Министерство природных ресурсов и экологии Российской Федерации

Федеральная служба по гидрометеорологии и мониторингу окружающей среды

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ   
“ДАЛЬНЕВОСТОЧНЫЙ РЕГИОНАЛЬНЫЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ ИНСТИТУТ”  
(ФГБУ “ДВНИГМИ”)

«УТВЕРЖДАЮ»

Диpектоp ДВНИГМИ

канд. физ.-мат. наук

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Ю.Н. Волков

ОТЧЕТ

О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

ПО ТЕМЕ ВНУТРЕННЕГО ПЛАНА

РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ И ТЕХНОЛОГИИ ОПЕРАТИВНОГО ПРОГНОЗА МЕТЕОРОЛОГИЧЕСКИХ ПАРАМЕТРОВ НА СТАНЦИЯХ ПРИМОРСКОГО КРАЯ НА ОСНОВЕ ПОДХОДА MOSИ ВЫХОДНЫХ ДАННЫХ МОДЕЛИ WRF (ПУГМС).

Зам. диpектоpа по научной pаботе

канд. геогр. наук Е.В. Каpасёв

Исполнитель

вед. инженер ОДПП Л.В. Гончуков

Владивосток 2014

Содержание

Введение 3

Схема данных 4

Методология MOS 5

Построение регрессий и отбор предикторов 9

Модель и потенциальные предикторы 10

Составление прогнозов 18

Процедура оценки 19

Заключение 20

Приложение 1. Параметры конфигурации модели 21

Список использованных источников 23

# Введение

Использование результатов расчета численных моделей не дает автоматического роста качества прогнозов в подразделениях УГМС. Еслиоставить результаты расчетов необработанными и по их результатам составить прогноз элементов погоды (экстремальная температура воздуха на высоте 2 м., максимальная скорость ветра, преобладающее направление ветра, сумма осадков за полусутки) согласно наставлению[1], то оправдываемость их окажется на уровне 60-70%, что явно не может удовлетворить ни синоптиков, ни потребителей.

Поэтому первоочередной задачей в деле модернизации гидрометеорологического производства является улучшение качества прогнозов элементов погоды, производимых с помощью автоматизированных систем. Наибольшее распространение в этом направлении получили методы физико-статистической интерпретации результатов расчета численных моделей[2]: метод идеального прогноза (PP, perfectprog)[3] и статистическая интерпретация гидродинамического прогноза (MOS, modeloutputstatistics)[4,5] в основе которых лежит предположение, что поля воспроизводимые моделями и данные наблюдений связаны линейной регрессией.

Целью работ по теме является разработка технологии и автоматизированной информационной системы составления прогнозов элементов погоды на метеостанциях Приморского края по методологии MOS для результатов оперативного расчета модели WRF, использующейся в Приморском УГМС.

# Схема данных

При разработке схемы хранения данных мы должны исходить из задач решаемых системой:

1. Построение регрессионных зависимостей между предикторами и предиктантами;
2. Подготовка прогнозов предиктантов на основании ранее полученных зависимостей;
3. Оценка успешности полученных прогнозов.

Исходя из этого схема должна предусматривать хранение следующих банков данных:

1. Массив значений потенциальных предикторов для построения регрессий;
2. Массив значений предиктантов для построения регрессий;
3. Массив регрессионных зависимостей;
4. Массив значений предикторов для подготовки прогнозов;
5. Массив значений подготовленных прогнозов;
6. Массив результатов наблюдений и измерений для проведения оценки успешности прогнозов;
7. Массив оценок успешности.

Кроме того, в схеме должны быть предусмотрены отношения для хранения описаний предикторов и предиктантов, пунктов прогнозирования и прочих справочников.

Поскольку в разрабатываемой технологии в качестве **предиктантов** выступают значения метеоэлементов, измеряемые на метеостанциях в стандартные сроки, логичным выглядит объединение массива результатов наблюдений и измерений с массивом значений предикторов, что должно позволить в будущем вовлекать поступающие данные для актуализации регрессионных зависимостей. Аналогично можно объединить и массивы **предикторов**, однако необходимо помнить, что выборки значений предикторов будут происходить при каждом составлении прогноза и большое количество записей может способствовать замедлению процесса.

Для описания **регрессионных взаимосвязей** необходимо, чтобы схема позволяла для каждого предиктанта однозначно сопоставить набор предикторов с соответствующими мультипликативными и степенными коэффициентами. Регрессии строятся применительно к совокупности предиктанта, пункта и заблаговременности прогноза, т.е. для каждого предиктанта, для каждого пункта и для каждой заблаговременности должны быть построены свои зависимости. Также схема должна предусматривать возможность хранения **нескольких версий преобразований** дляодногоитогоженаборапредиктанта, пунктаизаблаговременности. Поддержкаверсионностинеобходимадлявозможностираздельнойоценкирезультатоввслучае, когдаразныепроцедурыпостроениярегрессиивыдаютразныевариантыпреобразований.Схемадолжнапредусматриватьвозможностьпримененияпроцедур**классификации**ситуациииобеспечиватьхранениеотдельныхрегрессионныхзависимостейдлякаждогокласса.

Результатом применения выявленных регрессионных зависимостей к массиву значений предикторов является **прогноз** значения метеоэлемента в заданной точке, отнесенный к определенному времени. Так как прогноз может быть составлен с помощью различных версий регрессий, необходимо однозначно сопоставлять полученный прогноз с конкретной версией преобразования. Таким образом, в схеме данных прогноз будет характеризоваться значением, временем актуальности исходных данных, временем актуальности спрогнозированного значения и описанием того преобразования, с помощью которого он был получен. Имея описание преобразования, можно однозначно установить для какого пункта, какого метеоэлемента и с какой заблаговременностью был составлен прогноз.

