показать, что справедливо утверждение 4.5.

Утверждение 4.5. В условиях данного параграфа частная производная $\frac{\partial corr(\xi_t, \xi_{t+\tau})}{\partial \tau}$ в точке $\tau = 1 - \{t\}$ существует тогда и только тогда, когда $F'(0) = F'(1)$.

В рассмотренном выше примере функция $F(x)$ не является дифференцируемой в точках $x = 0$ и $x = 1$, потому $corr(\xi_t, \xi_{t+\tau})$ не имеет частных производных по t и τ при $\{t\} + \{\tau\} = 1$.

Пример 4.2. Пусть, как и в примере 4.1, $r_i = e^{-0.9i}$, $i = 0, 1, 2, \dots$, а функция распределения имеет вид

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{2} \sin\left(\pi x - \frac{\pi}{2}\right) + \frac{1}{2}, & 0 < x \leq 1 \\ 1, & 1 < x. \end{cases}$$

На рисунках 4.4, 4.5 изображены графики корреляционной функции $corr(\xi_t, \xi_{t+\tau})$ процесса ξ_t , как функции t и τ при фиксированном значении второй переменной.

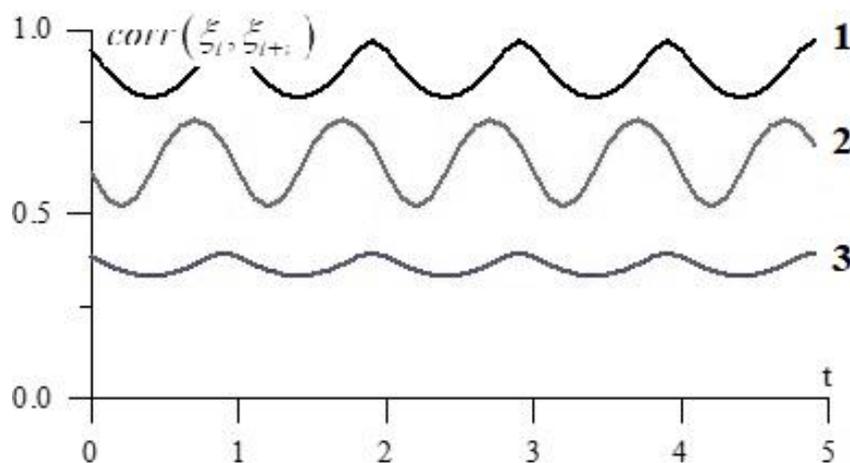


Рис. 4.4. Корреляционная функция $corr(\xi_t, \xi_{t+\tau})$ процесса ξ_t при фиксированных значениях τ . Кривая 1: $\tau = 0.2$, кривая 2: $\tau = 0.6$, кривая 3: $\tau = 1.2$.

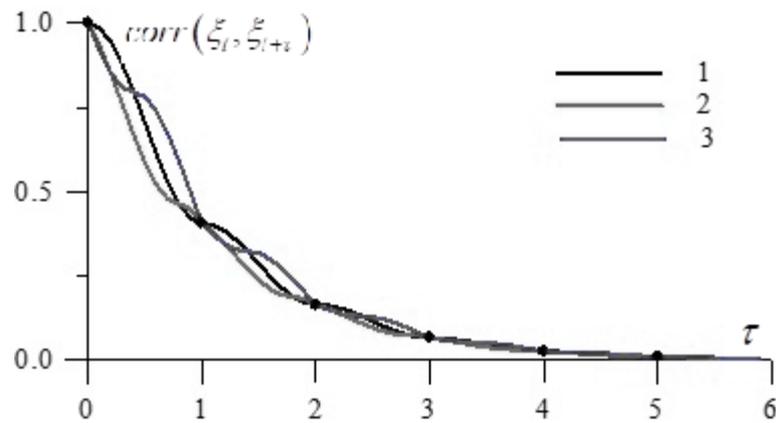


Рис. 4.5. Корреляционная функции $corr(\xi_t, \xi_{t+\tau})$ при фиксированных значениях t .

Кривая 1: $t = 0$, кривая 2: $t = 0.2$, кривая 3: $t = 1.6$. Точками (\bullet) обозначены значения r_τ .

4.1.2. Модель периодически коррелированного кусочно-постоянного процесса на основе стационарного случайного процесса и регулярного потока точек

Определение процесса

Алгоритм 4.2.

Пусть $\eta_i, i = -1, 0, 1, 2, \dots$ – стационарный в широком смысле случайный процесс с функцией одномерного распределения $F(x), x \in (a, b)$, конечными математическим ожиданием μ и дисперсией σ^2 , корреляционной функцией

$$corr(\eta_i, \eta_j) = corr(i, j) = corr(|i - j|) = corr(h) = r_h.$$

Пусть α – независимая от $\eta_i, i = -1, 0, 1, 2, \dots$ случайная величина, распределенная равномерно на отрезке $[0, 1]$.

Рассмотрим случайный кусочно-постоянный процесс ξ_t непрерывного времени $t \geq 0$

$$\xi_t = \begin{cases} \eta_{i-1}, & t \in [i, i + \alpha) \\ \eta_i, & t \in [i + \alpha, i + 1) \end{cases}$$

где $i = [t]$. На рисунке 4.6 представлен пример реализации процесса ξ_t .

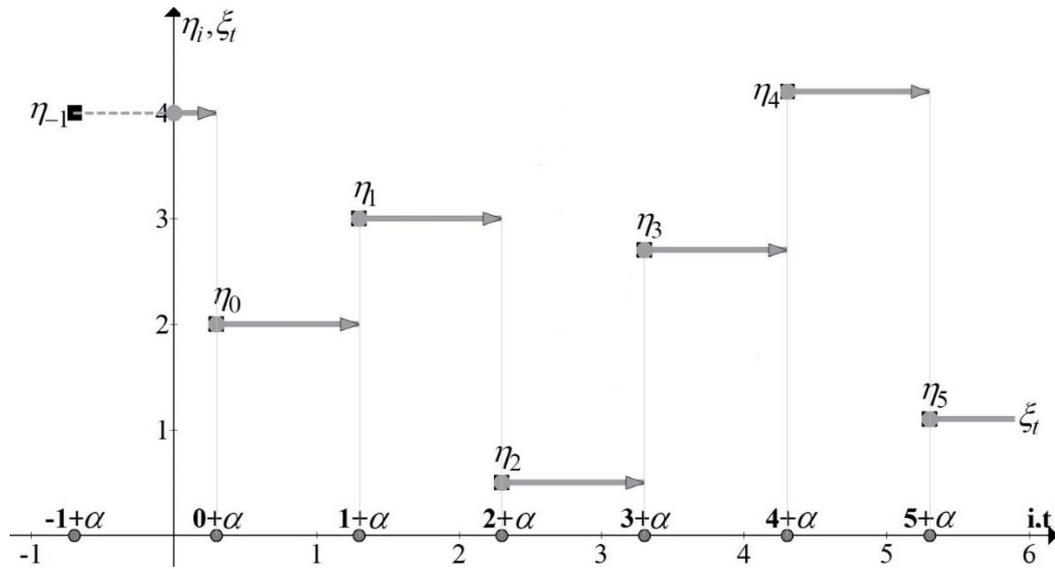


Рис. 4.6 Траектория процесса ξ_t .

Очевидно, что значение процесса ξ_t в точке t определяется равенством

$$\xi_t = \tilde{p}\eta_{i-1} + (1 - \tilde{p})\eta_i,$$

где

$$\tilde{p} = P(t - i < \alpha) = 1 - P(\alpha \leq t - i) = 1 - P(\alpha \leq \{t\}) = 1 - \{t\}.$$

Свойства процесса

Лемма 4.1. Математическое ожидание процесса ξ_t не зависит от t и

$$E\xi_t = \mu.$$

Дисперсия процесса ξ_t определяется равенством

$$D\xi_t = \sigma^2 \left(1 - 2\{t\} + 2\{t\}^2 + 2r_1\{t\}(1 - \{t\}) \right)$$

и является периодической функцией по t , $t \geq 0$ с периодом равным 1.

Доказательство. Используя свойства математического ожидания и то, что $E\eta_i = \mu$, $i = -1, 0, 1, 2, \dots$, вычислим математическое ожидание процесса ξ_t :

$$E\xi_t = E\left((1 - \{t\})\eta_{i-1} + \{t\}\eta_i\right) = (1 - \{t\})E\eta_{i-1} + \{t\}E\eta_i = (1 - \{t\})\mu + \{t\}\mu = \mu.$$

Используя равенства

$$E\eta_i = \mu, \sigma^2 = E\eta_i^2 - \mu^2, r_1 = \frac{E\eta_{i-1}\eta_i - \mu^2}{\sigma^2},$$

найдем $E\xi_t^2$:

$$\begin{aligned} E\xi_t^2 &= E\left((1-\{t\})\eta_{i-1} + \{t\}\eta_i\right)^2 = \\ &= (1-\{t\})^2 E\eta_{i-1}^2 + 2(1-\{t\})\{t\} E\eta_{i-1}\eta_i + \{t\}^2 E\eta_i^2 = \\ &= \left((1-\{t\})^2 + \{t\}^2\right)(\sigma^2 + \mu^2) + 2(1-\{t\})\{t\}(\sigma^2 r_1 + \mu^2) = \\ &= \mu^2 + \sigma^2(1 - 2\{t\} + 2\{t\}^2) + \sigma^2 r_1 2\{t\}(1-\{t\}). \end{aligned}$$

Следовательно,

$$D\xi_t = E\xi_t^2 - \mu^2 = \sigma^2(1 - 2\{t\} + 2\{t\}^2 + r_1 2\{t\}(1-\{t\})).$$

Периодичность $D\xi_t$ следует из периодичности функции $y(x) = \{x\}$.

Лемма доказана.

Лемма 4.2. Корреляционная функция $\text{corr}(\xi_t, \xi_p)$ процесса ξ_t для $t \in [i, i+1)$, $p \in [j, j+1)$, $i, j = 0, 1, 2, \dots$ имеет вид

$$\begin{aligned} \text{corr}(\xi_t, \xi_p) &= \\ &= \frac{1}{Z_t Z_p} \left(\begin{aligned} &(\{p\} - \{t\}\{p\})\eta_{|i-j-1|} + (1-\{t\} - \{p\} + 2\{t\}\{p\})\eta_{|i-j|} + \\ &+ (\{t\} - \{t\}\{p\})\eta_{|i-j+1|} \end{aligned} \right), \end{aligned}$$

где

$$Z_t = \sqrt{1 - 2\{t\} + 2\{t\}^2 + r_1 2\{t\}(1-\{t\})}, \quad Z_p = \sqrt{1 - 2\{p\} + 2\{p\}^2 + r_1 2\{p\}(1-\{p\})}.$$

Доказательство. Вычислим $E\xi_t \xi_p$:

$$\begin{aligned} E\xi_t \xi_p &= E\left((1-\{t\})\eta_{i-1} + \{t\}\eta_i\right)\left((1-\{p\})\eta_{j-1} + \{p\}\eta_j\right) = \\ &= (1-\{t\})\{p\} E\eta_{i-1}\eta_j + (1-\{t\})(1-\{p\}) E\eta_{i-1}\eta_{j-1} + \\ &+ \{t\}\{p\} E\eta_i\eta_j + \{t\}(1-\{p\}) E\eta_i\eta_{j-1} = \\ &= (\{p\} - \{t\}\{p\})\left(\sigma^2 \eta_{|i-j-1|} + \mu^2\right) + \end{aligned}$$

$$\begin{aligned}
& + \left((1 - \{t\})(1 - \{p\}) + \{t\}\{p\} \right) \left(\sigma^2 r_{|i-j|} + \mu^2 \right) + \\
& + \left(\{t\} - \{t\}\{p\} \right) \left(\sigma^2 r_{|i-j+1|} + \mu^2 \right) = \\
& = \mu^2 + \sigma^2 \left(\begin{aligned} & \left((\{p\} - \{t\}\{p\}) r_{|i-j-1|} + (1 - \{t\} - \{p\} + 2\{t\}\{p\}) r_{|i-j|} + \right. \\ & \left. + (\{t\} - \{t\}\{p\}) r_{|i-j+1|} \right) \end{aligned} \right).
\end{aligned}$$

Используя вычисленные выше значения $E\xi_t$, $D\xi_t$ и $E\xi_t\xi_p$, получаем

$$\begin{aligned}
\text{corr}(\xi_t, \xi_p) &= \frac{E\xi_t\xi_p - E\xi_tE\xi_p}{\sqrt{D\xi_t}\sqrt{D\xi_p}} = \\
&= \frac{\mu^2 + \sigma^2 \left(\begin{aligned} & \left((\{p\} - \{t\}\{p\}) r_{|i-j-1|} + (1 - \{t\} - \{p\} + 2\{t\}\{p\}) r_{|i-j|} + \right. \\ & \left. + (\{t\} - \{t\}\{p\}) r_{|i-j+1|} \right) \end{aligned} \right) - \mu^2}{\sqrt{\sigma^2 \left(1 - 2\{t\} + 2\{t\}^2 + r_1 2\{t\}(1 - \{t\}) \right)} \sqrt{\sigma^2 \left(1 - 2\{p\} + 2\{p\}^2 + r_1 2\{p\}(1 - \{p\}) \right)}} = \\
&= \frac{\left((\{p\} - \{t\}\{p\}) r_{|i-j-1|} + (1 - \{t\} - \{p\} + 2\{t\}\{p\}) r_{|i-j|} + (\{t\} - \{t\}\{p\}) r_{|i-j+1|} \right)}{\sqrt{\left(1 - 2\{t\} + 2\{t\}^2 + r_1 2\{t\}(1 - \{t\}) \right)} \sqrt{\left(1 - 2\{p\} + 2\{p\}^2 + r_1 2\{p\}(1 - \{p\}) \right)}} = \\
&= \frac{1}{Z_t Z_p} \left(\begin{aligned} & \left((\{p\} - \{t\}\{p\}) r_{|i-j-1|} + (1 - \{t\} - \{p\} + 2\{t\}\{p\}) r_{|i-j|} + \right. \\ & \left. + (\{t\} - \{t\}\{p\}) r_{|i-j+1|} \right) \end{aligned} \right).
\end{aligned}$$

Лемма доказана.

Утверждение 4.6. Случайный процесс ξ_t является периодически коррелированным с периодом равным 1.

Доказательство. Доказательство этого утверждения следует из лемм 4.1, 4.2 и того, что для корреляционной функции процесса ξ_t справедливо равенство

$$\text{corr}(\xi_t, \xi_p) = \text{corr}(\xi_{t+1}, \xi_{p+1}).$$

Утверждение доказано.

Модификация процесса с независимыми опорными точками

Если при построении процесса ξ_t вместо стационарного в широком смысле случайного процесса η_i , $i = -1, 0, 1, 2, \dots$ использовать множество независимых друг

от друга одинаково распределённых случайных величин $\tilde{\eta}_i, i = -1, 0, 1, 2, \dots$ с функцией одномерного распределения $F(x), x \in (a, b)$, то для построенного по Алгоритму 4.2 процесса ξ_t справедливо утверждение 4.7.

Утверждение 4.7. Справедливы равенства

$$E\xi_t = \mu,$$

$$D\xi_t = \sigma^2 \left(1 - 2\{t\} + 2\{t\}^2 \right),$$

$$\text{corr}(\xi_t, \xi_p) = \begin{cases} \frac{\{p\} - \{t\}\{p\}}{\tilde{Z}_t \tilde{Z}_p}, & j = i - 1 \\ \frac{1 - \{t\} - \{p\} + 2\{t\}\{p\}}{\tilde{Z}_t \tilde{Z}_p}, & j = i \\ \frac{\{t\} - \{t\}\{p\}}{\tilde{Z}_t \tilde{Z}_p}, & j = i + 1 \\ 0, & j \notin \{i - 1, i, i + 1\} \end{cases}$$

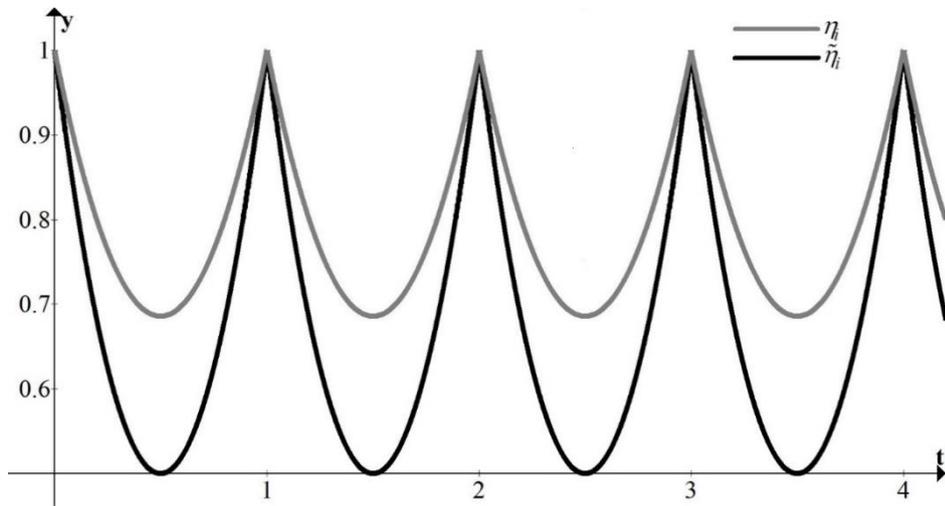
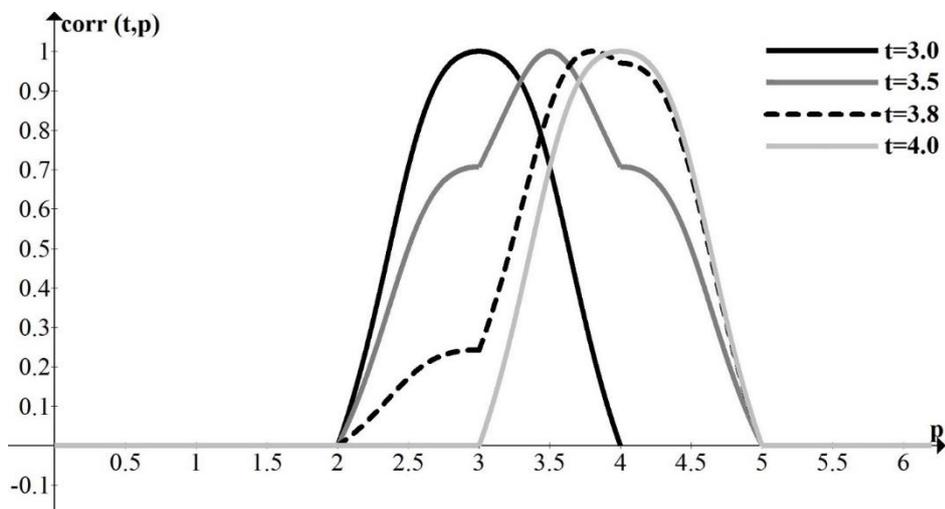
где

$$\tilde{Z}_t = \sqrt{1 - 2\{t\} + 2\{t\}^2}, \quad \tilde{Z}_p = \sqrt{1 - 2\{p\} + 2\{p\}^2}.$$

В качестве иллюстрации на рисунке 4.7 приведены графики дисперсии процесса ξ_t , построенного на основе случайного процесса $\eta_i, i = -1, 0, 1, 2, \dots$ с $F(x) = \Phi_{0,1}(x)$, $r_h = e^{-0.99h}$ и на основе случайных величин $\tilde{\eta}_i, i = -1, 0, 1, 2, \dots$ при $F(x) = \Phi_{0,1}(x)$. На рисунке 4.8 приведены графики $\text{corr}(\xi_t, \xi_p)$ как функции p при фиксированных значениях t в случае использования $\tilde{\eta}_i, i = -1, 0, 1, 2, \dots$

Модификация процесса с неравномерным распределением случайной точки

По аналогии с рассмотренным выше процессом ξ_t , можно рассмотреть кусочно-постоянный процесс ζ_t , построенный на основе стационарного в широком смысле случайного процесса с функцией одномерного распределения

Рис. 4.7. Дисперсия $D\xi_t$.Рис. 4.8. Корреляционная функция $corr(\xi_t, \xi_p)$.

$F(x)$, $x \in (a, b)$, конечными математическим ожиданием μ и дисперсией σ^2 , корреляционной функцией

$$corr(\eta_i, \eta_j) = corr(i, j) = corr(|i - j|) = corr(h) = r_h$$

и независимой от η_i , $i = -1, 0, 1, 2, \dots$ случайной величины α , распределение которой на отрезке $[0, 1]$ отлично от равномерного. Пусть $F_\alpha(x)$, $x \in [0, 1]$ – функция одномерного распределения случайной величины α . В этом случае значение процесса ζ_t в точке t определяется равенством

$$\zeta_t = (1 - F_\alpha(\{t\}))\eta_{i-1} + F_\alpha(\{t\})\eta_i.$$

Несложно проверить, что справедливы утверждения 4.8 и 4.9.

