

Министерство природных ресурсов и экологии Российской Федерации
Федеральная служба по гидрометеорологии и мониторингу окружающей среды

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ
«СИБИРСКИЙ РЕГИОНАЛЬНЫЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ
ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ ИНСТИТУТ»
(ФГБУ «СибНИГМИ»)

УДК 551.509.1/.5

Per. № НИОКТР АААА-А17-117022250099-8

Per. № ИКРБС АААА-Б19-219121790066-0



УТВЕРЖДАЮ
Директор ФГБУ «СибНИГМИ»
А.Б. Колкер
2019 г

ОТЧЕТ

О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

Метод и технология краткосрочного прогнозирования переходов температуры
через ноль и связанных с ними заморозков и гололедных явлений
на территории Урало-Сибирского региона

по теме:


РАЗВИТЬ ТЕХНОЛОГИИ ДЕТАЛИЗИРОВАННОГО ЧИСЛЕННОГО
КРАТКОСРОЧНОГО ПРОГНОЗА МЕТЕОРОЛОГИЧЕСКИХ ПАРАМЕТРОВ,
ЭЛЕМЕНТОВ И ЯВЛЕНИЙ ПОГОДЫ, ВКЛЮЧАЯ ОПАСНЫЕ,
ПО ТЕРРИТОРИИ СИБИРИ
(заключительный)

Тема 1.1.9.1 п.1

Новосибирск 2019


СПИСОК ИСПОЛНИТЕЛЕЙ

Отв. исполнитель,
ведущий научный
сотрудник, канд.
геогр. наук


13.12.2019

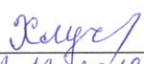
М.Я.Здерева
научное руководство, постановка задач,
подготовка программ распознавания образов,
программ оценок, проведение расчетов,
анализ результатов, написание отчета

Зав. отделом
СибНИГМИ


13.12.2019


В.М.Токарев
научное руководство, разработка алгоритмов
кодировки гололедных явлений, написание
программ обработки данных, разработка
критериев и вариантов фильтрации решений,
написание отчета

Научный сотрудник


12.12.2019


Н.А. Хлучина
написание программ для форматирования
данных КН-01, написание программ для
форматирования выходных полей моделей,
разработка технологических блоков

Младший научный
сотрудник


13.12.19


Л.П.Воробьева
подготовка базовой модельной продукции,
разработка технологических блоков

Младший научный
сотрудник


13.12.19

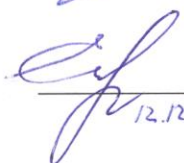
Т.В. Будз
ГИС-картирование результатов диагноза и
прогноза, работа с приложениями, анализ
результатов, оформление отчета

Инженер


11.12.2019


М.С. Котов
разработка и сопровождение архивирования
информации в SQL базе

Инженер


12.12.19

Н.В.Елфимова
подготовка диаграмм, таблиц, графиков,
форматирование и редактирование отчета

Нормоконтролер


16.12.19

Т.П. Панькова

РЕФЕРАТ

Отчет 56 с., 20 рис., 12 табл., 20 источн.

ГОЛОЛЕД, ЗАМОРОЗКИ В ВОЗДУХЕ И НА ПОЧВЕ, ЛОГИЧЕСКИЕ БИНАРНЫЕ ДЕРЕВЬЯ РАЗДЕЛЕНИЯ ОБРАЗОВ, ИНТЕРПРЕТАЦИЯ ГИДРОДИНАМИЧЕСКИХ МОДЕЛЕЙ

В основе работ лежит интерпретация выходной продукции моделей атмосферы COSMO-RU_Sib13 и GFS(NCEP), на базе которой строятся бинарные логические деревья распознавания образов по DW-алгоритму. Продолжена работа над оптимизацией данного алгоритма, направленная на повышение устойчивости решений за счет исключения случайных или редких связей факторов с предиктантом. Апробированы новые варианты сбалансированного критериального упрощения решений.

Разработана типизация гололедных явлений по приземным метеорологическим условиям их образования и продолжительности, всего 9 типов. Для каждого типа рассмотрена обеспеченность на архивной выборке 2014-17 годов и распределение по территории.

Для заморозков в воздухе, заморозков на почве и для гололеда по оптимизированному DW алгоритму построены деревья решений. Получены и проанализированы сравнительные оценки для всех видов прогнозов на зависимом и независимом материале. Методы и технология прогнозов гололеда, заморозков в воздухе и на почве передаются на оперативные испытания в подразделения Гидрометслужбы Урало-Сибирского региона.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	5
1 Состояние вопроса	8
2 Построение логических решающих правил на базе бинарных деревьев.....	12
2.1 Базовая методика распознавания образов.....	12
2.2 Методология получения комплекса решающих правил и выбора оптимального.....	15
2.3 Подготовка базовых данных.....	17
3. Прогноз перехода температуры через ноль в весенне-осенние периоды и связанных с ним заморозков.....	20
3.1 Оценка прогнозов перехода температуры через ноль на уровне двух метров по модельным выходным параметрам.....	20
3.2 Сравнительные оценки прогнозов заморозков в воздухе на независимой выборке ...	26
3.3 Сравнительные оценки прогнозов заморозков на почве.....	32
4 Гололед и гололедица – факт.....	37
4.1 Авторская типизация явлений.....	37
4.2 Обеспеченность данными гололеда и гололедицы по разработанным кодам на обучающей выборке.....	39
4.3 Оценка прогнозов гололеда.....	48
5 Технологические блоки прогнозов.....	51
ЗАКЛЮЧЕНИЕ.....	52
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	55

ВВЕДЕНИЕ

Целью данного исследования является разработка методов прогноза опасных явлений погоды, связанных с переходом температуры воздуха через ноль, и встраивание их в оперативную технологию получения расчетов, доведения результатов до потребителей и их оценок.

Переходы температуры через ноль не были бы столь значимы для условий жизни и всех сфер жизнеобеспечения, если бы не сопровождались фазовыми переходами воды. Отсюда следует и прямая климатическая зависимость частот и интенсивности гололедно-изморозевых явлений от близости незамерзающих океанов и морей. Урало-Сибирский регион в целом далек от таких побережий, за исключением полярных районов Урала и Западной Сибири, которые в теме не рассматриваются, но потери от заморозков и гололедицы, тем не менее, весьма значительны и растут по мере развития городской и транспортной инфраструктуры.

В данной теме рассматриваются три различных типа явлений, связанных с переходами температуры через ноль:

- осенне-весенние заморозки,
- гололед,
- гололедица.

Каждое из них имеет свои морфологические и физические особенности, которым должна соответствовать своя методика обработки данных и прогнозирования.

Формально заморозки могут повторяться несколько раз за сезон. Понятно, что наибольшую значимость имеют самые ранние заморозки летом-осенью и самые поздние весной-летом, но не все так однозначно. Пространственный фон температуры поверхности в условиях радиационного выхолаживания, а это основная форма заморозков, весьма изменчив в подсеточном для сети метеостанций масштабе и сильно зависит от рельефа, близости водоемов, инфраструктуры и прочих неметеорологических факторов. Поэтому не только прогноз, но и диагноз этого опасного явления должен рассматриваться как фоновый в масштабе десятков километров. Это означает, что гарантировать отсутствие и наличие заморозков можно только за пределами слишком широкого, а потому неприемлемого потенциального диапазона температур выше и ниже нуля.

К сожалению, нормативные документы и отсутствие опыта потребителей метеоинформации пока не позволяют использовать вероятностные формулировки прогнозов, которые наиболее адекватны природе пространственно-временной изменчивости метеорологических характеристик, условий и явлений погоды. В этом случае детерминированные оценки принято сопровождать формальными дополнениями вида "временами, местами", и каждый потребитель волен трактовать эти скрытые вероятности по своему усмотрению. Не меньшие сложности это создает и при разработке методов прогнозов, и при их оценке. Очевидно также, что при неравнозначности потенциальных потерь от пропущенного опасного явления и ложного предупреждения методики вынужденно должны подстраиваться в сторону большей предупрежденности и мириться с низкой оправдываемостью прогнозов с явлениями. Для редких ситуаций и явлений асимметрия оценок еще больше, а статистическая обеспеченность прогностических связей еще меньше.

Именно в силу пространственно-временной изменчивости подсеточного масштаба заморозок, отмечаемый отдельными станциями, может отсутствовать на части территории и наоборот, неотмечаемый станциями, наблюдаться на отдельных участках. Поэтому для тех хозяйств, культур, которые не были затронуты текущим заморозком, прогноз последующих столь же важен.

Гололед - очень опасное метеорологическое явление, поскольку покрывает все поверхности на определенной территории, и защититься от этого невозможно, только частично снизить ущерб при быстром реагировании. К счастью, для Урало-Сибирского региона это весьма редкое явление (за исключением районов с наветренной стороны Уральских гор), требующее тонкого сочетания термодинамических условий в нижних слоях атмосферы:

- отрицательные температуры в приземном слое и на поверхности,
- жидкие (возможно переохлажденные) осадки.

Такие условия могут создаваться в зоне теплого фронта, над вершиной теплого сектора активного циклона и даже в теплом секторе с моросью. Для фиксации и прогнозирования гололеда как редкого явления также характерны описанные выше методические проблемы. Различие только в том, что чаще преобладает не столько пространственная, сколько временная изменчивость за счет скорости смещения синоптических объектов. Исключение составляют особые формы рельефа - перевалы, водоразделы на дорогах и наветренные склоны, которые усиливают и сохраняют условия для гололеда более длительное время, чем на равнинах. Столь существенное и

специфическое влияние орографии вместе с отсутствием детальных профилей температуры и влажности в пограничном слое сильно осложняют работу по формированию прогностических признаков, и это объективно сказывается на качестве статистического распознавания гололеда.

Гололедица - отдельное следствие определенных погодных условий во всем списке гололедно-изморозевых явлений. Фактически гололедица - явление антропогенного происхождения. В незатронутой человеком природе она не заметна, поскольку вода стекает, впитывается или испаряется быстрее, чем наступает похолодание и замерзание. Редкие лужицы могут оставаться и замерзать только в скальных углублениях или глубоких следах животных. А вот разновидностей гололедицы в городах может быть много и наблюдается она слишком часто для коммунальных служб, чтобы с ней могли быстро справляться.

Видимо, неопределенность классификации гололедицы из-за множества антропогенных факторов и невозможность их воспроизведения на метеоплощадке не позволила включить ее в список наблюдений и в код КН-01. Поэтому авторам пришлось разработать собственный вариант типизации гололедицы в виде комплексов метеоусловий, способствующих ее появлению и сохранению.

Подробнее об этом ниже в соответствующем разделе, а пока отметим принципиальные отличия двух физических процессов формирования гололедицы. Первый, самый простой и очевидный - это замерзание лужиц и просто мокрых поверхностей независимо от происхождения этой воды. Второй - это уплотнение свежеснежного покрова пешеходами и автомобилями. Поверхностное подтаивание все еще неубранного уплотненного снега с последующим замерзанием обладает чертами и первого, и второго процессов, а также имеет совсем прямое отношение к коммунальным службам.

1 СОСТОЯНИЕ ВОПРОСА

Различают внутримассовый и фронтальный гололед. Исследования метеорологических условий образования гололеда на Урале и юго-востоке Западной Сибири проводились в 60-70-х годах двадцатого столетия.

Были описаны необходимые условия для образования гололеда:

- переохлажденный дождь, морось, туман или низкая слоистая облачность;
- слои инверсии в пограничном слое атмосферы;
- сочетание температуры воздуха у земли (от 0,5 до -3⁰С) и на уровне АТ-850 (от 5 до -5⁰С);
- температура воздуха на нижней границе первого слоя инверсии от -0,5 до -8⁰С;
- суммарный дефицит точки росы в слое Земля - 850 гПа \leq 3-5⁰С;
- очаг теплого воздуха на АТ-850гПа.

Анализ такого сочетания условий при отсутствии надежных прогностических атмосферных полей в 60-х годах был весьма затруднителен. Тем не менее, разработанные в то время графические монограммы для прогноза гололеда и его интенсивности Р.А. Ягудиным [1] по сочетанию ожидаемой температуры воздуха у земли и высоты изотермы -10⁰С по фактической аэрологической диаграмме являются до сих пор руководством [2] для оперативных подразделений гидрометслужбы.

В предлагаемой разработке будет представлен первый в России опыт автоматизированной методики для прогноза гололеда.

Гораздо шире методическое обеспечение прогнозами заморозков.

Охлаждение приземного слоя воздуха или поверхности почвы до нуля и ниже актуально предупредить в вегетационном периоде. Как правило, процесс понижения температуры до отрицательных значений происходит в течение нескольких часов. Однако даже кратковременный заморозок способен погубить теплолюбивые культуры. Заморозки могут принести значительный ущерб как агропромышленному комплексу в целом, так и индивидуальным хозяйствам. Желание садоводов и огородников как можно раньше получить урожай сдвигает сроки посевов в период большого риска заморозков. Раннее предупреждение о наступлении заморозков позволит принять ряд защитных мер.

Заморозки могут возникать при разных синоптических процессах. В зависимости от условий образования их делят на радиационные, адвективные или радиационно-адвективные. Адвективные заморозки связаны с вторжением полярной или арктической воздушной массы на умеренные широты. Как правило, они более продолжительны по

времени (до нескольких суток) и охватывают относительно большую территорию. К радиационным заморозкам приводит постепенное охлаждение почвы и приземного слоя воздуха после захода солнца, то есть по окончании притока тепла. Необходимое условие при этом: наличие инверсии и слабое перемешивание воздуха в пограничном слое атмосферы (слабый ветер или штиль), что препятствует поступлению теплого воздуха из более высоких слоев. Такая ситуация происходит в случаях с высоким приземным давлением, то есть в антициклонах. На продолжительность и интенсивность заморозков существенное значение оказывают влажность воздуха (влияет на инерцию охлаждения), состояние поверхности почвы и рельеф местности (стекание и застой холодного воздуха даже в небольших низинах).

Оперативные прогностические подразделения гидрометслужбы обязаны в периоды активной вегетации растений или уборки урожая давать прогноз заморозков [3].

Агроклиматическим анализом и изучением синоптических условий образования заморозков активно занимались в 40-70х годах двадцатого века [4,5].

Параллельно разрабатывались прогностические правила для данного явления. Первая автоматизированная статистическая модель прогноза заморозков в Сибири разработана в 1983 году (автор Н.К. Шарифуллин) [6]. Она включала несколько поэтапных блоков. В модели использовались алгоритмы минимизации ошибки скользящего контроля, регрессии для получения количественных значений ожидаемых температур, линейный дискриминант Фишера для получения линейных альтернативных правил, метод последовательной минимизации числа ошибок для отбора лучших правил и кусочно-линейная модель разделения гиперплоскостей. Громоздкое число базовых параметров в условиях отсутствия автоматизированной электронной архивации ограничивало разработку решений по пунктам (только Новосибирск). А построение решений по исходным фактически наблюдаемым данным ограничивало заблаговременность (максимум одни сутки).

Первый метод прогноза заморозков для юго-востока Западной Сибири, базирующийся на выходных полях гидродинамической модели атмосферы ЕЦСПП, был разработан в 1994 году (автор И.Г.Храмцова). [7, 8] Заблаговременность прогнозов увеличена до пяти суток, территория расширена до 90 пунктов. На базе трех выходных модельных параметров: давление на уровне моря (P_0), геопотенциал на АТ-500 (H_{500}) и температура воздуха на АТ-850 (T_{850}) дополнительно рассчитывались горизонтальный градиент геопотенциала, лапласиан давления, средние значения H_{500} , P_0 , T_{850} на заданной долготе, разность широт, на которых планетарная высотная фронтальная зона

пересекает меридианы 50⁰ и 80⁰ и др. На фактическом материале в каждой точке определены средние значения всех предикторов в обеих группах x_k (k=1/2: заморозки/без). Решающее правило отнесения ожидаемой ситуации к той или иной группе построено на минимизации расстояния Махаланобиса:

$$R_k = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n \sigma_{x_i}^2}} \dots\dots\dots(1)$$

Окончательный прогноз формировался по административным областям Западно-Сибирского УГМС по трем составляющим: прогноз класса явления, прогноз минимальной температуры воздуха (Т_{мин}) (по разработанной ранее методике) и прогноз температуры поверхности почвы (Т_п) (по уравнениям линейной регрессии, связывающим Т_п с Т_{мин}).

Аналогичный подход был использован при разработке методов и технологии прогнозов заморозков по другим областям Урало-Сибирского региона [9-10].

На базе этой же модели ЕЦСПП был разработан объективный метод прогноза заморозков для юга Красноярского края (автор В.Г.Токарев) [11]. В методе определены пороговые показатели сходства значений каждого предиктора для класса заморозков (аналогично формуле (1)).

Все вышеперечисленные методики прогнозировали явление по административным областям, с указанием принятой в оперативной службе терминологии «местами» или на «большей части территории». Первая попытка разработать детализированный по конкретным пунктам прогноз заморозков была осуществлена в 2005 году для метеостанций Новосибирской области [12-13]. Схема также основана на интерпретации гидродинамических полей, но теперь уже добавлена модель из центра Экзетер с меньшим шагом (2,5⁰*2,5⁰ вместо 5⁰*5⁰) и дополнительным списком выходных параметров. В методе также учитывается прогноз минимальной температуры воздуха по усовершенствованной физико-статистической схеме (адаптация рабочей выборки с циклическим пополнением архивов, подбор аналогов для рабочей выборки, изменение списка предикторов согласно ожидаемым амплитудам изменений модельных температур и др.). Логические правила для альтернативных решений (с заморозками/без) построены по иерархическим ступеням с уменьшением рассматриваемого кластера. Сначала получается решение по области в целом, в случае положительного ответа продолжается расчет по секторам области и далее по станциям внутри секторов. Процесс направлен на

повышение устойчивости полученных решений. Здесь же был апробирован метод построения бинарных решений по алгоритму DW [14-16], позволяющий получить логические бинарные деревья для каждой станции.

Развитие гидродинамической продукции дает новые возможности для ее физико-статистической интерпретации. Расширяется список выходных параметров, уменьшается шаг в выходных полях, повышается качество прогнозов. Тем не менее, остается ряд задач, направленных либо на уточнение гидродинамических прогнозов, либо на получение прогнозов явлений, основанных на анализе и учете комплекса метеопараметров. Природные явления, возникающие в результате перехода температуры через ноль, а в теплом периоде это заморозки, в холодном – гололедные, также требуют дополнительных разработок для их прогнозирования.

Результатом данного исследования являются построенные методы и технология альтернативных прогнозов гололеда, заморозков в воздухе и заморозков на почве на 1-3 суток по территории Урало-Сибирского региона.

