

**Министерство образования и науки Российской Федерации
САРАТОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО**

Кафедра метеорологии и климатологии

Лапина С.Н., Иванова Г.Ф., Семенова Н.В.

**Лабораторный практикум по курсу
« Обработка и анализ метеорологической
информации»**

Саратов 2011

Предисловие

Курс «Обработка и анализ метеорологической информации» объемом 32 часа, предназначен для студентов старших курсов бакалавриата специальности «Прикладная информатика в географии», предусматривает проведение лабораторного практикума.

Для методического обеспечения лабораторного практикума по данному курсу подготовлено учебно-методическое пособие, которое включает в себя вводную часть, и четыре лабораторные работы, сгруппированные в две части.

Часть первая включает лабораторные работы 1 и 2, в процессе выполнения которых, студенты знакомятся с общими положениями теории статистической обработки и анализа метеорологической информации, осваивают расчет основных характеристик распределения (средней, медианы, моды, дисперсии, среднего квадратического отклонения, коэффициентов вариации, асимметрии и эксцесса, а также их статистических ошибок).

При этом обращается внимание студентов на то, что статистические характеристики случайных процессов выступают в роли климатических показателей. Климат же является одним из компонентов географической оболочки.

Часть вторая включает лабораторные работы 3 и 4. При выполнении работы 3 студенты осваивают возможности статистической оценки метеорологической информации, полученной с помощью автоматической метеостанции Vantage Pro (WS), установленной в учебном метеоцентре кафедры метеорологии и климатологии географического факультета.

Станция Vantage Pro это автоматический метеокомплекс, состоящий из двух блоков: измерительного (блок датчиков) и приемного (консоль соединенная с компьютером). Блок датчиков выведен наружу и измеряет следующие параметры:

1. скорость и направление ветра,
2. температуру воздуха,
3. осадки,
4. влажность,
5. атмосферное давление.

Консоль – это мини компьютер, который осуществляет сбор, хранение и демонстрацию архивных и текущих данных погоды, в том числе и в графическом режиме.

В работе 4 в качестве инновационного подхода предлагается статистический анализ полей и явлений погоды проводить по комплекту карт, принятых в режиме реального времени по сети Интернет с использованием программы ГИС-Метео, позволяющей проводить обработку и анализ карт как в электронном, так и печатном вариантах.

В каждой работе дано краткое изложение теоретических вопросов, сформулированы задачи, указаны исходные данные, даны рекомендации по их выполнению, приведены тесты для контрольной проверки усвоения предложенного материала.

Данное пособие рассчитано на студентов двухуровневого образования, поскольку приобретенные первичные навыки обработки и анализа метеорологической информации являются базовой основой для подготовки магистров.

Вводная часть

Метеорологическая информация и способы ее представления

Вся работа по анализу и прогнозу атмосферных процессов и погоды производится на основе сведений о параметрах атмосферы, которые получают в результате метеорологических наблюдений. Совокупность таких сведений представляет первичную метеорологическую информацию.

Источником получения первичной метеоинформации являются метеорологические станции, большинство из которых входит в международную сеть. Каждой станции присваивается индекс (номер), по которому ее можно найти на бланке карты.

В дальнейшем эта информация обрабатывается, т.е. приводится к виду удобному для пользования. Такую информацию называют - вторичной. Существует целая система наблюдений, сбора, обработки метеоинформации – глобальная система Всемирной службы погоды. Данные наблюдений в атмосфере используются для различных целей, поэтому актуально встает

вопрос их сохранения. Проблема хранения информации решается путем создания баз и банков данных. При временном хранении используются базы данных, при долговременном – банки данных, в которых информация может быть востребована потребителем в течение неограниченного срока. Крупнейшими банками метеорологической информации являются банки мировых центров (Москва, Вашингтон, Мельбурн). Расширяющийся доступ в Интернет позволяет получать разнообразную информацию по различным регионам Земного Шара.

Системы получения первичной метеорологической информации дают дискретные или непрерывные распределения во времени и пространстве значений многих характеристик атмосферы. Это распределение, т.е. поля метеорологических величин и явлений могут быть описаны различным образом: словесно, таблично, графически.

Достоинством словесной информации является понимание широким кругом потребителей ее содержания. Это делает возможным ее распространение всеми средствами массовой информации. Обзоры погоды передаются по радио, телевидению, публикуются в газетах.

Табличная форма – это систематизированный по району и времени наблюдения набор информационных телеграмм, в которых в закодированном виде содержатся результаты метеорологических наблюдений. Количественные и качественные характеристики метеовеличин кодируются в виде пятизначных цифровых групп в соответствии с принятыми международными кодами. В табличном (закодированном) виде результаты наблюдений выглядят так:

41322 82211 10197 20183 40034 56613 78122 88900

При кодировании достигается существенное (по сравнению со словесным) уплотнение информации, что удобно при международном обмене информацией, делает доступным ввод ее в ЭВМ и для архивного хранения.

Еще более удобным и наглядным для анализа является графическое представление метеорологической информации в виде карт, графиков, диаграмм. Широко распространено в службе погоды использование географических карт с нанесенными на них данными метеорологических наблюдений и называемые синоптическими картами. На такую карту по определенной схеме вокруг каждой станции наносятся условными знаками значения метеорологических величин и явлений, которые наблюдатель фиксирует в данный момент времени. Такие карты называют приземными

картами погоды. На рисунке 1 показана схема нанесения метеорологических данных на карту погоды, а на рисунке 2 приведен пример нанесения метеорологической информации на приземную карту.

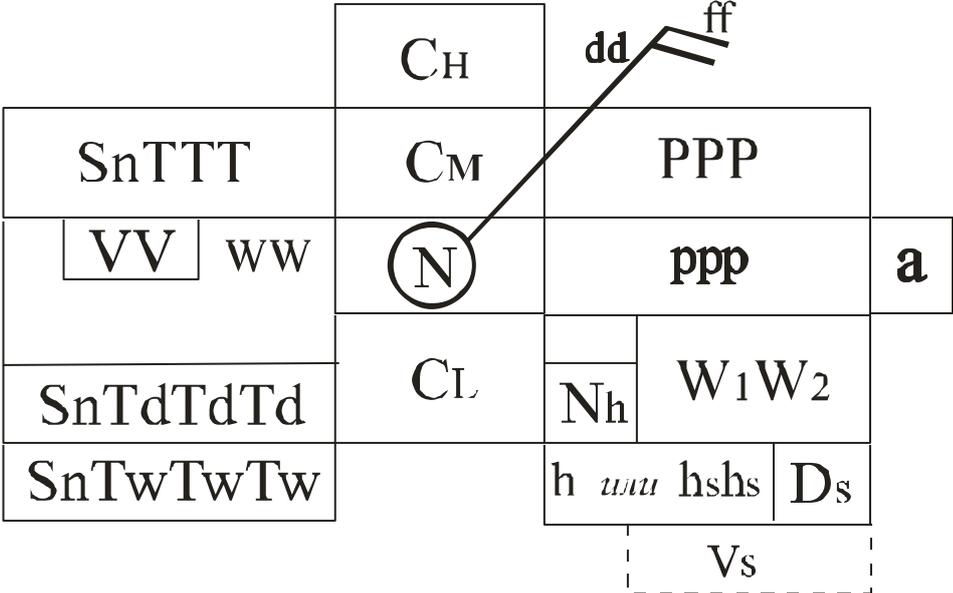


Рисунок 1 Схема нанесения метеорологических данных на карту погоды

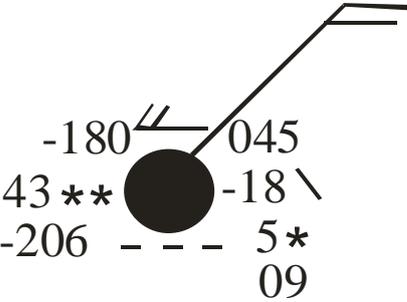


Рисунок 2 Пример нанесения метеорологических данных на карту погоды

В последние годы, для обработки метеорологической информации, стали активно применяться компьютерные технологии. Так на стыке информатики и метеорологии были созданы гидрометеорологические ГИС (геоинформационные системы). В организация Росгидромета получила широкое распространение геоинформационная система ГИС-Метео. Программный комплекс ГИС-Метео позволяет создавать метеорологические карты в любой картографической проекции и масштабе для любого региона с

использованием данных, распространяемых по глобальной телесвязи и через Интернет. Программа ГИС-Метео состоит из информационной системы, ее различных компонент, метеорологической база данных, технологических средств сбора, распространения и обработки данных. ГИС-Метео по заранее подготовленному сценарию автоматически или в интерактивном режиме подготавливает многочисленные слои информации на фоне географической карты. Такое совмещение информационных слоев на мониторе компьютера позволяет осуществлять как «безбумажную» (электронную) технологию работы метеоролога, так и подготавливать метеорологические карты к печати для дальнейшего анализа. Все образцы карт и их анализ представлен в работе 4.

