



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ
ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра Экспериментальной физики атмосферы

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(бакалаврская работа)

На тему «Анализ временных рядов метеорологических величин по данным метеорологической станции Константиновский»

Исполнитель Ахиза Анастасия Павловна
(фамилия, имя, отчество)

Руководитель кандидат физико-математических наук, профессор
(ученая степень, ученое звание)

Кузнецов Анатолий Дмитриевич
(фамилия, имя, отчество)

«К защите допускаю»
Заведующий кафедрой

(подпись)

доктор физико-математических наук, профессор
(ученая степень, ученое звание)

Кузнецов Анатолий Дмитриевич
(фамилия, имя, отчество)

Санкт-Петербург
2020

Оглавление

Введение	3
Глава 1 Организация проведения метеорологических наблюдений на метеорологической станции Константиновский.....	5
1.1 Обзор исследований и разработок по теме	5
1.2 Общая характеристика Мыса Константиновский	6
1.2.1 География и климат в районе метеорологической станции Константиновский	10
1.2.2 Техническое оснащение метеорологической станции Константиновский.....	10
1.2.3. Методика проведения метеорологических наблюдениям.....	14
Глава 2 Формирование архива метеорологической информации.....	17
Глава 3. Анализ временных рядов метеорологических величин.....	18
3.1. Характеристика временных рядов	18
3.2. Анализ временных рядов	19
Литература.....	46

Введение

Предсказание погоды с научной точки зрения – одна из сложнейших задач физики атмосферы. Существуют различные методы для прогнозирования метеорологических явлений и их величин, например, синоптические, численные, статистические методы, но в полном объеме ни один метод не обеспечивает пока точного прогноза. Именно поэтому исследования в области прогнозирования погодных условий являются важными и полезными, а тема – актуальной.

Температура воздуха у поверхности земли является одной из важнейших метеорологических величин. Без знания фактического значения этой метеорологической величины, а также ее прогностического значения трудно себе представить жизнь и хозяйственную деятельность современного общества. В своей повседневной жизни мы интересуемся прогностическим ее значением вечером, планируя как свою одежду, так и свою работу, если она будет протекать на открытом воздухе. Утром мы выясняем фактическое значение этой температуры, тем самым уточняя ее прогностическое значение. В своей хозяйственной деятельности трудно себе представить тот случай, когда этот метеорологический параметр или его прогноз различной заблаговременности не учитывался бы в той или иной мере. Это коммунальные службы, авиация, дорожные службы, прогностические подразделения гидрометеослужбы, строительство, энергетика и многие другие виды хозяйственной деятельности /1,2/.

Все выше сказанное определяет важность и актуальность исследования «поведения» временных рядов /3/ приземных значений температуры воздуха и этим же определяется актуальность и практическая значимость данной работы.

Цель дипломной работы – изучение возможности прогнозирования метеопараметров по их временным рядам и, в идеале, формирование и отработка на их основе методики расчета краткосрочного прогноза погоды.

Такие временные ряды предоставляются, в частности, автоматизированными метеостанциями типа VantagePro 2 [2]. Одна из таких установлена на Мысе Константиновский.

Данная работа заключается в анализе и использовании временных рядов метеовеличин для создания модели прогноза погоды адаптированной под регион Ненецкого автономного округа.

Задачи, решаемые в работе:

- изучение литературных источников по методам прогнозирования по временным рядам;
- сбор и анализ метеоданных для установления статистических закономерностей;
- разработка математической прогностической модели;
- разработка системы прогнозирования, оценка ее адекватности.

Объектом исследования является система метеорологических показателей Мыса Константиновский.

Предметом исследования являются модели прогнозирования погоды.

Методы исследований – визуальный, статистический и спектральный анализ данных, метод ложных соседей определения аргументов модельных функций, методы аппроксимации функций многих переменных с помощью обобщенных многочленов, радиальных базисных функций, искусственных нейронных сетей и т.д.

Работа состоит из Введения, трех глав, заключения.

В первой главе мы рассматриваем обзор исследований и разработок, даем общую характеристику нашего объекта, в данном случае Мыс Константиновский,

Глава 1 Организация проведения метеорологических наблюдений на метеорологической станции Константиновский

1.1 Обзор исследований и разработок по теме

В середине XIX в. возникают государственные сети станций. А в начале века трудами А. Гумбольдта и Г. Дове в Германии закладываются основы климатологии. Около 1820 года Г. В. Брандесу в Германии пришла мысль нанести на географические карты наблюдения Маннгеймской сети станций. Таким образом, появились первые синоптические карты, позволившие обнаружить области высокого и низкого давления. Они оказались подвижные и двигались, как правило, с запада на восток. После изобретения телеграфа, с 50-х годов, по инициативе астронома У. Леверье во Франции и адмирала Р. Фицроя в Англии синоптический метод исследования атмосферных фронтов быстро вошел в общее употребление. На его основе возникла новая отрасль метеорологии – синоптическая метеорология.

К середине XIX в. относится организация первых метеорологических институтов, в том числе Главной физической обсерватории в Петербурге (1849г.). Во второй половине XIX столетия были заложены основы динамической метеорологии, т. е. применения законов гидромеханики и термодинамики к исследованиям атмосферных процессов. Большой вклад в эту область метеорологии был сделан Кориолисом и Пуассоном во Франции, В. Феррелем в США, Г. Гельмгольцем в Германии, Г. Моном и К. Гульдбергом в Норвегии. К концу столетия усилилось изучение радиационных и электрических процессов в атмосфере [3].

Проводимые в настоящее время Правительством РФ мероприятия в Арктике, направленные на повышение экономического потенциала России, а также учения с использованием авиации как основы в обеспечении, переброске и боевой поддержке войск показали, что специфические климатические условия региона оказывают существенное влияние на результаты деятельности людей и работу техники, объединенных между собой в организационно-техническую систему [19].

Вместе с тем, существует проблема гидрометеорологического обеспечения в Арктическом регионе, заключающаяся в отсутствии надежных и достоверных методик прогноза погодных условий из-за редкой сети наблюдательных станций и специфики климата Арктической зоны, в отсутствии научно-методического аппарата, позволяющего разрабатывать оптимальные рекомендации лицам, принимающим решения, на основе методов оценивания влияний экстремальных условий климата на эксплуатацию техники и жизнедеятельность человека.

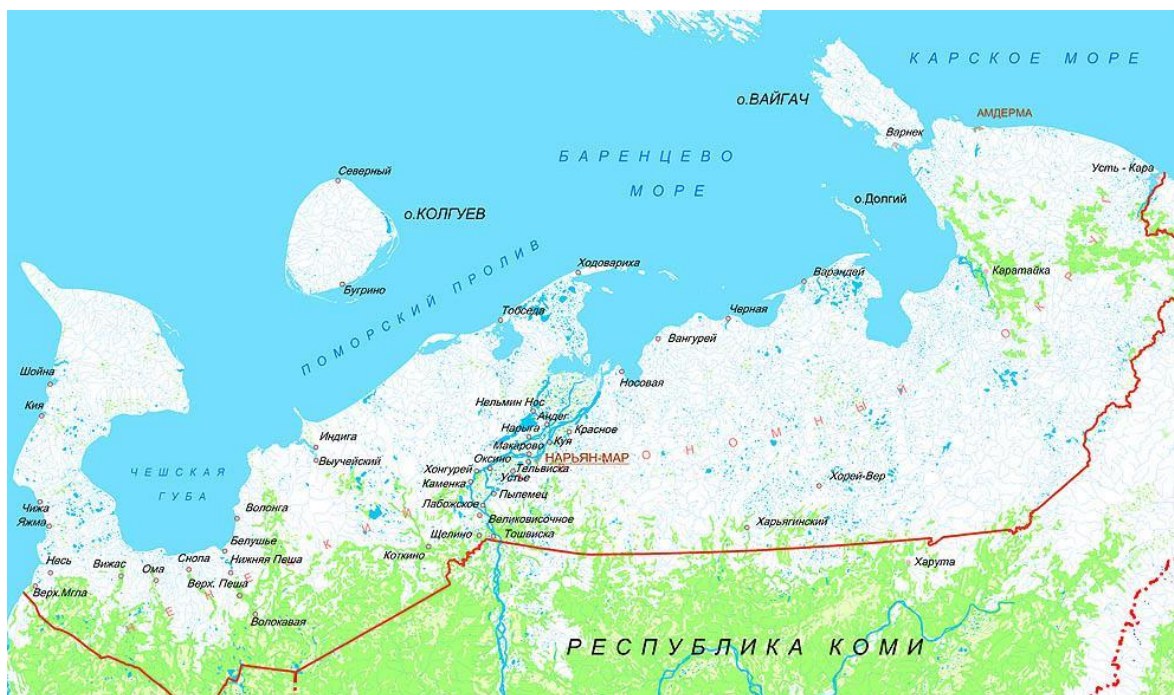
Для решения этой проблемы в высокоширотных и полярных районах предполагает проведение постоянного мониторинга погодно-климатических условий, тенденции к изменению которых отчетливо проявляются в Арктической зоне РФ [8-14]. Данные такого мониторинга являются основой для получения оценок экстремальных свойств гидрометеорологических условий при определении типов состояния погоды и разработки сценариев проводимых мероприятий, эксплуатации объектов инфраструктуры, военной и специальной техники [92-95].

1.2 Общая характеристика Мыса Константиновский

1 сентября 2019 года исполнилось 60 лет морской гидрометеорологической станции Мыс Константиновский.

В связи с закрытием ранее существовавшей станции Мыс Шведский и началом бурного освоения богатой запасами полезных ископаемых тундры Ненецкого автономного округа, развития судоходства на Европейском Севере страны, была открыта станция по программе 2 разряда МГ-2 Мыс Константиновский.

Станция расположена на юго-восточном побережье Печорского залива Баренцева моря на мысе того же названия в устье р. Печора.



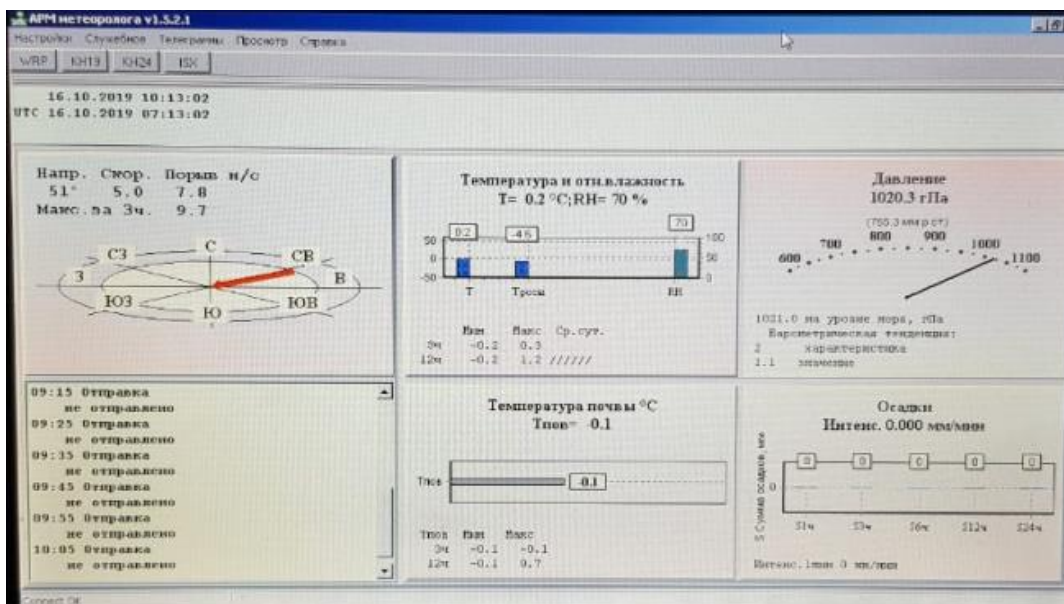
С момента открытия станция была оснащена современными по тому времени приборами: ртутным чашечным барометром для измерений атмосферного давления, термометрами и гигрометрами для измерений температуры и влажности воздуха и поверхности почвы, осадкомером Третьякова, гололедным станком. В дальнейшем проводилась модернизация станции. Были введены в эксплуатацию анеморумбометр М-63М-1 для измерений параметров ветра, ИВО-1 - для измерений высоты облачности, приборы М-53А и М-71 - измерители метеорологической дальности видимости в дневное и ночное время, ГМ-12 - измеритель параметров морского волнения.