К решению задачи подключена выходная продукция глобальной модели GFS (NCEP) с разрешением $0,5^0*0,5^0$ и мезомасштабной модели COSMO-Ru_Sib13 с разрешением 13,2км. Результаты первой скачиваются с ftp-сервера NCEP , второй – получаются в ЗапСибРВЦ в оперативном режиме. Новые вычислительные возможности позволяют разрабатывать метод и технологию сразу по всей территории Урало-Сибирского региона, с вариантами детализации результатов прогноза: по 430 метеостанциям или по кластерам заданного радиуса.

2 ПОСТРОЕНИЕ ЛОГИЧЕСКИХ РЕШАЮЩИХ ПРАВИЛ НА БАЗЕ БИНАРНЫХ ДЕРЕВЬЕВ

2.1 Базовая методика распознавания образов

Для бинарного варианта прогноза есть смысл применить процедуру распознавания образов. В нашем случае в качестве базового использован алгоритм DW [14-16], опыт работы с которым имеется в нескольких разработках СибНИГМИ [17,18].

Алгоритм построен на последовательном делении исходной выборки, содержащей архивные данные двух классов (образов) - с прогнозируемыми явлениями (1-й образ) и без них (2-й образ). Доля случаев с явлениями в исходной выборке принимается за безусловную ("климатическую") оценку вероятности 1-го образа. Связь частот (вероятностей) с каким-либо синхронным признаком определяется делимостью условных распределений двух образов. Параметрические методы распознавания используют сравнения статистических оценок (средние, дисперсии). Алгоритм DW непараметрический и ищет условие разделения выборки среди всех значений вариационного ряда признака.

Отметим, что все авторские варианты алгоритма DW используют удобную бинарную форму - "матрицу сопряженности" (таблица 1), из которой в метеорологии получают различные оценки бинарных прогнозов (2.1-2.9). В алгоритме DW для каждого значения признака как порогового (сравнение на " \leq " или " $>$ " порога) для разделения выборки суммируются счетчики клеток матрицы сопряженности в логичном предположении, что полученная подвыборка с большей вероятностью 1-го образа содержит "прогнозы с явлением", а другая - "прогнозы без явления".

Таблица 1 – Таблица сопряженности прогноз-факт

Прогноз	Факт		
	Да	Нет	
Да	k11	k12	k10
Нет	k21	k22	k20
	k01	k02	k00

Оценки успешности для альтернативных прогнозов согласно таблице 1:

$$pr1=k11/k01 - \text{предупрежденность наличия явления} \quad (2.1)$$

$$pr2=k22/k02 - \text{предупрежденность отсутствия явления} \quad (2.2)$$

$$vr1=k11/k10 - \text{оправдываемость прогнозов наличия явления} \quad (2.3)$$

$$vr2=k11/k10 - \text{оправдываемость прогнозов отсутствия явления} \quad (2.4)$$

$$vr=(k11+ k22)/k00 - \text{общая оправдываемость прогнозов} \quad (2.5)$$

$$LT=k12/k10 - \text{доля ложных предупреждений (тревог)} \quad (2.6)$$

$$TSS=k11/k01-k12/k02 - \text{критерий Пирси-Обухова} \quad (2.7)$$

$$PRV=0,5(pr1+vr1)-0,2(pr1-vr1), \text{ если } pr1 \geq vr1$$

$$PRV=0,5(pr1+vr1)-0,4(vr1-pr1), \text{ если } pr1 < vr1 - \text{критерий баланса} \quad (2.8)$$

$$MPR= \min(pr1,pr2) - \text{критерий минимальной предупрежденности} \quad (2.9)$$

Осталось свести матрицу сопряженности к какому-то одночисловому критерию качества разделения для простой максимизации. Авторами ранее уже был выбран критерий Пирси-Обухова как приемлемый именно для сравнения разделенных подвыборок, поскольку эксперименты с различными метеорологическими данными показали хорошее взвешенное разнесение вероятностей при максимизации данного критерия.

Итак, на первом шаге алгоритм сначала находит для заданного признака лучшее по критерию Пирси-Обухова пороговое значение для разделения выборки на две со взвешенным разнесением исходной вероятности 1-го образа на большую ("прогнозы с явлением") и меньшую ("прогнозы без явления"). Полученный результат уже является готовой мини-методикой прогноза по единственному априори заданному признаку. Далее та же процедура повторяется для каждого проверяемого признака с одновременным отбором признака с максимальным значением критерия Пирси-Обухова.

В итоге на первом шаге алгоритма выбран лучший признак из списка с лучшим для него порогом разделения исходной выборки и получены две подвыборки с большей и меньшей вероятностями 1-го образа относительно исходной. Как показал наш опыт при разработке прогнозов гроз даже такой на вид простой результат может оказаться "лучшим" для очень редких явлений по одному пункту (1-2 случая за 4 года). Понятно, что обеспеченность такой статистики очень низка и нужно менять подход к формированию выборки.

Второй шаг алгоритма повторяет первый два раза - для каждой из двух полученных подвыборок. Последующие шаги выполняются аналогично с удвоением количества обрабатываемых подвыборок.

Таким образом, получаем бинарное дерево цепочек логических выводов вида "если ..то..иначе.." с различными вероятностями явления в каждом узле дерева. Деление каждой ветви может заканчиваться по исчерпанию одного из образов, но разумнее предусмотреть

критерий останова раньше. Надо хорошо понимать, что с уменьшением объема делимых подвыборок снижается статистическая надежность, устойчивость получаемых результатов вплоть до случайных совпадений.

В принципе, в таком виде "деревом решений" можно пользоваться для выдачи прогнозов в вероятностной форме, поскольку каждый конечный узел имеет вероятность явления, но чтобы перейти к бинарной формулировке, надо определить вероятностный порог для отнесения прогноза к 1-му или 2-му образу. Один из простых вариантов - сравнивать с начальной "климатической" вероятностью. А далее объединить все матрицы сопряженности конечных парных веток в одну согласно выбранному порогу. При этом вполне могут найтись лишние деления, поскольку они для этого порога дают совпадающие образы, хотя и с различной вероятностью. В любом случае, полученная суммарная матрица сопряженности даст оценку качества прогнозов на этой обучающей выборке с заданным вероятностным порогом отнесения любого сочетания используемых признаков к одному из образов.

Кажется очевидным, что оценивать суммарную матрицу сопряженности можно тем же критерием, что и при делении ветвей, но эксперименты с глубокими деревьями показали, что это не лучший вариант.

Существенная авторская модификация алгоритма предусматривает еще 2 шага алгоритма для оптимизации "глубины" дерева, т.е. числа делений каждой ветви и автоматического выбора пороговой вероятности. Для этого вместо произвольного задания глубины и вероятностного порога предложена следующая процедура.

Для каждой глубины дерева, начиная с максимальной, вычисляются матрицы сопряженности для всех вероятностей из конечных ветвей, используя их как пороговые. В результате максимизации критерия находим как лучшие пороги для каждого варианта глубины дерева, так и лучший вариант его глубины. Вот именно такое детальное сравнение суммарных матриц сопряженности и соответствующих им критериев натолкнуло на необходимость использования критериев, отличных по свойствам от критерия Пирси-Обухова. Первым был вариант с предложенным критерием PRV (баланс предупрежденности и прогностической вероятности явления) (2.8) [18,19] а позже - критерий минимальной предупрежденности двух образов MPR (2.9). Оба использовались для сравнения различных вариантов построения прогностических решений, выбора лучших для последующей селекции на независимой выборке и оценки оперативных испытаний методик прогнозов.

2.2 Методология получения комплекса решающих правил и выбора оптимального

Описанный выше модифицированный алгоритм DW на выходе получает субоптимальное (оптимальное потребовало бы перебора всех возможных сочетаний признаков, что неприемлемо) последовательное решение одновременно и для пороговых вероятностей, и для глубины дерева, формализуя субъективные правила останова. Решение единственно, что для математика - идеальный законченный вариант. А для целей прогноза - это плохо. Важно понимать - почему?

В метеорологии давно принято, как очевидное, методологическое правило: после получения статистического (обычно прогностического) решения на некоторой многолетней архивной выборке проверять его на другой, независимой. Это делается для оценки достоверности и устойчивости найденных статистических связей, которые могут быть и вовсе случайными. Оценки на независимой выборке ожидаемо снижаются и остается лишь оценить их приемлемость для оперативной работы. Но не стоит думать, что математика об этом "не знала".

Изящное обоснование проблем познания дала знаменитая теорема Гёделя о неполноте. Из нее, в частности, следует, что выводы, полученные в замкнутой системе не обязательно справедливы за ее пределами. А ведь ограниченная выборка данных - это и есть наша "замкнутая система". Конечно, и без Гёделя было понятно, что оценочные (полученные на ограниченной выборке) статистические связи - это не законы природы, и сила статистики резко слабеет при невыполнении многих оговоренных в ее основах условий: однородности, эргодичности, нормальности и прочего, что обеспечивает строгость ее выводов. Но с Гёделем комфортнее.

Можно понять математиков, которые развивают математический аппарат, оперируя модельными данными с заданными свойствами. Так, если выборка данных внутренне неоднородна, то ее надо оценить, описать математически, отфильтровать и прогнозировать отдельно, как, например, тренд временного ряда. А поскольку в метеорологии приходится оперировать с данными "природы", а не заданного эксперимента, то приходится решать дилемму: или подбирать/разрабатывать адекватный "грязным" данным алгоритм, или подготавливать особым образом данные для применимости выбранного алгоритма. Оба подхода имеют свои преимущества и недостатки, но ничто не запрещает их комплексировать для получения максимального результата.

Примером такого комплексного подхода служат алгоритмические исследования, связанные аббревиатурой МГУА (метод группового учета аргументов) [20] В обоснование предложенных подходов к обработке данных и прогнозированию положены два

принципа: "внешнего дополнения" (явный отклик на теорему Гёделя) и "свободы выбора" Габора. Первый принцип относится к формированию данных, а конкретно, к делению выборки на части с применением к каждой различных алгоритмов, а второй дает механизм "селекции, самоорганизации", позаимствованный из "дерева жизни". Вот такой любопытный алгоритмический комплекс.

Еще один, более близкий по времени и тематике, пример – моделирование ансамблей в оперативных гидродинамических моделях. Здесь подход несколько иной, но близок в главном - вариации входных данных имитируют различающиеся выборки ("внешнее дополнение"), а соответствующее им множество выходных решений позволяет выбрать наиболее вероятное ("свобода выбора", селекция).

Аналогичные подходы использованы нами для алгоритмического расширения методики построения логических решающих правил для прогнозирования.

Использовались 3 выборки многолетнего архива: 2014-2017гг, 2018г., 2019г. Первая, 4-х летняя - для построения базовых деревьев решений глубиной от 1 до 5 уровней для различной полноты списков (9вариантов) входных параметров. Для последующей селекции по критерию MPR из полученной матрицы (5*9) решений отбирались 9 лучших - матрица (3*3) Вторая - для селекции лучшего решения по критерию MPR из матрицы решений(3*3). Третья - для независимой оценки полученных решений.

Алгоритм получения редуцированных списков входных параметров состоял из последовательных шагов построения деревьев, начиная с полного начального списка - 35 прогностических модельных параметров. Для каждого последующего шага список сокращался исключением наименее часто попадавших в деревья параметров с учетом весов, обратно пропорциональных глубинам дерева, на которых признак включался базовым алгоритмом DW. Число одновременных сокращений числа параметров уменьшалось примерно пропорционально текущим размерам списка от 5-7 на первом шаге до 1 на последнем. Таким образом было получено 9 частотных списков входных параметров.

Алгоритм получения редуцированных деревьев различной глубины (сложности) построен на последовательном обходе узлов выходных деревьев, полученных базовым алгоритмом DW, по уровням глубины, начиная с максимального и пересчете суммарной матрицы сопряженности для редуцированного на единицу глубины дерева. Таким образом получено 5 вариантов деревьев решений различной глубины (сложности).

В результате получилась матрица (5*9) вариантов деревьев решений, из которых по критерию MPR отбирались (3*3) лучших и подавались на вход селекции на независимой выборке для получения единственного решения по максимуму MPR.

Описанная иерархическая структура алгоритма направлена на повышение устойчивости решений за счет исключения случайных или редких связей факторов с предиктантом и механизма сбалансированного критериального упрощения решений.

Кроме алгоритмических механизмов на различных этапах исследований и вычислительных экспериментов менялись подходы к формированию входных выборок предикторов и предиктанта. Это определялось как частотными различиями (климат) предиктантов (гололед оказался самым редким из исследуемых явлений), так и надежностью получаемых связей. Поэтому деревья решений строились и по отдельным станциям, и по кластерам с радиусом 200км, и по территории региона в целом. Различные подходы к формированию адекватной особенностям предиктанта выборки зачастую дают гораздо больше, чем самые навороченные алгоритмические ухищрения. Что и подтверждалось всегда в наших исследованиях.

2.3 Подготовка базовых данных

В нашем случае в качестве признаков выступают выходные поля гидродинамических моделей (таблица 2) с целью получения адаптивных решений для каждой заблаговременности прогнозов. Наиболее наполненные количеством параметров на выходе модели COSMO и GFS. Первая имеет более высокое разрешение (13.2 км на сегодняшний день), вторая, несмотря на выдачу данных по сетке в 0.5*0.5 градусов по широте-долготе, показывает более качественные результаты. Значения параметров выбираются в ближайшем к искомой метеостанции узле сетки. Набор признаков определяется физической или потенциально возможной связью с заданным предиктантом (явлением погоды). Для наших задач отобрано 39 потенциальных предикторов (таблица 3), часть из них непосредственно берется из модельных выпусков, часть является расчетными (температура смоченного термометра, градиенты, часовые изменения). Исходный список включает температуру в верхнем слое почвы, на поверхности и на уровнях 2м, АТ-925; характеристики влажности на уровнях до АТ-850, ожидаемое изменение давления на уровне моря за 6 часов, вертикальные градиенты геопотенциала и его временные изменения, ветер у земли, осадки

Таблица 2 – Список используемых выходных параметров моделей GFS и COSMO

№	Параметр	Уровень	Описание параметра
1	P0	средний уровень моря	Давление [гПа]
2	HGT500	500 гПа	Геопотенциал [Дам]
3	HGT925	925 гПа	Геопотенциал [Дам]
4	T850	850 гПа	Температура [С]
5	T925	925 гПа	Температура [С]
6	T950	950 гПа	Температура [С]
7	T0	2 м	Температура [С]
8	Ts	0 м	Температура поверхности почвы[С]
9	RH925	925 гПа	Относительная влажность [%]
10	RH2m	2 м	Относительная влажность [%]
11	U0	10 м	U-компонента ветра [м/с]
12	V0	10 м	V-компонента ветра [м/с]

Таблица 3 - Параметры – предикторы для построения DW решений

№	Параметр	Описание параметра
1	Tsm850	температура смоченного термометра на уровне АТ-850 (в срок)
2	2Tsm850	температура смоченн. термометра на уровне АТ-850 (в срок+6ч)
3	T9-T0	вертикальный градиент температуры, АТ-925-2м (в срок)
4	2T9-T0	вертикальный градиент температуры, АТ-925-2м (в срок+6)
5	T8-T0	вертикальный градиент температуры, АТ-850-2м (в срок)
6	2T8-T0	вертикальный градиент температуры, АТ-850-2м (в срок+6)
7	RH0	относительная влажность на высоте 2м (в срок)
8	2RH0	относительная влажность на высоте 2м (в срок+6ч)
9	RH9	относительная влажность на уровне АТ-925 (в срок)
10	2RH9	относительная влажность на уровне АТ-925 (в срок+6ч)
11	UV0	скаляр вектора скорости на высоте 10м (в срок)
12	2UV0	скаляр вектора скорости на высоте 10м (в срок+6ч)
13	Tsm0	температура смоченного термометра на высоте 2м (в срок)
14	2Tsm0	температура смоченного термометра на высоте 2м (в срок+6ч)
15	dTd0	дефицит точки росы на высоте 2м (в срок)
16	2dTd0	дефицит точки росы на высоте 2м (в срок+6ч)
17	dTd8	дефицит точки росы на уровне АТ-850 (в срок)
18	2dTd8	дефицит точки росы на уровне АТ-850 (в срок+6ч)
19	T(0)-T(-6)	тенденция температуры на высоте 2м (срок - (срок-6ч))
20	T(+6)-T(0)	тенденция температуры на высоте 2м ((срок+6ч) - срок)

№	Параметр	Описание параметра
21	T(+6)-T(-6)	тенденция температуры на высоте 2м ((срок+6ч) - (срок-6ч))
22	P(0)-P(-6)	тенденция давления, средний уровень моря (срок - (срок-6ч))
23	P(+6)-P(0)	тенденция давления, средний уровень моря ((срок+6ч) – срок)
24	H500-H925	вертикальный градиент геопотенциала, АТ500-АТ925 (в срок)
25	2H500-H925	вертикальный градиент геопотенциала, АТ500-АТ925 (в срок+6ч)
26	dH59(+6)-dH59(0)	тенденция вертикального градиента геопотенциала, АТ500-АТ925 ((срок+6ч) - (срок-6ч))
27	AP0n	осадки (в срок-6ч)
28	AP00	осадки (в срок)
29	AP+6	осадки (в срок+6ч)
30	AP0+00+6	сумма осадков за 12ч (в срок + (срок+6ч))
31	T2m	минимальная температура на высоте 2м (ночь)
32	T925	температура на уровне АТ-925 (в срок)
33	T850	температура на уровне АТ-850 (в срок)
34	Ts	температура поверхности почвы (в срок)
35	Ts(+6)-Ts(0)	тенденция температуры поверхности почвы ((срок+6ч) - срок)
36	min(T0+5)*10	смещенная минимальная температура на 2м (ночь)
37	min(T8-5)*10	смещенная минимальная температура на АТ-850 (ночь)
38	abs(T0+5)*10	модуль смещенной минимальной температуры на 2м (ночь)
39	abs(T8-5)*10	модуль смещенной минимальной температуры на АТ-850 (ночь)
<p>Примечание – Параметры: T – температура; Td – температура точки росы; Tsm - температура смоченного термометра; Ts - температура поверхности почвы; P – давление, AP – осадки, dTd - дефицит точки росы. Уровни: 0м, 2м, 950 гПа, 925 гПа, 850 гПа, 500 гПа</p>		

3 ПРОГНОЗ ПЕРЕХОДА ТЕМПЕРАТУРЫ ЧЕРЕЗ НОЛЬ В ВЕСЕННЕ-ОСЕННИЕ ПЕРИОДЫ И СВЯЗАННЫХ С НИМ ЗАМОРОЗКОВ

Переходы температуры через ноль в теплом периоде приводят к заморозкам и поверхность почвы, при этом заморозки могут произойти в приземном слое воздуха (заморозки в воздухе), на почве (заморозки на почве), либо в воздухе и на почве.

