



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

Кафедра метеорологических прогнозов

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(бакалаврская работа)

На тему: «Прогноз ливней для Санкт-Петербурга: оценка успешности»

Исполнитель Манёнок Алевтина Дмитриевна
(фамилия, имя, отчество)

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

Дробжева Яна Викторовна
(фамилия, имя, отчество)

«К защите допускаю»
заведующий кафедрой

(подпись)

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

Анискина Ольга Георгиевна
(фамилия, имя, отчество)

« 22 » июня 2024 г.

Санкт-Петербург
2024

СОДЕРЖАНИЕ

	Стр.
СОКРАЩЕНИЯ	4
ВВЕДЕНИЕ	5
1 Осадки	6
1.1 Понятие атмосферных осадков	6
1.2 Способы наблюдения осадков	7
1.3 Методы измерения осадков	8
1.4 Методы прогноза ливней	10
2 Архив данных ливней	21
2.1 Описание метеорологической станции Санкт-Петербург	20
2.2 Теория построения матриц сопряжённости	24
2.3 Критерии успешности прогнозов	30
2.3.1 Критерий независимости Пирсона	30
2.3.2 Критерий надёжности прогнозов Багрова	34
2.3.3 Критерий точности прогнозов Обухова	36
2.3.4 Меры Гутмана	36
2.3.5 Информационное отношение	38
3 Расчёт критериев.....	40

3.1 Расчёт критерия независимости Пирсона	40
3.2 Расчёт критерия надёжности прогнозов Багрова	41
3.3 Расчёт критерия точности прогнозов Обухова	41
3.4 Расчёт мер Гутмана	42
3.5 Расчёт информационного отношения	42
ЗАКЛЮЧЕНИЕ	43
Список использованных источников	46
Приложение 1	48

Сокращения

СПБ – Санкт-Петербург

гг. – года(-ы)

ОЯ – опасные явления

ГПа – гектопаскаль

СЗ УГМС – ФГБУ "Северо-Западное управление по гидрометеорологии и мониторингу окружающей среды"

См. – смотрите (далее приводится соответствующая глава)

ВВЕДЕНИЕ

Опасность ливней заключается в создании условий для возникновения наводнений, оползней, обвалов, лавин, других стихийных бедствий. Они приводят к гибели посевов сельхоз культур и урожая. Ливневые дожди парализуют движение транспорта, заливают подземные переходы, складские помещения.

Цель работы – оценка успешности прогнозов ливней для города Санкт-Петербурга.

Для достижения поставленной цели, необходимо решить следующие задачи:

1. Сформировать архив суточных прогностических и фактических данных ливней для г. Санкт-Петербурга за период 2013-2015 гг.
2. Построить матрицы сопряженности альтернативных методических, инерционных и случайных прогнозов ливней на основе имеющихся данных.
3. Рассчитать критерии успешности прогнозов ливней для г. Санкт-Петербург
4. Оценить успешность прогнозов ливней для г. Санкт-Петербург.

1. Осадки

1.1 Понятие атмосферных осадков

Атмосферные осадки — это атмосферные явления, связанные с присутствием в атмосфере воды в жидком или твёрдом состоянии, выпадающей из облаков или осаждающейся из воздуха на земную поверхность и какие-либо предметы.

Различают:

1. обложные осадки, связанные преимущественно с тёплыми фронтами;
2. ливневые осадки, связанные преимущественно с холодными фронтами.

Уместно выделить два основных определения ливней.

1. Сильный дождь, обычно крупнокапельный, интенсивность которого (т.е. количество осадков, выпавших за 1 мин) не ниже определённого предела. Предел этот тем ниже, чем больше продолжительность дождя. Например, при продолжительности в 5 мин ливнем следует считать дождь со средней интенсивностью в 0,50 мм/мин; в 30 мин -, 23 мм/мин; в 1 ч – 0,20 мм/мин; в 6 ч – 0,09 мм/мин и т.д.
2. Осадки, выпадающие из кучево-дождевых облаков. В этом значении чаще говорят: ливневые осадки.

1.2 Способы наблюдения осадков

Измерение атмосферных осадков является важной задачей в метеорологии и климатологии. Существует несколько методов, которые используются для измерения осадков. Рассмотрим некоторые из них:

1. Дождемеры – это приборы, которые используются для измерения количества дождя. Они обычно состоят из цилиндрического сосуда с масштабom, который позволяет измерить количество осадков в миллиметрах. Дождемеры устанавливаются на открытых площадках, чтобы получить точные данные о количестве осадков в данной области.

2. Снегомеры используются для измерения количества снега. Они обычно представляют собой цилиндрические сосуды с масштабom, которые устанавливаются на открытых площадках. Когда снег выпадает, он накапливается в снегомере, и его количество измеряется в миллиметрах. Эти данные помогают ученым и метеорологам понять, сколько снега выпало в данной области и как это может повлиять на климатические условия.

3. Градусники используются для измерения температуры осадков. Они могут быть установлены в дождемерах или снегомерах, чтобы измерить температуру дождя или снега. Градусники обычно имеют шкалу в градусах Цельсия или Фаренгейта, которая позволяет измерить температуру с высокой точностью.

4. Радары используются для измерения интенсивности осадков и их распределения в пространстве. Радары работают на основе отражения радиоволн от осадков. Они могут определить тип осадков (дождь, снег, град) и их интенсивность. Радары позволяют получить детальную информацию о погодных условиях в реальном времени и использовать ее для прогнозирования погоды.

Это лишь некоторые из методов, которые используются для измерения атмосферных осадков. Каждый метод имеет свои преимущества и ограничения, и их выбор зависит от конкретных условий и целей измерений.

1.3 Методы измерения осадков

Технологические инновации в измерении атмосферных осадков.

Измерение атмосферных осадков является важной задачей в метеорологии и климатологии. Технологические инновации в этой области помогают улучшить точность и надежность измерений, а также расширить возможности сбора данных. Вот некоторые из новых технологий, применяемых в измерении атмосферных осадков:

Автоматические погодные станции.

Автоматические погодные станции (АПС) представляют собой комплексные системы, оснащенные различными датчиками для измерения различных параметров погоды, включая осадки. Эти станции автоматически регистрируют и передают данные о количестве и интенсивности осадков. Они могут быть установлены на различных местах, включая открытые поля, горные вершины и океанские платформы.

Радары для измерения осадков.

Радары для измерения осадков (РЗИО) используются для определения количества и интенсивности осадков на больших территориях. Эти радары излучают радиоволны, которые отражаются от осадков и возвращаются обратно на радар. По времени задержки и интенсивности возвращенного сигнала можно определить количество и тип осадков. РЗИО позволяют

получать данные о распределении осадков на больших расстояниях и в реальном времени.

Спутниковые системы наблюдения.

Спутниковые системы наблюдения позволяют получать данные о количестве и распределении осадков на больших территориях. Спутники оборудованы специальными приборами, которые измеряют радиацию, отраженную от земной поверхности и атмосферы. По этим данным можно определить количество и тип осадков. Спутниковые системы наблюдения обеспечивают широкий охват и высокую пространственную разрешающую способность.

Дистанционное зондирование.

Дистанционное зондирование позволяет измерять осадки с помощью специальных приборов, установленных на самолетах или беспилотных летательных аппаратах (дронах). Эти приборы излучают электромагнитные волны и регистрируют их отражение от осадков. По этим данным можно определить количество и интенсивность осадков. Дистанционное зондирование позволяет получать данные о местных осадках и измерять их в труднодоступных или опасных местах.

Технологические инновации в измерении атмосферных осадков играют важную роль в улучшении наших знаний о погоде и климате. Они помогают улучшить прогнозы погоды, изучить климатические изменения и принять эффективные меры для управления водными ресурсами.

Прогноз морозящих осадков. Морозящие осадки выпадают, главным образом, из низкой слоистой облачности, образующиеся в воздухе, имеющем устойчивую стратификацию (тёплые воздушные массы). Выпадение морозящих осадков обусловлено неоднородностью размеров капель в облаке, причем более крупные капли появляются вследствие

коагуляции мелких капель. Поскольку моросящие осадки образуются в водяных облаках, они отмечаются в случаях, когда верхняя граница облаков не достигает температуры кристаллизации, у поверхности же земли при выпадении мороси температура воздуха чаще всего близка к 0°C. Моросящие осадки наплетаются преимущественно в холодное время года при адвекции влажного тёплого воздуха с моря на материк в тёплых секторах углубляющихся циклонов. Но морось иногда отмечается и в области антициклонов (чаще на периферии), сформированных воздушной массой морского происхождения, а также в зоне размытых старых тёплых фронтов и фронтов окклюзии. Моросящие осадки часто усиливаются в ночные часы в результате добавочного радиационного охлаждения.