Анеморумбометр М-63М-1

В 2007 году на станции построен новый служебный дом. Передача метеорологической информации осуществляется мобильной и спутниковой связью. В 2007 году на станции внедрен безртутный барометр БРС-1М-1. В июле 2012года установлен автоматический метеорологический комплекс (АМК).





Автоматический метеорологический комплекс АМК

Первым начальником станции был Ивахин В.И. В дальнейшем начальники неоднократно менялись (Бороновский В.С., Чижихин В.М., Длуцкий Э.К., Коваленко В.А., Колмыченко П.А., Авдеев В.И., Куроптев Я.Д., Слесарев А.С.). Длительное время возглавлял станцию Большаков Анатолий Георгиевич. За многолетний добросовестный труд в Северном УГМС впоследствии он был награжден Орденом Великой Октябрьской социалистической революции. Около 10 лет (1987-1996) начальником станции был Вдовкин Игорь Иванович. Более 20 лет трудится на станции семья Нечаевых, Сергей Николаевич и Татьяна Афанасьевна. Более 10 лет начальником станции был Нечаев Сергей Николаевич, затем руководить станцией стала Татьяна Афанасьевна. За добросовестную работу неоднократно награждались почетными грамотами Росгидромета и ФГБУ «Северное УГМС». Нечаевы Сергей Николаевич и Татьяна Афанасьевна награждены знаками «Почетный работник гидрометслужбы России». Вместе с ними работает с 24.11.10 года Ахиза Анастасия Петровна, перспективный молодой сотрудник, неоднократно замещала начальника станции на период трудовых отпусков. Знает все виды наблюдений, обслуживание дизельгенераторов и бензоагрегатов, обеспечивает содержание приборов и оборудования в надлежащем виде.

Коллектив станции обеспечивает полное выполнение плановых заданий, отличное качество материалов наблюдений и информации. Информация станции используется при составлении прогнозов погоды, а также для обслуживания отраслей экономики и хозяйства страны. Данные гидрометеорологических наблюдений помещаются в режимно-справочные пособия.

1.2.1 География и климат в районе метеорологической станции Константиновский

Станция расположена на юго-восточном побережье Печорского залива Баренцева моря на мысе того же названия в устье р. Печора.

Полярная ночь длится 4 месяца, охватывая период с ноября по январь. Замерзает море в октябре. Полностью льдом покрывается редко. Лед держится на воде до самого конца июня. Максимальная ледовитость достигается в апреле, потом лед начинает отступать на восток, исчезая только в июле. Западная часть моря всегда от льда свободно. Максимальная температура воды наблюдается в августе, когда термометр будет показывать +12 градусов по Цельсию.

1.2.2 Техническое оснащение метеорологической станции Константиновский.

Метеорологической станцией можно назвать любое место, где на протяжении длительного времени (годы) производятся простейшие метеорологические наблюдения (над ветром, облачностью, температурой и влажностью воздуха, атмосферным давлением, явлениями погоды).

Оборудование метеорологической станции и его размещение

Барометр. На станции барометр находится в помещении, вдали от нагревательных приборов. Высота станции над уровнем моря указывается вместе с показаниями барометра.



Барометр ртутный

Приборы для измерения относительной влажности воздуха

Гигрометр и психрометр. При оснащении метеорологической станции гигрометр предпочтительнее, так как в отличие от психрометра гигрометр может работать круглый год. В то же время психрометр позволяет производить более точные наблюдения за относительной влажностью воздуха. Приборы устанавливаются в метеорологической будке на высоте 2 метров над землей вдали от строений и деревьев.



Психрометрическая будка



Гигрометр и психрометр

Флюгер. На метеостанции на высоте 10 метров должен быть установлен флюгер (допускается установка на высоте 2 и 5 метров над землей). Для установки выбирают место, открытое действию ветра всех направлений.

Осадкомер. На метеорологической станции должен быть установлен осадкомер. Место установки должно быть удалено от деревьев, высоких стен и т.п., а так же чтобы он не был заслонен при косом выпадении дождя или снега. В осадкомер не должен быть подвергнут сметанию снега с близ

лежащих предметов. От осадкомера до зданий и деревьев расстояние должно превышать их высоту в 3...5 раз.



Осадкомер

1.2.3. Методика проведения метеорологических наблюдений

Порядок проведения метеорологических наблюдений

Метеорологические наблюдения проводятся по срокам, распределенным в течение суток (1,7,13,19) часов, над метеорологическими элементами: ветром, облачностью и облаками, температурой и влажностью воздуха, отмечается атмосферное давление и описываются текущие явления погоды, а так же явления погоды, наблюдавшиеся между сроками. Срочные наблюдения производятся в следующем порядке: в течение пяти минут от срока наблюдений отмечается температура и влажность воздуха, описываются форма облаков, облачность, оптическая плотность, далее на

протяжении пятнадцати минут производятся наблюдения над ветром, отмечается текущее значение атмосферного давления. Так же описываются текущие явления погоды и явления происшедшие за последние шесть часов. В 19 часов измеряется количество осадков выпавших за сутки.

Температура и влажность воздуха

Количество водяного пара, находящегося в воздухе, называется влажностью воздуха. Для характеристики влажности употребляются следующие величины:

1. Абсолютная влажность.

2. Относительная влажность.

Порядок наблюдений по психрометру:

1. За 5 минут до срочного часа смачивают ткань на термометре, для этого берут дистиллированную воду, за неимением таковой можно пользоваться чистой снеговой водой или использовать дождевую воду, предварительно пропущенную через фильтровальную бумагу или вату.

2. Через 4 минуты производят отсчет сухого и смоченного термометров психрометра.

Наблюдения по психрометру при температуре воздуха около нуля имеют следующие особенности:

1. Ткань, в этом случае, смачивают за 30 минут до наступления срока наблюдения.

2. После отсчета термометров определяется состояние ткани – «лед» или «вода». Для этой цели не отточенным концом карандаша или тонкой деревянной палочкой осторожно касаются лоскутка ткани на смоченном термометре и в зависимости от того, мягкая или твердая ткань, отмечают «в» или «л».

Волосяной гигрометр

Волосяной гигрометр предназначен для измерения относительной влажности воздуха. Действие прибора основано на свойстве обезжиренного человеческого волоса, менять свою длину в зависимости от изменения

относительной влажности окружающего воздуха. Основное назначение волосяного гигрометра – измерять влажность в морозное время, когда по психрометру влажность не определяется.

Температура воздуха отсчитывается по сухому термометру психрометра.

Давление воздуха

Для измерения давления воздуха служит барометр aneroid, в циферблат которого вмонтирован термометр для отсчета температуры прибора.

При определении атмосферного давления следует:

1. Отметить показание барометра.
2. Отметить показание термометра при барометре.
3. Исправить показание барометра за высоту места наблюдения, температуру барометра.

Глава 2 Формирование архива метеорологической информации

К первичным источникам метеорологической информации относятся различные таблицы, составляемые на метеорологических станциях. Основной таблицей метеорологических наблюдений является ТМ-1 [7]. В зависимости от программы наблюдений на станциях могут формироваться различные виды таблиц, например:

- 1) ежечасных данных температуры, влажности и атмосферного давления – ТМ-2;
- 2) ежечасных данных скорости и направления ветра – ТМ-2б, ТМ-2в;
- 3) данных по температуре почвы – ТМ-3, ТМ-3а;
- 4) данных инструментальных наблюдений над обледенением проводов – ТМ-5;
- 5) данных количества росы – ТМ-10;
- 6) данных интенсивности дождя – ТМ-14;
- 7) данных о продолжительности солнечного сияния – ТМ-15.

Для оперативных целей первичная информация передается в виде телеграмм в оперативные прогностические центры с установленной регулярностью. Режимная информация за различные периоды времени заносится на технический носитель, проверяется и пересылается в территориальный УГМС и Госфонд. Способы передачи зависят от степени технической оснащённости пункта наблюдений. В нашей стране основным «хранилищем» метеорологических данных является Всероссийский научно-исследовательский институт гидрометеорологической информации — Мировой центр данных (ВНИИГМИ-МЦД).

Глава 3. Анализ временных рядов метеорологических величин

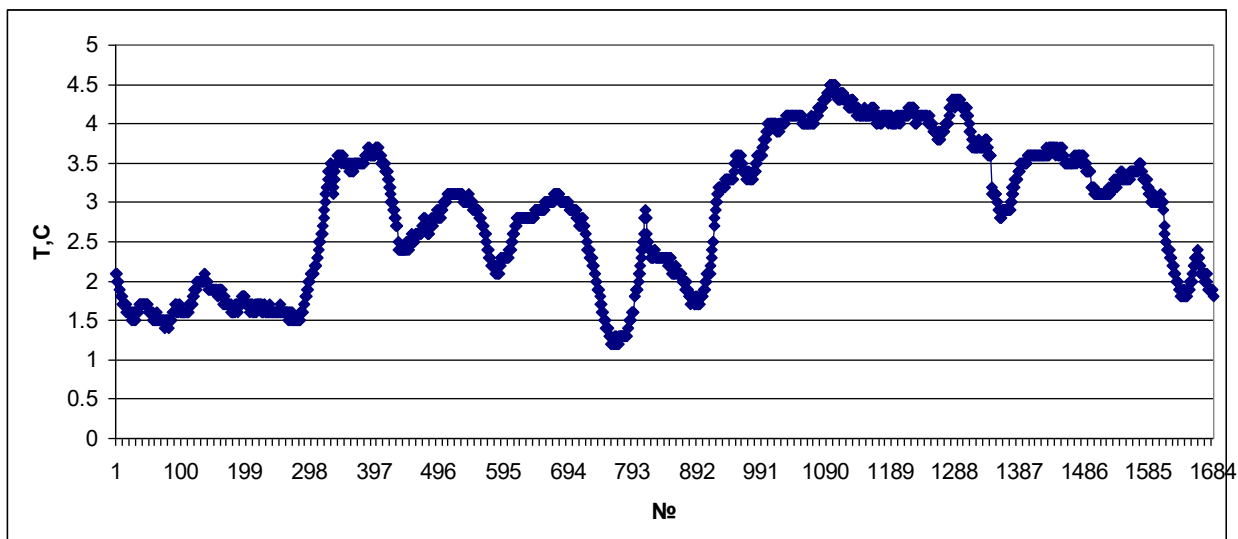
3.1. Характеристика временных рядов

За продолжительное время на станциях накапливается материал метеорологических наблюдений, который служит основой для составления метеорологических рядов. Метеорологический ряд – статистическая совокупность числовых значений метеорологических величин или характеристик атмосферного явления. Статистическая совокупность может быть представлена:

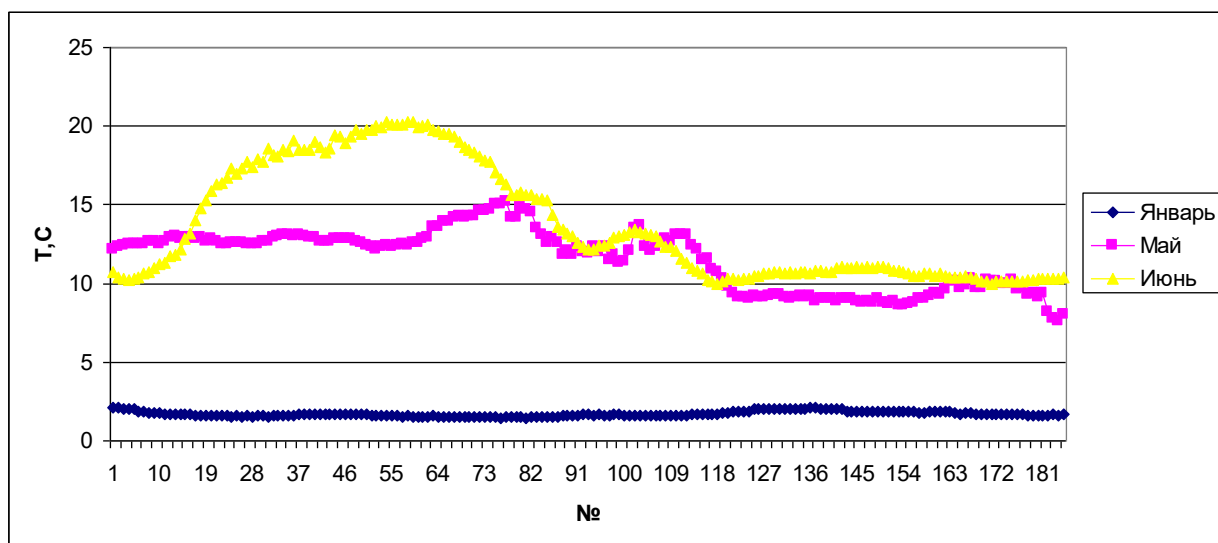
1. В виде простого статистического ряда (простая статистическая совокупность).
2. В виде статистического распределения (группированный ряд).
3. В виде ранжированного ряда.

