



# Метеорологические наблюдения: распределенные в пространстве и во времени (продолжение)

**В.ГОРДИН**

## **Всякое действие рождает ошибки. Проблема контроля**

*Попробуем географию! Лондон – столица Парижа,  
а Париж – столица Рима, а Рим...  
Нет, все не так, все неправильно!*

Л.Кэрролл. Алиса в Стране Чудес

Информация, полученная в результате измерений, передается в метеорологические центры. При передаче информации на непосредственные ошибки измерений могут накладываться иные ошибки. Поскольку передаваемых телеграмм тысячи, набираются сотни ошибок. Например, порядок, в котором передаются индекс<sup>18</sup> (номер) станции и результаты измерений, нестандартен. Человек такую ошибку исправит легко, но если в алгоритме обработки и контроля информации не предусмотреть, как такая ошибка может быть обнаружена, то, очевидно, дальнейшие действия компьютера будут следующие. Программа прогноза будет считать, что в точке с индексом, получившимся из чисел, описывающих температуру, наблюдается температура, получившаяся из чисел, описывающих индекс. Работа программы далее будет напоминать работу системы шестеренок, в которую щедрой рукой высыпали горсть шурупов. Чтобы этого не произошло, метеоинформация, проходящая по каналам связи, проходит первичную обработку и контроль.

В результате этой первичной обработки числа, описывающие координаты станции, срок наблюдения, температуру, ветер и влажность на различных высотах, будут стоять в «макете» строго на своих местах. Место отсутствующей информации (например, если зонд не долетел до положенной высоты) заполняется (кодируется) в макете специальным символом. Про-

<sup>18</sup> Координаты станции определяются уже в метеорологическом центре по ее индексу. Иногда приборы переносят на новое место, а в справочник координат и высот забывают внести исправление. С помощью уравнения гидростатики нетрудно оценить, к какой систематической ошибке в давлении приведет подъем синоптической станции, скажем, на 20 м.

грамма контроля должна обнаружить<sup>19</sup> грубые ошибки и, если возможно, исправить их, а если нет, то отбросить. Для обнаружения, например, полученные данные можно сравнить с их климатическими характеристиками (средними значениями в этой точке за данный месяц). Если зимой из Арктики сообщат о температуре +30 °С, то человек, знающий географию, этому не поверит. Нужно, чтобы и компьютер «не поверил». Такой вид контроля называется климатическим. Для его применения необходимо иметь в памяти машины не только поля первого приближения, в данном случае – поля климатических величин (рис.7,а и 8,а), но и поля допустимых (с точки зрения машинного контроля) отклонений метеорологических величин от этих средних. Можно сравнить проверяемое значение метеорологической величины с аналогичными на соседних станциях. Такой вид контроля называется горизонтальным. Можно сравнить значение с аналогичными за предыдущий срок наблюдений – такой контроль называется временным.

Разрабатывая процедуру прогноза, рассчитывают, что полученные результаты будут больше соответствовать действительности, чем климатические поля (иначе, зачем бы огород, т.е. прогноз, городить?). Следовательно, в качестве первого приближения для программ контроля целесообразно использовать не климатические, а прогностические поля. В этом случае можно уменьшить ширину «окна», т.е. значения отклонений, которые считаются максимально допустимыми. Таким образом, при использовании вместо климата хорошего прогноза первое приближение в алгоритме контроля улучшается. При этом уменьшается вероятность возникновения ошибок двух родов: пропуск содержащейся в телеграмме ошибки и бра-

<sup>19</sup> В таком случае хорошо бы из метеоцентра, где производится контроль, «переспросить» отправителя, т.е. автоматически попросить его уточнить подозрительное число. Пока такой системы обратной связи нет.