Факт перехода температуры от плюса к минусу определен по наблюдаемому суточному ходу температуры из КН01 с апреля по октябрь 2014-2019 гг. Случаи понижения температуры до нуля и ниже в ночные сроки после положительных температур днем отнесены к первому классу, остальные ко второму. Для заморозков вегетативный период определен в среднем для всей территории: с мая по сентябрь. Если переход через ноль отмечен по температуре воздуха на уровне 2 метров, то это определение для заморозков в воздухе; если рассматривается температура поверхности почвы – то для заморозков на почве. Как правило, наблюдаемые заморозки на почве сопровождаются заморозками и в воздухе. Но бывают исключения, когда температура воздуха близка к нулю, но не достигает его. Для исследования создан архив альтернативных заморозков в воздухе и на почве за теплый период 2014-2019 годов по метеостанциям Урало-Сибирского региона, включая республику Саха. В архиве 432-метеостанции для заморозков в воздухе и 369 – на почве (не везде измеряется температура поверхности почвы). При этом 4 года (2014-2017) использованы для рабочей выборки, 2018-для отбора решений и 2019 – для независимых оценок.

По алгоритму, описанному в разделе 2 настоящего отчета, построены бинарные деревья логических решений отдельно для прогнозов заморозков в воздухе и заморозков на почве. Деревья построены для каждой станции, заблаговременности (00+18, 00+42, 00+66) по каждой модели (GFS, COSMO).

В следующих разделах будет рассмотрена успешность как новых методических прогнозов, так и непосредственно гидродинамических прогнозов по базовым моделям.

3.1 Оценка прогнозов перехода температуры через ноль на уровне двух метров по модельным выходным параметрам

На первом этапе произведена оценка прогнозов перехода температуры через ноль по выходным модельным данным в теплом периоде. Период оценки: апрель 2014 - октябрь 2017 года, исключая июль, всего порядка 840 случаев для каждой из 432 станции.

Модельные данные (GFS, COSMO): прогноз температуры воздуха на уровне 2м и температуры поверхности земли.

Данные для метеостанций взяты по значению в ближайшем узле. Заблаговременности: 00+24, 00+48, 00+72. Класс явления определен аналогично факту на наблюдениях.

Получены матрицы сопряженности прогноз-факт для каждой заблаговременности. Представленный на рисунках 1-3 комплексный показатель успешности прогнозов - критерий Пирси-Обухова демонстрирует разделение рассматриваемой территории по качеству модельных прогнозов. Прогнозы переходов через ноль на западе и юге рассматриваемой территории несомненно успешнее по модели COSMO. На территории Иркутской области и Забайкалья оценки примерно равны, а в Якутии и по северу Красноярского края преимущество за моделью GFS. Особо заметная разница в пользу модели COSMO на территории Кемеровской, Новосибирской, Омской областей, Алтайского края.

Следующий численный эксперимент направлен на выяснение влияния учета дополнительного выходного модельного параметра, а именно температуры поверхности почвы при прогнозе перехода температуры через ноль в воздухе.

Для этого проведено два варианта расчетов:

- только по модельной температуре на уровне 2м;
- с подключением температуры поверхности почвы по алгоритму: либо один, либо второй параметр переходит через ноль.

По всем станциям и по всей выборке за 2014-2017 годы получено порядка 36000 прогнозов на каждые сутки по модели GFS. Из данных таблицы 1 видно, что подключение второго параметра увеличивает число ложных тревог ($k12$), однако в среднем оправдываемость прогнозов с наличием явления ($vg1$) уменьшается лишь на 2%, тогда как их предупреденность ($pr1$) возрастает на 5-6% (таблица 4). На общую оправдываемость этот результат не сказывается, однако комплексные показатели качества (PRV, TSS) немного возрастают. Подтверждают положительный вклад модельной температуры поверхности почвы в успешность прогнозов переходов через ноль в воздухе диаграммы распределения значений критерия Пирси-Обухова по градациям, приведенные на рисунке 4. Видно, что добавление параметра улучшает оценки на все рассматриваемые сроки прогноза. Динамика качества прогнозов показывает незначительные изменения практически по всем показателям, можно лишь говорить об уменьшении на 0,1 PRV и TSS ко вторым суткам.

Таким образом, комплексная оценка перехода температуры через ноль (на высоте 2 метров и на поверхности почвы) демонстрирует лучшие результаты.

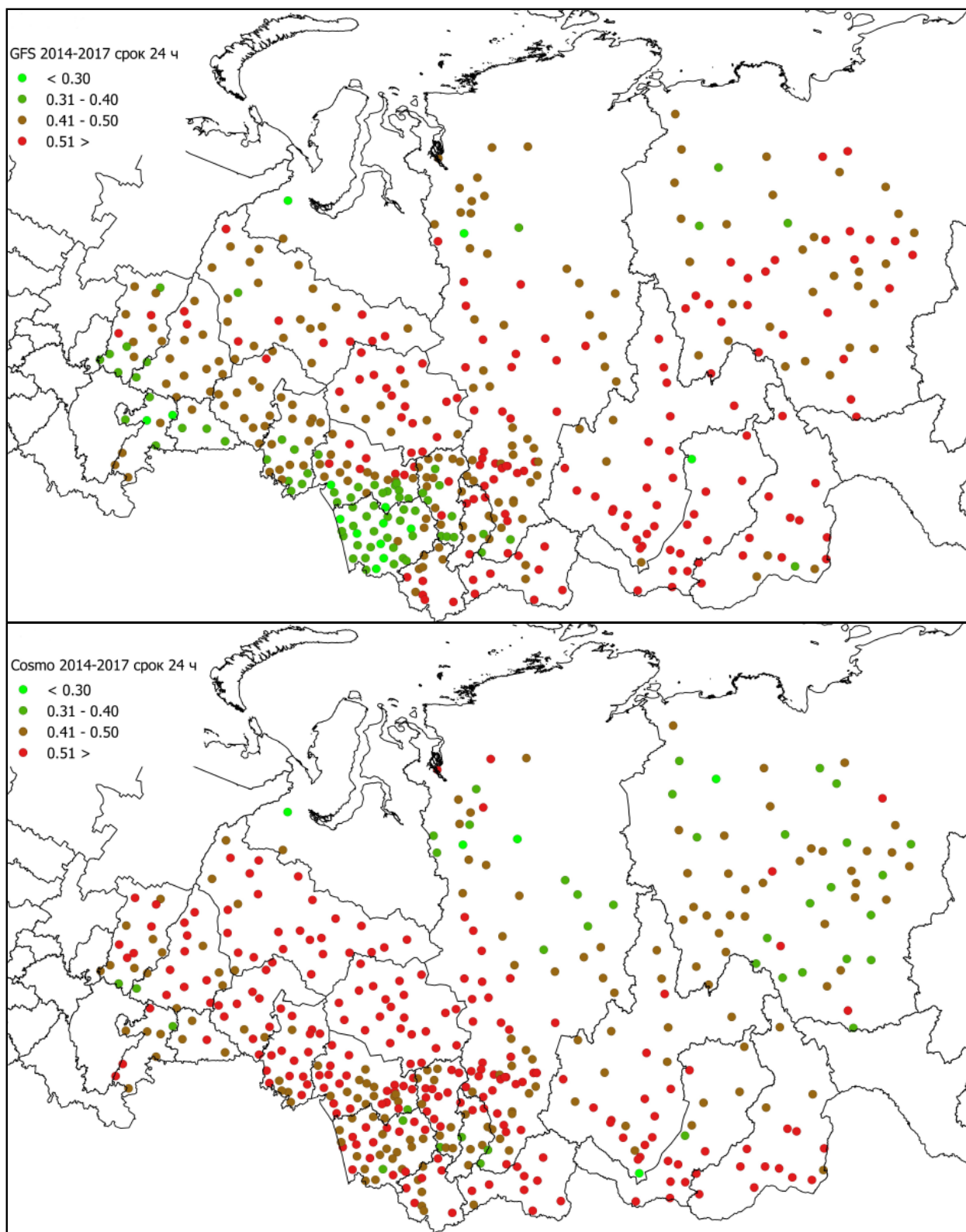


Рисунок 1 – Значения критерия Пирси-Обухова для прогноза перехода приземной температуры воздуха через ноль в теплом периоде по моделям GFS(вверху), COSMO(внизу). 2014-17гг. Заблаговременность 00+24

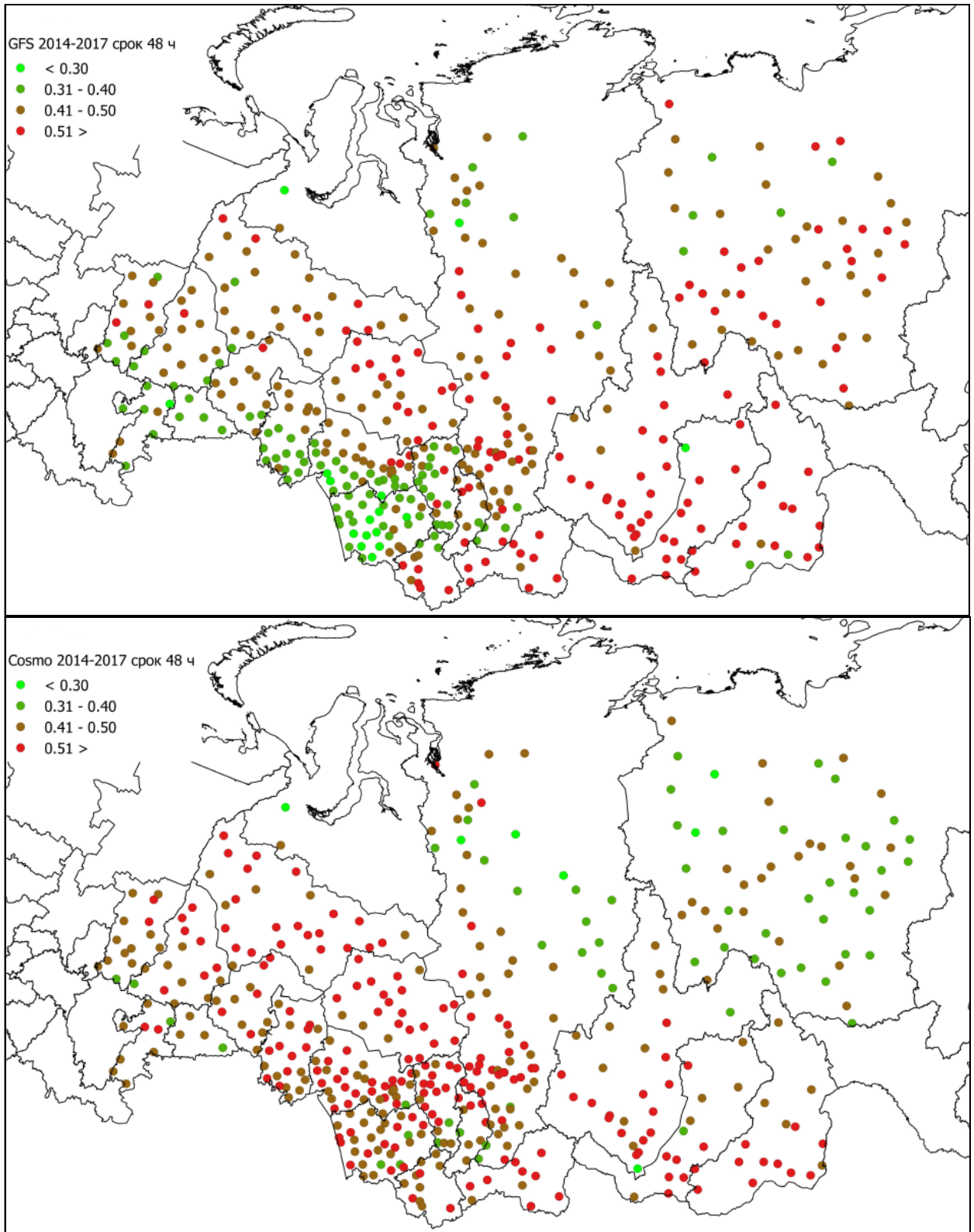


Рисунок 2 – Значения критерия Пирси-Обухова для прогноза перехода приземной температуры воздуха через ноль в теплом периоде по моделям GFS(вверху), COSMO(внизу). 2014-17гг. Заблаговременность 00+48

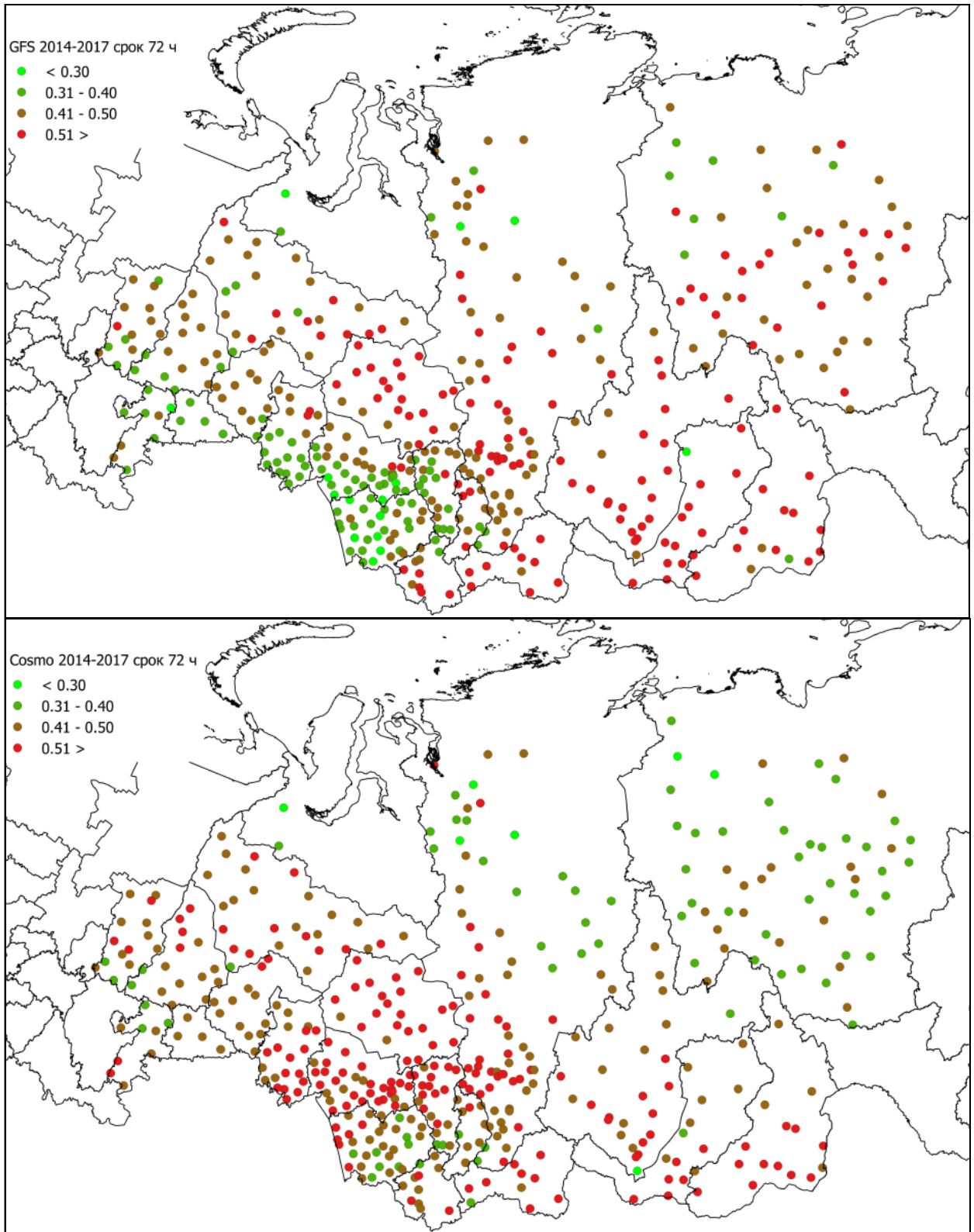


Рисунок 3 – Значения критерия Пирси-Обухова для прогноза перехода приземной температуры воздуха через ноль в теплом периоде по моделям GFS(вверху) COSMO(внизу). 2014-17гг. Заблаговременность 00+72

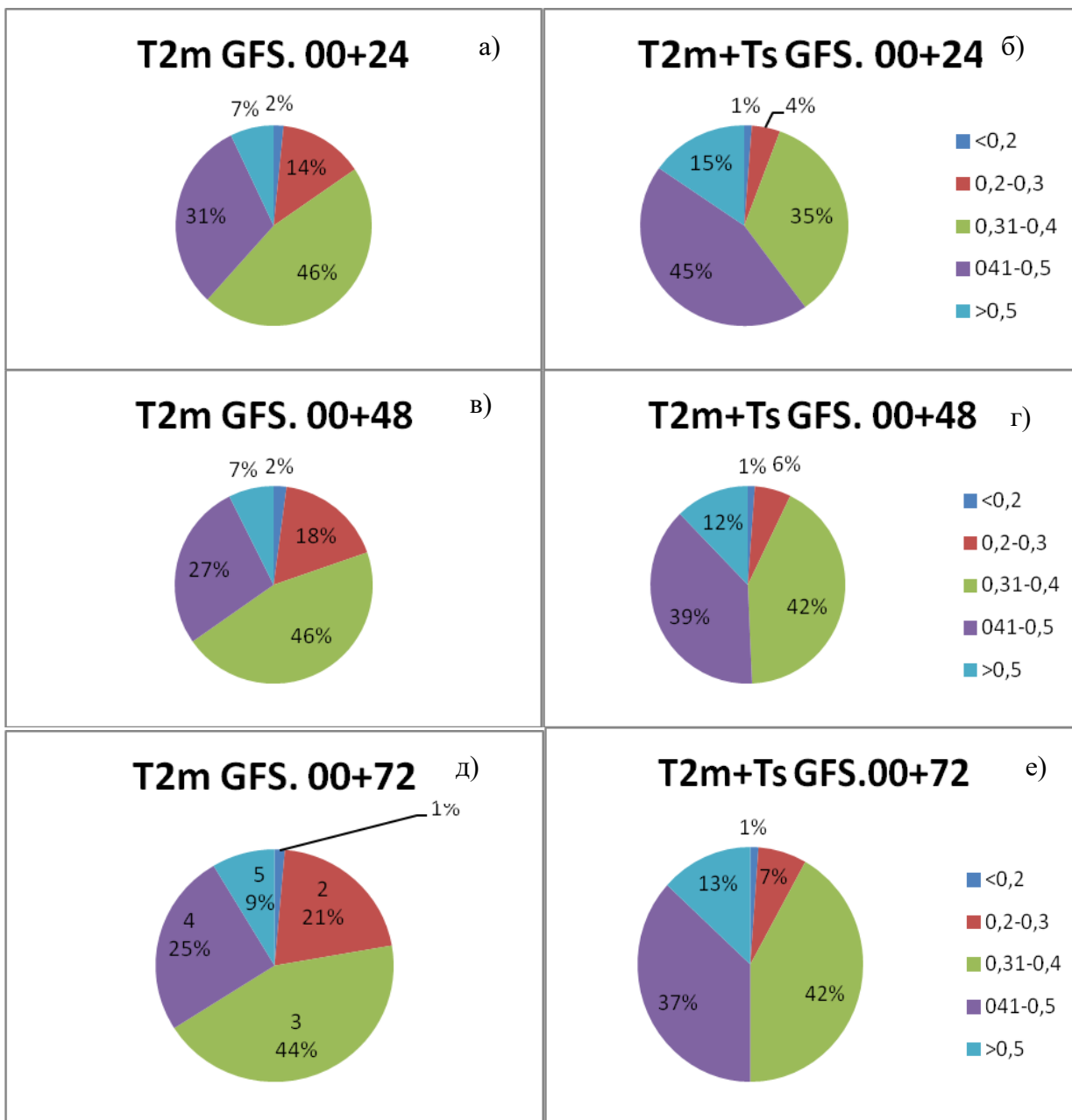


Рисунок 4 – Распределение по территории (процент станций в градации TSS) значений критерия Пирси-Обухова при оценке прогнозов перехода температуры через ноль по модели GFS на базе только температуры воздуха (рисунки а, в, д - слева) и с подключением температуры поверхности почвы (рисунки б, г, е – справа). Теплый период 2014-2017 годы, 432 метеостанции.

