

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

по дисциплине

«СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ГИДРОМЕТЕОРОЛОГИЧЕСКОЙ ИНФОРМАЦИИ»

Направление подготовки 05.03.05. — Прикладная гидрометеорология
Профиль подготовки — Прикладная метеорология
Квалификация (степень) — Бакалавр академический

Санкт-Петербург
РГГМУ
2016

УДК 551.509.314(072)
ББК 26.236я73
М54

*Одобрены учёным советом
метеорологического факультета РГГМУ*

Методические указания по дисциплине «Статистические методы анализа гидрометеорологической информации» для бакалавров по направлению подготовки 05.03.05. — Прикладная гидрометеорология, профиль подготовки — Прикладная метеорология / Сост. Л.О. Неёлова, Ю.В. Ефимова. — СПб.: РГГМУ, 2016. — 24 с.

Методические указания составлены в соответствии с программой дисциплины «Статистические методы анализа гидрометеорологической информации» для бакалавров по направлению подготовки 05.03.05. — Прикладная гидрометеорология, профиль подготовки — Прикладная метеорология. Даются рекомендации по изучению разделов дисциплины. Приводятся вопросы для самопроверки, рекомендуемая литература, контрольная работа.

Составители:

Л.О. Неёлова, к.ф.-м.н., доц. каф. метеорологических прогнозов РГГМУ;
Ю.В. Ефимова, к.ф.-м.н., доц. каф. метеорологических прогнозов РГГМУ.

Ответственный редактор А.И. Погорельцев, д.ф.-м.н., зав. каф. метеорологических прогнозов РГГМУ

© Неёлова Л.О., сост., 2016
© Ефимова Ю.В., сост., 2016
© Российский государственный гидрометеорологический университет (РГГМУ), 2016

ПРЕДИСЛОВИЕ

Целью дисциплины «Статистические методы анализа гидрометеорологической информации» является подготовка бакалавров по направлению 05.03.05. — Прикладная гидрометеорология, профиль подготовки — Прикладная метеорология, владеющих глубокими теоретическими знаниями и практическими навыками, необходимыми для использования статистических методов обработки и анализа метеорологических наблюдений.

Изучение дисциплины «Статистические методы анализа гидрометеорологической информации» базируется на знаниях студентов, полученных в результате усвоения курсов математики, теории вероятностей и математической статистики, физики, информатики, физики атмосферы, океана и вод суши, методов и средств измерений гидрометеорологической информации и др.

Основная задача курса — изучение методов статистического анализа временных рядов и метеорологических полей, основных положений объективного анализа метеорологической информации и физико-статистических методов прогноза состояния атмосферы.

Компетенции обучающегося, формируемые в результате освоения дисциплины

Процесс изучения дисциплины направлен на формирование следующих компетенций:

ОК-1 — Способность к логическому мышлению, обобщению, анализу, систематизации профессиональных знаний и умений, а также закономерностей исторического, экономического и общественно-политического развития.

ОК-2 — Способность решать стандартные профессиональные задачи на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом требований информационной безопасности.

ОПК-2 — Способность к проведению измерений и наблюдений, составлению описания проводимых исследований, подготовке данных для составления обзоров, отчетов и научных публикаций, составлению отчета по выполненному заданию, участие по внедрению результатов исследований и разработок.

ПК-2 — Способность анализировать явления и процессы, происходящие в природной среде, на основе экспериментальных данных и массивов гидрометеорологической информации, выявлять в них закономерности и отклонения.

ПК-5 — Способность реализовать решения гидрометеорологических задач и анализа полученных результатов.

ПК-6 — Владение профессиональной гидрометеорологической терминологией, формами отчетности, кодами и единицами.

В результате освоения компетенций в рамках дисциплины «Статистические методы анализа гидрометеорологической информации» обучающийся должен:

знать:

- виды и источники метеорологической информации;
- требования, предъявляемые к метеорологической информации;
- принципы статистических методов обработки и анализа метеорологической информации;
- место и роль объективного анализа в оперативной обработке информации о состоянии атмосферы;
- цели и задачи многомерного анализа метеорологической информации;
- перспективные направления развития статистических методов обработки информации в прогностических целях;

уметь:

- применять на практике статистические методы обработки метеорологической информации;
- обрабатывать и интерпретировать получаемую информацию о физическом состоянии атмосферы и гидросферы;
- анализировать результаты и грамотно применять полученные выводы для решения научных и практических задач в области метеорологии;

владеть:

- методикой обработки и интерпретации гидрометеорологической информации;
- методикой обработки архивных данных.

ОБЩИЕ УКАЗАНИЯ

По дисциплине «Статистические методы анализа обработки гидрометеорологической информации» на четвертом курсе факультета заочного обучения предусматривается изучение следующих разделов: «Современные системы сбора, обработки и хранения метеорологической информации», «Аналитические функции распределения, используемые в метеорологии», «Интервальное оценивание параметров и проверка статистических гипотез», «Методы статистического анализа временных рядов», «Построение и анализ эмпирических зависимостей», «Статистическая структура метеорологических полей», «Контроль данных наблюдений», «Численный анализ», «Диагноз и прогноз состояния атмосферы методами многомерного анализа». Студент должен выполнить одну контрольную работу. Дисциплина завершается зачетом.