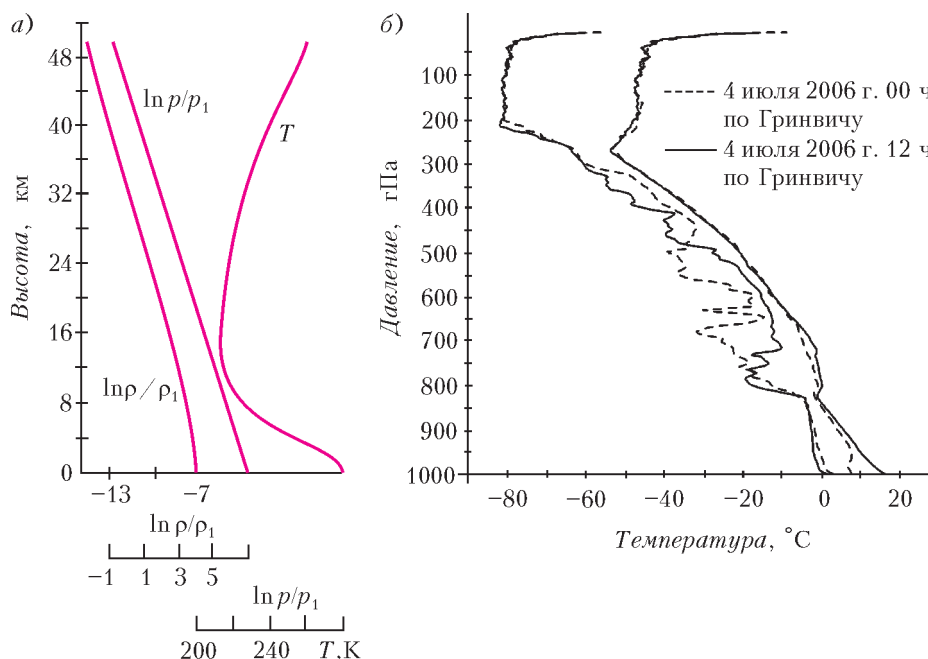


Рис.7. Можно сравнить проверяемую величину, соответствующую какой-то высоте измерения радиозонда, с аналогичными величинами на соседних высотах, которые также приводятся в аэрологической телеграмме. Такой вид контроля называется вертикальным. Чтобы им пользоваться, необходимо знать максимально возможную скорость изменения контролируемого поля по вертикали. На рисунке а) приведены средние профили температуры, плотности, давления. Здесь  $\rho_1 = 1 \text{ г/см}^3$ ,  $p_1 = 1 \text{ гПа}$ . На рисунке б) представлены профили температуры (справа) и температуры точки росы (слева), содержащиеся в двух аэрологических телеграммах. Проанализировав совместно эти два рисунка, можно грубо оценить допустимые (нужно, чтобы контроль их пропускал) отклонения температуры от климатических значений

ковка верного значения, которое существенно отличается от климатического поля (возможно, конкретные значения существенно отличаются от среднего<sup>20</sup>).

Существующие алгоритмы контроля постоянно совершенствуются, хотя и сейчас они неплохо устраняют наиболее распространенные ошибки. Наилучшие результаты дает комплексный контроль, который состоит из нескольких видов контроля, перечисленных выше. При этом ширина различных «окон» должна быть взаимно согласована. Самый сложный блок такого алгоритма – блок принятия решения: при наличии таких-то подозрений окончательно считать число верным или ошибочным. Качество контроля характеризуется двумя числами – ошибками обоих родов. Субъективный анализ результатов работы алгоритма контроля позволяет понять, в каких ситуациях (например, при контроле данных из каких географических районов) алгоритм работает неудовлетворительно (пропускает какие-то ошибки или отбрасывает верные данные) и как можно улучшить сам алгоритм.

Впрочем, ошибки бывают и экзотические. Одна станция, расположенная в Закавказье, постоянно передавала данные, сильно отличающиеся от данных соседних станций. Синоптики к этому привыкли и при

<sup>20</sup> Важный вопрос: как количественно описать вариабельность (отличие от среднего значения) того или иного метеорологического элемента в какой-то области в какой-то месяц? Мы обсудим его.

анализе поступающей информации данные этой станции не учитывали. Автоматический контроль также часто браковал эту информацию. Однако «расследование» показало, что это ... японская станция. Когда-то при набивке координат на перфокарты в долготу этой станции пропустили единичку, и долгота уменьшилась ровно на  $100^\circ$ , а широта не изменилась. В результате состояние атмосферы над Хоккайдо приписывалось Кавказу.

Другой пример. Норвегия (одна из стран – членов ВМО) в восьмидесятых годах установила в Арктике полсотни буев, автоматически передававших информацию. Четыре бую оказались с дефектом – они передавали очень низкое давление на уровне моря: 950 гПа (среднее значение давления на уровне моря равно 1013 гПа). При сравнении с прогнозом контроль эти данные браковал. Все было в порядке. Но когда в этот район пришел сильный циклон (область низкого давления) и «честный» прогноз дал значение давления 975 гПа, то данные этих буев попали в «окно» допустимых отклонений от прогноза.

Кроме того, они оказались идеально согласованными между собой (так как были одинаковыми) и потому прошли горизонтальный и временной контроль. Прежде чем была выявлена причина и установлен специальный алгоритм «защиты от данных» этих буев, было дано несколько прогнозов весьма низкого (особенно для Арктики) качества.

### Статистика или геометрия: коэффициент корреляции – это косинус

*Вдруг серый волк ударился о сырую землю и стал конем золотогривым.*

Русская народная сказка «Иван-царевич и серый волк»

Предположим, имеется два набора чисел одинаковой длины  $\{p_j\}_{j=1}^N$  и  $\{q_j\}_{j=1}^N$  и мы хотим понять, можно ли по одному ряду предсказать другой.<sup>21</sup> Например, по высоте столбика термометра за нашим окном мы предсказываем температуру: а) около термометра; б) в том месте, куда мы планируем отправиться. Понятно, что в случае а) зависимость линейная (для этого-то термометр и совершенствовались несколько веков):  $q_j = \alpha p_j + \beta$ , где  $j$  – номер момента времени, а константы  $\alpha$  и  $\beta$  характеризуют данный термометр и температурную шкалу. В случае б) результат зависит от подробностей. Если собираетесь в

<sup>21</sup> Здесь буква  $p$  не имеет отношения к давлению, это просто обозначение некоторого числа.

булочную за углом – ситуация почти такая же, что и в случае а). А если на другой континент – нечего тогда и в окно смотреть.

Важная характеристика числового ряда – его средняя величина  $Sp = \frac{1}{N} \sum_{j=1}^N p_j$ . Но если у нас рядом висят два термометра – один со шкалой Цельсия, а другой со шкалой Реомюра или Фаренгейта, то показания у них разные и среднее значение этих показаний у каждого термометра свое. Как же убедиться, что они показывают одну и ту же температуру?

