

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра математической кибернетики и компьютерных наук

**РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ АНАЛИЗА
ОСОБЕННОСТЕЙ ВЗАИМОСВЯЗИ КЛИМАТИЧЕСКИХ ЯВЛЕНИЙ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 451 группы
направления 09.03.04 — Программная инженерия
факультета КНиИТ
Антонова Виктора Владимировича

Научный руководитель
доцент, к. ф.-м. н.

Ю. Н. Кондратова

Заведующий кафедрой
к. ф.-м. н.

С. В. Миронов

Саратов 2019

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Основное содержание работы	5
ЗАКЛЮЧЕНИЕ	11
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	13

ВВЕДЕНИЕ

В XXI веке остро встает проблема глобальных климатических изменений. Поэтому в последние годы ей уделяется большое внимание. Межправительственная группа экспертов по изменению климата (МГЭИК) делает следующие выводы: одной из причин глобальных изменений климата в течение последнего столетия является быстрое развитие промышленности, которое способствует увеличению выбросов в атмосферу. Продукты сжигания углеродного топлива приводят к увеличению концентрации углекислого газа в воздушной оболочке Земли. Какие изменения климата связаны с этими процессами, пока предсказать не возможно. Единственным способом получения более или менее достоверных оценок будущих климатических изменений является построение модели климатической системы.

Однако в настоящее время, в веку постиндустриального общества, развитие широкого класса современных теоретических и прикладных областей научного исследования, связанных с климатом, стало в полной мере возможно только благодаря появлению компьютерного моделирования процессов в атмосфере. Примерами наиболее важных областей исследования, в которых компьютерное моделирование атмосферы играет основную роль, являются задачи численного прогноза погоды, изменения климата и вопросы, связанные с загрязнением окружающей среды и других компонентов в атмосфере. Необходимость использования ИТ при этом обусловлена целым набором причин. Одной из таких причин является решение уравнений, описывающие различные процессы в атмосфере. На основании этого возникла необходимость в создании компьютерных систем моделирования атмосферы.

Современные компьютерные системы атмосферного моделирования — это досконально проработанные комплексы, которые имеют в себе широкий спектр применений. Более точный прогноз для конкретной территории позволяет минимизировать убытки для экономики.

Погодные явления оказывают значительное влияние на состояние не только экономики страны, но и отдельные её регионы. В связи с этим управление гидрометеорологическими рисками является условием обеспечения гидрометеорологической безопасности и должно осуществляться как на микро-, так и на макроэкономическом уровне.

Актуальность настоящей работы определяется необходимостью изуче-

ния климатических изменений для уменьшения гидрометеорологических рисков в экономике.

Исследование связи между крупномасштабными модами климатической изменчивости необходимо для адекватной оценки роли различных естественных и антропогенных факторов в формировании изменений глобального климата. К наиболее значимым на глобальном уровне климатическим процессам относятся явления Эль-Ниньо — южное колебание (ЭНЮК), Ла-Нинья — это процесс обратный по отношению к Эль-Ниньо, североатлантическое колебание, арктическая осцилляция, тихоокеанское колебание и другие [1]. В данной работе акцент сделан на анализе взаимосвязи важнейших естественных климатических процессов в бассейнах Тихого, Атлантического и Индийского океанов: ЭНЮК — в экваториальных широтах, индийского муссона (ИМ) — в тропических и субтропических широтах и САК — в средних и более высоких широтах. Значимые климатические эффекты, связанные с этими процессами, проявляются в самых разных регионах Земли [2]– [3].

Цель данной работы — исследование направленных связей между процессами ЭНЮК, ИМ и САК методами анализа временных рядов.

Для реализации поставленной цели необходимо решить следующие задачи:

1. изучение методов установления взаимосвязи между парами климатических процессов;
2. разработка программного комплекса.

Бакалаврская работа содержит следующие разделы:

- определения, обозначения и сокращения;
- введение;
- построение математической модели;
- реализация программного комплекса;
- заключение;
- 20 источников;
- 3 приложения.