Литература

Основная:

1. *Корниенко В.С.* Математическая статистика. Решение задач по теме: «Проверка статистических гипотез». Методическая разработка. — Волгоград: Волгогр. гос. с.-ч. академия, 2010. — 68 с.
2. *Болдырев Н.П., Болдырева Н.В.* Статистика в схемах и таблицах. Ч. 2.: учебное пособие. — ФЛИНТА, 2014. — 135 с.
3. *Вуколов Э.А.* Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL: учебное пособие. — 2-е изд. — М.: Форум: НИЦ Инфра-М, 2013. — 464 с.: <http://znanium.com/catalog.php?bookinfo=369689>
4. *Долженков В.А.* Microsoft Office Excel 2007. — СПб.: БХВ-Петербург, 2007. — 1200 с.: <http://znanium.com/catalog.php?bookinfo=489481>

Дополнительная:

1. *Секан А.В.* Методы статистической обработки гидрометеорологической информации: учебник. — СПб., 2007. — 278 с.
2. *Аргучинцева А.В.* Методы статистической обработки и анализа гидрометеорологических наблюдений: учебное пособие. — Иркутск: Иркут. гос. ун-т, 2007. — 105 с.
3. *Малинин В.Н.* Статистические методы анализа гидрометеорологической информации. — СПб.: РГГМУ, 2008. — 407 с.
4. *Груза Г.В., Рейтенбах Р.Г.* Статистика и анализ гидрометеорологических данных. — Л.: Гидрометеоиздат, 1982. — 216 с.
5. *Казакевич Д.И.* Основы теории случайных функций в задачах гидрометеорологии. — Л.: Гидрометеоиздат, 1989. — 230 с.
6. *Гандин Л.С., Каган Р.Л.* Статистические методы интерпретации метеорологических данных. — Л.: Гидрометеоиздт, 1976. — 360 с.
7. *Рожков В.А.* Теория и методы статистического оценивания вероятностных характеристик случайных величин и функций с гидрометеорологическими примерами. Кн. 1. — СПб.: Гидрометеоиздат, 2001. — 340 с.

Интернет-ресурсы:

1. *Анализ* временных рядов:
http://chaos.phys.msu.ru/loskutov/PDF/Lectures_time_series_analysis.pdf
2. *Методы* и модели анализа временных рядов:
<http://www.tstu.ru/book/elib/pdf/2008/tatarenko-1.pdf>
3. *Использование* анализа временных рядов в изучении многолетних температурных изменений:
<https://www.google.ru/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CBsQFjAAahUKEwiYpbuH48zIAhWmEHIKHcEdA70&url=http%3A%2F%2Fcyberleninka.ru%2Farticle%2Fn%2Fispolzovanie-analiza-vremennyh-ryadov-v-izuchenii-mnogoletnih-temperaturnykh-izmeneniy.pdf&usg=AFQjCNGCwK3HfbimmF6ipGLyBnlK2FNpBA&bvm=bv.105454873,d.bGQ>
4. *Прогнозирование* временных рядов в пакете Statistica:
https://www.google.ru/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CBsQFjAAahUKEwje8N7X4szIAhVF_nIKHStrDTg&url=http%3A%2F%2Fvital.lib.tsu.ru%2Fvital%2Faccess%2Fservices%2Fdownload%2Fvitals%3A000439644%2FSOURCE1&usg=AFQjCNEVMk0xzpQUpbKNR9pPPRYZsYuYzA&bvm=bv.105454873,d.bGQ
5. *Анализ* временных рядов. Электронный учебник по статистике:
<http://www.statsoft.ru/home/textbook/modules/sttimser.html>

УКАЗАНИЯ ПО РАЗДЕЛАМ

Введение

Введение в дисциплину позволяет получить общие представления о роли статистических методов в решении задач обработки метеорологической информации. Выясните основные задачи дисциплины, ее связь с другими науками. Проанализируйте современное состояние и перспективы применения статистических методов в метеорологии. Это необходимо для понимания важности статистического подхода к описанию и прогнозу состояния атмосферы, понимания основных тенденций развития методов прогноза погоды на различные сроки. Обратите внимание на важность статистического подхода для климатологии, для охраны атмосферы как составной части окружающей среды, для разработки методов принятия оптимальных хозяйственных решений на основе метеорологической информации.

Вопросы для самоконтроля

1. Какие основные задачи решает данная дисциплина?
2. Чем обусловлена целесообразность применения статистического подхода в метеорологии?
3. В каких дисциплинах метеорологического цикла используются статистические методы обработки и анализа информации?

Современные системы сбора, обработки и хранения метеорологической информации

При изучении этого раздела важно уяснить, что статистическая обработка применяется к информации разного вида, получаемой от различных источников (данные наземных наблюдений и аэрологического зондирования, спутниковая информация, данные МРЛ и другие). Следует обратить внимание на существующие требования к точности наблюдаемых и прогнозируемых характеристик состояния атмосферы.

