

Глава 3

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ В ФУНДАМЕНТАЛЬНЫХ И ПРИКЛАДНЫХ ИССЛЕДОВАНИЯХ, ГЕОИНФОРМАЦИОННО КАРТОГРАФИРОВАНИЕ. СПЕЦИАЛЬНЫЕ БАЗЫ ПРОБЛЕМНО-ОРИЕНТИРОВАННЫХ И ПРЕДМЕТНО-ОРИЕНТИРОВАННЫХ ДАННЫХ И ЗНАНИЙ

3.1. ИСПОЛЬЗОВАНИЕ ПРОСТРАНСТВЕННЫХ МЕТЕОРОЛОГИЧЕСКИХ ДАННЫХ В ПРОГНОСТИЧЕСКИХ МЕТОДАХ И ПРИМЕР ПРИМЕНЕНИЯ WEB-ТЕХНОЛОГИЙ ДЛЯ ПРЕДСТАВЛЕНИЯ РЕЗУЛЬТАТОВ

Постановка задачи

Метеорологические и аэрологические станции, гидрологические посты, пункты радиолокационного зондирования и спутниковые снимки предоставляют широкий спектр базовых метеорологических данных наблюдений. Их регулярность и трехмерная пространственная структура позволяют строить гидродинамические модели атмосферных процессов и рассчитывать прогностические метеорологические поля. Развитие данного направления существенно облегчает работу современных оперативных синоптиков-прогнозистов, предоставляя будущие состояния атмосферы как в глобальном, так и в мезо-масштабе. В настоящее время гидродинамические модели успешно рассчитывают прогнозы барических и термических полей на сроки до 7-10 суток, однако ожидаемые значения элементов погоды в конкретных точках прогнозируются значительно менее успешно. Решают эту проблему исследования, связывающие элементы локальной погоды с термобарическими системами. До сих пор оперативные прогностические службы частично руководствуются синоптическими разработками. Автоматизировать трудоемкий и субъективный процесс составления прогнозов погоды призваны методы интерпретации модельных полей, базирующиеся на статистических алгоритмах с учетом физических связей между параметрами.

Непрерывность рядов данных по времени и пространству для большинства метеорологических параметров предоставляет возможность применения статистического анализа. Задача ставится следующим образом: для каждой точки (или пространства) на обучающих выборках получить решения, дающие возможность восстановить требуемый результат по модельным прогностическим полям. При этом различают несколько вариантов построения решений: на реальных фактических данных (PP-Perfect Prognose), на результатах модельных расчетов (MOS-Model Output Statistics) или на их комбинации. В первом случае полученные решения для восстановления являются независимыми по отношению к применяемой модели или к изменениям внутри одной модели. В варианте MOS в конечном результате заложен учет особенностей в реализациях конкретной модели. Исследователь выбирает тот или иной подход согласно технологическим возможностям и эффективности работы метода на выходе.

Сибирские разработки автоматизированных прогнозов погоды на базе физико-статистической интерпретации гидродинамической продукции берут начало в 70-х годах прошлого века, родоначальником их является И.Г. Храмова. Первые методические

прогнозы представляли приземную температуру воздуха в среднем на ожидаемую пентаду и декаду. Эволюция численных моделей атмосферных процессов способствовала развитию методов интерпретации в следующих направлениях:

- расширение списка прогнозируемых метеоэлементов;
- детализация прогнозов по времени;
- детализация прогнозов по пространству.

К настоящему времени разработаны и внедрены в оперативную практику методы среднесрочного прогноза не только для отдельных метеопараметров - температура воздуха, осадки, скорость ветра, но и для комплексных - метели, заморозки, метеоусловия для загрязнения воздуха, показатель пожароопасности в лесах [1-4]. Прогнозы детализированы по полусуткам и по заданным пунктам на территории от Урала до Байкала. В качестве базовых модельных прогнозов в наших разработках выступают преимущественно расчетные поля по модели ECMF (Европейского центра среднесрочных прогнозов погоды) и UKMO (Экзетер). Первая имеет наиболее высокие оценки точности, вторая предоставляет более широкий спектр параметров.