Для **оценки успешности** прогноза необходимо сравнить полученное в результате преобразований значение с реально измеренным или наблюденным и измерить разницу между ними. Каждому полученному прогнозу в соответствие ставится отклонение от фактического значения (ошибка) и показатель успешности в процентах. Набора таких данных должно быть достаточно для определения успешности полученных статистических зависимостей на длительном промежутке.

# Методология MOS

Гидродинамическиемодели атмосферыизначально,всилуограниченностивычислительныхресурсовинеполнотыматематическогоописания,применялисьдляпрогнозасиноптическойситуации (общейциркуляции. Исследуястатистические связи между переменными ОА и фактически наблюдавшимися погодными условиями, Клаин и др. [3] в 1959 году обнаружили устойчивую статистическую связь между осреднённой за 5 дней фактически измеренной приземной температурой и осреднённым за 5 дней полем H700. Применение полученных уравнений регрессии для прогностических полей H700показало хороший прогноз среднепентадной приземной температуры. Так была предложена концепция «идеального прогноза» (PerfectProg.,PP). Согласно ей, найденные статистические связи между измеренными значениями различных метеоэлементов (выражение одного предиктанта путём линейной комбинации предикторов) применяются к результатам расчёта гидродинамической модели для составления прогноза. В качестве предикторов должны использоваться данные наблюдений, полученных либо прямыми измерениями (инструментальными), либо опосредованными (из результатов объективного анализа).

Схожая методика позднее была Гланом и Лори[4,5], только в качестве предикторов использовались не данные фактических наблюдений, а результаты расчёта модели. Подход был назван MOS – modeloutputstatistics.

На Рисунке 1 изображены схемы двух версий MOS: на одной и предиктор и предиктант берутся в один и тот же момент времени (MOS1), а на второй (MOS2) допускается расхождение между временем актуальности предиктора (t) и действительным временем предиктанта (T). Левая часть рисунка относится к стадии построения уравнений регрессии, а на правой изображен способ, по которому полученные уравнения регрессии используются для получения прогноза. Чтобы разобраться какая из версий MOS лучше, разумно предположить, что лучшим предиктором, объясняющим yo, является не значение xm, рассчитанное моделью, а наблюденное xo, так как только в лучшем случае смоделированное значение xm может оказаться равным xo. Другими словами xm=xoтолько в том случае, когда модель является идеальной. Поэтому, в случае идеальной модели, применение MOS1 даст лучший результат, нежели MOS2, так как в каждый момент времени лучшим предиктором для yявляется xo взятый в тот же момент времени.

Однако идеальных моделей не существует, и xm согласуется с xo только в начальный момент времени, после чего будет наблюдаться расхождение, вызванное не только неидеальностью модели, но также и хаотичностью самой атмосферы. Это расхождение, в свою очередь, приведет к дальнейшему расхождению между xoи y, которое будет увеличиваться во времени. В таком случае, если искажения информации, вызванные несовершенной природой моделей, превышают потери вызванные хаотичностью атмосферы, то качество MOS2 превзойдет качество MOS1.

Исходя из того, что в качестве основных причин искажения информации при составлении прогноза рассматриваетсяа) несовершенство модели и б) природа атмосферы, логично разделить учёт их вклада в общий результат. Марзбан[2]предложил следующую схему. На первом этапе можно разработать уравнение регрессии для перевода «предикторов из модели» в «предикторы из наблюдений», что должно элиминировать как вклад самой модели, так и её систем усвоения данных наблюдений. А на втором этапе привлечь уже имеющиеся уравнения для перевода «предикторов из наблюдений» в «предиктанты из наблюдений». Этот перевод уже не зависит от каких-либо моделей и отражает физическую сущность атмосферы, кроме того здесь можно устранить инструментальные ошибки измерений. Такой метод можно считать гибридом MOS и PP, поскольку и наблюдённые значения предиктора и смоделированные участвуют в построении прогноза для предиктанта. Однако у этого метода есть один существенный недостаток. При таком подходе предиктор и предиктант не могут являться одной и той же величиной, так как на втором шаге необходимо будет перевести её саму на себя. Уйти от этого недостатка можно, заменив наблюдённое значение значением, полученным из реанализа (см. Рисунок 2). Так как реанализ строится по результатам наблюдений, то можно считать, что информация об измерениях не теряется.

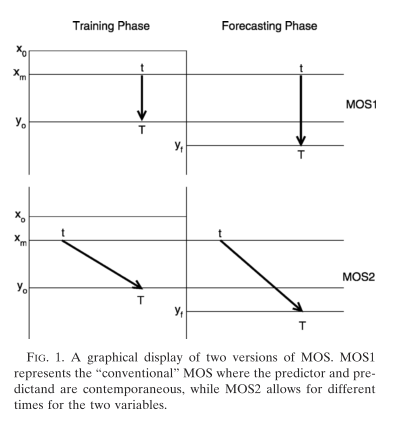


Рисунок 1 Различные версии MOS

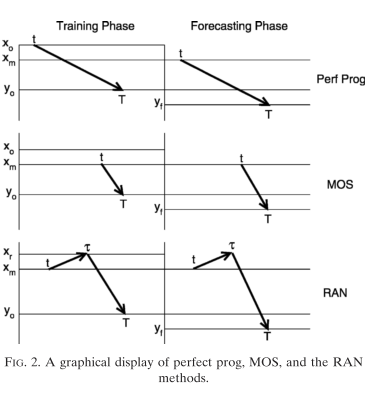


Рисунок 2Сравнение PP, MOSи RAN

Все вышеприведённые методы объединяет необходимость определения процедуры перевода (трансляции) предикторов в предиктант. Исторически первыми методами такого перевода были методы множественной линейной регрессии.Множественная линейная регрессия связывает зависимую переменную (предиктант) *Y* с *k* независимых переменных (предикторов) *Xi* Результатом является уравнение, с помощью которого можно оценить предиктант, как линейную комбинацию предикторов:

являетсяоценкой*Y,*а*ai* -- константыикоэффициентырегрессии*ai* определяются таким образом, чтобы сумма квадратов отклонений оценки являлась минимальной на обучающей (зависимой) выборке.

