

На правах рукописи



Каргаполова Нина Александровна

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

01.01.07 – Вычислительная математика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
доктора физико-математических наук

Новосибирск – 2022

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте вычислительной математики и математической геофизики Сибирского отделения Российской академии наук (ИВМиМГ СО РАН).

Научный консультант:

Огородников Василий Александрович, доктор физико-математических наук

Официальные оппоненты:

Журавлева Татьяна Борисовна, доктор физико-математических наук, Федеральное государственное бюджетное учреждение науки Институт оптики атмосферы им. В.Е. Зуева Сибирского отделения Российской академии наук, ведущий научный сотрудник лаборатории атмосферной радиации

Лемешко Борис Юрьевич, доктор технических наук, профессор, Федеральное государственное бюджетное образовательное учреждение высшего образования «Новосибирский государственный технический университет», профессор кафедры теоретической и прикладной информатики

Черемисин Александр Алексеевич, доктор физико-математических наук, профессор, Федеральное государственное бюджетное учреждение науки Институт химической кинетики и горения им. В.В. Воеводского Сибирского отделения Российской академии наук, заведующий лабораторией дисперсных систем

Ведущая организация:

Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет»

Защита состоится ____ 2022 г. в ____ ч. ____ мин. на заседании диссертационного совета Д 003.061.01 на базе Федерального государственного бюджетного учреждения науки Института вычислительной математики и математической геофизики Сибирского отделения Российской академии наук (ИВМиМГ СО РАН) по адресу: 630090, Новосибирск, проспект Академика Лаврентьева, 6, конференц-зал ИВМиМГ; тел. +7(383)330-71-59.

С диссертацией можно ознакомиться в библиотеке и на сайте ИВМиМГ СО РАН <https://icmmg.nsc.ru/ru>.

Автореферат разослан

2022 г.

Ученый секретарь диссертационного совета Д 003.061.01, д.ф.-м.н.

Рогазинский Сергей
Валентинович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования и степень её разработанности.

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 достаточно хорошо развит. 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 с заданными свойствами 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. В этот период стохастический подход чаще всего применялся для моделирования стационарных временных рядов суточного разрешения отдельных метеорологических параметров. Позднее появились модели нестационарных временных рядов, пространственных и пространственно-временных полей отдельных метеопараметров и их комплексов. 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. Об актуальности разработки стохастических «генераторов погоды», их теоретической и практической значимости свидетельствует то, что «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), в гляциологии (при исследовании процессов формирования снежного покрова), в гидрологии (при исследовании процессов массопереноса и биогенной нагрузки в системе водосбор – водоём), при решении задач агрометеорологии, связанных с исследованием продукционного процесса сельскохозяйственных растений, при анализе транспортных и транспортно-технологических систем. Следует отметить, что численное стохастическое моделирование метеорологических процессов часто является чрезвычайно трудоёмким. 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. Одним из природных факторов, 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. 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. Биоклиматические индексы (БИ) нашли широкое применение. Значения некоторых из них можно увидеть в ежедневном прогнозе погоды (соответствующие значения обозначаются, например, «Ощущается как», «Real feel»). В ряде стран многие из этих индексов используются в работе оперативных систем предупреждения о наступлении некомфортных погодных условий. БИ используют также при исследовании связей между погодными условиями и обострениями различных заболеваний человека, для медико-климатической классификации территорий и их климато-рекреационного районирования. 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 исследование 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. В рамках второго, динамического, подхода для исследования применяются динамические модели метеорологических процессов. В диссертационной работе рассмотрен новый подход к изучению биоклиматических процессов. Суть этого подхода заключается в том, что биоклиматические процессы рассматриваются как случайные процессы, и, следовательно, могут быть исследованы с использованием аппарата стохастического моделирования. Впервые этот подход был предложен при непосредственном участии автора диссертационной работы в 2019 году. Позднее стохастический подход был использован и другими исследователями.

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. Задача моделирования

периодически коррелированных процессов решается различными способами. Разработаны, например, алгоритмы моделирования с использованием векторных процессов авторегрессии, модели с использованием некоторых типов точечных потоков, спектральные модели. 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 векторного процесса авторегрессии. В связи с этим, представляется значимым разработать новые алгоритмы моделирования некоторых классов периодически коррелированных рядов.

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 работы определяется совокупностью полученных результатов:

- 1) разработаны новые стохастические модели временных рядов и пространственно-временных полей комплексов метеорологических параметров и их специальных характеристик, учитывающие временную нестационарность и пространственную неоднородность реальных метеопроцессов;
- 2) впервые предложены и разработаны стохастические модели временных рядов, пространственных и пространственно-временных полей

биоклиматических индексов, учитывающие временную нестационарность и пространственную неоднородность реальных процессов

- 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:

- 1) аппарат теории методов Монте-Карло, включая методы численного моделирования различных классов случайных процессов (в том числе, гауссовских и негауссовских, стационарных, периодически коррелированных и нестационарных);
- 2) аппарат теории вероятностей и математической статистики, в частности, теория случайных процессов, методы оценивания параметров стохастических моделей по реальным данным, критерии проверки статистических гипотез;
- 3) аппарат вычислительной математики, например, численные методы решения нелинейных уравнений и проблемы собственных значений, методы численного интегрирования;
- 4) аппарат математического анализа и линейной алгебры (включая теорию пределов, теорию интегрирования, теорию матриц);
- 5) методы статистической метеорологии и климатологии;
- 6) языки программирования C++ и FORTRAN для написания вычислительных программ.

Основные положения, выносимые на защиту:

- 1) алгоритмы стохастического моделирования временных рядов и пространственно-временных полей комплексов метеорологических процессов, позволяющие учитывать их нестационарность и неоднородность;
- 2) алгоритмы стохастического моделирования временных рядов, пространственных и пространственно-временных полей биоклиматических индексов, позволяющие учитывать их нестационарность по времени и неоднородность по пространству;
- 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.

Достоверность полученных 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.

Апробация работы. Результаты, представленные в диссертации, докладывались и обсуждались на Общеинститутском научном семинаре Института вычислительной математики и математической геофизики СО РАН (ИВМиМГ СО РАН), на объединенном семинаре ИВМиМГ 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, в том числе на Международной конференции «Марчуковские научные чтения» (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), 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), 6th International Conference on Time Series and Forecasting (Испания, Гранада, 25-27 сентября 2019).