Методические алгоритмы в схемах прогноза

Одним из важных факторов качественного решения является задание первоначального списка предикторов. Они должны отражать физические связи с предиктантом, и слабо коррелировать между собой. В нашем случае список параметров-предикторов ограничен гидродинамической продукцией, представленной в узлах регулярной сетки. А потребителей интересует ожидаемая погода в конкретных пунктах, то есть в нерегулярных точках. Интерполяция в данном случае малоэффективна или невозможна вовсе для получения отсутствующих в выходной модельной продукции требуемых элементов погоды. Задача решается в рамках обучающих статистических алгоритмов.

Самыми распространенными в метеорологии аппаратами для построения решений являются корреляционный, регрессионный, дискриминантный анализы с модификациями. Однако современные вычислительные возможности позволяют применять алгоритмы, которые не только заменяют предварительный глубокий анализ информативности тех или иных предикторов, но и способствуют более устойчивым решениям. Одним из таких алгоритмов является метод группового учета аргументов (МГУА), используемый в последних разработках СибНИГМИ.

МГУА построен на принципе самоорганизации, суть которого состоит в том, что при переборе аргументов и при постепенном усложнении априорных статистических моделей некоторые критерии проходят свой минимум. Алгоритм определяет этот минимум и таким образом указывает на оптимальную модель. Принципиальным отличием МГУА является привлечение внешних критериев. В исследованиях, посвященных прогнозу или восстановлению тех или иных метеорологических полей, до настоящего времени незаслуженно редко используются индуктивные методы математического моделирования, основанные на принципе самоорганизации моделей по внешним критериям. Критерий среднеквадратической ошибки, используемый в методе регрессионного анализа, по определению является внутренним, поскольку рассчитан по той же информации, по которой получена сама модель. В [5] показано, что любой внутренний критерий приводит к ложному правилу: чем сложнее модель - тем она точнее. Согласно теореме неполноты Геделя из математической логики только внешние критерии позволяют выбрать единственную модель оптимальной сложности.

Итак, в методе группового учета аргументов осуществляется целенаправленный перебор постепенно усложняющихся структур моделей и их отбор по ряду целесообразных эвристических критериев. При этом по основной части исходных фактических данных

получаются коэффициенты моделей- претендентов, а по другой - оценки этих моделей по выбранным критериям селекции.

Из предложенных в [6] внешних критериев селекции нами использованы регулярность и минимум смещения. В качестве критерия регулярности служит относительная среднеквадратическая ошибка

$$D^2(B) = \frac{\sum_{i=1}^{N_B} (q_i - q_i^*)^2}{\sum_{i=1}^{N_B} (q_i)^2} \rightarrow \min, \quad (1)$$

где q_i и q_i^* - полученное по модели и действительное значение выходной величины в i -той точке, N_B - число точек проверочной последовательности.

Чем меньше величина относительной среднеквадратической ошибки, тем модель более регулярна.

При большом уровне шумов в исходных данных более устойчивыми к помехам являются различные формы критерия минимума смещения. Суть этого критерия состоит в том, что модели, построенные на разных частях исходной последовательности, должны как можно меньше отличаться друг от друга. Сначала первая последовательность данных является обучающей, а вторая – проверочной (решения $q_i(A)$). Затем, наоборот, вторая является проверочной, а первая – обучающей (решения $q_i(B)$). Близость этих уравнений оценивается по величине среднеквадратического расхождения их выходов по всем точкам последовательности:

$$n_{sm}^2 = \frac{\sum_{i=1}^N [q_i(A) - q_i(B)]^2}{\sum_{i=1}^N q_{ifakt}^2}, \quad (2)$$

Таким образом, основным достоинством метода группового учета аргументов является построение непротиворечивых моделей, то есть дающих устойчивый результат независимо от исходных данных. Если к этому прибавить оптимальность по структуре и по множеству учитываемых переменных в решающих уравнениях, то применение МГУА к метеорологическим данным с целью получения прогнозов разной заблаговременности должно дать более надежный результат, чем регрессионный подход.