В реальности отдельные привлечённые предикторы (или даже все) могут быть бинарными, т.е. принимать значения 0 или 1.Если предположить, что предиктант может иметь только два состояния, тогда его можно оценивать, сопоставив «0» с первым состоянием, а «1» -- со вторым. Получаемую с помощью уравнения регрессии оценку можно рассматривать как вероятность получения второго состояния при заданной комбинации предикторов.

Если же предиктант может быть только в одном из нескольких (*K*) состояний, то его можно преобразовать в *K* бинарных предиктантов и интерпретировать их, как описывалось выше. Однако следует отметить, что во всех *K* регрессионных уравнениях следует использовать один и тот же набор предикторов, а также сумма всех *K* оценок должна быть равна 1. Для бинарных предиктантов рекомендуется использовать и бинарные предикторы[3].

В[6] для вероятностных оценок наступления тех или иных событий предлагается использовать модель логит-регрессии. Для бинарного предиктанта уравнение принимает довольно простой вид:

где

Либо

Также для вероятностных прогнозов широкое распространение получили методы дискриминантного анализа.

Вопрос определения количества и выбора предикторов является основным при построении систем MOS.Несмотря на то, что предиктант может коррелировать с сотнями предикторов, в уравнение регрессии включают далеко не все из них. Обусловлено это в первую очередь их высокой взаимной корреляцией. Также, если включить большое количество предикторов, можно получить достаточно близкие оценки предиктанта на обучающей выборке. Однако в полученных уравнениях будет отражаться не только физические взаимосвязи, но также и статистические характеристики самой обучающей выборки, которые не останутся в прежнем виде для других наборов данных. Однако значения различных метеорологических элементов, привлекаемых в качестве предикторов, не являются независимым набором данных. Для них характерна довольно высокая корреляция, как в пространственных измерениях, так и во времени. Кроме того, довольно высокая взаимная корреляция между предикторами, а также в некоторых случаях наличие существенной автокорреляции, делают непростой задачу определения критерия значимости для отбора предикторов. В качестве одного из критериев предлагается [7] использовать F-критерий, который может быть вычислен по формуле:

где -- квадрат коэффициента множественной корреляции, вычисленный при включённых *k* предикторах, *n* -- длина выборки. Критические значения для F-критерия берутся из стандартных таблиц критерия Фишера для степеней свободы 1 и *(n-k-*1*)*. Если F является значимым с заданным уровнем достоверности, то *k-*й предиктор включается в модель, и процедура отбора продолжается. В противном случае процедура прекращается, и уравнение регрессии получается с использованием только *k-*1 предикторов.

Другим критерием значимости, как уже указывалось, может служить уменьшение дисперсии на каждом шаге, которое может быть записано как . Критические значения обычно выбираются в пределах от 3 до 0,1 %. Ещё более простым критерием может служить количество предикторов включённых в конечное уравнение. Обычно в уравнения включается от 2-х до 30-ти предикторов, но рекомендуется, чтобы их количество не превышало 10-ти процентов от длины выборки.

# Построение регрессий и отбор предикторов

Построение регрессионных зависимостей решается с применением библиотеки методов линейной регрессии ALGLIB[8]. Предварительно проводится процедура синхронизации значений предикторов и предиктанта. В качестве исходных данных для построения модели выступает матрица из Nстрок и (M+1)столбцов, где N–количество случаев, M – количество предикторов. Для каждого случая в первые M столбцов записываются значения соответствующего предиктора, а в (M+1)-й значения предиктанта. Пропуски в данных заполняются средним значением соответствующего предиктора. Результатом работы алгоритма является регрессионная модель, которая содержит в себе коэффициенты линейного преобразования для каждого предиктора, матрицу ковариации, значения средней, средней относительной и среднеквадратичной ошибок преобразования.

Для отбора значимых предикторов реализовано два метода: прямого отбора и обратного исключения. В методе прямого отбора на первом шаге для вышеописанной матрицы Nx(M+1) строится матрица корреляции и выбирается предиктор, наиболее коррелирующий с предиктантом (независимо от знака).Условием включения предиктора является удовлетворение значения коэффициента взаимной корреляции критерию значимости Стьюдента с уровнем 95%. Кроме того, отобранный предиктор не должен коррелировать ни с одним из уже включенных в схему предикторов с тем же уровнем значимости.

После включения в схему производится построение регрессионной модели для отобранного предиктора и предиктанта. Полученная модель применяется к соответствующему предиктору, результатом чего является массив из N значений оценок предиктанта. Поэлементно вычитая из массива оценок массив значений предиктанта, получаем массив ошибок. Есливариация результата расчета уменьшается в сравнении с исходной вариацией предиктанта, то отобранный предиктор оставляют в модели и повторяют процедуру уже для (M-2) предикторов и массива ошибок оценки, полученной на предыдущем этапе. Итерации продолжают до тех пор, пока уменьшение вариации на не станет меньше порогового значения, заданного на уровне 5%. Также установлено ограничение на общее количество включенных предикторов в размере 10% от размера обучающей выборки, но не более 10 шт.