Утверждение 4.8. Математическое ожидание, дисперсия и корреляционная функция процесса ζ_t определяются равенствами

$$\begin{aligned} E\zeta_t &= \mu, \\ D\zeta_t &= \sigma^2 \left(1 - 2F_\alpha(\{t\}) + 2F_\alpha(\{t\})^2 \right) + 2\sigma^2 r_1 F_\alpha(\{t\})(1 - F_\alpha(\{t\})), \\ \text{corr}(\zeta_t, \zeta_p) &= \\ &= \frac{1}{Z_t Z_p} \left(\begin{aligned} &(F_\alpha(\{p\}) - F_\alpha(\{t\})F_\alpha(\{p\}))r_{|i-j-1|} + \\ &+ (1 - F_\alpha(\{t\}) - F_\alpha(\{p\}) + 2F_\alpha(\{t\})F_\alpha(\{p\}))r_{|i-j|} + \\ &+ (F_\alpha(\{t\}) - F_\alpha(\{t\})F_\alpha(\{p\}))r_{|i-j+1|} \end{aligned} \right), \end{aligned}$$

где

$$\begin{aligned} Z_t &= \sqrt{1 - 2F_\alpha(\{t\}) + 2F_\alpha(\{t\})^2 + r_1 2F_\alpha(\{t\})(1 - F_\alpha(\{t\}))}, \\ Z_p &= \sqrt{1 - 2F_\alpha(\{p\}) + 2F_\alpha(\{p\})^2 + r_1 2F_\alpha(\{p\})(1 - F_\alpha(\{p\}))}. \end{aligned}$$

Утверждение 4.9. Кусочно-постоянный процесс ζ_t является периодически коррелированным с периодом равным 1.

4.1.3. Модель периодически коррелированного кусочно-линейного процесса на основе стационарного случайного процесса и регулярного потока

Определение процесса

Пусть $\eta_i, i = -1, 0, 1, 2, \dots$ – стационарный в широком смысле случайный процесс с функцией одномерного распределения $F(x), x \in (a, b)$, конечными математическим ожиданием μ и дисперсией σ^2 , корреляционной функцией

$$\text{corr}(\eta_i, \eta_j) = \text{corr}(i, j) = \text{corr}(|i - j|) = \text{corr}(h) = r_h.$$

Пусть α – независимая от η_i , $i = -1, 0, 1, 2, \dots$ случайная величина, распределенная равномерно на отрезке $[0, 1]$. Рассмотрим случайный процесс ξ_t непрерывного времени $t \geq 0$, который представляет собой кусочно-линейную функцию, проходящую через точки $(-1 + \alpha; \eta_{-1}), (\alpha; \eta_0), (1 + \alpha; \eta_1), (2 + \alpha; \eta_2), (3 + \alpha; \eta_3), \dots$. На рисунке 4.9 представлен пример реализации процесса ξ_t .

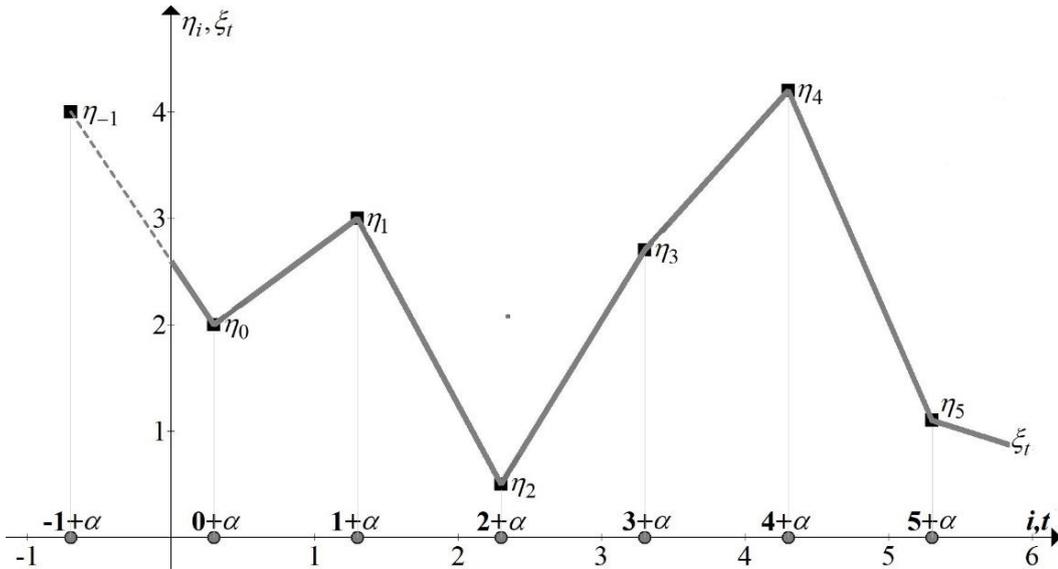


Рис. 4.9 Траектория процесса ξ_t .

Определим значение процесса ξ_t в точке $t \in [i, i + 1)$, $i = 0, 1, 2, \dots$. На интервале $[i, i + 1)$ значение рассматриваемого процесса с вероятностью \tilde{p} определяется уравнением прямой $y_-(t)$, проходящей через точки $(i - 1 + \alpha; \eta_{i-1}), (i + \alpha; \eta_i)$, а с вероятностью $(1 - \tilde{p})$ – уравнением прямой $y_+(t)$, проходящей через точки $(i + \alpha; \eta_i), (i + 1 + \alpha; \eta_{i+1})$. Следовательно,

$$\xi_t = \tilde{p}y_-(t) + (1 - \tilde{p})y_+(t), \quad t \in [i, i + 1), \quad i = 0, 1, 2, \dots \quad (4.1)$$

Уравнения вышеуказанных прямых имеют вид

$$\begin{aligned} y_-(t) &= (\eta_i - \eta_{i-1})t + (\eta_{i-1} - \eta_i)i + (\eta_{i-1} - \eta_i)\alpha + \eta_i, \\ y_+(t) &= (\eta_{i+1} - \eta_i)t + (\eta_i - \eta_{i+1})i + (\eta_i - \eta_{i+1})\alpha + \eta_i. \end{aligned} \quad (4.2)$$

Поскольку случайная величина α равномерно распределена на $[0,1]$, вероятность \tilde{p} определяется цепочкой равенств

$$\tilde{p} = P(t-i \leq \alpha) = 1 - P(\alpha < t-i) = 1 - P(\alpha < \{t\}) = 1 - \{t\}.$$

Используя полученное значение \tilde{p} и уравнения (4.2), равенство (4.1) можно переписать в следующем виде

$$\begin{aligned} \xi_t = & \eta_{i-1}(\{t\}^2 - \{t\}) + \eta_{i-1}\alpha(1 - \{t\}) + \eta_i(-2\{t\}^2 + \{t\} + 1) + \\ & + \eta_i\alpha(2\{t\} - 1) + \eta_{i+1}\{t\}^2 - \eta_{i+1}\alpha\{t\}, \\ & t \in [i, i+1), i = 0, 1, 2, \dots \end{aligned}$$

Далее, для краткости записи, будем использовать следующие обозначения:

$$\begin{aligned} A_t = & (\{t\}^2 - \{t\}), B_t = (1 - \{t\}), C_t = (-2\{t\}^2 + \{t\} + 1), \\ G_t = & (2\{t\} - 1), H_t = \{t\}^2, I_t = -\{t\}. \end{aligned}$$

В этих обозначениях

$$\xi_t = A_t\eta_{i-1} + B_t\eta_{i-1}\alpha + C_t\eta_i + G_t\eta_i\alpha + H_t\eta_{i+1} + I_t\eta_{i+1}\alpha.$$

Свойства процесса

Лемма 4.3. Математическое ожидание процесса ξ_t не зависит от t и $E\xi_t = \mu$.

Дисперсия рассматриваемого процесса

$$D\xi_t = \sigma^2 \left(\begin{aligned} & 6\{t\}^4 - 12\{t\}^3 + 6\{t\}^2 + \frac{2}{3} + \\ & + \frac{1}{3}r_1(2\{t\} - 1)^2(-6\{t\}^2 + 6\{t\} + 1) + \frac{2}{3}r_2\{t\}(3\{t\}^3 - 6\{t\}^2 + 4\{t\} + 1) \end{aligned} \right)$$

является периодической функцией t , $t \geq 0$ с периодом равным 1.

Доказательство. Используя свойства математического ожидания, независимость случайной величины α от η_i , $i = -1, 0, 1, 2, \dots$ и то, что

$E\alpha = \frac{1}{2}$, $E\eta_i = \mu$, $i = -1, 0, 1, 2, \dots$, вычислим математическое ожидание процесса ξ_t :

$$E\xi_t = E(A_t\eta_{i-1} + B_t\eta_{i-1}\alpha + C_t\eta_i + G_t\eta_i\alpha + H_t\eta_{i+1} + I_t\eta_{i+1}\alpha) =$$

$$\begin{aligned}
&= A_t E \eta_{i-1} + B_t E \eta_{i-1} \alpha + C_t E \eta_i + G_t E \eta_i \alpha + H_t E \eta_{i+1} + I_t E \eta_{i+1} \alpha = \\
&= A_t E \eta_{i-1} + B_t E \eta_{i-1} E \alpha + C_t E \eta_i + G_t E \eta_i E \alpha + H_t E \eta_{i+1} + I_t E \eta_{i+1} E \alpha = \\
&= A_t \mu + B_t \mu \frac{1}{2} + C_t \mu + G_t \mu \frac{1}{2} + H_t \mu + I_t \mu \frac{1}{2} = \\
&= \mu \left(\begin{aligned} &\{t\}^2 - \{t\} + \frac{1}{2}(1 - \{t\}) - 2\{t\}^2 + \{t\} + 1 + \\ &+ \frac{1}{2}(2\{t\} - 1) + \{t\}^2 - \frac{1}{2}\{t\} \end{aligned} \right) = \mu.
\end{aligned}$$

Найдём $E \xi_t^2$:

$$\begin{aligned}
E \xi_t^2 &= E (A_t \eta_{i-1} + B_t \eta_{i-1} \alpha + C_t \eta_i + G_t \eta_i \alpha + H_t \eta_{i+1} + I_t \eta_{i+1} \alpha)^2 = \\
&= A_t^2 E \eta_{i-1}^2 + B_t^2 E \eta_{i-1}^2 E \alpha^2 + C_t^2 E \eta_i^2 + G_t^2 E \eta_i^2 E \alpha^2 + H_t^2 E \eta_{i+1}^2 + I_t^2 E \eta_{i+1}^2 E \alpha^2 + \\
&+ 2A_t B_t E \eta_{i-1}^2 E \alpha + 2A_t C_t E \eta_{i-1} \eta_i + 2A_t G_t E \eta_{i-1} \eta_i \alpha + 2A_t H_t E \eta_{i-1} \eta_{i+1} + \\
&+ 2A_t I_t E \eta_{i-1} \eta_{i+1} E \alpha + 2B_t C_t E \eta_{i-1} \eta_i E \alpha + 2B_t G_t E \eta_{i-1} \eta_i E \alpha^2 + \\
&+ 2B_t H_t E \eta_{i-1} \eta_{i+1} E \alpha + 2B_t I_t E \eta_{i-1} \eta_{i+1} E \alpha^2 + 2C_t G_t E \eta_i^2 E \alpha + \\
&+ 2C_t H_t E \eta_i \eta_{i+1} + 2C_t I_t E \eta_i \eta_{i+1} E \alpha + 2G_t H_t E \eta_i \eta_{i+1} E \alpha + \\
&+ 2G_t I_t E \eta_i \eta_{i+1} E \alpha^2 + 2H_t I_t E \eta_{i+1}^2 E \alpha.
\end{aligned}$$

Используя то, что

$$\begin{aligned}
E \alpha &= \frac{1}{2}, \quad E \alpha^2 = \frac{1}{3}, \\
E \eta_{i-1} &= E \eta_i = E \eta_{i+1} = \mu, \\
E \eta_{i-1}^2 &= E \eta_i^2 = E \eta_{i+1}^2 = \sigma^2 + \mu^2, \\
r_1 &= \frac{E \eta_{i-1} \eta_i - \mu^2}{\sigma^2} = \frac{E \eta_i \eta_{i+1} - \mu^2}{\sigma^2}, \quad r_2 = \frac{E \eta_{i-1} \eta_{i+1} - \mu^2}{\sigma^2}, \\
E \eta_{i-1} \eta_i &= E \eta_i \eta_{i+1} = \sigma^2 r_1 + \mu^2, \quad E \eta_{i-1} \eta_{i+1} = \sigma^2 r_2 + \mu^2,
\end{aligned}$$

получаем

$$E \xi_t^2 = \mu^2 \left(\begin{aligned} &A_t^2 + C_t^2 + H_t^2 + A_t B_t + C_t G_t + H_t I_t + A_t G_t + B_t C_t + C_t I_t + G_t H_t + \\ &+ A_t I_t + B_t H_t + \frac{1}{3}(B_t^2 + G_t^2 + I_t^2) + \frac{2}{3}(B_t G_t + G_t I_t + B_t I_t) + \\ &+ 2(A_t C_t + C_t H_t + A_t H_t) \end{aligned} \right) +$$

$$\begin{aligned}
& +\sigma^2\left(A_t^2 + C_t^2 + H_t^2 + A_t B_t + C_t G_t + H_t I_t + \frac{1}{3}(B_t^2 + G_t^2 + I_t^2)\right) + \\
& +\sigma^2 r_1\left(A_t G_t + B_t C_t + C_t I_t + G_t H_t + \frac{2}{3}(B_t G_t + G_t I_t) + 2(A_t C_t + C_t H_t)\right) \\
& +\sigma^2 r_2\left(A_t I_t + B_t H_t + \frac{2}{3}B_t I_t + 2A_t H_t\right) = \\
& = \mu^2 + \sigma^2\left(6\{t\}^4 - 12\{t\}^3 + 6\{t\}^2 + \frac{2}{3}\right) + \sigma^2 \frac{1}{3} r_1 (2\{t\} - 1)^2 (-6\{t\}^2 + 6\{t\} + 1) + \\
& + \sigma^2 \frac{2}{3} r_2 \{t\} (3\{t\}^3 - 6\{t\}^2 + 4\{t\} + 1).
\end{aligned}$$

Следовательно,

$$\begin{aligned}
D\xi_t &= E\xi_t^2 - (E\xi_t)^2 = \\
&= \sigma^2 \left(\begin{aligned} & 6\{t\}^4 - 12\{t\}^3 + 6\{t\}^2 + \frac{2}{3} + \\ & + \frac{1}{3} r_1 (2\{t\} - 1)^2 (-6\{t\}^2 + 6\{t\} + 1) + \frac{2}{3} r_2 \{t\} (3\{t\}^3 - 6\{t\}^2 + 4\{t\} + 1) \end{aligned} \right).
\end{aligned}$$

Периодичность $D\xi_t$ следует из периодичности функции $y(t) = \{t\}$.

Лемма доказана.

В качестве иллюстрации, на рисунке 4.10 приведен график $D\xi_t$ при $F(x) = \Phi_{0,1}(x)$, $r_h = e^{-0.99h}$.

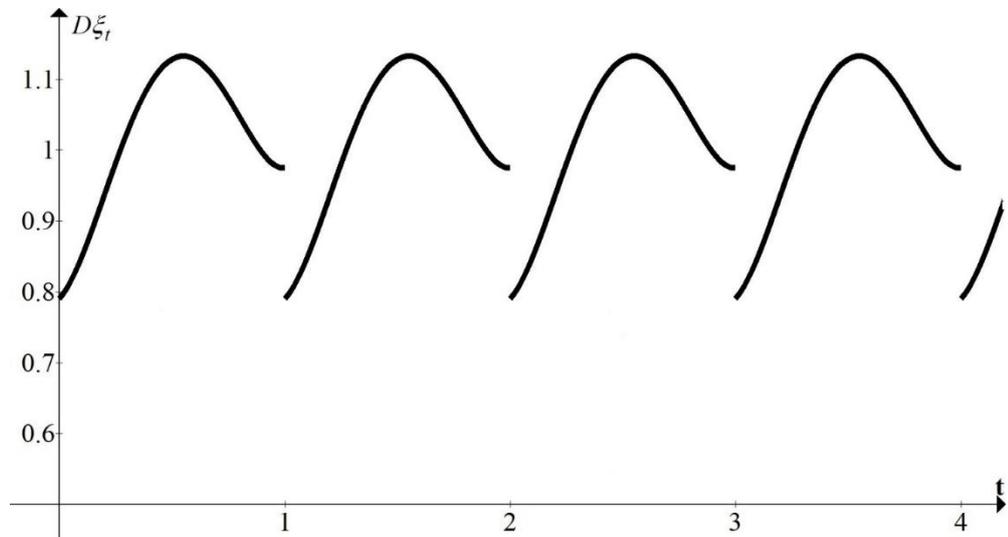


Рис. 4.10. Дисперсия $D\xi_t$ при $F(x) = \Phi_{0,1}(x)$, $r_h = e^{-0.99h}$.

В рассмотренном примере функция $D\xi_t$ имеет разрывы первого рода при всех целых положительных значениях t .

Лемма 4.4. Функция $D\xi_t$ непрерывна при всех целых положительных значениях t тогда и только тогда, когда $r_2 = 0$.

Доказательство. Действительно, для $i = 1, 2, 3, \dots$

$$\begin{aligned} \lim_{t \rightarrow i-0} D\xi_t &= \sigma^2 \left(\lim_{t \rightarrow i-0} \left(6\{t\}^4 - 12\{t\}^3 + 6\{t\}^2 + \frac{2}{3} \right) + \right. \\ &\quad \left. + \lim_{t \rightarrow i-0} \left(\frac{1}{3} r_1 (2\{t\} - 1)^2 (-6\{t\}^2 + 6\{t\} + 1) \right) + \right. \\ &\quad \left. + \lim_{t \rightarrow i-0} \left(\frac{2}{3} r_2 \{t\} (3\{t\}^3 - 6\{t\}^2 + 4\{t\} + 1) \right) \right) \\ &\stackrel{x=\{t\}}{=} \sigma^2 \left(\lim_{x \rightarrow 1-0} \left(6x^4 - 12x^3 + 6x^2 + \frac{2}{3} \right) + \right. \\ &\quad \left. + \lim_{x \rightarrow 1-0} \left(\frac{1}{3} r_1 (2x - 1)^2 (-6x^2 + 6x + 1) \right) + \right. \\ &\quad \left. + \lim_{x \rightarrow 1-0} \left(\frac{2}{3} r_2 x (3x^3 - 6x^2 + 4x + 1) \right) \right) = \sigma^2 \left(\frac{2}{3} + \frac{1}{3} r_1 + \frac{4}{3} r_2 \right), \\ \lim_{t \rightarrow i+0} D\xi_t &= \sigma^2 \left(\lim_{t \rightarrow i+0} \left(6\{t\}^4 - 12\{t\}^3 + 6\{t\}^2 + \frac{2}{3} \right) + \right. \\ &\quad \left. + \lim_{t \rightarrow i+0} \left(\frac{1}{3} r_1 (2\{t\} - 1)^2 (-6\{t\}^2 + 6\{t\} + 1) \right) + \right. \\ &\quad \left. + \lim_{t \rightarrow i+0} \left(\frac{2}{3} r_2 \{t\} (3\{t\}^3 - 6\{t\}^2 + 4\{t\} + 1) \right) \right) \\ &\stackrel{x=\{t\}}{=} \sigma^2 \left(\lim_{x \rightarrow 0+0} \left(6x^4 - 12x^3 + 6x^2 + \frac{2}{3} \right) + \right. \\ &\quad \left. + \lim_{x \rightarrow 0+0} \left(\frac{1}{3} r_1 (2x - 1)^2 (-6x^2 + 6x + 1) \right) + \right. \\ &\quad \left. + \lim_{x \rightarrow 0+0} \left(\frac{2}{3} r_2 x (3x^3 - 6x^2 + 4x + 1) \right) \right) = \sigma^2 \left(\frac{2}{3} + \frac{1}{3} r_1 \right). \end{aligned}$$

Очевидно, что равенство

$$\lim_{t \rightarrow i-0} D\xi_t = \lim_{t \rightarrow i+0} D\xi_t,$$

выполнение которого необходимо и достаточно для непрерывности функции $D\xi_t$ в точках $t = i$, $i = 1, 2, 3, \dots$ при том, что $D\xi_i = \sigma^2 \left(\frac{2}{3} + \frac{1}{3} r_1 \right)$, справедливо только при $r_2 = 0$.

Лемма доказана.