1.4 Методы прогноза ливней

Прогноз моросящих осадков осуществляется путём прогноза перемещения имеющиеся зоны осадков по изогипсам АТ850 с учётом эволюции облачности. В тех случаях, когда ожидается интенсивная джержа очень тёплого и влажного воздуха, моросящие осадки могут быть предсказаны и тогда, когда на исходной карте они ещё отсутствуют. Прогноз обложных осадков. Обложные осадки выпадают из облаков As-Ns и связаны с упорядоченными восходящими движениями воздуха, охватывающими большое пространство и большую толщину атмосферы. Выпадение этих осадков в наиболее яркой форме проявляется в циклонах ложбинах, то есть в области тёплых и медленно движущихся холодных фронтов, а также фронтов окклюзии. Зона фронтальных осадков, как правило, имеет наибольшую ширину у тех участков фронта, которые находятся в области хорошо выраженной циклонической циркуляции, особенно вблизи

центральной части циклона. Интенсивность осадков тем больше, чем больше влажность воздушной массы и величина вертикальной скорости. Успешность прогноза обложных осадков зависит от того, насколько точно дал прогноз перемещения фронта. В большинстве случаев зоны обложных осадков связанных с инверсией тёплого воздуха, поэтому эти зоны обычно совпадают с областями наибольшего падения давления у поверхности земли. К зоне фронтальных осадков иногда присоединяется область внутри массовых осадков, выпадающих из подинверсионных облаков St и Sc. В этих случаях зона осадков значительно расширяется так как известно, эти облака являются преимущественно водяными и поэтому сами по себе чаще всего осадков не дают. В области фронта кристаллы, вызывающие осадки из данных облаков, выпадают из более высокой облачности Cs-As. В холодное время года нередко наблюдаются осадки, не связанные с какими-либо фронтами.

Наблюдаются они обычно в области антициклона. Такие осадки выпадают из подинверсионных облаков St и Sc, располагающихся в зоне температур -10°C и ниже и находящиеся в смешанной фазе. Переходные сезоны, а также нередко во время зимних оттепелей осадки могут выпадать как в виде дождя, так и в виде снега. Фазовое состояние осадков зависит в основном от высоты изотермической поверхности 0°C и температуры у поверхности земли. Анализ аэрологических данных показал, что осадки выпадают в виде дождя, если температура поверхности земли выше 3°C , а на уровне 500 м выше 0°C , и в виде снега, если температуру поверхности земли ниже 3°C , а на высоте 500 м ниже 0°C .

Прогноз обложных осадков, фронтальных и внутримассовых, разрабатывают в следующем порядке.

1. Предсказывают образование, перемещение и эволюцию облаков, дающих и обложные осадки, особенно фронтальных систем облаков.

2. Предсказывают перемещение существующей зоны обложных осадков одновременно с прогнозом перемещения атмосферного фронта и циклона, с которыми эта зона связана.

При этом применяют правила ведущего потока, используя прогностическую и фактическую карты АТ700.

1. Предсказывают эволюцию зоны обложных осадков в соответствии с прогнозом эволюции циклоны и фронта.

В качестве дополнительных факторов, способствующих или препятствующих выпадению обложных осадков, следует учитывать:

Высоту верхней границы системы облаков As-Ns, толщину слоя облаков и положения изотермы $T = -10^{\circ}\text{C}$. Выпадение осадков весьма вероятно, если изотерма -10°C будет расположена внутри системы облаков. Осадки наиболее вероятны при толщине слоя облаков более 2 км;

Знак и величину вертикальной составляющей скорости. Восходящее движение воздуха способствуют выпадению осадков, нисходящие-препятствуют;

Знак и величину адвективных изменений температуры и влажности. Положение зоны обложных осадков в общем совпадает с положением зоны наиболее интенсивной адвекцией тепла и влаги перед тёплым фронтом;

Неоднородность подстилающей поверхности. Областные осадки усиливаются с наветренной стороны даже относительно небольших возвышенностей. В настоящее время ведутся исследования по разработке численных методов прогноза осадков.

Прогноз ливневых осадков, гроз и града. Грозы и ливневые осадки чаще всего наблюдаются в слабовыраженных и заполняющихся циклонах, на перифериях малоподвижных антициклонов, в барических седловинах и в малоградиентном барическом поле, а также на фронтах, особенно на холодных. Ливневые осадки и грозы обязаны своим происхождением конвекцией влажного воздуха в условиях неустойчивой стратификации и связаны с развитием кучево-дождевой облачности.

Благоприятными условиями для возникновения ливневых осадков и гроз являются:

1. Высокое влагосодержание воздуха во всей толще тропосферы;
2. Значительная неустойчивость воздушной массы;
3. Высокая относительная влажность воздуха у поверхности земли и на высотах;
4. Относительно низкое положение изотермы - 10°C около которой располагается уровень интенсивной кристаллизации;
5. Большая вертикальная протяжённость кучево-дождевых облаков, верхняя граница которых в умеренных широтах достигает уровня с температурой примерно -23°C и ниже, высота конвекции простирается до высоты 6-7 км., и мощность облака превышает 4,5 км.;
6. Неоднородность подстилающей поверхности.

В кучево-дождевом облаке скорость восходящих потоков увеличивается с высотой, достигает максимума примерно в средней его части и затем убывает к вершине. Над уровнем максимальной скорости восходящих потоков образуется зона аккумуляции, где происходит накопление крупных облачных элементов. Если максимальная скорость восходящих потоков составляет более 10 м/с и отмечается выше уровня

нулевой изотермы, а вершина облака находится выше уровня кристаллизации имеет температуру ниже -25°C , то в таком облаке возможно образование града. Град образуется в тёплое время года, когда сильно развита тепловая конвекция. Особенно интенсивный град выпадает из фронтальных СВ.

При разработке метода прогноза максимального количества конвективных осадков использовались экспериментальные данные интенсивности и количества осадков, о верхней границе кучево-дождевых облаков к моменту начала выпадения осадков, измеренной радиолокатора (МРЛ). Также использовались данные метеорологических наблюдений станций и постов, карты барической топографии. Основой данного метода является определение количества прогнозируемых осадков по влагосодержанию облака с учетом того известного факта, что большая часть сконденсированной в облаке влаги остается в атмосфере и не участвует в осадкообразовании. Такой учёт осуществляется с помощью коэффициента генерации осадков k_0 – отношения фактически выпавших осадков к общему количеству генерируемой облаком влаги. Интенсивность осадков J (мм/мин) из крупномасштабных облачных систем определяется произведением параметров: максимальной конвективной скорости W_m (м/с) и относительной площади, занятой конвективными ячейками в облаке,

$$\frac{\Delta t_l}{\Delta t} ,$$

где

Δt_l - продолжительность выпадения ливневого дождя (мин.),

Δt - общая продолжительность дождя(мин.):

$$J = 0.024 \cdot Wm \cdot \frac{\Delta t_{л}}{\Delta t}$$

(коэффициент корреляции $r = 0.8$)

Коэффициент генерации осадков, согласно, зависит от продолжительности выпадения ливневого дождя $\Delta t_{л}$, характеризующей горизонтальные размеры суммы конвективных ячеек, из которых выпадает основная масса осадков:

$$k_0 = 0.0055 \cdot \Delta t_{л}$$

(коэффициент корреляции $r=0.92$).

Количество осадков определяется как

$$Q = J \cdot \Delta t$$

Из проведенных ранее исследований известно, что коэффициент генерации осадков для крупномасштабной и мезомасштабной облачности различен. Коэффициент генерации осадков зависит от горизонтального и вертикального масштаба облачности, разности значений удельной влажности при насыщении на уровнях конденсации и конвекции. Наибольший коэффициент генерации осадков – в циклонической облачности, наименьший – во внутримассовой. Различные значения коэффициента генерации осадков для разных типов облачности, по-видимому, связаны с различными размерами конвективных ячеек, характерных для данной облачности, из которой выпадают осадки. Из облаков одной и той же мощности и влагосодержания может выпасть разное количество осадков. Это объясняется разными условиями баланса влаги в кучево-дождевых облаках (Cb). Баланс влаги в Cb зависит от зон

крупномасштабной конвергенции. Упорядоченная вертикальная скорость τ мала на уровне моря и меняется монотонно до численного максимума на бездивергентном уровне. При её среднем значении в слое от уровня моря до бездивергентного уровня $\bar{\tau} > 0$ наблюдается дивергенция на уровне моря, при $\bar{\tau} < 0$ - конвергенция. Этот вывод следует из уравнения неразрывности и граничного условия на уровне моря $\tau_0 = 0$.

Зависимость коэффициента генерации осадков от среднего значения упорядоченных вертикальных токов $\bar{\tau}$:

$$k_0 = 0.2 - 0.0053 \cdot \tau$$

(коэффициент корреляции $r=0.76$),

где τ в гПа/12ч.

Таким образом, оказалось, что коэффициент генерации осадков зависит от вертикальной скорости синоптического масштаба.