Метод обратного исключения подобен предыдущему с той лишь разницей, что изначально в модель включены все предикторы и затем из нее исключаются наименее коррелирующие с предиктантом. Здесь итерации продолжаются, пока уменьшение вариации не станет больше порогового значения.

# Модель и потенциальные предикторы

Модель WRF(WeatherResearchandForecasting) – численная мезомасштабная система прогнозирования погоды следующего поколения, разработанная как для оперативного прогнозирования, так и для исследовательских целей. Ее отличительными чертами являются несколько динамических ядер, 3-х мерная система усвоения данных и программная архитектура, нацеленная на параллельность вычислений и расширяемость системы. WRF может использоваться в широком спектре приложений на масштабах от нескольких метров до тысяч километров.

В модели WRF численно решается полная система уравнений гидротермодинамики атмосферы, которая включает в себя уравнения неразрывности, переноса импульса (с учетом эффектов сжимаемости и негидростатичности), переноса влаги и внутренней энергии. Для модели предлагается библиотека различных параметризаций процессов в пограничном слое атмосферы, формирования облачности и осадков, явлений конвекции и турбулентности. В модели реализованы возможности создавать вложенные вычислительные области и использовать алгоритмы усвоения данных наблюдений. В WRF используется σ-система вертикальных координат, которая позволяет описывать сложный рельеф, и сетка Аракавы C-типа по горизонтали (то есть объемная ячейка задается прямоугольным параллелепипедом; все скалярные величины задаются в его центре, а компоненты векторов в центрах советующих граней). При интегрировании уравнений гидродинамики используется расщепление по направлениям (в горизонтальных направлениях используется явная схема, а при вычислении вертикальных потоков и скоростей – неявная) и по физическим процессам. WRF является региональной моделью, поэтому для ее инициализации и задания граничных условий необходимы результаты расчетов глобальной модели численного прогноза погоды. Цикл работы модели WRF состоит из подготовки данных, включая их горизонтальную и вертикальную интерполяцию на сетку модели, улучшение интерполированных данных с помощью усвоения наблюдений метеостанций и радиозондов и численное интегрирование.

В нашем случае расчет проводился по расчетной области с шагом сетки 5 км на 27-ми вертикальных уровнях. Для решения задачи метеорологического прогноза система уравнений гидротермодинамики атмосферы должна быть дополнена начальными и граничными условиями. В качестве начальных условий должны задаваться трехмерные распределения полей трех компонент скорости, температуры, давления и влажности. В качестве граничных условий должны задаваться значения температуры, влажности и компонент скорости на боковых и на верхней границах области, а также значения потоков тепла, влаги и импульса на нижней границе области, прилегающей к поверхности Земли. В нашем случае для этих целей используются данные оперативных прогнозов Национального центра прогнозирования атмосферы NCEP (США). Прогностические поля модели NCEP сроком на 72 ч. прогнозного времени определяют начальные и граничные условия для метеорологических переменных. Граничные условия на подстилающей поверхности зависят от свойств подстилающей поверхности. Используются следующие данные о подстилающей поверхности: высота поверхности над уровнем моря, категория подстилающей поверхности в соответствии с классификацией Геологической службы США USGS (например, лес, водная поверхность, и т.д.), тип растительности, тип почвы, средняя сезонная температура, среднемесячное альбедо. Соответствующие данные брались с сайта USGS.

Основной задачей при составлении регрессионных зависимостей для прогноза элементов погоды является выбор величин (переменных), используемых в качестве предикторов. Их условно можно разделить на несколько групп: расчётные переменные модели, интерполированные переменные модели, производные переменные, геоклиматические и данные наблюдений. В первую группу входят величины, непосредственно рассчитанные моделью в её системе координат. Туда же можно отнести величины, полученные из непосредственно рассчитанных с помощью преобразований по физическим законам (температура в градусах Кельвина, относительная влажность и пр.). Во вторую группу входят величины, полученные из первой путём математической интерполяции в другие системы координат (вертикальные уровни, проекции). К третьей группе относятся все величины, полученные из первых двух групп путём математических преобразований (градиенты, лапласианы, дивергенции, роторы и пр.). Логично предположить, что чем меньше математических преобразований было произведено над рассчитанной величиной, тем в более согласованном состоянии она находится по отношению к самой себе и другим переменным в пространстве и времени.

Рассмотрим некоторые примеры отобранных предикторов. В [9] для построения MOS для температуры, точки росы, относительной влажности и скорости ветра на основе модели MM5 были использованы предикторы всех вышеуказанных типов (см. Таблица 1).

Таблица 1. Вариант перечня использовавшихся предикторов для прогноза температуры, точки росы, относительной влажности и ветра

|  |  |  |
| --- | --- | --- |
| Предиктор | Уровни | Кол-во |
| T, Td, RH, U, V, WSPD | σ=0.995; 800, 700, 600, 500 hPa | 30 |
| H, Verticalvelocity | 800, 700, 600, 500 hPa | 8 |
| 3h acc. precip. |  | 1 |
| Sfc. -700 hPa lapse rate |  | 1 |
| Avg. sfc. – 500hPa RH |  | 1 |
| Sin, Cos дня года |  | 2 |
| Наблюдения за ближайший срок |  | 4 |
|  |  | 47 |

В [10]дляпрогнозатемпературыиспользовалисьпредикторыизТаблицы 2.