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, в том числе 22 работы в изданиях, зарегистрированных в базах данных Web of Science и/или Scopus, из которых

12 статей – в журналах из перечня ВАК. В опубликованных работах отражено основное содержание, результаты и выводы диссертационного исследования.

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

СОДЕРЖАНИЕ РАБОТЫ

В Главе 1 предложены алгоритмы моделирования некоторых классов случайных процессов и полей.

В §1.1 предложены две модификации известного алгоритма «по строкам и столбцам» моделирования двумерных однородных случайных полей с корреляционными функциями вида $r(x, y) = \exp\left(-(\alpha x^2 + \gamma y^2)\right)$.

Модификации основаны на специальной перестановке элементов моделируемого поля на втором шаге алгоритма «по строкам и столбцам» и на рандомизации параметра моделируемого поля, что позволяет существенно расширить класс корреляционных функций, для которых применим алгоритм «по строкам и столбцам». Показано, что с помощью предложенных модификаций можно моделировать двумерные случайные поля с корреляционной функцией

$$R(x, y) = \int_a^b f(z) \exp(-Dz) dz,$$

где $D = D(x, y) = (x - dy)^2 + \left(\frac{2d}{c} - d^2 + 1\right)y^2$, $f(z) > 0$, $z \in (a, b)$,

$0 < a < b$ – произвольная плотность одномерного распределения.

В §1.2 рассмотрены вопросы, связанные со стохастической интерполяцией стационарных и нестационарных случайных процессов.

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 η_{i-j} , $i, j \in \mathbb{Z}$. Случайный процесс η_t , $t \in \mathbb{R}$ определяется по правилу:

$$\eta_t = \xi_j$$

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

$$\eta_t = \xi_i$$

при $t = i, i \in \mathbb{Z}$. Показано, что справедливы леммы 1, 2.

Лемма 1. Процесс $\eta_t, t \in \mathbb{R}$ имеет плотность одномерного распределения $f(x)$.

Лемма 1. При $m = 2k, k \in \mathbb{Z}^+$ справедливы равенства

$$\text{corr}(\eta_t, \eta_p) = \begin{cases} \eta_{|i-n|}, & t = i, p = n, i, n \in \mathbb{Z}, \\ \sum_{j=i-k+1}^{i+k} \omega_j^t \eta_{|j-n|}, & t = i + h, p = n, i, n \in \mathbb{Z}, 0 < h < 1, \\ \sum_{l=n-k+1}^{n+k} \omega_l^p \eta_{|i-l|}, & t = i, p = n + g, i, n \in \mathbb{Z}, 0 < g < 1, ; \\ \sum_{j=i-k+1}^{i+k} \sum_{l=n-k+1}^{n+k} \omega_j^t \omega_l^p \eta_{|j-l|}, & t = i + h, p = n + g, \\ & i, n \in \mathbb{Z}, 0 < h, g < 1. \end{cases}$$

Дано также аналитическое описание корреляционной функции $\text{corr}(\eta_t, \eta_p)$ при $m = 2k + 1, k \in \mathbb{Z}^+$ и доказано следующее утверждение.

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

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

$$\psi_t = \begin{cases} \sum_{j=i-k+1}^{i+k} \omega_j^t \xi_j, & t = i + h, i \in \mathbb{Z}, 0 < h < 1, \\ \psi_t = \xi_i, & t = i, i \in \mathbb{Z}. \end{cases}$$

В связи с эти доказана следующая лемма.

Лемма 3. В рассмотренных условиях

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

$$\text{и } D\eta_t = D\psi_t \Leftrightarrow \forall i, j \in \mathbb{Z} \quad r_{|i-j|} = 1.$$

Отдельно рассмотрен случай, когда интерполируемый процесс $\dots, \xi_{-2}, \xi_{-1}, \xi_0, \xi_1, \xi_2, \xi_3 \dots$ является нестационарным с плотностями одномерных распределений $f_i(x)$ и корреляционной функцией $\text{corr}(\xi_i, \xi_j) = r_{ij}, i, j \in \mathbb{Z}$.

В §1.3 предложены алгоритмы моделирования условных негауссовских процессов с точечными и интервальными условиями.

Для численного моделирования на основе реальных данных условных негауссовских временных рядов $\vec{\zeta} = (\zeta_1, \zeta_2, \dots, \zeta_N)$, удовлетворяющих точечным условиям

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

предложен приближенный Алгоритм 1, основанный на методе нормализации. Здесь и далее для удобства предполагается, что $\Omega = \{1, 2, \dots, K\}, K < N$.

Алгоритм 1.

1. Пусть $\zeta_i^j, i=1, 2, \dots, N, j=1, 2, \dots, NY$ – выборка данных. Преобразуем эти данные по правилу

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

где $\Phi(x)$ – 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 $\vec{\eta} = (\vec{\eta}_{(1)}, \vec{\eta}_{(2)})$, удовлетворяющий условиям $\eta_i = c_i, i \in \Omega$. 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

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

Для моделирования вектора $\vec{\zeta}$, удовлетворяющего интервальным условиям

$$\zeta_i \in (a_i, b_i), i \in \Omega, -\infty \leq a_i, b_i \leq +\infty,$$

предложен Алгоритм 2, основанный на методе обратных функций распределения (МОФР) и методе исключения.

Алгоритм 2.

1. По данным $\zeta_i^j, i = 1, 2, \dots, N, j = 1, 2, \dots, NY$ оцениваем выборочную корреляционную матрицу $R_\zeta = (\rho_{ij})_{i,j=1}^N$ вектора $\vec{\zeta}$ и в рамках МОФР вычисляем корреляционную матрицу R'_ζ вспомогательного безусловного гауссовского процесса. Элементы ρ'_{ij} матрицы R'_ζ есть решения уравнений

$$\rho_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_i^{-1}(\Phi(y)) G_j^{-1}(\Phi(z)) \phi_\rho(y, z) dy dz,$$

$$\phi_\rho(y, z) = \left(2\pi \sqrt{1 - (\rho'_{ij})^2} \exp \left(\frac{y^2 + z^2 - 2\rho'_{ij} yz}{2(1 - (\rho'_{ij})^2)} \right) \right)^{-1}.$$

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

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 и повторяем шаг 2.