В случаях, когда дождь ливневого характера вносит основной вклад в общую сумму осадков, максимальное количество осадков рассчитывается как функция только максимальной конвективной скорости, и уравнение регрессии имеет вид (мм/12ч.):

$$Q = c \cdot Wm$$

где c - эмпирический коэффициент.

При этом решение, по какой из зависимостей рассчитывать количество осадков, определяется пороговыми значениями влажности в слое 850–700 гПа (средним дефицитом точки росы), максимальной конвективной скорости, лапласиана приземного давления и вертикальных упорядоченных движений на уровне 850 гПа.

Ранее проведенные исследования показали: при сильных ливнях чем больше величина максимальной конвективной скорости, тем меньше разница между максимальной и средней интенсивностью осадков, продолжительностью ливневой части дождя и общей продолжительностью дождя. В связи с этим в формулу для прогноза максимального количества осадков введен коэффициент на интенсивность конвекции

$$L = f(Wm)$$

определенный из соотношения между максимальной и средней интенсивностями осадков, зависящих от максимальной конвективной скорости Wm .

Считается, что одну часть времени за дождь осадки выпадают с максимальной интенсивностью, а другую часть продолжительности осадков – со средней интенсивностью осадков. Для осадков со средней интенсивностью уравнения регрессии имеют, соответственно, вид:

$$Q' = 4.36 \cdot L \cdot Wm(0.2 - 0.0053\bar{t}),$$

$$Q' = c \cdot L \cdot Wm$$

Исследования, проведенные А.А. Акулиничевой показали, что отношение продолжительности ливневой части дождя к общей продолжительности дождя характеризуется плотностью конвективных струй χ , средние значения которой для летних месяцев находятся в интервале 0,29–0,38.

С учётом этого параметра окончательная формула прогноза максимального количества осадков имеет вид:

$$Q_{max} = Q \cdot \chi + Q' \cdot (1 - \chi).$$

Расчет максимального количества осадков физико-статистическим подходом производится с ежечасной детальностью, а затем на основе ежечасных данных прогнозируется количество осадков по полусуткам. Ежечасные прогнозы максимального количества осадков используются в рассматриваемом комплексном методе прогноза.

Данный метод является основой прогноза осадков во всех градациях. Альтернативный метод прогноза ливней градации ОЯ является уточняющим для градации осадков 50 мм/12 ч и более.

Прогноз ливней в градации ОЯ осуществляется по дискриминантным функциям, позволяющим по набору признаков с помощью выбранного решающего правила определить, к какому классу относится рассматриваемый объект: к классу «ливни в градации ОЯ» или к классу «без ливней в градации ОЯ». Для оценки качества разделения классов использовались такие характеристики, как: выборочное расстояние Махаланобиса и предупреденность и оправдываемость каждого из классов.

Отбор предикторов, определяющих выпадение ливней в градации ОЯ, производился на основе физической сущности происходящего процесса, т.е. исходя из условий и механизма образования ливней в градации ОЯ.

Так как главной причиной образования ливней в градации ОЯ является вертикальная неустойчивость атмосферы, мерой интенсивности

поднятия восходящего воздуха за счет конвекции выбрана величина максимальной конвективной скорости Wm .

Одним из основных факторов, влияющих на количество осадков, является влажность у поверхности Земли и степень увлажнения восходящего воздуха в слое тропосферы 850–700 гПа.

Точка росы у поверхности Земли ($Td3$) характеризует массовую долю влаги у поверхности Земли и влияет на высоту уровня конденсации. Степень увлажнения восходящего воздуха оценивается средним дефицитом влажности в слое 850-700 гП ($\bar{d} 8.7$). Лапласиан приземного давления $\nabla P3$ характеризует конвергенцию воздушных масс в нижнем слое тропосферы, благоприятствующую развитию суперячейковых кучево-дождевых облаков. Он косвенно определяет знак и модуль упорядоченных вертикальных движений на уровне 850 гПа, которые вносят вклад в интенсивность и количество осадков при ливнях в градации ОЯ.

Наибольшая информативность этих параметров подтверждена их ранжированием с помощью коэффициента точечно-биссеральной корреляции, который наиболее удобен для выявления связи между предикторами и двумя классами явлений, исключаящими друг друга.

Таким образом, прогностические дискриминантные функции для двух классов явлений:

ливневых осадков с количеством ≥ 50 мм /12ч. и без них $Lя$ и $Lбя$, имеют вид:

$$Lя = 0.242 \cdot Wm + 1.621 \cdot Td3 + 0.068 \cdot \nabla^2 P3 + 0.475 \cdot \overline{d8.7} - 18.0$$

$$Lбя = 0.117 \cdot Wm + 1.425 \cdot Td3 - 0.02 \cdot \nabla^2 P3 + 0.618 \cdot \overline{d8.7} - 12.4$$

Прогноз ливней градации ОЯ осуществляется по наибольшей по значению функции.

В комплексном методе осадки в градации ОЯ (≥ 50 мм/12ч) прогнозируются, если одновременно рассчитанное количество осадков ≥ 50 мм/12ч и прогнозируются ливни градации ОЯ. Если одно из условий не выполняется, прогнозируются осадки в соседней градации – неблагоприятных явлений (НЯ) от 30 до 50 мм/12ч.

2. Архив данных о ливнях

Для оценки прогнозов ливней необходимо было сформировать архив суточных прогностических и фактических данных ливней для г. Санкт-Петербург.

Прогностические данные были представлены ФГБУ "Северо-Западное управление по гидрометеорологии и мониторингу окружающей среды" за период апрель - октябрь 2013-2015 гг., всего 1277 данных.

Фактические данные были взяты с сайта gr5.ru. Информация о работе этого сайта схематически представлена на рисунках 2.1-2.4.

The screenshot shows the website gr5.ru displaying weather data for Saint-Petersburg. The current temperature is 22 °C. A forecast for the next 7 days is provided, including temperature ranges, precipitation, and wind speed. The website also features a search bar, navigation links, and a footer with contact information.

7 суток	Сегодня	Завтра	Ср	Чт	Пт	Сб	Вс	Пн			
Сегодня, вк, 24 июня	Завтра, шг, 25 июня	Ср, 26 июня	Чт, 27 июня	Пт, 28 июня	Сб, 29 июня	Вс, 30 июня	Пн				
Местное время	15 21	03 09 15 21	03 09 15 21	03 09 15 21	03 09 15 21	03 09 15 21	03 09 15 21	03			
Облачность, %											
Осадки, мм											
Туман, %	-1					-3					
Температура, °C	+21 +19	+20 +22	+19 +18	+21 +23	+20 +17	+22 +26	+22 +19	+23 +28	+29 +19	+20 +22	+20 +18
Давление, мм рт.ст.	760 761	762 763 763 763	762 762 761 761	761 762 762 762	763 763 762 761	761 761 759 757	758 758 757 757	756			
Ветер: скорость, м/с	1 3 2 2	3 5 4 4	2 3 6 4	2 1 2 3	2 2 3 3	1 1 2 3	4 3 4 3	4 2 1			
Порывы, м/с	6 8 5 7	12 10 5 7	13 10		7 7 5 5	7 6	6 7 9 7	11 6			
направление	С-З С-З	З С-З С-З	С-З С-З	С-З С-З	С-З С-З	С-З С-З	Ю-В Ю Ю-З	С-З С-З	С-З С-В		