Таблица2.Вариантперечняиспользовавшихсяпредикторов для прогноза температуры

| Предиктор | Уровни | Кол-во |
| --- | --- | --- |
| totalcloudcover |  | 1 |
| T | surface, 1000, 925, and 850 hPa | 4 |
| Tdd | 1000 hPa | 1 |
| geopotentialheight | 850 and700 hPa | 2 |
| wind speed | 850hPa | 1 |
| temperature advection | 1000hPa | 1 |
| RH | 925, 950, 700, and 500 hPa | 4 |
| OT | (1000–850) and (1000–925) hPa | 2 |
| George K index |  | 1 |
| sun elevation function |  | 1 |
| latest observation |  | 1 |
|  |  | 19 |

Тамжедлясоставленияпрогнозаветрасоставлялосьтрирегрессионныхмодели (дведлякомпонентвектораскоростииоднадляскаляраскорости) с использованием предикторов из Таблицы 3. Направлениеветрарассчитываласьизполученныхпрогнозовкомпонентвектора

Таблица 3. Вариант перечня использовавшихся предикторов для прогноза ветра

| Предиктор | Уровни | Кол-во |
| --- | --- | --- |
| Laplacianofpressuretendency |  | 1 |
| E–Wtemperaturegradient | 1000 hPa | 1 |
| E–Wdewpointdepressiongradient | 925 hPa | 1 |
| N–S dewpoint depression gradient | 1000 hPa | 1 |
| U | surface, 1000, 925, 850, and 700 hPa | 5 |
| V | 1000, 925, 850, and 700 hPa | 4 |
| N–S shear of U | 700 hPa | 1 |
| E–W shear of V | 925 hPa | 1 |
| vorticity advection | surface and 925 hPa | 2 |
| Laplacian of temperature advection | 925 hPa | 1 |
| geopotential height | 850 hPa | 1 |
| geostrophic vorticity | 1000 and 925 hPa | 2 |
| orographic wind speed |  | 1 |
| wind speed | 1000, 925, 850, and 700 hPa | 4 |
| temperature advection | 1000 hPa | 1 |
| Geostrophic U | 850 and 700 hPa | 2 |
| geostrophic wind speed | 1000, 925, and 850 hPa | 3 |
| observed wind speed, U, V |  | 3 |
|  |  | 35 |

Количественныйпрогнозосадков (КПО) досихпорявляетсянетривиальнойзадачей[11]. В качестве предиктантов для КПО используются количество выпавших осадков в пункте, разница между спрогнозированными и выпавшими осадками.

НарядусдетерминированнымКПОиспользуетсявероятностныйКПО. Детерминированный КПО, воспроизводя только одно значение, является стандартным полем в результатах расчёта численных моделей.

Вероятностныйпрогнозосадковдаётсяобычновтерминахвероятностивыпаденияколичестваосадковвышеопределённогокритическогозначения. Длякаждоготакогокритериядолжны строиться свои регрессионные модели. Основными способами построения регрессий являются дискриминантный анализ и логит-регрессия.

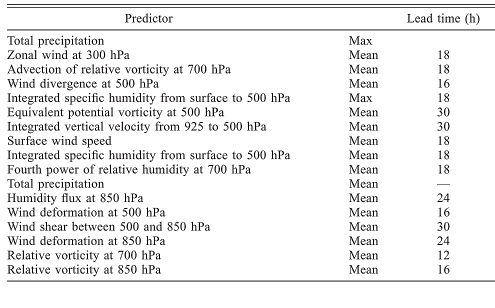
В качестве меры оценки качества вероятностных прогнозов предлагается [12,13]показатель успешности Брайера (Briersskillscore, BSS)и показатель Брайера (Briersscore).

BSпринимает значения от 0 (идеальный прогноз) до 1 (полностью неверный прогноз)

Относительнокритерияотборапредикторов. В[14]методомпробиошибокбылообнаружено, чтолучшеекачествомоделинанезависимойвыборке достигается, когда процесс отбора предикторов останавливается, если при добавлении очередногопредиктора BSSна обучающей выборке увеличивается менее чем на 5 %.

В качестве предикторов для прогноза осадков чаще берутся синтетические величины. Так в [11] используются предикторы, указанные в Таблице 4.

Таблица 4Вариант перечня использовавшихся предикторов для прогноза осадков



В [10]дляполучениявероятностивыпаденияосадковприменяли предикторы, причисленные в Таблице 5.

Таблица 5. Вариант перечня использовавшихся предикторов для прогноза вероятности выпадения осадков

| Предиктор | Уровни | Кол-во |
| --- | --- | --- |
| cloudcover |  | 1 |
| convectiveandtotalprecipitationrates |  | 2 |
| Laplacian of pressure tendency |  | 1 |
| E–W temperature gradient | 850 hPa | 1 |
| dewpoint depression | 700 hPa | 1 |
| N–S shear of U | 925, 850 and 700 hPa | 3 |
| E–W shear of V | 925 and 700 hPa | 2 |
| Laplacian of temperature advection | 925 and 500 hPa | 2 |
| vertical velocity | 1000, 925, 850, 700, and 500 hPa | 5 |
| geopotential height | 850 and 700 hPa | 2 |
| N–S height gradient | 1000 hPa | 1 |
| geostrophic vorticity | 1000 and 500hPa | 2 |
| surface wind speed; divergence | 500hPa | 2 |
| relative humidity | 925, 850, 700, and 500 hPa | 4 |
| vertical shear of V | 1000-850hPa | 1 |
| George K index |  | 2 |
| observed precipitation at analysis time (binary) |  | 1 |
| observed precipitation in last 6 h (binary) |  | 1 |
|  |  | 34 |