Прежде всего вычтем средние величины, т. е. перейдем к центрированным числовым рядам:  $\tilde{p}_j = p_j - Sp$ ,  $\tilde{q}_j = q_j - Sq$ . У каждого из этих числовых рядов среднее значение нулевое. Предположим, что первый центрированный ряд состоит из чисел 1 и -1 в равных количествах, а второй получается из первого умножением на число  $\alpha \neq 0$ . Тогда стоит считать, что это один и тот же набор измерений, только описанный в разных шкалах (масштабах). Однако это нужно проверить. Пусть сначала  $N = 3$ , т.е. произведено лишь по три измерения. Тогда каждый из двух наборов содержит по 3 числа и каждый набор можно интерпретировать как вектор в трехмерном пространстве  $\mathbb{R}^3$ . Теперь надо узнать, являются ли эти векторы пропорциональными, а если нет, то насколько велика мера этой непропорциональности. Такой мерой является угол между векторами, точнее косинус угла. Когда он равен 1, векторы пропорциональны, а когда 0 – ортогональны. Если же косинус равен -1, векторы пропорциональны, но со знаком «минус». Это означает, что когда на одном термометре температура выше средней, на другом она ниже и наоборот. О таких температурных шкалах упоминалось в первой части статьи.

Чтобы вычислить косинус угла между векторами в  $\mathbb{R}^3$ , существует замечательный способ – нужно вычислить скалярное произведение векторов и поделить на произведение их норм:

$$\cos \vartheta = \frac{(\tilde{p}, \tilde{q})}{|\tilde{p}| |\tilde{q}|} = \frac{\left( \sum_{j=1}^3 \tilde{p}_j \tilde{q}_j \right)}{\left( \sum_{j=1}^3 \tilde{p}_j^2 \right)^{1/2} \left( \sum_{j=1}^3 \tilde{q}_j^2 \right)^{1/2}}.$$

Если векторы  $\tilde{p}$  и  $\tilde{q}$  пропорциональны и коэффициент пропорциональности положителен, то (проверьте)  $\cos \vartheta = 1$ , а если коэффициент отрицателен, то  $\cos \vartheta = -1$ . Эту формулу для косинуса легко обобщить на случай произвольного  $N$  – нужно лишь поменять верхний предел у всех сумм. Неравенство Коши – Бунаковского для числовых рядов гласит:

$$|(\tilde{p}, \tilde{q})| \leq |\tilde{p}| |\tilde{q}| \Leftrightarrow \left| \sum_{j=1}^N \tilde{p}_j \tilde{q}_j \right|^2 \leq \sum_{j=1}^N \tilde{p}_j^2 \sum_{j=1}^N \tilde{q}_j^2.$$

Значит, модуль величины, которую мы назвали косинусом, не превосходит 1. По этой величине можно восстановить и угол между двумя векторами – самый обычный угол, хотя векторы и  $N$ -мерные.

Число  $r = \cos \vartheta$  имеет еще одно название: *коэффициент корреляции* между числовыми рядами  $\{p_j\}_{j=1}^N$  и

$\{q_j\}_{j=1}^N$ . Этот термин используют, когда говорят о случайных величинах, вероятностях и т.п. Он характеризует похожесть колебаний этих числовых рядов. Пример случайной величины дает обыкновенный кубик с числами на гранях от единицы до шести. Каждая грань, если кубик не испорчен, выпадает с одинаковой вероятностью, равной  $1/6$ . Если кубик подбросить много раз, то среднее значение случайной величины (выпавшего числа) будет близко к 3,5. Пусть  $p_j$  – число на грани кубика при  $j$ -м подбрасывании, а  $q_j = 3,5$  при всех  $j$ . У этих величин одинаковые средние, а коэффициент корреляции равен (проверьте) нулю.

**Статистические связи и предсказания**

*Иван Николаевич смутился, но ненадолго, потому что вдруг сообразил, что профессор непременно должен оказаться в доме №13 и обязательно в квартире 47.*

М.А.Булгаков. Мастер и Маргарита

Будем считать, что в каждой точке поля  $\vec{x}$  отклонение поля  $\zeta(\vec{x})$  от климата есть случайная величина, принимающая числовые значения. В теории вероятностей среднее всех значений называют математическим ожиданием и обозначают символом  $M\zeta(\vec{x})$ . Если же нужно подчеркнуть, что оценка выполнена по конечному ряду, то используют символ  $S\zeta(\vec{x})$ .

Если ряды  $\{p_j\}_{j=1}^N$  и  $\{q_j\}_{j=1}^N$  составлены, скажем, из значений давления и температуры в какой-то точке  $\vec{x}$  в моменты  $t_j$ , то коэффициент корреляции  $r$  характеризует статистическую связь между ними. Например, если бы оказалось, что  $r > 0$ , то это означало бы, что чем выше давление, тем, как правило, выше температура. Чем ближе был бы коэффициент  $r$  к 1, тем реже нарушалось бы это статистическое правило. В реальности же коэффициент корреляции между температурой и давлением в одной и той же точке и в один и тот же момент времени близок к нулю. А насколько длинным должен быть ряд, т.е. каким следует выбирать  $N$ ? Ответ на этот простой вопрос весьма сложен, а неправильный ответ может привести к принципиальным ошибкам.