Однако не всегда прогнозируемый метеоэлемент имеет количественные значения. В последние годы особое внимание уделяется прогнозам опасных, а значит достаточно редких явлений. В этих случаях эффективнее использовать аппараты распознавания образов. В разработках СибНИГМИ для выделения ситуаций с резкими изменениями температуры воздуха, для кластеризации осадков, для определения условий для заморозков, сильного ветра, повышенного загрязнения воздуха и других успешно применяется так называемый DW-алгоритм, разработанный в Институте математики СО РАН СССР [7]. Алгоритм предполагает на входе генеральную совокупность признаков, соответствующих анализируемым объектам (ситуациям), которые в свою очередь относятся к первому или второму образу (классу). Необходимо на основе обучающей матрицы построить в классе логических деревьев решающее правило, по которому каждому вектору информативных признаков соответствует образ 1 или образ 2. Алгоритм перебора при построении дерева решений основан на принципе "лучший к лучшему". На первом шаге строится лучшее согласно некоторому заданному критерию разделение совокупности на две группы (ветки). Затем процесс повторяется уже для каждой полученной группы объектов и т.д.. В качестве условий останова выступают либо ограничение количества вершин дерева, либо минимально допустимое число объектов на ветке. Отметим, что выбор критерия

сбалансированного разделения ветвей дерева является ключевым моментом алгоритма, поэтому он корректируется исследователем согласно поставленной задаче. Реализованный на ПЭВМ алгоритм достаточно быстро выдает готовое решение. Это позволяет произвести большое количество экспериментов, меняя вид и количество определяющих признаков и критерии для разделения образов.

Технологическая линия прогноза

Методические программные модули являются лишь частью всей системы получения автоматизированного прогноза. Полная технологическая линия включает следующие подсистемы (рисунок 1):

- прием и обработка метеорологических данных оперативных наблюдений;
- прием гидродинамических прогностических метеополей из ведущих мировых центров погоды;
- расчет объективных (статистических) прогнозов явлений и элементов погоды;
- система доведения результатов до потребителей;
- система оценок качества прогнозов.