Рисунок 2.1 – Общий вид сайта gr5.ru

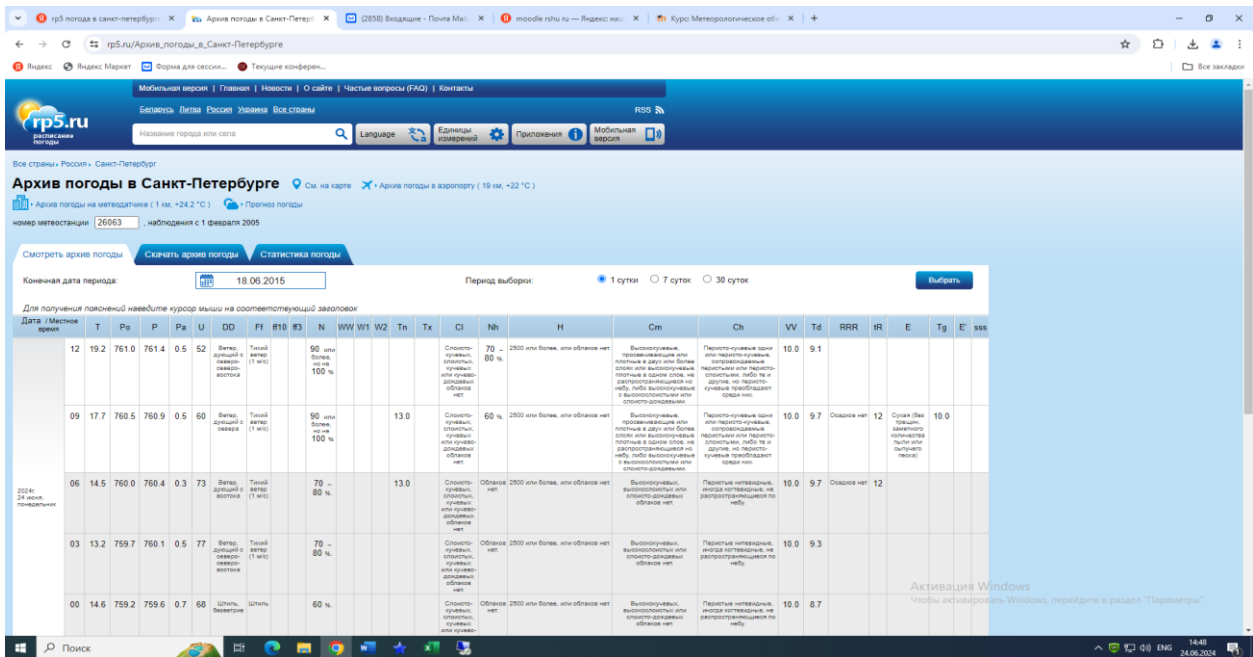


Рисунок 2.2 – Пример окна для скачивания фактических данных

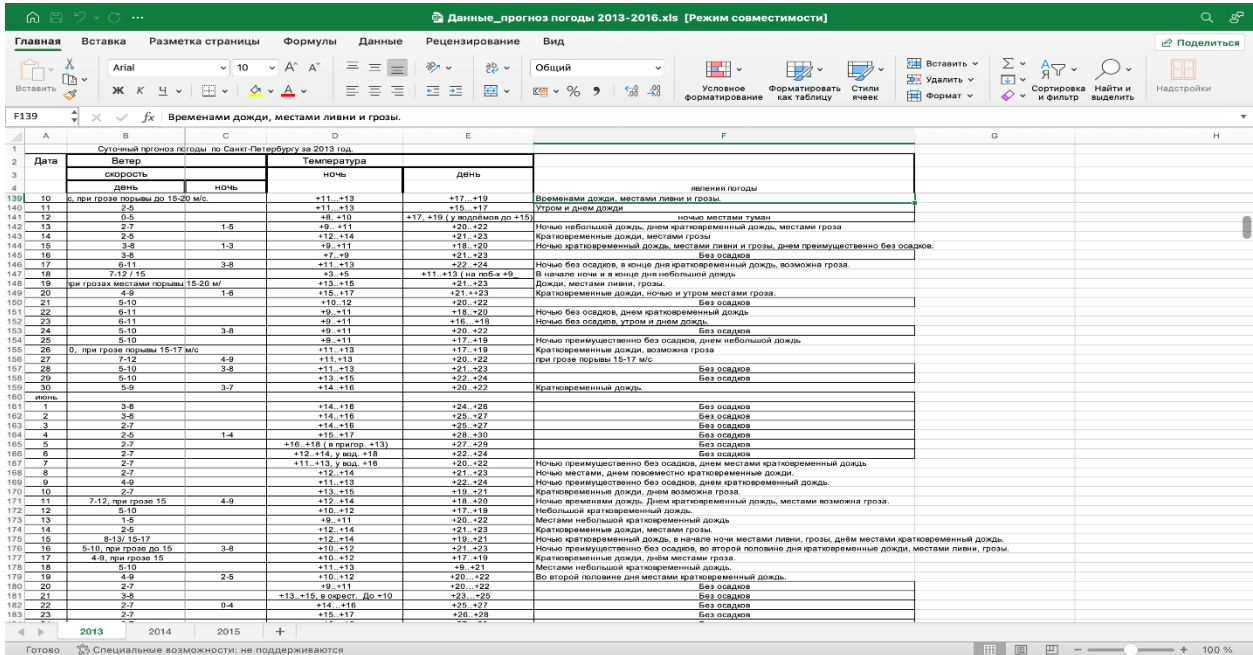


Рисунок 2.3 – Общий вид архива прогностических данных ливней в г. Санкт-Петербург за период 2013-2015гг. с сайта rp5.ru

The image shows a screenshot of an Excel spreadsheet titled 'ФАКТ_2013-2015.xls'. The spreadsheet contains a table of meteorological data for Saint-Petersburg. The columns include time (T), temperature (Po, P, Pa), wind speed (U, DD), and various weather indicators (Ff, #T0, #T3, N, WW, W1, W2, Tn, Tx, Cl, Nh, H, Cm, Ch, VV, Td, RRR, IR). The data is organized in rows, with the first row being a header and subsequent rows representing individual time intervals. The data shows fluctuations in temperature and wind speed, along with various weather conditions like 'Облаков нет', 'Облаков н', and 'Снег негр'.

Рисунок 2.4 – Общий вид фактических данных ливней в г. Санкт-Петербург за период 2013-2015гг. с сайта rp5.ru

2.1 Описание метеорологической станции Санкт-Петербург

СЗ УГМС находится по адресу: 199106, г. Санкт-Петербург, В.О., 23 линия, 2а.

В настоящее время Росгидромет является органом исполнительной власти, выполняющим функции по оказанию государственных услуг в области гидрометеорологии и иных связанных с ней областях, мониторинга окружающей среды и её загрязнения. Деятельность Росгидромета в сфере его полномочий направлена на повышение качества жизни населения, обеспечение высоких темпов устойчивого экономического развития страны и на повышение уровня гидрометеорологической безопасности населения и экономики РФ.

В число важных задач Росгидромета входит гидрометеорологическое обеспечение деятельности Российской Федерации в Арктике, Антарктике и Мировом океане.

2.2 Теория построения матриц сопряжённости

Матрица сопряжённости является обобщённой формой представления реализации прогноза погоды.

В виде матрицы сопряжённости представляется информация о прогнозируемых и фактических условиях погоды за выбранный отрезок времени (месяц, сезон, год). Она содержит число случаев (n_{ij}) соответствия этих условий ($\Pi_j \sim \Phi_i$).

Итак, матрица сопряжённости имеет вид матрицы размером $m \times n$, которая содержит два прогностических условия ($m=2$): «явление прогнозируется» (Π) или «не прогнозируется» ($\bar{\Pi}$) и два условия исхода ($n=2$): «явление было» (Φ), или «не было» ($\bar{\Phi}$).

Первая цифра при n – номер строки (i), вторая – номер столбца (j).

Π_j – прогноз данной градации, Φ_i – фактическая погода той же градации.

Итак, n_{ij} определяется по условию: что ожидалось по прогнозу (Π_j) и что фактически наблюдалось (Φ_i) или n_{ij} – число случаев сопряженности $\Phi_i \sim \Pi_j$. Соответствующие сочетания «прогноза» и «факта» характеризуют следующее:

n_{11} — число случаев оправдавшихся прогнозов наличия явления: явление (или состояния погоды) прогнозировалось Π и наблюдалось Φ ;

n_{21} — число случаев неоправдавшихся прогнозов наличия явления: явление (или состояния погоды) прогнозировалось Π , но фактически не наблюдалось $\bar{\Phi}$

n_{12} — число случаев неоправдавшихся прогнозов отсутствия явления : явление (или состояние погоды) не прогнозировалось $\bar{\Pi}$, но фактически наблюдалось Φ ;

n_{22} — число случаев оправдавшихся прогнозов отсутствия явления: явление (или состояния погоды) не прогнозировалось $\bar{\Pi}$ и фактически не наблюдалось $\bar{\Phi}$.

Здесь n_{12} — ошибки первого рода — ошибки пропуски; n_{21} — ошибки второго рода — ошибки страховки; n_{01} — число случаев прогнозов наличия явления (или состояния погоды) — число прогнозов с текстом Π ;

n_{02} — число случаев прогнозов отсутствия явления (или состояния погоды) — число прогнозов с текстом $\bar{\Pi}$,

n_{10} — число случаев наличия явления (или состояния погоды) - столько раз явление фактически наблюдалось Φ ; n_{20} — число случаев отсутствия явления (или состояния погоды) — столько раз явление фактически не наблюдалось $\bar{\Phi}$,

N — общее число прогнозов (за месяц, сезон и т.п.).

Построение матрицы сопряжённости методического прогноза.

Матрица сопряжённости методических прогнозов разрабатывается по следующей схеме:

1. За выбранный период времени устанавливается общее число прогнозов N и число случаев, когда явление фактически наблюдалось Φ , то есть n_{10} .
2. Устанавливается число случаев отсутствия явления $\bar{\Phi}$, то есть n_{20} .
3. Находим число случаев, когда явление прогнозировалось Π , то есть n_{01} .

4. Находим число случаев, когда явление не прогнозировалось $\bar{П}$, то есть $n_{02} = N - n_{01}$.

5. Из таблицы находим число случаев, когда явление не прогнозировалось $\bar{П}$, но наблюдалось Φ , то есть n_{12} .