Познакомиться с современным состоянием систем сбора и обработки гидрометеорологической информации можно на примере как глобальных систем Всемирной службы погоды, так и специальных систем, обеспечивающих решение региональных задач мониторинга состояния окружающей среды.

Проблема хранения гидрометеорологической информации решается путем создания гидрометеорологических баз и банков данных. При временном хранении информации, когда срок хранения ограничен и информация должна постоянно обновляться, используют базы данных. При долговременном хранении в банке данных информация хранится и может быть восстановлена потребителем в течение неограниченного срока.

Гидрометеорологические банки данных выполняют разнообразные функции: включение в информационную базу новых сведений после проведения контроля, структурные преобразования массивов данных, их стандартная обработка, информационное обслуживание потребителей, в том числе нестандартная обработка данных по запросам.

Следует познакомиться с различными системами банков данных, функционирующих в нашей стране и за рубежом.

Доступ в Интернет позволит получить разнообразную метеорологическую информацию по различным регионам земного шара.

Вопросы для самопроверки

1. Какие виды метеорологической информации можно выделить?
2. Перечислите основные источники метеорологической информации.
3. От чего зависит точность исходной информации?
4. Перечислите элементы Единой системы сбора, обработки и хранения информации.
5. В чем состоит основная задача гидрометеорологических банков данных?
6. Какие возможности получения информации открываются при использовании сети Интернет?

Аналитические функции распределения, используемые в метеорологии

В последние десятилетия наблюдается широкое использование аппарата теории случайных функций в геофизических науках. Основой этого является идея рассмотрения фиксированных мгновенных значений гидрометеорологических процессов и пространственных полей как отдельных реализаций некоторого случайного процесса или случайного поля. Такой подход позволяет отказаться от рассмотрения особенностей отдельных мгновенных значений гидрометеорологических полей, зависимость которых от пространственных координат, а также их временной ход носят весьма сложный и запутанный характер, и перейти к рассмотрению некоторых осредненных свойств *статистической совокупности* их реализаций.

Основным понятием в теории вероятностей является *случайная величина*. Случайные величины определяются своими вероятностными характеристиками — законами распределения и моментами распределения.

Закон распределения случайной величины является ее исчерпывающей характеристикой. Однако его не всегда удается установить, и часто используют отдельные числовые характеристики, выражающие некоторые существенные черты распределения.

Различают начальные моменты k -го порядка и центральные моменты k -го порядка.

Для случайной величины, которая распределена по *нормальному закону*, дифференциальный закон распределения имеет вид

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(x - m_x)^2}{2\sigma_x^2}\right).$$

Это так называемая кривая распределения Гаусса. Она характеризуется только двумя моментами распределения — математическим ожиданием и дисперсией. Эта кривая является симметричной относительно вертикальной прямой, проходящей через точку $x = m$ и имеет максимум в этой точке, равный $\frac{1}{\sigma\sqrt{2\pi}}$. Для кривой Гаусса асимметрия и эксцесс равны нулю. Для многих метеорологических случайных величин данный закон распределения вполне приемлем.

Наряду с нормальным законом распределения в метеорологии используются и другие законы: равномерного распределения, логарифмически нормального распределения, гамма-распределения и др.

Сложные метеорологические явления чаще всего бывают обусловлены совокупным воздействием ряда различных случайных величин. Например, возникновение грозы связано и с температурой воздуха у земли и на высотах, с влажностью воздуха, с энергетическими характеристиками воздушной массы и т.д. Если рассматривать перечисленные характеристики, как случайные величины, то следует исследовать не только законы их распределения, но и законы распределения целой *системы случайных величин*.

Вопросы для самопроверки

1. Какая величина называется случайной? Приведите примеры случайных величин в метеорологии.
2. Что понимается под законом распределения случайной величины?
3. Укажите примеры законов распределения непрерывных и дискретных случайных величин.
4. Приведите примеры метеорологических величин, которые могут рассматриваться как распределенные по нормальному закону.

Интервальное оценивание параметров и проверка статистических гипотез

Параметры законов распределения оцениваются по конечному числу наблюдений за метеорологической величиной (по выборке конечного объема). Поэтому необходимо использовать однородные данные в достаточном объеме, который зависит от вида оцениваемого параметра. Обычно для оценок параметров используется не менее 30 случаев наблюдений. С целью

получения достоверных статистических результатов необходимо проверить предположения (статистические гипотезы) о свойствах выборок, в частности, об однородности данных, о соответствии аналитической функции распределения эмпирическим данным и др. Эта проверка проводится с применением специальных теоретических законов распределения.

В частности, для оценки соответствия случайной величины X нормальному закону распределения используют критерий Пирсона χ^2 . Следует детально рассмотреть вопросы получения интервальных оценок таких параметров распределения, как математическое ожидание, дисперсия и среднее квадратическое отклонение.