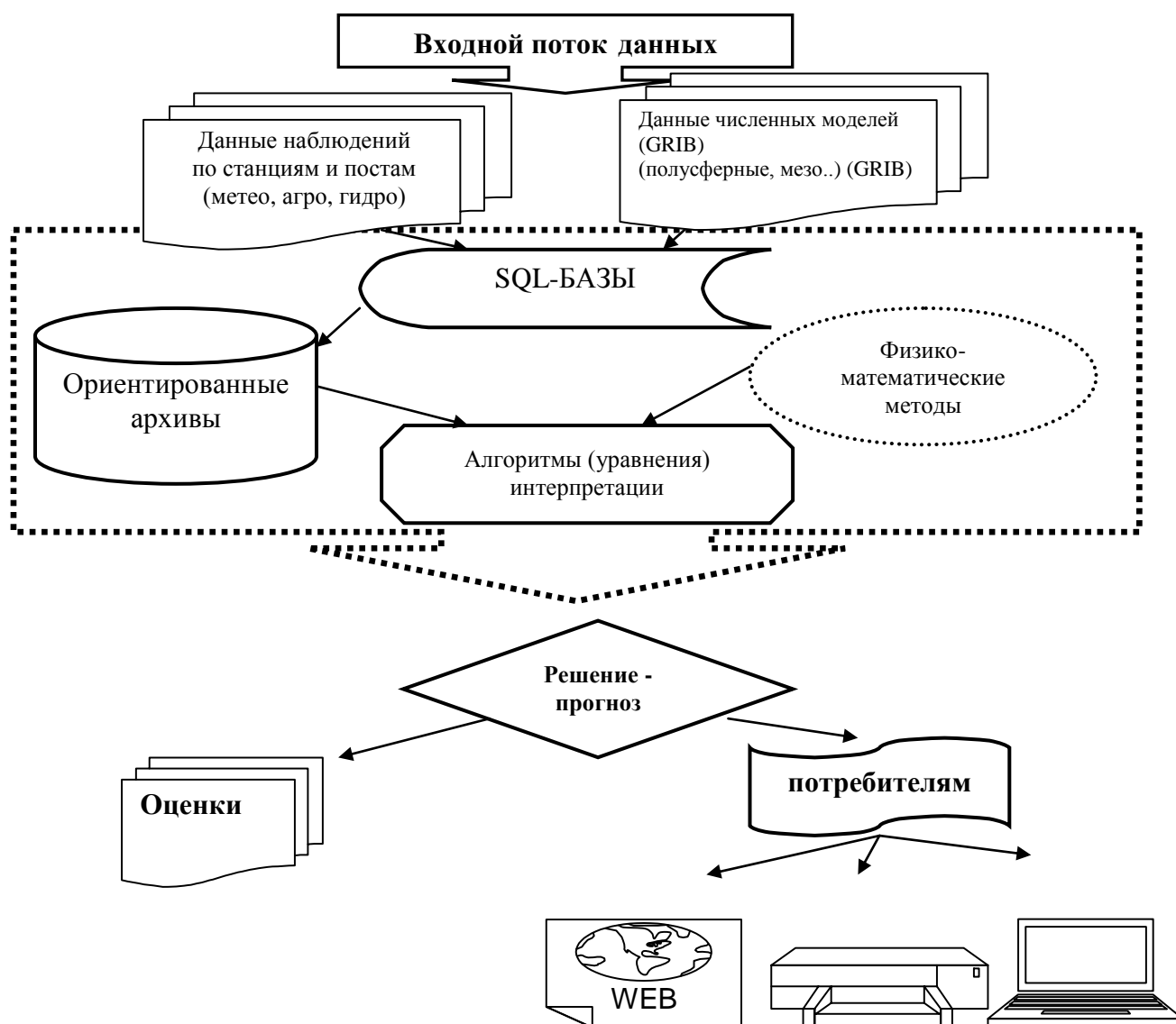


Рис.1 Технологическая линия вычислительного процесса получения прогнозов погоды в Новосибирске

В Новосибирском Центре по гидрометеорологии и в СибНИГМИ подготовлены и внедрены программные модули для всех блоков расчета. Существует несколько разработанных систем сбора и хранения метеоданных, в том числе SQL-база СибНИГМИ. Исходные данные записываются в базу, проходят первичную обработку. На этом этапе подключены алгоритмы контроля и восстановления, разработанные с учетом специфики метеопараметра. Проверенные данные одновременно поступают в блок расчета прогнозов явлений и элементов погоды и пополняют текущие архивные выборки и передаются в систему визуализации.

Визуализация данных реализована при помощи Гис-ресурса «Погода в реальном времени», разработанного в СИБНИГМИ. Блок-схема системы визуализации данных показана на рис.2. Система построена по технологии динамической генерации веб-страниц «на лету» и реализована на базе модуля mod-perl к популярному серверу Apache. Графический контент создается при помощи известной библиотеки манипулирования графикой ImageMagick. Вся серверная часть функционирует под управлением ОС Linux.

Интерфейсы построены по принципу вложенных страниц с системами горизонтальных и вертикальных меню (рисунок 3). Прогнозы погоды отображаются в соответствии с профилем доступа текущего зарегистрированного пользователя. Комплекс программ позволяет не только работать в интерактивном режиме, но и создавать автоматические рассылки данных клиентам по электронной почте и коротких текстовых сообщений по GSM сетям сотовой связи.