1 Основное содержание работы

Математический аппарат

Были проанализированы взаимосвязи трех климатических процессов: ЭНЮК (процесс, происходящий в экваториальных широтах), ИМ (индийский муссон, процесс, происходящий в тропических широтах) и САК (среднеарктическое колебание, процесс, происходящий в широтах выше субтропических) [4]. Эти три процесса рассматривались как три временных ряда.

Существуют следующие методы анализа временных рядов:

- корреляционный анализ позволяет выявить существенные периодические зависимости и их лаги (задержки) внутри одного процесса (автокорреляция) или между несколькими процессами (кросскорреляция) [5];
- спектральный анализ позволяет находить периодические и квазипериодические составляющие временных рядов [6];
- модели авторегрессии и скользящего среднего оказываются особенно полезными для описания и прогнозирования процессов, проявляющих однородные колебания вокруг среднего значения [7];

В данной работе использовались однокомпонентная и двухкомпонентная авторегрессия с последующим применением причинности по Грэнджеру, так как нам необходимо было строить предсказания значений без учета посторонних факторов. Однокомпонентная авторегрессия позволяет анализировать значения в отдельном регионе, а двухкомпонентная — понять, существует ли связь между ними в разных регионах и таким способом скорректировать предсказания. Причинность по Грэнджеру в свою очередь требует и однокомпонентную, и двухкомпонентную регрессии, и является основным методом определения взаимосвязей [7, 8].

Кроме того, для более точного определения взаимосвязей процессов в совокупности использовалась модель трехкомпонентной авторегрессии, которая позволила выявить зависимость одного процесса от другого даже с учетом влияния третьего процесса, искажающего или нейтрализующего влияние второго процесса. Для оценки влияния по трехкомпонентной модели также использовалась причинность по Грэнджеру [7].

Заданные временные ряды представляют собой значения колебаний климатических процессов, упорядоченные по годам и месяцам. Так как эти процессы подвержены сезонным изменениям (в особенности процессы, происхо-

дящие широтах, удаленных от экватора, а именно САК и в меньшей степени ИМ), временные ряды были обработаны следующим образом: было выбрано несколько периодов, анализ которых нас интересовал (один из примеров такого периода — теплое полугодие, то есть промежуток с апреля по сентябрь), и значения в одном периоде были заменены на их среднее арифметическое по каждому году. Ряды средних арифметических значений климатических процессов в определенный период года анализировались вышеописанными методами авторегрессии.

Для анализа временных рядов использовались модели однокомпонентной, двухкомпонентной и трехкомпонентной авторегрессии.

Для заданного временного ряда a_1, a_2, \dots, a_n мы определяем линейную авторегрессию k -го порядка следующим образом:

$$\forall i \in [k+1, n] \ a_i = f(a_{i-1}, a_{i-2}, \dots, a_{i-k}), \quad (1)$$

где f — некоторая линейная комбинация со свободным коэффициентом, т. е. $a_i = a_{i-1}c_1 + a_{i-2}c_2 + \dots + a_{i-k}c_k + x$ [9].

Определим двухкомпонентную авторегрессию. Пусть даны два временных ряда a_1, a_2, \dots, a_n и b_1, b_2, \dots, b_n . Двухкомпонентная регрессия — это зависимость значений a_i от предыдущих значений a_i и b_i одновременно. Если порядок регрессии по первой компоненте d_1 , а по второй d_2 , то мы хотим приблизить $\forall i \in [\max(d_1, d_2), n] \ a_i = f(a_{i-1}, a_{i-2}, \dots, a_{i-d_1}, b_{i-1}, b_{i-2}, \dots, b_{i-d_2})$, т. е.,

$$a_i = c_1a_{i-1} + c_2a_{i-2} + \dots + c_{d_1}a_{i-d_1} + e_1b_{i-1} + e_2b_{i-2} + \dots + e_{d_2}b_{i-d_2} + x. \quad (2)$$

Двухкомпонентная авторегрессия позволяет определить влияние одного процесса на другой, однако если существует третий процесс, влияние которого либо более сильно выражено, чем влияние второго процесса, либо его (частично или полностью) нейтрализует, для того, чтобы распознать влияние, необходимо было исследовать все три процесса в совокупности.