Лабораторная работа 1.

Тема: Распределение частот. Гистограмма частот.

Цель работы: Научиться анализировать и оценивать метеорологический ряд на вероятность появления события ниже или выше заданной величины.

Основные понятия математической статистики

Статистика - наука о закономерностях массовых явлений или событий. Объектом исследования математической статистики является статистическая совокупность, а предметом изучения - методы статистической обработки. **Статистическая совокупность** - всякое множество сходных в некоторых отношениях, но индивидуально различающихся объектов. Совокупностями будут, например, результаты наблюдений за температурой воздуха на какой-либо станции за определенный (два, пять, сто лет) период. Основным методом математической статистики является **выборочный метод**. Сущность его заключается в том, что сплошное обследование массовых однородных объектов заменяется частным, причем без существенных ошибок в выводах.

Согласно положениям выборочного метода различают **генеральную совокупность** и **выборочную совокупность**. **Генеральная совокупность** это объединение множества значений случайной величины, которое охватывает весь диапазон ее изменения. **Выборочная совокупность** - часть генеральной совокупности. Выборку используют для того, чтобы изучить ее статистические свойства и по ним сделать выводы о статистических свойствах генеральной совокупности

Выборочная совокупность в свою очередь делится на малую ($n < 30$) и большую ($n > 30$). Отдельные значения изменяющейся величины, входящие в совокупность, называются **вариантами** или **случайной величиной**. Если различие между отдельными вариантами выражается в качествах, то вариацию называют **качественной** (например, различия по цвету, вкусу и т.п.). Если варианты принимают числовые значения, то вариация считается **количественной**. Случайные величины могут быть **дискретными** (число дней с явлениями), **непрерывными** (значения температуры, влажности, давления), **ограниченными с одного конца** (осадки, скорость не могут быть

отрицательными величинами), **ограниченными с обоих концов** (облачность может быть от 0 до 10 баллов) или неограниченными (нельзя указать скорость ветра, которая превысит максимальную скорость в данном пункте).

Задание 1.1. Вычисление абсолютных и относительных частот

Абсолютная повторяемость или частота m_i это число, показывающее, сколько раз то или иное значение элемента или градации встречается в общей совокупности n . Термин частота в статистике означает число случаев.

Относительная повторяемость или частота p_i есть относительная частота (часть) соответствующего интервала (градации) к сумме всех частот $n = \sum m_i$ и определяется по формуле:

$$p_i = \frac{m_i}{n} 100, \% \quad (1.1)$$

Накопленная (кумулятивная) повторяемость (P) это суммарное значение относительных частот:

$$P = \sum p_i, \% \quad (1.2)$$

Для группировки выборки существуют практические правила:

А. Число классов (градаций) k можно оценить по объему выборки n по уравнению:

$$k = 5 \lg n. \quad (1.3)$$

Б. Ширина класса ΔX приближенно равна:

$$\Delta X = (x_{\max.} - x_{\min.})/k, \quad (1.4)$$

где $x_{\max.}$, $x_{\min.}$ - значения максимального и минимального члена выборки.

Распределение частот (повторяемостей) применяется для такого упорядочения данных, которое позволяет легко и быстро их проанализировать.

Проиллюстрируем это на следующем примере: пусть 59 температур расположены по датам, скажем, с 1 мая по 28 июня. В статистике анализируемый элемент называют случайной переменной и обозначают X или другой какой-либо заглавной буквой. В нашем примере температура — случайная переменная. Найти из необработанных данных самую низкую и самую высокую температуру или такое ее значение, вокруг которого группируются другие температуры,— задача довольно трудоемкая. Если эти

температуры перегруппировать в порядке их возрастания или убывания, то уже сама запись позволит без труда увидеть самые высокие, самые низкие и средние значения температуры. В этом случае самой меньшей величине присваивают ранг 1, следующей — 2 и т. д. до последней (ранг 59). Если две температуры одинаковы, им присваивают средний ранг. Например, если 5-я и 6-я температуры от начала ряда равны 30° С, то их ранг будет 5,5.

Дальнейшее упорядочение данных можно произвести путем объединения отдельных фактических температур в группы. Такая методика обобщения данных позволит в дальнейшем пользоваться не отдельными наблюдениями, а их группами. **Интервал температуры (либо любой другой случайной переменной), на который распространяется данная группа, называют градацией или классом.** В метеорологии обычно пользуются следующими типами градаций:

- 1) градации числовые и равные по величине,
- 2) градации числовые и неравные по величине,
- 3) градации нечисловые (выраженные словесно). Рассмотрим указанные выше типы градаций.

Тип 1. Градации числовые и равные по величине.

Это наиболее удобный тип градаций. В таблице 1.1 (1) 59 значений температуры представлены в виде распределения частот для шести градаций. Размер градации есть интервал между последовательными математическими границами. В таблице 1.1 (1) размер градации равен 4. Частота (число случаев) обозначается буквой n_i и представляет собой сумму числа случаев, вошедших в данную градацию. Сумма частот всех градаций равна общему числу наблюдений: $\sum m_i = n$. Третий столбец показывает повторяемость градаций, в %, которая называется относительной частотой.

Таблица 1.1

Три типа распределения частот

1. Температура, °С			2. Осадки, мм			3. Облачность		
Градации	m_i	$p_i, \%$	Градации	m_i	$p_i, \%$	Градации	m_i	$p_i, \%$
16,0-19,9	3	5	0	33	49	Ясно	29	37
20,0-23,9	9	15	0,1-0,3	16	24	Малооблачно	18	23
24,0-27,9	12	21	0,4-1,0	7	10	С просветами	13	16
28,0-31,9	15	25	1,1-5,0	5	7	Сплошная	19	24
32,0-35,9	12	21	5,1-8,0	4	6			
36,0-39,9	8	13	8,1-12,0	3	4			

Тип 2. Градации числовые, но не равные по величине.

В таблице 1.1 (2) дано распределение частот для сумм осадков. Если пользоваться здесь градациями, равными по величине, то из-за большого числа случаев, когда осадки не наблюдаются или выпадают в малом количестве, повторяемости в градациях будут распределены чрезвычайно непропорционально. При выборе градаций, не равных по величине, руководствуются теми же принципами, что и при выборе равных по величине градаций.

Тип 3. Градации нечисловые, выраженные в словесных характеристиках.

В таблице 1.1(3) дано распределение частот для облачности, выраженной в словесных характеристиках. На основании распределения частот, выраженного в такой форме, численные расчеты произвести невозможно. Однако такое распределение позволяет быстро интерпретировать данные 79 упомянутых наблюдений. Такая форма распределения частот может быть преобразована в числовую путем придания словесным характеристикам «ясно», «малооблачно», «с просветами», «сплошная» определенных числовых значений.

Таблица 1.2

Распределение накопленных частот

Градации температуры	m_i	НЧ	ВГ	$p_i, \%$	$P=\sum p_i, \%$
16,0-19,9	3	3	19,95	5	5
20,0-23,9	9	12	23,95	15	20
24,0-27,9	12	24	27,95	21	41
28,0-31,9	15	39	31,95	25	66
32,0-35,9	12	51	35,95	21	87
36,0-39,9	8	59	39,95	13	100

Примечание: НЧ- накопленные частоты; ВГ- верхняя граница

В таблице 1.2 дано распределение накопленных частот. В ней приведены следующие данные: градации случайной переменной, частоты градаций (m_i), верхние математические границы градаций (ВГ), относительные частоты p_i , накопленные частоты (НЧ), полученные путем последовательного

суммирования абсолютной частоты (m_i), т. е. $3 + 9=12$, $3 + 9+12 = 24$ и т. д.. Накопленные относительные частоты ($P=\sum p_i, \%$), определяются величинами столбца p_i и определяются путем последовательного суммирования относительных частот p_i , т.е. $5+15=20$, $5+15+21=41$ и т.д.