Особое внимание следует обратить на понятие статистической гипотезы, нуль-гипотезы и альтернативной гипотезы. Необходимо понять, что собой представляет уровень значимости. В качестве примера следует ознакомиться со схемой проверки статистических гипотез, провести проверку гипотез о соответствии аналитической функции распределения эмпирическим данным (критерии согласия).

Вопросы для самопроверки

1. Каким требованиям должна удовлетворять выборка значений случайной величины?
2. Какие статистические гипотезы проверяются с применением распределения χ^2 , распределения Стьюдента и распределения Фишера?
3. Чем отличается нуль-гипотеза от альтернативной гипотезы?
4. Опишите схему проверки статистических гипотез.
5. В чем заключается критерий согласия χ^2 ?
6. В чем заключается проверка метеорологического ряда на соответствие модели случайной величины?

Методы статистического анализа временных рядов

В теории случайных функций под случайным процессом или случайной функцией понимают обобщение понятия случайной величины. В случае случайного процесса результатом опытов является не число или определенная величина, а некоторая функция времени, которая при повторении опытов в одинаковых условиях каждый раз случайным образом меняет свой вид. Случайная функция является функцией одного или нескольких переменных.

Неслучайная функция, получающаяся в результате каждого опыта, называется *реализацией* случайной функции. При каждом повторении опыта будем получать новую реализацию. Следовательно, случайную функцию (случайный процесс) $X(t)$ можно рассматривать как множество всех ее реализаций. Такой статистический подход очень удобен при исследовании

многих метеорологических процессов и полей. Например, при обработке данных самописцев (термографа или барографа). Каждая лента самописца может быть рассмотрена как реализация случайного процесса. Наглядным примером случайной функции может служить турбулентная диффузия. Одним из методов изучения турбулентной диффузии в реальной атмосфере является применение трансзондов — уравновешенных шаров-пилотов, вес которых подбирается так, чтобы они свободно плавали в воздухе вблизи какой-нибудь изобарической поверхности. Каждый такой зонд рассматривается как частица в потоке газа. Регистрируя значение какой-нибудь координаты одной из таких частиц через определенные промежутки времени, мы получим реализацию случайного процесса.

С другой стороны, случайный процесс можно рассматривать как совокупность всех его случайных сечений, то есть как совокупность всех случайных величин, соответствующих моментам времени $t = t_1, t_2, t_3 \dots t_i \dots$

Случайный процесс приближенно считается полностью определенным, если известна совместная функция распределения множества его сечений. И чем ближе сечения расположены друг относительно друга, тем функция распределения будет полнее описывать случайный процесс.

Наиболее простыми для изучения и статистического описания являются такие случайные процессы, статистические свойства которых не изменяются с изменением аргумента. Такие процессы называются *стационарными*.

Для метеорологических процессов предположение о стационарности достаточно хорошо подтверждается для сравнительно небольших интервалов времени. С увеличением интервалов изменения аргумента наблюдается нарушение стационарности.

Особое внимание следует обратить на такое понятие, как эргодичность случайного процесса, так как эргодические случайные процессы допускают оценку статистических характеристик по одной реализации — временному ряду.

В связи с этим, можно определить статистические оценки временного ряда как реализации эргодического случайного процесса.

Необходимо изучить влияние ошибок наблюдений на значения оценок.

Следует ознакомиться с основами спектрального анализа стационарного случайного процесса, с фильтрацией и сглаживанием временных рядов метеорологических величин и с определением оценок корреляционной функции и спектральной плотности.

Необходимо подробно рассмотреть применение аппарата спектральных разложений к стационарным случайным процессам.

Изучить зависимость корреляционной функции и соответствующей ей спектральной плотности от параметра α . С ростом параметра α корреляционная функция резко уменьшается, так как связь между сечениями убывает. При малых значениях α имеем пример узкополосного спектра.

При увеличении α имеет место широкополосный спектр. Случайный процесс, спектральная плотность которого постоянная во всем частотном интервале, называется «белым шумом» по аналогии с белым светом, у которого спектральный состав примерно однороден. Такой процесс физически нереализуем, так как его дисперсия обращается в бесконечность. Однако его можно рассматривать как предел реального широкополосного случайного процесса при стремлении α к бесконечности. Примером чисто случайного процесса, сечения которого не коррелированы между собой, может служить случайный процесс ошибок измерений.

Вопросы для самопроверки

1. Как связаны понятия временного ряда наблюдений и случайного процесса?
2. Что понимается под ансамблем реализаций и сечений случайного процесса?
3. Какие статистические характеристики рядов наблюдений используются в качестве климатических показателей?
4. Какими свойствами обладают статистические характеристики стационарного случайного процесса?
5. Какие процессы являются эргодическими?
6. Как рассчитывать корреляционную функцию эргодического случайного процесса?
7. Как определить спектральную плотность стационарного случайного процесса?
8. Каковы цели фильтрации и сглаживания временных рядов?

Построение и анализ эмпирических зависимостей

Эмпирической называется зависимость между значениями двух или более метеорологических величин, установленная на основе опытов (экспериментов). Эмпирическая стохастическая зависимость имеет место, когда каждому значению одной переменной с определенной вероятностью соответствуют значения другой переменной. Эта зависимость может быть линейной или нелинейной. Количественной мерой тесноты связи двух переменных является коэффициент корреляции. Необходимо изучить его свойства и методику расчета с оценкой достоверности полученного значения коэффициента корреляции.