6. Находим остальные элементы матрицы как разностные значения:

$$n_{11} = n_{10} - n_{12}; \quad n_{21} = n_{01} - n_{11}; \quad n_{22} = n_{20} - n_{21}.$$

В результате было найдено:

n_{01} — число прогнозов наличия явления $П$;

n_{02} — число прогнозов отсутствия явления $\bar{П}$,

n_{10} — число случаев, когда явление фактически наблюдалось Φ ;

n_{20} — число случаев, когда явление фактически не наблюдалось $\bar{\Phi}$,

N — общее число суточных прогнозов.

Построение матрицы сопряжённости инерционного прогноза.

Данная матрица устанавливается только по фактическим данным об явлении.

Исходное состояние или значение явления рассматривается как прогностическое $П_{ин}$ и выполняется сравнение с последующим состоянием или значением, которое принимается за факт $\Phi_{ин}$.

Последнее состояние или значение снова принимается за прогноз и т.д.

Такая цепочка сопоставления «прогноза - факта» выполняется в таблице данных «по вертикали». Общее число методических и инерционных прогнозов должно быть одинаково. Поэтому в начале ряда Φ_i следует условно добавить одно значение предшествовавшей погоды Φ_0 , допустить $\Phi_0 = \Phi_1$. Или сделать такую операцию в конце ряда $\Phi_{n+1} = \Phi_n$.

Порядок построения.

1. Фактически наблюдавшаяся частота явления погоды не зависит от способа прогнозирования. Поэтому n_{10} и n_{20} в матрице инерционного прогноза будут те же, что и в матрице методической прогноза.
2. Инерционные прогнозы обладают свойством несмещённости, то есть количество фактически наблюдавшихся фаз погоды (n_{10} и n_{20}) равно количеству текстов прогнозов (n_{01} и n_{02}).
3. Находим число ошибок – пропусков n_{12} . Для этого по таблице проводим сопоставление значений Φ_i «по вертикали». Так же благодаря свойству несмещённости в инерционной матрице $n_{12} = n_{21}$.
4. Находим соответствующие разности, чтобы определить остальные значения.

$$n_{11} = n_{10} - n_{12}; \quad n_{22} = n_{02} - n_{12}.$$

Построение матрицы сопряжённости случайного прогноза.

Прогнозы, при которых предсказанные значения метеорологических величин не зависят от распределения повторяемости осуществившихся значений, называются случайными. Такой прогноз статистически не связан с наблюдавшимся явлением, а следовательно, и вероятность осуществления признака Φ или $\bar{\Phi}$ не зависит от формулировки прогноза.

Рассмотрим более общее пояснение случайного прогноза. Независимость этих признаков означает, что знание признака Π не даёт нам никакой информации о признаке Φ . независимость признаков Π_j и Φ_i формулируется следующим образом:

1. доля элементов с признаком Φ среди элементов, обладающих Π , то есть $n_{11}/(n_{11}+n_{21})$, должно быть равно доле элементов с признаком Φ среди элементов, не обладающих признаком Π (иначе $\bar{\Pi}$), то есть

$$n_{12}/(n_{12}+n_{22});$$

$$\frac{n_{11}}{n_{11} + n_{21}} = \frac{n_{12}}{n_{12} + n_{22}} = \frac{n_{11} + n_{12}}{N}$$

2. доля элементов с признаком $\bar{\Phi}$ среди элементов, обладающих признаком Π , то есть $n_{21}/(n_{11}+n_{21})$, должна быть равна доле элементов с признаком $\bar{\Phi}$ среди элементов не обладающих признаком Π (иначе $\bar{\Pi}$), то есть $n_{22}/(n_{12}+n_{22})$; отсюда получаем

$$\frac{n_{21}}{n_{11} + n_{21}} = \frac{n_{22}}{n_{12} + n_{22}} = \frac{n_{21} + n_{22}}{N}$$

Если в качестве исходного признака рассматривать Π и $\bar{\Pi}$ и находить их доли соответственно среди элементов Φ и $\bar{\Phi}$, то можно записать

$$\frac{n_{11}}{n_{11} + n_{12}} = \frac{n_{21}}{n_{21} + n_{22}} = \frac{n_{11} + n_{21}}{N}$$

$$\frac{n_{12}}{n_{11} + n_{12}} = \frac{n_{22}}{n_{21} + n_{22}} = \frac{n_{12} + n_{22}}{N}$$

На основании данных соотношений устанавливается случайность сопряженности признаков Π_j и Φ_i . Таким образом,

$$n_{11\text{сл}} = \frac{(n_{11} + n_{22})(n_{11} + n_{12})}{N} = \frac{n_{01}n_{10}}{N}$$

$$n_{12\text{сл}} = \frac{(n_{11} + n_{22})(n_{11} + n_{12})}{N} = \frac{n_{02}n_{10}}{N}$$

Из других соотношений находим:

$$n_{21\text{сл}} = \frac{(n_{11} + n_{21})(n_{21} + n_{22})}{N} = \frac{n_{01}n_{20}}{N}$$

$$n_{22\text{сл}} = \frac{(n_{12} + n_{22})(n_{21} + n_{22})}{N} = \frac{n_{02}n_{20}}{N}$$

На основании (n_{ij}) частоты матрицы сопряженности методического прогноза определяются частоты при случайном выборе текста прогноза.

Таблица 2.1 Матрица сопряжённости методических прогнозов ливней

	П	$\bar{\text{П}}$	Σ
Ф	32	137	169
$\bar{\text{Ф}}$	14	457	471
Σ	46	594	640

Таблица 2.2 Матрица сопряженности инерционных прогнозов ливней

	П	$\bar{\text{П}}$	Σ
Ф	64	105	169
$\bar{\text{Ф}}$	105	366	471
Σ	169	471	640

Таблица 2.3 Матрица случайных альтернативных прогнозов ливней

	П	$\bar{П}$	Σn
Ф	12	157	169
$\bar{Ф}$	34	437	471
Σn	46	594	640

2.3 Критерии успешности прогнозов

2.2.1 Критерий независимости Пирсона

Оценка успешности альтернативных прогнозов.

Успешность прогнозирования погоды есть степень соответствия значений метеорологических величин или явлений погоды, которые содержатся в тексте прогноза и которые затем фактически наблюдались.

Критерии успешности позволяют дать количественную оценку сходства прогноза факту и тем самым установить метеорологическую результативность прогнозирования.

Есть популяция, характеризующаяся двумя качественными показателями, обозначим их X_1 и X_2 . X_1 может принимать k значений, X_2 может принимать m значений. Каждая наблюдаемая единица i характеризуется парой значений (x_{1i}, x_{2i}) .

Получается таблица: n_{ij} – это количество наблюдений, в которых зафиксировано значение $X_{1,i}$ первого знака и значение $X_{2,j}$ второго признака.

Таблица 2.4 Матрица сопряженности многофазовых прогнозов

	$X_{2,1}$	$X_{2,2}$...	$X_{2,m}$
$X_{1,1}$	n_{11}	n_{12}	...	n_{1m}
$X_{1,2}$	n_{21}	n_{22}	...	n_{2m}
...
$X_{1,k}$	n_{k1}	n_{k2}	...	n_{km}

Вопрос: есть ли зависимость между значениями переменных X_1 и X_2 .

Таблица 2.5 Матрица сопряженности многофазовых прогнозов

	$X_{2,1}$	$X_{2,2}$...	$X_{2,m}$	
$X_{1,1}$	n_{11}	n_{12}	...	n_{1m}	$N_1^1 = \sum_{j=1}^m n_{1j}$
$X_{1,2}$	n_{21}	n_{22}	...	n_{2m}	$N_2^1 = \sum_{j=1}^m n_{2j}$
...
$X_{1,k}$	n_{k1}	n_{k2}	...	n_{km}	$N_k^1 = \sum_{j=1}^m n_{kj}$
	$N_1^2 = \sum_{i=1}^k n_{i1}$	$N_2^2 = \sum_{i=1}^k n_{i2}$...	$N_m^2 = \sum_{i=1}^k n_{im}$	$N = \sum_{i=1}^k N_i^1 = \sum_{j=1}^m N_j^2$

Критерий Пирсона – это критерий проверки гипотезы о согласованности двух распределений.

Есть теоретическое распределение, согласно которому вероятности попадания в группу i равны p_i .

Тогда логично ожидать, что среди n наблюдений должно быть примерно np_i наблюдений – это теоретические частоты. На самом же деле мы пронаблюдали частоты n_i .

Интуиция подсказывает, что если нулевая гипотеза о том, что наблюдаемое распределение согласуется, верна, то наблюдаемые частоты n_i должны быть похожи на np_i .

Тестирование гипотез.

Формулируем некоторое утверждение, которое хотим проверить. Называем его нулевой гипотезой H_0 . Формулируем альтернативную гипотезу H_1 (H_a) – что происходит, если нулевая гипотеза нарушается. Далее собираем информацию (проводим опросы/наблюдения/замеры/...) и согласно β , вероятность несовершения ошибки II рода равна $(1 - \beta)$ – это называется какому-то правилу делаем вывод, следует ли отклонить нулевую гипотезу или нет.