Лемма 4.5. Совместное математическое ожидание $E\xi_t \xi_p$ процесса ξ_t для $t \in [i, i+1)$, $i = 0, 1, 2, \dots$ имеет вид

1) при $p \in [j, j+1)$, $0 \leq j \leq i-2$ или $j \geq i+2$

$$\begin{aligned}
E\xi_t \xi_p &= \mu^2 + \sigma^2 r_{|i-j-2|} \frac{1}{6} \{p\} (\{t\} - 1) (2 - 3\{p\} - 3\{t\} + 6\{t\}\{p\}) + \\
&+ \sigma^2 r_{|i-j-1|} \frac{1}{6} \left(\begin{aligned} &1 + 6\{p\} - 3\{p\}^2 + \{t\}^2 (3 + 18\{p\} - 24\{p\}^2) + \\ &+ \{t\} (-4 - 26\{p\} + 30\{p\}^2) \end{aligned} \right) + \\
&+ \sigma^2 r_{|i-j|} \left(\begin{aligned} &\{t\}\{p\} (7 - 6\{t\}) + \frac{1}{6} (4 - 3\{t\}^2) + \{p\}^2 \left(-\frac{1}{2} - 6\{t\} + 6\{t\}^2 \right) \end{aligned} \right) + \\
&+ \sigma^2 r_{|i-j+1|} \frac{1}{6} \left(\begin{aligned} &1 + 6\{t\} - 3\{t\}^2 + \{p\}^2 (3 + 18\{t\} - 24\{t\}^2) + \\ &\{p\} (-4 - 26\{t\} + 30\{t\}^2) \end{aligned} \right) + \\
&+ \sigma^2 r_{|i-j+2|} \frac{1}{6} \{t\} (\{p\} - 1) (2 - 3\{p\} - 3\{t\} + 6\{t\}\{p\}),
\end{aligned}$$

2) при $p \in [i-1, i)$

$$\begin{aligned}
E\xi_t \xi_p &= \mu^2 + \sigma^2 \frac{1}{6} \left(\begin{aligned} &1 + 6\{p\} - 3\{p\}^2 + \{t\}^2 (3 + 18\{p\} - 24\{p\}^2) + \\ &+ \{t\} (-4 - 26\{p\} + 30\{p\}^2) \end{aligned} \right) + \\
&+ \sigma^2 r_1 \frac{1}{6} \left(\begin{aligned} &4 - 3\{t\}^2 + 3\{p\}^2 (14\{t\}^2 - 15\{t\}) + \{p\} (-2 + 47\{t\} - 39\{t\}^2) \end{aligned} \right) + \\
&+ \sigma^2 r_2 \frac{1}{6} \left(\begin{aligned} &1 + 6\{t\} - 3\{t\}^2 + \{p\}^2 (3 + 18\{t\} - 24\{t\}^2) + \\ &+ \{p\} (-4 - 26\{t\} + 30\{t\}^2) \end{aligned} \right) +
\end{aligned}$$

$$+\sigma^2 r_3 \frac{1}{6} \{t\} (\{p\} - 1) (2 - 3\{p\} - 3\{t\} + 6\{t\}\{p\}),$$

3) при $p \in [i, i+1)$

$$E_{\xi_t \xi_p} = \mu^2 + \sigma^2 \left(\begin{aligned} &\{t\}\{p\}(7 - 6\{t\}) + \frac{1}{6}(4 - 3\{t\}^2) + \\ &+ \{p\}^2 \left(-\frac{1}{2} - 6\{t\} + 6\{t\}^2 \right) \end{aligned} \right) + \\ + \sigma^2 r_1 \frac{1}{3} \left(1 + \{t\} - 24\{t\}\{p\}^2 (\{t\} - 1) + \{p\} (1 - 26\{t\} + 24\{t\}^2) \right) + \\ + \sigma^2 r_2 \frac{1}{6} \left(3\{p\}^2 (2\{t\} - 1) - 2\{t\} + 3\{t\}^2 - 2\{p\} (1 - 5\{t\} + 6\{t\}^2) \right),$$

4) при $p \in [i+1, i+2)$

$$E_{\xi_t \xi_p} = \mu^2 + \sigma^2 \frac{1}{6} \left(\begin{aligned} &1 + 6\{t\} - 3\{t\}^2 + \{p\}^2 (3 + 18\{t\} - 24\{t\}^2) + \\ &+ \{p\} (-4 - 26\{t\} + 30\{t\}^2) \end{aligned} \right) + \\ + \sigma^2 r_1 \frac{1}{6} \left(4 - 3\{p\}^2 + 3\{t\}^2 (14\{p\}^2 - 15\{p\}) + \{t\} (-2 + 47\{p\} - 39\{p\}^2) \right) + \\ + \sigma^2 r_2 \frac{1}{6} \left(\begin{aligned} &1 + 6\{p\} - 3\{p\}^2 + \{t\}^2 (3 + 18\{p\} - 24\{p\}^2) + \\ &+ \{t\} (-4 - 26\{p\} + 30\{p\}^2) \end{aligned} \right) + \\ + \sigma^2 r_3 \frac{1}{6} \{p\} (\{t\} - 1) (2 - 3\{p\} - 3\{t\} + 6\{t\}\{p\}).$$

Доказательство. Заметим, что в четырёх случаях, указанных в формулировке утверждения, значение ξ_p определяется равенствам

$$1) \quad \xi_p = A_p \eta_{j-1} + B_p \eta_{j-1} \alpha + C_p \eta_j + G_p \eta_j \alpha + H_p \eta_{j+1} + I_p \eta_{j+1} \alpha,$$

$$2) \quad \xi_p = A_p \eta_{i-2} + B_p \eta_{i-2} \alpha + C_p \eta_{i-1} + G_p \eta_{i-1} \alpha + H_p \eta_i + I_p \eta_i \alpha,$$

$$3) \quad \xi_p = A_p \eta_{i-1} + B_p \eta_{i-1} \alpha + C_p \eta_i + G_p \eta_i \alpha + H_p \eta_{i+1} + I_p \eta_{i+1} \alpha,$$

$$4) \quad \xi_p = A_p \eta_i + B_p \eta_i \alpha + C_p \eta_{i+1} + G_p \eta_{i+1} \alpha + H_p \eta_{i+2} + I_p \eta_{i+2} \alpha.$$

Докажем утверждение в первом случае (при $p \in [j, j+1)$, $0 \leq j \leq i-2$ или $j \geq i+2$). Запишем произведение $\xi_t \xi_p$:

$$\begin{aligned}
& \xi_t \xi_p = (A_t \eta_{i-1} + B_t \eta_{i-1} \alpha + C_t \eta_i + G_t \eta_i \alpha + H_t \eta_{i+1} + I_t \eta_{i+1} \alpha) \times \\
& \times (A_p \eta_{j-1} + B_p \eta_{j-1} \alpha + C_p \eta_j + G_p \eta_j \alpha + H_p \eta_{j+1} + I_p \eta_{j+1} \alpha) = \\
& = A_t \eta_{i-1} A_p \eta_{j-1} + A_t \eta_{i-1} B_p \eta_{j-1} \alpha + A_t \eta_{i-1} C_p \eta_j + A_t \eta_{i-1} G_p \eta_j \alpha + A_t \eta_{i-1} H_p \eta_{j+1} + \\
& + A_t \eta_{i-1} I_p \eta_{j+1} \alpha + B_t \eta_{i-1} \alpha A_p \eta_{j-1} + B_t \eta_{i-1} B_p \eta_{j-1} \alpha^2 + B_t \eta_{i-1} \alpha C_p \eta_j + \\
& + B_t \eta_{i-1} G_p \eta_j \alpha^2 + B_t \eta_{i-1} \alpha H_p \eta_{j+1} + B_t \eta_{i-1} I_p \eta_{j+1} \alpha^2 + C_t \eta_i A_p \eta_{j-1} + \\
& + C_t \eta_i B_p \eta_{j-1} \alpha + C_t \eta_i C_p \eta_j + C_t \eta_i G_p \eta_j \alpha + C_t \eta_i H_p \eta_{j+1} + C_t \eta_i I_p \eta_{j+1} \alpha + \\
& + G_t \eta_i \alpha A_p \eta_{j-1} + G_t \eta_i B_p \eta_{j-1} \alpha^2 + G_t \eta_i \alpha C_p \eta_j + G_t \eta_i G_p \eta_j \alpha^2 + G_t \eta_i \alpha H_p \eta_{j+1} + \\
& + G_t \eta_i I_p \eta_{j+1} \alpha^2 + H_t \eta_{i+1} A_p \eta_{j-1} + H_t \eta_{i+1} B_p \eta_{j-1} \alpha + H_t \eta_{i+1} C_p \eta_j + H_t \eta_{i+1} G_p \eta_j \alpha + \\
& + H_t \eta_{i+1} H_p \eta_{j+1} + H_t \eta_{i+1} I_p \eta_{j+1} \alpha + I_t \eta_{i+1} \alpha A_p \eta_{j-1} + I_t \eta_{i+1} B_p \eta_{j-1} \alpha^2 + \\
& + I_t \eta_{i+1} \alpha C_p \eta_j + I_t \eta_{i+1} G_p \eta_j \alpha^2 + I_t \eta_{i+1} \alpha H_p \eta_{j+1} + I_t \eta_{i+1} I_p \eta_{j+1} \alpha^2.
\end{aligned}$$

Используя свойства случайной величины α , свойства математического ожидания и равенства

$$E \eta_{i-1} \eta_{j-1} = E \eta_i \eta_j = E \eta_{i+1} \eta_{j+1} = \sigma^2 r_{|i-j|} + \mu^2,$$

$$E \eta_{i-1} \eta_j = E \eta_i \eta_{j+1} = \sigma^2 r_{|i-j-1|} + \mu^2,$$

$$E \eta_i \eta_{j-1} = E \eta_{i+1} \eta_j = \sigma^2 r_{|i-j+1|} + \mu^2,$$

$$E \eta_{i-1} \eta_{j+1} = \sigma^2 r_{|i-j-2|} + \mu^2,$$

$$E \eta_{i+1} \eta_{j-1} = \sigma^2 r_{|i-j+2|} + \mu^2,$$

найдем $E \xi_t \xi_p$:

$$\begin{aligned}
E \xi_t \xi_p &= \left(\sigma^2 r_{|i-j-2|} + \mu^2 \right) \left(A_t H_p + \frac{1}{2} (A_t I_p + B_t H_p) + \frac{1}{3} (B_t I_p) \right) + \\
& + \left(\sigma^2 r_{|i-j-1|} + \mu^2 \right) \left(A_t C_p + C_t H_p + \frac{1}{3} (B_t G_p + G_t I_p) + \right. \\
& \left. + \frac{1}{2} (A_t G_p + B_t C_p + C_t I_p + G_t H_p) \right) + \\
& + \left(\sigma^2 r_{|i-j|} + \mu^2 \right) \left(A_t A_p + C_t C_p + H_t H_p + \frac{1}{3} (B_t B_p + G_t G_p + I_t I_p) + \right. \\
& \left. + \frac{1}{2} (A_t B_p + B_t A_p + C_t G_p + G_t C_p + H_t I_p + I_t H_p) \right) +
\end{aligned}$$

$$\begin{aligned}
& + \left(\sigma^2 \eta_{|i-j+1|} + \mu^2 \right) \left(C_t A_p + H_t C_p + \frac{1}{3} (G_t B_p + I_t G_p) + \right. \\
& \quad \left. + \frac{1}{2} (C_t B_p + G_t A_p + H_t G_p + I_t C_p) \right) + \\
& + \left(\sigma^2 \eta_{|i-j+2|} + \mu^2 \right) \left(H_t A_p + \frac{1}{2} (H_t B_p + I_t A_p) + \frac{1}{3} (I_t B_p) \right) = \\
& = \mu^2 + \sigma^2 \eta_{|i-j-2|} \frac{1}{6} \{p\} (\{t\} - 1) (2 - 3\{p\} - 3\{t\} + 6\{t\}\{p\}) + \\
& + \sigma^2 \eta_{|i-j-1|} \frac{1}{6} \left(1 + 6\{p\} - 3\{p\}^2 + \{t\}^2 (3 + 18\{p\} - 24\{p\}^2) + \right. \\
& \quad \left. + \{t\} (-4 - 26\{p\} + 30\{p\}^2) \right) + \\
& + \sigma^2 \eta_{|i-j|} \left(\{t\}\{p\} (7 - 6\{t\}) + \frac{1}{6} (4 - 3\{t\}^2) + \{p\}^2 \left(-\frac{1}{2} - 6\{t\} + 6\{t\}^2 \right) \right) + \\
& + \sigma^2 \eta_{|i-j+1|} \frac{1}{6} \left(1 + 6\{t\} - 3\{t\}^2 + \{p\}^2 (3 + 18\{t\} - 24\{t\}^2) + \right. \\
& \quad \left. + \{p\} (-4 - 26\{t\} + 30\{t\}^2) \right) + \\
& + \sigma^2 \eta_{|i-j+2|} \frac{1}{6} \{t\} (\{p\} - 1) (2 - 3\{p\} - 3\{t\} + 6\{t\}\{p\}).
\end{aligned}$$

Доказательство утверждения при $p \in [i-1, i)$, $p \in [i, i+1)$, $p \in [i+1, i+2)$ проводится аналогично.

Лемма доказана.

Утверждение 4.10. Случайный процесс ξ_t является периодически коррелированным с периодом равным 1.

Доказательство. Доказательство этого утверждения следует из лемм 4.3, 4.5 и свойств функции $y(t) = \{t\}$.

Модификация процесса с независимыми опорными точками

Если при построении процесса ξ_t вместо стационарного в широком смысле случайного процесса η_i , $i = -1, 0, 1, 2, \dots$ использовать множество независимых друг от друга одинаково распределённых случайных величин $\tilde{\eta}_i$, $i = -1, 0, 1, 2, \dots$ с

функцией одномерного распределения $F(x)$, $x \in (a, b)$, то для процесса ξ_t справедливы леммы 4.6 и 4.7.

Лемма 4.6.

$$E\xi_t = \mu,$$

$$D\xi_t = \sigma^2 \left(6\{t\}^4 - 12\{t\}^3 + 6\{t\}^2 + \frac{2}{3} \right).$$

Лемма 4.7. Корреляционная функция $\text{corr}(\xi_t, \xi_p)$ процесса ξ_t для $t \in [i, i+1)$, $i = 0, 1, 2, \dots$ тождественно равна 0 при $p < i-2$ и $p \geq i+3$, а для $p \in [i-2, i+3)$ справедливы равенства

$$\text{corr}(\xi_t, \xi_p) = \begin{cases} \frac{\{p\}(\{t\}-1)(2-3\{p\}-3\{t\}+6\{t\}\{p\})}{6\tilde{Z}_t\tilde{Z}_p}, & p \in [i-2, i-1), \\ \frac{1+6\{p\}-3\{p\}^2+\{t\}^2(3+18\{p\}-24\{p\}^2)}{6\tilde{Z}_t\tilde{Z}_p} + \\ \quad + \frac{\{t\}(-4-26\{p\}+30\{p\}^2)}{6\tilde{Z}_t\tilde{Z}_p}, & p \in [i-1, i), \\ \frac{\{t\}\{p\}(7-6\{t\})+\frac{1}{6}(4-3\{t\}^2)}{\tilde{Z}_t\tilde{Z}_p} + \\ \quad + \frac{\{p\}^2\left(-\frac{1}{2}-6\{t\}+6\{t\}^2\right)}{\tilde{Z}_t\tilde{Z}_p}, & p \in [i, i+1), \\ \frac{1+6\{t\}-3\{t\}^2+\{p\}^2(3+18\{t\}-24\{t\}^2)}{6Z_tZ_p} + \\ \quad + \frac{\{p\}(-4-26\{t\}+30\{t\}^2)}{\tilde{Z}_t\tilde{Z}_p}, & p \in [i+1, i+2), \\ \frac{\{t\}(\{p\}-1)(2-3\{p\}-3\{t\}+6\{t\}\{p\})}{6\tilde{Z}_t\tilde{Z}_p}, & p \in [i+2, i+3), \end{cases}$$

где

$$\tilde{Z}_t = \sqrt{6\{t\}^4 - 12\{t\}^3 + 6\{t\}^2 + \frac{2}{3}}, \quad \tilde{Z}_p = \sqrt{6\{p\}^4 - 12\{p\}^3 + 6\{p\}^2 + \frac{2}{3}}.$$

В качестве иллюстрации на рисунке 4.11 приведены графики $\text{corr}(\xi_t, \xi_p)$ как функции p при фиксированных значениях t .

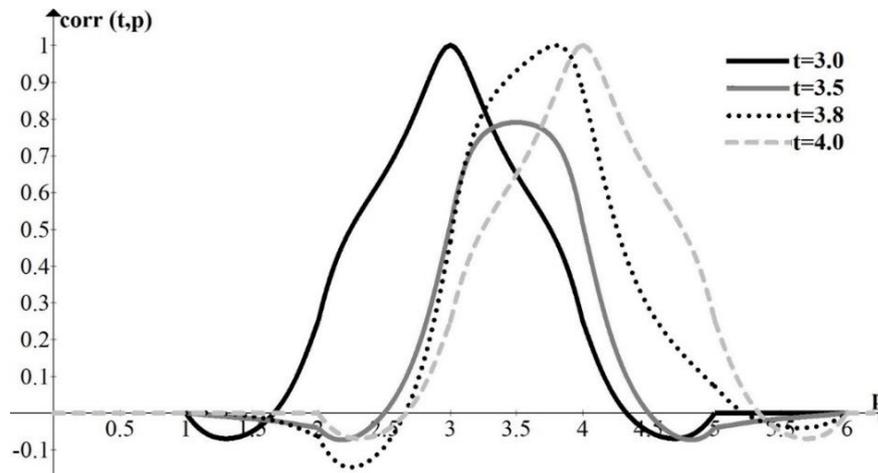


Рис. 4.11 Корреляционная функция $\text{corr}(\xi_t, \xi_p)$.

4.2. Асимптотически периодически коррелированные процессы

В Главах 2 и 3, а также в параграфе 4.1 Главы 4 рассматривались случайные процессы, относящиеся к классу периодически коррелированных. Следует отметить, что помимо периодически коррелированных процессов выделяют и другие классы случайных процессов с периодическими свойствами. Например, процессы, удовлетворяющие определениям 4.1-4.3.

Определение 4.1. Случайный процесс $\xi(t)$, $t \in \mathbb{R}$ называется периодическим с вероятностью 1 [21], если с вероятностью 1 выполняется равенство

$$\xi(t+T) = \xi(t), \quad t \in \mathbb{R}.$$

Определение 4.2. Случайный процесс $\xi(t)$, $t \in \mathbb{R}$ периодический в узком смысле [21, 70], если все его n -мерные распределения инвариантны при сдвиге каждого из аргументов на величину T , т.е.

$$P(\xi(t_1+T) < x_1, \dots, \xi(t_n+T) < x_n) = P(\xi(t_1) < x_1, \dots, \xi(t_n) < x_n).$$

Определение 4.3. Случайный процесс $\xi(t)$, $t \in \mathbb{R}$ называется почти-периодически коррелированным [14], если его ковариационная функция $cov(\xi(t_1), \xi(t_2))$ существует и равномерно непрерывна, а при любых фиксированных значениях t_1, t_2 функции $cov(\xi(t_1+T), \xi(t_2+T))$ являются почти-периодическими функциями T .

В определении 4.3 используется понятие почти-периодической функции. Напомним, что непрерывная функция $f(x)$, $x \in \mathbb{R}$ называется почти-периодической, если для каждого $\varepsilon > 0$ можно указать такое $l = l(\varepsilon)$, что в каждом интервале оси x длиной l найдётся хотя бы одно число $\tau = \tau(\varepsilon)$, для которого при любом x выполняется неравенство $|f(x+\tau) - f(x)| < \varepsilon$ [7]. Другими словами, $f(x)$ – функция, значения которой при добавлении к аргументу надлежащим образом выбранных постоянных чисел (почти периодов) приближённо повторяются.

Определим ещё один класс случайных процессов с периодическими характеристиками.

Определение 4.4. Случайный процесс $\xi(t)$ аргумента t называется асимптотически периодически коррелированным случайным процессом (АПКСП) с периодом T , если существуют периодические с периодом T по аргументу t функции $m(t), v(t), c(t, t+h)$, удовлетворяющие условию: для всех h и для любого $\varepsilon > 0$ существуют такие T_E, T_D, T_c (зависящие от ε, h), при которых

$$\begin{aligned} \forall t > T_E \quad |E\xi(t) - m(t)| &< \varepsilon, \\ \forall t > T_D \quad |D\xi(t) - v(t)| &< \varepsilon, \\ \forall t > T_c \quad |corr(\xi(t), \xi(t+h)) - c(t, t+h)| &< \varepsilon. \end{aligned}$$

Покажем, что для АПКСП справедливо следующее утверждение.

Утверждение 4.11. Если $\xi(t)$ – АПКСП с периодом T , то для всех h и для любого $\varepsilon > 0$ выполняется

$$\begin{aligned} \forall t > T_E \quad \forall k \in \mathbb{N} \quad |E\xi(t) - E\xi(t+kT)| < \varepsilon, \\ \forall t > T_D \quad \forall k \in \mathbb{N} \quad |D\xi(t) - D\xi(t+kT)| < \varepsilon, \\ \forall t > T_c \quad \forall k \in \mathbb{N} \quad |corr(\xi(t), \xi(t+h)) - corr(\xi(t+kT), \xi(t+kT+h))| < \varepsilon. \end{aligned}$$

Доказательство. Покажем, что выполнено первое неравенство. Пусть значение ε зафиксировано. Тогда, по определению АПКСП,

$$\forall t > T_E \quad |E\xi(t) - m(t)| < \varepsilon.$$

Следовательно, для всех $k \in \mathbb{N}$ верно неравенство

$$\forall t > T_E \quad |E\xi(t+kT) - m(t+kT)| < \varepsilon.$$

Поскольку функция $m(t)$ периодическая и $m(t) = m(t+kT)$, справедлива цепочка неравенств

$$\begin{aligned} |E\xi(t) - E\xi(t+kT)| &= |E\xi(t) - E\xi(t+kT) - m(t) + m(t+kT)| = \\ &= |(E\xi(t) - m(t)) - (E\xi(t+kT) - m(t+kT))| \leq \\ &\leq |E\xi(t) - m(t)| + |E\xi(t+kT) - m(t+kT)| < 2\varepsilon. \end{aligned}$$

Что и требовалось доказать. Доказательство выполнения соответствующих неравенств для $D\xi(t)$ и $corr(\xi(t), \xi(t+h))$ проводится аналогичным образом.

Утверждение доказано.

Покажем, что если процесс $\xi(t)$ – процесс дискретного аргумента t , то справедливо утверждение 4.12.

Утверждение 4.12. Для того чтобы процесс $\xi(t)$ дискретного аргумента t являлся АПКСП с периодом T , необходимо и достаточно, чтобы существовало $T > 0$, удовлетворяющее условию: для всех h и для любого $\varepsilon > 0$ существуют такие T_E, T_D, T_c (зависящие от ε, h), при которых

$$\begin{aligned} \forall t > T_E \quad \forall k \in \mathbb{N} \quad |E\xi(t) - E\xi(t+kT)| < \varepsilon, \\ \forall t > T_D \quad \forall k \in \mathbb{N} \quad |D\xi(t) - D\xi(t+kT)| < \varepsilon, \end{aligned} \tag{4.3}$$

$$\forall t > T_c \quad \forall k \in \mathbb{N} \quad \left| \text{corr}(\xi(t), \xi(t+h)) - \text{corr}(\xi(t+kT), \xi(t+kT+h)) \right| < \varepsilon.$$

Доказательство. Необходимость выполнения условия (4.3) следует из утверждения 4.11. Покажем, что условие (4.3) является достаточным.