Допустим, температура и давление измеряются ежесекундно в течение минуты, т.е.  $N = 60$ . Вполне возможно, что обе величины растут или одна из них убывает. Поскольку минута – сравнительно небольшой (в метеорологическом смысле) отрезок времени, давление и температура хорошо описываются линейными функциями времени. Тогда коэффициент корреляции этих рядов близок (проверьте) к 1 (или к -1, если одна величина растет, а другая убывает). Если сделать опрометчивый вывод, что такая связь будет продолжаться и далее, на больших временных интервалах, то результат получится совершенно неверный. Известно, например, что днем бывает, как правило, теплее, чем ночью, а зимой – холоднее, чем летом. Но колебаний давления с большой амплитудой, аналогичных колебаниям температуры, ни с суточным, ни с годовым периодом в природе не существует. Нельзя было по

ряду на коротком отрезке (1 минута) делать выводы о столь сильной корреляции температуры и давления.

Оставив вопрос о методах определения достаточности длины ряда без ответа, предположим, что наши ряды достаточны и что они хорошо представляют

бесконечные последовательности  $\{p_j\}_{j=1}^{\infty}$  и  $\{q_j\}_{j=1}^{\infty}$ . Будем считать, что существуют пределы при  $N \rightarrow +\infty$

$$\text{сумм } \frac{1}{N} \sum_{j=1}^N p_j, \frac{1}{N} \sum_{j=1}^N q_j, \frac{1}{N} \sum_{j=1}^N \tilde{p}_j \tilde{q}_j, \frac{1}{N} \sum_{j=1}^N \tilde{p}_j^2, \frac{1}{N} \sum_{j=1}^N \tilde{q}_j^2$$

и что суммы  $S$  «недалеки» от этих пределов. Две

последние суммы (точнее, их пределы при  $N \rightarrow +\infty$ ) называются дисперсиями величин  $p$  и  $q$  соответственно. Корни из дисперсий называют среднеквадратичными отклонениями – СКО – и обозначают  $\sigma_p$  или  $\sigma_q$ . Они характеризуют вариабельность соответствующих рядов вокруг среднего значения.

### Интерполяция, климат, корреляции

*Иногда приходится полагаться на случай; ни в чем нельзя быть вполне уверенным в морском сражении.*

Г.Нельсон

Вооруженные понятиями среднего, дисперсии и корреляции, вернемся к земной атмосфере. Полученные в результате измерений большие массивы метеорологической информации (рис.8) уже позволяют, до определенной степени, оценивать климат планеты. А еще можно оценить, как убывает с расстоянием между двумя точками связь (корреляция) между значениями температуры в этих точках. Эта скорость убывания также характеризует нашу атмосферу.

Кроме того, знание статистических характеристик позволяет подстроить численные алгоритмы, оптимальные для поля именно с такими характеристиками, и минимизировать среднюю ошибку алгоритмов. Объясним эти идеи подробнее.

Рассмотрим поле (функцию пространственных переменных, время зафиксировано) отклонений измеренных значений, например температуры, от средних климатических значений. Если нанести эти отклонения в точках станций на географическую карту, то станет видно, что полученное поле отклонений зависит от горизонтальных координат сравнительно плавно: имеются зоны положительных отклонений и зоны отрицательных отклонений с характерными размерами  $10^3$  км. Другими словами, если на окружающих точку станциях отклонения от климатических значений в какой-то момент времени положительные, то с большой вероятностью в этой точке и в этот момент (а также и в близкие моменты времени) отклонение от климатического значения также будет положительным. И величину этого отклонения можно оценить, например вычислив коэффициент корреляции между отклонениями в какой-то паре точек. Если эти две точки близки, то коэффициент корреляции близок к единице, однако по мере удаления точек друг от друга коэффициент будет убывать и даже может стать отрицательным.

Если известны коэффициенты корреляции между отклонениями в двух (или нескольких) точках, то можно оценить отклонение еще в одной точке, после чего можно будет вычислить и само значение температуры в данной точке.

Рассмотрим в качестве случайной величины отклонение в данной точке  $\bar{x}$  метеорологического поля от климата. Ее математическое ожидание будет равно нулю, так как климат это и есть математическое ожидание метеорологического поля. Строго следуя такому определению климата, нужно положить  $N = \infty$ , провести измерения метеорологического поля от начала мира до конца света, а затем взять среднее по всему этому ряду. Это, конечно же, нереально. Поэтому приходится использовать ряд наблюдений покороче. Кстати, это может дать более полезный результат. Действительно, для того чтобы оценить вероятность той или иной погоды на завтра или через неделю, хорошо бы иметь информацию, скажем, за последние сто лет – узнать, что за это время бывало часто, а что редко. А если миллион лет назад на месте пустыни было болото или ледник, то вот эти древние факты несущественны. Еще раз подчеркнем: несущественны для задачи прогноза (или для оценок погоды) на день или неделю.