В идеале нужно придумать некоторый механизм принятия решения относительно нулевой гипотезы, который

- i) Будет отклонять её, если она неверна
- ii) не будет отклонять её, если она верна.

Но бывают ошибки следующего типа:

- i) Ошибка I рода – ошибочно отклонить нулевую гипотезу, которая на самом деле верна.
- ii) Ошибка II рода – ошибочно не суметь отклонить нулевую гипотезу, которая на самом деле неверна (ложна).

Вероятность ошибки первого рода – это уровень значимости α .

Вероятность ошибки второго рода – это мощностью теста.

Обычно большее беспокойство доставляет ошибка I рода, поэтому будем требовать, чтобы вероятность ее совершения была не очень большой. Обычно, берут уровень значимости 5% (иногда 1%, иногда 10%).

Нулевую гипотезу можно отклонить или не отклонить. Но нельзя принять. Никакую гипотезу нельзя доказать статистическими методами.

Пример. H_0 : все лебеди белые.

H_1 : не все лебеди белые.

Мы пронаблюдали 1000 лебедей и все они оказались белыми.

Вопрос: означает ли это, что нулевая гипотеза верна ?

Нет, это означает только то, что мы увидели 1000 белых лебедей. Может быть, 1001-й оказался бы серым или каким-то еще. Это означает, что пока мы не нашли противоречий с нулевой гипотезой, т.е. мы не можем её отклонить.

Тестовая статистики определяется по формуле

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{(n_{ij} - n_{ij}^{сл.})^2}{n_{ij}^{сл.}}$$

Для определения теоретического (табличного) значения χ^2 рассчитывается число степеней свободы ν в таблице сопряженности.

(Число степеней свободы этого распределения определяется как разность между числом событий и числом связей, налагаемых моделью).

$$\nu = (k_1 - 1) \cdot (k_2 - 1)$$

где k_1 – число строк, k_2 – число столбцов.

Общая оправдываемость - это отношение числа оправдавшихся прогнозов к общему числу прогнозов.

Общая оправдываемость прогнозов $P_{\text{общ}}$, в соответствии с частотой n_{ij} реализации текста альтернативного прогноза определяется по формуле

$$P_{\text{общ}} = \frac{n_{11} + n_{22}}{N} \cdot 100\%$$

Г

Д

Е

2.3.2 Критерий надёжности прогнозов Багрова

Н

И

оцениваемых прогнозов к максимально возможному.

Н

Критерий N лишён недостатков, присущих критерию «общая оправдываемость». Значения N меняются в пределах от 0 (прогнозы на уровне случайных) до 1 (прогнозы на уровне идеальных). Мы прогнозируем наступление какого-то события. Известна его безусловная вероятность $p_{\text{сл}}$ – если мы не пользуемся никакими предсказательными моделями, то вероятность правильно сделать прогноз равна $p_{\text{сл}} : p_{\text{идеал}} = 1$ – это был бы идеальный вариант предсказания.

Л

Отношение N показывает какую часть потенциально возможного улучшения точности от наивного случайного прогноза, который оказывается правильным с вероятностью $p_{\text{сл}}$, до идеального, который оказывается верным с вероятностью $p_{\text{идеал}}$, мы достигаем при помощи оцениваемого метода, который дает правильность предсказания p .

У

$p_{\text{сл}}$ – это базовая точность, относительно которой мы всё измеряем.

Ч

p – это точность оцениваемого метода.

А

$p_{\text{идеал}}$ – это точность идеального метода.

Е

В

О

$$H = \frac{P - P_{сл}}{P_{идеал} - P_{сл}}$$

Интерпретация критерия H:

1. Безусловная вероятность ($P_{сл}$):
 - Это вероятность правильно предсказать событие, если не использовать никаких предсказательных моделей. Это наша базовая точность, или точность случайного прогноза.
2. Точность оцениваемого метода (P):
 - Это вероятность правильного предсказания события, используя оцениваемый метод прогноза.
3. Точность идеального метода ($P_{идеал}$):
 - Это идеальная точность, которая равна 1 (или 100%). Это максимально возможная точность прогноза.
4. Отношение H:
 - Показатель H показывает, какую часть потенциально возможного улучшения точности от наивного случайного прогноза ($P_{сл}$) до идеального прогноза ($P_{идеал}$) мы достигаем с помощью оцениваемого метода (P).
 - Чем ближе H к 1, тем лучше ваш метод прогнозирования по сравнению со случайным прогнозом и тем ближе он к идеальному прогнозу.
 - Если H=1, то ваш метод идеален.
 - Если H=0, то ваш метод не лучше случайного угадывания.

Таким образом, H действительно показывает, насколько эффективно ваш метод улучшает точность прогнозов по сравнению с базовой случайной точностью и как он приближается к идеальной точности.

$H(\Phi)$ – энтропия фактического распределения, она показывает неопределённость исходного распределения событий без учёта прогнозов. В данном контексте, $H(\Pi)$ действительно является условной энтропией.

2.3.3 Критерий надёжности прогнозов Q Обухова

Критерий точности прогнозов Q , предложенный М.А. Обуховым, рассчитывается по формуле

$$Q = 1 - \left(\frac{n_{12}}{n_{10}} + \frac{n_{21}}{n_{20}} \right).$$

Показатель Q выражает долю точных, успешных прогнозов при известной повторяемости фаз явлений. Он меняется в пределах от 0 (прогнозы на уровне случайных) до 1 (прогнозы на уровне идеальных). Если все прогнозы ошибочны $Q = -1$. Критерий отражает Q успех прогнозирования относительно идеальных прогнозов и рассчитывается по следующей формуле

$$Q = 1 - \left(\frac{n_{12}}{n_{10}} + \frac{n_{21}}{n_{20}} \right) = 1 - (FNR + FPR)$$

Чем меньше вероятности ошибок I и II рода, тем ближе значение Q к 1.

2.3.4 Меры Гутмана

Предназначены для оценки связи между известными признаками Π и Φ .

Есть три меры Гутмана. Меры Гутмана характеризуют относительный прирост успешности прогноза выбранного признака в частных или общем случае. У нас есть совокупность случаев сочетания

признаков Π и Φ (таблица сопряженности). В ней n строк и m столбцов. Оценим влияние признака Π на признак Φ .

Оценим насколько изменится прогнозирование признака Φ при знании признака Π или насколько изменится вероятность осуществления Φ при условии, что известен текст прогноза Π ?

Критерий Гутмана измеряет насколько точно прогнозы согласуются с фактическими наблюдениями. Чем ближе значение λ к 1, тем выше надежность прогноза. Значение, близкое к 0, указывает на то, что прогнозы не лучше случайных догадок.

Числитель формулы представляет собой корректировку суммы максимальных значений с учетом максимальной суммы по столбцам. Знаменатель формулы корректирует общее количество наблюдений с учётом максимальной суммы по столбцам.

n_{imax} - максимальное значение частоты в i -й строке, n_{0imax} - максимальное значение в строке сумм, определяющих повторяемость текстов прогнозов.

Меры Гутмана – ассиметричные меры связи. Область их изменения находится в пределах от 0 до 1. Если $\lambda(\Pi/\Phi) = 1$, то вероятность статистического прогноза Π по Φ максимальна. Если $\lambda(\Phi/\Pi) = 1$, то вероятность предсказания Φ при известном Π максимальна. Формулы в определенной степени являются аналогами условной вероятности.

Меры Гутмана $\lambda(\Pi/\Phi)$, $\lambda(\Phi/\Pi)$ и λ характеризуют относительный прирост успешности прогноза выбранного признака в частных или общем случаях.

2.3.5 Информационное отношение

Универсальный критерий - информационное отношение v .

$$\vartheta = \frac{H(\Phi) - H\left(\frac{\Phi}{\Pi}\right)}{H(\Phi)} = 1 - \frac{H\left(\frac{\Phi}{\Pi}\right)}{H(\Phi)}.$$

$$H(\Phi) = - \sum_{i=1}^n p(\Phi_i) \log p(\Phi_i)$$

Величина ϑ показывает, какая часть неопределённости климатологических прогнозов (природной неопределённости осуществления явления) устраняется с помощью методических прогнозов.

$\vartheta = 0$ – прогнозы на уровне случайных;

$\vartheta = 1$ – прогнозы на уровне идеальных, если $n_{12} = n_{21} = 0$.

И также $\vartheta = 1$, если все прогнозы ошибочны.

Если, например, $\vartheta = 0$, то прогнозы на уровне случайных.

До появления прогноза была энтропия на уровне $H(\Phi)$.

Энтропия – это мера неопределённости, хаоса.

Почему берётся знак минус? Потому что мы работаем с вероятностями, которые всегда лежат в отрезке $[0, 1]$, а значит, $\log_{10}(p(\Phi_i)) \leq 0$. Чтобы сумма была неотрицательной, перед знаком суммы ставится «–». Максимум энтропии достигается на равномерном распределении, т.е. в тех случаях, когда $p(\Phi_i) = 1/n$.