Таблица 4 - Средние оценки прогнозов перехода температуры через ноль по выходным параметрам модели GFS. Март-октябрь 2014-17 гг. 432 станции.

	k11	k12	k21	k22	k00	vr1	pr1	vr	PRV	TSS
00+24h										
T2m	36256	21910	40821	263460	362447	62	47	83	0.486	0.394
T2m+Ts	40546	27290	36531	258080	362447	60	53	82	0.533	0.430
00+48h										
T2m	34537	21590	40634	263093	359854	62	46	83	0.475	0.384
T2m+Ts	38592	26954	36579	257729	359854	59	51	82	0.521	0.419
00+72h										
T2m	34970	22182	40846	264448	362446	61	46	83	0.476	0.384
T2m+Ts	39209	27504	36607	259126	362446	59	52	82	0.524	0.421

3.2 Сравнительные оценки прогнозов заморозков в воздухе на независимой выборке

Фактические и модельные данные весенне-осеннего сезона апрель-сентябрь 2019 годов (около 305 случаев) использованы в качестве независимой выборки для сравнительных оценок по разным вариантам и подходам.

В первую очередь, отобранные на первом этапе 9 (3*3) лучших на обучающей выборке вариантов решений различных по глубине деревьев и списков предикторов (см. раздел 2) использованы для селекции единственного лучшего решения по независимой выборке 2018 года. Таким образом, имеем итоговые решения для каждой станции и для каждых 1-3 суток: 00+18-первая ночь, 00+42- вторая ночь, 00+66-третья ночь.

Как изменяется качество прогнозов на предварительно полученных вариантах решений по сравнению с окончательным продемонстрировано на рисунке 5. Максимизация показателя MPR (2.9) явилась хорошим критерием к отбору оптимальных решений (в нашем случае res1, res2, ... res9 - сжатие по числу веток и числу параметров). После применения этого алгоритма на решения по каждой станции на каждую заблаговременность получаем для всей территории результат восстановления на независимой выборке, названный best. На этом же графике приведены оценки MPR для

исходных модельных прогнозов температуры с учетом прогноза температуры на поверхности почвы. Даже в этом случае прогноз перехода температуры через ноль по модели значительно уступает результатам по любому из построенных логических решений. На рисунке 6 видно, что этот вывод относится практически ко всем станциям на рассматриваемой территории, за исключением шести пунктов на северо-востоке Красноярского края, где самые неудачные прогнозы как по моделям, так и по статистическому распознаванию.

Сравнительные оценки на независимом периоде методических прогнозов по бинарным логическим деревьям, построенным по выходной продукции разных моделей, показывают преимущество DW-решений на базе COSMO. Из таблицы 5 видно, что в среднем по территории за весь период предупреденность явления по модели COSMO на 5-7% выше, при этом ложных тревог больше лишь на 3-4%, что сказывается на более высоких оценках критерия Пирси-Обухова. Однако из-за небольших преимуществ в вероятности явления, другой критерий качества – PRV, напротив, выше на 0,1 у GFS. Распределение оценок по территории, приведенное на рисунках 7-8, показывает, что на большинстве станций уровень успешности сопоставимый, а если и есть разница, то она в пользу решений на базе COSMO. Этот вывод сохраняется на все сроки.

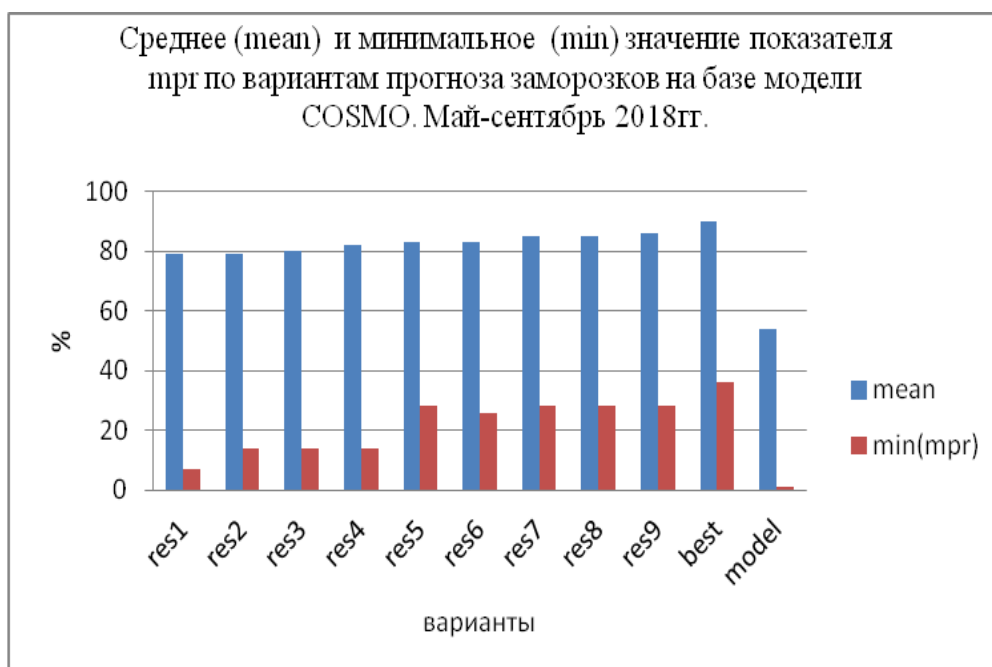


Рисунок 5 - Средний (mean) и минимальный(min) показатель mpr по территории региона на выборке 2018 г. по модели COSMO без интерпретации (model), по предварительным (1-9) и окончательному (best) DW-решениям.

Таблица 5 - Средние по территории оценки прогнозов заморозков в воздухе по разработанному методу. Теплый период 2019гг

00+h	k11	k12	k21	k22	k00	vr1	vr2	vr	pr1	pr2	TSS	PRV
Базовая модель COSMO												
18	12930	9257	1505	105468	129160	58	99	92	90	92	0.815	0.677
42	12566	9656	1646	105293	129161	57	98	91	88	92	0.800	0.661
66	12227	10719	1820	104395	129161	53	98	90	87	91	0.777	0.634
Базовая модель GFS												
18	12057	7253	2454	110420	132184	62	98	93	83	94	0.769	0.686
42	11858	7757	2442	110128	132185	60	98	92	83	93	0.763	0.672
66	11497	8690	2633	108933	131753	57	98	91	81	93	0.740	0.643

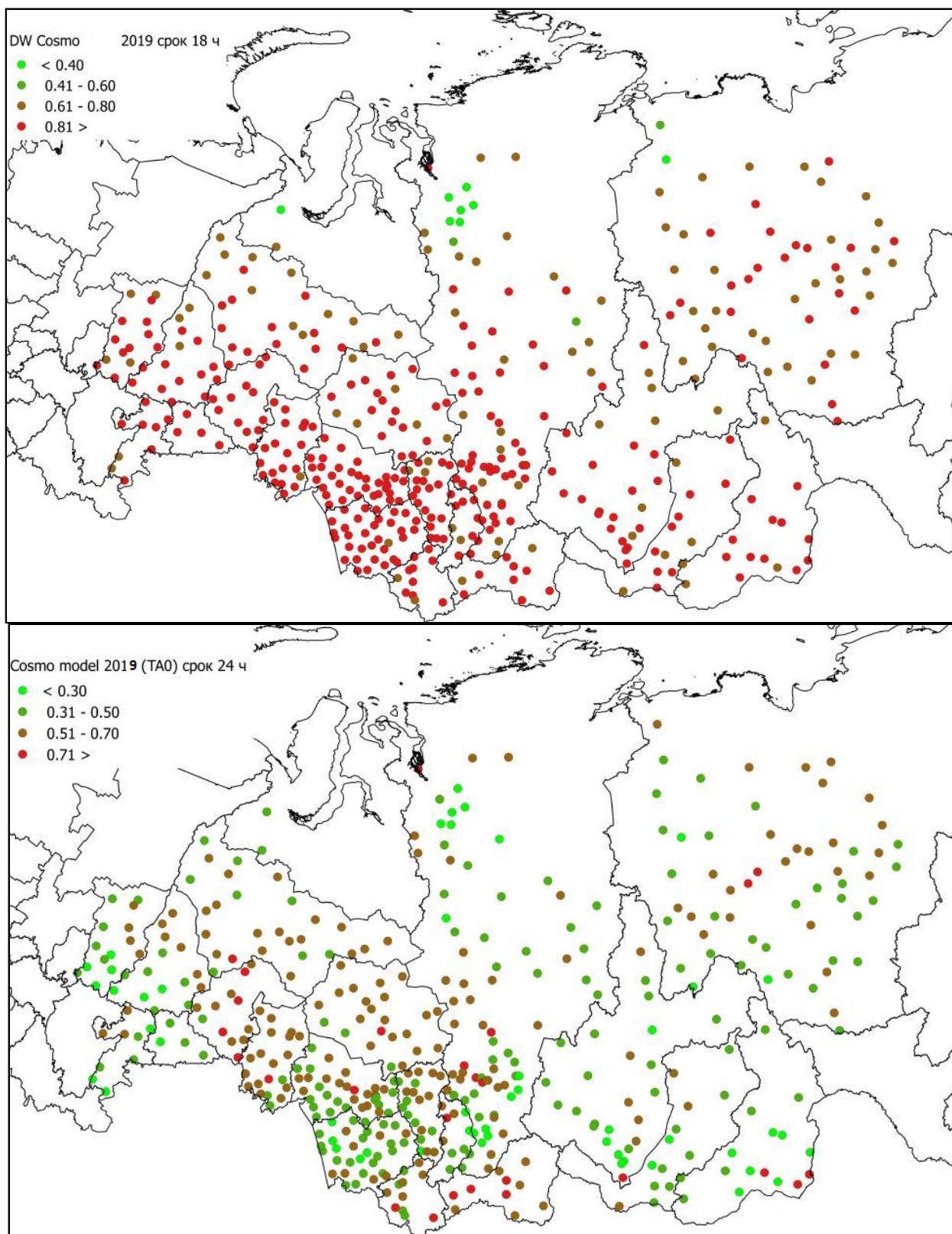


Рисунок 6 - Критерий Пирси-Обухова для альтернативных прогнозов заморозков по DW-решениям (вверху) и по модельной температуре без интерпретации (внизу) на первые сутки. Май-сентябрь 2019г.

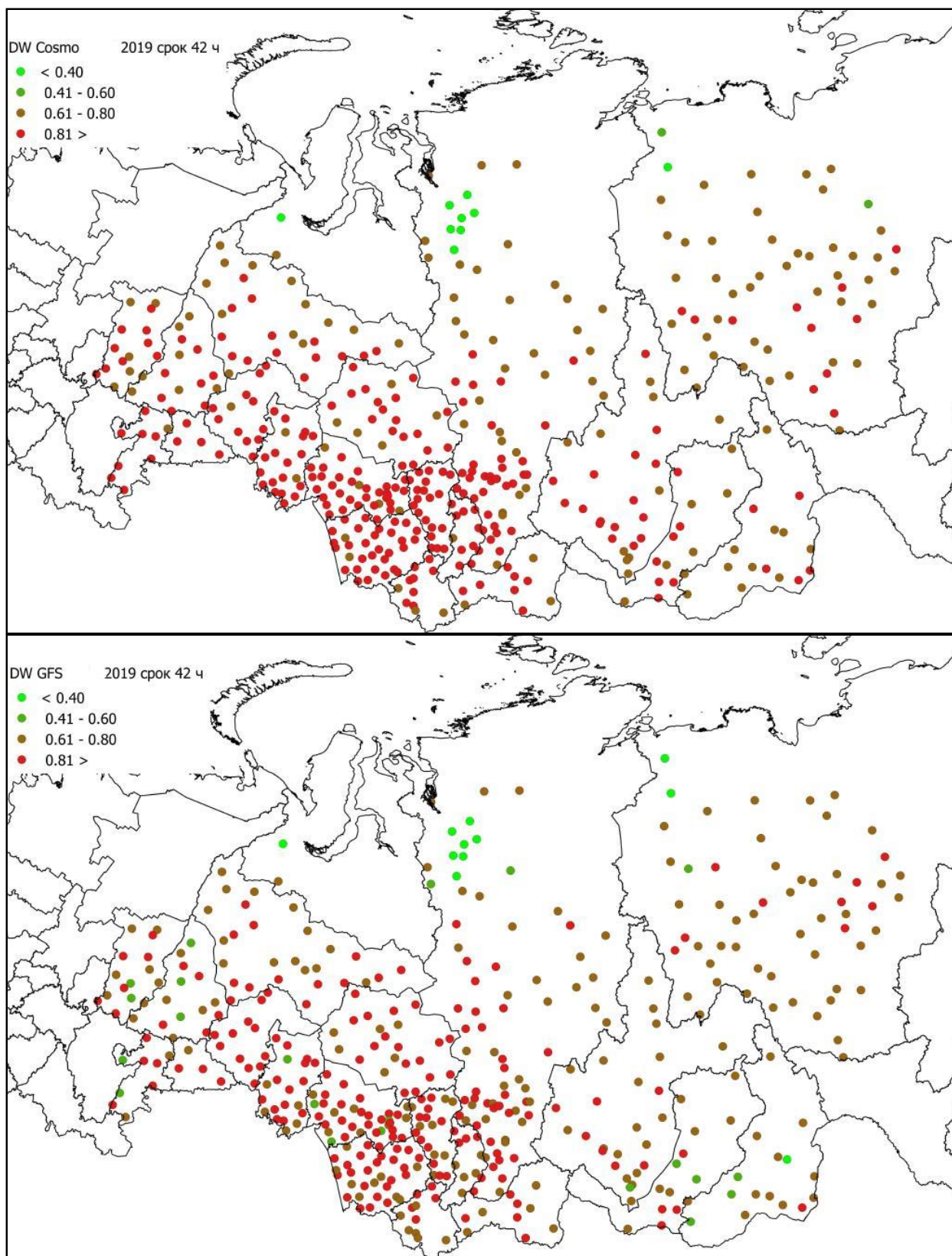


Рисунок 7 - Значение критерия Пирси-Обухова прогнозов заморозков в воздухе по бинарным деревьям, построенным на базе модели COSMO(вверху) и GFS (внизу) на вторые сутки. Май-сентябрь 2019г.

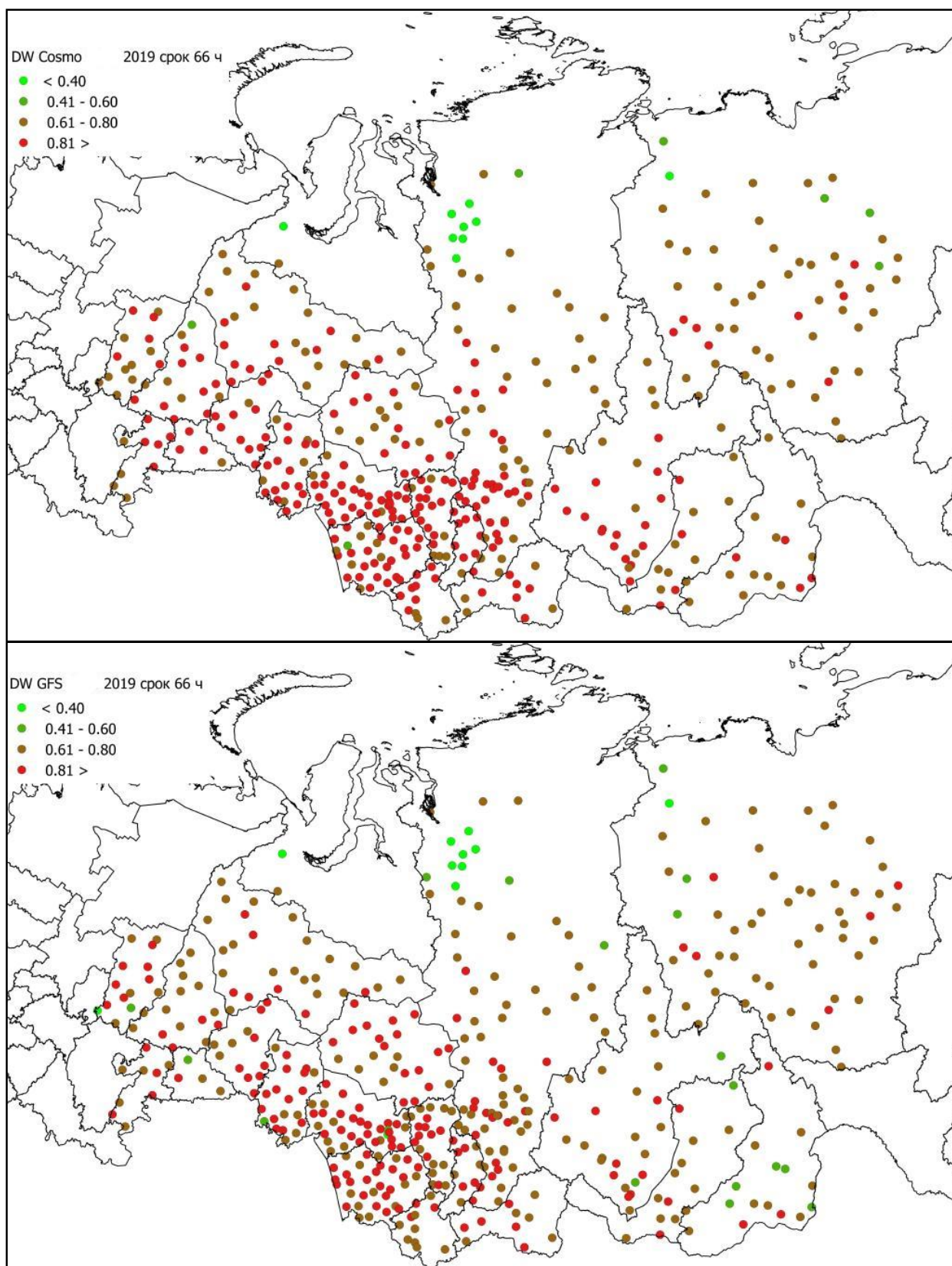


Рисунок 8 - Значение критерия Пирси-Обухова прогнозов заморозков в воздухе по бинарным деревьям, построенным на базе модели COSMO(вверху) и GFS (внизу) на третьи сутки. Май-сентябрь 2019г.