Если же нас интересует, какой была средняя за каждый век температура воздуха по Европе и росла она или убывала, когда века сменяли друг друга, то столетнего архива для этого уже явно недостаточно. Информацию об изменениях такого временного масштаба, по-видимому, будет полезнее искать не в небесах (запуская радиозонды), а под землей (исследуя останки животных и растений, льды, геологические породы и т.п.). Среднее за сто лет зависит от того, когда эти сто лет начинаются. За миллион лет это средне-вековое значение температуры может сильно измениться.

Назовем климатическим полем, или просто климатом<sup>22</sup>, метеорологическое поле, осредненное на некотором временном интервале  $\Delta t$ . Климатическое поле  $\zeta(\bar{x})$  (давление, компоненты ветра и т.п.) есть функция пространственных переменных, зависящая от ин-

<sup>22</sup> Слово «климат» произошло от греческого «наклон». Понятие введено Гиппархом из Никей (190–120 до н.э.). Земля была им разделена условно на пять широтных зон, на которые солнечные лучи падают под разным наклоном. Там, где наклон силен, солнце слабо нагревает воздух. Греция и Малая Азия лежат во втором (т. е. умеренном) северном климате, каковой наиболее благоприятен для жизни человека. На самом деле основной нагрев воздуха производит свет, не идущий непосредственно от Солнца, а сначала поглощенный и затем излученный поверхностью планеты (в современных прогнозах учитывают альбедо поверхности, определяющее, какая часть светового потока будет поглощена поверхностью суши или воды, а какая будет отражена сразу в атмосферу). В XIX веке А.Гумбольдт (1769–1859) предложил при анализе причин различия климата в разных частях планеты кроме широты местности учитывать также близость океанов и т. п. географические факторы, влияющие на климат в современном его понимании.

<sup>23</sup> Поскольку на расстояниях порядка радиуса Земли корреляция близка к нулю, кривизна Земли в данной статистической модели несущественна – Землю можно считать плоской!