**В качестве потенциальных предикторов были выбраны следующие (см. Таблицу 6):**

Таблица 6 Предикторы, привлекаемые для разрабатываемой системы MOS

|  |  |  |
| --- | --- | --- |
| № | имя | Описание |
| 1 | advRh2 | Адвекция относительной влажности на высоте 2м |
| 2 | advRh850 | Адвекция относительной влажности на высоте 850 гПа |
| 3 | advT2 | Адвекция температуры на высоте 2м |
| 4 | advT850 | Адвекция температуры на высоте 850 гПа |
| 5 | advTdd2 | Адвекция дефицита точки росы на высоте 2м |
| 6 | advTdd850 | Адвекция дефицита точки росы на высоте 850 гПа |
| 7 | cape | Доступная потенциальная энергия конвекции |
| 8 | cin | Энергия подавления конвекции |
| 9 | defSh500 | Деформация сдвига поля ветра на высоте 500 гПа |
| 10 | defSt500 | Деформация сжатия поля ветра на высоте 500 гПа |
| 11 | defV500 | Полная деформация поля ветра на высоте 500 гПа |
| 12 | dudx500 | Частная производная компоненты скорости ветра на высоте 500 гПа |
| 13 | dudy500 | --//-- |
| 14 | dvdx500 | --//- |
| 15 | dvdy500 | --//- |
| 16 | lcl | Уровень конденсации |
| 17 | lfc | Уровень свободной конвекции |
| 18 | rainc | Количество жидких осадков из конвективных схем |
| 19 | rainnc | Кол-во жидких осадков |
| 20 | rh2 | Относительная влажность на высоте 2 м. |
| 21 | rh500 | Относительная влажность на высоте 500 гПа |
| 22 | rh850 | Относительная влажность на высоте 850 гПа |
| 23 | rh925 | Относительная влажность на высоте 925 гПа. |
| 24 | rot500 | Ротор поля ветра на высоте 500 гПа |
| 25 | slptend | Барическая тенденция на уровне моря |
| 26 | slp | Давление на уровне моря |
| 27 | snownc | Количество твердых осадков |
| 28 | t2 | Температура воздуха на высоте 2 м. |
| 29 | t500 | Температура воздуха на высоте 500 гПа |
| 30 | t850 | Температура воздуха на высоте 850 гПа |
| 31 | t925 | Температура воздуха на высоте 925 гПа. |
| 32 | td2 | Точка росы на высоте 2 м. |
| 33 | td500 | Точка росы на высоте 500 гПа |
| 34 | td850 | Точка росы на высоте 850 гПа |
| 35 | td925 | Точка росы на высоте 925 гПа. |
| 36 | tdd2 | Дефицит точки росы на высоте 2 м. |
| 37 | tdd500 | Дефицит точки росы на высоте 500 гПа |
| 38 | tdd850 | Дефицит точки росы на высоте 850 гПа |
| 39 | tdd925 | Дефицит точки росы на высоте 925 гПа. |
| 40 | totalprecip | Общее количество осадков |
| 41 | u10 | Зональная составляющая скорости ветра на высоте 10 м |
| 42 | u500 | Зональная составляющая скорости ветра на высоте 500 гПа |
| 43 | u850 | Зональная составляющая скорости ветра на высоте 850 гПа |
| 44 | v10 | Меридиональная составляющая скорости ветра на высоте 10 м |
| 45 | v500 | Меридиональная составляющая скорости ветра на высоте 500 гПа |
| 46 | v850 | Меридиональная составляющая скорости ветра на высоте 850 гПа |
| 47 | w500 | Вертикальная составляющая скорости ветра на высоте 500 гПа |
| 48 | w850 | Вертикальная составляющая скорости ветра на высоте 850 гПа |
| 49 | wdir10 | Направление ветра на высоте 10 м |
| 50 | wdir500 | Направление ветра на высоте 500 гПа |
| 51 | wdir850 | Направление ветра на высоте 850 гПа |
| 52 | wspd10 | Скорость ветра на высоте 10 м |
| 53 | wspd500 | Скорость ветра на высоте 500 гПа |
| 54 | wspd850 | Скорость ветра на высоте 850 гПа |

Потенциальные предикторы были получены путем обработки выходных файлов с результатами расчетов модели WRF за период 2005-2009 гг.Параметры применяемой конфигурации приведены в приложении 1. Кроме того, для каждого предиктора были рассчитаны минимум, максимум и среднее значение за полусутки (с 9 до 21 и с 21 до 9 часов ВСВ).

Общая выборка разбивалась на различные классы исходя из различных способов классификации текущей ситуации. Первая классификация опирается на календарный принцип, в котором рассматривается 19 классов. 12 «месячных» классов (янв., фев., март, …, дек.), 6 «двухмесячных» классов (янв.-фев., март-апр., … , ноя.-дек.) и 1 общий (весь год). В основу второй классификации было положено значение направления приземного ветра, было выделено 5 классов: северный, восточный, южный, западный и маловетрие. В последний класс попадали случаи, в которых скорость ветра была ниже 2 м/с. Третья классификация была построена аналогично предыдущей, но для ветра на высоте 850 гПа. В качестве критерия для попадания в класс «маловетрено» была выбрана скорость 5 м/с. Для заданной точки и заданной заблаговременности класс считается пригодным к дальнейшему использованию, если содержит не менее 40 случаев.