Для определения коэффициентов уравнения регрессии, описывающего линейную зависимость двух случайных величин, используется метод наименьших квадратов.

Построение нелинейной зависимости целесообразно проводить методом выравнивания, т.е. перехода к новым переменным, между которыми имеется линейная зависимость. Для этих переменных применяется линейный регрессионный анализ, а затем осуществляется обратное преобразование к исходным переменным.

Особое внимание следует обратить на непараметрический подход при установлении эмпирических зависимостей и ознакомиться с понятием порядковых статистик, рассмотреть непараметрические коэффициент связи двух переменных, к которым относятся ранговые коэффициенты Кендалла и Спирмена.

Вопросы для самопроверки

1. Какие виды стохастической связи отмечаются между двумя метеорологическими величинами?
2. Какие свойства имеет коэффициент корреляции? Приведите диаграммы рассеяния для случаев, когда коэффициент корреляции имеет значения, близкие к -1 , к нулю и к $+1$?
3. Когда возникает ложная корреляция?
4. В чем сущность метода наименьших квадратов?
5. Опишите линейную регрессионную модель двух переменных?
6. Как строится нелинейная регрессионная модель двух переменных?
7. Что такое порядковые статистики?

Статистическая структура метеорологических полей

Помимо случайных процессов в метеорологии очень часто применяется такое понятие, как случайное поле. Случайное поле — это случайная функция от нескольких независимых переменных. Рассмотрим случайное поле $U(x, y, z, t)$, где x, y, z — пространственные координаты точки, а t — время. Пространственные координаты и временную координату можно рассматривать как координаты некоего четырехмерного вектора $\rho(x, y, z, t)$ и обозначать случайное поле в виде $U(\rho)$. Аналогично случайным процессам, случайное поле можно рассматривать, с одной стороны, как совокупность всех его реализаций, а с другой стороны, как совокупность всех его сечений. Под реализацией случайного поля можно, например, представить себе карту распределения какой-нибудь метеорологической величины за определенный срок наблюдений, а под случайным сечением — характеристики этой величины при фиксированных значениях x, y, z в различные моменты времени.

Как и для случайного процесса, на практике редко удается определить n -мерные функции распределения или плотности распределения, поэтому для характеристики случайного поля используются главным образом моменты распределения — математическое ожидание, дисперсия, ковариационные и структурные функции.

При изучении случайных процессов весьма важным является условие стационарности, существенно упрощающее описание случайного процесса. Для пространственных полей аналогичными условиями являются условия

однородности и изотропности. Для однородного случайного поля математическое ожидание не зависит от координат точек поля, т.е. является постоянной величиной, а корреляционная функция зависит только от разности векторов $l = \rho_2 - \rho_1$.

Для *однородного* и *изотропного* случайного поля в широком смысле слова математическое ожидание есть величина постоянная, а корреляционная функция зависит только от модуля l . Значит, для однородного и изотропного случайного поля $m_u = \text{const}$, $R_u(|l|)$.

Многочисленные исследования структуры метеорологических полей указывают на существенное различие изменений метеорологических элементов в горизонтальном и вертикальном направлениях. Поэтому при изучении мезо- и макромасштабных метеорологических полей свойства однородности относят только к горизонтальным координатам. При этом предполагается, что однородными являются только центрированные случайные поля. Само математическое ожидание нельзя считать постоянным.

Для более полного описания особенностей метеорологических полей, а также для сжатия информации часто применяют аппарат разложения метеорологических полей в ряды по естественным ортогональным функциям (ЕОФ). Этот аппарат был разработан и внедрен в практику такими видными учеными, как А.Н. Обухов, Н.А. Багров, М.И. Юдин.

Наряду со скалярными метеорологическими полями в метеорологии используется и векторное представление случайного поля. Это приводит к необходимости расчета статистических характеристик не только для каждой составляющей в отдельности, но и взаимного статистического анализа составляющих.

Вопросы для самопроверки

1. Какие характеристики используются для описания статистической структуры метеорологического поля?
2. В чем заключается гипотеза локального однородного и изотропного поля?
3. Каковы особенности статистической структуры реальных метеорологических полей?
4. Как определяются естественные ортогональные составляющие метеорологических полей?
5. Какие статистические характеристики применяются для описания векторного поля?

Контроль данных наблюдений

Сведения об окружающей нас природной среде и, в частности, о состоянии атмосферы и гидросферы получаются на основе обработки и последующего анализа гидрометеорологических данных, собранных по

различным территориям Земли и за различные периоды времени. Ввиду того, что физические величины, характеризующие состояние атмосферы, очень изменчивы в пространстве и во времени, для изучения происходящих процессов необходимо собрать метеорологические данные в единые комплексы, причем в максимально короткие сроки с момента проведения наблюдений. Важнейшей характеристикой первичной метеорологической информации является *точность ее измерения*. Большое значение имеют также требования к *разрешению в пространстве и времени (периодичности)* поступающей информации. Поскольку основным потребителем первичной метеорологической информации является служба погоды, отметим общие требования, предъявляемые этой службой. Эти требования включают: *глобальность, комплексность, регулярность, синхронность*.