Это согласуется с идеей неопределённости – когда исходы равновероятны сложнее что-либо предугадать. Если условная энтропия мала, то это означает, что прогноз хорошо предсказывает фактическое состояние, т.е. неопределённость достаточно малая. Если прогноз является эффективным, то он сильно снижает неопределённость, т.е. делает условную энтропию $H(\Phi | \Pi)$ достаточно маленькой.

С точки зрения информационного отношения мы смотрим насколько удалось снизить неопределённость при помощи прогноза.

3 Расчёт критериев качества прогнозов

3.1 Расчёт критерия независимости Пирсона

$$P_{\text{ин}}^{\text{общ}} = \frac{64+366}{640} \cdot 100\% = 0.671$$

$$P_{\text{общ}}^{\text{м}} = \frac{32 + 457}{640} \cdot 100\% = 0,764$$

$$P_{\text{общ}}^{\text{сл}} = \frac{12 + 437}{640} \cdot 100\% = 0,7$$

Рассчитаем критерий независимости

$$\chi^2 = \frac{(32 - 12)^2}{12} + \frac{(14 - 34)^2}{34} + \frac{(137 - 157)^2}{157} + \frac{(457 - 437)^2}{437} = 48,5.$$

$$\vartheta = (2 - 1)(2 - 1) = 1$$

Задавая уровень значимости $\alpha=0,05$ (в 5% случаях, возможна независимость признаков - подтверждение нулевой гипотезы), получаем $\chi^2 \alpha \vartheta = 3,84$,

$$48,5 > 3,84$$

По правилу $\chi^2 > \chi^2 \alpha \vartheta$, то гипотеза о независимости признаков П и Ф отклоняются - между признаками взаимозависимость существует.

Аналогично рассчитаем критерий независимости к инерционному прогнозу

$$\chi^2 = \frac{(64 - 12)^2}{12} + \frac{(105 - 34)^2}{34} + \frac{(105 - 157)^2}{157} + \frac{(366 - 437)^2}{437} = 402,3.$$

$$402,3 > 3,84$$

По правилу $\chi^2 > \chi^2 \alpha \vartheta$, то гипотеза о независимости признаков П и Ф отклоняются - между признаками взаимозависимость существует.

3.2 Расчёт критерия надёжности прогнозов Багрова

$$H(\Phi) = -\sum p(\Phi) \cdot \lg p(\Phi) = -\left(\frac{n_{10}}{N} \cdot \lg \frac{n_{10}}{N} + \frac{n_{20}}{N} \cdot \lg \frac{n_{20}}{N}\right) = -\frac{1}{640} \left(169 \lg \frac{169}{640} + 471 \lg \frac{471}{640}\right) = 0,251.$$

$$\begin{aligned} H(\Pi) &= -\left[\frac{n_{01}}{N} \left(\frac{n_{11}}{n_{01}} \cdot \lg \frac{n_{11}}{n_{01}} + \frac{n_{21}}{n_{01}} \cdot \lg \frac{n_{21}}{n_{01}}\right) + \frac{n_{02}}{N} \left(\frac{n_{12}}{n_{02}} \cdot \lg \frac{n_{12}}{n_{02}} + \frac{n_{22}}{n_{02}} \cdot \lg \frac{n_{22}}{n_{02}}\right)\right] \\ &= -\left[\frac{46}{640} \left(\frac{32}{46} \cdot \lg \frac{32}{46} + \frac{14}{46} \cdot \lg \frac{14}{46}\right) + \frac{594}{640} \left(\frac{137}{594} \cdot \lg \frac{137}{594} + \frac{457}{594} \cdot \lg \frac{457}{594}\right)\right] \\ &= 0,07 \cdot (0,7 \cdot (-0,15) + 0,3 \cdot (-0,5)) + 0,9 \cdot (0,23 \cdot (-0,63) + 0,8 \cdot (-0,1)) = 0,237. \end{aligned}$$

3.3 Расчёт критерия точности прогнозов Обухова Q

$$Q^m = 1 - \left(\frac{137}{169} + \frac{14}{471}\right) = 0.15.$$

$$Q^{un} = 1 - \left(\frac{105}{169} + \frac{105}{471}\right) = 0.15.$$

$$Q^{cl} = 1 - \left(\frac{157}{169} + \frac{34}{471}\right) = -0.001.$$

3.4 Расчёт мер Гутмана

$$\lambda_{\Phi} \left(\frac{\Phi}{\Pi} \right) = \frac{\sum_{j=1}^m n_{jmax} - n_{iomax}}{N - n_{iomax}} = \frac{n_{1max} + n_{2max} - n_{iomax}}{N - n_{iomax}} = \frac{137 + 457 - 594}{640 - 594} = 0.$$

$$\lambda_{\Phi} \left(\frac{\Pi}{\Phi} \right) = \frac{105 + 366 - 471}{640 - 471} = 0.$$

$$\lambda_{\Phi} \left(\frac{\Phi}{\Phi} \right) = 0.$$

$$\lambda_{\Phi} \frac{((137 + 457) - 594) + ((32 + 457) - 471)}{2 \cdot 640 - 47 - 594} = 0.084.$$

$$\lambda_{\Phi} \frac{((105 + 366) - 471) + ((105 + 366) - 471)}{2 \cdot 640 - 471 - 471} = 0.$$

$$\lambda_{\Phi} \left(\frac{\Pi}{\Phi} \right) = \frac{\sum_{j=1}^m n_{jmax} - n_{iomax}}{N - n_{iomax}} = \frac{n_{1max} + n_{2max} - n_{iomax}}{N - n_{iomax}} = \frac{32 + 457 - 471}{640 - 471} = 0,1.$$

3.5 Расчёт информационного отношения

$$V = 1 - \frac{0.237}{0.251} = 0.0557768924 = 5.5777\% \approx 6\%$$

$$V = 1 - \frac{0.22}{0.251} = 0.123505976 = 12.35\% \approx 12\%$$

ЗАКЛЮЧЕНИЕ

Общий вывод.

На основании критерия Пирсона мы отклоняем нулевую гипотезу о независимости показателей Ф и П и делаем вывод, что между переменными существует связь. Однако Критерий Пирсона не говорит ни о силе этой связи, ни о её направленности.

Методические (м) и инерционные (ин) прогнозы показывают некоторую степень точности и являются лучше случайных, но значительно отстают от идеальных прогнозов.

Меры Гутмана.

Методическая оценка (м)

1. Влияние признака П (прогноз) на признак Ф (факт):

Значение $\lambda(\Phi/\Pi)=0.1$. Это указывает на то, что прогноз (методическая оценка) оказывает слабое влияние на фактические результаты. То есть,

точность прогнозов незначительно выше, чем случайные прогнозы.

2. Влияние признака Ф (факт) на признак П (прогноз):

Значение $\lambda(\Pi/\Phi)=0$.

Это означает, что фактические результаты не оказывают никакого влияния на методическую оценку прогнозов.

Иными словами, методические прогнозы не улучшаются в зависимости от

фактических данных.

3. Взаимное влияние признаков П (прогноз) и Ф (факт):

Значение $\lambda=0.084$. Это указывает на слабое взаимное влияние между прогнозами и фактическими результатами. Они оказывают

незначительное влияние друг на друга, что свидетельствует о низкой степени взаимосвязи.

Инерционная оценка (ин)

1. Влияние признака П (прогноз) на признак Ф (факт):

Значение $\lambda(\Phi/\Pi)=0$.

Это указывает на то, что прогнозы (инерционная оценка) не оказывают влияния на фактические результаты, и точность прогнозов не отличается от случайных прогнозов.

2. Влияние признака Ф (факт) на признак П (прогноз):

Значение $\lambda(\Pi/\Phi)=0$. Это означает, что фактические результаты не оказывают влияния на инерционные прогнозы. Прогнозы не улучшаются в зависимости от фактических данных.

3. Взаимное влияние признаков П (прогноз) и Ф (факт):

Значение $\lambda=0$. Это указывает на отсутствие взаимного влияния между прогнозами и фактическими результатами. Они не оказывают влияния друг на друга, что свидетельствует о полной независимости.

Методическая оценка:

Прогнозы оказывают слабое влияние на фактические результаты, но фактические результаты не оказывают влияния на точность прогнозов.

Взаимное влияние между прогнозами и фактическими результатами также слабое.

• Инерционная оценка:

Прогнозы и фактические результаты не оказывают влияния друг на друга, что свидетельствует о полной независимости этих признаков.

Информационное отношение.

Первый расчет ($v=6\%$): Это значение указывает на то, что методические прогнозы устраняют 6% неопределённости по сравнению с полностью случайными прогнозами.

Это слабый результат, показывающий, что методические прогнозы лишь немного лучше случайных.