Для того чтобы найти, как часто повторяются значения метеорологических элементов, большие или меньшие тех или иных заданных величин, пользуются распределением накопленных (кумулятивных) частот. Ответ на вопрос, сколько или какой процент из 59 температур ниже, например, $31,95^0$ С, можно легко получить непосредственно из табл. 1.2, откуда следует, что в 39 случаях или в 66%, температура ниже $31,95^0$. Для нахождения числа случаев с температурой выше $31,95^0$ нужно просто вычесть из общего количества наблюдений число случаев с температурой ниже $31,95^0$ ($59-39=20$ случаев или $100- 66=34\%$). Таким образом, в 20 случаях или в 34 % случаев температура выше $31,95^0$.

Порядок выполнения задания 1.1

Исходные данные:

Для примера взят метеорологический ряд среднесуточной температуры за 59 дней - с 1 мая по 28 июня:

Дата	1.06	2.06	3.06	4.06	5.06	6.06	7.06	8.06	...	25.07	26.07	27.07	28.07
T^0 С	16,8	19,7	24,1	18,4	28,3	27,9	25,8	29,1	...	38,4	26,1	33,4	39,7

1. Ранжировать исходный метеорологический ряд в возрастающем порядке:
16,8 18,4 19,7... 24,1 25,8 26,1... 27,9 28,3 29,1. 33,4 38,4 39,7.
2. Оценить ряд на тип градаций 1, 2 или 3 по табл. 1.1. Для приведенного примера ряда среднесуточной температуры градации выбраны по типу 1 – числовые и равные по величине.
3. Рассчитать возможное число классов (градаций) k по формуле 1.3. В примере $k = 6$.
4. Рассчитать ширину градации ΔX по формуле 1.4. В примере $\Delta X = 4^0$ С.
5. Построить таблицу 1.2 для полученных результатов.
6. Разбить метеорологический ряд на градации (см. табл.1.2).
7. Рассчитать абсолютные m_i и относительные частоты $p_i, \%$ по формулам 1.1 и 1.2.

8. Рассчитать накопленные абсолютные НЧ и относительные Р частоты по формуле 1.3.
9. Проанализировать полученные результаты.

Задание 1.2. Построение гистограммы абсолютных и относительных частот

Распределение частот обычно представляется графически с помощью гистограммы или полигона частот. На рисунке 1.1 дана гистограмма распределения абсолютных частот, приведенных в таблице 1.2. При построении гистограммы по оси абсцисс откладываются значения переменной (в данном случае - температуры), по оси ординат - частоты градаций (m_i). Гистограмма состоит из смежных прямоугольников, основаниями которых являются отрезки между последовательными математическими границами, а высотами - частоты каждой градации.

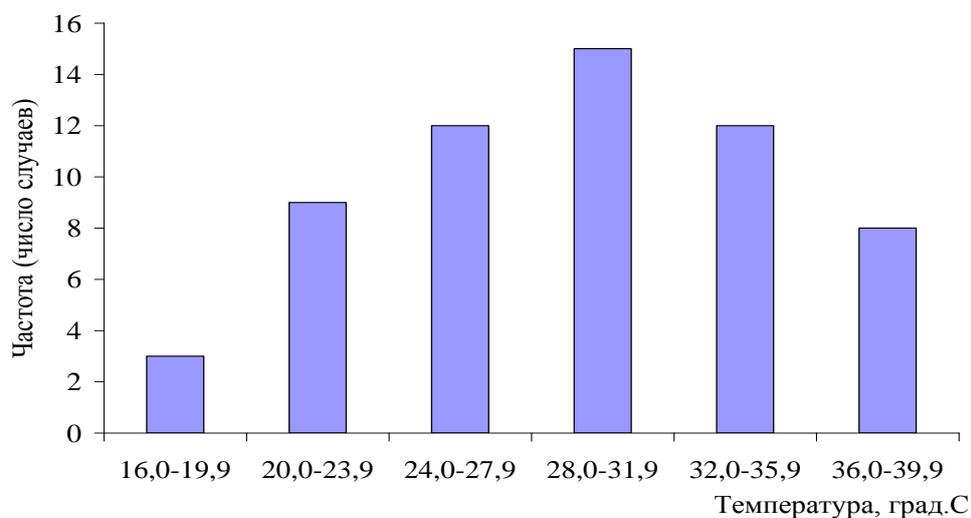


Рисунок 1.1 Гистограмма абсолютных частот, построенная по данным таблице 1.2

На рисунке 1.2 представлена гистограмма распределения накопленных относительных частот ($P=\sum p_i, \%$), приведенных в таблице 1.2.

При построении гистограммы по оси абсцисс откладываются значения переменной в выбранных градациях (в данном случае - температуры), по оси ординат - накопленные относительные частоты градаций ($P=\sum p_i, \%$). Гистограмма состоит из смежных прямоугольников, основаниями которых

являются отрезки между последовательными математическими границами выбранных градаций, а высотами - накопленные частоты каждой градации.

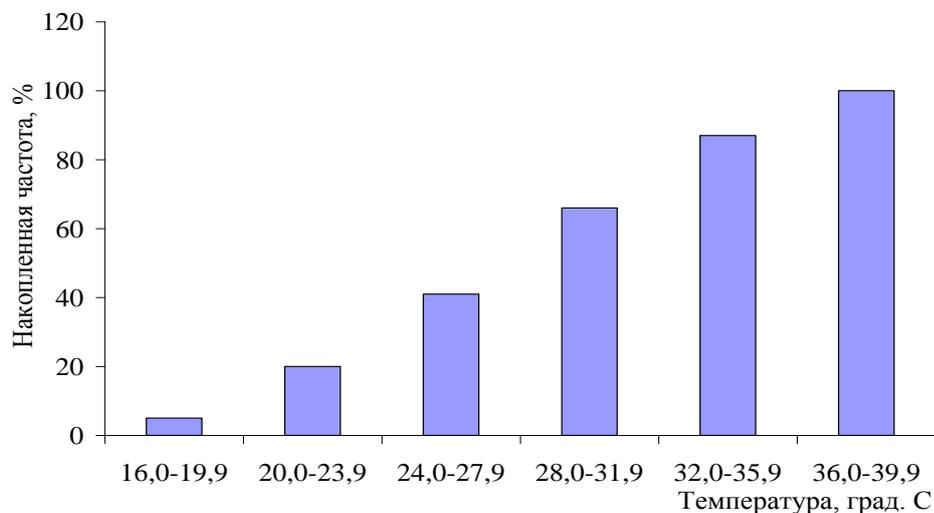


Рисунок 1. 2 Кумулятивная гистограмма, построенная по накопленным относительным частотам по данным таблице 1.2.

При построении гистограммы по оси абсцисс откладываются значения переменной в выбранных градациях (в данном случае - температуры), по оси ординат - накопленные относительные частоты градаций ($P = \sum p_i, \%$). Гистограмма состоит из смежных прямоугольников, основаниями которых являются отрезки между последовательными математическими границами выбранных градаций, а высотами - накопленные частоты каждой градации.

Порядок выполнения задания 1.2

1. Рассчитать абсолютные и относительные частоты в соответствии с заданием 1.
2. Построить гистограмму абсолютных частот.
3. Построить гистограмму относительных частот.

Проверочные тесты по лабораторной работе 1.

1. По какой формуле рассчитываются накопленные частоты?
 - a) $P = \sum p_i, \%$
 - b) $P = \sum p_i / n, \%$

c) $P = \sum p_i \cdot n \%$

Ответ: 522

2. По какой формуле рассчитываются относительные частоты?

a) $p_i = m_i \cdot n \%$

b) $p_i = m_i \cdot n \%$

c) $p_i = \frac{m_i}{n} 100 \%$

Ответ: 225

3. По какой формуле рассчитывается ширина класса?

a) $\Delta X = (X_{\max} - X_{\min}) - k$,

b) $\Delta X = (X_{\max} - X_{\min}) / n$

c) $\Delta X = (X_{\max} - X_{\min}) / k$

Ответ: 225

Лабораторная работа 2.

Тема: Параметры распределения

Цель работы: Научиться рассчитывать различные параметры распределения, их ошибки и анализировать их.

Задание 2.1. Рассчитать статистики положения центра распределения случайной величины - среднюю, медиану, моду

Метеорология и климатология оперирует данными фактических наблюдений – метеорологическими рядами – за имеющийся период наблюдений от нескольких лет до 100 иногда до 150-200 лет. Считается, что ансамбль погодных условий, описывающих климат заданного момента времени, остается практически постоянным в течение 30-50 лет (в среднем около 40 лет), а между смежными 40-летиями может уже существенно отличаться, обуславливая изменчивость климата. Поэтому, различного рода многолетние средние, называемые **климатическими нормами**, обычно рассчитываются за период 30 лет и более.