Сетевые аэрологические наблюдения служат основой оперативных гидродинамических прогнозов. В последние годы требования к надежности этого вида метеорологической информации непрерывно возрастают в связи с тем, что для прогностических целей используются все более сложные математические модели. С помощью этих моделей можно прогнозировать основные метеорологические величины, как для небольших регионов, так и для территории всего Земного шара. Поэтому контроль аэрологических сообщений оказывается обязательным этапом в технологической схеме каждого крупного метеорологического прогностического центра.

Как известно, результаты наблюдений не могут быть абсолютно точными. Они всегда содержат множество погрешностей. Устранение всех незначительных погрешностей совершенно необязательно. Но обнаружение и исправление грубых просчетов и ошибок, являющихся следствием неисправности измерительной аппаратуры, искажений в процессе передачи сообщений по каналам связи и т.д. совершенно необходимо. Несмотря на то, что так называемые грубые ошибки встречаются, не очень часто, они могут существенно снизить точность анализа и прогноза, метеорологических полей.

Главной целью контроля является именно предупреждение искажающего влияния грубых ошибок.

В условиях современной технической революции и всеобщей компьютеризации прогностической метеорологической службы необходимо рассматривать автоматизированные методы контроля. Такие методы в отличие от субъективных приемов, используемых непосредственно при ручной обработке метеорологической информации, позволяют подвергнуть контролю весь объем начальных данных и устранить большую часть грубых ошибок.

В настоящее время для контроля аэрологической информации используется вертикальный статический контроль, основанный на проверке выполнения уравнения статики в пределах каждого слоя между стандартными изобарическими поверхностями.

Вопросы для самопроверки

1. Каковы цели и задачи объективного контроля?
2. Как проводится контроль грубых ошибок?
3. В чем заключается процедура вертикального статического контроля?
4. Чем объясняется необходимость комплексного контроля метеорологических полей?
5. Как используется прогностическая информация в процедурах контроля?

Численный анализ

Численный анализ гидрометеорологической информации статистическими методами можно условно подразделить на два вида: объективный анализ и многомерный статистический анализ.

Необходимо обратить внимание на место и роль объективного анализа в оперативной обработке информации о состоянии атмосферы, изучить этапы объективного анализа метеорологических полей. Также рассмотреть и изучить методы горизонтальной интерполяции полей метеорологических величин: полиномиальная интерполяция, оптимальная интерполяция, интерполяция методом взвешенного среднего, весовая анизотропная интерполяция.

Особое внимание необходимо уделить особенностям интерполяции по вертикали и применению для этих целей сплайнов для интерполяции по вертикали.

Вопросы для самопроверки

1. Каковы цели и задачи численного анализа метеорологических полей?
2. Какие этапы можно выделить в объективном анализе?
3. В чем заключается процедура горизонтальной интерполяции метеорологических полей?
4. Чем объясняются особенности интерполяции полей по вертикали?
5. Перечислите основные методы, используемые для интерполяции по вертикали?

Диагноз и прогноз состояния атмосферы методами многомерного анализа

Для ознакомления с данной темой необходимо уяснить себе цели и задачи многомерного анализа метеорологической информации, а также методы построения множественной регрессии.

Необходимо изучить линейные и нелинейные модели множественной регрессии. Обратит внимание на выбор информативных переменных для моделей множественной регрессии и использование множественной регрессии при прогнозе метеорологических величин. Изучить метод группового учета аргументов и его использование в прогностических моделях.

В настоящее время для диагноза и прогноза состояния атмосферы широко используются методы дискриминантного анализа — параметрические и непараметрические модели. При разработке прогноза указанным методом предварительно производится статистическая обработка архивного материала.

В теории статистических прогнозов характеристики состояния атмосферы, x_1, x_2, \dots , используемые при составлении прогноза, принято называть прогностическими признаками, или *предикторами*, а прогностические характеристики y_1, y_2, \dots — *предиктантами*. И предиктанты и предикторы могут быть как количественными (температура воздуха, влажность, скорость и направление ветра и др.), так и качественными (форма облаков, вид осадков и др.) метеорологическими характеристиками. В перечень предикторов обычно включаются характеристики исходного состояния атмосферы — так называемые первичные прогностические признаки.

Стоит обратить внимание на оценку информативности переменных для создания оптимального описания в задачах дискриминантного анализа, а также на построение дискриминантных функций.

Вопросы для самопроверки

1. Какие метеорологические задачи решаются методами многомерного статистического анализа?
2. Какие требования предъявляются к архивным данным?
3. На чем основан выбор информативных предикторов при построении уравнения множественной регрессии?
4. Как строится разделяющая граница классов в методе эталонов?
5. Как определить дискриминантную функцию при нормальном законе распределения?
6. Для каких целей используется расстояние Махаланобиса?
7. Как осуществляется разбиение множества метеорологических полей на классы?
8. Какой критерий качества классификации чаще всего используется в кластерном анализе?