4. Трансформируем реализацию условного вектора

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

В §1.3 приведены результаты численных экспериментов, связанных с исследованием трудоёмкости Алгоритма 2, а также приближенный алгоритм с трудоёмкостью меньшей, чем у Алгоритма 2.

Глава 2 посвящена разработке, исследованию и верификации стохастических «генераторов погоды».

В §2.1 рассмотрены стохастические модели совместных нестационарных временных рядов метеорологических параметров. В п. 2.1.1 предложена стохастическая модель совместных нестационарных временных рядов трех метеорологических характеристик: индикатора наличия/отсутствия осадков в течение суток, минимальной и максимальной за сутки приземной температуры воздуха. Моделирование рядов температуры на временном интервале длительностью N суток осуществляется в предположении гауссовости их одномерных распределений, а для моделирования индикаторного ряда осадков $\vec{E}^T = (E_1, E_2, \dots, E_N)^T$ используется пороговое преобразование

$$E_j = \begin{cases} 1, & \zeta_j^E \leq c_j, \\ 0, & \zeta_j^E > c_j, \end{cases}$$

латентного нестационарного гауссовского процесса

$(\vec{\zeta}^E)^T = (\zeta_1^E, \zeta_2^E, \dots, \zeta_N^E)^T$. Коэффициенты корреляции $g(i, j)$ процесса

$(\vec{\zeta}^E)^T$ определяются по известным коэффициентам корреляции $r(i, j)$

индикаторного процесса и вероятностям p_j наличия осадков в j -е сутки $(i, j = 1, 2, \dots, N)$ на основании следующей леммы.

Лемма 4. Если $i \neq j$, то $g(i, j)$ есть решение уравнения

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

где $T(c, a)$ – функция Оуэна

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

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

а пороговое значение c_j определяется из равенства

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c_j} \exp\left(-\frac{t^2}{2}\right) dt = p_j.$$

Предложенная модель успешно прошла верификацию по реальным данным с метеостанций, расположенных в различных климатических зонах на территории РФ. В качестве иллюстрации, в таблице 1 приведены оценки среднего числа AN_0 дней в месяце с минимальной температурой за сутки

ниже и максимальной выше 0°C . Здесь и далее σ – среднеквадратическое отклонение оценки рассматриваемой характеристики по выборке реальных данных имеющегося объема, РД – оценки по реальным, а МД – по модельным данным.

В п. 2.1.2 предложены 2 подхода к моделированию совместных нестационарных негауссовских временных рядов различных комбинаций метеорологических параметров, например, временного ряда

Таб. 1. Оценки AN_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

$$(T_1, T_2, \dots, T_{8d}, H_1, H_2, \dots, H_{8d}, P_1, P_2, \dots, P_{8d})^T$$

температуры воздуха (измеряемой в градусах Цельсия), его относительной влажности (в долях единицы) и атмосферного давления (единица измерения – гПа) на временном интервале длительностью d суток с 3-часовым разрешением (при таком разрешении определяется 8 значений в сутки каждого из метеопараметров).

В рамках первого подхода, для учёта суточного хода реальных метеопроцессов использовано предположение о периодической коррелированности временных рядов с периодом равным 1 суткам.

Определение 1. Случайный процесс $\xi(t)$, $t \in \mathbb{R}$ называют периодически коррелированным случайным процессом (ПКСП), если существует такое число $T > 0$, что математическое ожидание $E\xi(t)$, дисперсия $D\xi(t)$ и корреляционная функция $\text{corr}(\xi(t_1), \xi(t_2))$ процесса $\xi(t)$ удовлетворяют условиям

$$\forall t \quad E\xi(t+T) = E\xi(t), \quad D\xi(t+T) = D\xi(t),$$

$$\forall t_1, t_2 \quad \text{corr}(\xi(t_1+T), \xi(t_2+T)) = \text{corr}(\xi(t_1), \xi(t_2)).$$

Предположение о периодической коррелированности рассматриваемых временных рядов было использовано как при оценивании по реальным данным корреляционной матрицы

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

совместных рядов с блочно-теплицевыми подматрицами R_X, R_{XY} , $X, Y \in \{T, H, P\}$, так и при определении параметров плотностей $t_k(x), h_k(x), p_k(x)$, $k = 1, 2, \dots, 8d$ одномерных распределений температуры

воздуха, его относительной влажности и атмосферного давления, соответственно. При построении модели предполагалось, что

- плотности $t_k(x)$ есть смеси

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

двух нормальных распределений,

- плотности $h_k(x)$ распределения относительной влажности в зависимости от метеостанции, данные с которой были использованы, суть плотности бета-распределения, смеси бета-распределений или смеси усеченных на $[0,1]$ нормальных распределений,
- плотности $p_k(x)$ – смеси двух усеченных на $(0,+\infty)$ нормальных распределений.

В рамках второго подхода, была построена модель нестационарных совместных рядов (НР) метеопараметров без использования предположения о периодической коррелированности временных рядов отдельных метеопараметров. При построении модели использовались плотности распределения такого же вида, как и в случае модели ПКСП, а их параметры изменялись не только внутри суток, но и от суток к суткам. Корреляционная матрица процесса также оценивалась по реальным данным, однако её подматрицы R_X, R_{XY} , $X, Y \in \{T, H, P\}$ не блочно-тёплицевы.

При реализации обеих моделей, численное моделирование негауссовских временных рядов с заданной корреляционной матрицей плотностями одномерных распределений осуществлялось методом обратных функций распределения. Было проведено сравнение двух моделей по трудоёмкости, показано, что справедлива лемма 5.

Лемма 5. Трудоёмкость s_1 и s_2 численной реализации МОФР для моделирования нестационарного процесса и моделирования ПКСП, соответственно, связаны соотношением

$$\frac{s_1}{s_2} = \frac{12d(24d-1)}{360d} \approx 0.8d.$$

Обе модели прошли верификацию. В таблице 2 приведен пример характеристики, использованной для верификации обеих моделей. Показано, что модель ПКСП, в отличие от модели нестационарного процесса, применима только на относительно коротких временных интервалах длительностью d суток (в пределах 11-15 суток).