Статистиками положения центра распределения случайной величины являются следующие характеристики - средняя, мода, медиана.

Одной из важнейших обобщающих характеристик метеорологического ряда является **средняя величина** (X_{cp}):

- для несгруппированной выборки

$$X_{cp} = (x_1 + x_2 + x_3 + \dots + x_n) / n = \sum x_i / n, \quad (2.1)$$

- для сгруппированной выборки

$$X_{cp} = 1/n \sum x_{i\ cp} \cdot m_i$$

где n – сумма членов в ряду, i - номер члена в ряду; $x_{i\ cp}$ - середина i –того класса сгруппированной выборки, m_i – объем класса.

Значение **средней** заключается в ее свойстве нивелировать индивидуальные различия, в результате чего выступает более или менее устойчивая числовая характеристика признака. Средняя величина характеризует групповые свойства и является центром распределения.

Во многих случаях в качестве обобщающих характеристик совокупности оказываются полезными так называемые структурные средние, которые занимают особое место в ряду распределения. К структурным средним обычно относят медиану (Me), моду (Mo).

Медианой называется значение случайной величины (Me), которое делит распределение пополам по накопленной частоте или вероятности, т.е. так, что слева и справа от нее лежит 50% объема распределения. Медиана это значение метеорологической величины, накопленная частота которой равна 50%

Медиана стоит в центре ранжированного ряда, который расположен в порядке убывания или возрастания значений x_i . При этом число единиц совокупности с большим или меньшим, чем медиана значением ряда одинаково.

Если всем единицам ряда придать порядковые номера, то номер медианы в ряду с нечетным числом членов « n » определяется как $(n+1)/2$. Так, в ряду из 81 члена номер медианы $(81 + 1)/2 = 41$, т. е. медианой является значение ряда, стоящее под номером 41.

Если число членов в ряду четное, то медиану приходится определять как среднюю из двух центральных значений ранжированного ряда, порядковые номера которых $n/2$ и $n/2+1$. Так, если в ряду 80 значений, то центральными будут ранжированные значения с порядковыми номерами $80/2 = 40$ и $80/2+1=41$. При большем числе случаев (более 100) выбирают значение медианы как x_i с порядковым номером $n/2$.

Медиану рекомендуется определять в дополнение к средней арифметической при асимметричных распределениях.

Медианный интервал, определяется по накопленным частотам. Внутри интервала медиана находится по формуле:

$$Me = x_n + k * (\sum m/2 - \sum m_{n-1}) / m_{me} \quad (2.2)$$

где Me — медиана, x_n — нижняя граница медианного интервала, k — величина интервала (градации), $\sum m$ - сумма частот, $\sum m/2$ – номер медианы в ряду, $\sum m_{n-1}$ – накопленная частота для конца интервала, предшествующего медианному, m_{me} – частота медианного интервала.

Таблица 2.1

Оценка модального интервала при ширине градации $\Delta x = 0,5^\circ\text{C}$

Температура воздуха, $^\circ\text{C}$	Число лет (частота градации) m_i	Накопленные частоты, $\sum m_i$	Относительная повторяемость $p_i, \%$	Накопленная повторяемость, $\sum p_i, \%$
16,0-16,5	3	3	6	6
16,5-17,0	7	10	14	20
17,0-17,5*	16	26	32	52
17,5-18,0	10	36	20	72
18,0-18,5	9	45	18	90
18,5-19,0	3	48	6	96
19,0-19,5	2	50	4	100
Всего	$\sum m_i = n = 50$		100	100

* - модальный интервал

По данным среднемесячных температур из таблицы 2.1 рассчитаем медиану по формуле 2.2:

$$Me = 17,0 + 0,5(50/2 - 10)/16 = 17,4$$

Формула медианы (2.2) применима для любого интервального ряда — с равными и неравными градациями.

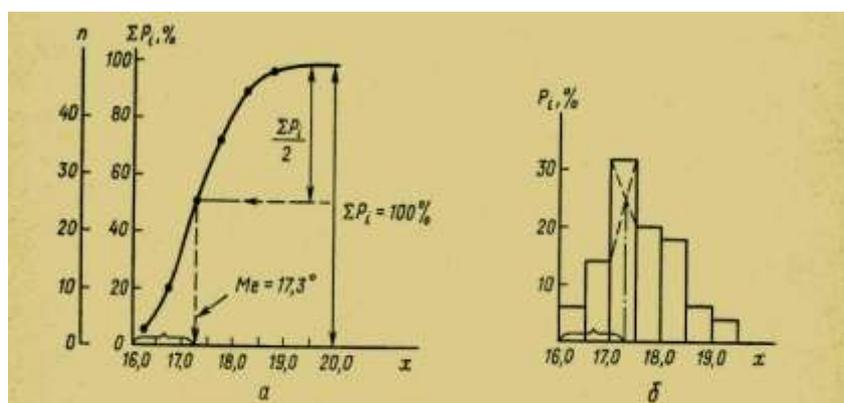


Рис.2.1. Графическое определение медианы (а) и моды (б)

Мода – наиболее часто встречающееся в совокупности значение. Моду рекомендуется определять при резко асимметричных распределениях, для которых среднее арифметическое не является типичным значением элемента в том смысле, что наибольшая повторяемость не приходится на интервалы значений случайной величины, близкие к средней арифметической.

Модальный интервал — это интервал с наибольшей повторяемостью. В таблицы 2.1 модальный интервал: 17,0° - 17,5°С. Внутри интервала мода определяется по формуле (2.3):

$$M_o = x_n + k * (m_2 - m_1) / [(m_2 - m_1) + (m_2 - m_3)], \quad (2.3)$$

где M_o — мода, x_n - нижняя граница модального интервала, k – величина интервала, m_1 – частота интервала, предшествующего модальному, m_2 – частота модального интервала, m_3 – частота интервала, следующего за модальным. По данным таблицы 2.1. получаем:

$$M_o = 17,0 + 0,5 (16 - 7) / [(16 - 7) + (16 - 10)] = 17,3.$$

Формула (2.3) пригодна только для расчета моды в рядах с равными интервалами.

Для графического определения медианы последнюю ординату кумулятивной кривой делят пополам. Через полученную точку проводят прямую, параллельную оси X , до пересечения ее с кумулятой. Точка пересечения абсциссы с перпендикуляром и будет медианой (рисунок 2.1 а).

Для графического определения моды используют три соседних столбца гистограммы: самый высокий и два прилегающих к нему слева и справа (рисунок 2.1 б). Внутри столбца, соответствующего градации с наибольшей частотой, проводят прямые линии так, как это сделано на рисунке 2.1 б. Абсцисса пересечения этих прямых будет модой распределения.

Для не очень асимметричных и одновершинных распределений связь между модой, медианой и средней определяется соотношением:

$$M_o = x_{cp.} + 3 (M_e - x_{cp.}). \quad (2.4)$$

Порядок выполнения задания 2.1.

1. Исходный метеорологический ряд ранжировать в возрастающем порядке.
2. Ранжированный метеорологический ряд разбить по градациям.
3. Подготовить и заполнить таблицу (аналогично таблице 2.1).
4. Рассчитать накопленные частоты Σm_i .
5. Рассчитать относительные частоты p_i .
6. Рассчитать накопленные относительные частоты P .
7. По накопленным относительным частотам построить гистограмму и по ней графически определить численное значение медианы (рисунок 2.1).
8. Рассчитать медиану Me по формуле 2.2.
9. Рассчитать моду по формуле 2.3.
10. Полученные результаты представить в таблице 2.2.
11. Рассчитать среднее арифметическое значения по формуле 2.1.

Таблица 2.2

Рассчитанные статистики по метеорологическому ряду
 _____ за период _____
 по метеостанции _____

Статистики	\bar{X}_{cp}	Me	Mo	D	σ	As	Ex	R	d	C_v
Значение										
Ошибка										

Задание 2.2. Рассчитать статистики рассеивания, асимметрии и эксцесса распределения случайной величины

Характеристиками степени рассеивания случайной величины около центра являются **дисперсия и среднее квадратическое отклонение**. **Дисперсия распределения является мерой разброса распределения случайных величин относительно среднего значения** данного ряда и определяется по формуле:

$$D = \frac{\sum (x_i - x_{cp.})^2}{n} \quad (2.5)$$

Дисперсия имеет размерность квадрата случайной величины $[x^2]$, что неудобно на практике. Поэтому, если из дисперсии извлечь квадратный корень с положительным знаком, то получим характеристику рассеивания с размерностью случайной величины. Эта характеристика рассеивания получила название **среднего квадратического отклонения σ** :

$$\sigma = D^{\frac{1}{2}} = \sqrt{\frac{(x_i - x_{cp.})^2}{n}}. \quad (2.6)$$

Большинство распределений не симметричны относительно $x_{cp.}$. Чтобы охарактеризовать степень асимметрии вводится специальная безразмерная характеристика - коэффициент асимметрии или скошенности распределений. Обычно он обозначается через As .