Заключение

В заключение следует ознакомиться с перспективными направлениями развития методов статистической обработки и анализа метеорологической информации. Эти направления тесно связаны с развитием систем получения и передачи метеорологической информации, с расширением возможностей применения вычислительной техники, с совершенствованием математического аппарата и применением новых подходов к статистической обработке и анализу метеорологической информации.

КОНТРОЛЬНАЯ РАБОТА

Общие указания

Перед выполнением контрольной работы следует изучить рекомендованные литературные источники. После самостоятельного изучения необходимо выполнить контрольную работу, состоящую из восьми заданий. При выполнении заданий используются теоретические положения и расчетные формулы по каждому из соответствующих разделов дисциплин.

В заданиях 1, 3 и 4 используются данные наблюдений в пункте проживания студента-заочника за любой непрерывный тридцатидневный интервал времени в течение года. Если такие данные у студента не имеются, то можно взять из литературных источников.

Каждое из заданий 2, 5–8 дано в 10 вариантах, способы выбора которых указаны в заданиях.

Ответы на вопросы контрольной работы должны быть достаточно подробными. Изложение следует начинать с постановки задачи, необходимо приводить расчетные формулы с пояснением используемых величин и давать анализ полученных результатов.

Задание 1

В пункте проживания студента-заочника провести ежедневные наблюдения в течение тридцати дней за температурой воздуха (с точностью до десятых долей градуса) в выбранный срок местного времени, например, в 9 или 15 часов. Для получения временного ряда из 30 значений температуры воздуха вычислить среднее значение, дисперсию, среднее квадратическое отклонение и проверить гипотезу соответствия закона распределения вероятности нормальному закону.

Задание 2

Ответьте подробно на вопрос, номер которого определяется по последней цифре номера зачетной книжки.

1. Какие статистические характеристики используются для описания случайного процесса и являются климатическими показателями?
2. Какие свойства многомерных законов распределения и моментных характеристик присущи стационарному случайному процессу?
3. Укажите свойства корреляционной и структурной функций стационарного случайного процесса и связь между ними.
4. Какие случайные процессы являются эргодическими?

5. Как осуществляется вычисление корреляционной функции эргодического стационарного случайного процесса?
6. Как влияет наличие ошибок в данных наблюдений на оценки значений структурной функции стационарного случайного процесса?
7. Как влияет наличие ошибок в данных наблюдений на оценки значений корреляционной функции стационарного случайного процесса?
8. Как определить дискретный спектр стационарного случайного процесса, реализации которого заданы на промежутке времени $[-T, T]$?
9. Как определить спектральную плотность стационарного случайного процесса?
10. Каковы особенности корреляционных функций и спектральных плотностей узкополосного и широкополосного стационарных случайных процессов?

Задание 3

Предполагая, что ряд ежедневных наблюдений за температурой воздуха в пункте наблюдений (см. задание 1) представляет собой реализацию эргодического стационарного случайного процесса, произвести расчеты нормированной корреляционной функции и структурной функции для значений сдвига во времени, не превышающего 10 суток. Построить соответствующие графики. Аппроксимировать нормированную корреляционную функцию аналитическим выражением с указанием средней квадратической погрешности аппроксимации.

Задание 4

Используя аналитическое выражение для нормированной корреляционной функции, полученное в предыдущем задании, рассчитать значения спектральной плотности при частотах, равных 0; 0.5; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0. Определить вид спектра. Провести качественный анализ его предсказуемости.

Задание 5

Ответьте подробно на вопрос, номер которого определяется по предпоследней цифре (число десятков) номер зачетной книжки.

1. В чем заключаются свойства однородности и изотропности случайного поля?
2. Как осуществляется разложение метеорологических полей по естественным ортогональным составляющим?
3. Опишите подробно, что собой представляют собственные вектора и собственные числа корреляционной матрицы случайного поля.

4. Укажите цели и основные этапы объективного анализа метеорологических полей.
5. Какими способами осуществляется контроль данных в процедурах объективного анализа?
6. Какие методы используются для интерполяции по горизонтали?
7. Чем отличаются метод взвешенного среднего от метода весовой анизотропной интерполяции? Опишите подробно эти два метода интерполяции по горизонтали.
8. В чем заключается метод оптимальной интерполяции?
9. Какие методы используются для интерполяции по вертикали?
10. Опишите алгоритм построения сплайн полинома.

Задание 6

Провести вертикальный статический контроль данных радиозондирования, пользуясь сведениями табл. 1. Номер варианта определяется по последней цифре номера зачетной книжки студента-заочника. Дополнительно к значениям, приведенным в табл. 1 для каждого варианта, вводится поправка K , учитывающая год выполнения задания:

$$K = \Gamma - 10 - 2000,$$

где Γ — текущий год.

Например, в 2013 г. K будет иметь значение $K = 2013 - 10 - 2000 = 3$. Ко всем значениям H в табл. 1 следует прибавлять величину K , а ко всем значениям t следует добавлять $0,1 K$.