3.3 Сравнительные оценки прогнозов заморозков на почве

Аналогично решениям для заморозков в воздухе выбран лучший вариант деревьев для каждой станции на каждый срок заблаговременности по каждой модели с подключением подвыборки 2018 года.

Результаты проведения расчетов на независимой выборке 2019года показали следующее:

- в отличие от заморозков в воздухе общий уровень прогнозов более низкий за счет увеличения числа «ложных тревог»: примерно в два раза снижается оправдываемость прогнозов наличия заморозков. При этом средняя предупрежденность уменьшается лишь на 3-5% (таблица 6);

- практически все средние показатели уровня успешности не имеют разницы между решениями по моделям COSMO и GFS;

- появились различия в распределении ошибок по территории (рисунки 9 - 11)

Таблица 6 - Средние по территории оценки прогнозов заморозков на почве по разработанному методу. Теплый период 2019гг.

00+h	k11	k12	k21	k22	k00	vr1	vr2	vr	pr1	pr2	TSS	PRV
COSMO												
18	3224	9273	471	93102	106070	26	99	91	87	91	0.782	0.442
42	3069	9153	566	93278	106066	25	99	91	84	91	0.755	0.429
66	2996	10024	568	92475	106063	23	99	90	84	90	0.743	0.413
GFS												
18	3096	8744	628	96437	108905	26	99	91	83	92	0.748	0.432
42	2989	8500	678	96738	108905	26	99	92	82	92	0.734	0.427
66	2922	9349	676	95592	108539	24	99	91	81	91	0.723	0.410

Видно, что успешность прогнозов заметно снижается в районах с гористой местностью, на востоке (Забайкалье, Бурятия, Иркутская область, Республика Саха) и на севере Красноярского края. На западе территории, прилегающей к Уралу прогнозы на базе GFS в среднем точнее, но только на первые сутки, в последующие разница нивелируется. Пример оценок для Новосибирской области дан в таблице 7: оценки примерно равны, лишь предупрежденность второго класса (отсутствия явления) на 2-4% выше у GFS.

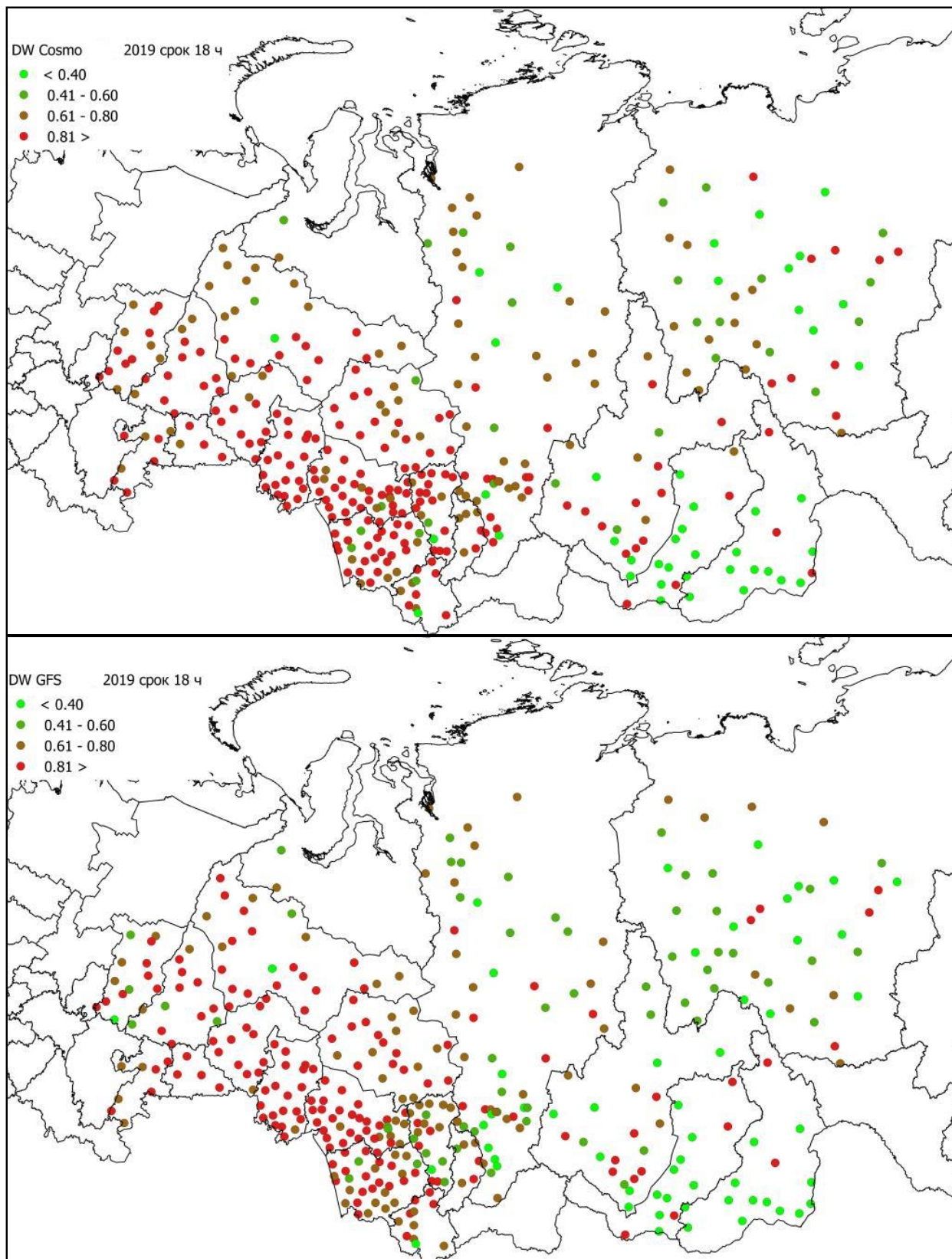


Рисунок 9 - Значение критерия Пирси-Обухова прогнозов заморозков на почве по бинарным деревьям, построенным на базе модели COSMO(вверху) и GFS (внизу) на первые сутки. Май-сентябрь 2019г

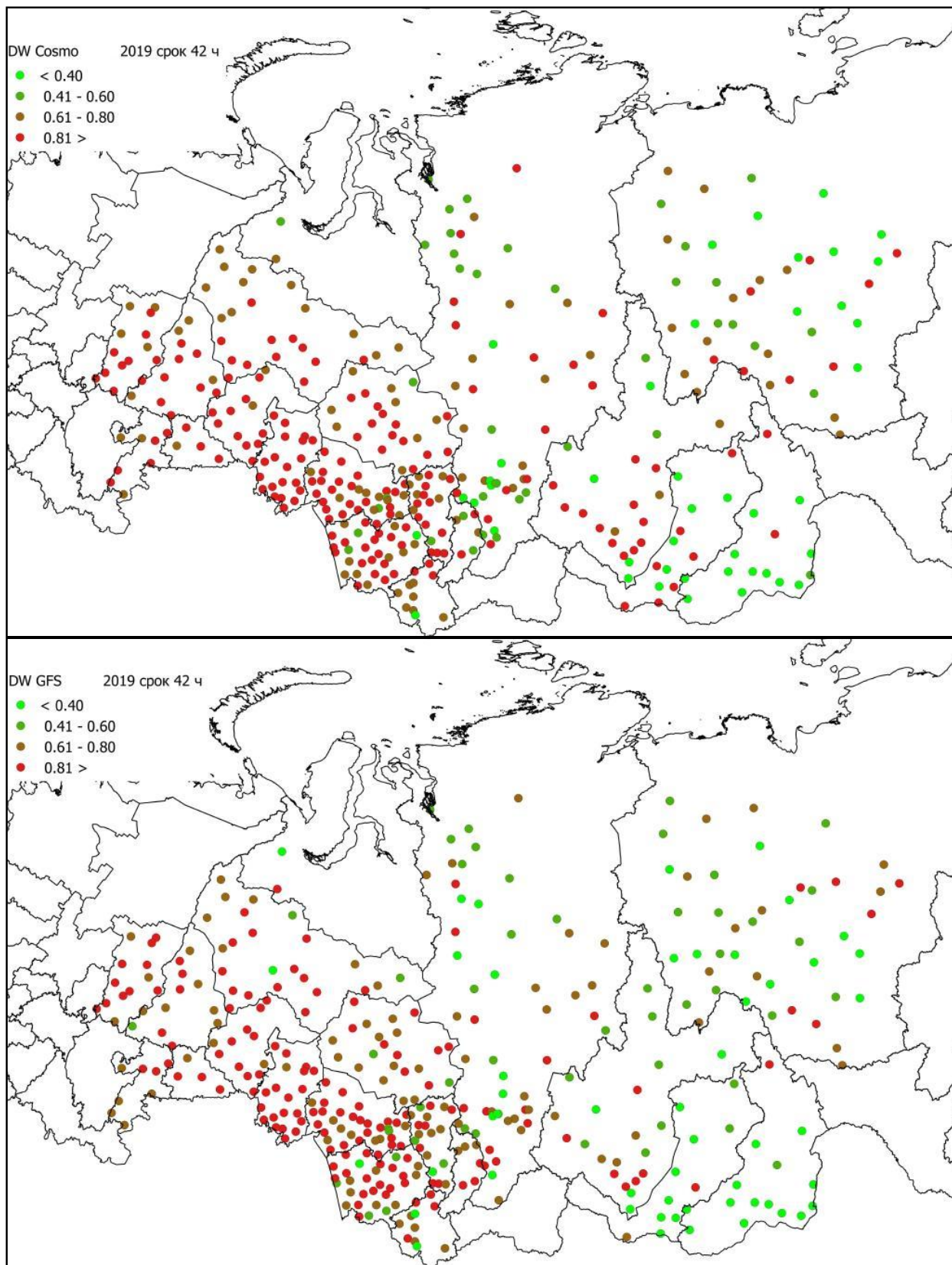


Рисунок 10 - Значение критерия Пирси-Обухова прогнозов заморозков на почве по бинарным деревьям, построенным на базе модели COSMO(вверху) и GFS (внизу) на вторые сутки. Май-сентябрь 2019г

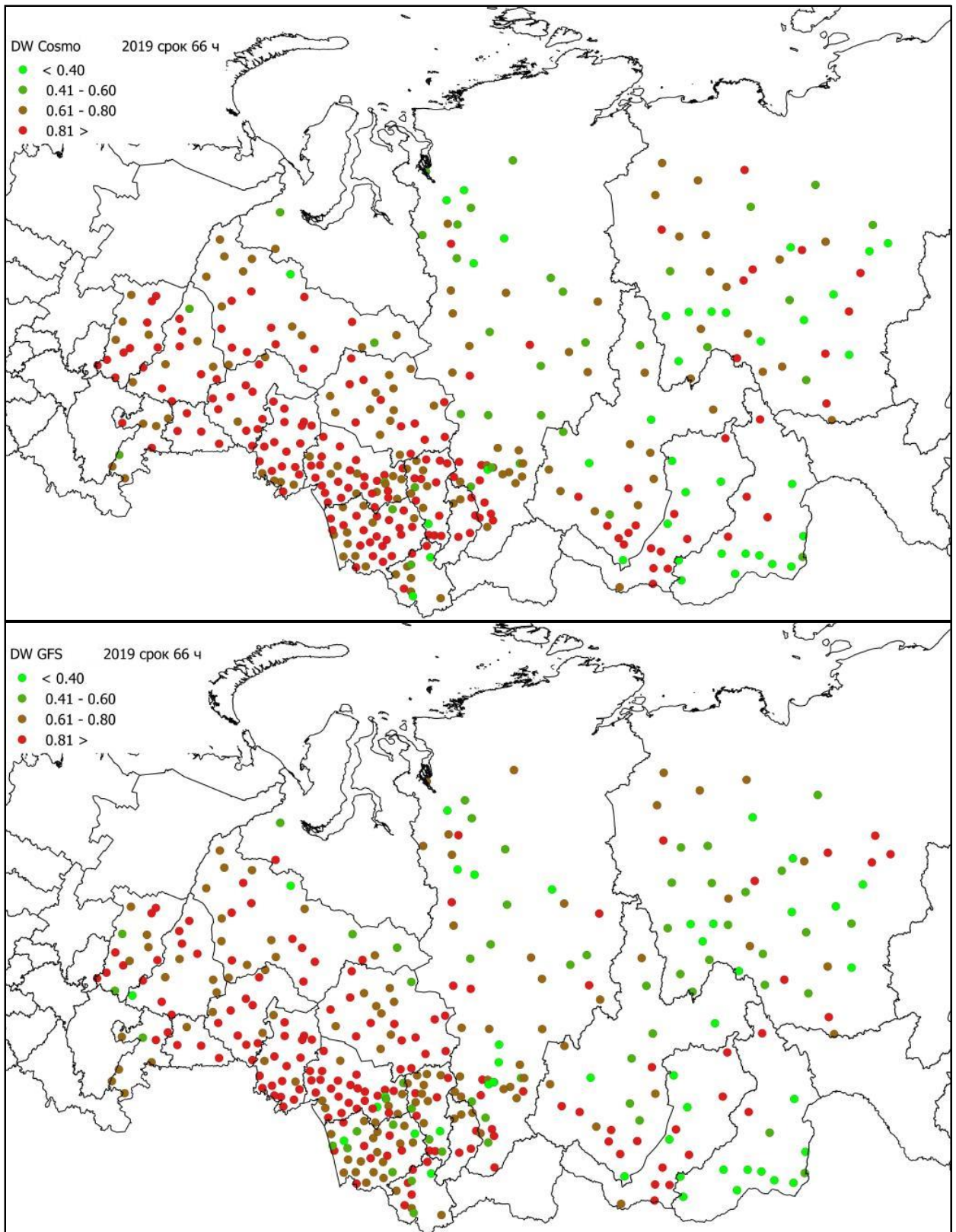


Рисунок 11 - Значение критерия Пирси-Обухова прогнозов заморозков на почве по бинарным деревьям, построенным на базе модели COSMO(вверху) и GFS (внизу) на третьи сутки. Май-сентябрь 2019г

Таблица 7 – Матрица сопряженности и средние оценки прогнозов заморозков на почве для территории Новосибирской области. Май-сентябрь 2019г. Алгоритм распознавания образов. Базовые модели COSMO и GFS

Модель	K11	K12	K21	K22	vr1	vr2	vr	pr1	pr2	TSS
Заблаговременность +18										
GFS	305	458	62	2049	40	97	82	83	82	0,648
COSMO	313	465	54	1908	40	97	81	85	80	0,657
Заблаговременность +42										
GFS	291	405	61	2091	42	97	84	83	84	0,664
COSMO	294	449	57	1928	40	97	81	84	81	0,649
Заблаговременность +66										
GFS	262	465	70	1987	36	97	81	79	81	0,600
COSMO	263	526	69	1744	33	96	77	79	77	0,560

Основные выводы из раздела:

- 1) Прогноз перехода температуры через ноль по модельным выходным полям температуры воздуха на высоте 2 метра в теплом периоде занижает число таких ситуаций на факте. Комплексование прогнозов с температурой на уровне поверхности почвы повышает предупреденность на 5-8 %.
- 2) Применение разработанного метода распознавания переходов температуры через ноль на базе модифицированного алгоритма DW значительно увеличивает результативность прогнозов. Для модели GFS на независимой выборке средний по территории критерий Пирси-Обухова возрастает в два раза по сравнению с оценками по модельным параметрам.
- 3) Прогнозы заморозков в воздухе по DW-решениям на базе модели COSMO в среднем более успешны, чем по модели GFS. Для прогнозов заморозков на почве качество снижается по обеим моделям и выравнивается.
- 4) По разработанным решениям прогнозы заморозков в воздухе на независимой выборке май-сентябрь 209 г. показали среднюю предупреденность явления около 90% при вероятности около 55%, средний критерий Пирси-Обухова около 0,8. Для прогнозов заморозков на почве в том же периоде вероятность ниже примерно в два раза – около 25%, а предупреденность снизилась примерно на 5% -85-88%.
- 5) По срокам заблаговременности качество прогнозов меняется несущественно.

4 ГОЛОЛЕД И ГОЛОЛЕДИЦА – ФАКТ

4.1 Авторская типизация явлений

В процессе работы по теме при подготовке и первичном анализе имеющихся данных наблюдений (в коде КН-01) стало ясно, что с точки зрения потребителя фиксации только процесса образования гололеда на очень неравномерной сети станций недостаточно, а наблюдений на станциях за весьма актуальным для городов явлением гололедицы вообще нет по известным объективным причинам. Авторы приняли нелегкое, увеличивающее сложность и объем работ решение разработать дополнительную типизацию гололеда и первую экспериментальную - для гололедицы, взяв за основу априорные потенциальные условия и динамику погоды, влияющие на образование, сохранение и ослабление указанных явлений.

В таблице 8 представлена рабочая кодировка с расшифровкой и соответствующая отобранная за 2014-2017гг элементарная статистика по станциям Урала, Сибири и Якутии.

Из сокращенной интерпретации видно, что дополнительная типизация настроена на временную динамику погодных условий, определяющих ответную динамику "жизненного цикла" гололеда и гололедицы. Очевидное слабое место такой типизации видно из анализа колоссального разброса частот кодов по станциям, и это только подтверждает "введение" в проблемы подсеточного и орографического масштабов в начале отчета. Дальнейшие исследования, вероятно, будут уточнять наши результаты, рассматривая отдельно экстремальные географические места, пункты, выделяя кластеры и т.п.