Коэффициентом асимметрии As называется безразмерное отношение, численно равное отношению среднего куба отклонения величины x_i от среднего арифметического к кубу среднего квадратического отклонения:

$$As = \frac{\sum (x_i - x_{cp.})^3}{n\sigma^3}. \quad (2.7)$$

Если $As > 0$, то правая ветвь распределения длиннее левой и говорят, что оно правоасимметрично или имеет положительную асимметрию. Если $As < 0$ то распределение левоасимметрично или имеет отрицательную асимметрию. Для симметричных распределений $As = 0$

На практике различают почти симметричные распределения, слабо, умеренно и сильно асимметричные распределения. В эти качественные понятия в зависимости от типа распределения, характера решаемой задачи и объема исходных данных вкладываются различные численные значения модулей As . Ориентировочно во многих реальных метеорологических задачах можно условно считать:

$As < 0$ - практическое отсутствие асимметрии, $As = 0,1 - 0,2$ - слабая асимметрия, $As = 0,3 - 0,5$ - умеренная асимметрия, $As > 0,6-1,0$ - сильная асимметрия, $As > 1,0$ - очень сильная асимметрия

Еще одна важная характеристика формы распределений - коэффициент эксцесса, который характеризует крутость или островершинность $f(x)$.

Коэффициентом эксцесса E_x называется безразмерное выражение:

$$A_s = \frac{\sum (x_i - x_{cp.})^4}{n\sigma^4} - 3 \quad (2.8)$$

Для нормального закона распределения, который имеет самое широкое применение в теории и на практике, $E_x = 3$. Таким образом, для нормального закона $E_x = 0$, для всех остальных законов распределений E_x может быть больше или меньше нуля.

Приняты следующие градации для характеристики островершинности эмпирических распределений:

$|E_x| < 0,2$ - практически эксцесс отсутствует,

$|E_x| = 0,2-0,3$ - слабый эксцесс, $|E_x| = 0,3-0,6$ - умеренный эксцесс,

$|E_x| = 0,6-1,0$ - сильный эксцесс, $|E_x| > 1$ - очень сильный эксцесс.

Если $|E_x| < -2$, то возникает многовершинность (поли-modalность) распределения и для анализа выборку следует дифференцировать на более простые «однопричинные» выборки.

Порядок выполнения задания 2.2

1. Рассчитать для исходного ряда параметры рассеяния: дисперсию, среднее квадратическое отклонение, коэффициенты асимметрии и эксцесса по формулам 2.5-2.8.
2. Полученные результаты представить в таблице 2.2.

Задание 2.3. Рассчитать меры изменчивости ряда - размах варьирования, среднее абсолютное отклонение, коэффициент вариации

Для того чтобы судить, насколько типична средняя для данной совокупности наблюдений, ее следует дополнять показателями, характеризующими изменчивость (колеблемость, рассеивание) значений. Наиболее простыми мерами изменчивости являются R (амплитуда колебаний A) и среднее абсолютное отклонение d :

$$R = x_{\max.} - x_{\min.}, \quad (2.9)$$

$$d = (\sum |x_i - x_{cp}|) / n, \quad (2.10)$$

где x_{\max} , x_{\min} - максимальное и минимальное значение в ряду;

$|x_i - x_{cp}|$ - модуль отклонения, т. е. величина отклонения без знака, n - число наблюдений.

Размах варьирования (амплитуда колебаний) дает самое общее представление об изменчивости, так как показывает лишь, насколько отличаются друг от друга крайние значения, но не указывает, насколько велики отклонения значений внутри ряда.

Наиболее распространенными показателями изменчивости являются среднее квадратическое отклонение σ - «сигма» и дисперсия d . Эти показатели рассмотрены выше. Однако часто возникает необходимость сравнения изменчивости отдельных частей одного ряда или разнородных метеоэлементов (рядов), имеющих неодинаковую размерность. В этих случаях сопоставление осуществляется с помощью безразмерного **коэффициента вариации C_v** :

$$C_v = \sigma/x_{cp}. \quad (2.11)$$

Порядок выполнения задания 2.3

1. Для исходного ряда рассчитать параметры изменчивости: размах варьирования, среднее арифметическое отклонение, коэффициент вариации по формулам 2.9-2.11.
2. Полученные результаты представить в таблице 2.2.

Задание 2.4. Рассчитать средние квадратические ошибки основных статистик

Все численные значения статистик, получаемых по выборкам (эмпирическим распределениям), называются оценками или выборочными показателями в отличие от их точных значений или параметров, соответствующих генеральным совокупностям. Поэтому оценка любой статистики должна непременно сопровождаться оценкой ее **ошибки**. **Величину отклонения выборочного показателя от его генерального**

параметра называют статистической ошибкой или ошибкой репрезентативности. Размеры выборочных ошибок зависят главным образом от объема выборки и от размера варьирования признака. При объемах выборки $n > 50$ ошибки могут быть записаны следующим образом:

1. Ошибка среднего значения $\sigma_{x_{cp.}}$:

$$\sigma_{x_{cp.}} = \frac{\sigma(x)}{\sqrt{n}} . \quad (2.12)$$

2. Ошибка среднего квадратического отклонения σ_{σ} :

$$\sigma_{\sigma} = \frac{\sigma(x)}{\sqrt{2n-1}} . \quad (2.13)$$

3. Ошибка коэффициента асимметрии σ_{As} :

$$\sigma_{As} = \sqrt{\frac{6}{n}} . \quad (2.14)$$

4. Ошибка коэффициента эксцесса - σ_{Ex} :

$$\sigma_{Ex} = \sqrt{\frac{24}{n}} . \quad (2.15)$$

5. Ошибка коэффициента вариации - σ_{Cv} :

$$\sigma_{Cv} = C_v \sqrt{\frac{1+c_v^2}{2n}} . \quad (2.16)$$

Приведенные формулы ошибок предназначены для нормального закона. Для любых распределений этими формулами также можно пользоваться. Но необходимо знать, что рассчитанные величины ошибок будут приближенными.

Синтаксис записи оценок статистики и ее ошибки имеет вид: $x_{cp.} \pm \sigma(x)$; $\sigma(x) \pm \sigma_{\sigma}$ и т.д.

Порядок выполнения задания 2.4.

1. Рассчитать средние квадратические ошибки вычисленных статистик по формулам 2.12-2-16.
2. Полученные результаты представить в таблице 2.2.

Проверочные тесты по лабораторной работе 2.

1. По какой формуле рассчитывается дисперсия?

a) $D = (x_i - x_{cp.})^2 / n$

b) $D = \Sigma (x_i - x_{cp.})^2 / n$

c) $D = \Sigma (x_i - x_{cp.}) / n$

Ответ: 522

2. По какой формуле рассчитываются среднее квадратическое отклонение?

a) $\sigma = \Sigma (x_i - x_{cp.})^2 / n$

b) $\sigma = [\Sigma (x_i - x_{cp.})^2 / n]^{1/2}$

c) $\sigma = [\Sigma (x_i - x_{cp.}) / n]^{1/2}$

Ответ: 252

3. По какой формуле рассчитывается коэффициент вариации?

a) $C_v = 1/x_{cp.}$

b) $C_v = \sigma / \Sigma x_i$

c) $C_v = \sigma / x_{cp.}$

Ответ: 225

Лабораторная работа 3

Тема: Статистическая оценка метеорологической информации, полученной с помощью автоматической метеостанции Vantage Pro (WS)

Цель работы: получить представление о работе автоматического метеорологического комплекса Vantage Pro. Освоить технологию получения некоторых гидрометеорологических параметров и провести их статистическую оценку.

Общие сведения:

Автоматическая станция **Vantage Pro** позволяет проводить постоянный мониторинг метеорологических величин. Многофункциональным элементом станции является консоль, которая осуществляет сбор, хранение, и демонстрацию архивных и текущих данных погоды. На дисплее консоли (рисунок 3.1) отображаются в цифровом и графическом виде некоторые метеорологические параметры: атмосферное давление, температура воздуха, относительная влажность воздуха, скорость и направление ветра, количество и интенсивность осадков и другие, необходимые для статистической оценки и прогноза.