Для этого ввели понятие трехкомпонентной авторегрессии.

Пусть даны три временных ряда:

1. a_1, a_2, \dots, a_n ;

2. b_1, b_2, \dots, b_n ;

3. t_1, t_2, \dots, t_n .

Трехкомпонентная авторегрессия — это зависимость значений a_i от предыдущих значений a_i , b_i и t_i одновременно. Если порядок регрессии по первой компоненте d_1 , по второй d_2 , а по третьей d_3 то мы хотим приблизить $\forall i \in [\max(d_1, d_2), n]$

$a_i = f(a_{i-1}, a_{i-2}, \dots, a_{i-d_1}, b_{i-1}, b_{i-2}, \dots, b_{i-d_2}, t_{i-1}, t_{i-2}, \dots, t_{i-d_3})$, т. е.,

$$a_i = c_1 a_{i-1} + c_2 a_{i-2} + \dots + c_{d_1} a_{i-d_1} + e_1 b_{i-1} + e_2 b_{i-2} + \dots + e_{d_2} b_{i-d_2} + \\ + k_1 t_{i-1} + k_2 t_{i-2} + \dots + k_{d_3} t_{i-d_3} + x. \quad (3)$$

В случае, если порядок авторегрессионных моделей не позволяет подобрать параметры для точного выполнения вышеописанных равенств, будет минимизироваться сумма квадратов ошибок модели. Для этого строится система линейных алгебраических уравнений, которая решается методом Гаусса.

Для подбора оптимального порядка модели используются статистические критерии:

- для однокомпонентной модели — байесовский информационный критерий;
- для двухкомпонентной и трехкомпонентной моделей вычисляется влияние побочной компоненты на главную с использованием причинности по Грэнджеру; при помощи критерия Фишера вычисляется вероятность ошибочного вывода о наличии влияния, и порядок модели выбирается так, чтобы вероятность ошибки была минимальна.

Реализация программного комплекса и результаты работы

Программный комплекс для решения поставленной задачи был реализован на языке C++ в среде разработки Microsoft Visual Studio 2019. Были реализованы функции, решающие СЛАУ, строящие регрессионные модели (однокомпонентные, двухкомпонентные и трехкомпонентные) с заданным порядком, а также самостоятельно подбирающие оптимальный порядок. При построении двухкомпонентных и трехкомпонентных моделей вычисляется вероятность влияния побочной компоненты на главную по Грэнджеру. Также были реализованы функции, строящие предсказания следующих значений времен-

ных рядов по всем видам регрессионных моделей.

На языке C++ с использованием вышеописанных функций было реализовано приложение, данные в которое вводятся из текстовых файлов, а управление осуществляется при помощи консольного ввода-вывода. Текстовые файлы с данными имеют структуру таблицы, в каждой строке которых записан номер года и 12 показателей (отдельно по каждому месяцу). Программа поддерживает возможность также анализировать не отдельные месяцы, но и периоды из нескольких месяцев (сезоны и полугодия) в зависимости от настроенного количества месяцев в периоде и смещения. Например, для анализа показателей по теплому и холодному полугодиям (теплое — с апреля, т. е. 4 месяца, по сентябрь, т. е. по 9 месяц) необходимо настроить, что анализируемые периоды состоят из 6 месяцев, и надо пропустить первые 3 месяца (чтобы отсчет начинался с апреля) (рис. 1).