В первом случае значения x_i величины x , наблюдавшиеся в момент времени t_j , обычно располагаются в хронологической последовательности (хронологический ряд). Примером хронологического ряда служат различного рода таблицы последовательных записей результатов метеорологических наблюдений (например, таблицы месячной отчетности). Метеорологический ряд, представленный в виде хронологической последовательности, имеет важное значение, так как только по хронологическому ряду возможно рассчитать различные характеристики динамики ряда (например, изменения и колебания климата), изучить закономерности изменения случайной величины во времени – от года к году, от одного десятилетия к другому десятилетию и т.д.

3.2. Анализ временных рядов



Временной ряд температуры: январь



Временные ряды температуры: январь, май, июнь

Таблица

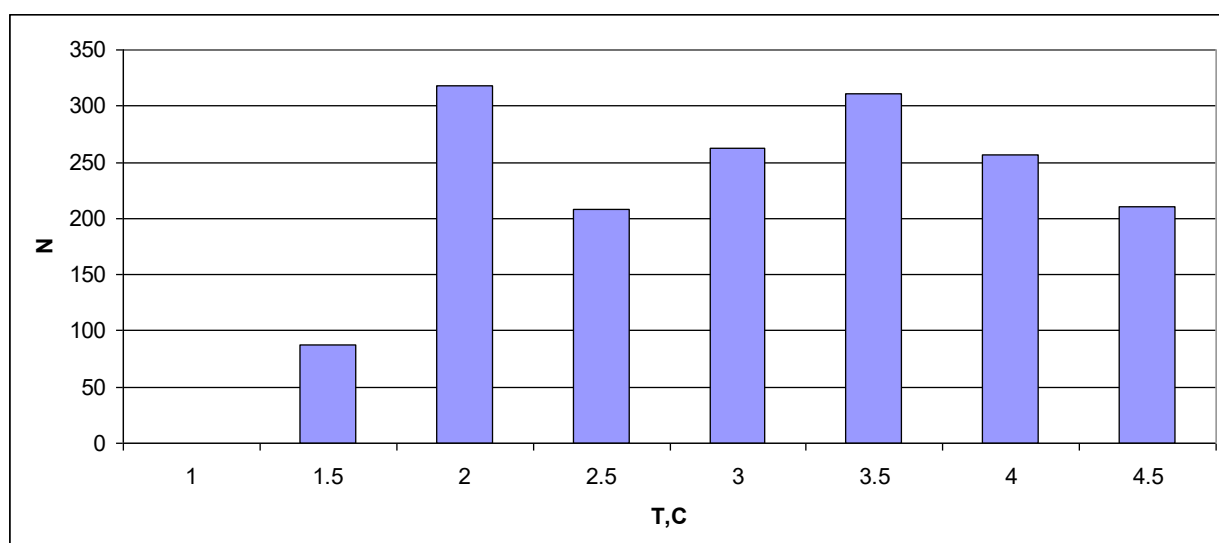
Статистические характеристики рядов температуры: N = 185

<i>Январь</i>		<i>Май</i>		<i>Июнь</i>	
Среднее	1.71	Среднее	11.53	Среднее	13.74
Стандартная ошибка	0.01	Стандартная ошибка	0.14	Стандартная ошибка	0.27
Медиана	1.70	Медиана	12.30	Медиана	12.30
Мода	1.60	Мода	12.70	Мода	10.30
Стандартное отклонение	0.17	Стандартное отклонение	1.91	Стандартное отклонение	3.63
Дисперсия выборки	0.03	Дисперсия выборки	3.63	Дисперсия выборки	13.16
Эксцесс	-0.73	Эксцесс	-1.18	Эксцесс	-1.29
Асимметричность	0.55	Асимметричность	-0.25	Асимметричность	0.59
Интервал	0.70	Интервал	7.60	Интервал	10.30
Минимум	1.40	Минимум	7.60	Минимум	10.00
Максимум	2.10	Максимум	15.20	Максимум	20.30
Сумма	316.00	Сумма	2132.70	Сумма	2541.90
Счет	185.00	Счет	185.00	Счет	185.00

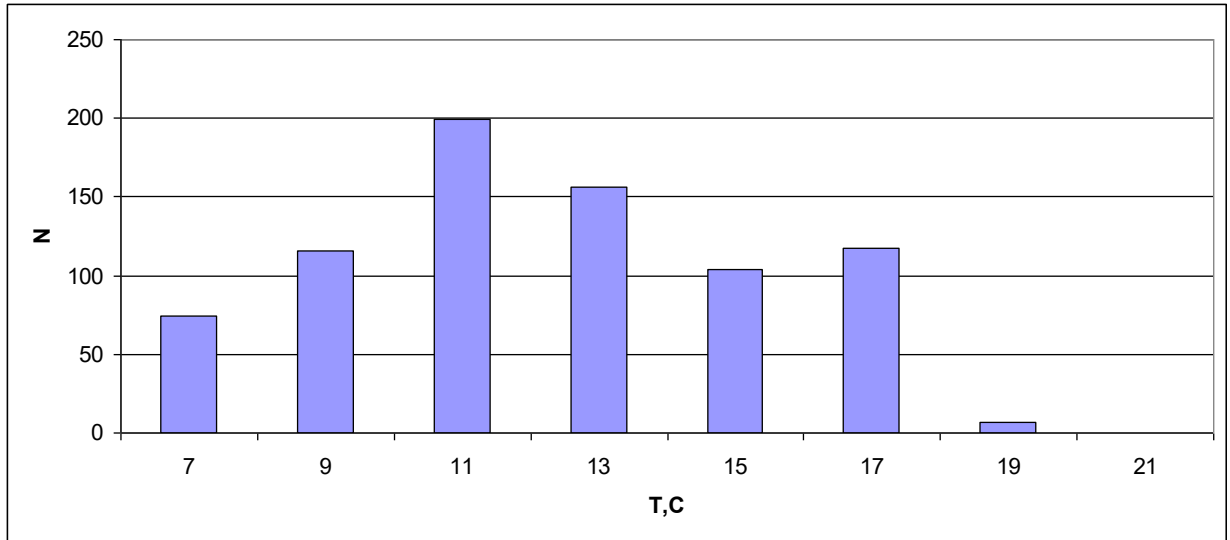
Таблица

Статистические характеристики рядf температуры: N = 1684

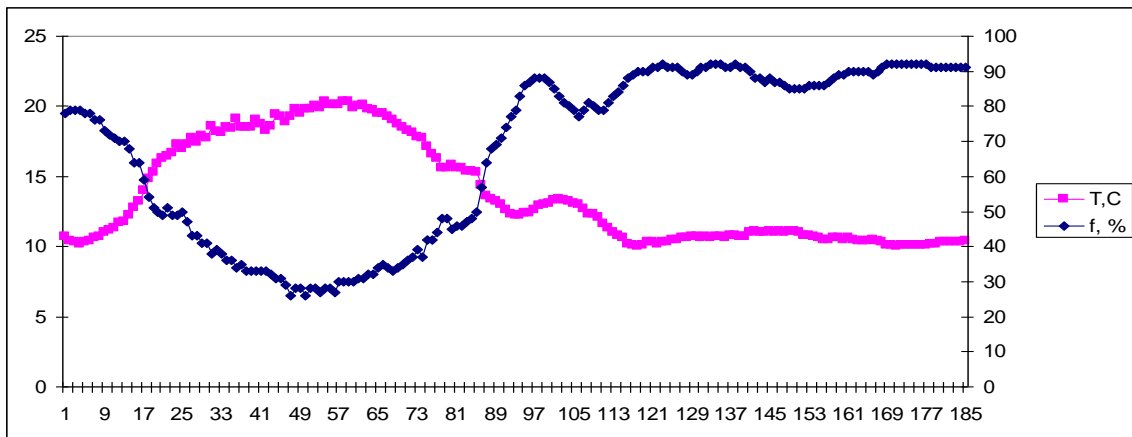
<i>Январь</i>	
Среднее	2.87
Стандартная ошибка	0.02
Медиана	3.00
Мода	4.10
Стандартное отклонение	0.90
Дисперсия выборки	0.81
Эксцесс	-1.25
Асимметричность	-0.08
Интервал	3.30
Минимум	1.20
Максимум	4.50
Сумма	4835.70
Счет	1684.00