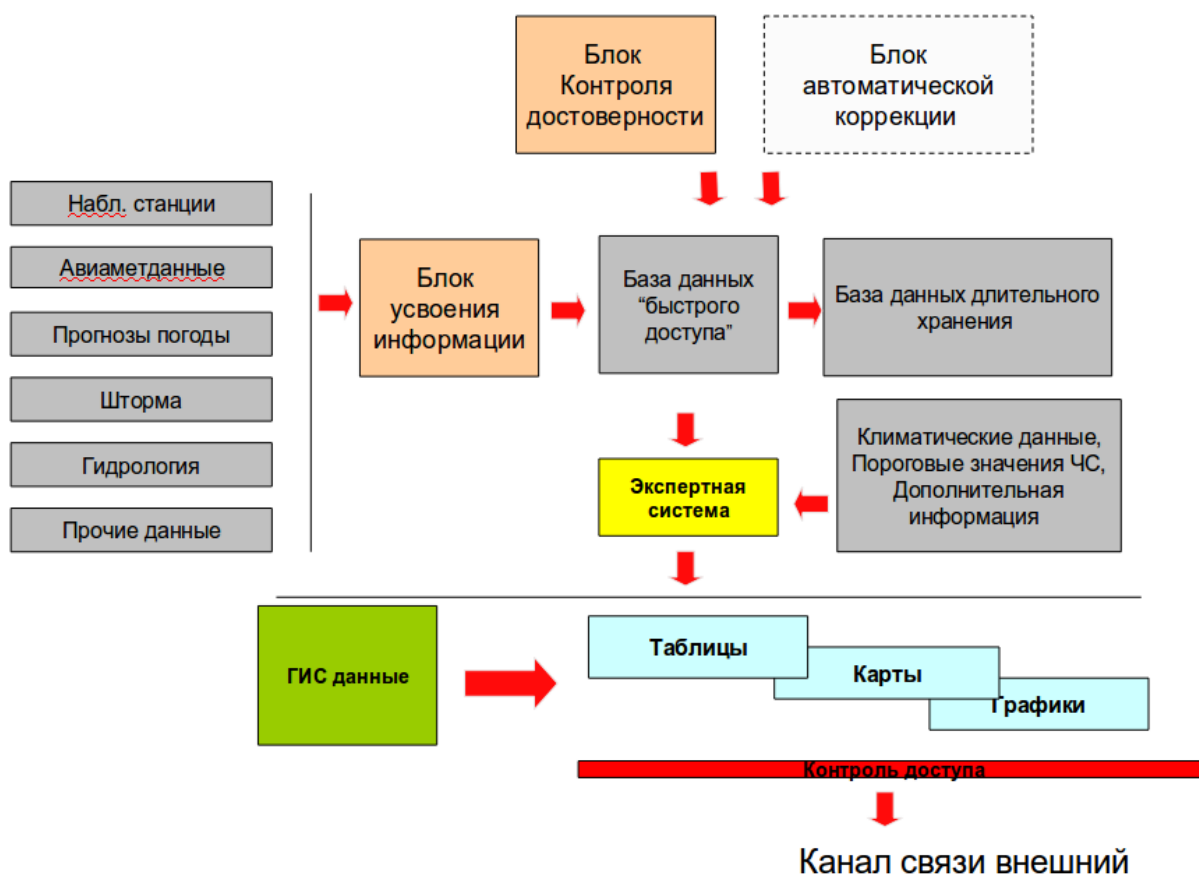


Рис.2. Блок-схема системы визуализации



а)

прогноз горимости, температуры, осадков

б)

Дата	Класс по Нестерову	Температура С	Вероятные осадки мм
2010-10-05 00:00:00	1	+5	1
2010-10-06 00:00:00	1	+4	0
2010-10-07 00:00:00	1	+4	1
2010-10-08 00:00:00	1	+8	3
2010-10-09 00:00:00	1	+9	1

Станция фактическая погода:

Данные за 5 октября - 13:0

Видимость: 4км
 Температура: 3.9 С растет
 Ветер: 3 4 м/с
 Атм. давл: 748 мм падает
 Текущие явления: дождь слабый
 Явления за прошлые 3 ч.: морось.