В §2.2 предложен стохастический «генератор погоды», предназначенный для моделирования пространственно-временного поля

Таб. 2. Оценки вероятностей p_j того, что хотя бы 1 раз за рассматриваемый временной интервал длительностью d дней температура была ниже $lt^{\circ}C$, а относительная влажность – выше lh . Сочи, 1-10 февраля.

$lt^{\circ}C$	lh	РД, $p_j \pm \sigma$	МД (ПКСП)	МД (нестационарный процесс)
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

$$\vec{\xi} = \left(\vec{\xi}_1^T, \dots, \vec{\xi}_{ND}^T \right)^T;$$

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

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

приземной температуры воздуха T_{ji}^k , ортогональных компонент $u_{ji}^k, v_{ji}^k, j=1,2,\dots,8$ вектора скорости ветра с 3-часовым разрешением и полусуточного количества $P_{li}^k, l=1,2$ осадков на временном интервале длительностью ND суток на сети из $NS=8$ метеорологических станций, расположенных в южной части Байкальской природной территории (БПТ). Поле моделируется в предположении о его пространственной неоднородности и временной периодической коррелированности с периодом равным 1 суткам. Предполагается также, что одномерные распределения $T_{ji}^k, u_{ji}^k, v_{ji}^k$ суть смеси нормальных величин, функция одномерного распределения осадков имеет вид

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

со специально выбранными параметрами, а совместная корреляционная матрица поля $\vec{\xi}$ блочно-тёплицева и оценена по данным многолетних наблюдений на метеостанциях. Для моделирования негауссовского поля $\vec{\xi}$ были применены метод нормализации и итерационный алгоритм, предложенный в работе 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. Приведены результаты сравнения использованных алгоритмов и результаты верификации построенной модели реального поля комплекса метеопараметров. В качестве примера, на рисунке 1 приведены полученные по реальным и модельным данным оценки вероятности

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

того, что средняя за первую половину суток температура воздуха T_{cp}^8 превышает заданный уровень lt °C и средний за эти же полусутки модуль скорости ветра $wind_{cp}^8 = \sqrt{(u_{cp}^8)^2 + (v_{cp}^8)^2}$ больше $lw = 2$ м/с на метеостанции Кяхта.

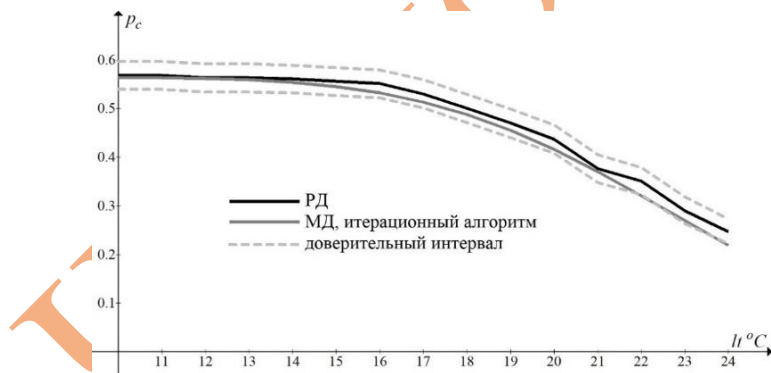


Рис. 1. Оценки вероятностей p_c . 10-20 июля.

В **Главе 3** приведены результаты разработки, исследования, верификации и сравнения различных моделей временных рядов, пространственных и пространственно-временных полей ряда биоклиматических индексов.

В §3.1 даны определения рассматриваемых в диссертации биоклиматических индексов. Так, например, по определению, индекс жары HI °C есть функция приземной температуры воздуха T °C и его относительной влажности H , измеряемой в долях единицы:

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

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

В диссертации рассмотрены 2 подхода к моделированию биоклиматических процессов, суть которых схематично изображена на рисунке 2.

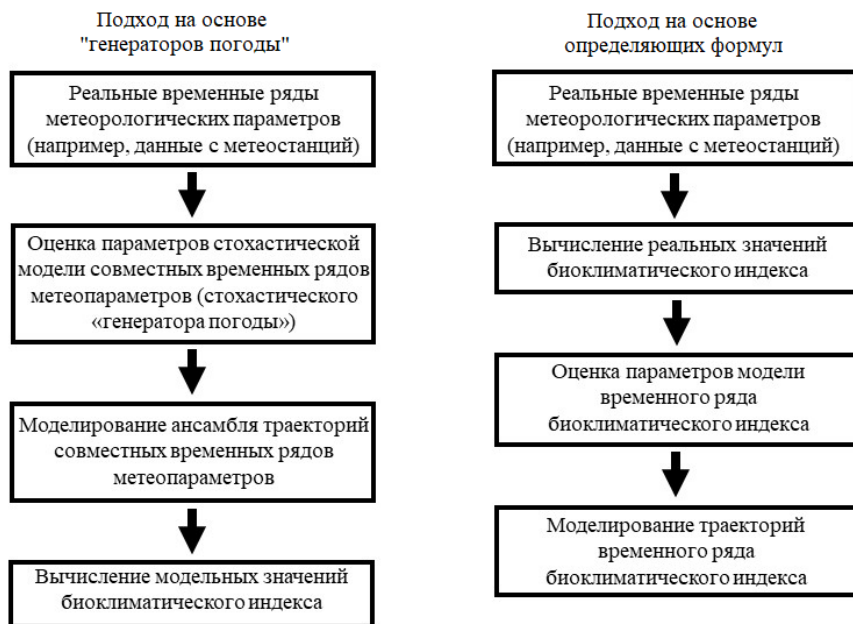


Рис. 2. Блок-схемы 2 подходов к моделированию рядов биоклиматических индексов.

В §3.2 предложены и верифицированы модели временных рядов биоклиматического индекса суровости метеорежима (БИСМ), индекса жары (ИЖ) и энтальпии влажного воздуха (ЭВВ), основанные на использовании предложенных в Главе 2 стохастических «генераторов погоды». В качестве примера, ниже приведён алгоритм моделирования временных рядов ЭВВ с трёхчасовым разрешением на интервале длительностью d суток.