Таблица 1

Варианты исходных данных

№ варианта	H_{850}	t_{850}	H_{700}	t_{700}	H_{500}	t_{500}	H_{400}	t_{400}
1	157	21	311	10	591	-11	—	—
2	154	19	315	9	592	-10	—	—
3	—	—	311	10	591	-11	758	-25
4	—	—	315	9	592	-10	754	-27
5	153	14	314	6	582	-10	—	—
6	142	6	300	1	562	-16	—	—
7	—	—	314	6	582	-10	749	-29
8	—	—	300	1	562	-16	740	-31
9	150	9	301	3	580	-12	—	—
10	—	—	301	3	580	-12	751	-28

Задание 7

Провести горизонтальный контроль значения метеорологического поля в узле путем интерполяции значений поля с трех окружающих узел станций. Использовать метод взвешенного среднего, весовой анизотропной интерполяции и оптимальной интерполяции. Номер варианта соответствует предпоследней цифре (число десятков) номера зачетной книжки студента-заочника. Выбор интерполируемой метеорологической величины осуществляется по табл. 2. Исходные данные приведены в табл. 3. Ко всем значениям H в табл. 3 добавляется значение K (см. задание 6), а ко всем значениям t добавляется $0,1 K$.

Таблица 2

Варианты выбора интерполируемой метеорологической величины

№ варианта	Интерполируемая величина
1	H_{850}
2	t_{850}
3	H_{700}
4	T_{700}
5	H_{500}
6	T_{500}
7	H_{400}
8	T_{400}
9	H_{300}
10	T_{300}

Таблица 3

Значения геопотенциальной высоты H (в дам) и температуры t (в °С)

Давление (P), гПа	Номер станции							
	1		2		3		0(узел)	
	H	t	H	t	H	t	H	t
850	149	9	153	-14	152	10	151	11
700	309	1	312	4	311	3	310	2
500	573	-14	577	-12	576	-13	575	-13
400	753	-25	754	-23	749	-27	751	-24
300	940	-43	945	-40	943	-41	942	-42

Принять расстояния между узлом и станциями равными $\rho_{01} = 220$ км, $\rho_{02} = 240$ км, $\rho_{03} = 200$ км, а между станциями: $\rho_{12} = 420$ км, $\rho_{13} = 350$ км, $\rho_{23} = 360$ км.

Нормированная корреляционная функция имеет вид

$$r(\rho) = e^{-\rho} (1 + 0,9\rho),$$

где ρ в тыс. км.

Провести сопоставление результатов интерполяции различными методами.

Задание 8

Ответить подробно на вопрос, номер которого соответствует последней цифре номера зачетной книжки.

1. Как построить уравнение множественной регрессии?
2. Как построить уравнение линейной регрессии?
3. Выписать систему уравнений метода наименьших квадратов для определения коэффициентов уравнения регрессии второй степени на основе двух предикторов.
4. Выписать систему уравнений метода наименьших квадратов для определения коэффициентов уравнения регрессии второй степени на основе трех предикторов.
5. Рассмотреть математическую постановку задачи дискриминантного анализа.
6. Рассмотреть математическую постановку задачи непараметрического дискриминантного анализа.
7. В чем различие параметрического и непараметрического дискриминантного анализа?
8. Какие методы относятся к непараметрическому дискриминантному анализу?
9. Каковы цели и задачи кластерного анализа?
10. Как определить наиболее информативные предикторы для построения уравнения регрессии?

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ	3
ОБЩИЕ УКАЗАНИЯ	5
Литература	5
УКАЗАНИЯ ПО РАЗДЕЛАМ	7
Введение	7
Современные системы сбора, обработки и хранения метеорологической информации	7
Аналитические функции распределения, используемые в метеорологии	8
Интервальное оценивание параметров и проверка статистических гипотез	9
Методы статистического анализа временных рядов	10
Построение и анализ эмпирических зависимостей	12
Статистическая структура метеорологических полей	13
Контроль данных наблюдений	14
Численный анализ	16
Диагноз и прогноз состояния атмосферы методами многомерного анализа	16
Заключение	17
КОНТРОЛЬНАЯ РАБОТА	18
Общие указания	18
Задание 1	18
Задание 2	18
Задание 3	19
Задание 4	19
Задание 5	19
Задание 6	20
Задание 7	21
Задание 8	22

УЧЕБНОЕ ИЗДАНИЕ

Составители:

Людмила Олеговна Неёлова,
Юлия Викторовна Ефимова

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
по дисциплине
«СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА
ГИДРОМЕТЕОРОЛОГИЧЕСКОЙ ИНФОРМАЦИИ»

Редактор Н.И. Афанасьева
Технический редактор Ю.И. Климов

Подписано в печать 03.10.16. Формат 60×90 1/16. Гарнитура Newton.
Печать цифровая. Усл. печ. л. 1,5. Тираж 100 экз. Заказ № 531.
РГГМУ, 195196, Санкт-Петербург, Малоохтинский пр., 98.
Отпечатано в ЦОП РГГМУ