Без потери общности будем считать, что $t=1,2,3,\dots$. Рассмотрим T последовательностей точек

$$\begin{aligned} x_0^1 &= 1, x_1^1 = 1+T, \dots, x_k^1 = 1+kT, \dots \\ x_0^2 &= 2, x_1^2 = 2+T, \dots, x_k^2 = 2+kT, \dots \\ &\dots \\ x_0^T &= T, x_1^T = T+T, \dots, x_k^T = T+kT, \dots, \end{aligned}$$

т.е. последовательности вида $x_k^i = i+kT$, $k=0,1,2,\dots$, $i=1,2,\dots,T$, и T соответствующим им последовательностей

$$y_k^i = E\xi(x_k^i), \quad k=0,1,2,\dots, \quad i=1,2,\dots,T.$$

Из условия (4.3) следует, что последовательности y_k^i , $k=0,1,2,\dots$, $i=1,2,\dots,T$ являются фундаментальными [88], т.е.

$$\forall \varepsilon > 0 \exists T_E^i : \forall k_1, k_2 > T_E^i \quad \left| y_{k_2}^i - y_{k_1}^i \right| < \varepsilon,$$

и, следовательно, по критерию Коши [88] существует

$$\lim_{k \rightarrow \infty} y_k^i = \lim_{k \rightarrow \infty} E\xi(x_k^i) = \lim_{k \rightarrow \infty} E\xi(i+kT) = E_i, \quad i=1,2,\dots,T.$$

Определим функцию $m(t)$ по правилу

$$m(t) = m(x_k^i) = m(i+kT) = E_i, \quad k=0,1,2,\dots, \quad i=1,2,\dots,T.$$

Функция $m(t)$ по построению является периодической, а из определения предела последовательности следует, что при заданном $\varepsilon > 0$

$$\exists \tilde{T}_E^i : \forall k > \tilde{T}_E^i \quad \left| E\xi(x_k^i) - m(x_k^i) \right| < \varepsilon, \quad i=1,2,\dots,T.$$

Пусть $\tilde{T}_E = \max_{i=1,2,\dots,T} \tilde{T}_E^i$. Тогда $\forall k > \tilde{T}_E \quad \left| E\xi(x_k^i) - m(x_k^i) \right| < \varepsilon$, $i=1,2,\dots,T$, откуда

следует выполнение соответствующего неравенства из определения 4.4 для $E\xi(t)$.

Определение функций $v(t)$ и $c(t, t+h)$ и доказательство требуемых неравенств для $D\xi(t)$, $\text{corr}(\xi(t), \xi(t+h))$ проводится аналогичным образом.

Утверждение доказано.

В следующих параграфах приведены примеры случайных асимптотически периодически коррелированных процессов.

4.3. Неоднородные марковские последовательности с матрицей переходных вероятностей, периодически зависящей от времени

Описание процесса

Рассмотрим бинарную марковскую цепь ξ_t , $t = 0, 1, 2, \dots$ с множеством значений $M = \{0, 1\}$, начальным вектором распределений A

$$A = (a_1, a_0) = (a, 1-a)$$

и матрицей переходных вероятностей $\Pi(t)$, являющейся периодической функцией дискретного аргумента t («времени»), принимающей значения

$$P_k = \begin{pmatrix} p_{11}[k] & p_{10}[k] \\ p_{01}[k] & p_{00}[k] \end{pmatrix} = \begin{pmatrix} p_k & 1-p_k \\ 1-q_k & q_k \end{pmatrix}$$

при $t = un + k > 0$, где $n \geq 2$ – период функции $\Pi(t)$, $u = 0, 1, 2, \dots$, $k = 0, 1, \dots, n-1$.

Здесь $p_{fg}[k]$ – это вероятность перехода из состояния f в состояние g ($f, g \in \{0, 1\}$) при $t = un + k$, a_f – начальная вероятность состояния f . Всюду далее будем считать, что матрицы P_k , $k = 0, 1, \dots, n-1$ не являются единичными.

С учётом определения марковской цепи и свойств условных вероятностей, задача определения одномерного распределения цепи ξ_t , $t = 0, 1, 2, \dots$ как функции

переменной t сводится к задаче вычисления произведений вида

$$\left(\prod_{i=0}^{n-1} P_i \right)^m, \quad \left(\prod_{i=0}^{n-1} P_i \right)^m \prod_{j=0}^l P_j, \quad j = 0, 1, \dots, n-2.$$

Далее будут использованы следующие обозначения:

$$d_k = \det \left(\prod_{i=0}^{k-1} P_i \right), \quad k = 1, 2, \dots, n,$$

$$t_k = \begin{cases} \det \left(\prod_{i=1}^{k-1} P_i \right), & k \geq 2, \\ 1, & k = 1, \end{cases}$$

$$c_k = \begin{cases} \left\{ \prod_{i=1}^{k-1} P_i \right\}_{11}, & k \geq 2, \\ 1, & k = 1, \end{cases}$$

$$b_k = \frac{c_k - q_0 t_k}{1 - d_k}.$$

Здесь и далее символом $\{M\}_{ij}$ обозначается элемент с номером ij матрицы M .

Одномерное распределение процесса ξ_t

В работах [32, 152] методом математической индукции доказаны леммы 4.8, 4.9 и утверждения 4.13, 4.14.

Лемма 4.8. В условиях данного параграфа, при $n \geq 2$ матрицу $\prod_{i=0}^{n-1} P_i$ можно

представить в виде

$$\prod_{i=0}^{n-1} P_i = S = \begin{pmatrix} b_n + (1 - b_n)d_n & 1 - b_n - (1 - b_n)d_n \\ b_n - b_n d_n & 1 - b_n + b_n d_n \end{pmatrix}.$$

Лемма 4.9. При $m \geq 1$ выполняется тождество

$$\left(\prod_{i=0}^{n-1} P_i \right)^m = S^m = \begin{pmatrix} b_n + (1 - b_n)d_n^m & 1 - b_n - (1 - b_n)d_n^m \\ b_n - b_n d_n^m & 1 - b_n + b_n d_n^m \end{pmatrix}.$$

Утверждение 4.13. При $t = mn$

$$\begin{aligned} P(\xi_t = 1) &= P(\xi_{mn} = 1) = p_{mn} = b_n + (a - b_n)d_n^m, \\ P(\xi_t = 0) &= P(\xi_{mn} = 0) = 1 - p_{mn} = 1 - b_n - (a - b_n)d_n^m. \end{aligned}$$

При $t = mn + k, 1 \leq k < n$

$$\begin{aligned} P(\xi_t = 1) &= P(\xi_{mn+k} = 1) = p_{mn+k} = b_n d_k + b_k - b_k d_k + (a - b_n) d_k d_n^m, \\ P(\xi_t = 0) &= P(\xi_{mn+k} = 0) = 1 - p_{mn+k} = 1 - b_n d_k - b_k + b_k d_k - (a - b_n) d_k d_n^m. \end{aligned}$$

Доказательство утверждения 4.13 основано на леммах 4.8, 4.9 и том, что

$$\begin{aligned} (P(\xi_{mn} = 1), P(\xi_{mn} = 0)) &= (p_{mn}, 1 - p_{mn}) = AS^m = \\ &= (a, 1 - a) \begin{pmatrix} b_n + (1 - b_n)d_n^m & 1 - b_n - (1 - b_n)d_n^m \\ b_n - b_n d_n^m & 1 - b_n + b_n d_n^m \end{pmatrix}, \\ (P(\xi_{mn+k} = 1), P(\xi_{mn+k} = 0)) &= (p_{mn+k}, 1 - p_{mn+k}) = AS^m P_0 \dots P_{k-1} = \\ &= (b_n + (a - b_n)d_n^m, 1 - b_n - (a - b_n)d_n^m) \left(\prod_{i=0}^{k-1} P_i \right). \end{aligned}$$

Утверждение 4.14. Справедливы равенства

$$\begin{aligned} f_0 &= \lim_{m \rightarrow \infty} p_{mn} = b_n, \\ f_k &= \lim_{m \rightarrow \infty} p_{mn+k} = b_n d_k + b_k - b_k d_k, \quad 1 \leq k < n. \end{aligned}$$

Корреляционная функция процесса ξ_t

Корреляционная функция $\text{corr}(\xi_t, \xi_{t+h})$ процесса ξ_t определяется равенством

$$\text{corr}(\xi_t, \xi_{t+h}) = \frac{P(\xi_t = 1, \xi_{t+h} = 1) - P(\xi_t = 1)P(\xi_{t+h} = 1)}{\sqrt{P(\xi_t = 1) - P^2(\xi_t = 1)} \sqrt{P(\xi_{t+h} = 1) - P^2(\xi_{t+h} = 1)}}.$$

В качестве примера вычисления значения функции с рассмотрим 2 случая:

$t = mn, h = un$ и $t = mn, h = un + v, 1 \leq v < n$. В первом случае,

$$\begin{aligned} P(\xi_{mn} = 1) &= b_n + (a - b_n)d_n^m, \\ P(\xi_{mn+un} = 1) &= P(\xi_{(m+u)n} = 1) = b_n + (a - b_n)d_n^{m+u}, \\ P(\xi_{mn} = 1, \xi_{mn+un} = 1) &= P(\xi_{mn} = 1)P(\xi_{mn+un} = 1 | \xi_{mn} = 1) = \\ &= P(\xi_{mn} = 1) \left\{ S^u \right\}_{11} = P(\xi_{mn} = 1) (b_n + (1 - b_n)d_n^u), \end{aligned}$$

$$\text{corr}(\xi_{mn}, \xi_{mn+un}) = \frac{\sqrt{(b_n + (a - b_n)d_n^m)(1 - b_n - (a - b_n)d_n^m)}}{\sqrt{(b_n + (a - b_n)d_n^{m+u})(1 - b_n - (a - b_n)d_n^{m+u})}} d_n^u.$$

Во втором случае справедливы равенства

$$\begin{aligned} P(\xi_{mn} = 1) &= b_n + (a - b_n)d_n^m, \\ P(\xi_{mn+un+v} = 1) &= P(\xi_{(m+u)n+v} = 1) = \\ &= b_n d_v + b_v - b_v d_v + (a - b_n)d_v d_n^{m+u}, \\ P(\xi_{mn} = 1, \xi_{mn+un+v} = 1) &= P(\xi_{mn} = 1)P(\xi_{mn+un+v} = 1 | \xi_{mn} = 1) = \\ &= P(\xi_{mn} = 1) \left\{ S^u P_0 \dots P_{v-1} \right\}_{11}, \\ \left\{ S^u P_0 \dots P_{v-1} \right\}_{11} &= (b_n + (1 - b_n)d_n^u)(b_v + (1 - b_v)d_v) + \\ &+ (1 - b_n - (1 - b_n)d_n^u)(b_v - b_v d_v) = b_n d_v + d_v d_n^u - b_n d_v d_n^u + b_v - b_v d_v. \end{aligned}$$

Следовательно,

$$\begin{aligned} \text{corr}(\xi_{mn}, \xi_{mn+un+v}) &= \\ &= \frac{\sqrt{(b_n + (a - b_n)d_n^m)(1 - b_n - (a - b_n)d_n^m)}}{\sqrt{(b_n d_v + b_v - b_v d_v + (a - b_n)d_v d_n^{m+u})(1 - b_n d_v - b_v + b_v d_v - (a - b_n)d_v d_n^{m+u})}} d_v d_n^u. \end{aligned}$$

При $t = mn + w$, $h = un$ и $t = mn + w$, $h = un + v$ вычисление функции $\text{corr}(\xi_t, \xi_{t+h})$ проводится аналогичным образом

Пример 4.3. На рисунках 4.12 и 4.13 представлены графики коэффициентов корреляции $\text{corr}(\xi_0, \xi_h)$ и $\text{corr}(\xi_t, \xi_{t+h})$ процесса ξ_t при

$$n = 2, A = (0.5, 0.5), P_0 = \begin{pmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{pmatrix}, P_1 = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}.$$

Несложно убедиться в том, что справедливо следующее утверждение.

Утверждение 4.15. Процесс ξ_t является асимптотически периодически коррелированным с периодом n .

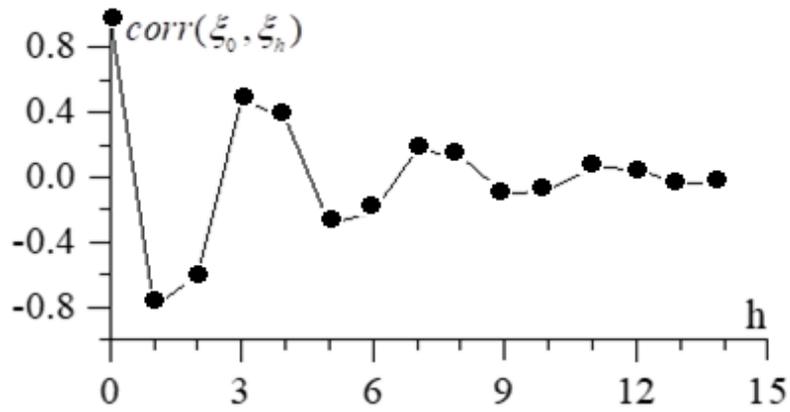


Рис. 4.12. Коэффициентов корреляции $corr(\xi_0, \xi_h)$.

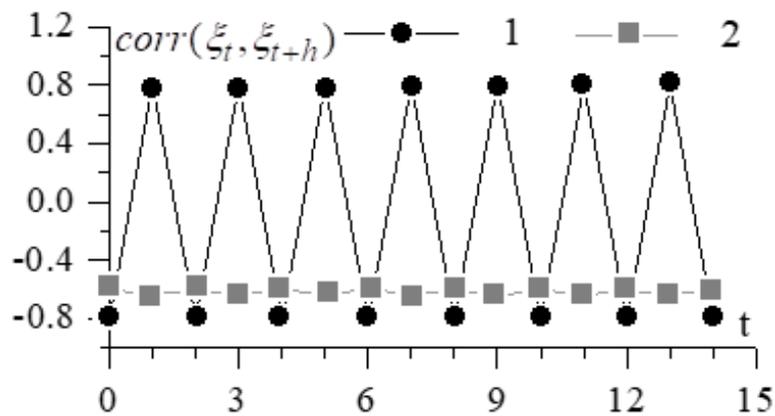


Рис. 4.13. Коэффициенты корреляции.: кривая 1: $corr(\xi_t, \xi_{t+1})$, кривая 2: $corr(\xi_t, \xi_{t+2})$.

Характеристики серий постоянных значений

По аналогии с работой [75], определим 1-серию длины k марковской последовательности ξ_t .

Определение 4.5. Если $\xi_{t-1} = 0, \xi_t = 1, \xi_{t+1} = 1, \dots, \xi_{t+k-1} = 1, \xi_{t+k} = 0$, то $\xi_t, \xi_{t+1}, \dots, \xi_{t+k-1}$ образуют 1-серию длины k , начинающуюся в момент $t > 0$.

Обозначим определенную таким образом 1-серию через $L(t, k)$. Тогда распределение длины 1-серии определяется равенством

$$P(L(t, k)) = \frac{P(\xi_{t-1} = 0, \xi_t = 1, \xi_{t+1} = 1, \dots, \xi_{t+k-1} = 1, \xi_{t+k} = 0)}{P(\xi_{t-1} = 0, \xi_t = 1)}.$$

Процесс ξ_t является марковским, следовательно,

$$\begin{aligned}
P(\xi_{t-1} = 0, \xi_t = 1) &= P(\xi_t = 1 | \xi_{t-1} = 0)P(\xi_{t-1} = 0), \\
P(\xi_{t-1} = 0, \xi_t = 1, \xi_{t+1} = 1, \dots, \xi_{t+k-1} = 1, \xi_{t+k} = 0) &= \\
= P(\xi_{t+k} = 0 | \xi_{t+k-1} = 1)P(\xi_{t+k-1} = 1 | \xi_{t+k-2} = 1) \dots P(\xi_{t+1} = 1 | \xi_t = 1) \times \\
&\times P(\xi_t = 1 | \xi_{t-1} = 0)P(\xi_{t-1} = 0), \\
P(L(t, k)) &= P(\xi_{t+k} = 0 | \xi_{t+k-1} = 1) \times \\
&\times P(\xi_{t+k-1} = 1 | \xi_{t+k-2} = 1) \times \dots \times P(\xi_{t+1} = 1 | \xi_t = 1).
\end{aligned}$$

Используя метод математической индукции легко проверить, что для различных t и k вероятность $P(L(t, k))$ выражается следующим образом (при $m \geq 0, u > 0$):

- при $t = mn, k = un$

$$P(L(t, k)) = \left(\prod_{i=0}^{n-2} p_i^u \right) p_{n-1}^{u-1} (1 - p_{n-1}),$$

- при $t = mn, k = un + v, 1 \leq v < n$

$$P(L(t, k)) = \left(\prod_{i=0}^{v-2} p_i^{u+1} \right) \left(\prod_{j=v-1}^{n-1} p_j^u \right) (1 - p_{v-1}),$$

- при $t = mn + w, k = un, 1 \leq w < n$

$$P(L(t, k)) = \left(\prod_{\substack{i=0 \\ i \neq w-1}}^{n-1} p_i^u \right) p_{w-1}^{u-1} (1 - p_{w-1}),$$

- при $t = mn + w, k = un + v, 1 \leq w < n, 1 \leq v < n, w + v \leq n$

$$P(L(t, k)) = \left(\prod_{i=0}^{w-1} p_i^{u-1} \right) \left(\prod_{j=w}^{n-1} p_j^u \right) \left(\prod_{l=0}^{w+v-2} p_l \right) (1 - p_{w+v-1}),$$

- при $t = mn + w, k = un + v, 1 \leq w < n, 1 \leq v < n, w + v > n$

$$P(L(t, k)) = \left(\prod_{i=0}^{w-1} p_i^u \right) \left(\prod_{j=w}^{n-1} p_j^{u+1} \right) \left(\prod_{l=0}^{w+v-n-2} p_l \right) (1 - p_{w+v-n-1}),$$

- при $t = mn, k = v, 1 \leq v < n$

$$P(L(t,k)) = (1 - p_{v-1}) \prod_{i=0}^{v-2} p_i,$$

- при $t = mn + w$, $k = v$, $w + v \leq n$

$$P(L(t,k)) = (1 - p_{w+v-1}) \prod_{i=w}^{w+v-2} p_i,$$

- при $t = mn + w$, $k = v$, $w + v > n$

$$P(L(t,k)) = (1 - p_{w+v-n-1}) \prod_{i=w}^{n-1} p_i \prod_{j=0}^{w+v-n-2} p_j.$$

Здесь и далее

$$\prod_{l=i}^j x_l = 1, \quad j < i.$$

В [32] доказано следующее утверждение.

Утверждение 4.16. Среднее значение $EL(t)$ и дисперсия $DL(t)$ длительности 1-серий, начавшихся в момент времени t есть периодические функции с периодом n , причём

$$\begin{aligned} EL(mn) &= \\ &= \sum_{v=1}^{n-1} v(1 - p_{v-1}) \prod_{i=0}^{v-2} p_i + n(1 - p_{n-1}) \left(\prod_{i=0}^{n-2} p_i \right) \frac{1}{\left(1 - \prod_{i=0}^{n-1} p_i \right)^2} + \\ &+ \sum_{v=1}^{n-1} (1 - p_{v-1}) \prod_{i=0}^{n-1} p_i \left(\prod_{i=0}^{v-2} p_i \right) \left\{ n \frac{1}{\left(1 - \prod_{i=0}^{n-1} p_i \right)^2} + v \frac{1}{\left(1 - \prod_{i=0}^{n-1} p_i \right)} \right\}, \\ EL(mn + w) &= n(1 - p_{w-1}) \left(\prod_{\substack{i=0 \\ i \neq w-1}}^{n-1} p_i \right) \frac{1}{\left(1 - \prod_{i=0}^{n-1} p_i \right)^2} + \end{aligned}$$

$$\begin{aligned}
& + \sum_{v=1}^{n-w} (1-p_{w+v-1}) \left(\prod_{j=w}^{n-1} p_j \right) \left(\prod_{l=0}^{w+v-2} p_l \right) \left[n \frac{1}{\left(1 - \prod_{i=0}^{n-1} p_i\right)^2} + v \frac{1}{\left(1 - \prod_{i=0}^{n-1} p_i\right)} \right] + \\
& + \sum_{v=n-w+1}^{n-1} (1-p_{w+v-n-1}) \left(\prod_{j=w}^{n-1} p_j \right) \left(\prod_{l=0}^{w+v-n-2} p_l \right) \left(\prod_{i=0}^{n-1} p_i \right) \times \\
& \times \left[n \frac{1}{\left(1 - \prod_{i=0}^{n-1} p_i\right)^2} + v \frac{1}{\left(1 - \prod_{i=0}^{n-1} p_i\right)} \right] + \\
& + \sum_{v=1}^{n-w} v (1-p_{w+v-1}) \prod_{j=w}^{w+v-2} p_j + \sum_{v=n-w+1}^{n-1} v (1-p_{w+v-n-1}) \left(\prod_{j=w}^{n-1} p_j \right) \left(\prod_{l=0}^{w+v-n-2} p_l \right).
\end{aligned}$$

Пример 4.4. На рисунке 4.14 приведены графики среднего $EL(t)$ и дисперсии $DL(t)$ при $n=3$, $p_0=0.9$, $p_1=0.5$, $p_2=0.1$.