# Составление прогнозов

Разрабатываемая технология MOSв первую очередь нацелена на составление прогнозов экстремальных температур, максимального ветра и суммы осадков в полусуточном разрешении, что является максимально приближенным к продукту, выпускаемому прогностическими подразделениями УГМС

После процедуры отбора для каждого пункта, каждого метеоэлемента и каждой заблаговременности имеем набор предикторов с установленными коэффициентами линейных преобразований, которые можно применять в прогностических целях. Результаты оперативного расчета модели WRF должны пройти процедуру постобработки для вычленения необходимого набора значений предикторов, к которым будут применены выявленные зависимости.

Разрабатываемая технология предусматривает наличие нескольких версий преобразований, что должно в некоторой степени сделать систему более устойчивой к отсутствию необходимых значений предикторов. Если для основной версии преобразований ввиду какого-либо сбоя недостаточно данных, то можно попытаться найти допустимую для использования версию, для которой будет достаточно имеющихся данных.

Результатом процедура прогноза является значение метеоэлемента, отнесенное к конкретному пункту и к конкретной дате.

Для всех классов был произведен расчет прогнозов минимальной температуры за ночь с заблаговременностью 33 и 57 часов и максимальной температуры за день с заблаговременностью 21, 45 и 69 часов. Расчет производился на независимой выборке, которая была построена способом, аналогичным таковому для обучающей выборки, по результатам расчета за период 2010-2011 гг. Если наблюдаемая ситуация попадала в класс, основанный на классификации по направлению ветра и недостаточно наполненный на этапе построения регрессий, то прогноз составлялся по уравнению из соответствующей календарной классификации (общегодовой).

# Процедура оценки

Для оценки результат прогнозирования сравнивается с фактически наблюденным или измеренным значением. Для каждого прогноза фиксируется абсолютное отклонение и успешность (попадание в допустимые интервалы).

На основании набора отклонений и успешности согласно [15] рассчитываются дополнительные величины, характеризующие успешность метода: средняя абсолютная и средняя относительная ошибки, среднеквадратическая ошибка, коэффициент Пирси-Обухова (для альтернативного прогноза явлений), доля попадания в заданные интервалы и пр. Кроме того, будет проводиться оценка согласно критериям для синоптических прогнозов.[1] Прогноз температуры по пункту считается оправдавшимся, если зафиксированные значения экстремальных температур за период не вышли из прогнозируемого интервала более, чем на 2 °С. Для ветра допускается отклонение максимальной скорости от прогнозируемого интервала на 2 м/с. Оценка прогноза осадков производится согласно Таблице 7..

Таблица 7Соответствие значений прогностического количества выпавших осадков (мм/ 12 часов) и фактически зафиксированных, дающее оправдавшийся прогноз количества осадков

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Жидкие и смешанные осадки | | Твердые осадки | |
| Прогноз | Факт | Прогноз | Факт |
| Без осадков | 0 мм | 0÷3 мм | 0 мм | 0÷2 мм |
| Небольшие осадки | 0÷3 мм | 0.1÷10 мм | 0÷2 мм | 0.1÷5 мм |
| Умеренные | 4÷14 мм | 3÷35 мм | 3÷6 мм | 2÷14 мм |
| Сильные | 5÷49 мм | 7÷50 мм | 7÷19 мм | 7÷25 мм |
| Очень сильные | ≥50 мм | ≥35 мм | ≥20 мм | ≥15 мм |

# Заключение

В отчёте изложены результаты работ по теме. Основные выводы заключаются в следующем.

* 1. Произведен обзор технология и методик MOSс докладом на двух семинарах. Разработана схема хранения данных и реализована в виде БД под управлением MSSLQServer. Разработан прототип информационной системы, реализованы модули построения регрессий, отбора предикторов, составления прогнозов. Реализован модуль постобработки результатов расчета для получения значений необходимых предикторов и последующей передачи их в информационную систему.
  2. Произведен расчет однородного ряда прогнозов модели WRF с шагом сетки 5 км и заблаговременность 72 часа за период 2005-2012 гг.
  3. Произведена оценка прогнозов экстремальной температуры за полусутки
  4. Тема требует продления до конца 2014 года.

# Приложение 1. Параметры конфигурации модели

// globalattributes:

:TITLE = " OUTPUT FROM WRF V3.4.1 MODEL" ;

:START\_DATE = "2012-12-28\_12:00:00" ;

:SIMULATION\_START\_DATE = "2012-12-28\_12:00:00" ;

:WEST-EAST\_GRID\_DIMENSION = 150 ;

:SOUTH-NORTH\_GRID\_DIMENSION = 150 ;

:BOTTOM-TOP\_GRID\_DIMENSION = 28 ;

:DX = 5000.f ;

:DY = 5000.f ;

:GRIDTYPE = "C" ;

:DIFF\_OPT = 1 ;

:KM\_OPT = 4 ;

:DAMP\_OPT = 0 ;

:DAMPCOEF = 0.2f ;

:KHDIF = 0.f ;

:KVDIF = 0.f ;

:MP\_PHYSICS = 8 ;

:RA\_LW\_PHYSICS = 1 ;

:RA\_SW\_PHYSICS = 1 ;

:SF\_SFCLAY\_PHYSICS = 2 ;

:SF\_SURFACE\_PHYSICS = 2 ;

:BL\_PBL\_PHYSICS = 2 ;

:CU\_PHYSICS = 5 ;

:SURFACE\_INPUT\_SOURCE = 1 ;

:SST\_UPDATE = 0 ;

:GRID\_FDDA = 0 ;

:GFDDA\_INTERVAL\_M = 0 ;

:GFDDA\_END\_H = 0 ;

:GRID\_SFDDA = 0 ;

:SGFDDA\_INTERVAL\_M = 0 ;

:SGFDDA\_END\_H = 0 ;