Приоритет кодов: проверка кодов от 1 до 9 (до первого выполнения кода), иначе код 0.

Список необходимых метеорологических параметров для описания данных типов представлен в таблице 9. Создан архив по данным КН01 за холодный период 2014-2019 годов по 432 метеостанциям Урало-Сибирского региона, включая республику Саха.

Предварительная кодировка (кода) явлений (таблица 10), оценивает наличие гололеда и наличие и тип осадков.

Алгоритм кодировки: проверка условий для кодов от 1 до 9 (до первого выполнения кода), иначе код 0.

По описанной ступенчатой кодировке получены архивы за холодный период 2014-2019 годов для каждой станции Урало-Сибирского региона.

Таблица 8 - Частота кодов авторской типизации гололеда и гололедицы по погодным условиям по территории Урало-Сибирского региона и Якутии. Частоты по полусуткам за 4 года.

Код	Максимальное число случаев (индекс станции)	Среднее число случаев	Минимальное число случаев	Явление
1	175 (23912)	6	1	гололед (КН-01)
2	629 (30961)	66	2	гололедица сильная: жидкие/смешанные осадки + резкий переход через 0 ($t \leq -10/12ч$)
3	94 (23912)	3	3	гололед сохраняется: после кода 1, пока сроки снег $\leq 1см$, $t_n \leq -1н$, $t_d \leq -5$, (до +2-х суток)
4	699 (30961)	115	4	гололедица умеренная: жидкие/смешанные осадки + постепенный переход через 0 ($t < -7/24ч$)
5	57 (23912)	1	5	гололед скрытый: после кодов 1,3 лег снег $\leq 5см$ (до +1-х суток)
6	199 (23527)	47	6	гололедица местами: устойчивый снежный покров + оттепель + резкий переход через 0 ($t \leq -5/12ч$)
7	58 (28144)	13	7	гололедица местами: устойчивый снежный покров + солнце + $t_d \geq -5 + t_n < -5$
8	26 (23912)	1	8	гололед слабеет: после кодов 1,3,5 лег_снег $\geq 5см$ (до +2-х суток)
9	29 (23912)	1	9	гололед тает: после кодов 1,2,3 $t_n > 0$, $t_d > -5$, (до +1-х суток)
0	Ни одно условие не выполняется			

Таблица 9 – Параметры наблюдений из КН01

Параметр	Описание параметра
G	код гололеда – (0, 1,2,3, 4) (кодА согласно таблице 2)
ТТ	температура воздуха на уровне 2м
Tmin	минимальная температура за ночь
Tmax	максимальная температура за день
Ttr	температура точки росы
RR	количество осадков за 12 ч (в коде КН-01), в срок "00" - за ночь, в срок "12" -
FF	скорость ветра в срок
FP	максимальный порыв ветра (в срок и между сроками)
N	октант нижней облачности (в коде КН01)
HE	состояние подстилающей поверхности (в коде КН01)
HSS	высота снежного покрова (в коде КН01)
E51	состояние поверхности: 2-сырая, 3-вода, 4-замерзшая, 5-покрыта льдом

Таблица 10 - Кодирование архива (кода) по кодам явлений погоды КН01

кода (W)	Явление	Коды КН-01 (WW, W1W2)
0	нет жидких, смешанных осадков и гололеда	-
1	Гололед	24, 56, 57, 66, 67
2	смешанные осадки	23, 68, 69, 79, 83, 84, 87, 88, 95, 97
3	жидкие осадки	20, 21, 25, 50, 51, 52, 53, 54, 55, 58, 59, 60, 61, 62, 63, 64, 65, 80, 81, 82, 91, 92, 93, 94
4	жидкие осадки между сроками	5, 6, 8

4.2 Обеспеченность данными гололеда и гололедицы по разработанным кодам на обучающей выборке

На рисунках 12-20 представлено распределение числа случаев по каждому коду по станциям территории за период 2014-17 гг. На рисунке 12 видно, что на восточной и северо-восточной части данные по основному коду гололеда в основном отсутствуют. Данных о наличии явления нет на 225 метеостанциях, это больше половины территории. Вопрос о причинах этого остается открытым: не бывает гололеда или нет станков, если есть добросовестно ли передаются данные. Климатическая вероятность гололеда максимальна в Пермском крае, который расположен с наветренной стороны Уральских гор и подвержен намерзанию льда. По суточному ходу: в дневное время явление чаще, но на небольшой процент. Сохранение гололеда (код3, рисунок 14) актуально лишь на юге Омской области и в Пермском крае. Еще реже продолжение гололедной ситуации (коды5 и 9, рисунки 16 и 20). Данных, когда гололед слабеет в связи с выпавшим снегом (код 8, рисунок 19) совсем мало, лишь на единичных станциях их число достигает 10 раз за 4 года.

Сильная гололедица, когда жидкие или смешанные осадки сопровождаются резким переходом температуры через ноль (код2, рисунок 13) имеет максимальную вероятность из всех рассматриваемых типов явления. В ночное время код2 отмечен на большинстве станций на несколько случаев чаще. Умеренная гололедица (код4, рисунок 15) чаще всего отмечена на юго-востоке Западной Сибири: юг Красноярского края, Кемеровская, Новосибирская области, Алтайский край. Гололедицы местами при снежном покрове с оттепелью и переходами температуры через ноль по кодам 6 и 7 чаще наблюдаются также на юге и западе региона.

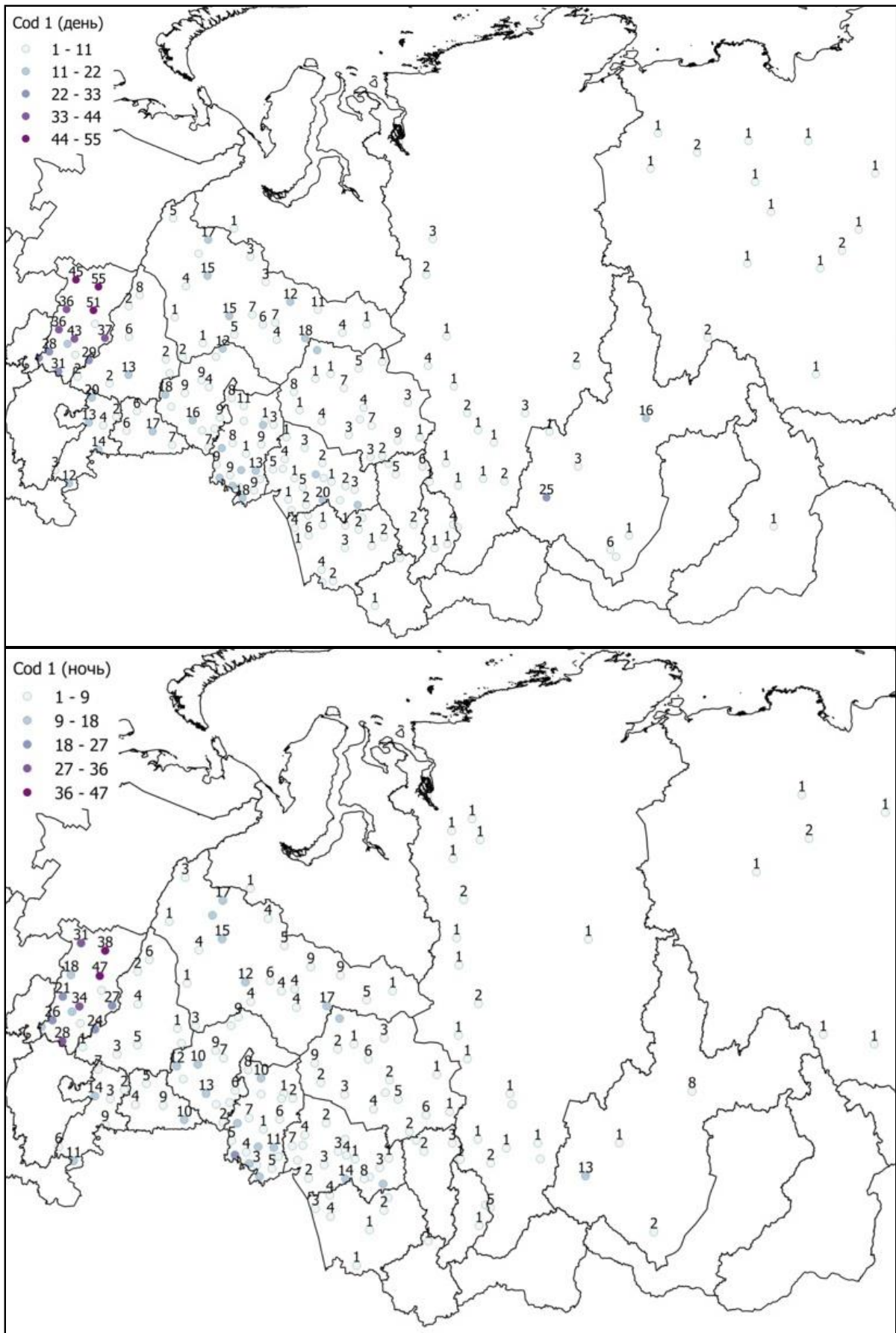


Рисунок 12-Распределение числа случаев с кодом=1 по территории региона.
Холодный период 2014-2017 гг.

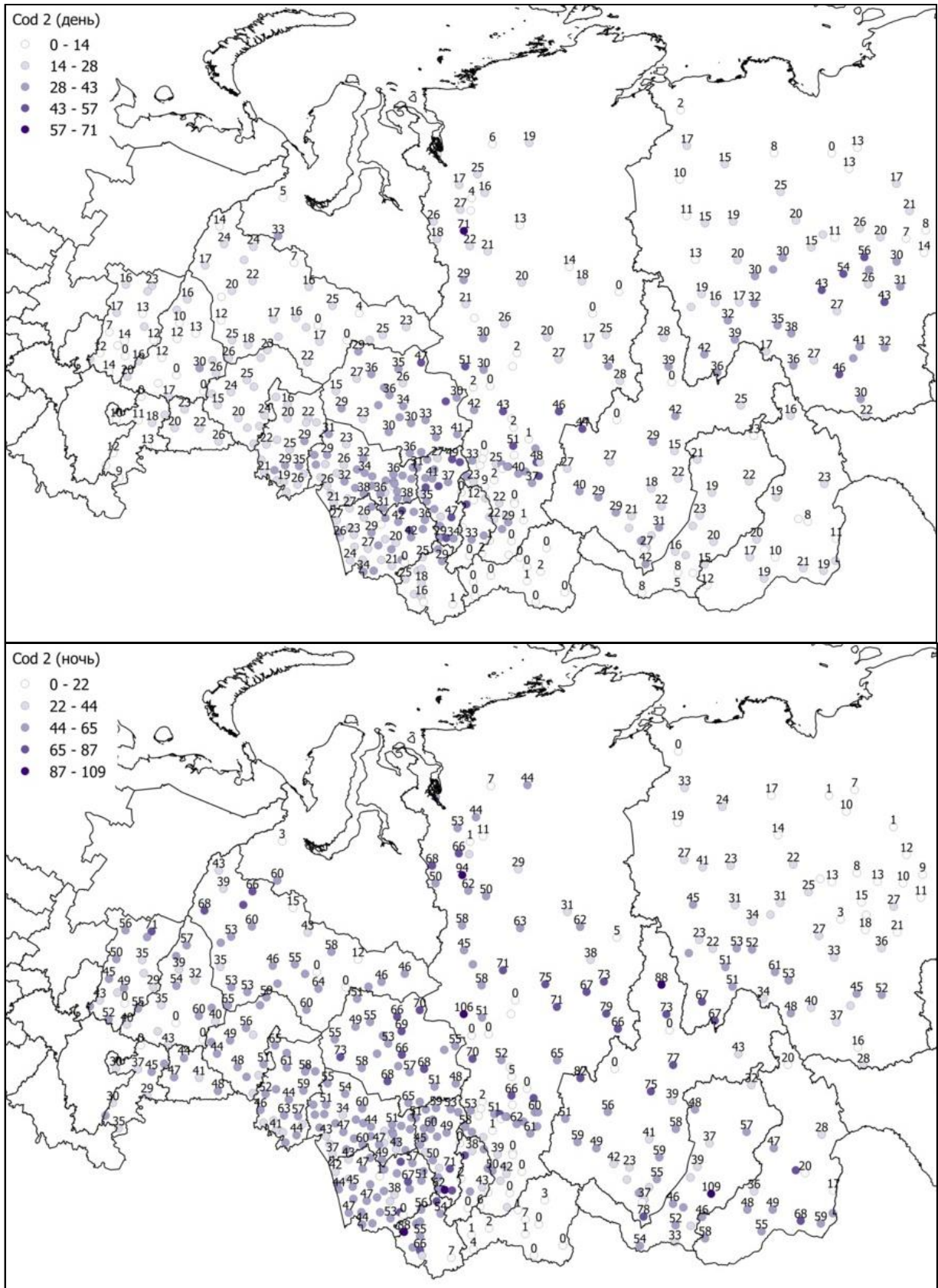


Рисунок 13-Распределение числа случаев с кодом=2 по территории региона.
Холодный период 2014-2017 гг.

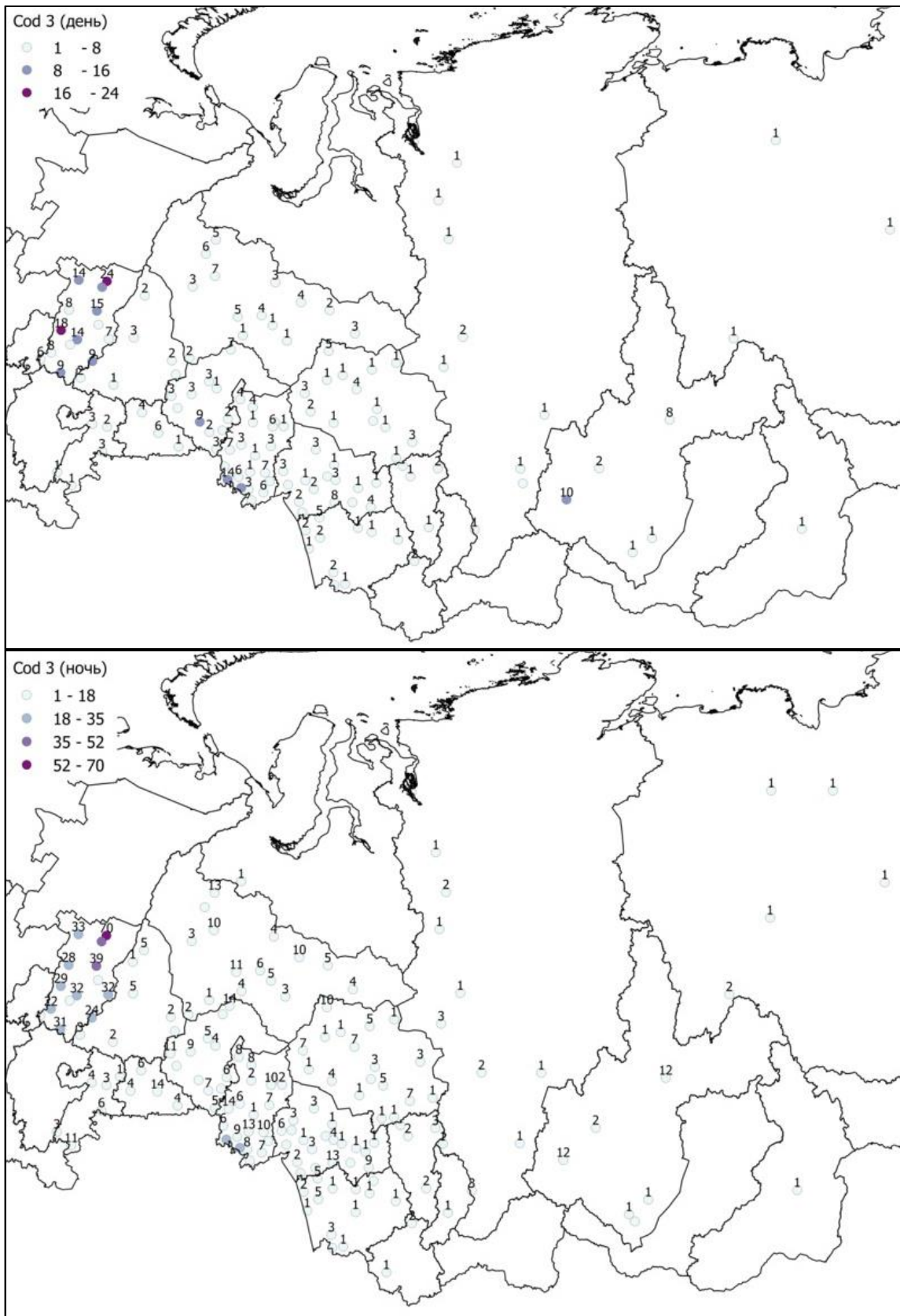


Рисунок 14-Распределение числа случаев с кодом=3 по территории региона. Холодный период 2014-2017гг