Количество выпавших осадков посуточно
 Суммарное количество осадков за месяц
 Высота снежного покрова

Данные за 10 дней:

Ход температуры за 10 дней
 Осадки за 10 дней

Справка погоды:

январь

Рис.3 Изображение прогноза температуры и пожароопасности (а) по территории Западной Сибири и пример всплывающего окна с прогнозом по станции (б)

Список литературы

1. Храмцова И.Г., Шустова Г.А. Опыт использования в оперативных условиях модели прогноза экстремальной температуры на 1 - 6 суток // Труды ЗапСибНИГМИ. - 1990. - Вып. 93. - С. 106 - 112.
2. Храмцова И.Г. и др. Модель прогноза осадков на 1 - 5 дней в тёплое полугодие по территории Урала и Сибири // Труды СибНИГМИ, 1992, вып. 100, С. 167 - 178.
3. 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
4. Здерева М.Я., Виноградова М.В. Среднесрочный прогноз степени пожарной опасности в лесах по метеорологическим условиям. // Метеорология и гидрология.-2009.-№1.- С.16-26.
5. Ивахненко А.Г., Юрачковский Ю.П. Моделирование сложных систем по экспериментальным данным.- М.: "Радио и связь", 1987. -120с.
6. Ивахненко А.Г., Коппа Ю.В., Степашко В.С. и др. Справочник по типовым программам моделирования. - Киев:Теника, 1980. - 184с.
7. Манохин А.Н. Алгоритм DW для распознавания образов: Пакет прикладных программ ОТЭКС.-Новосибирск:изд-во Новосибирского государственного университета, 1981.-С.3-30.

Аннотация.

ИСПОЛЬЗОВАНИЕ ПРОСТРАНСТВЕННЫХ МЕТЕОРОЛОГИЧЕСКИХ ДАННЫХ В ПРОГНОСТИЧЕСКИХ МЕТОДАХ И ПРИМЕР ПРИМЕНЕНИЯ WEB-ТЕХНОЛОГИЙ ДЛЯ ПРЕДСТАВЛЕНИЯ РЕЗУЛЬТАТОВ

М.Я. Здерева,¹ А.Б.Колкер,¹ В.М.Токарев.¹

¹ Сибирский региональный научно-исследовательский гидрометеорологический институт, (ГУ СибНИГМИ) Новосибирск

В прогностических метеорологических схемах широко используются пространственные данные как в регулярной, так и в нерегулярной географической сетке. На примере физико-статистических разработок Сибирского гидрометеорологического института для прогнозов элементов погоды на средние сроки рассмотрены некоторые алгоритмические примеры синтеза пространственных данных. Новые прогностические методы являются частью WEB- технологии автоматизированного расчета и представления результатов. Технология включает обработку поступающих в оперативном режиме модельных и фактических данных, блок расчетов прогнозов по разработанным методам, вывод результатов во внутреннюю и внешнюю сеть, оформление для представления в автоматизированной информационной системе (АИС) «Погода в реальном времени».

USING THE SPATIAL METEOROLOGICAL DATA IN PROGNOSTIC METHODS AND EXAMPLE OF WEB TECHNOLOGIES APPLICATION IN RESULT PRESENTATION

M.Ya.Zdereva, A.B.Kolker, V.M.Tokarev,

Siberian Regional Research Institute Hydrometeorological Institute, Novosibirsk

Spatial data, both in regular and in irregular geographical mesh, are widely used in prognostic meteorological schemes. Algorithmic synthesis of the spatial data is considered with an example of physical/statistical developments in medium-range weather forecast performed in the Siberian Hydrometeorological Institute. New prognostic methods are the part of WEB technology for computerized calculation and result presentation. Technology includes processing of on-line model and actual data, block of forecast calculation in accordance with developed methods, output of results to internal and external networks, and formalization for presentation in the automatic information system "Real-Time Weather".