Рисунок 3.1 Внешний вид консоли в режиме текущей погоды

Клавиатура, расположенная справа (рисунок 3.1), позволяет просматривать текущую и архивную информацию, изменять режимы работы станции, осуществлять калибровку, настраивать и просматривать графики, выбирать датчики, читать прогноз и т.д. Некоторые данные на экране могут отображаться в разных единицах измерения. Например, температура в $^{\circ}\text{C}$ (градусы Цельсия), $^{\circ}\text{F}$ (градусы Фаренгейта) и в К (Кельвинах). Атмосферное давление может быть показано в дюймах (in), миллиметрах (мм), миллибарах (мб) или гектопаскалях (гПа). Экран текущей погоды на мониторе показывает до десяти погодных переменных одновременно.

Станция Vantage Pro сопряжена с компьютером, на котором установлено программное обеспечение Weather Link, что дает возможность использовать дополнительные функции станции по обработке и накоплению материалов. Программа Weather Link визуально представляет картину распределения метеорологических величин в течение различных временных диапазонов (час, 12 часов, сутки, трое суток, неделя, месяц, год). Например, на рисунке 3.2 показано распределение температуры и давления в течение суток 4 ноября 2007 года, а на рисунке 3.3 – в течение месяца (октябрь 2007).



Рисунок 3.2 Распределение температуры (красная линия) и атмосферного давления (серая линия) в течение суток

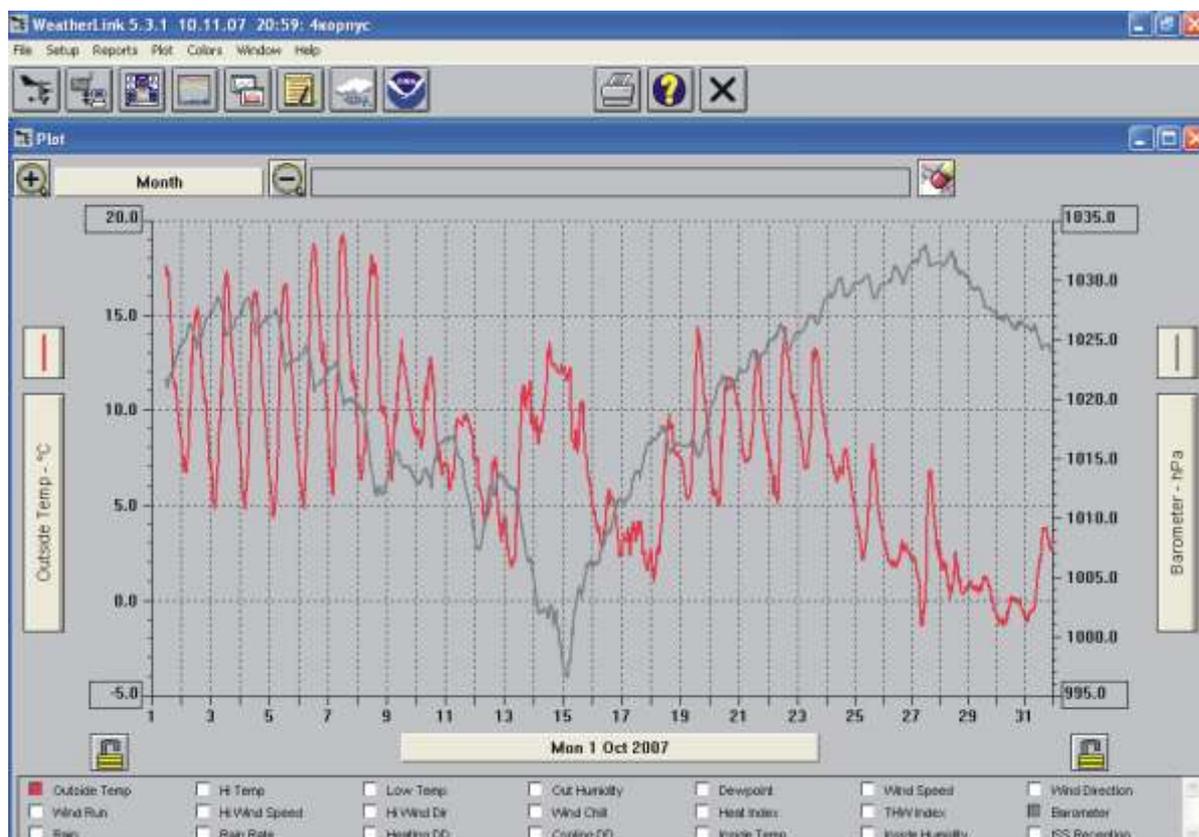


Рисунок 3.3 Распределение температуры (красная линия) и атмосферного давления (серая линия) в течение октября 2007 года

Задание 3.1 *Анализ метеорологической информации полученной с помощью автоматической метеорологической станции WP*

Содержание задания:

1. С консоли снять текущие данные основных метеорологических параметров: давления, температуры, влажности, скорости и направления ветра.
2. Осуществить выборку сроков наблюдений и метеорологических величин с помощью программы Weather Link.
3. По графику суточного хода температуры определить среднесуточную, максимальную и минимальную (экстремальные значения) температуры и их амплитуду.
4. Провести аналогичный анализ для распределения давления.
5. По графику изменения температуры и давления за месяц определить среднемесячные величины, выбрать максимум и минимум, дать общий

анализ. Оценить изменение метеорологических величин в течение трех суток, недели или года.

- По рекомендации преподавателя сопоставить суточный ход метеоэлементов (температуры, давления) в разных барических системах (циклонах, антициклонах). Через архив программы ГИС-Метео подобрать карты положения разных барических образований, занимающих район Поволжья. С помощью архива программы Weather Link построить графики суточного хода метеорологических величин в соответствии с датами подобранных карт. При анализе акцентировать внимание на среднесуточные значения метеовеличин, их амплитуду, возможные колебания времени наступления экстремальных значений этих величин.

Выполнение пункта 6 задания 3.1 может быть рекомендовано на второй ступени двухуровневого образования.

Исходные данные: Фактические данные различных метеорологических величин. Графики суточного, месячного хода температуры и давления. Приземные карты погоды.

Порядок выполнения задания 3.1

- Обратить внимание на графики суточного хода температуры и давления. На шкале времени (ось X) после 12 часов временные интервалы обозначаются числами от 1 до 12, что соответствует интервалу времени от 13 до 24 часов. На рисунке 3.2 шкала температуры (°C) показана слева, а шкала давления (гПа) справа.
- Для вычисления среднесуточных величин необходимо снять с графика (рисунок 3.2) значения температуры (давления) за основные синоптические сроки 00, 03, 06.....21, 24 часа, т.е. через каждые три часа. Полученные значения температуры просуммировать и разделить на 8.
- Для расчета амплитуды необходимо выбрать самое высокое значение температуры (давления) и низкое за сутки (рисунок 3.2). Разница между полученными величинами и есть амплитуда.
- Для общей оценки изменения температуры и давления посмотреть, как ведет себя температура (давление) в утренние, дневные, вечерние и ночные часы.

5. Для вычисления среднемесячной температуры с графика (рисунок 3.3) снять значения температуры за каждые сутки, которые затем суммируются и делятся на число дней.
6. Анализ давления проводится по аналогичной схеме.

Проверочные тесты по лабораторной работе 3

1. На графике суточного хода метеорологических величин визуально можно обнаружить

- 1) Среднесуточные значения;
- 2) Значения экстремальных величин;
- 3) Амплитуду значений;

Ответ: 252

2. Годовая амплитуда температуры воздуха это

- 1) Разность температур в течение одного месяца;
- 2) Разность среднемесячных температур самого холодного и самого теплого месяца;
- 3) Разность температур в течение суток

Ответ: 252

3. Годовая амплитуда давления воздуха это

- 1) Разность давления между самого холодным и самым теплым месяцем;
- 2) Разность между самым высоким и самым низким значением давления в течение года;
- 3) Разность давления в течение суток

Ответ: 252

4. Суточный ход изменения метеоэлементов - это представление информации

- 1) в графическом виде
- 2) текстовом
- 3) логическом

Ответ: 522

Лабораторная работа 4

Тема: Статистический анализ полей и явлений погоды по метеорологическим картам

Цель работы: Познакомить студентов с возможностями геоинформационной системы ГИС-Метео представлять метеорологическую информацию на картах в различных формах, освоить простейшие способы обработки и анализа синоптических карт.

Общие сведения:

В метеорологии при обработке синоптических карт анализируется большой объем информации, поэтому на современном этапе в сети Росгидромета используются ГИС технологии.