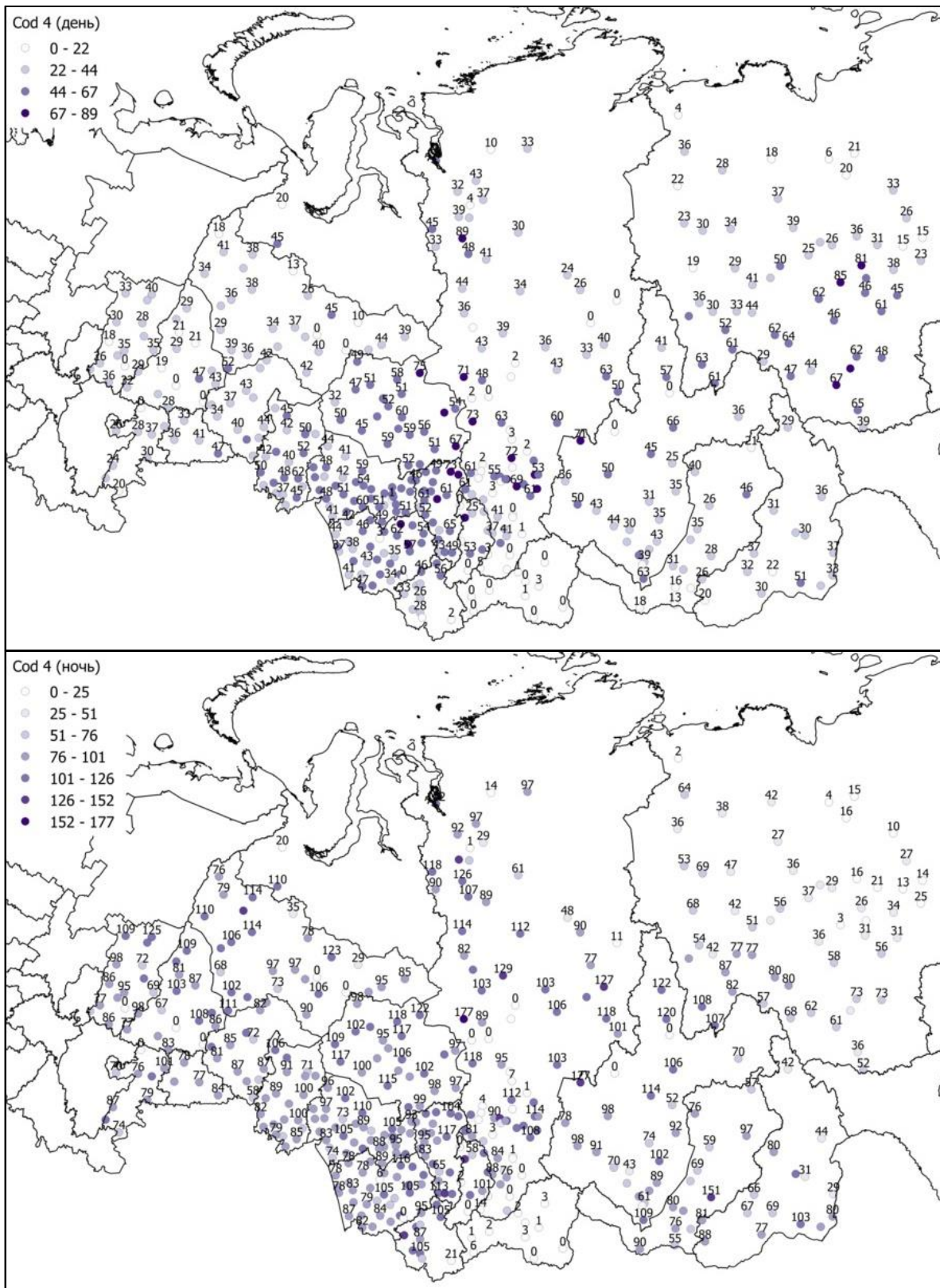


Рисунок 15-Распределение числа случаев с кодом=4 по территории региона.
Холодный период 2014-2017гг

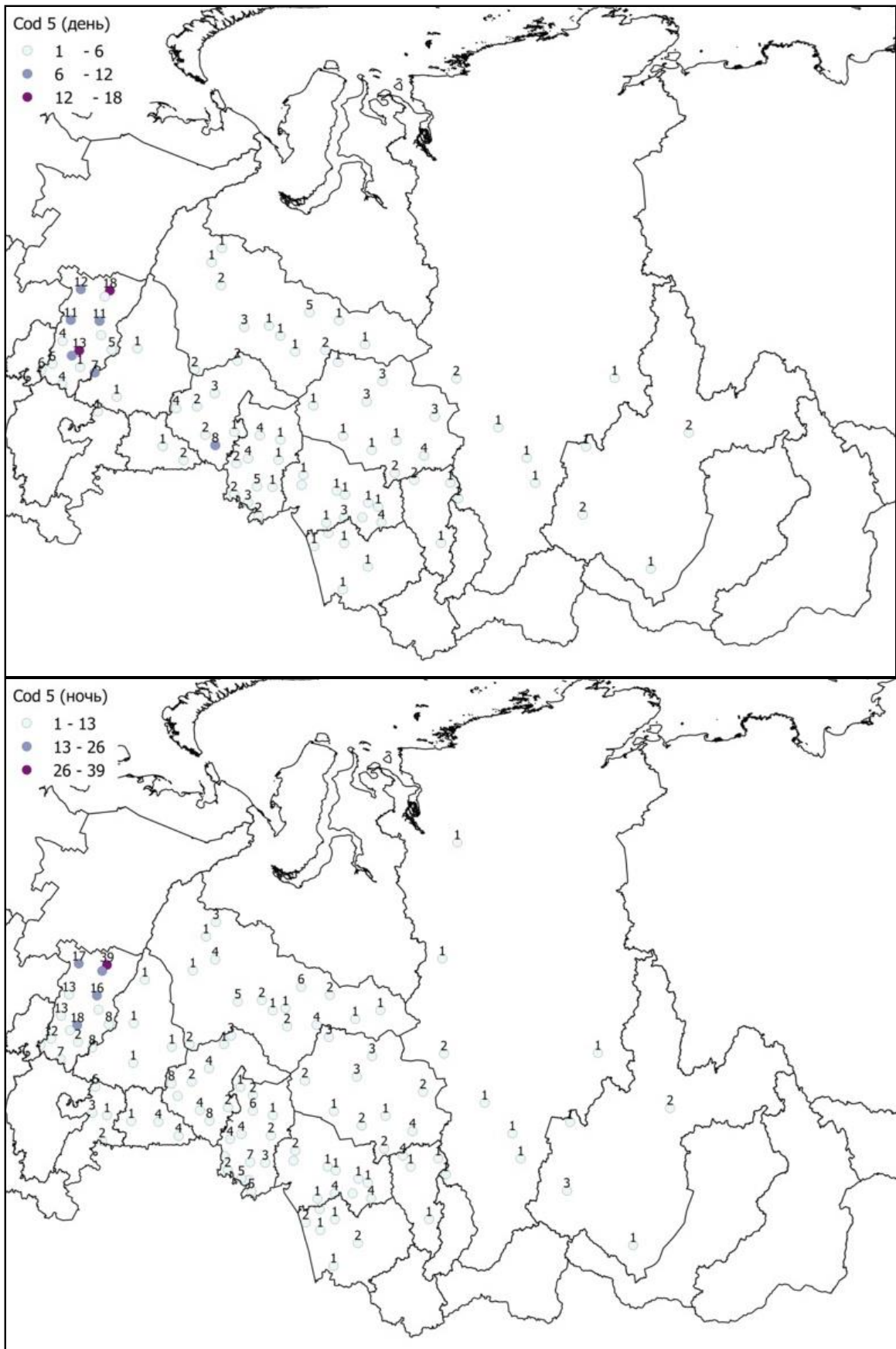


Рисунок 16-Распределение числа случаев с кодом=5 по территории региона.
Холодный период 2014-2017 гг.

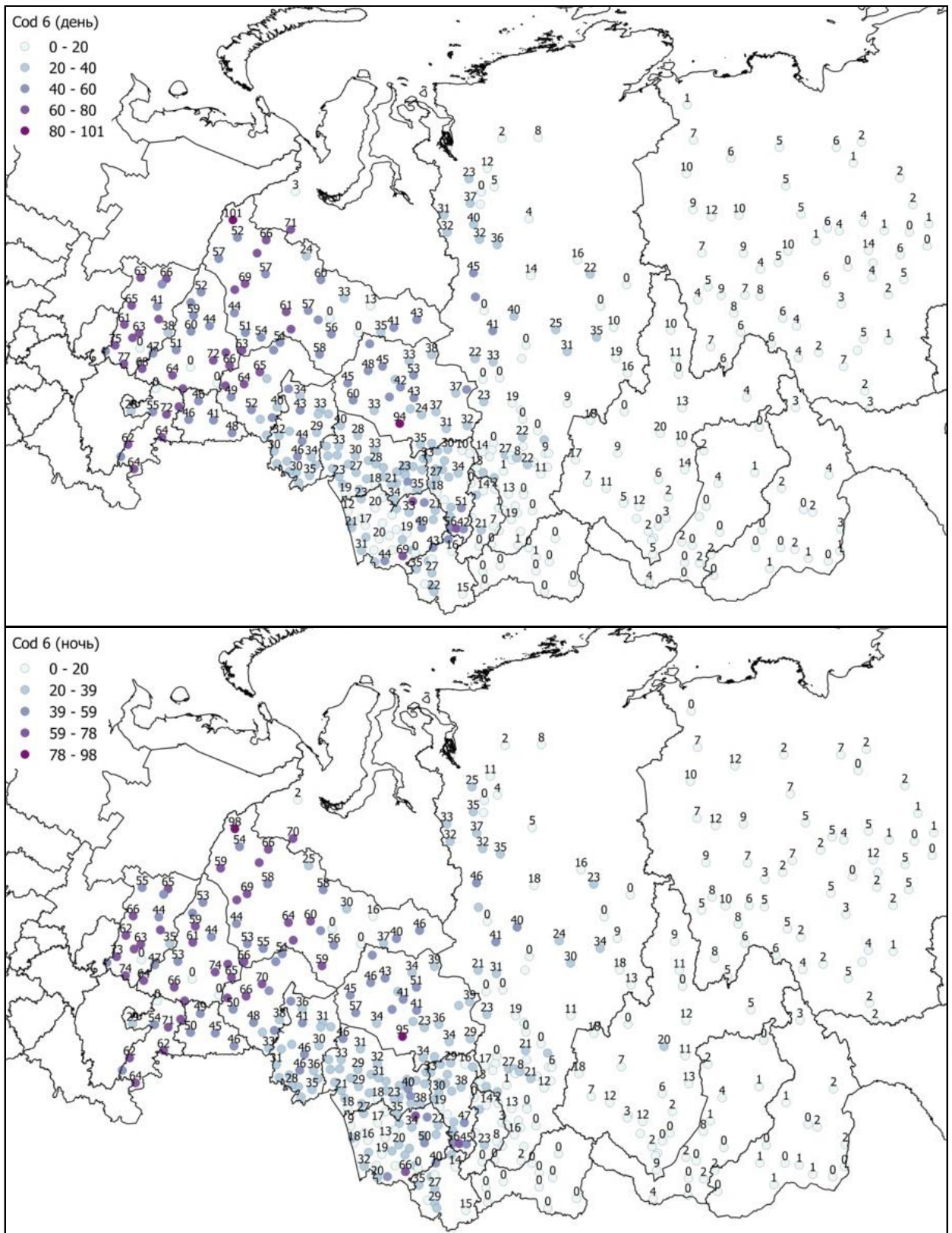


Рисунок 17-Распределение числа случаев с кодом=6 по территории региона.
Холодный период 2014-2017 гг.

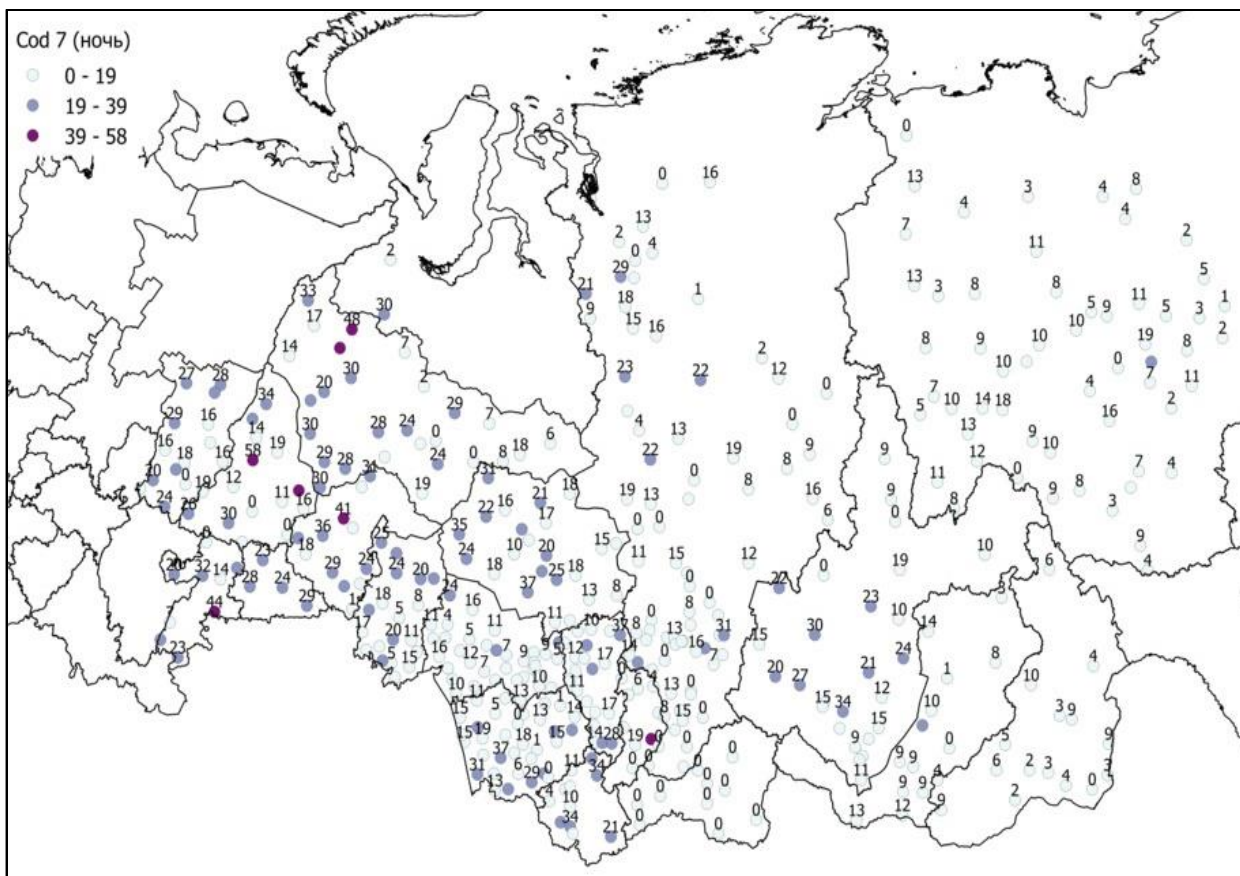


Рисунок 18-Распределение числа случаев с кодом=7 по территории региона.
Холодный период 2014-2017 гг.

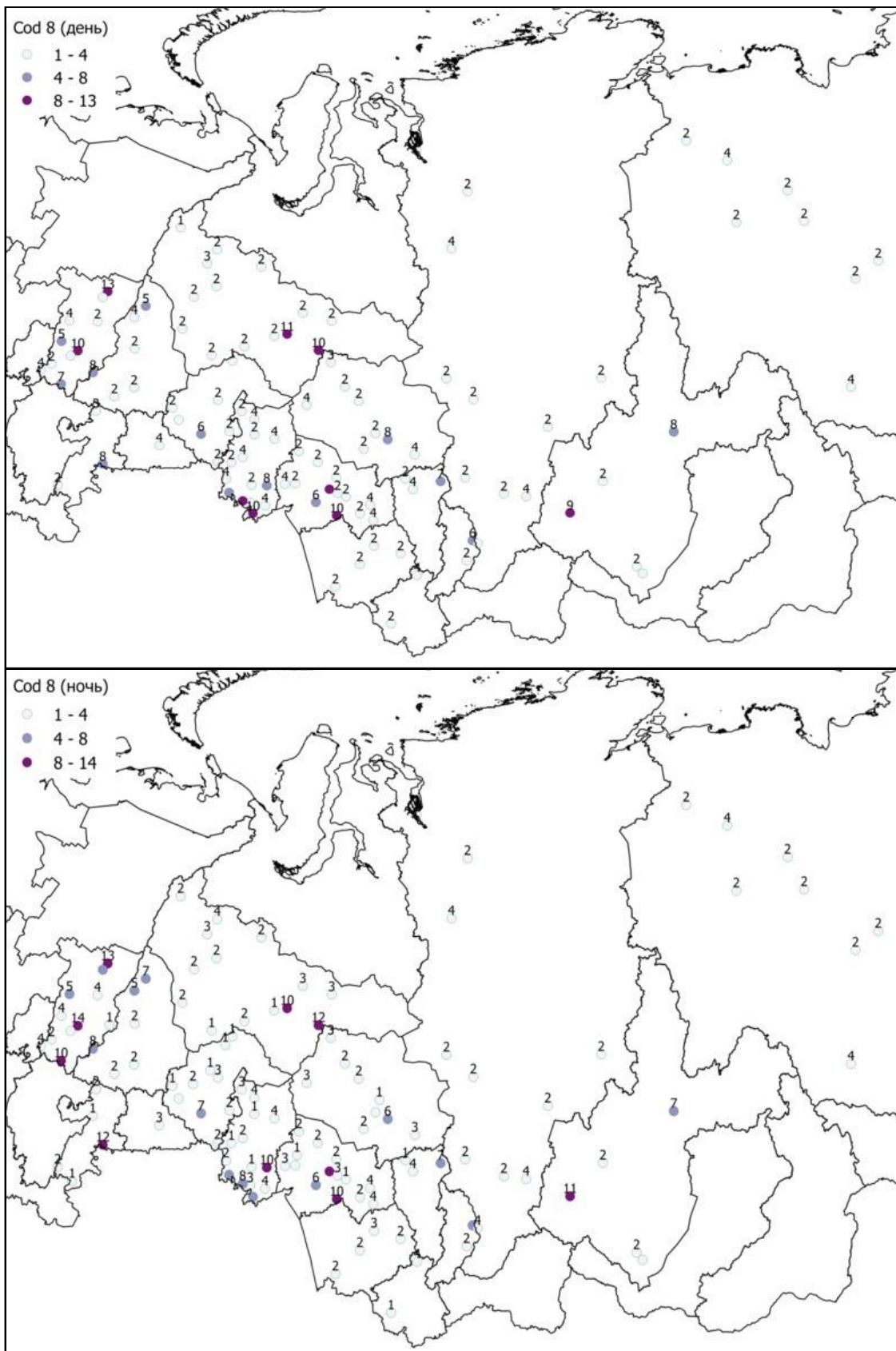


Рисунок 19-Распределение числа случаев с кодом=8 по территории региона.
Холодный период 2014-2017 гг.