Так же, как и для бинарных однородных цепей Маркова в работе [76], для рассматриваемых в данном параграфе марковских цепей ξ_t можно определить начальные, конечные и накрывающие 1-серии. Распределения таких серий, их моменты и особенности подробно описаны в [32]. В частности, показано, что средняя длительность и дисперсия начальных, конечных и накрывающих 1-серий являются осциллирующими функциями переменной t .

Марковские цепи рассмотренного в этом параграфе вида находят своё применение при решении практических задач. Так в работе [173] продемонстрирован пример использования неоднородных бинарных марковских цепей с матрицей переходов, являющейся периодической функцией, для моделирования и исследования индикаторных рядов температуры воздуха и векторных индикаторных рядов, характеризующих одновременное превышение температурой воздуха и его относительной влажностью заданных уровней. В следующем параграфе марковская цепь рассмотренного вида будет использована

при построении асимптотически периодически коррелированного процесса непрерывного времени.

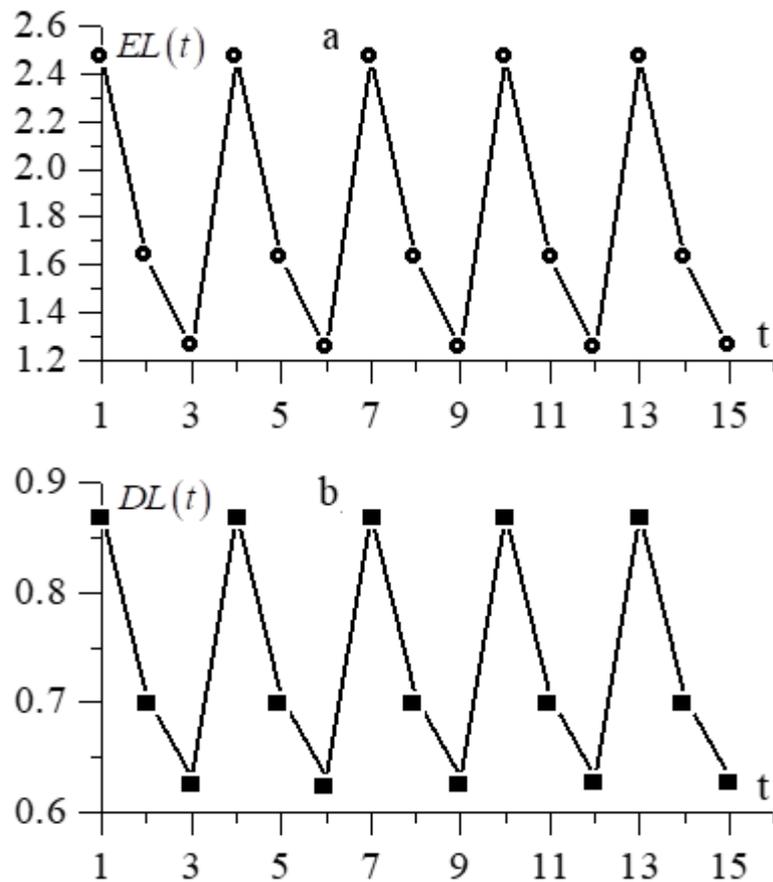


Рис. 4.14. Графики среднего и дисперсии длительности серий: график «а» – среднее $EL(t)$, график «б» – дисперсия $DL(t)$.

4.4. Модели асимптотически периодически коррелированных кусочно- постоянных процессов

В этом параграфе предложены 2 алгоритма конструктивного определения асимптотически периодически коррелированных кусочно-постоянных случайных процессов. В качестве основы для построения случайных процессов непрерывного

времени использованы неоднородная бинарная марковская цепь и неоднородный точечный поток, параметры которых периодически изменяются по времени.

4.4.1. Модель асимптотически периодически коррелированного кусочно-постоянного процесса на основе бинарной неоднородной цепи Маркова

В предыдущем параграфе были рассмотрены некоторые свойства неоднородной бинарной марковской цепи с матрицей переходных вероятностей, являющейся периодической функцией времени. Марковскую цепь такого типа можно использовать при моделировании асимптотически периодически коррелированного кусочно-постоянного процесса [134, 137].

Определение процесса

Пусть $\eta_i, i = 0, 1, 2, \dots$ – цепь Маркова, удовлетворяющая требованиям параграфа 4.3. Построим кусочно-постоянный процесс $\xi_t, t \geq 0$ непрерывного аргумента («времени») t согласно Алгоритму 4.3.

Алгоритм 4.3.

1. $\xi_0 = \alpha_0$;
2. Если $\eta_i = \eta_{i-1}, i \geq 1$, то $\xi_t = \xi_{i-1}, t \in (i-1, i]$;
3. Если $\eta_i \neq \eta_{i-1}, i \geq 1$, то $\xi_t = \alpha_i, t \in (i-1, i]$,

где $\alpha_0, \alpha_1, \dots$ – независимые между собой и от $\eta_i, i = 0, 1, 2, \dots$ случайные величины с функцией одномерного распределения $F(x)$, математическим ожиданием μ , дисперсией σ^2 и $|E\alpha_i^2| < +\infty$.

Пример 4.5. Пусть $F(x) = \Phi_{5,4}, n = 2$, а матрица переходных вероятностей принимает значения

$$P_0 = \begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \end{pmatrix}, P_1 = \begin{pmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{pmatrix}.$$

На рисунке 4.15 приведен пример реализации марковского процесса η_i , $i = 0, 1, 2, \dots$ и соответствующего ему кусочно-постоянного процесса ξ_t , $t \geq 0$. На рисунках 4.16 и 4.17 представлены оценки корреляционной функции $\text{corr}(\xi_t, \xi_{t+h})$ процесса ξ_t как функции времени t и сдвига h . Оценки проводились по 10^6 реализаций.

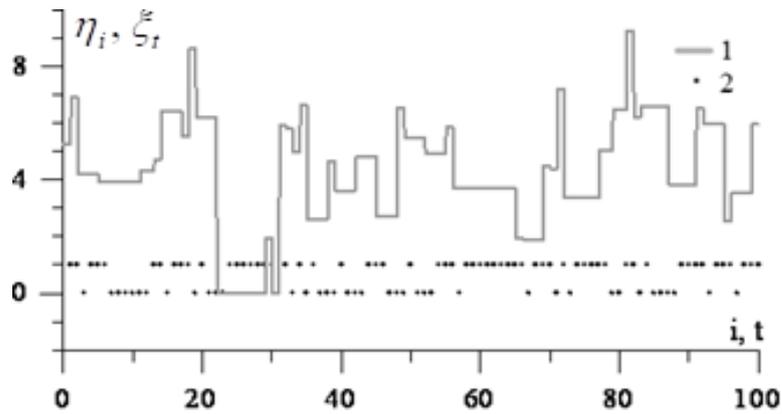


Рис. 4.15. Примеры реализаций. Кривая 1 – кусочно-постоянный процесс ξ_t , последовательность точек 2 – марковская цепь η_i .

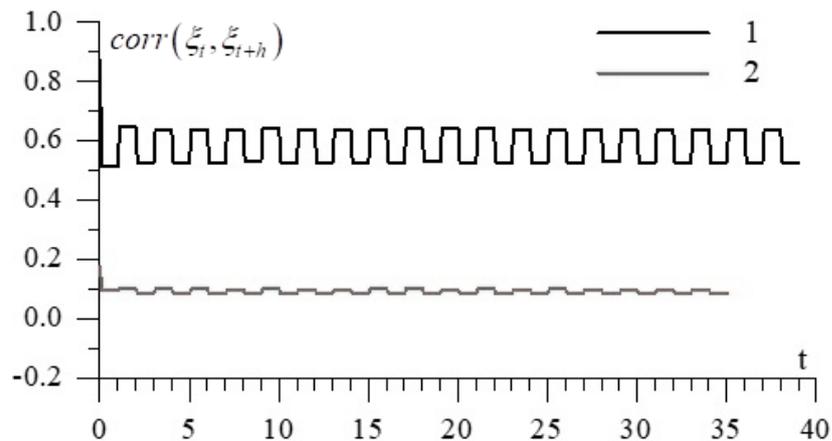


Рис. 4.16. Оценка корреляционной функции $\text{corr}(\xi_t, \xi_{t+h})$ процесса ξ_t при фиксированных значениях h . Кривая 1: $h = 1$, кривая 2: $h = 5$.

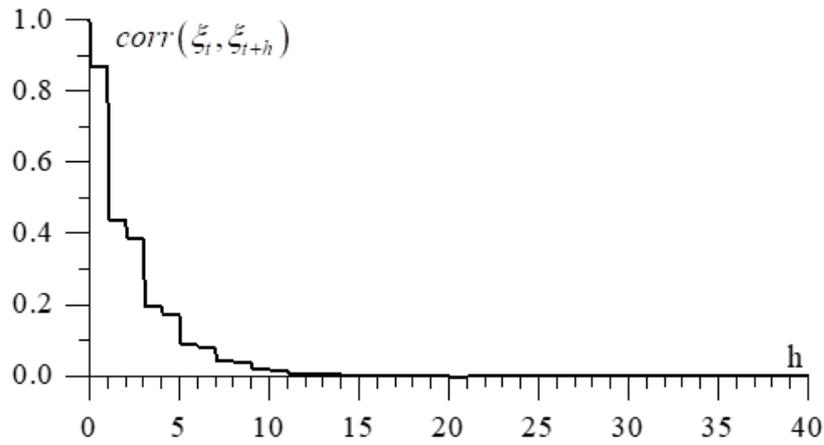


Рис. 4.17. Оценка корреляционной функции $\text{corr}(\xi_t, \xi_{t+h})$ процесса ξ_t при $t=0$.

Корреляционная функция процесса ξ_t

Исследуем корреляционную структуру процесса ξ_t . Так как ξ_t и ξ_{t+h} одинаково распределены, то

$$\text{corr}(\xi_t, \xi_{t+h}) = \frac{E\xi_t\xi_{t+h} - \mu^2}{\sigma^2}.$$

Обозначим

$$P_\eta(t, h) = \begin{cases} P(\eta_{[t]} = \dots = \eta_{[t+h]}), & t \in \mathbb{Z}, t+h \in \mathbb{Z}, \\ P(\eta_{[t]} = \dots = \eta_{[t+h]+1}), & t \in \mathbb{Z}, t+h \notin \mathbb{Z}, \\ P(\eta_{[t]+1} = \dots = \eta_{[t+h]}), & t \notin \mathbb{Z}, t+h \in \mathbb{Z}, \\ P(\eta_{[t]+1} = \dots = \eta_{[t+h]+1}), & t \notin \mathbb{Z}, t+h \notin \mathbb{Z}. \end{cases}$$

Верны равенства

$$E\xi_t\xi_{t+h} = P_\eta(t, h)E\xi_t^2 + (1 - P_\eta(t, h))(E\xi_t)^2,$$

$$\text{corr}(\xi_t, \xi_{t+h}) = \frac{P_\eta(t, h)E\xi_t^2 + (1 - P_\eta(t, h))\mu^2 - \mu^2}{\sigma^2} = P_\eta(t, h).$$

Очевидно, что функция $\text{corr}(\xi_t, \xi_{t+h})$ как функция одного аргумента (t или h) является кусочно-постоянной. Благодаря марковости процесса $\eta_i, i=0,1,2,\dots$

несложно вычислить вероятности $P_\eta(t, h)$. В качестве примера, найдём вероятность

$P(\eta_{[t]} = \dots = \eta_{[t+h]})$ в случае $[t] = mn$, $[t+h] = (m+u)n$:

$$\begin{aligned} & P(\eta_{[t]} = \dots = \eta_{[t+h]}) = \\ & = P(\eta_{[t]} = 1, \eta_{[t]+1} = 1, \dots, \eta_{[t+h]} = 1) + P(\eta_{[t]} = 0, \eta_{[t]+1} = 0, \dots, \eta_{[t+h]} = 0) = \\ & = P(\eta_{[t+h]} = 1 | \eta_{[t+h]-1} = 1) P(\eta_{[t+h]-1} = 1 | \eta_{[t+h]-2} = 1) \times \\ & \quad \times \dots \times P(\eta_{[t]+1} = 1 | \eta_{[t]} = 1) P(\eta_{[t]} = 1) + \\ & + P(\eta_{[t+h]} = 0 | \eta_{[t+h]-1} = 0) P(\eta_{[t+h]-1} = 0 | \eta_{[t+h]-2} = 0) \times \\ & \quad \times \dots \times P(\eta_{[t]+1} = 0 | \eta_{[t]} = 0) P(\eta_{[t]} = 0) = \\ & = (b_n + (a - b_n)d_n^m) \prod_{k=1}^n p_k^u + (1 - b_n - (a - b_n)d_n^m) \prod_{k=1}^n q_k^u. \end{aligned}$$

Из того, что функция одномерного распределения процесса ξ_t есть $F(x)$ (она не зависит от t), и вида функции $\text{corr}(\xi_t, \xi_{t+h})$ следует справедливость утверждения 4.17.

Утверждение 4.17. Процесс ξ_t является асимптотически периодически коррелированным.

4.4.2. Модель асимптотически периодически коррелированного кусочно-постоянного процесса на основе нерегулярного точечного потока

В этой параграфе рассмотрен ещё один алгоритм моделирования асимптотически периодически коррелированного кусочно-постоянного процесса [30, 178]. Этот алгоритм основан на использовании процедуры кусочно-постоянного восполнения неоднородных точечных потоков. Схожий алгоритм был предложен Михайловым Г.А. для моделирования стационарных процессов на основе пуассоновских потоков [52].

Пусть $\lambda(t) > 0$ – ограниченная по модулю периодическая функция с периодом T . Определим точечный поток $\{\tau_k\}$:

$$\tau_0 = 0, \quad \tau_k = \sum_{i=1}^k \eta_i, \quad k = 1, 2, \dots,$$

где η_i – независимые неотрицательные случайные величины, распределенные с плотностями

$$f_i(x) = \lambda(\tau_{i-1}) e^{-\lambda(\tau_{i-1})x}.$$

Нестационарный кусочно-постоянный процесс $\xi_t, t \geq 0$ построим согласно Алгоритму 4.4.

Алгоритм 4.4.

1. Моделируется последовательность $\{\tau_k\}$.
2. Для каждого интервала $[\tau_{i-1}; \tau_i), i = 1, 2, \dots$ полагаем $\xi_t = \nu_i$, где ν_i – независимые случайные величины с заданной функцией одномерного распределения $F(x)$ и конечным вторым моментом.

Пример 4.6. На рисунках 4.18 и 4.19 приведены графики оценок корреляционной функции $\text{corr}(\xi_t, \xi_{t+h})$ как функции одной переменной при фиксированном значении другой переменной при $F(x) = \Phi_{0,1}$, $\lambda(t) = 0.9 \sin 2t + 1$.

Из построения процесса ξ_t следует, что его одномерное распределение не зависит от t и определяется функцией распределения $F(x)$. Очевидно, что при $t = 0$ выполняется равенство

$$\text{corr}(\xi_0, \xi_{0+h}) = P(\tau_1 > h) = P(\eta_1 > h) = \exp(-\lambda(0)h).$$

Опишем аналитически корреляционную функцию $\text{corr}(\xi_t, \xi_{t+h})$ построенного процесса ξ_t при $t > 0, h > 0$. Пусть μ, σ^2 – математическое ожидание и дисперсия случайной величины с функцией одномерного распределения $F(x)$. Поскольку

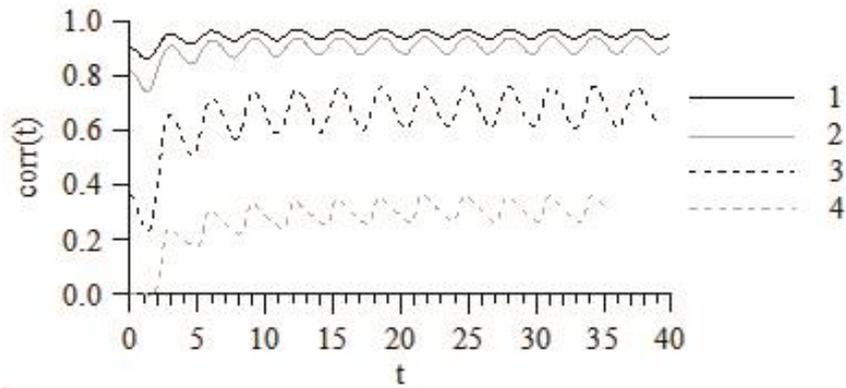


Рис. 4.18. Оценка корреляционная функция $corr(\xi_t, \xi_{t+h})$ процесса ξ_t при фиксированных значениях h . Кривая 1: $h = 0.1$; кривая 2: $h = 0.2$; кривая 3: $h = 1$; кривая 4: $h = 5$.

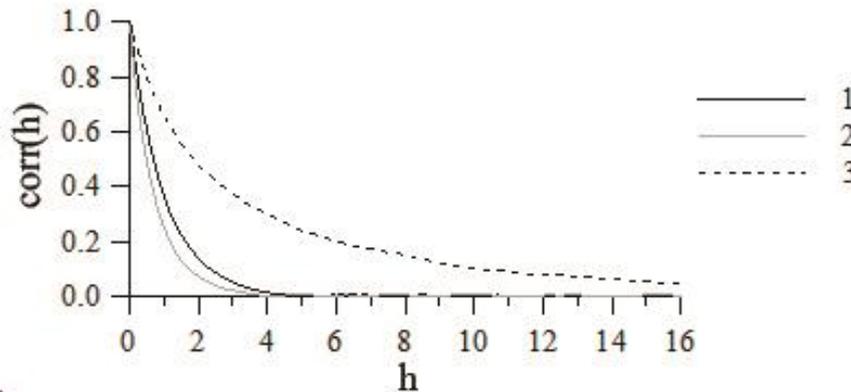


Рис. 4.19. Оценка корреляционная функция $corr(\xi_t, \xi_{t+h})$ процесса ξ_t при фиксированных значениях t . Кривая 1: $t = 0$; кривая 2: $t = 1$; кривая 3: $t = 3$.

$$E\xi_t\xi_{t+h} = p\xi_t^2 + (1-p)\xi_t\xi_{t+h} = pE\xi_t^2 + (1-p)E\xi_tE\xi_{t+h} = p\sigma^2 + \mu^2,$$

то

$$corr(\xi_t, \xi_{t+h}) = \frac{E\xi_t\xi_{t+h} - E\xi_tE\xi_{t+h}}{\sqrt{D\xi_t}\sqrt{D\xi_{t+h}}} = p,$$

где

$$\begin{aligned} p &= P(\exists k : \tau_k \leq t, t+h < \tau_{k+1}) = \sum_{k=0}^{\infty} P(\tau_k \leq t, t+h < \tau_{k+1}) = \\ &= P(\tau_0 \leq t, t+h < \tau_1) + \sum_{k=1}^{\infty} P(\tau_k \leq t, t+h < \tau_{k+1}) = \end{aligned}$$

$$\begin{aligned}
&= P(t+h < \eta_1) + \sum_{k=1}^{\infty} P(\tau_k \leq t, t+h < \tau_{k+1}) = \\
&= \exp(-\lambda(0)[t+h]) + \sum_{k=1}^{\infty} P(\tau_k \leq t, t+h < \tau_{k+1}).
\end{aligned}$$

Используя формулу полной вероятности, можно получить, что для $k \geq 1$ выполняется цепочка равенств

$$\begin{aligned}
P(\tau_k \leq t, t+h < \tau_{k+1}) &= P\left(\sum_{i=1}^k \eta_i \leq t, \sum_{i=1}^{k+1} \eta_i > t+h\right) = \\
&= \int_0^t \left(P\left(\sum_{i=2}^k \eta_i \leq t-t_1, \sum_{i=2}^{k+1} \eta_i > t+h-t_1 \mid \eta_1 = t_1\right) \times \lambda(0) \exp(-\lambda(0)t_1) \right) dt_1 = \\
&= \int_0^t \int_0^{t-t_1} \left(P\left(\sum_{i=3}^k \eta_i \leq t-t_1-t_2, \sum_{i=3}^{k+1} \eta_i > t+h-t_1-t_2 \mid \eta_1 = t_1, \eta_2 = t_2\right) \times \right. \\
&\quad \left. \times \lambda(0) \exp(-\lambda(0)t_1) \lambda(t_1) \exp(-\lambda(t_1)t_2) \right) dt_2 dt_1 = \\
&= \int_0^t \dots \int_0^{t-t_1-\dots-t_{k-1}} \left(\lambda(0) \lambda(t_1) \dots \lambda(t_{k-1}) \times \right. \\
&\quad \left. \times \exp(-\lambda(0)t_1 - \dots - \lambda(t_{k-1})t_k - \lambda(t_k)(t+h-t_1-\dots-t_k)) \right) dt_k \dots dt_1.
\end{aligned}$$

Численные эксперименты показывают, то при различных функциях $\lambda(t)$ процесс ξ_t является асимптотически периодически коррелированным. Численные эксперименты также показывают, что свойство асимптотической периодической коррелированности процесса ξ_t сохраняется при замене экспоненциальных плотностей $f_i(x)$ на широкое семейство плотностей $g_i(x, \lambda(t))$, $x > 0$ с периодическим параметром $\lambda(t)$.