- Второй расчёт ($v=12\%$): Это значение указывает на то, что методические прогнозы устраняют 12% неопределённости по сравнению с полностью случайными прогнозами. Это чуть более высокий результат, показывающий, что методические прогнозы оказываются немного более точными, но все еще далеко от идеальных.

Значения информационного отношения v показывают, что методические прогнозы лишь незначительно лучше случайных прогнозов, устраняя от 6% до 12% неопределенности. Это указывает на необходимость улучшения методических прогнозов для достижения более высокой точности и надёжности.

Список использованных источников

1. Хандожко Л.А. Оптимальные погодо-хозяйственные решения, РГГМУ, СПб, 1999 - 162 с.
2. Дробжева Я.В., Волобуева О.В. Метеорологические прогнозы и их экономическая полезность. [Текст]: учебное пособие – СПб.: Адмирал, 2016 – 116 с.
3. Алексеева А.А., Лосев В.М. Прогноз сильных летних осадков на основе выходных данных региональной модели Гидрометцентра России 2014 - 16с.
4. Наставление гидрометеорологическим станциям и постам выпуск 3 часть 1, СССР, Ленинград, Гидрометеиздат, 1985 - 300с.
5. Полякова В.В., Шаброва Н.В. Основы теории статистики, Екатеринбург, 2015 - 148с.
6. Измерение атмосферных осадков: методы, точность и проблемы [Электронный ресурс]. Режим доступа: <https://nauchniestati.ru/spravka/izmerenie-atmosfernyh-osadkov/>
7. Сайт Гидрометцентра России [Электронный ресурс]. Режим доступа: <https://meteoinfo.ru>
8. Электронная библиотека РГГМУ [Электронный ресурс]. Режим доступа:
9. <http://elib.rshu.ru/>
10. RSHU-MOODLE, курс Дробжева Я.В. Метеорологическое обеспечение народного хозяйства [Электронный ресурс]. Режим доступа:
11. [Курс: Метеорологическое обеспечение народного хозяйства | RSHU-MOODLE](#)

12.Погода в Санкт-Петербурге [Электронный ресурс]. Режим доступа:
gr5.ru

13.Савельев В. Статистика и котики, 2016 - 196 с.

Приложение 1

Таблица ливней в городе Санкт-Петербурге за период 2013-2015гг.

Ливни в СПб за 2013 год.

Месяц	Апрель		Май		Июнь		Июль		Август		Сент		Окт	
Дата	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф
1	-	-	-	-	-	-	~	~	-	~	-	~	-	~
2	-	-	-	-	-	-	-	~	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	~	-	-	-	-	-	-
5	-	-	-	~	-	-	~	-	-	-	-	-	-	-
6	-	-	-	~	-	-	~	-	-	-	-	-	-	-
7	-	-	-	-	-	-	~	~	-	-	-	-	-	~
8	-	~	-	-	-	~	-	-	-	-	-	-	-	~
9	-	-	-	-	-	~	-	-	-	-	-	-	-	~
10	-	-	~	~	-	-	-	-	-	~	-	-	-	-
11	-	~	-	-	-	~	-	~	~	~	-	-	-	-
12	-	-	-	-	-	~	-	-	-	~	-	~	-	-
13	-	~	-	~	-	-	-	-	-	-	-	-	-	-
14	-	-	-	~	-	-	-	-	~	~	-	-	-	-
15	-	-	~	-	~	~	-	-	~	~	-	-	-	-
16	-	-	-	-	~	~	-	~	-	~	-	-	-	-
17	-	~	-	~	-	~	-	-	-	-	-	-	-	-
18	-	-	-	-	-	-	-	~	-	-	-	-	-	-
19	-	-	~	-	-	-	-	~	-	~	-	-	-	~
20	-	-	-	~	-	-	-	~	-	-	-	-	-	-

21	-	-	-	-	-	-	-	-	-	~	-	-	-	-
22	-	-	-	-	-	-	~	~	-	-	-	~	-	-
23	-	-	-	~	-	-	-	~	-	-	-	~	-	~
24	-	-	-	-	-	-	-	~	-	-	-	~	-	~
25	-	-	-	~	-	-	-	-	-	-	-	-	-	-
26	-	-	-	-	-	-	-	-	-	-	-	-	-	-
27	-	~	-	-	~	~	-	-	-	-	-	-	-	~
28	-	-	-	-	-	-	-	-	-	-	-	~	-	~
29	-	-	-	-	-	~	-	-	-	-	-	-	-	~
30	-	~	-	-	~	~	-	-	-	~	-	-	-	~
31			-	~			-	-	-	-	-	-	-	-

Ливни в СПб за 2014 год.

Месяц	Апрель		Май		Июнь		Июль		Август		Сент		Окт	
Дата	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф
1	-	~	-	-	-	-	-	-	~	-	-	-	-	~
2	-	-	-	~	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	~	~	-	~	-	-	-	-	-	-
4	-	~	-	-	~	-	-	~	-	-	-	-	-	-
5	-	-	-	~	-	-	-	-	-	-	-	-	-	-
6	-	~	-	-	-	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	~	-	-	-	-	-	-	-	-
8	-	-	-	-	~	~	-	-	~	~	-	-	-	-
9	-	-	-	-	-	~	-	-	-	-	-	-	-	-
10	-	-	~	~	-	-	-	-	-	-	-	-	-	-
11	-	-	~	~	-	-	-	-	-	-	-	-	-	-
12	-	-	-	-	-	~	-	-	~	~	-	~	-	-
13	-	-	-	~	-	-	-	-	-	~	-	~	-	-
14	-	~	-	-	-	-	-	-	-	-	-	-	-	-
15	-	-	-	-	-	-	-	-	~	-	-	-	-	-
16	-	-	-	-	-	~	~	~	-	~	-	-	-	-
17	-	-	-	-	-	~	~	-	-	~	-	-	-	-
18	-	-	-	-	-	~	~	-	-	-	-	-	-	-
19	-	-	-	-	-	-	-	-	-	~	-	-	-	-
20	-	-	~	-	-	-	-	-	-	~	-	-	-	-
21	-	-	~	-	-	~	-	-	-	-	-	-	-	~
22	-	-	-	~	-	~	-	-	-	~	-	~	-	-

23	-	-	-	-	-	-	-	-	-	~	-	-	-	-
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25	-	-	-	~	-	-	-	-	-	-	-	~	-	-
26	-	-	-	-	-	-	-	-	-	~	-	-	-	-
27	-	-	-	-	-	-	-	-	~	-	-	~	-	-
28	-	-	-	-	-	-	-	-	-	~	-	~	-	-
29	-	-	-	-	-	-	-	-	-	~	-	-	-	-
30	-	-	-	-	-	-	-	~	-	-	-	~	-	-
31														

Ливни в СПб за 2015 год.

Месяц	Апрель		Май		Июнь		Июль		Август		Сент		Окт	
Дата	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф	П	Ф
1	-	-	-	-	-	-	-	~	-	-	-	-	-	~
2	-	~	-	~	-	-	-	-	-	~	-	-	-	~
3	-	-	-	-	-	-	-	-	-	~	-	~	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-	~	-	-
6	-	-	-	-	-	-	-	-	-	-	-	~	-	~
7	-	-	-	-	-	~	~	~	~	~	-	~	-	~
8	-	~	-	~	-	-	-	~	-	~	-	~	-	-
9	-	-	-	-	-	-	-	-	~	-	-	-	-	~
10	-	~	-	-	-	-	-	~	-	-	-	-	-	~
11	-	-	-	-	-	~	~	~	-	-	-	-	-	-
12	-	~	-	-	-	~	-	-	~	-	-	-	-	-
13	-	-	-	~	-	-	-	-	~	~	-	-	-	-
14	-	-	-	~	-	-	-	~	-	-	-	-	-	-
15	-	~	-	-	-	~	-	-	-	-	-	-	-	-
16	-	-	-	-	-	-	-	-	-	-	-	-	-	~
17	-	-	-	-	-	-	~	~	-	-	-	-	-	-
18	-	-	-	-	-	~	-	-	-	-	-	-	-	-
19	-	-	-	~	-	~	~	-	-	-	-	-	-	-
20	-	~	-	-	-	-	-	~	-	-	-	-	-	-
21	-	-	-	-	-	~	-	-	-	-	-	-	-	-
22	-	-	-	-	-	-	-	~	-	-	-	-	-	-

23	-	~	-	-	~	~	-	~	-	-	-	-	-	-
24	-	-	-	~	~	~	-	~	-	-	-	-	-	-
25	-	-	-	-	-	-	-	-	-	-	-	-	-	-
26	-	-	-	-	-	-	-	-	-	~	-	~	-	-
27	-	-	~	~	-	-	~	-	-	-	-	-	-	-
28	-	~	~	-	-	-	-	-	-	~	-	-	-	-
29	~	~	-	-	-	-	-	-	-	-	-	-	-	-
30	-	-	-	-	-	-	~	~	-	-	-	-	-	-
31							~	~	-	-	-		-	-