Гистограмма распределения температуры: январь



Гистограмма распределения температуры: июнь



Совместный график температуры и влажности, июнь

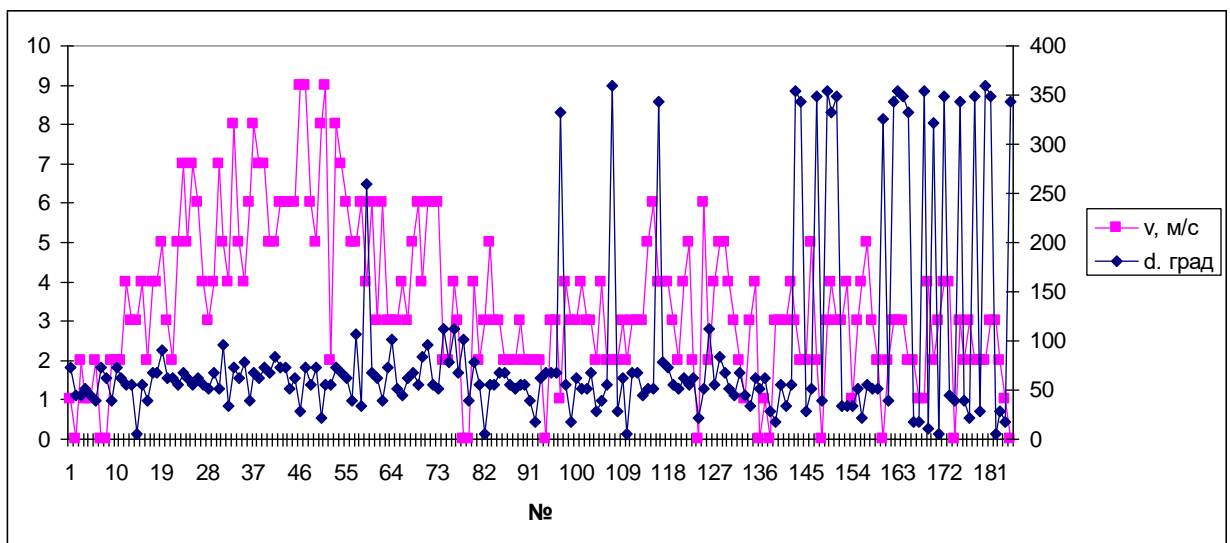


Рис Совместный график скорости и направления ветра, июнь

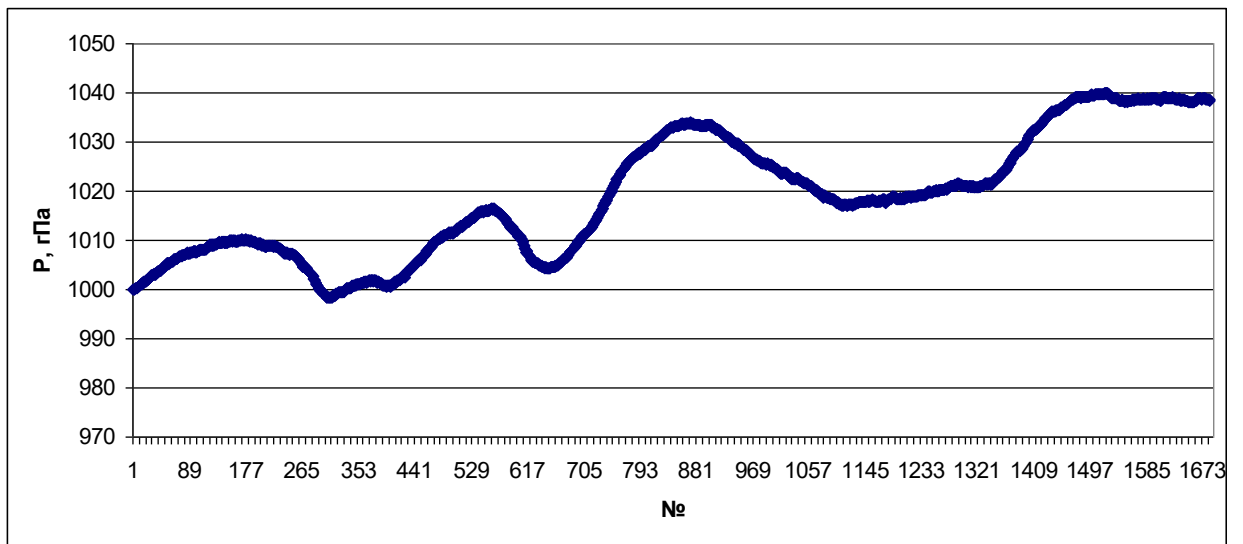


Рис. Временной ряд давления: январь

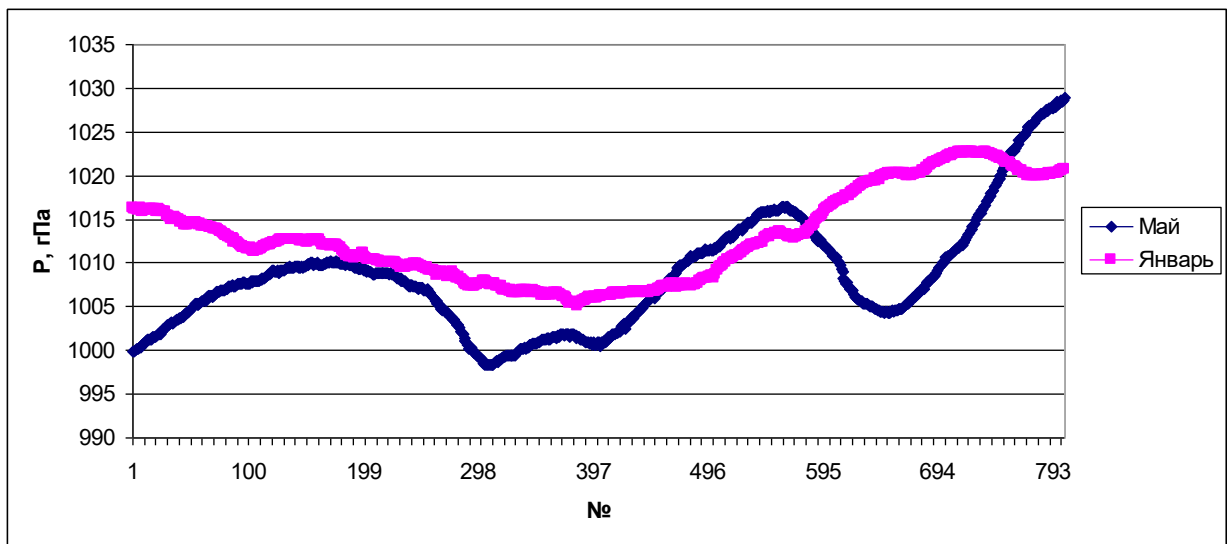
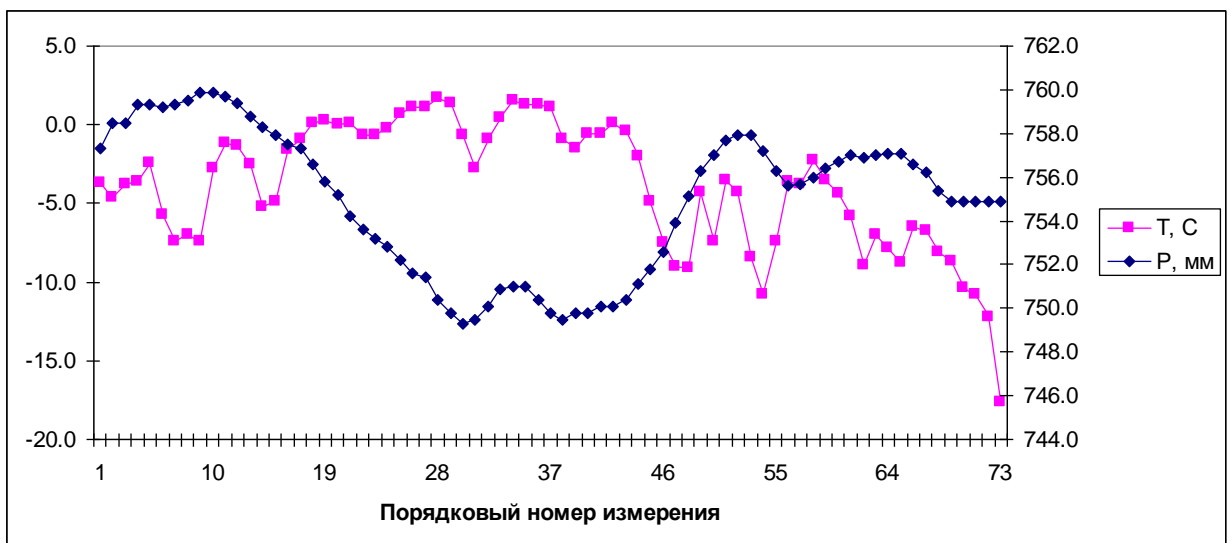
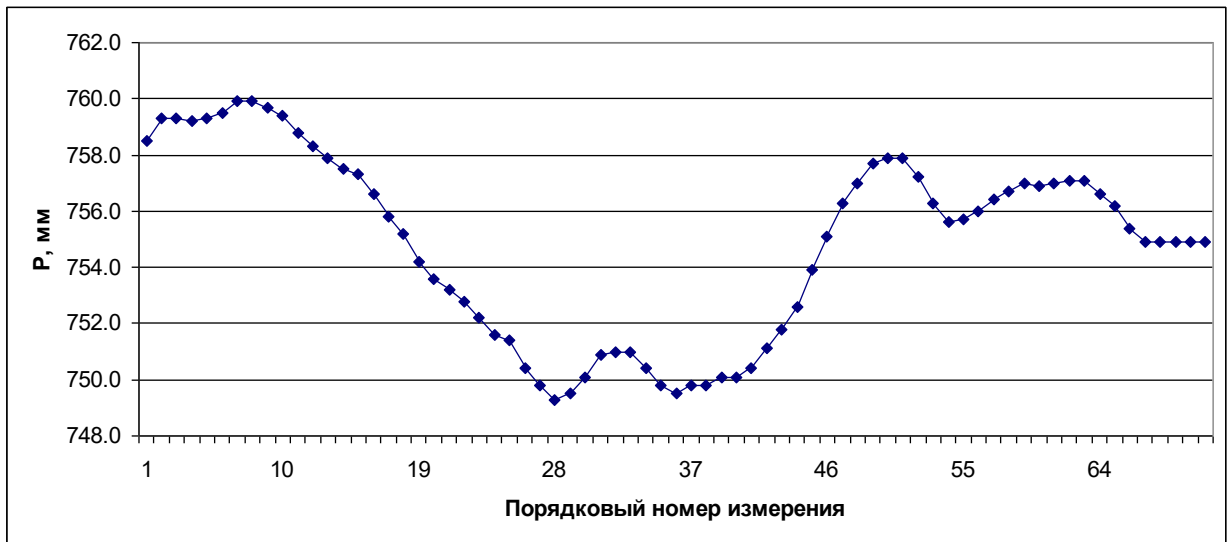
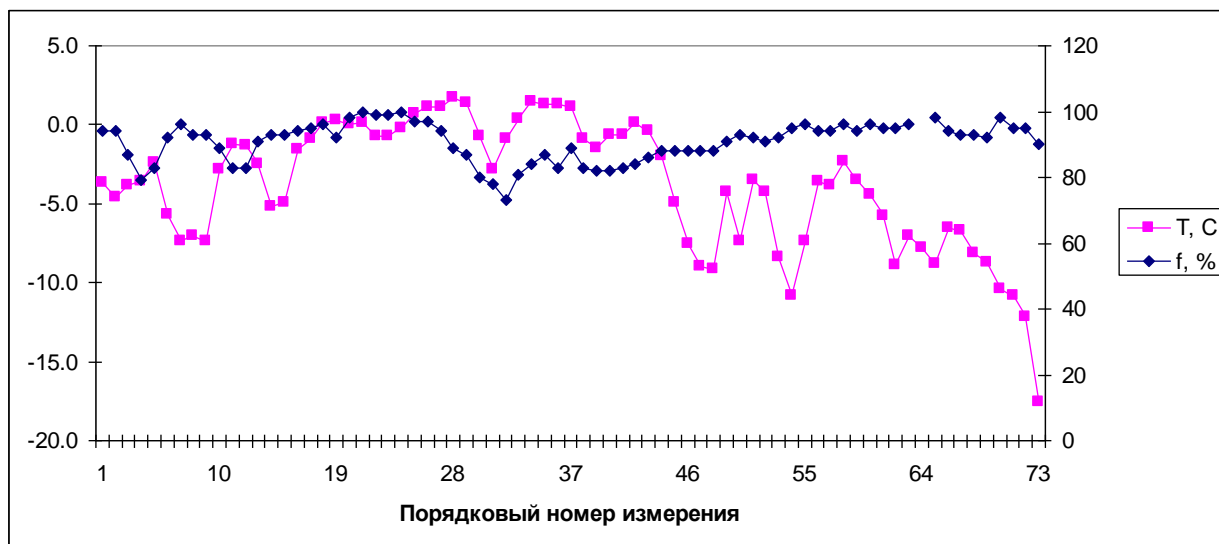


Рис. Совместный график давления: май и июнь

Визуализация временных рядов метеорологических величин







Контроль качества временных рядов

1. Была сформирована таблица с исходными данными.

	A	B	C	D	E	F	G	H
1	576	07.03.2010 0:12	-6,7	1021	3,3	89	2	51
2	577	07.03.2010 0:27	-6,8	1021,1	3,3	90	1	62
3	578	07.03.2010 0:42	-7,1	1021,2	3,2	90	2	157
4	579	07.03.2010 0:57	-7,5	1021,2	3,1	89	5	62
5	580	07.03.2010 1:12	-7,8	1021,4	3,1	90	5	45
6	581	07.03.2010 1:27	-7,7	1021,5	3,1	90	3	62
7	582	07.03.2010 1:42	-7,7	1021,7	3,1	90	3	22
8	583	07.03.2010 1:57	-7,6	1021,9	3,1	90	4	67
9	584	07.03.2010 2:12	-7,4	1022,1	3,2	91	2	73
10	585	07.03.2010 2:27	-7,4	1022	3,2	91	3	62
11	586	07.03.2010 2:42	-7,2	1022,2	3,2	91	2	79
12	587	07.03.2010 2:57	-7,2	1022,3	3,2	90	5	73
13	588	07.03.2010 3:12	-7,2	1022,6	3,2	90	3	62
14	589	07.03.2010 3:27	-7,3	1022,8	3,2	90	2	56
15	590	07.03.2010 3:42	-7,5	1022,9	3,1	89	4	45
16	591	07.03.2010 3:57	-7,7	1023	3,1	89	4	73

Рис. 1. Пример представления временных рядов, сформированных для последующего контроля на выбросы

2. Проверяем наши данные на наличие выбросов.

	A	B	C	D	E	F	G	H	I	J	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	-6,7									Номер колонки с рядом		1													
2	-6,8									Общее число значений		925													
3	-7,1									Начальное значение n1		1													
4	-7,5									Конечное значение n2		925													
5	-7,8									Длина контрольного от		50													
6	-7,7									Значение коэффициента		10													
7	-7,7																								
8	-7,6																								
9	-7,4									РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ															
10	-7,4									К а ф е д р а Э К С П Е Р И М Е Н Т А Л Ь Н О Й Ф И З И К И А Т М О С Ф Е Р Ы															
11	-7,2									Обнаружения пропусков и идентификация ВЫБРОСОВ разными методами															
12	-7,2									Профессор А.Д.Кузнецов															
13	-7,2																								
14	-7,3									Перед запуском программы введите:															
15	-7,5									№ колонки с исследуемым рядом															
16	-7,7									По отрезкам длины ndl рассчитываются среднее и СКО.															
17	-7,9									Наличие выброса проверяется методами															
18	-7,9									Метод 1: Контроль по разности двух значений: ndl и ndl-1 относительно СКО участка длиной ndl															
19	-8									Метод 2: Контроль по разности двух значений: значения ndl-1 и среднего на участке ndl и относительно СКО участка длиной ndl															
20	-7,9									Метод 3: в пределах контрольного участка разность значения ряда и среднего сравнивается с СКО (статистика-для контрольного участка).															
21	-8,2									Метод 4: метод Ирвина															
22	-8,3																								
23	-8,7									На листе 2 в колонку 1 записывается исходный ряд															
24	-8,6									На листе 2 в колонку 2 записывается индикатор выброса (метод 1)															
25	-8,7									На листе 2 в колонку 3 записывается индикатор выброса (метод 2)															
26	-8,3									На листе 2 в колонку 4 записывается индикатор выброса (метод 3)															
27	-8,1									На листе 2 в колонку 5 записывается индикатор выброса (метод Ирвина)															
28	-8,6																								
29	-8,5																								

Рис. 2. Проверка работы программы Макрос.