:HYPSOMETRIC\_OPT = 2 ;

:SF\_URBAN\_PHYSICS = 1 ;

:SHCU\_PHYSICS = 0 ;

:MFSHCONV = 0 ;

:FEEDBACK = 1 ;

:SMOOTH\_OPTION = 0 ;

:SWRAD\_SCAT = 1.f ;

:W\_DAMPING = 0 ;

:MOIST\_ADV\_OPT = 1 ;

:SCALAR\_ADV\_OPT = 1 ;

:TKE\_ADV\_OPT = 1 ;

:DIFF\_6TH\_OPT = 0 ;

:DIFF\_6TH\_FACTOR = 0.12f ;

:OBS\_NUDGE\_OPT = 0 ;

:BUCKET\_MM = -1.f ;

:BUCKET\_J = -1.f ;

:PREC\_ACC\_DT = 0.f ;

:OMLCALL = 0 ;

:ISFTCFLX = 0 ;

:ISHALLOW = 0 ;

:DFI\_OPT = 0 ;

:WEST-EAST\_PATCH\_START\_UNSTAG = 1 ;

:WEST-EAST\_PATCH\_END\_UNSTAG = 149 ;

:WEST-EAST\_PATCH\_START\_STAG = 1 ;

:WEST-EAST\_PATCH\_END\_STAG = 150 ;

:SOUTH-NORTH\_PATCH\_START\_UNSTAG = 1 ;

:SOUTH-NORTH\_PATCH\_END\_UNSTAG = 149 ;

:SOUTH-NORTH\_PATCH\_START\_STAG = 1 ;

:SOUTH-NORTH\_PATCH\_END\_STAG = 150 ;

:BOTTOM-TOP\_PATCH\_START\_UNSTAG = 1 ;

:BOTTOM-TOP\_PATCH\_END\_UNSTAG = 27 ;

:BOTTOM-TOP\_PATCH\_START\_STAG = 1 ;

:BOTTOM-TOP\_PATCH\_END\_STAG = 28 ;

:GRID\_ID = 1 ;

:PARENT\_ID = 0 ;

:I\_PARENT\_START = 1 ;

:J\_PARENT\_START = 1 ;

:PARENT\_GRID\_RATIO = 1 ;

:DT = 30.f ;

:CEN\_LAT = 45.f ;

:CEN\_LON = 134.5f ;

:TRUELAT1 = 45.f ;

:TRUELAT2 = 45.f ;

:MOAD\_CEN\_LAT = 45.f ;

:STAND\_LON = 134.5f ;

:POLE\_LAT = 90.f ;

:POLE\_LON = 0.f ;

:GMT = 12.f ;

:JULYR = 2012 ;

:JULDAY = 363 ;

:MAP\_PROJ = 1 ;

:MMINLU = "MODIFIED\_IGBP\_MODIS\_NOAH" ;

:NUM\_LAND\_CAT = 20 ;

:ISWATER = 17 ;

:ISLAKE = -1 ;

:ISICE = 15 ;

:ISURBAN = 13 ;

:ISOILWATER = 14 ;

# Список использованных источников

1. Наставление по краткосрочным прогнозам погоды общего назначения : РД 52.27.724 - 2009.

2. Marzban C., Sandgathe S., Kalnay E. MOS, Perfect Prog, and Reanalysis // Mon. Weather Rev. 2006. Vol. 134, № 2. P. 657–663.

3. Klein W.H., Lewis B., Enger I. Objective prediction of five-day mean temperatures during winter. // J. Meteorol. 1959. Vol. 16, № 6. P. 672–682.

4. Klein W.H., Glahn H.R. Forecasting local weather by means of model output statistics // Bull. Am. Meteorol. Soc. 1974. Vol. 55, № 10. P. 1217–1227.

5. Glahn H.R., Lowry D.A. The use of model output statistics (MOS) in objective weather forecasting. // J. Appl. Meteorol. 1972. Vol. 11, № 8. P. 1203–1211.

6. Lemcke C., Kruizinga S. Model output statistics forecasts: three years of operational experience in the Netherlands // Mon. Weather Rev. 1988. Vol. 116, № 5. P. 1077–1090.

7. Klein W.H. Objective Specification of Monthly Mean Surface Temperature from Mean 700 mb Heights in Winter // Mon. Weather Rev. 1983. Vol. 111, № 4. P. 674–691.

8. Bochkanov S. ALGLIB [Online]. URL: http://alglib.sources.ru.

9. Hart K.A. et al. An evaluation of mesoscale-model-based model output statistics (MOS) during the 2002 Olympic and Paralympic Winter Games // Weather Forecast. 2004. Vol. 19, № 2. P. 200–218.

10. Wilson L.J., Vallée M. The Canadian updateable model output statistics (UMOS) system: Design and development tests // Weather Forecast. 2002. Vol. 17, № 2. P. 206–222.

11. Sokol Z. MOS-Based Precipitation Forecasts for River Basins // Weather Forecast. 2003. Vol. 18. P. 769–781.

12. Brier G.W. Verification of forecasts expressed in terms of probability // Mon. Weather Rev. 1950. Vol. 78, № 1. P. 1–3.

13. Wilks D.S. Statistical methods in the atmospheric science. Academic Press, 1995. P. 467.

14. Applequist S. et al. Comparison of Methodologies for Probabilistic Quantitative Precipitation Forecasting // Weather Forecast. 2002. Vol. 17, № 4. P. 783–799.

15. Методические указания. Проведение производственных (оперативных) испытаний новых и усовершенствованных методов гидрометеорлогических и гелиофизических прогнозов : РД 52.27.284—91. 1991.