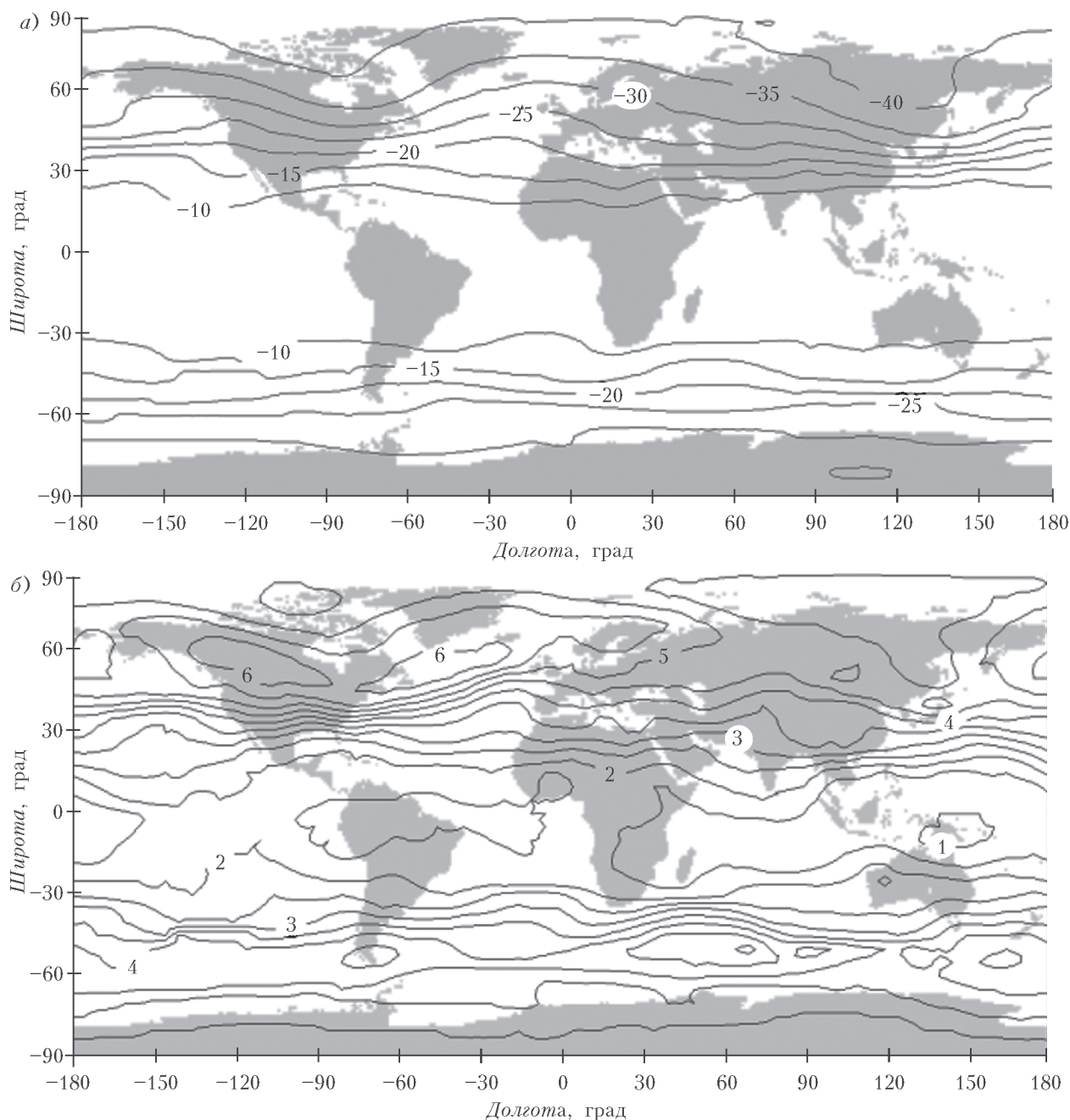


Рис.8. Климатические поля: а) средне-январские температуры (в °С) на барическом уровне 500 гПа; б) СКО январской температуры. Поля имеют выраженный зональный характер: зависимость от долготы существенно меньше, чем от широты. Разумеется, СКО, отвечающая какому-то определенному месяцу, меньше, чем для года в целом. Разбиение на месяцы перед осреднением не полностью отвечает существу дела – например, приход муссона и соответствующая перестройка циркуляции происходят в разные года в разные даты

тервала осреднения  $\Delta t$ . Климатическое поле зависит и от времени начала интервала (климат можно связывать с началом, серединой или концом интервала осреднения). Итак, климат похож на математическое ожидание и может заменять его во многих практических вычислениях, но в строгом смысле им не является.

Помимо средних значений поля полезно статистически оценивать и более сложные его характеристики, например дисперсию (или СКО; см. рис.8,б)

$$D\zeta(\bar{x}) = \sigma_{\zeta}^2(\bar{x}) = M(\zeta(\bar{x}) - M\zeta(\bar{x}))^2$$

или корреляционную функцию

$$K(\bar{x}_1, \bar{x}_2) = \frac{M(\zeta(\bar{x}_1) - M\zeta(\bar{x}_1))(\zeta(\bar{x}_2) - M\zeta(\bar{x}_2))}{\sigma_{\zeta}(\bar{x}_1)\sigma_{\zeta}(\bar{x}_2)},$$

которая характеризует синхронность колебаний  $\zeta(\bar{x})$  в двух заданных точках  $\bar{x}_1, \bar{x}_2$ . Но ведь в каждой точке не поставить термометр, так что для всех пар точек такую функцию не вычислить. Ограничимся меньшим.

Предположим, что корреляционная функция  $K$  шести переменных  $\bar{x}_1, \bar{x}_2$  (каждый вектор имеет по три координаты) на самом деле зависит только от трех

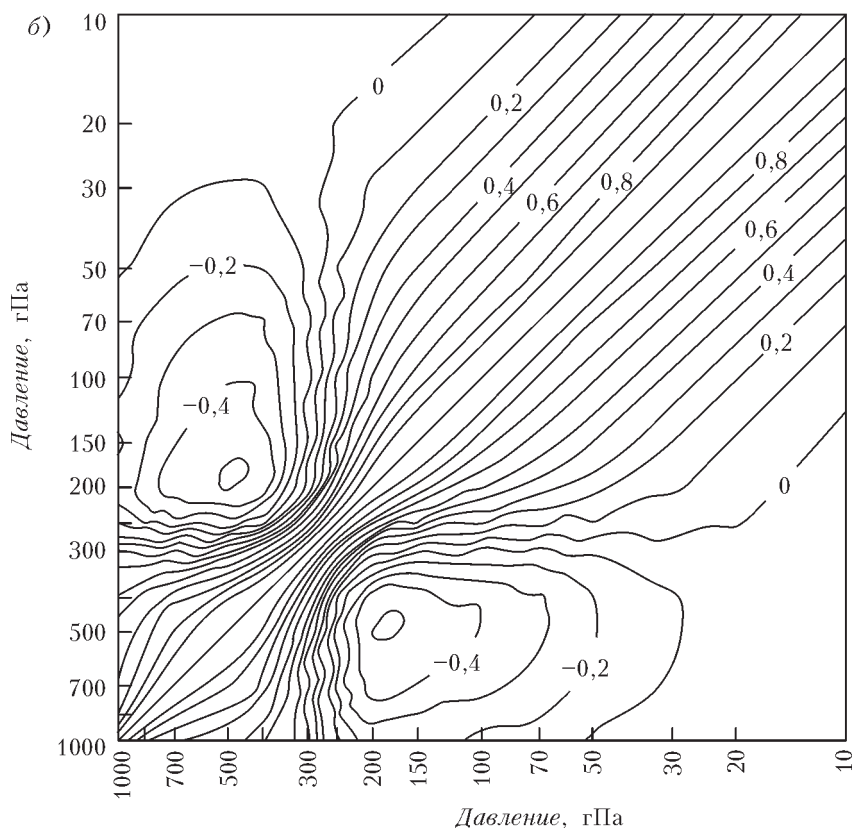
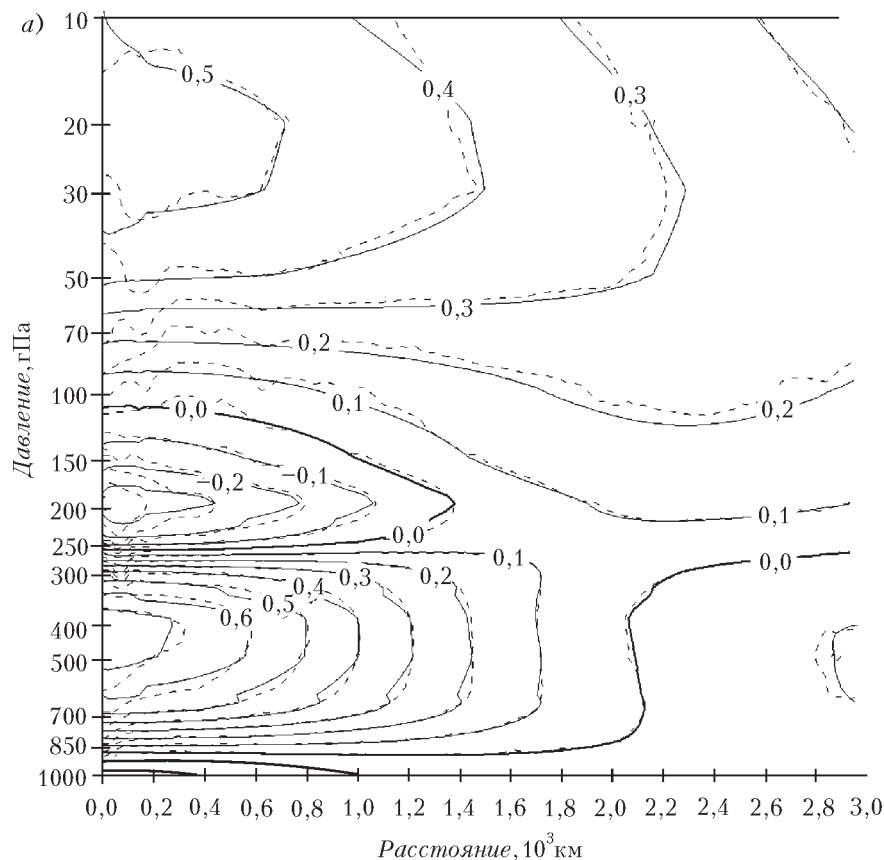