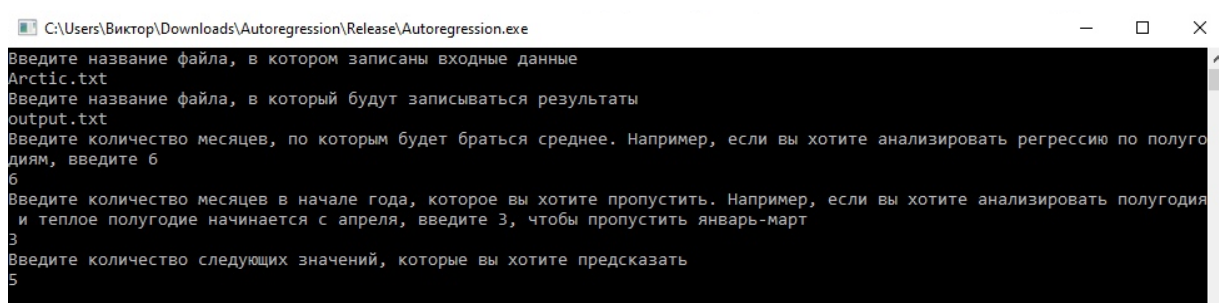


Рисунок 1 – Пример взаимодействия с программой.

Проанализировали вероятность влияния климатических процессов друг на друга по полугодиям при помощи двухкомпонентной и трехкомпонентной модели. Для этого разделили каждый год на два полугодия: теплое (месяца 4, 5, 6, 7, 8 и 9) и холодное (все остальные месяца) и усреднили значения в каждом из этих полугодий, после чего построили авторегрессионные модели отдельно по каждому полугодию.

Индекс ЭНЮК — аномалии температуры поверхности в экваториальной области Тихого океана, индекс САК — аномалии разности нормированных значений давления между северо-атлантическими центрами действия — азорским антициклоном и исландским циклоном, индекс активности ИМ — аномалии количества осадков над Индией.

Таблица 1 – Вероятности влияния климатических процессов в теплое полугодие по двухкомпонентной авторегрессии. Столбец — оказывает влияние, строка — зависимый процесс.

	ИМ	САК	ЭНЮК
ИМ	-	0.552527	0.752247
САК	0.91977	-	0.764789
ЭНЮК	0.982705	0.120845	-

Таблица 2 – Вероятности влияния климатических процессов в теплое полугодие по трехкомпонентной авторегрессии. Столбец — оказывает влияние, строка — зависимый процесс.

	ИМ	САК	ЭНЮК
ИМ	-	0.67471	0.788088
САК	0.951318	-	0.859109
ЭНЮК	1	0.0149046	-

В целом результаты, полученные по трехкомпонентной авторегрессионной модели, сходятся с теми, которые были получены по двухкомпонентной модели. Однако трехкомпонентная авторегрессия позволяет обнаруживать влияние процессов в том случае, если на процесс одновременно влияют оба других процесса, как, например, в случае с САК: на него оказывает влияние как ЭНЮК, так и ИМ, но двухкомпонентная авторегрессионная модель не может обнаружить влияние ЭНЮК, не учитывая оба фактора.

Таблица 3 – Вероятности влияния климатических процессов в холодное полугодие по двухкомпонентной авторегрессии. Столбец — оказывает влияние, строка — зависимый процесс.

	ИМ	САК	ЭНЮК
ИМ	-	0.769392	0.516208
САК	0.575928	-	0.499448
ЭНЮК	0.373407	0.168395	-

Таблица 4 – Вероятности влияния климатических процессов в холодное полугодие по трехкомпонентной авторегрессии. Столбец — оказывает влияние, строка — зависимый процесс.

	ИМ	САК	ЭНЮК
ИМ	-	0.834231	0.521519
САК	0.635479	-	0.607324
ЭНЮК	0.648266	0.125001	-

Во время холодного полугодия влияние процессов друг на друга выражено слабее. Трехкомпонентная модель позволила более точно обнаружить влияние САК на ИМ, однако влияния других процессов друг на друга обнаружить не удалось ни двухкомпонентной, ни трехкомпонентной авторегрессионной моделью.

Кроме того, аналогичные методы были использованы для проведения сезонного анализа влияния процессов друг на друга.

Анализ отдельных сезонов позволяет находить зависимости более точно, так как при анализе по полугодиям показатели усредняются, что может исказить данные в случае, если влияние выражено только в какую-то определенную часть полугодия.