Заключение

Ниже приведены **основные результаты выполненного исследования:**

- разработаны, численно реализованы и верифицированы стохастические «генераторы погоды», позволяющие численно моделировать временные ряды и пространственно-временные поля комплексов метеорологических параметров и их специальных характеристик с учётом временной нестационарности (в том числе, суточного хода) и пространственной неоднородности реальных метеопроцессов;
- разработаны, численно реализованы и верифицированы стохастические модели временных рядов, пространственных и пространственно-временных полей биоклиматических индексов, основанные как на применении «генераторов погоды», так и на непосредственном использовании определяющих формул. Разработанные модели позволяющие учитывать и воспроизводить нестационарность по времени и неоднородность по пространству реальных процессов;
- разработаны алгоритмы моделирования некоторых классов периодически коррелированных и асимптотически периодически коррелированных кусочно-постоянных и кусочно-линейных случайных процессов, алгоритмы моделирования условных негауссовских процессов с точечными и интервальными условиями, алгоритмы моделирования двумерных случайных полей с корреляционными функциями специального вида и метод стохастической интерполяции случайных процессов.

Рекомендации и перспективы дальнейшей разработки темы

Построенные стохастические «генераторы погоды» могут быть в дальнейшем использованы для решения достаточно широкого класса научных задач. Так, например, если преобразовать представленную в параграфе 2.2 Главы 2

модель пространственно-временного поля метеопараметров на сети метеостанций в модель поля на регулярной сетке, то полученная модель может быть объединена со стохастической или динамической моделью стока рек, бассейны которых находятся на рассматриваемой территории. Такая объединённая модель позволит исследовать зависимость речного стока от метеорологических условий, в том числе при экстремальных режимах выпадения осадков. В то же время, дальнейшая параметризация моделей временных рядов метеопараметров, представленных в параграфе 2.1 Главы 2, даст возможность проводить исследования различных сценариев изменения климата. На основе модельных траекторий можно будет оценивать различные характеристики комплексов метеорядов в рассматриваемых климатических условиях.

Предложенные в Главе 3 стохастические модели условных и безусловных временных рядов, пространственных и пространственно-временных полей биоклиматических индексов могут быть использованы при исследовании различных характеристик волн холода или жары, в том числе их длительности и скорости продвижения по территории. Кроме этого, рассмотренные модели могут быть применены для районирования территорий по признакам, связанным с вероятностями возникновения экстремальных и аномальных биоклиматических явлений (в реальных данных, полученных по метеонаблюдениям, может не содержаться информации о таких явлениях, однако она может быть получена по модельным реализациям). Научный интерес представляет также расширение построенных моделей для исследования связей между погодными условиями и распространением и обострением различных заболеваний человека. Для проведения таких исследований необходимо объединить метеорологические и медицинские данные, и построить совместную модель биоклиматических процессов и процессов, описывающих динамику заболеваемости.

Список литературы

1. Адаменко В.Н., Хайруллин К.Ш. Проблемы биоклиматической оценки суровости погоды и мелиорация микроклимата застройки // Труды ГГО им. А.И. Воейкова, 1973, вып. 306, – С. 3-18.
2. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физматлит, 1963, 500 с.
3. Андерсон Т. Статистический анализ временных рядов. – М.: МИР, 1976, 755 с.
4. Анисимова А.В. Численное моделирование индикаторных случайных полей жидких осадков // Труды конференции молодых учёных ИВМиМГ СО РАН, Новосибирск: ИВМиМГ СО РАН, 1997, – С. 3-15.
5. Бабичева Г.А., **Каргаполова Н.А.** Исследование двух алгоритмов моделирования однородных случайных полей с четырехпараметрической экспоненциальной корреляционной функцией // Труды международной научной конференции «Марчуковские научные чтения – 2017», Новосибирск, 25 июня – 14 июля 2017, – С. 78-82.
6. Бабичева Г.А., **Каргаполова Н.А.**, Огородников В.А. Рандомизированный алгоритм моделирования однородных случайных полей с невыпуклыми корреляционными функциями. Труды международной конференции «Актуальные проблемы вычислительной и прикладной математики», Новосибирск, 2015, – С. 91-95.
7. Большая советская энциклопедия. В 30-ти т. – 3-е изд. – М.: Совет. энцикл., 1969 – 1986.
8. Булыгина О.Н., Веселов В.М., Разуваев В.Н., Александрова Т.М. Описание массива срочных данных об основных метеорологических параметрах на станциях России // Свидетельство о государственной регистрации базы

данных № 2014620549. <http://meteo.ru/data/163-basic-parameters#описание-массива-данных>

9. Гавриловская Н.В., Топаж А.Г., Хворова Л.А. Моделирование погодных сценариев для оценки урожайности зерновых культур в условиях Западной Сибири // Известия АлтГУ. – 2011. – №1-1. – С. 71-77.
10. Гандин Л.С. Объективный анализ метеорологических полей. – Л.: Гидрометеорологическое изд-во, 1963, 287 с.
11. Гандин Л.С., Каган Р.Л. Статистические методы интерпретации метеорологических данных. – Л: Гидрометеоиздат, 1976, 360 с.
12. Гельфан А.Н. Динамико-стохастическое моделирование формирования талого стока. – М.: Наука, 2007, 279 с.
13. Гельфан А.Н., Морейдо В.М. Динамико-стохастическое моделирование формирования снежного покрова на Европейской территории России // Лёд и Снег. – 2014. – Т. 54, № 2. – С. 44-52.
14. Гладышев Е.Г. Периодически и почти-периодически коррелированные случайные процессы с непрерывным временем // Теория вероятн. и ее примен. – 1963. – Т. 8, вып. 2. – С. 184-189.
15. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 2003, 479 с.
16. Головина Е.Г., Трубина М.А. Методика расчетов биометеорологических параметров (индексов). – С.Пб., 1997, 110 с.
17. Гомбоева Н.Г. Климато-рекреационные ресурсы Забайкалья и их влияние на здоровье населения: автореф. дис. ... канд. биол. наук: 14.00.17 / Гомбоева Нина Гындуновна. – М., 1997. – 16 с.
18. ГОСТ Р ИСО 11079-2015 Эргономика термальной среды. Определение холодового стресса и его интерпретация на основе показателей требуемой термоизоляции одежды и локального охлаждающего воздействия. Национальный стандарт Российской Федерации. – М.: Стандартинформ. 2015. – 39 с.

19. Демьянов В.В., Савельева Е.А. Геостатистика: теория и практика. – М.: Наука, 2010, 327 с.
20. Деренок К.В. Численное моделирование сильных и длительных понижений температуры // Вычислительные технологии. – 2008. – Т.13, спец. вып. 4. – С. 27-34.
21. Драган Я.П., Рожков В.А., Яворский И.Н. Методы вероятностного анализа ритмики океанологических процессов. – Л.: Гидрометеиздат, 1987, 319 с.
22. Емелина С.В. Прогноз погодных условий, неблагоприятных для населения с сердечно-сосудистыми и аллергическими заболеваниями: дис. ... канд. геогр. наук: 25.00.30 / Емелина Светлана Валерьевна. – М., 2019. – 122 с.
23. Ермаков С.М., Михайлов Г.А. Статистическое моделирование. – М.: Наука, 1982, 296 с.
24. Зайкин Д.А., Крестьянцев А.Б., Таровик О.В., Топаж А.Г. Имитационная модель морской транспортно-технологической системы платформы «Приразломная» // ПРОНЕФТЬ. Профессионально о нефти. – 2017, № 2(4). – С. 61-68.
25. Иошпа А.Р. Особенности биоклиматического режима региона на примере Астраханской области: дис. ... канд. географ. наук: 25.00.30 / Иошпа Александр Рувимович. – С.-Пб., 2005. – 148 с.
26. Исаев А.А. Экологическая климатология. – М.: Науч. мир, 2003. – 470 с.
27. Каган Р.Л., Федорченко Е.И. К вопросу о статистическом моделировании двумерных метеорологических полей // Труды ГГО им. А.И. Воейкова. 1973, вып. 308. – С. 20-26.
28. **Каргаполова Н.А., Огородников В.А.** Вероятностная модель нестационарных совместных временных рядов приземной температуры, относительной влажности воздуха и атмосферного давления // Труды Труды международной научной конференции «Марчуковские научные чтения – 2018», 8-12 октября, 2018, Новосибирск, Россия. – С. 177-182.
29. **Каргаполова Н.А., Огородников В.А.** Моделирование комплекса метеорологических параметров с учетом их годовой нестационарности //

- Труды международной научной конференции «Марчуковские научные чтения – 2017», 25 июня – 14 июля 2017. – Новосибирск, 2017. – С. 388-393.
30. **Каргаполова Н.А.** Об одном алгоритме моделирования асимптотически периодически коррелированных случайных процессов на основе неоднородных пуассоновских точечных потоков // Труды конференции молодых ученых ИВМиМГ СО РАН, Новосибирск, 11 апреля 2016 года, 2017. – С. 43-48.
31. **Каргаполова Н.А.** Об одном алгоритме моделирования случайных процессов с периодическими свойствами // Сборник статей IX Международной научно-технической конференции «Аналитические и численные методы моделирования естественнонаучных и социальных проблем», Пенза, 2014. – С. 20-25.
32. **Каргаполова Н.А.** Численное моделирование и исследование нестационарных случайных процессов с периодическими характеристиками: дис. ... канд. физ.-мат. наук: 01.01.07 / Каргаполова Нина Александровна. – Новосибирск, 2013. – 95 с.
33. **Каргаполова Н.А.,** Огородников В.А. Статистическая структура совместных рядов индикаторов выпадения осадков, суточной минимальной и максимальной приземной температуры воздуха // Интерэкспо Гео-Сибирь. – 2017. – Т. 4, № 1. – С. 111-115.
34. Кондратьев С.А., Шмакова М.В. Детерминированно-стохастическое моделирование массопереноса в системе водосбор-водоем // Фундаментальная и прикладная гидрофизика. – 2018. – Т. 11, № 3. – С. 55-65.
35. Кондратьев С.А., Шмакова М.В. Математическое моделирование стока реки Невы в условиях возможного изменения климата // Уч. зап. РГГМУ. – 2016. – № 42. – С. 24–32.
36. Королев В.Ю. EM-алгоритм, его модификации и их применение к задаче разделения смесей вероятностных распределений. Теоретический обзор. – М.: ИПИ РАН, 2007, 94 с.

37. Кузякина М.В., Гура Д.А. Оценка комфортности биоклиматических условий Краснодарского края с применением ГИС-технологий / Юг России: экология, развитие. – 2020. Т. 15, № 3. – С. 66-76.
38. Май Р.И., Таровик О.В., Топаж А.Г. Моделирование морской погоды как входного сигнала имитационных моделей транспортных и экологических систем в Арктическом регионе // Проблемы экологического мониторинга и моделирования экосистем. – 2018. – Т. XXIX, № 3. – С. 20-38.
39. Марченко А.С. Аппроксимация эмпирического распределения вероятностей суточных сумм жидких осадков // Труды ЗапСибНИИ Госкомгидромета. – 1989. – вып. 86. – С. 66- 74.
40. Марченко А.С., Минакова Л.А. Вероятностная модель временных рядов температуры воздуха // Метеорология и гидрология. – 1980. – № 9. – С. 39-47.
41. Марченко А.С., Огородников В.А. Авторегрессионные процессы с заданной корреляционной структурой // Известия вузов: Математика. – 1985. – № 7. – С. 63-67.
42. Марченко А.С., Огородников В.А. Вероятностные модели последовательности сухих и дождливых суток. – Новосибирск, 1991, 22 с. (Препринт АН СССР. Сиб. отд-ние. ВЦ, № 933).
43. Марченко А.С., Огородников В.А. Моделирование стационарных гауссовских последовательностей большой длины с произвольной корреляционной функцией // Журн. вычисл. математики и матем. физики. – 1984. – Т. 24, № 10. – С. 1514-1519.
44. Марченко А.С., Романенко Т.П. Моделирование гамма-последовательностей и их использование для изучения выбросов скорости ветра // Метеорология и гидрология. – 1975. – № 7. – С. 54-62.
45. Марченко А.С., Семочкин А.Г. Изучение выбросов относительной влажности воздуха путем статистического моделирования бета-последовательностей // Труды ГГО им. А.И. Воейкова. –1977. – № 397. – С. 35-43.

46. Марченко А.С., Сёмочкин А.Г. ГФФФ- метод моделирования временных рядов по наблюдаемым реализациям // Сб. научных трудов «Численные методы статистического моделирования». – Новосибирск: Изд. ВЦ СО АН СССР, 1987. – С 14-22.
47. Марченко А.С., Сёмочкин А.Г. Модели одномерных и совместных распределений случайных неотрицательных величин // Метеорология и гидрология. – 1982. – № 3. – С. 52–56.
48. Марчук Г.И., Михайлов Г.А., Назаралиев М.А., Дарбинян Р.А., Елепов Б.С., Каргин Б.А. Метод Монте-Карло в атмосферной оптике. – Новосибирск: Наука, Сиб. отд-ние, 1976, 283 с.
49. Миннихметов И.Р. Стохастическое моделирование условных гауссовских процессов // Препринты ИПМ им. М.В.Келдыша. М. – 2011. – № 79.
50. Михайлов Г.А. Моделирование случайных процессов и полей на основе точечных потоков Пальма // Докл. АН СССР. – 1982. – Т. 3, № 3. – С. 531-535.
51. Михайлов Г.А. О методе «повторений» для моделирования случайных векторов и процессов (рандомизация корреляционных матриц) // Теория вероятностей и её применения. – 1974. –Т. 19, № 4. – С. 873-878.
52. Михайлов Г.А. Оптимизация весовых методов Монте-Карло. – М: Наука, 1987, 239 с.
53. Михайлов Г.А., Войтишек А.В. Численное статистическое моделирование. Методы Монте-Карло. – М.: Изд. центр «Академия», 2006, 368 с.
54. Михеев М.А., Михеева И.М. Основы теплопередачи. М.: Энергия, 1977, 343 с.
55. Ницис В.Э. Использование климатической информации при проектировании систем вентиляции и систем кондиционирования воздуха // Труды ГГО им. А.И. Воейкова. – 1983. – вып. 475. – С. 3-7.
56. Огородников В.А. Кусочно-линейная аппроксимация дискретной корреляционной функции // Сб. научных трудов «Теория и приложения статистического моделирования». Новосибирск. – 1992. – С. 20-25.

57. Огородников В.А. “Моделирование одного класса изотропных гауссовских полей” // Сб. научных трудов «Теория и приложения статистического моделирования». Новосибирск. – 1988. – С. 25-30.
58. Огородников В.А. О динамико-вероятностном прогнозе// Изв. АН СССР. Физика атмосферы и океана. – 1975. – Т. 11, № 8. – С. 851-853.
59. Огородников В.А., Деренок К.В., Толстых У.И. Специальные численные модели дискретных случайных рядов. – Новосибирск, 2009, 30 с. (Препринт РАН. Сиб. отд-ние. ИВМиМГ, № 1166).
60. Огородников В.А., Сересева О.В. Мультипликативная численная стохастическая модель полей суточных сумм жидких осадков и ее использование для оценки статистических характеристик экстремальных режимов их выпадения // Оптика атмосферы и океана. – 2015. – Т. 28, № 3. – С. 238-245.
61. Павлова В.Н. Продуктивность зерновых культур в России при изменении агроклиматических ресурсов в 20-21 веках: дис. ...докт. географ. наук: 25.00.30 / Павлова Вера Николаевна. –М., 2021, 271 с.
62. Палагин Ю.И., Федотов С.В., Шалыгин А.С. Параметрические модели для статистического моделирования векторных неоднородных случайных полей // Автоматика и телемеханика. – 1990. – № 6. – С. 79-89.
63. Переведенцев Ю.П., Шумихина А.В. Динамика биоклиматических показателей комфортности природной среды в Удмуртской Республике // Ученые записки Казанского университета. Серия Естественные науки. – 2016. – Т. 158, №. 4. – С. 531-547.
64. Пиранашвили З.А. Некоторые вопросы статистико-вероятностного моделирования случайных процессов // В кн. Вопросы исследования операций. – Тбилиси: Менциереба, 1966. С. 53-91.
65. Поляк И.И. Методы анализа случайных процессов и полей в климатологии. – Л.: Гидрометеиздат, 1979, 255 с.
66. Пригарин С.М. Методы численного моделирования случайных процессов и полей. – Новосибирск: ИВМиМГ СО РАН, 2005, 259с.

67. Пригарин С.М. Численное моделирование многомерных гауссовских распределений: учебное пособие. Новосибирск: ИПЦ НГУ, 2018, 82 с.
68. Разуваев В.Н., Финаев А.Ф. Оценка биоклиматического индекса суровости климатического режима на территории России // Тр. Всероссийского научно-исследовательского института гидрометеорологической информации – Мирового центра данных. – 2019. – № 185. – С. 18-39.
69. Рожков В.А. Методы вероятностного анализа океанологических процессов. – Л.: Гидрометеиздат, 1979, 280 с.
70. Рожков В.А., Трапезников Ю.А. Вероятностные модели океанологических процессов. – Л.: Гидрометеиздат, 1990, 272 с.
71. Рожков В.А., Трапезников Ю.А. К вопросу о построении моделей океанологических процессов // Труды ГОИН. – 1983. – № 169. – С. 46-59.
72. Руководство по специализированному климатологическому обслуживанию экономики. Под ред. Кобышевой Н.В. СПб.: Астерион, 2008, 336 с.
73. Румянцева С.А. Вероятностное моделирование ветрового волнения как полимодулированного полициклического случайного процесса: автореф. дис. ... канд. физ.-мат. наук: 11.00.08 / Румянцева Светлана Александровна. – С.-Пб., 1993, 24 с.
74. Русанов В.И. Комплексные метеорологические показатели и методы оценки климата для медицинских целей: Учебное пособие. – Томск: Издательство Томского университета, 1981, 86 с.
75. Савельев Л.Я., Балакин С.В. Совместное распределение числа единиц и числа 1-серий в двоичных марковских последовательностях // Дискрет. матем. – 2004. – Т. 16, № 3. – С. 43-62.
76. Савельев Л.Я., Балакин С.В., Хромов Б.В. Накрывающие серии в двоичных марковских последовательностях // Дискрет. матем. – 2003. – Т. 15, № 1. – С. 50-76.
77. Сванидзе Г.Г. Математическое моделирование гидрологических рядов. – Л.: Гидрометеиздат, 1977, 296 с.

78. Семенов Ю.В. Системы кондиционирования воздуха с поверхностными воздухоохладителями. М.: Техносфера, 2014, 272 с.
79. Сёмочкин А.Г. Функции распределения модуля и угла направления вектора ветра // Метеорология и гидрология. – 2000. – №6. – С. 22-28.
80. Смирнов Н.В., Большев Л.Н. Таблицы для вычислений функции двумерного нормального распределения. – М.: Изд. АН СССР, 1962, 204 с.
81. Соболев И.М. Численные методы Монте-Карло. – М.: Наука, 1973, 311 с.
82. Статистическая структура метеорологических полей. Сб. статей под ред. Гандин Л.С., Захариев В.И., Целнай Р. – Будапешт, 1976, 364 с.
83. Ткачук С.В. Обзор индексов степени комфортности погодных условий и их связь с показателями смертности // Труды Гидрометцентра России. – 2012. – Вып. 347. – С.194-214.
84. Ткачук С.В., Рубинштейн К.Г. Сравнительный анализ биоклиматических индексов для прогноза с использованием региональной модели // Избранные труды Международной молодежной школы и конференции CITES2011. Томск: Издательство Томского ЦНТИ. – 2011. – С.151-155.
85. Товстик Т.М. Моделирование векторного марковского процесса с произвольным одномерным распределением // Вестник ЛГУ. – 1985. – №8. – С. 10-120.
86. Топаж А.Г. Моделирование суточных метеоданных как входного сигнала модели продукционного процесса // В сб. науч. тр. «Почва и растение – процессы и модели». – С.-Пб.: АФИ, – 1992. – С. 79-86.
87. Фаддеев Д.К., Фадеева В.Н. Вычислительные методы линейной алгебры. С.-Пб.: Лань, 2009, 735 с.
88. Фихтенгольц Г.М. Курс дифференциального и интегрального исчисления. В 3-х т. – М.: Наука, 1970, 659 с.
89. Шлычков В.А., Огородников В.А., Сересева О.В. Совместная численная стохастическая модель временных рядов суточного стока реки и пространственно-временных полей суточных сумм жидких осадков // Интерэкспо Гео-Сибирь. – 2015. – Т. 4. – С. 145-149.

90. ЯндексПогода [Электронный ресурс]. <https://yandex.ru/pogoda/docs/glossary.html#feels-like> (последнее посещение: 12.12.2021)
91. Accuweather [Электронный ресурс]. What is the AccuWeather RealFeel Temperature? <https://www.accuweather.com/en/weather-news/what-is-the-accuweather-realfeel-temperature/156655> (последнее посещение: 14.09.2020)
92. Ailliot P., Allard D., Monbet V., Naveau P. Stochastic weather generators: an overview of weather type models // Journal de la société française de statistique. – 2015. – Tome 156, № 1. – P. 101-113.
93. Ailliot P., Bessac J., Monbet V., Pene F. Non-homogeneous hidden Markovswitching models for wind time series // Journal of Statistical Planning and Inference. – 2015. – 160. – P.75-88.
94. Akenteva M.S., **Kargapolova N.A.**, Ogorodnikov V.A. Numerical study of the bioclimatic index of severity of climatic regime based on a stochastic model of the joint meteorological time series // Proceedings of the 5th International Workshop «Applied Methods of Statistical Analysis. Statistical Computation and Simulation». – 2019. – P. 311-319.
95. Ambos A.Ju., Mikhailov G.A. Statistical modelling of the exponentially correlated multivariate random field // Rus. J. Numer. Analys. Math. Modeling. –2011. – V. 26. № 3. – P. 213-232.
96. Amiranashvili A, Mirianashvili K, Fedorova N, Levit V, Carnaúba F, Silva A. Comparative analysis of air equivalent-effective temperature in some cities of Georgia and Brazil // Proc of Int Conf “Environment and Global Warming”, Dedicated to the 100th Birthday Anniversary of Academician F. Davitaya, Collected Papers New Series. – 2011. – P. 105-110.
97. Anderson G.B., Bell M.L., Peng R.D. Methods to calculate the heat index as an exposure metric in environmental health research // Env Health Perspect. – 2013. – 121(10). – P. 1111-1119.
98. Andrade-Bejarano M., Longford N.T. Outliers in mixed models for monthly average temperatures // Aust J Stat. – 2010. – V. 39, № 3. – P. 203-221.