Алгоритм 3.

- 1) По данным многолетних наблюдений на метеостанции вычисляется среднее значение P атмосферного давления на рассматриваемом временном интервале длительностью d суток.
- 2) Моделируется реализация

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

предложенного в §2.1 стохастического «генератора погоды», т.е. реализация совместного периодически коррелированного временного ряда значений температуры воздуха и его относительной влажности.

3) Построенная реализация вектора $\left(\bar{T}^T, \bar{H}^T\right)^T$ и значение P используются

для вычисления значений вектора $\bar{J}^T = (J_1, J_2, \dots, J_{8d})^T$ ЭВВ по формуле

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

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

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

где L, c_c, c_v – удельная теплота парообразования, удельная теплоемкость сухого воздуха и удельная теплоемкость водяного пара, соответственно.

На рисунке 3 приведены оценки коэффициентов корреляции $\text{corr}(J_1, J_{1+h})$ и $\text{corr}(J_i, J_{i+1})$, полученные по реальным и модельным данным.

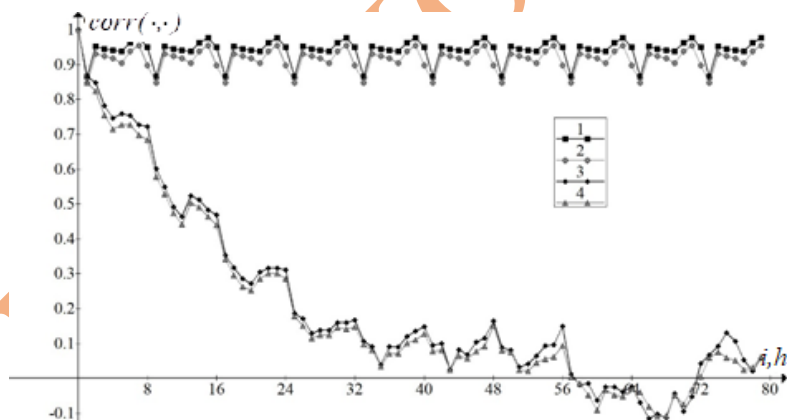


Рис. 3. Оценки коэффициентов корреляции. Кривые 1 и 2 – оценки $\text{corr}(J_i, J_{i+1})$ по реальным и модельным данным, соответственно; кривые 3, 4 – оценки $\text{corr}(J_1, J_{1+h})$ по реальным и модельным данным. Томск, 1-10 июля.

Подход к моделированию рядов БИ на основе «генераторов погоды» позволил численно исследовать зависимость оценок различных характеристик рядов ЭВВ и ИЖ от входных параметров моделей.

В §3.3 рассмотрен подход к моделированию временных рядов БИ, основанный на применении определяющих формул. Этот подход не включает в себя шага, на котором моделируются совместные временные ряды метеорологических параметров. В частности, построены модели временных рядов среднесуточного ИЖ и БИСМ, рядов индекса холодового стресса (ИХС) с 3- и 12-часовым разрешением. Так, например, временные ряды $\vec{B} = (B_1, B_2, \dots, B_{365})^T$ среднесуточного БИСМ на интервале годовой длины моделируются методом обратных функций распределения как нестационарный негауссовский процесс с корреляционной матрицей $\tilde{R}$, оцененной по реальным данным, и с плотностями одномерных распределений $s_i(x)$, $i = 1, 2, \dots, 365$ вида

$$s_i(x) = \theta_i \frac{1}{10B(\alpha_{1i}, \beta_{1i})} (x/10)^{\alpha_{1i}-1} (1-x/10)^{\beta_{1i}-1} + \\ + (1-\theta_i) \frac{1}{10B(\alpha_{2i}, \beta_{2i})} (x/10)^{\alpha_{2i}-1} (1-x/10)^{\beta_{2i}-1}, \\ 0 \leq \theta_i \leq 1, x \in [0, 10], i = 1, 2, \dots, 365,$$

которые в смысле критерия Пирсона хорошо аппроксимируют выборочные одномерные распределения $\tilde{s}_i(x)$, $i = 1, 2, \dots, 365$. В качестве иллюстрации того, насколько хорошо предложенная модель воспроизводит свойства реального временного ряда, на рисунке 4 приведены полученные по реальным и модельным данным оценки среднего минимального $Amin(i)$ и среднего максимального $Amax(i)$ БИСМ за месяц. Отклонения оценок этих характеристик по модельным данным от оценок по реальным данным для 16 из 18 рассмотренных метеостанций не превышают соответствующих значений σ (использованы данные за 1966-2015 гг.).

Было проведено сравнение основанных на определяющих формулах и на «генераторах погоды» моделей рассмотренных БИ по точности и трудоёмкости моделирования. Кроме этого, предложенные модели были использованы для исследования изменчивости различных характеристик временных рядов БИ во времени, и показано, например, что климатические условия в смысле БИСМ в 1991-2015 гг. были несколько мягче, чем в 1966-1990 гг.

С использованием предложенных в Главе 1 алгоритмов моделирования условных негауссовских процессов с точечными и интервальными условиями, разработанные модели временных рядов БИ были применены для их вероятностного прогнозирования. На основе Алгоритма 1 были построены реализации условных временных рядов $\vec{B} = (B_1, B_2, \dots, B_N)$ среднесуточного

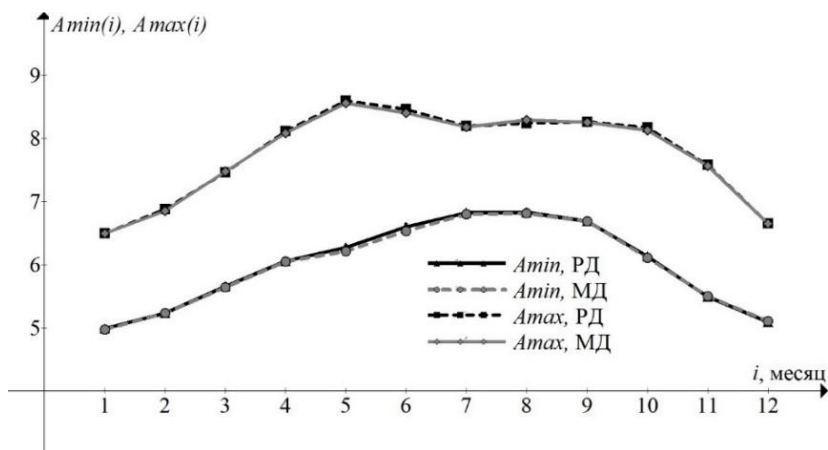


Рис. 4. Оценки $Amin(i)$ и $Amax(i)$. Пограничный.