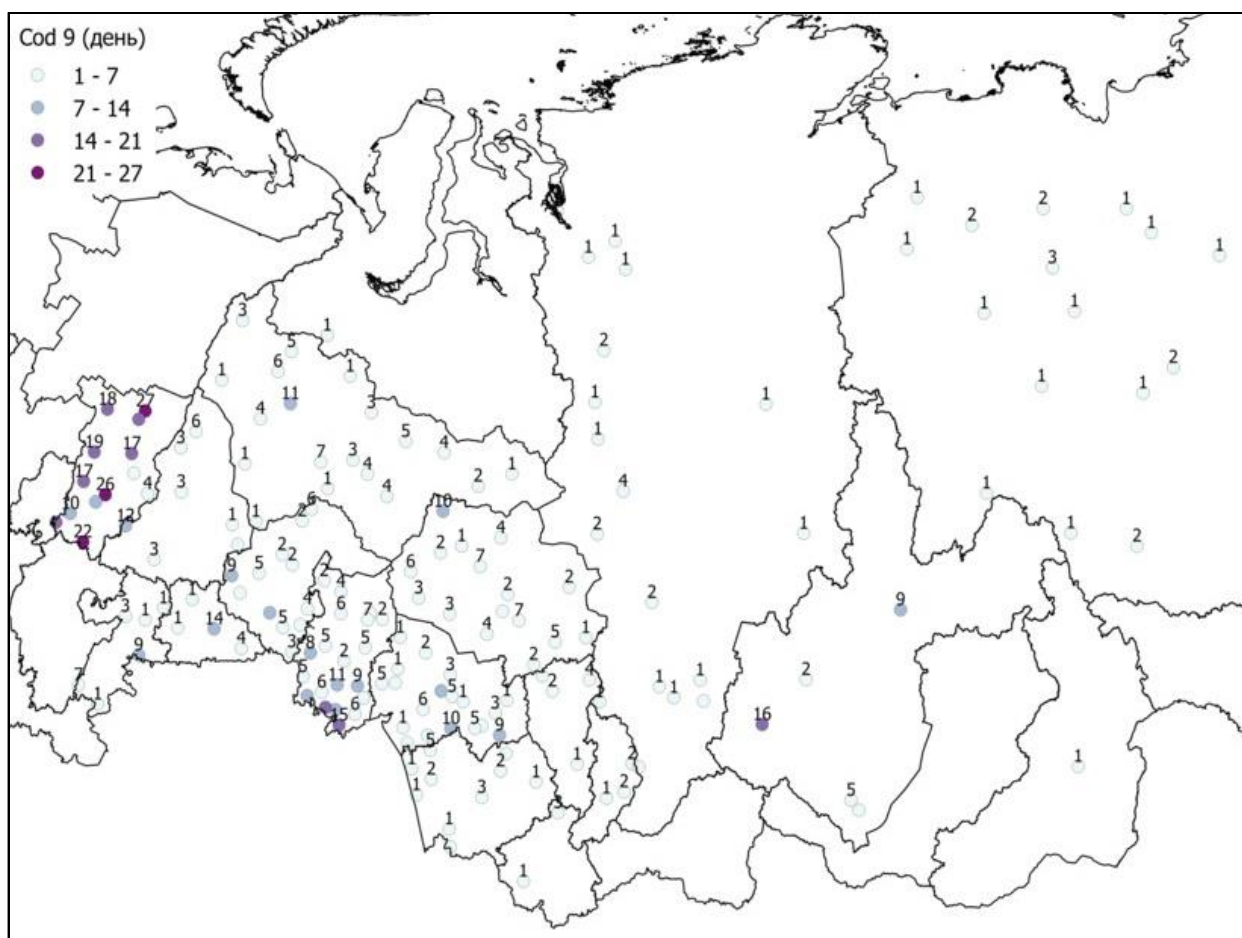


Рисунок 20-Распределение числа случаев с кодом=9 по территории региона. Холодный период 2014-2017 гг.

По алгоритму, описанному в разделе 2 настоящего отчета, построены бинарные деревья логических решений отдельно для гололеда и по каждому коду типизации гололедицы. Деревья построены для каждой станции, заблаговременности (00+18, 00+42, 00+66) по каждой модели (GFS, COSMO).

4.3 Оценка прогнозов гололеда

Как показано выше, фактическая вероятность гололеда на большей части территории региона низкая, за исключением Пермского края. Поэтому в ходе исследования были проведены различные эксперименты по формированию предиктанта с целью повышения устойчивости решений. Для каждого варианта по методологии, описанной во втором разделе настоящего отчета, построены бинарные логические деревья решений. По построенным деревьям получены результаты восстановления гололеда по выборке холодного периода 2018 года и рассчитаны оценки. Как следует из таблицы 11, прогнозы

гололеда, рассчитанные по каждой (из 298) станции, показывают низкую успешность. При усреднении оценок по территории предупреденность явления не превышает 55% при чрезвычайно большом количестве «ложных тревог», так что средняя оправдываемость явления не выше 3%. В связи с этим, был проверен следующий вариант: выборки по классам явление/без явления формируются по всей территории с соответствующим каждому случаю набором предикторов. Видно (таблица 11), что этот вариант существенно повысил предупреденность (до 80%), но оправдываемость явления сохранилась низкой. Наилучшим оказался следующий вариант, в котором данные предиктанта сформированы по кластерам радиусом 200км. Число ложных тревог уменьшилось в 10 раз при сохранении предупреденности на уровне 78-80% в среднем по территории.

Последний вариант принят за окончательный. В таблице 12 приведены оценки для этого варианта, но для деревьев на базе разных моделей. Видно, что результаты сопоставимы, оценки примерно одного уровня.

Таблица 11 - Сравнение оценок прогноза гололеда по разным вариантам решений. Модель GFS. Холодный период 2018 год.

	k11	k12	k21	k22	k0	vr1+	vr2-	vr	pr1	pr2	MPR	TSS
COSMO												
00+06	1902	6836	1032	93419	103189	21,8	98,9	92,4	64,8	93,2	64,8	0,580
00+18	1645	7186	861	110065	119757	18,6	99,2	93,3	65,6	93,9	65,6	0,595
00+30	1859	7693	1123	93627	104302	19,5	98,8	91,5	62,3	92,4	62,3	0,547
00+42	1603	9646	925	108377	120551	14,3	99,2	91,2	63,4	91,8	63,4	0,552
00+54	1885	9398	1066	92795	105144	16,7	98,9	90,0	63,9	90,8	63,9	0,547
00+66	1662	11509	852	106760	120783	12,6	99,2	89,8	66,1	90,3	66,1	0,564
GFS												
00+06	1813	7228	1040	90552	100633	20,1	98,9	91,8	63,5	92,6	63,5	0,562
00+18	1582	6865	762	111457	120666	18,7	99,3	93,7	67,5	94,2	67,5	0,617
00+30	1854	8037	1027	89849	100767	18,7	98,9	91,0	64,4	91,8	64,4	0,561
00+42	1578	8155	834	110113	120680	16,2	99,2	92,6	65,4	93,1	65,4	0,585
00+54	1865	7969	1054	89084	99972	19,0	98,8	91,0	63,9	91,8	63,9	0,557
00+66	1644	7438	847	110261	120190	18,1	99,2	93,1	66,0	93,7	66,0	0,597

Анализ полученных прогностических деревьев показал, что повторяемость выбранных алгоритмом параметров:

- в целом отражает известную физику формирования условий в нижней атмосфере и у поверхности для образования гололеда;

- частотные списки лучших параметров для моделей COSMO и GFS, в основном, пересекаются для ночного срока 00+18ч (см. таблицу 2, номера параметров):

9-RH9, 36-min(T0+5), 38-abs(T0+5), 28-AP00, 15-dTd0, 3-T9-T0;

- для дневного срока 00+30ч списки существенно различаются:

15-dTd0, 28-AP00, 7-Td0 - у COSMO вместо 36-min(T0+5), 24-H500-H92 у GFS, при частично совпадающих 38-abs(T0+5), 9-RH9, 3-T9-T0.

Для кодов расширенной типизации гололеда и гололедицы также были построены деревья решений, но прямая их проверка, к сожалению, невозможна, поскольку фактически прогнозируются условия погоды, для которых сформулированы ожидаемые, но не фиксируемые явления, в прогнозе которых заинтересованы потребители метеоинформации.

Для использования полученных результатов, вероятно, придется сначала решить некоторые организационные вопросы с ЦГМС, которые возьмутся их опробовать в опытном порядке и оценить дальнейшие перспективы. Передавать на испытания в традиционном принятом порядке наши новые результаты невозможно.

5 ТЕХНОЛОГИЧЕСКИЕ БЛОКИ ПРОГНОЗОВ

Полученные DW-решения для всех явлений и программы восстановления по ним встроены в оперативную технологию производства краткосрочных прогнозов в ФГБУ «ЗС РВЦ» и в ФГБУ «СибНИГМИ».

Технология включает следующие блоки:

1. Производство расчетов по COSMO_RU_Sib13. Форматирование выходных данных, выбор базовых параметров в ближайших узлах по списку станций региона, расчет дополнительных параметров.
2. Скачивание выходной продукции модели GFS(NCEP). Форматирование выходных данных, выбор базовых параметров в ближайших узлах по списку станций региона, расчет дополнительных параметров.
3. Прогноз явлений по станциям и кластерам на 1-3 суток.
4. Формирование результатов в таблицы и передача их пользователям через интернет-ресурсы.
5. Архивирование фактических метеорологических параметров из SQL-базы КН-01.
6. Архивирование результатов расчетов прогнозов явлений.
7. Автоматизированный расчет оценок прогнозов по разработанному методу.

Для каждого блока подготовлено программное обеспечение с целью работы в оперативном режиме.

ЗАКЛЮЧЕНИЕ

В ходе выполненного исследования получены следующие результаты:

1. В основе прогностических методов лежит уже используемый в предыдущих разработках DW алгоритм построения логических бинарных деревьев разделения образов. Продолжена работа над оптимизацией данного алгоритма, направленная на повышение устойчивости решений за счет исключения случайных или редких связей факторов с предиктантом. Апробированы новые варианты сбалансированного критериального упрощения решений.
2. Разработана типизация гололедных явлений по приземным метеорологическим условиям их образования и продолжительности, всего 9 типов. Для каждого типа рассмотрена обеспеченность на архивной выборке 2014-17 годов и распределение по территории.
3. Для выборки предиктантов использованы данные КН-01 по 430 метеостанциям Урало-Сибирского региона за период 2014-2019 годы. За синхронный период создан архив признаков (предикторов), включающий 39 параметров, выбранных и рассчитанных по выходным полям гидродинамических моделей атмосферы. В качестве базовых моделей использованы COSMO (модификация Ru_Sib13) с разрешением 13.2км и GFS(NCEP) с разрешением $0,5^0*0,5^0$. Результаты первой рассчитываются в ЗапСибРВЦ в оперативном режиме, второй – скачиваются с сервера NCEP по ftp-протоколу.
4. Для заморозков в воздухе, заморозков на почве и для гололеда построены DW – деревья. Решения получены для каждой станции на каждый срок заблаговременности (1-3 суток) по каждой модели. Рабочая выборка включает 2014-2017, данные 2018 года использованы в алгоритме селекции решений, а 2019 – для оценок на независимом материале. В силу малой обеспеченности данными о гололеде на территории региона получены дополнительно варианты решений для кластеров радиусом 200км, которые показали более качественный результат. По этой же причине на этапе отбора решений для гололеда из множества предварительных использованы данные за два года (2018-2019), поэтому оценить результаты на полностью независимой выборке невозможно. Анализ будет продолжен в ходе оперативных испытаний.
5. Получены сравнительные оценки для прогнозов перехода температуры через ноль в теплом периоде с переходом на прогноз заморозков.

Основные выводы:

- Прогноз перехода температуры через ноль по модельным выходным полям температуры воздуха на высоте 2 метра в теплом периоде занижает число таких ситуаций на факте. Комплексование прогнозов с температурой на уровне поверхности почвы повышает предупрежденность на 5-8 %.

- Применение разработанного метода распознавания переходов температуры через ноль на базе модифицированного алгоритма DW значительно увеличивает результативность прогнозов. Для модели GFS на независимой выборке средний по территории критерий Пирси-Обухова возрастает в два раза по сравнению с оценками по модельным параметрам.

- Прогнозы заморозков в воздухе по DW-решениям на базе модели COSMO в среднем более успешны, чем по модели GFS. Для прогнозов заморозков на почве качество снижается по обоим моделям и выравнивается.

- По разработанным решениям прогнозы заморозков в воздухе на независимой выборке май-сентябрь 209 г. показали среднюю предупрежденность явления около 90% при вероятности около 55%, средний критерий Пирси-Обухова около 0,8. Для прогнозов заморозков на почве в том же периоде вероятность ниже примерно в два раза – около 25%, а предупрежденность снизилась примерно на 5% -85-88%.

- По срокам заблаговременности качество прогнозов меняется незначительно.

6. Оценки прогнозов гололеда на выборке 2018-2019гг., не участвующей в обучении, но являющейся решающей при отборе лучших решений показали сопоставимость результатов по обоим базовым моделям. В среднем по территории предупрежденность явления находится на уровне 63-66% на все сроки прогноза. Вероятность явления по модели COSMO снижается к третьей ночи на 9 % от первых +6 часов, достигая 12,5%. По модели GFS на все сроки вероятность на уровне 18-20%. Средняя оправдываемость прогнозов около 90-93%, и критерий Пирси-Обухова 0,55-0,60 независимо от базовой продукции.

Выполненные исследования подхватили чуть не утерянную традицию и одно из основных предназначений СибНИГМИ - развивать самые актуальные для Урало-Сибирского региона и Якутии метеорологические направления. А что может быть актуальнее для нашего региона с полугодовой морозной зимой чем гололедные явления?

Работа выполнена на современном уровне с использованием выходной продукции актуальных оперативных моделей. К плюсам можно отнести и использование двух моделей различных научных школ для сравнения и оценки их применимости к условиям

самого резко-континентального климата на планете, для которого они вряд ли специально адаптировались.

Значительную часть работы содержит специальная подготовка и обработка входных данных для трех целевых задач. Авторы почти не использовали готовых стандартных алгоритмов и программ, поскольку предиктанты слишком неоднородны и даже трудноформализуемы для машинной обработки. К тому же в процессе разработки методики и алгоритмов зачастую рождается понимание границ возможного после экспериментирования с самими входными данными, а не только с алгоритмами построения решений.

Выполненная работа включает как законченные разделы, доведенные до готовых к передаче на оперативные испытания программ усвоения оперативных данных и полностью автоматизированных расчетов прогнозов (заморозки и гололед), так и открывает новое направление исследований (гололедица), для внедрения первых результатов которого нормативные документы пока не готовы.

Методы и технология прогнозов гололеда, заморозков в воздухе и на почве передаются на оперативные испытания в подразделения Гидрометслужбы Урало-Сибирского региона

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Ягудин Р.А. Рекомендации к прогнозу гололеда на юго-востоке Западной Сибири//Труды «ЗСРНИГМИ». - М.:Гидрометеиздат.-1977.-Вып.30.-С.68-69.
2. Руководство по краткосрочным прогнозам погоды. Часть II. Выпуск 2. Урал и Сибирь/под ред. В.М.Ярковой и И.П.Фадеевой/Л.:Гидрометеиздат.-1988.-С.181-185.
3. РД 52.27.724-2019. Наставление по краткосрочным прогнозам погоды общего назначения. //М., ФГБУ «Гидрометцентр России».- 2019.-66с.
4. Шевчук И.А. Синоптические условия осенних заморозков в юго-восточных районах Западной Сибири //Сб. работ по региональной синоптике.- 1964. - №9.- С.87-97
5. Храмцова И.Г. Климатическая характеристика заморозков (в том числе особо опасных) на юго-востоке Западной Сибири//Труды ЗСРНИГМИ – 1977–Вып. 27 – С.49-55
6. Шарифуллин Н.К. Модель прогноза заморозков//Труды ЗапСибНИИ Госкомгидромета СССР.- 1988.- Вып.83. – С.90-98.
7. Прогноз заморозков в воздухе и на поверхности почвы на 1-5 суток, детализированный по суткам для юго-востока Западной Сибири: Отчет о НИР (промежут.)/ СибНИГМИ Росгидромета; Руководитель И.Г.Храмцова.- Шифр темы 9.2.23; №ГР 01940001672; Инв. № 03950000278. – Новосибирск, 1994.-48с.
8. Храмцова И.Г., Шустова Г.А. Автоматизированный метод прогноза заморозков на 1-5 суток для юго-востока Западной Сибири // Труды СибНИГМИ.-2000.-Вып.102.-С.34-40.
9. Прогноз заморозков, детализированной для Омской и Тюменской областей: отчет по НИР (промежуточный)// СибНИГМИ Росгидромета: руководитель Храмцова И.Г. – Шифр темы 1.19.1.1; №ГР 0196009726; Новосибирск,1997.-38с.
10. Прогноз заморозков на 1-5 суток, детализированный по суткам для Свердловской, Пермской, Челябинской и Курганской областей: отчет по НИР (промежуточный) //СибНИГМИ Росгидромета; руководитель Храмцова И.Г. Шифр темы 1.19.1.1; №ГР 0196009726; Новосибирск, 1998.-41с.
11. Токарев В.Г. Метод прогноза заморозков по дням пентады для юга Красноярского края.//Труды СибНИГМИ. -2000.-Вып.103.-С.115-120
12. Разработка метода и технологии расчета прогноза погоды по полусуткам пентады для административных районов Новосибирской области: Отчет по НИР/СибНИГМИ Росгидромета: руководитель Здерова М.Я. – Шифр темы 1.1.2.9(р). Новосибирск, 2005.-76с.

13. Здерева М.Я., Торубарова Г.П., Шустова Г.А. Физико-статистическая схема прогноза экстремальной температуры воздуха по пунктам Новосибирской области на 1-5 суток // Труды СибНИГМИ. - 2006. – Вып.105. – С.40-46.
14. Лбов Г.С., Манохин А.Н. Распознавание образов при разнотипных признаках в условиях малой выборки // Статистические проблемы управления. - Вильнюс, 1976.- Вып.14.-С.57-63.
15. Манохин А.Н. Алгоритм DW для распознавания образов: Пакет прикладных программ ОТЭКС.-Новосибирск: изд-во Новосибирского государственного университета, 1981.- С.3-30.
16. Загоруйко Н.Г., Елкина В.Н., Емельянов С.В., Лбов Г.С. Пакет прикладных программ ОТЭКС. М.: Финансы и статистика, 1986. - 160с.
17. Здерева М.Я., Токарев В.М. Анализ и прогноз условий погоды, влияющих на концентрацию атмосферных примесей мегаполиса // Труды СибНИГМИ. 2011. - Вып.106. –С.152-158.
18. Здерева М.Я., Токарев В.М. Хлучина Н.А. Воробьева Л.П., Бабошина Н.А. Оперативная технология прогноза гроз в Сибири и результаты её испытаний // Труды Гидрометеорологического научно-исследовательского центра Российской Федерации, 2018. - №2(368). - С. 27-43.
19. Разработать технологию краткосрочного прогнозирования локальных неблагоприятных явлений погоды для Западной Сибири на базе физико-статистической интерпретации и комплексирования выходной продукции гидродинамического моделирования в РСМЦ Новосибирск // Отчет НИР (заключ.). - ФГБУ «СибНИГМИ»; Шифр темы 1.1.1.2; рук. Здерева М.Я. – Новосибирск, 2016.- 48с. - № ГР 01201454707; Рег. № ИКРБС: АААА-Б17-217011260064-6;
20. Ивахненко А.Г., Юрачковский Ю.П. Моделирование сложных систем по экспериментальным данным.- М.: Радио и связь, 1987. -120с.