Программа ГИС-Метео позволяет изготовить географическую основу карты (бланк) любой территории и с помощью компонент программы осуществить выборку метеорологических параметров из базы данных и нанести их на карту в различных формах представления (цифры, изолинии, цветное поле и др.).

Главными синоптическими компонентами формирующими карту являются: **Приземная наноска, Синоптика, Погода, Значки, Линии.**

Компонента **Приземная наноска** отображает на экране данные наблюдений, поступающие в коде КН-01 (SYNOP), в виде стандартных метеорологических пуансонов (рисунок 4.1). Информация отбирается из базы данных за указанный срок для каждой метеорологической станции, попавшей в границы текущего слайда. Поскольку наносимый пуансон имеет довольно большие размеры, количество станций, отображаемых на экране дисплея, существенно зависит от масштаба. Схема наноски (рисунки 1 и 2) и расшифровка параметров была указана в водной части.

С помощью компоненты **Синоптика** можно нанести на карту отдельные метеорологические параметры, используя любые формы представления: **Цифры, Цветное поле, Изолинии.** Одной командой **Синоптика** на слайд может быть нанесен только один из параметров. Для совмещения на слайде нескольких параметров необходимо выполнить команду **Синоптика** для каждого параметра отдельно. На рисунке 4.2 показан образец карты в форме представления **Изолинии.**

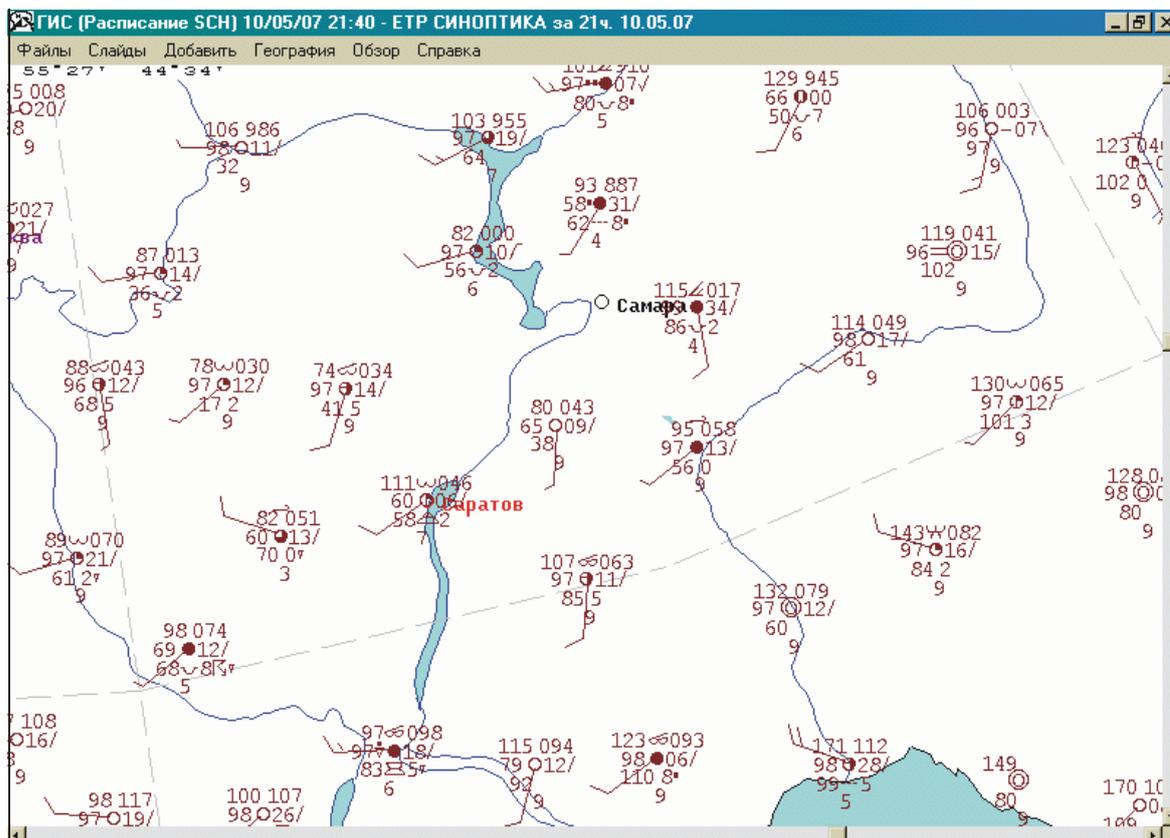


Рисунок 4.1 Приземная наноска метеорологической информации

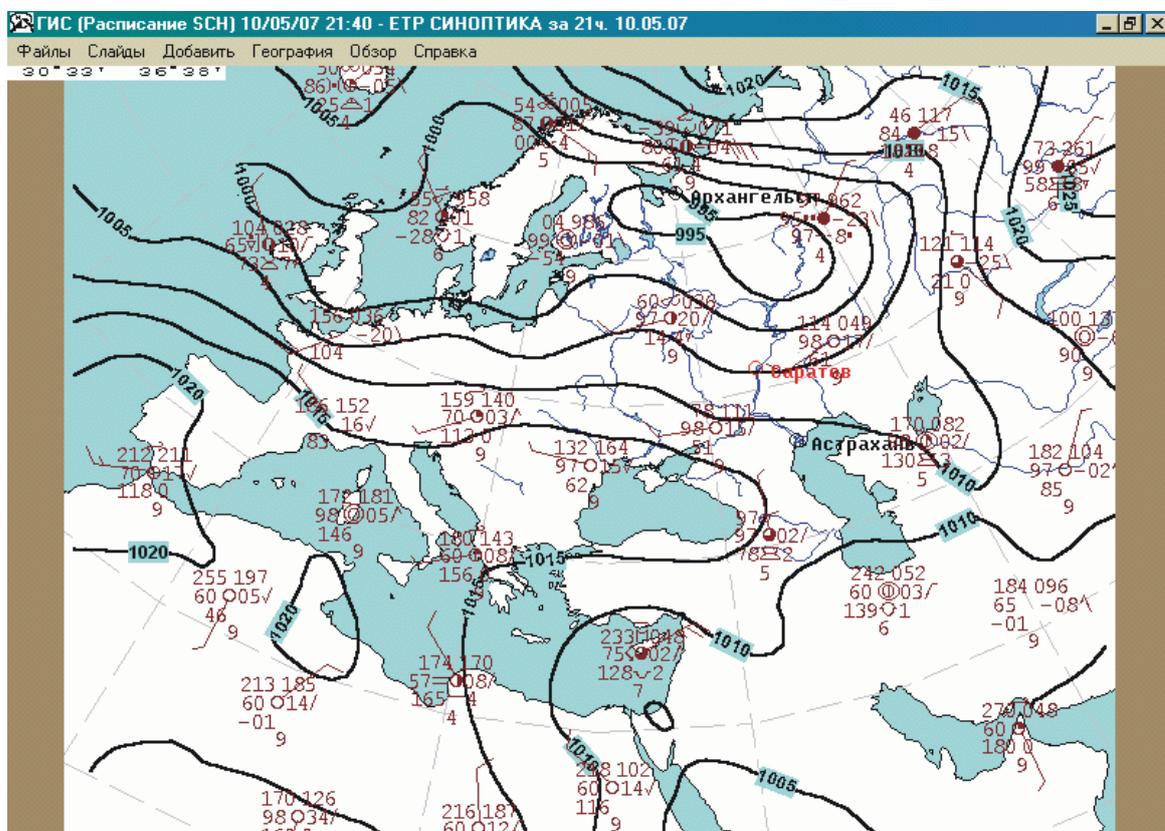


Рисунок 4.2 Проведение изолиний атмосферного давления (изобры) с помощью компоненты Синоптика программы Гис-Метео

Для обработки карты можно воспользоваться компонентой **Значки**, которая позволяет наносить явления погоды, выбирая их из списка.

Завершающим этапом обработки является нанесение на карту, в соответствии с правилами, атмосферных фронтов. В программе ГИС-Метео это можно осуществить с помощью компоненты **Линии**.

Следует отметить, что возможности представления информации в программе ГИС-Метео очень большие. Можно не только выбирать любые формы представления, но и совмещать их для одного или нескольких параметров. При этом можно изменять масштаб и выбирать любые регионы для рассмотрения в зависимости от запросов потребителей.