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 приведены примеры оценок этих вероятностей. Фактически, в таблице 3 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. 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 января ($j = 11$)	0.000	0.691	0.309	0.000
12 января ($j = 12$)	0.188	0.718	0.095	0.000
15 января ($j = 15$)	0.290	0.681	0.029	0.000
16 января ($j = 16$)	0.608	0.388	0.005	0.000
19 января ($j = 19$)	0.024	0.331	0.621	0.024
20 января ($j = 20$)	0.004	0.308	0.658	0.030

В §3.4 предложены модели пространственных и пространственно-временных полей БИ. Предложенный в §2.2 стохастический «генератор погоды» использован для моделирования неоднородного по пространству и нестационарного по времени поля ИХС на сети метеостанций в южной части БПТ. В таблице 4 приведены оценки вероятности

$$p_{min} = P\left(\exists i \in \{1, 2, \dots, ND\} : \min W_i^k = \min\{W_{1,i}^k, W_{2,i}^k, W_{3,i}^k, W_{4,i}^k\} \in (a, b]\right)$$

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 $(a, b]$.

Таб. 4. Оценки вероятности p_{min} . 10-20 января.

$(a, b]$	Иркутск ($k=1$)		Хомутово ($k=2$)	
	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

Подход на основе определяющих формул был использован для моделирования пространственно-временного поля

$$\overline{HI} = \left(HI_1^a, \dots, HI_{NS}^a, HI_1^{a+1}, \dots, HI_{NS}^{a+1}, \dots, HI_1^{a+N-1}, \dots, HI_{NS}^{a+N-1} \right)^T$$

среднесуточного ИЖ на сети $NS=18$ расположенных на юге России метеостанций (см. рисунок 5) на временном интервале длительностью N дней. Рассматриваемый временной интервал $[a, a+N-1]$ принадлежит летнему периоду (с 1 июня по 31 августа).

Моделирование негауссовского неоднородного и нестационарного поля $\overline{HI}$ с заданными одномерными распределениями в виде смесей нормальных распределений и заданной корреляционной матрицей осуществлялось с помощью МОФР и методом нормализации (МН). Показано, что, в целом, построенные обоими методами траектории обладают свойствами близкими, к свойствам реальных траекторий рассматриваемого поля. Ряд характеристик, например, вероятность $p_N(lev)$ того, что на протяжении N суток среднесуточный ИЖ 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



Рис. 5. Расположение метеорологических станций.

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 °C, длительности N и начала временного интервала a . При этом, моделирование методом нормализации существенно менее трудоёмко (см. таблицу 5).

Таб. 5. Время моделирования 10^4 траекторий поля $\overline{HI}$, 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

Методом обратных функций распределения также проводилось моделирование пространственного поля $\overline{W}_s^T$ и пространственно-временного

поля $\overrightarrow{W}_{st}^T$ среднесуточного ИХС на сети из NS метеостанций, расположенных на юге Западной Сибири. Моделирование осуществлялось в предположении, что одномерные распределения среднесуточного ИХС суть смеси $g_i(x)$, $i=1, \dots, NS$ (для случая пространственно-временного поля – смеси $g_i^j(x)$, $i=1, \dots, NS$, $j=1, \dots, ND$, ND – количество дней) двух гауссовских распределений, а корреляционная структура поля задаётся выборочной корреляционной матрицей $R = \{r(i, j)\}_{i,j=1}^{NS}$ (для поля $\overrightarrow{W}_{st}^T$ – $R = \{r(i, j)\}_{i,j=1}^{NS \times ND}$). Предложенные модели успешно прошли верификацию.

Построенные модели полей $\overrightarrow{W}_s^T$ и $\overrightarrow{W}_{st}^T$ 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 связи с этим, в п. 3.4.2 было проведено сравнение различных детерминированных и недетерминированных методов интерполяции ИХС со станций в узлы сетки. Методом кросс-валидации было показано, что наименьшую среднеквадратическую ошибку интерполяции даёт рассмотренная в Главе 1 стохастическая интерполяция с NS ближайших метеостанции с вероятностями, обратно пропорциональными расстояниям от узла до метеостанций. В качестве примера применения расширенной модели пространственного поля ИХС на сетке, в таблице 6 приведены оценки вероятностей $P_{Bolot\ reg}(lev)$ того, что ИХС будет ниже заданного уровня $lev^{\circ}C$ на территории Болотнинского района Новосибирской области. На территорию этого района попал 31 узел сетки с шагом 10 км в обоих направлениях. Для сравнения приведены оцененные по реальным данным вероятности $P_{Bolot}(lev)$ того, что ИХС ниже заданного уровня $lev^{\circ}C$ на метеостанции Болотное.

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 различных типов. Кроме этого, исследованы некоторые свойства асимптотически периодически коррелированных случайных процессов, в том числе, бинарных марковских цепей с матрицами переходных

Таб. 6. Оценки вероятностей $P_{Bolot\ reg}(lev)$. 20 декабря.

$lev\ ^\circ C$	$P_{Bolot\ reg}(lev)$	$P_{Bolot}(lev)$
-5	0.953	0.964
-10	0.845	0.869
-15	0.569	0.676
-20	0.269	0.424
-25	0.101	0.214
-30	0.050	0.102
-35	0.011	0.060

вероятностей, являющимися периодическими функциями времени, и кусочно-постоянных процессов, построенных с использованием точечных потоков.

В §4.1 был рассмотрен 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 $i = [t]$, η_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 , дисперсией σ^2 и корреляционной функцией

$$corr(\eta_i, \eta_j) = corr(i, j) = corr(|i - j|) = corr(h) = r_h,$$

α – независимая от η_i , $i = -1, 0, 1, 2, \dots$ случайная величина с функцией одномерного распределения $F_\alpha(x)$, $x \in [0, 1]$. Показано, что для процесса ξ_t справедлива лемма 6.