ЗАКЛЮЧЕНИЕ

В данной работе рассмотрен математический аппарат для установления взаимосвязи между климатическими явлениями. Для определения влияния были использованы двухкомпонентные и трехкомпонентные регрессионные модели и их анализ с использованием причинности по Грэнджеру. На языке C++ была разработана программа, автоматически обрабатывающая сезонные данные по климатическим процессам, группирующая их в зависимости от того, в каких временных периодах необходимо определить влияние процессов друг на друга, строящая двухкомпонентные и трехкомпонентные регрессионные модели, определяющая значимость влияния процессов и предсказывающая возможные следующие значения показателей, для которых строились регрессионные модели.

В результате работы при помощи регрессионных моделей были выявлены следующие сезонные зависимости процессов друг от друга (ЭНЮК — аномалии температуры поверхности в экваториальной области Тихого океана, САК — аномалии разности нормированных значений давления между североатлантическими центрами действия — азорским антициклоном и исландским циклоном, ИМ — аномалии количества осадков над Индией):

- в теплое полугодие ИМ значимо влияет как на САК, так и на ЭНЮК; кроме того, ЭНЮК малозначительно влияет на САК (влияние обнаружено только трехкомпонентной моделью);
- в холодное полугодие присутствует малозначительное влияние САК на ИМ;
- в весеннее время сильно выражено влияние САК на ИМ, в значительно меньшей степени — ЭНЮК на ИМ;
- в летнее время присутствует сильное влияние ИМ на САК и ЭНЮК, а также значительное влияние ЭНЮК на САК;
- в осеннее время наиболее выражено влияние ИМ на ЭНЮК;
- в зимнее время ИМ подвергается сильному влиянию САК, менее сильному — ЭНЮК. Кроме того, также выражено влияние ИМ на ЭНЮК.

Исходя из этого, можно сделать вывод, что аномалии индийского муссона оказывают наиболее сильное влияние на температурные аномалии в экваториальной области Тихого океана и на североатлантические колебания давления (в меньшей степени). Кроме того, ИМ также подвергается влиянию САК и

ЭНЮК. Наиболее сильно влияние процессов друг на друга выражено в летнее время.

Использование трехкомпонентных регрессионных моделей позволило обнаружить более слабо выраженные факторы влияния процессов друг на друга при наличии более сильных факторов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 *Stocker, T. F.* Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. / T. F. Stocker. — 2013.
- 2 *Воскресенская, Е. Н.* Классификация Эль-Ниньо и климатические аномалии в Черноморском регионе / Е. Н. Воскресенская. — 2010.
- 3 *Груза, Г. В.* О связи климатических аномалий на территории России с явлением Эль-Ниньо — Южное колебание. / Г. В. Груза, Э. Я. Ранькова, Л. К. Клещенко, Л. И. Аристова // *Метеорология и гидрология*. — 1999. — Т. 5. — С. 32–50.
- 4 Глобальные климатические индексы [Электронный ресурс]. — URL: <https://oko-planet.su/pogoda/pogodaday/47776-globalnye-klimaticheskie-indeksy.html> (Дата обращения 27.05.2019). Загл. с экр. Яз. рус.
- 5 *Харченко, М. А.* Корреляционный анализ. Учебное пособие для вузов / М. А. Харченко. — Воронеж: Издательство ВГУ, 2008.
- 6 *Грэнджер, К. У.* Спектральный анализ временных рядов в экономике / К. У. Грэнджер, М. Хатанака. — Москва: Статистика, 1972.
- 7 *Магнус, Я. Р.* Эконометрика. Начальный курс / Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий. — Москва: Дело, 2004.
- 8 *Мохов, И. И.* Трехкомпонентный анализ сезонных особенностей взаимосвязи между явлениями Эль-Ниньо, североатлантическим колебанием и индийским муссоном / И. И. Мохов, Д. А. Смирнов // *Метеорология и гидрология*. — 2016. — Т. 12. — С. 8–32.
- 9 Прикладной регрессионный анализ [Электронный ресурс]. — URL: <http://window.edu.ru/resource/636/19636/files/metod181.pdf> (Дата обращения 26.05.2019). Загл. с экр. Яз. рус.