Первичный анализ (обработка) синоптических карт включает в себя: проведение изолиний, в частности, проведение изобар, выявление положения центров циклонов, антициклонов, их размеров, определение барических ложбин и гребней, линий атмосферных фронтов, разграничивающих воздушные массы различных свойств. Анализ барического поля имеет особое значение, потому что закономерности пространственного распределения давления одновременно характеризуют и закономерности поля ветра, т.к. поле давления и ветра тесно связаны. Циклоны и антициклоны являются не только определенными барическими системами, но и системами ветра. В циклонах в северном полушарии ветры дуют против часовой стрелки, в антициклонах по часовой.

Задание 4.1 Проведение изобар на карте погоды

Содержание задания: Проведение изобар – линий равного давления является одной из важнейших операций обработки приземных карт погоды. При проведении изобар (или других изолиний) необходимо определить значение метеорологических величин в точках, не совпадающих с положением пунктов измерения. Другими словами, интерполировать значение давления между пунктами наблюдений. Обычно такая интерполяция при работе с картами на бумажной основе производится визуально. Результаты интерполяции зависят от квалификации специалиста.

Исходные данные: Фрагмент карты, полученной по программе ГИС-Метео с нанесенным данными атмосферного давления.

Порядок выполнения лабораторной работы 4

1. Обратить внимание на схему нанесения метеорологических данных на карте погоды (см. вводную часть). Расшифровать приземную наноску для нескольких станций.
2. Правильно прочесть давление на станциях. Оно нанесено справа сверху относительно пуансона станции с десятичными долями гПа.
3. Простым карандашом, интерполируя (находя нужное давление между значениями давления на соседних станциях) провести изобары через 5 гПа, кратные пяти, например, 1010, 1005, 1000, 995 гПа (рисунок 4.3). Изобары следует проводить гуще там, где сильнее ветер.

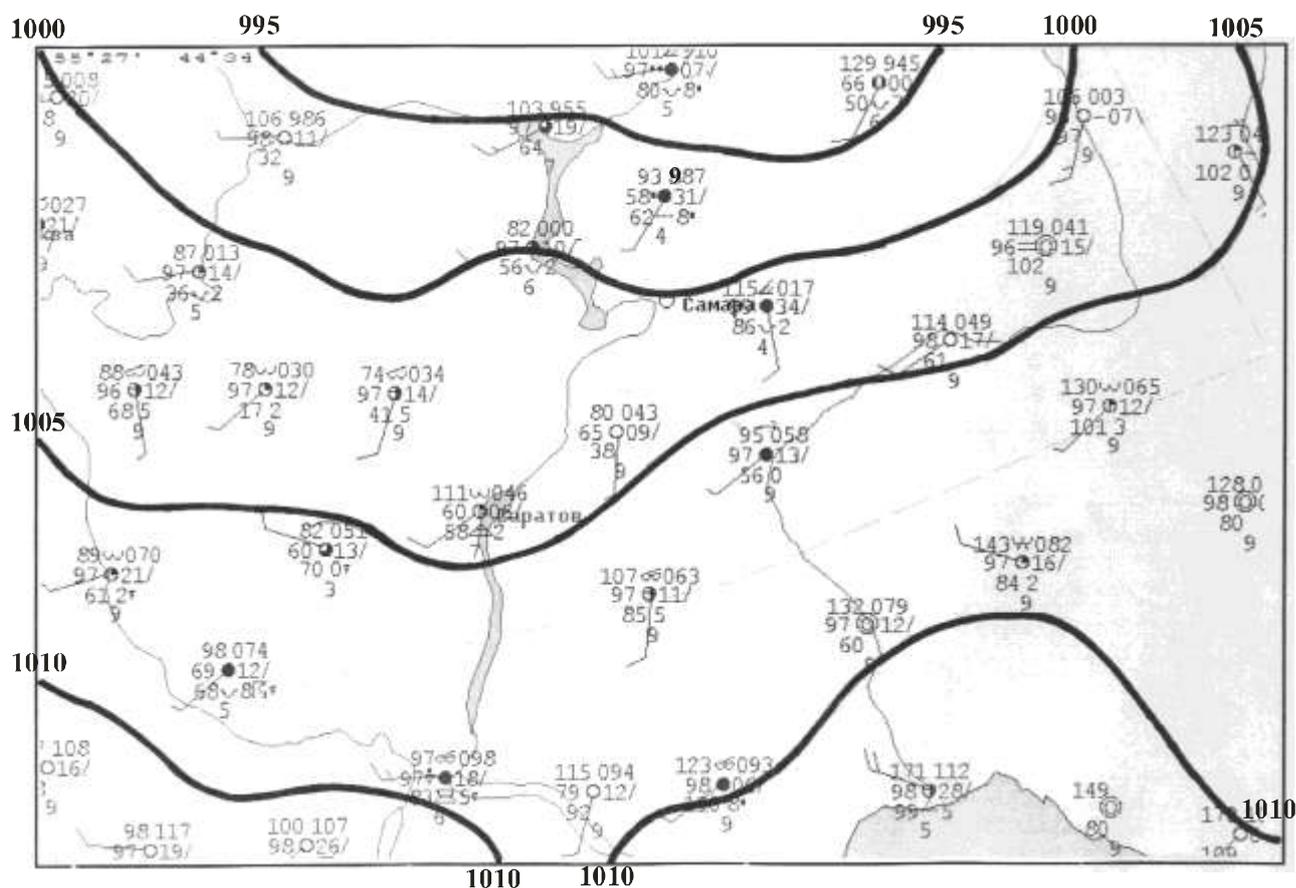


Рисунок 4.3 Пример проведения изобар на карте погоды принятой по ГИС-Метео

2. Выявить принципиальные различия в погоде в циклонах и антициклонах, обратив внимание на наличие (отсутствие) облачности, систему ветра и другие характеристики погоды по усмотрению преподавателя.
3. Сопоставить по картам положение барических образований за два последовательных срока и отметить произошедшие временные изменения. Определить интенсивность барических образований (по значению давления в центре), область распространения, характер погоды.

Выполнение данного пункта задания может быть рекомендовано на второй ступени двухуровневого образования.

Проверочные тесты по лабораторной работе 4

1. Изобары на приземных картах проводятся через
 - 1) 2 гПа
 - 2) 3 гПа
 - 3) 5 гПа
 - 4) 6 Гпа.Ответ: 2252

2. Циклон – это область
 - 1) пониженного давления
 - 2) повышенного давления
 - 3) давления без изменения.Ответ: 522

3. Давление воздуха у земной поверхности составляет 1000 гПа. Определить величину давления в миллиметрах ртутного столба.
 - 1) 760
 - 2) 750
 - 3) 745
 - 4) 765Ответ: 2522

4. Данные температуры – это данные следующего типа:

- 1) числового целого
- 2) числового действительного
- 3) логического

Ответ: 252

5. Какая компонента является первичной при формировании синоптической карты

- 1) Синоптика
- 2) География
- 3) Приземная наноска

Ответ: 252

6. Сколько параметров может быть нанесено одновременно при выполнении команды **Синоптика**

- 1) один
- 2) два
- 3) три
- 4) неограниченно

Ответ: 5222

Литература:

1. Верещагин М.А., Наумов Э.П., Шанталинский К.М. Статистические методы в метеорологии. – Казань. Изд-во Казгу, 1990. - 109 с.
2. Волынцева О.И., Смирнова А.А. Анализ и прогноз погоды с помощью ГИС Метео. Обнинск: Изд-во ГУ «ВНИИГМИ-МЦД», 2005. - 189с.
3. Исаев А.А. Статистика в метеорологии и климатологии. М.: Изд-во МГУ, 1988. - 245 с.
4. Кобышева Н.В., Наровлянский Г.Я. Климатологическая обработка метеорологической информации. - Л.: Гидрометеиздат, 1972. - 209 с.
5. Пановский Г.А., Брайер Г.В. Статистические методы в метеорологии. - Л.: Гидрометеиздат, 1972. - 209 с.
6. Подрезов О.А. Методы статистической обработки и анализа гидрометеорологических наблюдений.- Бишкек: Изд. КРСУ. 2003.- 262 с.
7. Сикан А.В. Методы статистической обработки гидрометеорологической информации. – СПб.: изд. РГГМУ. 2007.- 278с.
8. Тудрий В.Д. Методы статистической обработки гидрометеорологических наблюдений.- Казань. Изд-во Казгу, 2007. - 164 с.

9. Шкляева Л.С. Статистические методы в метеорологии. – Пермь: Перм. ун-т. - 2003. - 128 с.
10. Чичасов Г.Н. Численные методы обработки и анализа информации. – Алматы: Казгидромет, 1995. - 106 с.