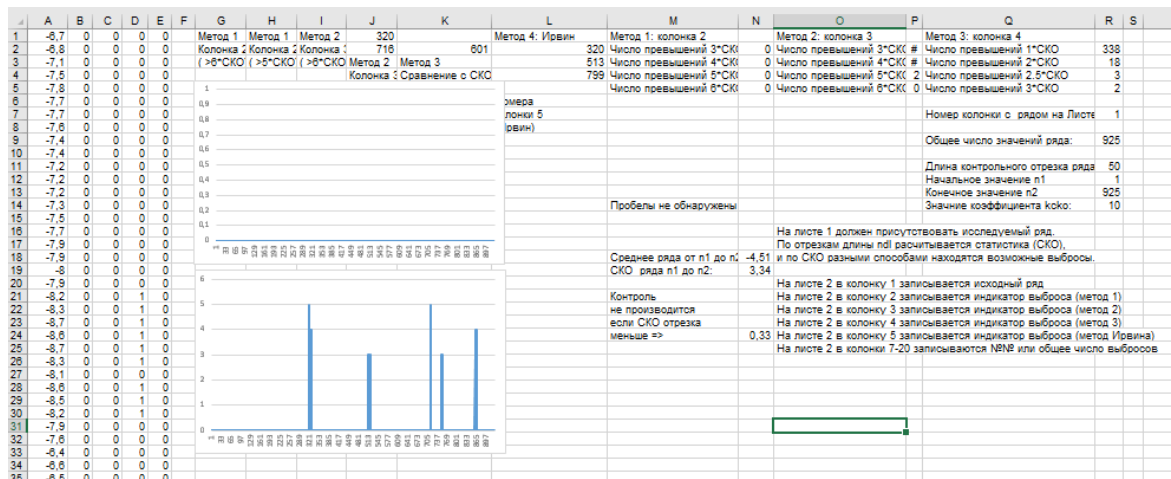


Рис. 3. Пример фрагмента Листа 2 после запуска файла «Контроль 4М по отрезкам (Ирвин 1).xls»

Анализируя наши данные, мы видим, что выбросов не обнаружено. Для чистоты эксперимента вводим искусственные данные в ячейку 51. Вместо температуры -5,1 вставляем температуру -15,1.

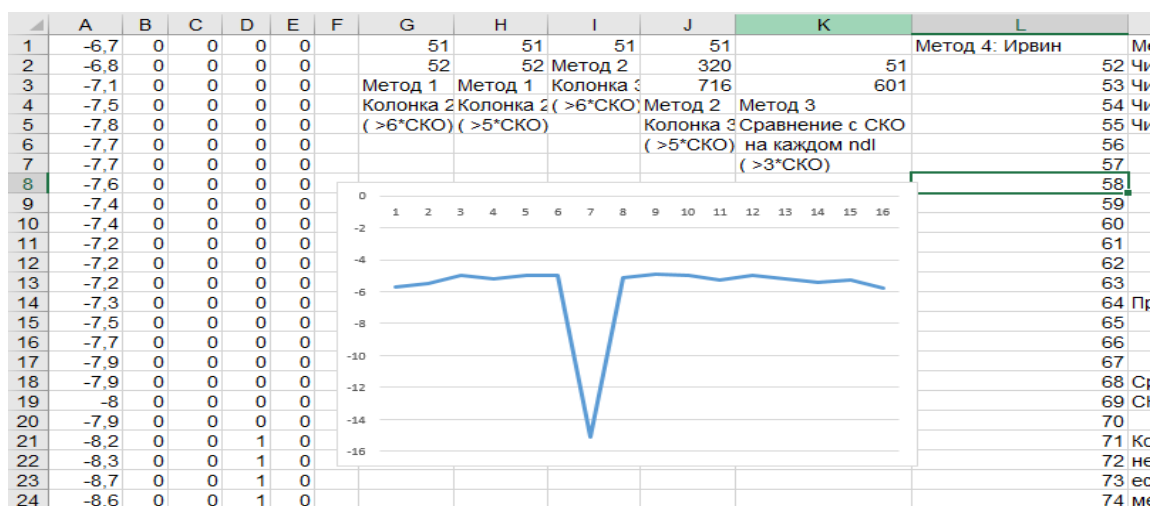


Рис. 3. Ввод искусственного выброса виден на графике.

Программа выдала ошибки в ячейках 51 и 52, т.к. там наблюдается резкое повышение температуры.

3. Вводим в наши данные искусственные пробелы.

Для этого удаляем данные в ячейке 51 и проверяем программу.

Мода как средняя величина употребляется чаще для данных, имеющих нечисловую природу.

б) **Дисперсия D** и стандартное отклонение σ характеризуют среднее рассеяние значений ряда от среднего арифметического значения.

Выборочной дисперсией называют среднее арифметическое отклонения квадратов значений временного ряда от их среднего значения для данной выборки.

Выборочная дисперсия является смещенной оценкой генеральной дисперсии. Для исправления выборочной дисперсии достаточно умножить ее на дробь.

7) **Коэффициент вариации** позволяет оценить, велика или нет изменчивость ряда:

- если $C > 33\%$ – изменчивость значительна,

- если $C < 33\%$ – изменчивость мала и выборка может считаться однородной.

Коэффициент вариации также используется для сравнения изменчивости двух выборок с разными единицами измерения.

8) **Асимметрия** характеризует симметричность эмпирической функции распределения (ЭФР) относительно среднего значения.

Значение $A_s = 0$ при полной симметрии ЭФР относительно среднего значения.

Если $A_s > 0$, то ЭФР обладает положительным «хвостом» и основная масса наблюдений (а также медиана) меньше среднего значения (это означает, что преобладают данные с меньшими значениями по сравнению со средним арифметическим).

Если $A_s < 0$, то ЭФР обладает отрицательным «хвостом» и основная масса наблюдений (а также медиана) больше среднего значения (это означает, что преобладают данные с большими значениями).

9) **Экцесс** характеризует относительную остроконечность или сглаженность распределения по сравнению с нормальным распределением.

Если $E_x > 0$, то ЭФР является островершинной и, как правило, у нее наблюдается два равнозначных «хвоста».

Если $E_x < 0$, то ЭФР является плосковершинной и распределение стремится к случайному распределению.

10) **Доверительный интервал** – это интервал значений статистической характеристики временного ряда, соответствующий доверительной области статистического критерия.

Таблица 1 - Результаты расчета статистических характеристик блока данных

<i>Все данные</i>	
Среднее	-4,33
Стандартная ошибка	0,10
Медиана	-4,30
Мода	-3,60
Стандартное отклонение	3,13
Дисперсия выборки	9,80
Эксцесс	-0,21
Асимметричность	0,15
Интервал	15,20
Минимум	-11,90
Максимум	3,30
Сумма	-4153,20
Счет	959

Анализируя таблицу 1 можно сделать вывод, что эмпирическая функция распределения (ЭФР) обладает положительным «хвостом» и основная масса наблюдений (а также медиана) меньше среднего значения (это означает, что преобладают данные с меньшими значениями по сравнению со средним арифметическим), так как асимметрия больше 0 (0,15). Так же ЭФР является плосковершинной и распределение стремится к случайному распределению, так как эксцесс меньше 0 (-0,21).

Таблица 2 - Результаты расчета статистических характеристик 1 половины блока данных

<i>Первая половина</i>	
Среднее	-6,26
Стандартная ошибка	0,09
Медиана	-6,20
Мода	-6,70
Стандартное отклонение	2,04
Дисперсия выборки	4,17
Эксцесс	0,12
Асимметричность	-0,52
Интервал	10,20
Минимум	-11,90
Максимум	-1,70
Сумма	-3004,10
Счет	480

Анализируя таблицу 2 можно сделать вывод, ЭФР обладает отрицательным «хвостом» и основная масса наблюдений (а также медиана) больше среднего значения (это означает, что преобладают данные с большими значениями), так как асимметрия меньше 0 (-0,52). Так же ЭФР является островершинной и, как правило, у нее наблюдается два равнозначных «хвоста», так как эксцесс больше 0 (0,12).

Таблица 3 - Результаты расчета статистических характеристик 2 половины блока данных

<i>Вторая половина</i>	
Среднее	-2,40
Стандартная ошибка	0,13
Медиана	-2,70
Мода	-3,60
Стандартное отклонение	2,83
Дисперсия выборки	7,98
Эксцесс	0,36
Асимметричность	-0,37
Интервал	14,00
Минимум	-10,70
Максимум	3,30
Сумма	-1149,10
Счет	479

Анализируя таблицу 3 можно сделать вывод, ЭФР обладает отрицательным «хвостом» и основная масса наблюдений (а также медиана) больше среднего значения (это означает, что преобладают данные с большими значениями), так как асимметрия меньше 0 (-0,37). Так же ЭФР является островершинной и, как правило, у нее наблюдается два равнозначных «хвоста», так как эксцесс больше 0 (0,36).

Таблица 4 - Результаты расчета характеристик 1, 2 половины и всех данных

t кр	1,962440705	t кр	1,964929	t кр	1,964918
dx	0,009644184	dx	-0,03338	dx	-0,04655
Пирсон 2	888,1185118	Пирсон 2	429,2513	Пирсон 2	430,1982
Пирсон 1		Пирсон 1	531,0222	Пирсон 1	532,0754
d2	0,926088125	d2	0,896141	d2	0,896246
d1		d1	1,108606	d1	1,10849

Таблица 5 - Результаты расчета характеристик 1, 2 половины и всех данных

	Хср-dx	Хср	Хср+dx
Весь ряд	-4,34	-4,33	-4,32
Первая половина	-6,21	-6,26	-6,31
Вторая половина	-2,37	-2,40	-2,43
Ряд нестационарен			
	Дисперсия		
Весь ряд		887,1561	821,5847
Первая половина	475,821	429,2513	384,7149
Вторая половина	474,8209	428,3045	383,821
	Ст.откл		
Весь ряд		1,9624	28,6633
Первая половина	21,81332	1,9649	19,61415
Вторая половина	21,79039	1,9649	19,59135

Проведя расчет доверительного интервала для среднего арифметического значения для всего ряда данных при уровне значимости 95% ($\alpha = 0.95$) и длине временного ряда 959 я получила 1,9624. Для первой и второй половины ряда получилось 1,9649.

Проведя расчет доверительного интервала для дисперсии для всего ряда данных при уровне значимости 95% ($\alpha = 0.95$) и длине временного ряда 959 я получила 887,1561. Для первой половины ряда получилось 429,2513, а для второй 428,3045.

Анализируя табл.5 можно сделать вывод, что ряд не стационарен.

Временной тренд

При исследовании временных рядов может оказаться важным нахождения момента времени, соответствующего резкому изменению характера протекающего процесса, т. е. точки бифуркации. При этом наступление

такого момента может быть достаточно легко установлено при графическом анализе временного ряда, так и носить «скрытый» характер.

Рассмотрим одну из возможных методик реализации нахождения положения точки бифуркации – метод «ступеньки».

Алгоритм №1.

программа «Ступенька1.xls»

Пусть имеется временной ряд их N значений. Разделим этот ряд на два отрезка: $[1, n_1]$ и $[n_1+1, N]$.

Вычислим средние значения для каждого отрезка:

$$Y_{1cp} = \sum_1^{n_1} Y_i \qquad Y_{2cp} = \sum_{n_1+1}^n Y_i$$

Тогда положение точки бифуркации будет определяться таким номером n_1 , для которого будет выполняться следующее условие:

$$SS = \sum_1^{n_1} (Y_i - Y_{1cp})^2 + \sum_{n_1+1}^n (Y_i - Y_{2cp})^2 = \min$$

При реализации данного алгоритма точка бифуркации определяется положением двух прямых, при котором сумма квадратов отклонений этих прямых от значений временного ряда будет минимальным.

Для оценки «эффективности» произведенного деления ряда на два отрезка можно воспользоваться сравнением дисперсии всего ряда $\sigma_{\text{ряда}}$ и дисперсии $\sigma_{\text{ступ}}$ его двух отрезков, вычисленных по формуле:

$$\sigma_{\text{ступ}} = \sqrt{\frac{\sigma_1^2 n_1 + \sigma_2^2 n_2}{(n_1 + n_2 - 1)}}$$

Для приведенного на предыдущем слайде примера:

$$\sigma_{\text{ряда}} = 1.703; \quad \sigma_{\text{ступ}} = 0.891$$

Алгоритм №2.