Лемма 6. Математическое ожидание, дисперсия и корреляционная функция процесса ξ_t определяются равенствами

$$E\xi_t = \mu,$$

$$D\xi_t = \sigma^2 \left(1 - 2F_\alpha(\{t\}) + 2F_\alpha(\{t\})^2 \right) + 2\sigma^2 r_1 F_\alpha(\{t\})(1 - F_\alpha(\{t\})),$$

$$corr(\xi_t, \xi_p) = \frac{1}{Z_t Z_p} \left(\begin{aligned} & (F_\alpha(\{p\}) - F_\alpha(\{t\})F_\alpha(\{p\}))r_{|i-j-1|} + \\ & + (1 - F_\alpha(\{t\}) - F_\alpha(\{p\}) + 2F_\alpha(\{t\})F_\alpha(\{p\}))r_{|i-j|} + \\ & + (F_\alpha(\{t\}) - F_\alpha(\{t\})F_\alpha(\{p\}))r_{|i-j+1|} \end{aligned} \right),$$

где

$$Z_t = \sqrt{1 - 2F_\alpha(\{t\}) + 2F_\alpha(\{t\})^2 + r_1 2F_\alpha(\{t\})(1 - F_\alpha(\{t\}))},$$

$$Z_p = \sqrt{1 - 2F_\alpha(\{p\}) + 2F_\alpha(\{p\})^2 + r_1 2F_\alpha(\{p\})(1 - F_\alpha(\{p\}))}.$$

В частности, если α распределена равномерно на отрезке $[0,1]$, то справедлива лемма 7.

Лемма 7. 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, а 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

$$\text{corr}(\xi_t, \xi_p) = \frac{1}{Z_t Z_p} \left(\left(\{p\} - \{t\} \{p\} \right) r_{|i-j-1|} + \left(1 - \{t\} - \{p\} + 2\{t\} \{p\} \right) r_{|i-j|} + \left(\{t\} - \{t\} \{p\} \right) r_{|i-j+1|} \right),$$

где

$$Z_t = \sqrt{1 - 2\{t\} + 2\{t\}^2 + r_1 2\{t\}(1 - \{t\})},$$

$$Z_p = \sqrt{1 - 2\{p\} + 2\{p\}^2 + r_1 2\{p\}(1 - \{p\})}.$$

Показано, что при произвольной функции $F_\alpha(x), x \in [0,1]$ справедливо следующее утверждение.

Утверждение 2. Случайный процесс ξ_t является периодически коррелированным с периодом равным 1.

В §4.1 был также рассмотрен случайный кусочно-линейный процесс ξ_t непрерывного времени $t \geq 0$, определяемый опорными точками $(-1 + \alpha; \eta_{-1}), (0 + \alpha; \eta_0), (1 + \alpha; \eta_1), (2 + \alpha; \eta_2), (3 + \alpha; \eta_3), \dots$, где $\eta_i, i = -1, 0, 1, 2, \dots$ рассмотренный выше стационарный процесс, а α – равномерно распределенная на $[0,1]$ случайная величина. Показано, что значение процесса ξ_t в произвольной точке $t, t \in [i, i+1), i = 0, 1, 2, \dots$ определяется равенством

$$\xi_t = \eta_{i-1} \left(\{t\}^2 - \{t\} \right) + \eta_{i-1} \alpha (1 - \{t\}) + \eta_i \left(-2\{t\}^2 + \{t\} + 1 \right) +$$

$$+ \eta_i \alpha (2\{t\} - 1) + \eta_{i+1} \{t\}^2 - \eta_{i+1} \alpha \{t\}$$

и справедливы лемма 8 и утверждение 3.

Лемма 8. Математическое ожидание процесса ξ_t не зависит от t и $E\xi_t = \mu$. Дисперсия рассматриваемого процесса

$$D\xi_t = \sigma^2 \left(6\{t\}^4 - 12\{t\}^3 + 6\{t\}^2 + \frac{2}{3} + \frac{1}{3}r_1(2\{t\} - 1)^2(-6\{t\}^2 + 6\{t\} + 1) + \right. \\ \left. + \frac{2}{3}r_2\{t\}(3\{t\}^3 - 6\{t\}^2 + 4\{t\} + 1) \right)$$

является периодической функцией t с периодом равным 1.

Утверждение 3. Случайный процесс ξ_t является периодически коррелированным с периодом равным 1.

В §4.2 введено понятие асимптотически периодически коррелированного случайного процесса.

Определение 2. Случайный процесс ξ_t непрерывного или дискретного аргумента t называется асимптотически периодически коррелированным случайным процессом (АПКСП) с периодом T , если существуют такие периодические с периодом T функции $m(t), v(t), c(t, t+h)$ аргумента t , что

$$\forall \varepsilon > 0 \exists T_1 : \forall t > T_1 \quad |E\xi(t) - m(t)| < \varepsilon, \quad |D\xi(t) - v(t)| < \varepsilon,$$

$$\forall h \quad |corr(\xi(t), \xi(t+h)) - c(t, t+h)| < \varepsilon.$$

Показано, что для АПКСП справедливы следующие утверждения.

Утверждение 4. Если ξ_t – АПКСП, то

$$\forall \varepsilon > 0 \exists T_1 : \forall t > T_1 \quad \forall k \in \mathbb{Z}^+$$

$$|E\xi(t) - E\xi(t+kT)| < \varepsilon, \quad |D\xi(t) - D\xi(t+kT)| < \varepsilon, \quad (*)$$

$$\forall h \quad |corr(\xi(t), \xi(t+h)) - corr(\xi(t+kT), \xi(t+kT+h))| < \varepsilon.$$

Утверждение 5. Для того чтобы процесс ξ_t дискретного аргумента t являлся АПКСП с периодом T , необходимо и достаточно, чтобы существовало такое $T > 0$, при котором выполняется условие (*).