99. Apaydin H., Sonmez F.K., Yildirim Y.E. Spatial Interpolation Techniques for Climate Data in the Gap Region in Turkey // *Climate Research*. – 2004. – V. 28, No. 1. – P. 31-40.
100. Ashraf M., Loftis J.C., Hubbard K.G. Application of geostatistics to evaluate partial weather station networks // *Agricultural and Forest Meteorology*. – 1997. – V. 84, № 3–4. – P. 255-271.
101. Babicheva G.A., **Kargapolova N.A.**, Ogorodnikov V.A. Special algorithms for the simulation of homogeneous random fields // *Numerical Analysis and Applications*. – 2016. – V. 9, № 2. – P. 95-106. DOI: 10.1134/S1995423916020014
102. Balouktsis A., Tsanakas D., Vachtsevanos G. Stochastic Simulation of Hourly and Daily Average Wind Speed Sequences // *Wind Engineering*. – 1986. V. 10, № 1. – P. 1-11.
103. Belkin V.S., Dyurgerov M.B., Finaev A.F., Soroko S.I. Bioclimatic evaluation of the human discomfort level for several Antarctic regions // *Human Physiology*. – 2016. – V. 42, № 2, – P. 119-127.
104. Bessac J., Ailliot P., Cattiaux J., Monbet V. Comparison of hidden and observed regime-switching autoregressive models for (u, v)-components of wind fields in the northeastern Atlantic // *Advances in Statistical Climatology, Meteorology and Oceanography*. – 2016. – V. 2. – P.1-16.
105. Blazejczyk K., Epstein Y., Jendritzky G., Staiger H., Tinz B. Comparison of UTCI to selected thermal indices // *Int J Biometeor*. – 2012. – V. 56. – P. 515-535.
106. Brissette F.P., Khalili M., Leconte R. Efficient stochastic generation of multi-site synthetic precipitation data // *Journal of Hydrology*. – 2007. – V. 345, № 3-4. – P. 121-133.
107. Buishand T.A. Stochastic modeling of daily rainfall sequences. Netherlands: Mededlingen Landbouwhogeschool Wageningen, 1977, 77–3.
108. Byrne M.P., OGorman P.A. Understanding decreases in land relative humidity with global warming: conceptual model and GCM simulations // *J Climate*. – 2016. – V. 29. – P. 9045-9061.

109. Cario M.C., Nelson B.L. Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Working paper. Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 1997, 19 p.
110. Cannistraro G., Cannistraro M., Restivo R. A Smart Thermo-hygrometric Global Index for the Evaluation of Particularly Critical Urban Areas Quality: the City of Messina Chosen as a Case Study // *Smart Science*. – 2014. – V. 2, № 1. – P. 29-35.
111. Chai H., Cheng W., Zhou C., Chen X., Ma X., Zhao S. Analysis and comparison of spatial interpolation methods for temperature data in Xinjiang Uygur Autonomous Region, China // *Natural Science*. – 2011. – V. 3. – P. 999-1010.
112. Chen S.S., Gopinath R.A. Gaussianization // *Advances in Neural Information Processing Systems*, MIT Press. – 2000. – V. 13. – P. 423-429.
113. Derenok K.V., Ogorodnikov V.A. Numerical simulation of significant long-term decreases in air temperature // *Russ J Num Anal Math Modell*. – 2008. – V. 23, №3. – P. 223-277.
114. Eccel E. Estimating air humidity from temperature and precipitation measures for modelling applications // *Meteorol Appl*. – 2012. – V.19. – 118128.
115. Emelina S.V., Makosko A.A., Matesheva A.V. Evaluation of changing weather and climate comfort conditions in Russia from 1980 to 2050 // *Turbulence, Atmosphere and Climate Dynamics. IOP Conference Series: Earth and Environmental Sciences*. – 2019. – V. 231. – 012015. – P. 1-7.
116. Environment Canada [Электронный ресурс]. Wind chill: the chilling facts. Gatineau - Quebec: Environment Canada. – 2014. http://publications.gc.ca/collections/collection_2014/ec/En56-222-2-2014-eng.pdf (последнее посещение: 29.04.2020)
117. Evstafieva A.I., Khlebnikova E.I., Ogorodnikov V.A. Numerical stochastic models for complexes of time series of weather elements // *Russ J Num Anal Math Modell*. – 2005. – V. 20, №6. – P. 535-548.
118. Forsythe N., Fowler H.J., Blenkinsop S., Burton A., Kilsby C.G., Archer D.R., Harpham C., Hashmi M.Z. Application of a stochastic weather generator to assess

- climate change impacts in a semi-arid climate: The Upper Indus Basin // *Journal of Hydrology*. – 2014. – V. 517. – P. 1019-1034.
119. de Freitas C.R., Grigorieva E.A. A comparison and appraisal of a comprehensive range of human thermal climate indices // *Int J Biometeorol.* – 2017. – V. 61. – P. 487-512.
120. Gabriel K.R., Neumann J. A Markov chain model for daily rainfall occurrences at Tel Aviv // *Quarterly Journal of the Royal Meteorological Society*. – 1962. – V. 88, № 375. – P. 90-95.
121. Gadzhev G., Ganey K. Computer Simulations of Air Quality and Bio-Climatic Indices for the City of Sofia // *Atmosphere*. – 2021. – V. 12, № 8. – 1078.
122. Gevorkyan M.N., Demidova A.V., Zaryadov I.S., Korolkova A.V., Kulyabov D.S., Sevastianov L.A., Sobilewski R.A. Approaches to Stochastic Modeling of Wind Turbines // *Proceedings of 31st European conference on modelling and simulation, ECMS 2017*. P. 622-627.
123. Głogowski A., Bryś K., Perona P. Bioclimatic conditions of the Lower Silesia region (South-West Poland) from 1966 to 2017 // *Int J Biometeorol.* – 2021. – V. 65, № 9. – P.1515-1527.
124. Głogowski A., Perona P., Bryś K., Bryś T. Nonlinear reconstruction of bioclimatic outdoor-environment dynamics for the Lower Silesia region (SW Poland) // *Int J Biometeorol.* – 2021. – V. 65. – P. 1189-1203.
125. *Guide to Meteorological Instruments and Methods of Observation: CIMO guide*. Switzerland: WMO, 2014.
126. Guo Y., Dong W., Ren F., Zhao Z.-C., Huang J. Surface air temperature simulations over China with CMIP5 and CMIP3 // *Adv Clim Change Res.* – 2013. – V. 4, № 3. – P.145-152.
127. Hartkamp A.D., De Beurs K., Stein A., White J.W. *Interpolation Techniques for Climate Variables*. NRG-GIS Series 99-01. Mexico, D.F.: CIMMYT, 1999.
128. Haugh M. *An introduction to copulas*. IEOR E4602: quantitative risk management. Lecture notes. Columbia University, 2016.

129. Hering A., Kazor K., Kleiber W. A Markov-Switching Vector Autoregressive Stochastic Wind Generator for Multiple Spatial and Temporal Scales // Resources. – 2015. – V. 4. – P. 70-92.
130. Houghten F.C., Teague W.W., Miller W.E., Yant W.P. Thermal exchanges between the human body and its atmospheric environment // American J of Physiology. – 1929. – V. 83, № 3. – P. 386-406.
131. Huth R., Kysel J., Dubrovsk M. Statistical downscaling and weather generator: higher-order statistical moments // Studia geophysica et geodaetica. – 2003. – V. 47, № 1. – P. 203-216.
132. Isaaks E.H., Srivastava R.M. An introduction to applied geostatistics. Oxford Univesity Press, New York, 1989.
133. Ivanov V.Y., Bras R.L., Curtis D.C. A weather generator for hydrological, ecological, and agricultural applications // Water Resour. Res. – 2007. – V. 43. – W10406.
134. **Kargapolova N.A.** Algorithms for numerical simulation of piecewise-constant periodically correlated processes based on different types of point flows // Proceedings of the 4th Stochastic Modeling Techniques and Data Analysis International Conference with Demographics Workshop Мальта, 2016. – P. 183-191.
135. **Kargapolova N.A.** Monte Carlo Simulation of Daily Precipitation and River Flow Conditional Spatio-Temporal Fields // Proceedings of the European Simulation and Modelling Conference 2017. Belgium, 2017. – P. 311-315.
136. **Kargapolova N.** Monte Carlo Simulation of Non-stationary Air Temperature Time-Series // Proc. of 8th Int. Conf. on Simulation and Modeling Methodologies, Technologies and Applications – SIMULTECH, Porto, 2018, edited by F.D. Rango, T. Ören and M.S. Obaidat. 2018. – P. 323-329.
137. **Kargapolova N.A.** Monte Carlo simulation of piecewise-constant periodically correlated processes // Proc. of 30th European Simulation and Modelling Conference 2016, Las Palmas, 2016. –P. 79-83.

138. **Kargapolova N.** Numerical Stochastic Model of Non-stationary Time Series of the Wind Chill Index // *Methodol Comput Appl Probab.* – 2021 (online first – 2020). – V. 23. – P. 257-271. <https://doi.org/10.1007/s11009-020-09778-x>
139. **Kargapolova N.A.** Numerical Study of the Conditional Time Series of the Average Daily Heat Index // *Proceedings of Papers International Conference on Time Series and Forecasting.* – 2019. – V. 1. – P. 226-234.
140. **Kargapolova N.** Stochastic model of the joint time-series of air temperature and atmospheric pressure // *Proc. of the 32nd European Modelling and Simulation Conference, Ghent, 2018.* – P. 199-204.
141. **Kargapolova N.A.** Stochastic model of the time series of the average daily bioclimatic index of severity of climatic regime // *Proceedings of the 33rd European Simulation and Modelling Conference, Spain, 2019.* – P. 185-189.
142. **Kargapolova N.** Stochastic Models of Non-stationary Time Series of the Average Daily Heat Index // *Proceedings of the 9th International Conference on Simulation and Modeling Methodologies, Technologies and Applications – 2019.* – V. 1. – P. 209-215. DOI: 10.5220/0007788502090215
143. **Kargapolova N.** Stochastic Simulation of Meteorological Non-Gaussian Joint Time-Series // *Simulation and Modeling Methodologies, Technologies and Applications. Advances in Intelligent Systems and Computing.* – 2019. – V. 873. – P. 117-127. https://doi.org/10.1007/978-3-030-01470-4_7
144. **Kargapolova N.** Stochastic Simulation of Non-stationary Meteorological Time-series: Daily Precipitation Indicators, Maximum and Minimum Air Temperature Simulation using Latent and Transformed Gaussian Processes // *Proceedings of the 7th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, 2017.* – P. 173-179.
145. **Kargapolova N.** Stochastic Model of Spatial Fields of the Average Daily Wind Chill Index // *Information.* – 2020. – V. 11, № 4. – P. 177. <https://doi.org/10.3390/info11040177>

146. **Kargapolova Nina** Stochastic simulation of the spatio-temporal field of the average daily heat index in Southern Russia // *Climate Research*. – 2020. – V. 82. – P.149-160. DOI: <https://doi.org/10.3354/cr01623>
147. **Kargapolova N.**, Khlebnikova E., Ogorodnikov V. Monte Carlo simulation of the joint non-Gaussian periodically correlated time-series of air temperature and relative humidity // *Statistical papers*. – 2018. – V. 59, № 4. – P. 1471-1481. doi 10.1007/s00362-018-1031-z
148. **Kargapolova N.A.**, Khlebnikova E.I., Ogorodnikov V.A. Numerical study of properties of air heat content indicators based on the stochastic model of the meteorological processes // *Russ J Num Anal Math Modelling*. – 2019. – V. 34, № 2. – P. 95-104. <https://doi.org/10.1515/rnam-2019-0008>
149. **Kargapolova N.A.**, Khlebnikova E.I., Ogorodnikov V.A. Stochastic models of joint non-stationary time-series of air temperature, relative humidity and atmospheric pressure // *Communications in Statistics - Simulation and Computation*. – 2021 (published online – 2019). – V. 50, №. 12. – P. 3972-3983. <https://doi.org/10.1080/03610918.2019.1635157>
150. **Kargapolova N.A.**, Ogorodnikov V.A. An algorithm for numerical simulation of isotropic random fields and its meteorological application // *Proceedings of International Workshop «Applied Methods of Statistical Analysis. Nonparametric approach»*. – 2015. – P. 392-399.
151. **Kargapolova N.A.**, Ogorodnikov V.A. Conditional stochastic model of daily precipitation and river flow joint spatial field // *Proceedings of International Workshop «Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis (AMSA'2017)»*. – 2017. – P. 298-302.
152. **Kargapolova N.A.**, Ogorodnikov V.A. Inhomogeneous Markov chains with periodic matrices of transition probabilities and their application to simulation of meteorological processes // *Russ. J. Numer. Anal. Math. Modelling*. – 2012. – V. 27, № 3. – P. 213-228. <https://doi.org/10.1515/rnam-2012-0012>
153. **Kargapolova Nina A.**, Ogorodnikov Vasily A. Numerical stochastic modelling of spatial and spatio-temporal fields of the wind chill index in the South of Western

- Siberia // Russ J Num Anal Math Modelling. – 2021. – V. 36, №. 1. – P. 33-42.
<https://doi.org/10.1515/rnam-2021-0003>
154. **Kargapolova N.**, Ogorodnikov V. Stochastic Model of Conditional Non-stationary Time Series of the Wind Chill Index in West Siberia // Methodol Comput Appl Probab. – 2021. <https://doi.org/10.1007/s11009-021-09861-x>
155. Katz R.W. Precipitation as a chain-dependent process // Journal of Applied Meteorology. – 1977. – V. 16, № 7. – P. 671-676.
156. Kleiber W., Katz R.W., Rajagopalan B. Daily minimum and maximum temperature simulation over complex terrain // Ann Appl Stat. – 2013. – V. 7, № 1. – P. 588-612.
157. Kleiber W., Katz R.W., Rajagopalan B. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes // Water Resources Research. – 2012. – V. 48, № 1. – W01523.
158. Lanciani A., Salvati M. Spatial interpolation of surface weather observations in Alpine meteorological services. FORALPS Technical Report, 2. Università degli Studi di Trento, Dipartimento di Ingegneria Civile e Ambientale, Trento, Italy, 2008.
159. Legesse A.N., Saha A.K., Carpanen R.P. Characterisation of wind speed series and power in Durban // Journal of Energy in Southern Africa. – 2017. – V. 28, № 3. – P. 66-78.
160. Lowe D., Ebi K.L., Forsberg B. Heatwave Early Warning Systems and Adaptation Advice to Reduce Human Health Consequences of Heatwaves // Int. J. Environ. Res. Public Health. – 2011. – V. 8. – P. 4623-4648.
161. Lukić M., Pecelj M., Protić B., Filipović D. An evaluation of summer discomfort in Nis (Serbia) using Humidex // J Geograp Institute Jovan Cvijic, SASA. – 2019. – V. 69, № 2. – P. 109-122.
162. Maatouk H., Bay X. A new rejection sampling method for truncated multivariate Gaussian random variables restricted to convex sets // Monte Carlo and Quasi-Monte Carlo Methods. – 2016. – V. 163. – P. 521-530.

163. Masterton J.M., Richardson F.A. Humidex. A Method of Quantifying Human Discomfort Due to Excessive Heat and Humidity. CLI 1-79, Environment Canada, Atmospheric Environment Service, Downsview, Ontario, 1979.
164. Medvyatskaya A.M., Ogorodnikov V.A. Approximate spectral models of random processes with periodic properties // Russ J Num Anal Math Modelling. – 2019. – V. 34, №. 6. – P. 353-360.
165. Mehan S., Guo T., Gitau M.W., Flanagan D.C. Comparative Study of Different Stochastic Weather Generators for Long-Term Climate Data Simulation // Climate. – 2017. – V. 5. – 26.
166. Mekis É., Vincent L.A., Shephard M.W., Zhang X. Observed Trends in Severe Weather Conditions Based on Humidex, Wind Chill, and Heavy Rainfall Events in Canada for 1953–2012 // Atmos Ocean/ – 2015. – V. 53. – P. 383-397.
167. Miñano R., Milano F. Construction of SDE-based wind speed models with exponential autocorrelation // Renewable Energy. – 2015. – V. 94.
168. Missenard A. L'Homme et le climat. Paris, 1937.
169. Moeletsi M.E., Shabalala Z.P., De Nysschen G., Walke S. valuation of an inverse distance weighting method for patching daily and dekadal rainfall over the Free State Province, South Africa // Water SA. – 2016. – V. 42, № 3. – P. 466-474.
170. Monbet V., Ailliot P. Sparse vector Markov switching autoregressive models. Application to multivariate time series of temperature // Computational Statistics & Data Analysis. – 2017. – V. 108. – P. 40-51.
171. Niederreiter H. Random number generation and Quasi-Monte Carlo method. Philadelphia: SIAM, 1992, 243 p.
172. Ogorodnikov V.A., **Kargapolova N.A.**, Sereseva O.V. Numerical stochastic model of spatial fields of daily sums of liquid precipitation // Russ J Num Anal Math Modelling. – 2013. – V. 28, №. 2. – P. 187-200. <https://doi.org/10.1515/rnam-2013-0011>
173. Ogorodnikov V., **Kargapolova N.**, Sereseva O. Numerical Stochastic Models of Meteorological Processes and Fields // Springer Proceedings in Mathematics and Statistics. – 2014. – P. 409-417. DOI 10.1007/978-1-4939-2104-1_40.

174. Ogorodnikov V.A., Khlebnikova E.I., Kosyak S.S. Numerical stochastic simulation of joint non-Gaussian meteorological series // *Russ J Num Anal Math Modelling*. – 2009. – V. 24, №. 5. – P. 467-480.
175. Ogorodnikov V.A., Prigarin S.M. *Numerical Modelling of Random Processes and Fields: Algorithms and Applications*, VSP, Utrecht, 1996.
176. Ogorodnikov V.A., Prigarin S.M. On stochastic interpolation of discrete random processes and fields // *Russ J Num Anal Math Modelling*. – 1996. – V. 11, №. 1. – P. 49-69.
177. Ogorodnikov V., Sereseva O. Approximate numerical modelling of inhomogeneous stochastic fields of daily sums of liquid precipitation // *Russ J Num Anal Math Modelling*. – 2014. – V. 29, №. 6. – P. 375-382.
178. Ogorodnikov V.A., Seresseva O.V., **Kargapolova N.A.** Stochastic models of piecewise-constant and piecewise-linear non-Gaussian processes based on Poisson flows // *Russ J Num Anal Math Modelling*. – 2016. – V. 31, №. 3. – P. 179-185. <https://doi.org/10.1515/rnam-2016-0018>
179. Oszcewski R., Bluestein M. The New Wind Chill Equivalent Temperature Chart // *Bulletin of the American Meteorological Society*. – 2005. – V. 86. – P. 1453-1458.
180. Owen D.B. Table for computing bivariate normal probabilities // *Ann. Math. Statist.* – 1956. – V. 27, № 4. – P. 1075-2000.
181. Pal R., Sehgal V., Misra A., Ghosh K., Mohanty U.C., Rana R. Application of Seasonal Temperature and Rainfall Forecast for Wheat Yield Prediction for Palampur, Himachal Pradesh // *International Journal of Agriculture and Food Science Technology*. – 2013. – V. 4. – P. 453-460.
182. Pan F., Nagaoka L., Wolverson S., Atkinson S., Kohler T., O’neill M. A Constrained Stochastic Weather Generator for Daily Mean Air Temperature and Precipitation // *Atmosphere*. – 2021. – V. 12., № 2. – 135.
183. Parlange M.B., Katz R.W. An extended version of the richardson model for simulating daily weather variables // *Journal of Applied Meteorology*. – 2000. – V. 39, № 5. – P. 610-622.