программа «Ступенька2.xls»

Пусть, по-прежнему, имеется временной ряд их N значений. Разделим этот ряд на два отрезка: $[1, n_1]$ и $[n_1+1, N]$.

Считая, что $x_i = i$, вычислим для каждого отрезка коэффициенты линейного временного тренда по формулам:

$$a_1 = \frac{n(\sum y_i x_i) - \sum y_i \sum x_i}{n(\sum x_i^2) - (\sum x_i)^2}; \quad a_0 = \frac{1}{n}(\sum y_i - a_1 \sum x_i)$$

При расчете коэффициентов линейного временного тренда для первого отрезка суммирование ведется от 1 до n_1 , а для второго отрезка – от n_1+1 до N . Соответственно: для первого отрезка в приведенных соотношениях $n = n_1$, а для второго: $n = N - n_1$.

Обозначим рассчитанные по двум отрезкам временного ряда коэффициенты линейного временного тренда как a_{01}, a_{11} и a_{02}, a_{12} .

Тогда положение точки бифуркации будет определяться таким номером n_1 , для которого будет выполняться следующее условие:

$$SS = \sum_{i=1}^{n_1} (a_{01} + a_{11} i - Y)^2 + \sum_{i=n_1+1}^N (a_{02} + a_{12} i - Y)^2 = \min$$

Оценка статистической значимости найденного положения точки бифуркации

Для оценки значимости проведенного разделения сравним значение критерия Фишера:

$$F = \frac{\sigma_{\text{ряда}}^2}{\sigma_{\text{ступ}}^2}$$

с его критическим значением $F_{\text{кр}}$.

Если выполняется неравенство $F > F_{\text{кр}}$, то будем считать, что проведенное разделение статистически значимо.

Таблица 1 – Ряд данных

-5,2	1,5	0,1
-5,3	1,5	0,1
-5,4	1,5	0
-5,4	1,5	0,1
-5,5	1,5	0,1
-5,6	1,5	0
-5,6	1,5	-0,1
-5,5	1,5	0
-5,5	1,5	0,1
-5,6	1,5	0
-5,6	1,5	0
-5,6	1,5	0
-5,6	1,5	0,1
-5,7	1,5	0
-5,7	1,5	0,1
-5,8	1,5	0,2
-6	1,5	0,2
-6,2	1,5	0,2
-6,4	1,5	0
-6,4	1,5	0
-6,4	1,5	0,2
-6,6	1,5	0
-6,6	1,5	0
-6,6	1,5	-0,1
-6,5	1,5	-0,4

График 1 – Распределение температуры

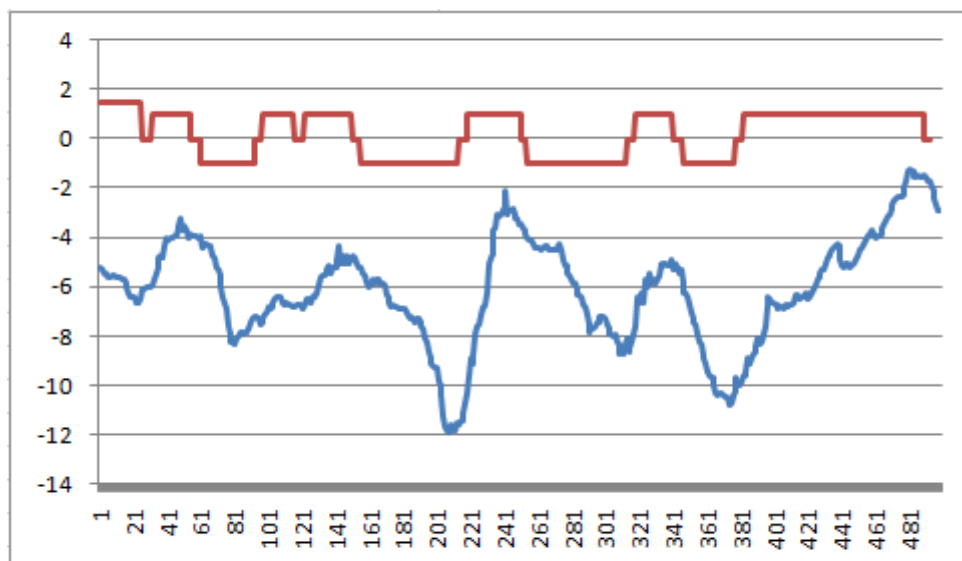


График 2 – Распределение температуры с линией тренда

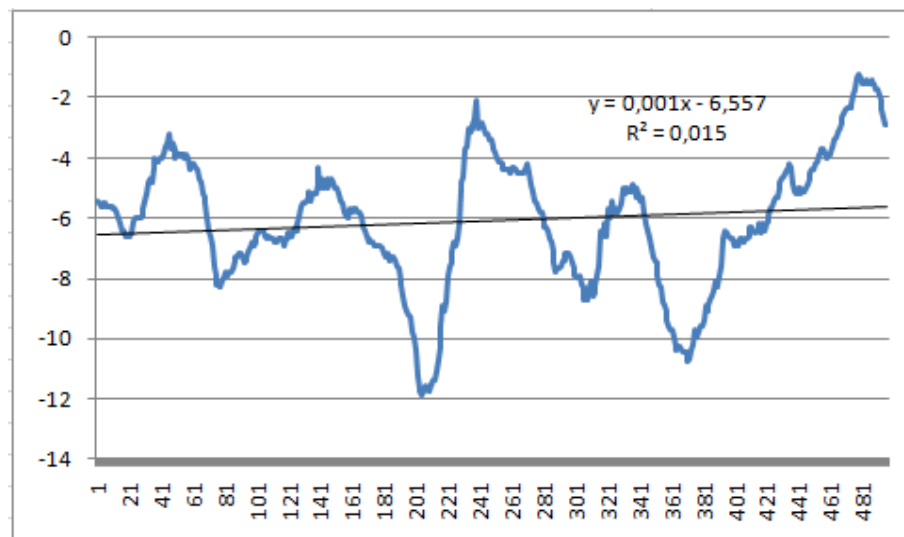


Таблица 2 – расчет данных для всего ряда

N	500
t	0,335112321
tkp	1,964729307

График 3 – Распределение восходящей температуры с линией тренда

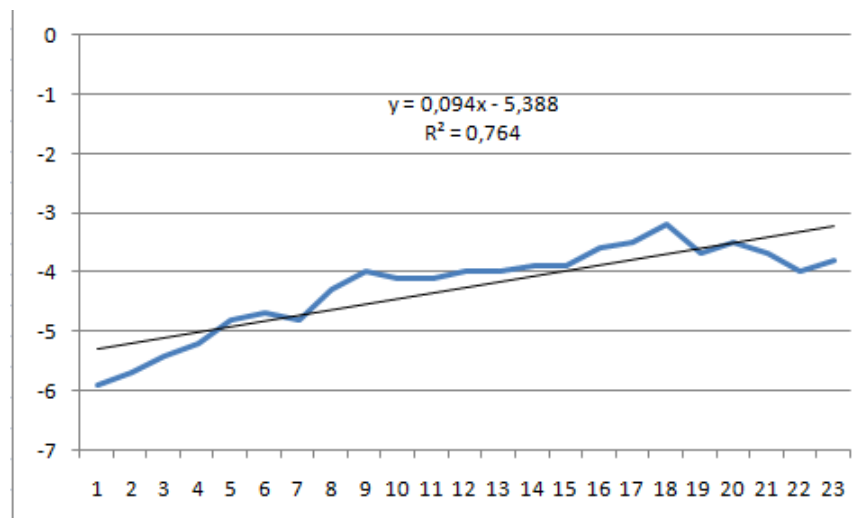


Таблица 3 – расчет данных для восходящих данных

N	22
t	5,426223952
tkp	2,079613837

График 4 – Распределение нисходящей температуры с линией тренда

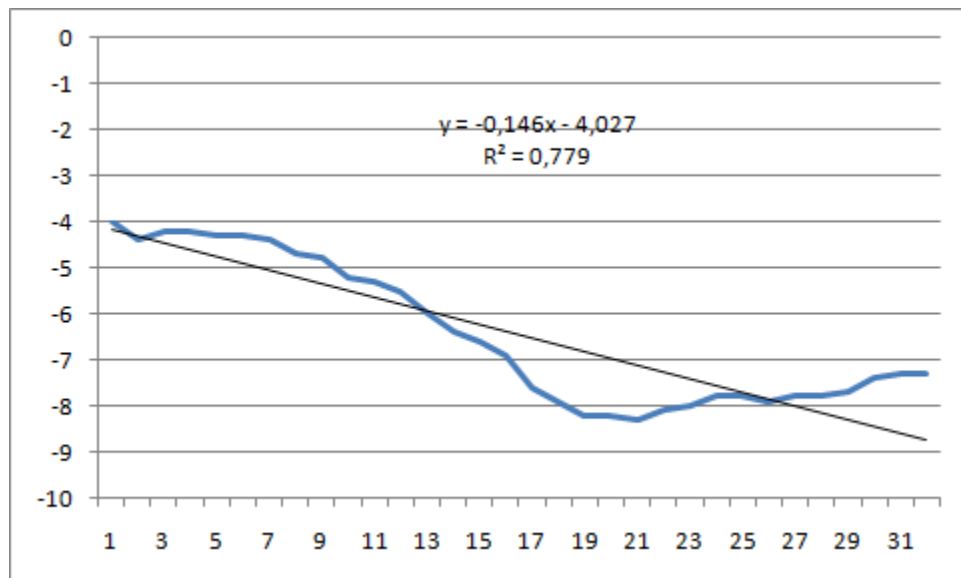


Таблица 4 – расчет данных для нисходящих данных

N	31
t	6,804778201
t _{кр}	2,042272449

- На основании таблицы 2 можно сделать вывод, что $t < t_{кр}$. Это значит, что весь ряд не является значимым.
- На основании таблицы 3 можно сделать вывод, что $t > t_{кр}$. Это значит, что восходящий ряд является значимым.
- На основании таблицы 4 можно сделать вывод, что $t > t_{кр}$. Это значит, что нисходящий ряд является значимым.

Временной тренд

При исследовании временных рядов может оказаться важным нахождение момента времени, соответствующего резкому изменению характера протекающего процесса, т. е. точки бифуркации. При этом наступление такого момента может быть достаточно легко установлено при графическом анализе временного ряда, так и носить «скрытый» характер.

Рассмотрим одну из возможных методик реализации нахождения положения точки бифуркации – метод «ступеньки».

Алгоритм №1.

программа «Ступенька1.xls»

Пусть имеется временной ряд их N значений. Разделим этот ряд на два отрезка: $[1, n_1]$ и $[n_1+1, N]$.

Вычислим средние значения для каждого отрезка:

$$Y_{1cp} = \sum_1^{n_1} Y_i \qquad Y_{2cp} = \sum_{n_1+1}^n Y_i$$

Тогда положение точки бифуркации будет определяться таким номером n_1 , для которого будет выполняться следующее условие:

$$SS = \sum_1^{n_1} (Y_i - Y_{1cp})^2 + \sum_{n_1+1}^n (Y_i - Y_{2cp})^2 = \min$$

При реализации данного алгоритма точка бифуркации определяется положением двух прямых, при котором сумма квадратов отклонений этих прямых от значений временного ряда будет минимальным.

Для оценки «эффективности» произведенного разделения ряда на два отрезка можно воспользоваться сравнением дисперсии всего ряда $\sigma_{\text{ряда}}$ и дисперсии $\sigma_{\text{ступ}}$ его двух отрезков, вычисленных по формуле:

$$\sigma_{\text{ступ}} = \sqrt{\frac{\sigma_1^2 n_1 + \sigma_2^2 n_2}{(n_1 + n_2 - 1)}}$$

Для приведенного на предыдущем слайде примера:

$$\sigma_{\text{ряда}} = 1.703; \quad \sigma_{\text{ступ}} = 0.891$$

Алгоритм №2.

программа «Ступенька2.xls»

Пусть, по-прежнему, имеется временной ряд их N значений. Разделим этот ряд на два отрезка: $[1, n_1]$ и $[n_1+1, N]$.

Считая, что $x_i = i$, вычислим для каждого отрезка коэффициенты линейного временного тренда по формулам:

$$a_1 = \frac{n(\sum y_i x_i) - \sum y_i \sum x_i}{n(\sum x_i^2) - (\sum x_i)^2}; \quad a_0 = \frac{1}{n}(\sum y_i - a_1 \sum x_i)$$

При расчете коэффициентов линейного временного тренда для первого отрезка суммирование ведется от 1 до n_1 , а для второго отрезка – от n_1+1 до N . Соответственно: для первого отрезка в приведенных соотношениях $n = n_1$, а для второго: $n = N - n_1$.