В §4.3 рассмотрена бинарная марковская цепь $\xi_t, t = 0, 1, 2, \dots$ с множеством значений $\{0, 1\}$, начальным вектором распределений $A = (a, 1-a)$ и матрицей переходных вероятностей $\Pi(t)$, являющейся периодической функцией t , принимающей значения

$$P_k = \begin{pmatrix} p_k & 1-p_k \\ 1-q_k & q_k \end{pmatrix}$$

при $t = un + k > 0$, где $n \geq 2$ – период функции $\Pi(t)$, $u = 0, 1, 2, \dots$, $k = 0, 1, \dots, n-1$. Приведены аналитические выражения для функции одномерного распределения марковской цепи ξ_t и её корреляционной функции. Показано, что справедливо утверждение 6.

Утверждение 6. Процесс ξ_t является АПКСП с периодом n .

Кроме этого, исследованы различные вероятностные характеристики серий постоянных значений цепи заданной длины. Так, например, показано, что вероятность начала серии из k единиц в момент t при $t = mn + w$, $k = un + v$, $1 \leq w < n$, $1 \leq v < n$, $w + v \leq n$ определяется равенством

$$P(L(t, k)) = \left(\prod_{i=0}^{w-1} p_i^{u-1} \right) \left(\prod_{j=w}^{n-1} p_j^u \right) \left(\prod_{l=0}^{w+v-2} p_l \right) (1 - p_{w+v-1}).$$

Рассмотренная бинарная марковская цепь использована в §4.4 для конструктивного определения кусочно-постоянного случайного процесса непрерывного времени. Пусть η_i , $i = 0, 1, 2, \dots$ – цепь Маркова с описанными выше свойствами. Кусочно-постоянный процесс ξ_t , $t \geq 0$ непрерывного аргумента t построен согласно Алгоритму 4.

Алгоритм 4.

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 η_i , $i = 0, 1, 2, \dots$ случайные величины с функцией одномерного распределения $F(x)$, математическим ожиданием μ , дисперсией σ^2 и $|E\alpha_i^2| < +\infty$.

Показано, что справедливо утверждение 7.

Утверждение 7. Процесс ξ_t является АПКСП с периодом n .

В §4.4 рассмотрен ещё один пример асимптотически периодически коррелированного кусочно-постоянного процесса.

Пусть $\lambda(t) > 0$ – ограниченная по модулю периодическая функция с периодом T , а точечный поток $\{\tau_k\}$ определён по правилу

$$\tau_0 = 0, \tau_k = \sum_{i=1}^k \eta_i, k=1,2,\dots,$$

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 $\xi_t, t \geq 0$ построен согласно Алгоритму 5.

Алгоритм 5.

- 1) Моделируется последовательность $\{\tau_k\}$.
- 2) Для каждого интервала $[\tau_{i-1}; \tau_i), i=1,2,\dots$ значение процесса ξ_t определяется равенством $\xi_t = \nu_i$, где ν_i – независимые случайные величины с заданной функцией одномерного распределения $F(x)$ и конечным вторым моментом.

Показано, что корреляционная функция процесса ξ_t определяется равенствами

$$\text{corr}(\xi_t, \xi_{t+h}) = \exp(-\lambda(0)[t+h]) + \sum_{k=1}^{\infty} P(\tau_k \leq t, t+h < \tau_{k+1})$$

и получены аналитические выражения для $P(\tau_k \leq t, t+h < \tau_{k+1}), k \geq 1$. Численные эксперименты показали, что при различных функциях $\lambda(t)$ процесс ξ_t является АПКСП с периодом T . В качестве иллюстрации, на рисунке 6 приведены графики корреляционной функции $\text{corr}(\xi_t, \xi_{t+h})$ при $F(x) = \Phi_{0,1}, \lambda(t) = 0.9 \sin 2t + 1$.

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 в Главах 2 и 3 для определения параметров и верификации стохастических моделей метеорологических и биоклиматических процессов.

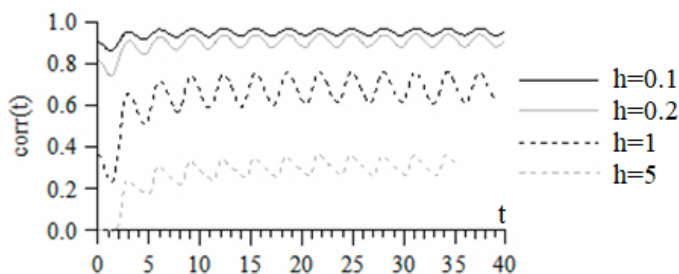


Рис. 6. Корреляционная функция $\text{corr}(\xi_t, \xi_{t+h})$ процесса ξ_t при фиксированных значениях h .

В **Приложении 3** представлены некоторые результаты верификации моделей, предложенных в §3.3, §3.4 Главы 3 диссертационной работы.

ЗАКЛЮЧЕНИЕ

Сформулируем основные результаты диссертационной работы:

- 1) 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;
- 2) 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 позволяют учитывать и воспроизводить нестационарность по времени и неоднородность по пространству реальных процессов;
- 3) разработаны алгоритмы моделирования некоторых классов периодически коррелированных и асимптотически периодически коррелированных кусочно-постоянных и кусочно-линейных случайных процессов, алгоритмы моделирования двумерных случайных полей с корреляционными функциями специального вида и метод стохастической интерполяции случайных процессов.

ОСНОВНЫЕ ПУБЛИКАЦИИ АВТОРА ПО ТЕМЕ ДИССЕРТАЦИИ

1. 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 (Web of Science: без кватилия, Scopus: Q3)
2. **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> (WoS: Q3, Scopus: Q2)
 3. **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> (WoS: без кватилия, Scopus: Q3)
 4. **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> (WoS: Q4, Scopus: Q2)
 5. **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 (WoS: Q2, Scopus: Q1)
 6. **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> (WoS: Q3, Scopus: Q2)
 7. **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> (WoS: Q3, Scopus: Q3)
 8. **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> (WoS: Q3, Scopus: Q2)
 9. **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> (WoS: Q3, Scopus: Q2)
 10. **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> (WoS: Q3, Scopus: Q2)

11. 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> (WoS: Q3, Scopus: Q2)
12. Ogorodnikov V.A., Sereseva 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> (WoS: Q3, Scopus: Q2)

Каргаполова Нина Александровна

**Алгоритмы численного стохастического моделирования
нестационарных метеорологических и биоклиматических процессов**

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
доктора физико-математических наук