184. Perry A.G., Korenberg M.J., Hall G.G., Moore K.M. Modeling and syndromic surveillance for estimating weather-induced heat-related illness // *J Environ Public Health*. – 2011. – 750236.
185. Pfliegerer P., Jézéquel A., Legrand J., Legrix N., Markantonis I., Vignotto E., Yiou P. Simulating compound weather extremes responsible for critical crop failure with stochastic weather generators // *Earth System Dynamics*. – 2021. – V. 12. – P. 103-120.
186. Prigarin S.M., Litvenko K.V. Conditional spectral models of extreme ocean waves // *Russ J Num Anal Math Modelling*. – 2012. – V. 27, № 3. – P. 289-302.
187. Richardson C.W. Stochastic simulation of daily precipitation, temperature and solar radiation // *Water Resources Research*. – 1981. – V. 17, № 1. – P. 182–190.
188. Richardson C.W., Wright D.A. WGEN: A Model for Generating Daily Weather Variables. – U. S. Department of Agriculture, Agricultural Research Service, ARS-8, 1984. – 83 p.
189. Romaszko-Wojtowicz A., Cymes I., Dragańska E., Doboszyńska A., Romaszko J., Glińska-Lewczuk K. Relationship between biometeorological factors and the number of hospitalizations due to asthma // *Sci Rep*. – 2020. – V. 10, № 1. – 9593.
190. Rothfus L.P. The Heat Index “Equation” (or, More Than You Ever Wanted to Know About Heat Index). SR 90–23. Fort Worth, TX: NOAA, 1990.
191. Schoen C. A new empirical model of the temperature– humidity index. *J Appl Meteorol*. – 2005. – V.44. – P. 1413-1420.
192. Semenov M.A., Barrow E.M. Use of a stochastic weather generator in the development of climate change scenarios // *Climatic Change*. – 1997. – V. 35. – P. 397-414.
193. Semenov M.A., Brooks R.J., Barrow E.M., Richardson C.W. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates // *Climate Research*. – 1998. – V. 10. – P. 95-107.
194. Shartova N., Shaposhnikov D., Konstantinov P., Revich B. Cardiovascular mortality during heat waves in temperate climate: an association with bioclimatic

- indices // *Int J of Environmental Health Research.* – 2018. – V. 28, № 5. – P. 522-534.
195. Sheuer E.M., Stoller D.D. On generation of normal random vectors // *Technometrics.* – 1963. – V. 4, № 2. – P. 278-281.
196. Siple P.A., Passel C.F. Measurements of dry atmospheric cooling in sub-freezing temperatures // *Proc Amer Philos Soc.* – 1945. – V. 89, № 1. – P. 177-199.
197. Sirangelo B., Caloiero T., Coscarelli R., Ferrari E., Fusto F. Combining stochastic models of air temperature and vapour pressure for the analysis of the bioclimatic comfort through the Humidex // *Sci Rep.* – 2020. – V. 10. – 11395.
198. Sluiter R. Interpolation methods for climate data (literature review). KNMI: De Bilt, The Netherlands, 2009.
199. Sohoni V., Gupta S.C., Nema R.K. A Critical Review on Wind Turbine Power Curve Modelling Techniques and Their Applications in Wind Based Energy Systems // *Journal of Energy.* – 2016. – V. 2016, – 8519785.
200. Srikanthan R., McMahon T.A. Stochastic generation of annual, monthly and daily climate data: A review // *Hydrology and Earth System Sciences.* – 2001. – V. 4, № 5. – P. 653-670.
201. Steadman R.G. The Assessment of Sultriness, Part I: A Temperature-Humidity Index Based on Human Physiology and Clothing Science // *J. Appl. Meteor.* – 1979. – V. 18. – P. 861-873.
202. Steadman R.G. A universal scale of apparent temperature // *J Climate Appl Meteorol.* – 1984. – V. 23. – P. 1674-1687.
203. Swiler L., Gulian M., Frankel A., Safta C., Jakeman J. A Survey of Constrained Gaussian Process: Approaches and Implementation Challenges // *J of Machine Learning for Modeling and Computing.* – 2020. – V. 1, № 2. – P. 119-156.
204. Todorovic P., Woolhiser D.A. A stochastic model of n-day precipitation. *Journal of Applied Meteorology.* – 1975. – V. 14. – P. 17-24.
205. Varouchakis E. Geostatistics: Mathematical and Statistical Basis (Book Chapter) // *Spatiotemporal Analysis of Extreme Hydrological Events*, 1st ed., ELSEVIER, 2019, p.1-38.

206. Vesely F.M. Paleari L., Movedi E., Bellocchi G., Confalonieri R. Quantifying Uncertainty Due to Stochastic Weather Generators in Climate Change Impact Studies // *Sci Rep.* – 2019. – V. 9, 9258.
207. Yagloglou C.P. The heat given up by the human body and its effect on heating and ventilating problem // *Transact Americ soc of heat and ventilat engineers.* – 1924. – V. 30. – P. 365-376.
208. Zare S., Hasheminejad N., Shirvan H.E., Hemmatjo R., Sarebanzadeh K., Ahmadi S. Comparing Universal Thermal Climate Index (UTCI) with selected thermal indices/environmental parameters during 12 months of the year // *Weather and Climate Extremes.* – 2018. – V. 19. – P. 49-57.
209. Zheng Z., Dai H. Simulation of multi-dimensional random fields by Karhunen-Loève expansion // *Computer Methods in Applied Mechanics and Engineering.* – 2017. – V. 324. – 10.1016/j.cma.2017.05.022.
210. Zheng Z., Dai H., Wang Y., Wang W. A sample-based iterative scheme for simulating non-stationary non-Gaussian stochastic processes // *Mechanical Systems and Signal Processing.* – 2021. – V. 151, 107420.

Приложение 1. Известные методы моделирования гауссовских и негауссовских случайных процессов

Моделирование гауссовских векторов

В общем случае, моделирование n -мерного гауссовского вектора $\vec{\xi}$ с нулевым средним и произвольной корреляционной матрицей

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{21} & \dots & r_{21} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nn} \end{pmatrix}$$

осуществляется с помощью линейного преобразования

$$\vec{\xi} = A\vec{\varphi},$$

где $\vec{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_n)^T$ – гауссовский вектор с нулевым средним и единичной корреляционной матрицей, а A – нижнетреугольная матрица, удовлетворяющая условию $AA^T = R$ [175]. Этот метод можно назвать методом моделирования гауссовского вектора на основе разложения Холецкого корреляционной матрицы.

Матрица A

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{21} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

может быть вычислена по хорошо известным рекуррентным формулам, приведённым, например, в [87],

$$\begin{aligned} a_{11} &= r_{11}^{1/2}, \quad a_{i1} = r_{i1}/a_{11}, \quad i = 2, \dots, n, \\ a_{ij} &= \frac{r_{ij} - \sum_{k=1}^{j-1} a_{ik} r_{jk}}{a_{ii}^{1/2}}, \quad i = 2, \dots, n, \quad j = 2, \dots, i, \\ a_{ij} &= 0, \quad i = 1, \dots, n-1, \quad j = i+1, \dots, n. \end{aligned}$$

$$d_k^2 = r_{k+1k+1} - \vec{b}^T [k] \vec{r} [k], \quad k = 1, \dots, n-1,$$

а J_k – матрица перестановок

$$J_{k+1} = \begin{pmatrix} 0 & 0 & \dots & 1 \\ 0 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots \\ 1 & 0 & \dots & 0 \end{pmatrix}.$$

В работе [175] описан ряд свойств рассмотренного метода условных распределений для моделирования гауссовского вектора с заданной корреляционной матрицей.

Моделирование негауссовских процессов

В Главах 1–3 известный метод обратных функций распределения был использован для моделирования безусловных негауссовских случайных процессов. Опишем этот метод подробнее.

Метод обратных функций распределения является одним из методов, применяемых для моделирования случайных процессов с заданными одномерными распределениями и корреляционной функцией [53, 66, 109]. Этот метод был предложен в [64]. В зарубежной литературе данный метод часто называют методом моделирования с помощью гауссовских копул (см., например, [128]).

Пусть $\eta(t)$ – стандартный ($E\eta(t)=0, D\eta(t)=1$) гауссовский процесс с корреляционной функцией $\rho(t,s)$. Тогда случайный процесс $\xi(t)$ с функциями одномерных распределений $F_t(x)$ моделируется по формуле

$$\xi(t) = F_t^{-1}(\Phi(\eta(t))), \quad (\text{П1})$$

где $\Phi(x)$ – функция стандартного нормального распределения. Ковариационная функция процесса $\xi(t)$ имеет вид

$$r(t,s) = R(t,s, \rho(t,s)),$$

$$R(t, s, \rho) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_t^{-1}(\Phi(y)) F_s^{-1}(\Phi(z)) \phi_{\rho}(y, z) dy dz,$$

где

$$\phi_{\rho}(y, z) = \left(2\pi \sqrt{1 - \rho^2} \exp \left(\frac{y^2 + z^2 - 2\rho yz}{2(1 - \rho^2)} \right) \right)^{-1}$$

– плотность двумерного гауссовского вектора с нулевым математическим ожиданием и единичной дисперсией у каждой из компонент вектора и коэффициентом корреляции между компонентами равным ρ [66]. Таким образом, метод обратных функций распределения сводится к вычислению функции

$$\rho(t, s) = R^{-1}(t, s, r(t, s)),$$

моделированию гауссовского процесса $\eta(t)$ с такой корреляционной функцией и его нелинейному преобразованию (П1).

Детальное изложение данного метода и область его применимости приведены в [66].

Приложение 2. Сведения о метеорологических станциях

В таблице П1 приведены основные сведения о метеорологических станциях, данные с которых использовались для определения входных параметров моделей и их верификации.

Таб. П1. Сведения о метеостанциях.

Синоптический индекс станции	Название станции и/или населённый пункт	Координаты		Высота над уровнем моря, м
		Широта, град с.ш.	Долгота, град. в.д.	
22028	Териберка	69°12'	35°07'	33
22113	Мурманск	68°58'	33°03'	57
22408	Калевала	65°13'	31°09'	118
22550	Архангельск	64°30'	40°44'	8
22641	Онега	63°54'	38°07'	11
22820	Петрозаводск	61°49'	34°16'	110
22845	Каргополь	61°31'	38°56'	124
22981	Великий Устюг	60°46'	46°18'	93
23445	Надым	65°28'	72°40'	14
23804	Сыктывкар	61°40'	50°52'	116
23933	Ханты-Мансийск	61°01'	69°07'	46
24641	Вилуйск	63°47'	121°37'	110
24688	Оймякон	63°15'	143°09'	740
24959	Якутск	62°01'	129°43'	98
25563	Анадырь	64°47'	177°34'	64
26063	Санкт-Петербург	59°58'	30°18'	3

26702	Калининград	54°42'	20°37'	20
27595	Казань	55°44'	49°12'	116
28367	Тюмень	57°07'	65°26'	101
28440	Екатеринбург	56°50'	60°38'	281
28698	Омск	55°01'	73°23'	121
29328	Бакчар	57°00'	82°04'	109
29348	Первомайское	57°04'	86°13'	114
29393	Червянка	57°39'	99°32'	219
29418	Северное	56°20'	78°22'	125
29430	Томск	56°30'	84°55'	141
29539	Болотное	55°40'	84°24'	192
29541	Тайга	56°04'	85°37'	250
29557	Тисуль	55°45'	88°18'	210
29605	Татарск	55°13'	75°58'	110
29612	Барабинск	55°20'	78°22'	119
29638	Огурцово (Новосибирск)	54°54'	82°57'	131
29645	Кемерово	55°15'	86°13'	148
29698	Нижеудинск	54°53'	99°02'	410
29829	Камень-на-Оби	53°49'	81°16'	127
29838	Барнаул	53°26'	83°31'	183
29849	Кузедеево	53°20'	87°11'	293
29915	Славгород	52°58'	78°39'	125
29939	Бийск-Зональная	52°41'	84°56'	222
29999	Орлик	52°30'	99°49'	1376
30209	Ершово	57°02'	102°18'	372
30230	Киренск	57°46'	108°04'	256
30309	Братск	56°17'	101°45'	410
30337	Казачинск	56°17'	107°37'	355
30356	Таксимо	56°23'	114°50'	513

30433	Нижнеангарск	55°47'	109°33'	477
30437	Карам	55°09'	107°37'	495
30504	Тулун	54°36'	100°38'	523
30521	Жигалово	54°48'	105°10'	416
30554	Багдарин	54°28'	113°35'	900
30555	Троицкий прииск	54°37'	113°08'	1315
30612	Балаганск	54°00'	103°04'	428
30627	Баяндай	53°06'	105°32'	757
30632	о. Большой Ушканий	53°51'	108°35'	458
30635	Усть-Баргузин	53°25'	109°01'	459
30636	Баргузин	53°37'	109°38'	488
30650	Романовка	53°12'	112°47'	920
30703	Инга	52°58'	101°59'	541
30710	Иркутск	52°16'	104°21'	467
30716	Хомутово	52°28'	104°22'	453
30727	Большое Голоустное	52°02'	105°25'	461
30745	Сосново-Озерское	52°32'	11°33'	952
30758	Чита	52°05'	113°29'	671
30811	Тунка	51°44'	102°32'	720
30815	Хамар-Дабан	51°32'	103°36'	1442
30822	Бабушкин	51°43'	105°51'	465
30823	Улан-Удэ	51°50'	107°36'	514
30825	Иволгинск	51°45'	107°17'	562
30844	Хилок	51°21'	110°28'	800
30859	Агинское	51°06'	114°31'	680
30915	Цакир	50°26'	103°36'	984
30925	Кяхта	50°22'	106°27'	791
30935	Красный Чикой	50°22'	108°45'	770
31369	Николаевск-на-Амуре	53°09'	140°42'	67

31510	Благовещенск	50°15'	127°30'	130
32061	Александровск-Сахалинский	50°54'	142°10'	30
32583	Петропавловск-Камчатский	52°59'	158°39'	32
32618	о. Беринга	55°12'	165°59'	14
34009	Курск	51°46'	36°10'	246
34172	Саратов	51°29'	45°56'	151
34834	Приморско-Ахтарск	46°02'	38°09'	3
34927	Краснодар	45°03'	39°02'	28
34949	Ставрополь	45°07'	42°05'	451
35121	Оренбург	51°41'	55°06'	115
36034	Рубцовск	51°35'	81°12'	216
36096	КЫЗЫЛ	51°43'	94°30'	626
36259	Кош-Агач	50°00'	88°40'	1759
37001	Анапа	44°53'	37°17'	30
37018	Туапсе	44°06'	39°04'	67
37031	Армавир	44°59'	41°07'	158
37054	Минеральные Воды	44°14'	43°04'	315
37099	Сочи	43°35'	39°46'	57
37107	Красная Поляна	43°41'	40°12'	566
37112	Зеленчукская	43°52'	41°34'	928
37123	Кисловодск	43°54'	42°43'	943
37196	Клухорский перевал	43°15'	41°50'	2037
37228	Владикавказ	43°02'	44°41'	708
37461	Сулак, высокогорная	42°22'	46°15'	2927
37470	Дербент	42°04'	48°18'	-19
37471	Буйнакск	42°49'	47°07'	472
37472	Махачкала	42°58'	47°33'	28
37663	АХТЫ	41°28'	47°45'	1016

**Приложение 3. Некоторые результаты верификации моделей,
предложенных в параграфах 3.3 и 3.4 Главы 3**

*Результаты верификации модели условных временных рядов
полусуточного ИХС*

Ниже в таблицах П2-П4 приведены результаты верификации стохастической модели условных временных рядов полусуточного индекса холодного стресса, предложенной в параграфе 3.3 Главы 3.

Таб. П2. Оценки вероятности $p(i,l) = P(W_i < l | W_1 < l, W_2 < l)$. Кош-Агач,
16-31 января.

		$l = -17 \text{ } ^\circ\text{C}$	$l = -23 \text{ } ^\circ\text{C}$	$l = -29 \text{ } ^\circ\text{C}$	$l = -35 \text{ } ^\circ\text{C}$
$i = 3$	РД, $p(i,l) \pm 3\sigma$	0.978 ± 0.045	0.917 ± 0.165	0.875 ± 0.411	0.833 ± 0.534
	МД	0.971	0.808	0.813	0.678
$i = 5$	РД, $p(i,l) \pm 3\sigma$	0.956 ± 0.059	0.944 ± 0.198	0.875 ± 0.428	0.667 ± 0.546
	МД	0.965	0.799	0.768	0.455
$i = 7$	РД, $p(i,l) \pm 3\sigma$	0.933 ± 0.082	0.806 ± 0.196	0.625 ± 0.417	0.333 ± 0.558
	МД	0.911	0.715	0.679	0.298
$i = 9$	РД, $p(i,l) \pm 3\sigma$	0.889 ± 0.103	0.806 ± 0.209	0.688 ± 0.432	0.333 ± 0.572
	МД	0.905	0.652	0.603	0.276

Таб. П3. Оценки среднего числа $AN(lev)$ компонент вектора

$\vec{W} = (W_1, W_2, \dots, W_{2d})$ меньших lev °C при условии $W_1 < EW_1, W_2 < EW_2$. Огурцово
(Новосибирск), 15-31 декабря.

lev °C	РД, $AN(lev) \pm \sigma$	МД
-5	30.000 ± 0.239	29.838
-11	27.750 ± 0.601	28.314
-17	22.500 ± 1.137	23.418
-23	15.550 ± 1.548	17.093
-29	8.000 ± 1.382	8.381
-35	3.250 ± 0.869	3.088
-41	1.100 ± 0.415	0.864

Таб. П4. Оценки среднего числа $AN(lev)$ компонент вектора

$\vec{W} = (W_1, W_2, \dots, W_{2d})$ меньших lev °C при условии $W_1 < EW_1, W_2 < EW_2$.
Александровское, 1-15 января.

lev °C	РД, $AN(lev) \pm \sigma$	МД
-5	27.952 ± 0.149	27.816
-11	27.000 ± 0.398	27.071
-17	25.095 ± 0.859	24.520
-23	22.048 ± 1.361	18.495
-29	17.571 ± 1.410	16.542
-35	11.857 ± 1.555	10.864
-41	7.381 ± 1.589	6.597

Результаты верификации модели пространственно-временного поля СИЖ

Далее в таблицах П5-П7 приведены некоторые результаты верификации модели пространственно-временного поля среднесуточного индекса жары, рассмотренной в параграфе 3.4 Главы 3.

По модельным и реальным траекториям оценивалась вероятность $p_N(lev)$ того, что в течение N суток СИЖ выше заданного уровня $lev^{\circ}C$ на всех 18 станциях. Проводились численные эксперименты, в рамках которых изменялись значения уровня $lev^{\circ}C$, длительности N и начала временного интервала a . Оценки по траекториям, построенным с помощью МОФР, попадают в доверительные интервалы $(p_N(lev) - \sigma, p_N(lev) + \sigma)$ и $(p_N(lev) - 3\sigma, p_N(lev) + 3\sigma)$ оценок по реальным данным в 75% и 99% проведенных численных экспериментов, соответственно, а оценки по траекториям, полученным при использовании МН, – в 58% и 75 % экспериментов. Также проводились оценки вероятности $\bar{p}_N(lev)$ того, что в течение N суток СИЖ выше заданного уровня $lev^{\circ}C$ на всех станциях за исключением 2 высокогорных метеостанций, расположенных на Клухорском перевале и в Сулаке. Оценки по траекториям, полученным с помощью МОФР и МН, попали в доверительные интервалы $(\bar{p}_N(lev) - \sigma, \bar{p}_N(lev) + \sigma)$ оценок по реальным данным в 83% и 77 % численных экспериментов. Оценки вероятности $pr(k, lev)$ того, что в заданный день хотя бы на k станциях СИЖ превысит уровень $lev^{\circ}C$ по траекториям, построенным с применением МОФР И МН попадают в интервалы $(pr(k, lev) - \sigma, pr(k, lev) + \sigma)$ в 88-89% проведенных экспериментов.

Таб. П5. Оценки вероятности $p_N(lev)$ того, что в течение N суток СИЖ выше заданного уровня lev °C на всех 18 станциях.

lev °C	16-20 июля			3-13 августа		
	Реальные данные $p_N(lev) \pm \sigma$	МОФР	МН	Реальные данные $p_N(lev) \pm \sigma$	МОФР	МН
0	1.000 ± 0.003	0.999	0.999	1.000 ± 0.009	0.983	0.974
3	0.951 ± 0.017	0.958	0.963	0.854 ± 0.025	0.862	0.742
6	0.675 ± 0.044	0.649	0.644	0.517 ± 0.034	0.494	0.176
9	0.276 ± 0.037	0.222	0.220	0.127 ± 0.025	0.134	0.002
12	0.008 ± 0.013	0.024	0.020	0.005 ± 0.007	0.011	0.000
15	0.000 ± 0.002	0.000	0.000	0.000 ± 0.001	0.000	0.000

Таб. П6. Оценки вероятности $\bar{p}_N(lev)$ того, что в течение N суток СИЖ выше заданного уровня lev °C на всех станциях за исключением 2 высокогорных.

lev °C	16-20 июля			3-13 августа		
	Реальные данные $\bar{p}_N(lev) \pm \sigma$	МОФР	МН	Реальные данные $\bar{p}_N(lev) \pm \sigma$	МОФР	МН
9	1.000 ± 0.002	1.000	1.000	1.000 ± 0.002	1.000	1.000
12	1.000 ± 0.020	0.995	0.997	1.000 ± 0.012	0.983	1.000
15	0.935 ± 0.042	0.897	0.893	0.863 ± 0.023	0.805	0.882
18	0.472 ± 0.041	0.441	0.442	0.371 ± 0.026	0.351	0.037
21	0.057 ± 0.024	0.072	0.076	0.063 ± 0.015	0.066	0.000
24	0.000 ± 0.006	0.005	0.005	0.000 ± 0.004	0.003	0.000

Таб. П7. Оценки вероятности $pr(k,lev)$ того, что в заданный день хотя бы на k станциях СИЖ превысит уровень lev °C.

lev °C	$pr(3,lev)$, 29 июня			$pr(6,lev)$, 28 июля		
	Реальные данные $pr(3,lev) \pm \sigma$	МОФР	МН	Реальные данные $pr(6,lev) \pm \sigma$	МОФР	МН
19	1.000 ± 0.003	0.998	0.999	1.000 ± 0.008	0.993	0.993
22	0.984 ± 0.019	0.949	0.954	0.951 ± 0.020	0.945	0.943
25	0.740 ± 0.040	0.754	0.760	0.780 ± 0.038	0.788	0.787
28	0.431 ± 0.047	0.442	0.450	0.561 ± 0.047	0.535	0.533
31	0.171 ± 0.034	0.167	0.167	0.358 ± 0.041	0.296	0.301
34	0.024 ± 0.014	0.028	0.026	0.138 ± 0.032	0.150	0.152
37	0.000 ± 0.003	0.002	0.001	0.057 ± 0.023	0.069	0.068
40	0.000 ± 0.001	0.000	0.000	0.016 ± 0.012	0.019	0.020
43	0.000 ± 0.000	0.000	0.000	0.000 ± 0.003	0.001	0.001