Обозначим рассчитанные по двум отрезкам временного ряда коэффициенты линейного временного тренда как a_{01} , a_{11} и a_{02} , a_{12} .

Тогда положение точки бифуркации будет определяться таким номером n_1 , для которого будет выполняться следующее условие:

$$SS = \sum_{i=1}^{n_1} (a_{01} + a_{11} i - Y)^2 + \sum_{i=n_1+1}^N (a_{02} + a_{12} i - Y)^2 = \min$$

Оценка статистической значимости найденного положения точки бифуркации

Для оценки значимости проведенного разделения сравним значение критерия Фишера:

$$F = \frac{\sigma_{\text{ряда}}^2}{\sigma_{\text{ступ}}^2}$$

с его критическим значением $F_{\text{кр}}$.

Если выполняется неравенство $F > F_{\text{кр}}$, то будем считать, что проведенное разделение статистически значимо.

Таблица 1 – Ряд данных

-5,2	1,5	0,1
-5,3	1,5	0,1
-5,4	1,5	0
-5,4	1,5	0,1
-5,5	1,5	0,1
-5,6	1,5	0
-5,6	1,5	-0,1
-5,5	1,5	0
-5,5	1,5	0,1
-5,6	1,5	0
-5,6	1,5	0
-5,6	1,5	0
-5,6	1,5	0,1
-5,7	1,5	0
-5,7	1,5	0,1
-5,8	1,5	0,2
-6	1,5	0,2
-6,2	1,5	0,2
-6,4	1,5	0
-6,4	1,5	0
-6,4	1,5	0,2
-6,6	1,5	0
-6,6	1,5	0
-6,6	1,5	-0,1
-6,5	1,5	-0,4

График 1 – Распределение температуры

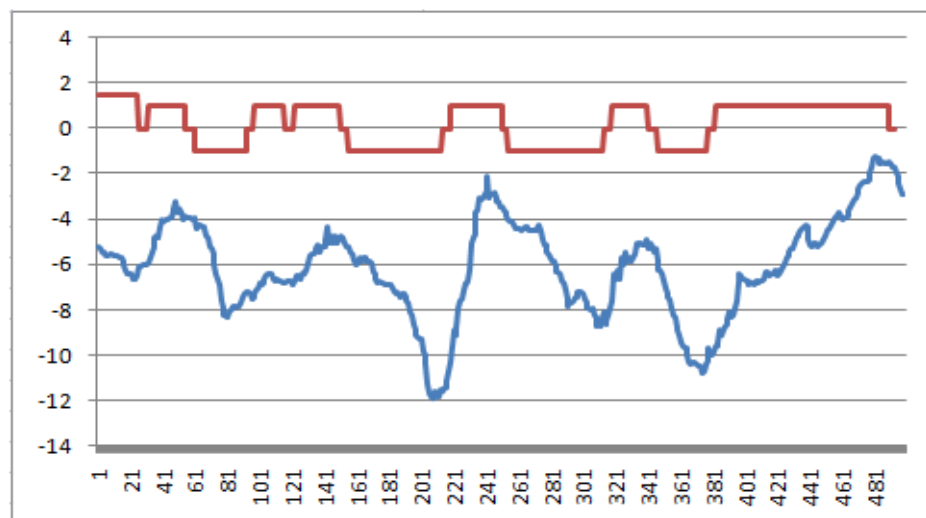


График 2 – Распределение температуры с линией тренда

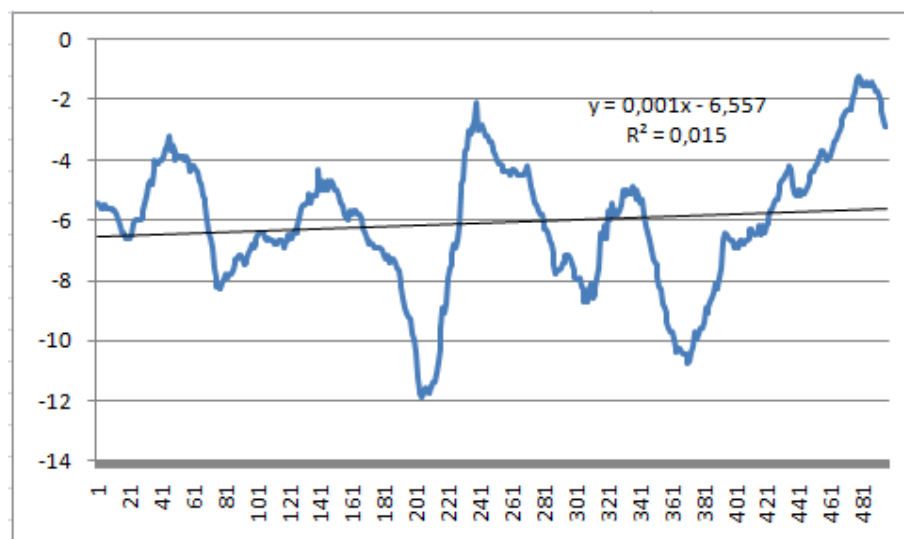


Таблица 2 – расчет данных для всего ряда

N	500
t	0,335112321
tkp	1,964729307

График 3 – Распределение восходящей температуры с линией тренда

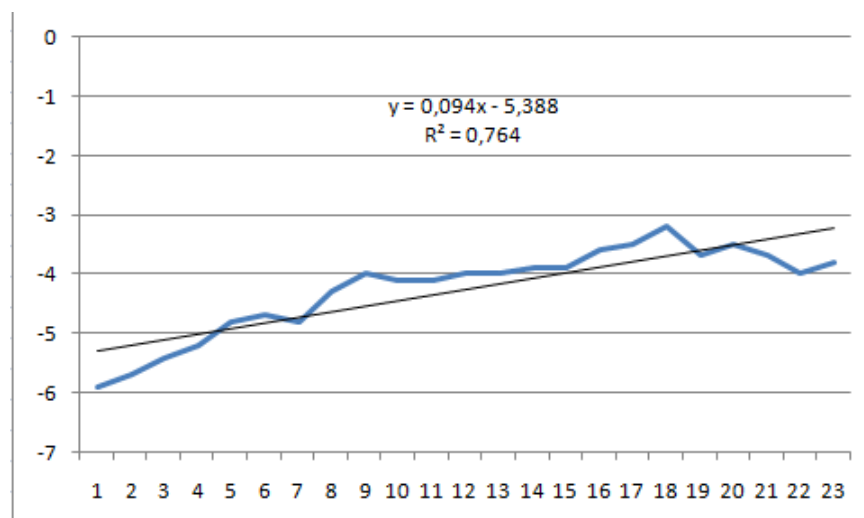


Таблица 3 – расчет данных для восходящих данных

N	22
t	5,426223952
tkp	2,079613837

График 4 – Распределение нисходящей температуры с линией тренда

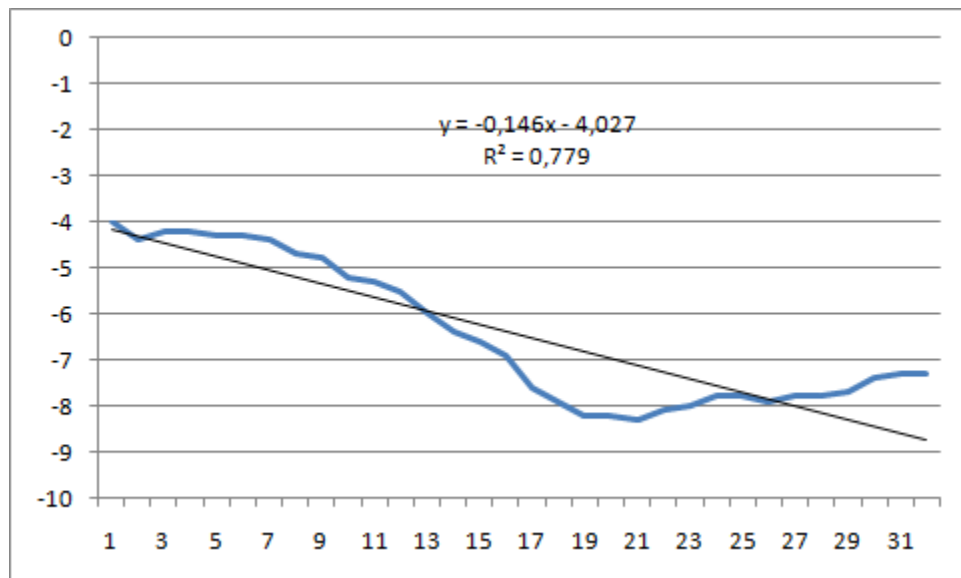


Таблица 4 – расчет данных для нисходящих данных

N	31
t	6,804778201
t _{кр}	2,042272449

- На основании таблицы 2 можно сделать вывод, что $t < t_{кр}$. Это значит, что весь ряд не является значимым.
- На основании таблицы 3 можно сделать вывод, что $t > t_{кр}$. Это значит, что восходящий ряд является значимым.
- На основании таблицы 4 можно сделать вывод, что $t > t_{кр}$. Это значит, что нисходящий ряд является значимым.

Эмпирическая функция распределения – гистограмма

Этапы построения гистограммы:

1. Рассчитываем минимальное и максимальное значение ряда (целого ряда, а также его первая и вторая часть)

Весь ряд		Первая половина ряда		Вторая половина ряда	
Среднее	-4,50422	Среднее	-3,42873	Среднее	-5,58203
Стандартная ошибка	0,110031	Стандартная ошибка	0,159328	Стандартная ошибка	0,134369
Медиана	-4,3	Медиана	-3,5	Медиана	-5,8
Мода	-3,6	Мода	-3,6	Мода	-1,1
Стандартное отклонение	3,346461	Стандартное отклонение	3,428322	Стандартное отклонение	2,888147
Дисперсия выборки	11,1988	Дисперсия выборки	11,75339	Дисперсия выборки	8,34139
Эксцесс	-0,59172	Эксцесс	-0,67794	Эксцесс	-0,84707
Асимметричность	0,151222	Асимметричность	-0,01898	Асимметричность	0,044629
Интервал	14,6	Интервал	14,1	Интервал	11,8
Минимум	-11,3	Минимум	-10,8	Минимум	-11,3
Максимум	3,3	Максимум	3,3	Максимум	0,5
Сумма	-4166,4	Сумма	-1587,5	Сумма	-2578,9
Счет	925	Счет	463	Счет	462

Рис. 1. Расчитанные минимумы и максимумы рядов

2. Производим разбивку интервалов «карманов»

карманы для всего ряда		Карман	Частота
-12		-12	0
-10		-10	50
-8		-8	107
-6		-6	180
-4		-4	155
-2		-2	242
0		0	85
2		2	77
4		4	29

Рис. 2. Интервал «карманов» для всего ряда

карманы для 1 части ряда		Карман	Частота
-11		-11	0
-9		-9	26
-7		-7	49
-5		-5	77
-3		-3	124
-1		-1	77
1		1	56
3		3	47
5		5	7

Рис. 3. Интервал «карманов» для первой части ряда

караны для 2 части ряда		Карман	Частота
-12		-12	0
-10		-10	36
-8		-8	72
-6		-6	113
-4		-4	90
-2		-2	101
0		0	44
2		2	6

Рис. 4. Интервал «карманов» для второй части ряда

3. Строим гистограммы по полученным данным

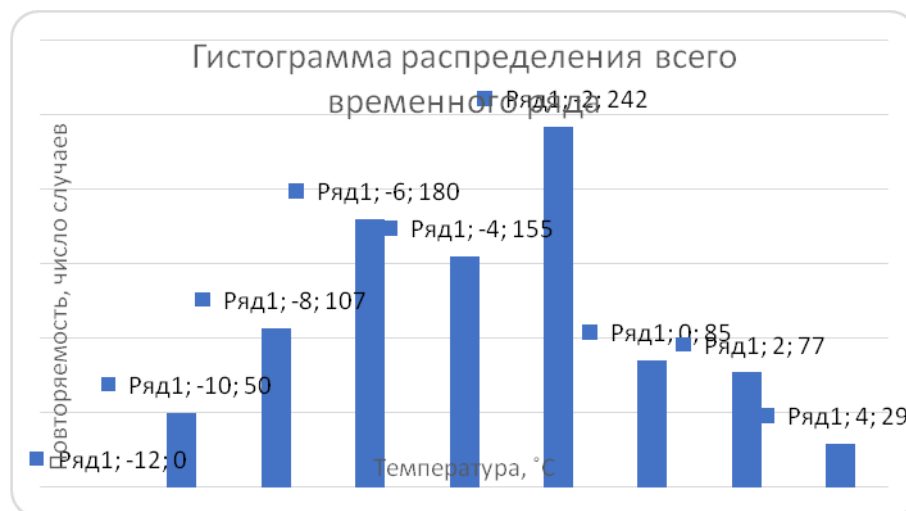


Рис. 5. Гистограмма всего временного ряда

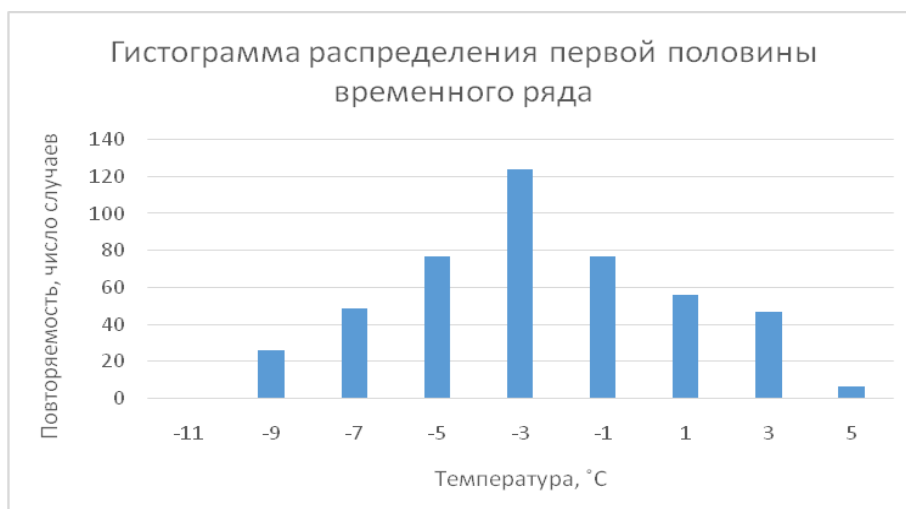


Рис. 6. Гистограмма первой половины ряда

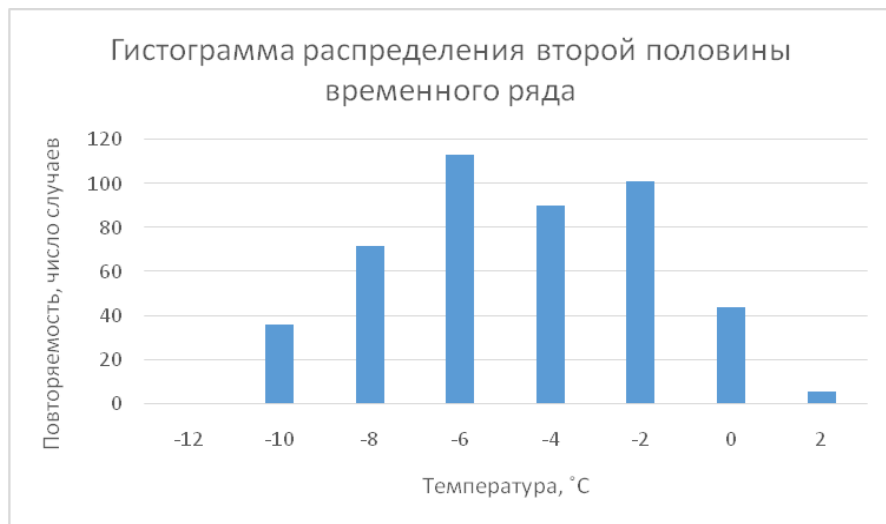


Рис. 7. Гистограмма второй половины ряда

4. Начинаем использовать Макросы для построения гистограмм

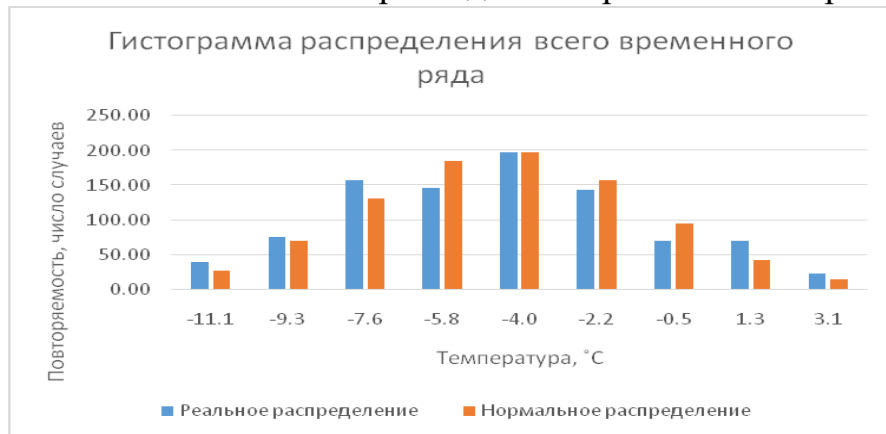


Рис. 8. Гистограмма всего временного ряда

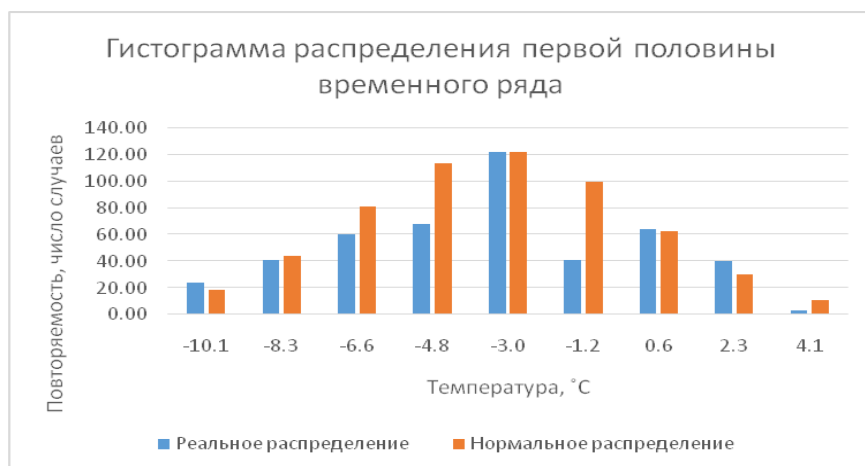


Рис. 9. Гистограмма первой половины временного ряда

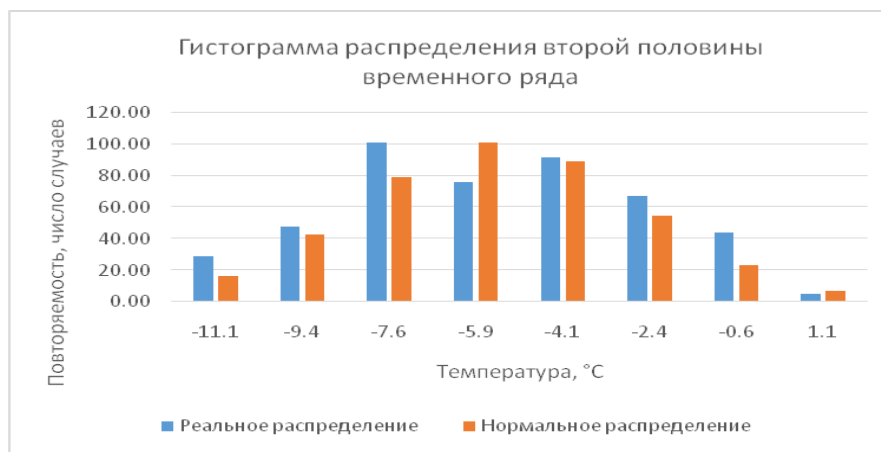
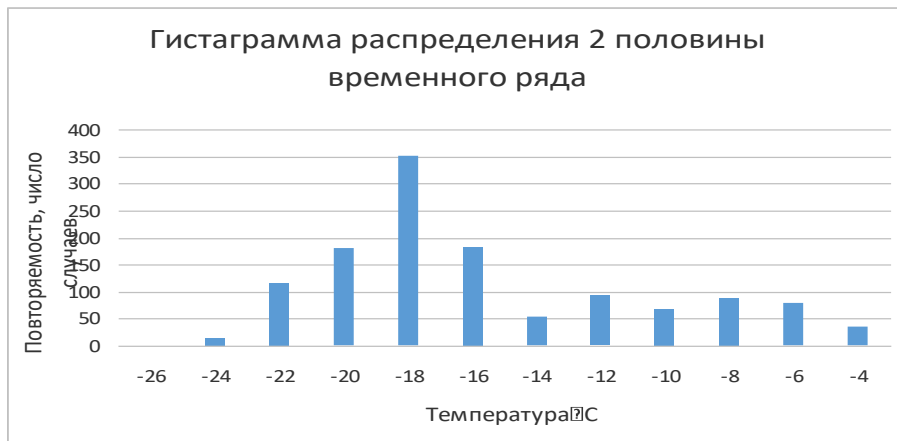


Рис. 10. Гистограмма второй половины временного ряда

Вывод: на всех трех гистограммах временных рядов мы наблюдаем симметричное распределение. Чаще всего «реальное» распределение выше «нормального» распределения.





Литература

1. Безручко Б.П., Смирнов Д.А. Математическое моделирование и хаотические временные ряды. Саратов: ГосУНЦ «Колледж», 2005. – 320 с.
2. Горбань А.Н. Функции многих переменных и нейронные сети // Соросовский образовательный журнал. 1998. № 12. С. 105-112.
3. Гордин В.А. Математика, компьютер, прогноз погоды. Л.: Гидрометеиздат, 1991.
4. Гриценко А.В. Использование локальных метеостанций для прогноза погоды. Метеокомплекс VANTAGE PRO 2. – Охрана окружающей среды и рациональное использование природных ресурсов 2009 / Материалы VIII международной научной конференции аспирантов и студентов. – Донецк, ДонНТУ – 2009.
5. Гриценко А.В. Прогнозирование метеопараметров по временным рядам. – Компьютерный мониторинг и информационные технологии 2008 / Материалы IV международной научной конференции студентов, аспирантов и молодых ученых. – Донецк, ДонНТУ – 2009.
6. Дроздов О.А., Васильев В.А. и др. Климатология. - Л.: Гидрометеиздат, 1989. – 568 с.
7. Дымников В.П., Филатов А.Н. Основы математической теории климата, М.: ВИНТИ, 1994.- 254 с.
8. Заенцев И.В. Нейронные сети: основные модели. Воронеж, 1999. 76 с.
9. Исаев А.А. Статистика в метеорологии и климатологии. – М.: Изд-во МГУ, 1988. – 245 стр.
10. Кобышева Н.В., Наровлянский Г.Я. Климатологическая обработка метеорологической информации. – Л.: Гидрометеиздат, 1978. - 96 с.
11. Кому и зачем нужен прогноз? Статья.
Источник: http://www.primpogoda.ru/articles/prosto_o_pogode/komu_i_za_chem_nuzhen_prognoz/
12. Кэйри П., Берк К. Анализ данных с помощью Microsoft Excel. – М.: Изд-во «Вильямс», 2005. – 560 с.
13. Лосев К.С. Климат: вчера, сегодня... и завтра? – Л.: Гидрометеиздат, 1985.- 176 с.
14. Малинецкий Г.Г., Потапов А.Б. Современные проблемы нелинейной динамики. М.: Эдиториал УРСС, 2000. 336 с.

15. Наставление гидрометеорологическим станциям и постам. Вып. Часть 2. Обработка материалов метеорологических наблюдений.- Спб.: Гидрометеиздат, 2000 – 48 с.
16. Немного истории о метеорологии. Статья.
Источник: http://www.primpogoda.ru/articles/prosto_o_pogode/o_meteorologii/
17. Носач В.В. Решение задач аппроксимации с помощью персональных компьютеров. – М.: МИКАП, 1994. – 382 с.
18. Сергин В.Я. Сергин С.Я. Системный анализ проблемы больших колебаний климата и оледенения Земли. – Л.: Гидрометеиздат, 1978. – Тюрин Ю.Н., Макаров А.А. статистический анализ данных на компьютере.- М.: ИНФА-М, 1998. – 528 с.
19. Чернышева Л.С., Платонова В.А. Закономерности временных изменений характеристик термической структуры холодного полугодия и их практическое использование. – Владивосток.: Изд-во ДВГУ, 2008. – 86 с.