Рис.9. Изолинии корреляционной функции температуры: а) на одном и том же барическом уровне  $p$  (в гПа), разделенных расстоянием  $r$  (в тыс. км); б) на разных барических уровнях при нулевом расстоянии по горизонтали. Использована статистика умеренных широт северного полушария для января

неотрицательных координат: вертикальных координат, точнее давлений, и расстояния между точками по горизонтали:

$$K = K(p_1, p_2, r), \text{ где}$$

$$r^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2.$$

На рисунке 9,а приведены изолинии корреляционной функции температуры при  $p_1 = p_2$ , когда  $K$  зависит только от одного давления и от расстояния<sup>23</sup>. На рисунке 9,б изображены изолинии для случая, когда  $r = 0$ , а давления различны. В первом случае нет изолинии  $K = 1$ , поскольку измерения, которыми мы располагаем, производятся с ошибками. Поэтому функция  $K = K(p, r)$  разрывна при  $r = 0$ . Глядя на рисунок 9, мы можем сделать вывод, что наименьшие статистические связи (корреляции) наблюдаются в районе тропопавзы (т.е. в диапазоне 300–100 гПа), которая отделяет тропосферу от стратосферы. Но тропопавза не бесконечно тонкая поверхность, а зона. Она меняет свою высоту со временем, и ее положение зависит от горизонтальных координат. Движение – преобразование, которое не меняет расстояний. Движение полупространства это всегда сдвиг этого полупространства по горизонтали и поворот вокруг вертикальной оси. В обоих случаях  $r$  в формуле для  $K$  не меняется. Если корреляционная функция какого-то случайного поля  $\zeta(\vec{x})$  инвариантна (т.е. не меняется) при сдвигах, то поле  $\zeta$  называется однородным, а если она инвариантна при поворотах, то – изотропным.

Реальные метеорологические поля (если рассматривать их как реализации случайных полей) в точности не являются однородными и изотропными по горизонтали, однако такая модель часто оказывается вполне пригодной для практического применения. А вот изменения полей по вертикальной координате в интересных для метеорологии масштабах существенно иные, чем по горизонтали. Отличия от изотропии, если они определены достаточно надежно, могут дать существенную информацию. Рассмотрим, например, кросс-корреляционные функции (рис.10), полученные по рисунку 6 (см. предыдущий номер журнала) осредненным:

$$K(x, y) = \frac{\sum_{j,k} f_0(j, k) f_1(j + x, k + y)}{\sum_{j,k} 1}.$$

Здесь в качестве функций  $f_0, f_1$  исполь-

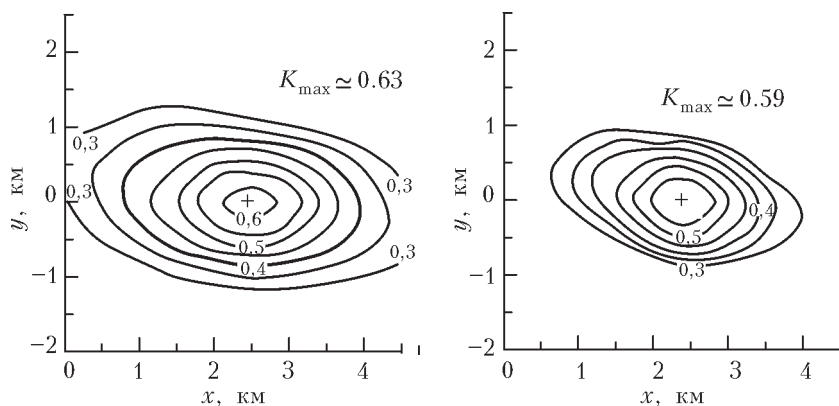


Рис.10. Кросс-корреляционные функции для компонент ветра ( $u$  – слева,  $v$  – справа) со сдвигом по времени примерно в 3,5 мин. Максимальная корреляция здесь достигается отнюдь не в начале координат (см. знак «плюс»). Она, как видно из обоих рисунков, смещена по потоку основного ветра, причем на расстояние, которое частицы воздуха как раз и проходят за 3,5 мин. Разумеется, на больших пространственно-временных масштабах основной поток ветра доминирует над флуктуациями не так сильно. Впрочем, в тропиках подобные явления (переноса максимума корреляций основным потоком) наблюдаются и на больших масштабах

зуются компоненты  $u$  или  $v$  отклонения скорости ветра на поверхности от среднего потока в два момента времени. Суммирование тут производится по всем точкам сетки  $(j, k)$  таким, что и она, и точка  $(j + x, k + y)$  лежат в указанном сеточном квадрате. Из рисунка 10 следует, что  $K(0,0) \approx 0,3$  для  $u$  и  $K(0,0) \approx 0$  для  $v$ , т.е. ветер слабо коррелирует с тем ветром, который дул всего 3,5 мин назад. Корреляция существенно больше, если учесть основной поток. Действительно,  $\max_{x,y} K(x,y) \approx K(0, 3,5 \text{ км}) \approx 0,6$  для обеих компонент.

Как уже неоднократно упоминалось, в нашем распоряжении имеется информация, поступающая из сети метеорологических станций, которые расположены крайне неравномерно. Для дальнейшего использования этой информации в численном прогнозе нужно переинтерполировать ее в регулярную сетку так называемых узлов. Здесь неважно, что эта сетка станет регулярной, важно то, что это уже будут совсем другие точки.

Будем интерполировать не само метеорологическое поле, а его отклонение от вычисленного заранее детерминированного климатического (или прогностического) поля. Статистическая гипотеза, которой мы будем пользоваться, состоит в том, что это нормированное отклонение (т.е. деленное на корень из дисперсии)

$$\xi(\bar{x}) = \frac{\zeta(\bar{x}) - M\zeta(\bar{x})}{\sigma_\zeta(\bar{x})}$$

есть центрированное, однородное и изотропное случайное поле. Следовательно, нам нужно заранее знать поля  $M\zeta(\bar{x})$  и  $\sigma_\zeta(\bar{x})$ , а также корреляционную функцию для нормированного случайного поля  $\xi$ .

### Оптимальная интерполяция

В этом безумии есть своя система.

У.Шекспир. Гамлет, принц датский

Важный блок алгоритма прогноза погоды составляет формирование начальных полей на регулярной (часто широтно-долготной) сетке. Но измерения имеются в других точках – там, где расположены станции, или над которыми пролетел спутник. Если нас интересуют не климатические задачи, а прогностические, то имеет смысл рассматривать отклонение не от среднего, климатического, поля, а от поля прогноза, который был сделан, скажем, сутки назад:

$$\xi(\bar{x}) = \frac{\zeta(\bar{x}) - \zeta_{\text{прогн}}(\bar{x})}{\sigma_{\text{прогн}}(\bar{x})},$$

где знаменатель – это СКО прогноза от фактических наблюдений. Для современных прогностических моделей дисперсия этого отклонения в несколько раз меньше, чем дисперсия отклонений от климата.

Будем считать, что корреляционные функции уже известны. Пусть  $\bar{x}_0$  – точка, в которой нужно определить отклонение  $\xi_0$  нормированного метеорологического поля от климата, а  $\{\bar{x}_j\}_{j=1}^N$  – множество точек, окружающее точку  $\bar{x}_0$ , в которых это отклонение  $\xi_j = \xi(\bar{x}_j)$  уже известно, поскольку в этих точках расположены метеорологические станции. Величину  $\xi_0$  будем оценивать по формуле  $\xi_0 \approx \sum_{j=1}^N a_j \xi_j$ , подбирая коэффициенты  $a_j$  оптимальным образом: чтобы средний квадрат ошибки  $M \left( \xi_0 - \sum_{j=1}^N a_j \xi_j \right)^2$  был минимальным.

Используя определение корреляционной функции и дифференцируя по всем  $a_j$ , получим систему линейных алгебраических уравнений

$$B \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_N \end{pmatrix}, \quad B = \|b_{jm}\|,$$

$$b_{jm} = K(\bar{x}_j, \bar{x}_m), \quad z_j = K(\bar{x}_j, \bar{x}_0).$$

Можно доказать, что симметричная матрица  $B$  неотрицательно определена. С помощью некоторых ухищрений (см. [1], [5], [6]; изолинии такой «подправленной» корреляционной функции  $K$  на рисунке 9 изображены пунктиром) можно обеспечить и положительную определенность, т.е. ее собственные числа вещественны и неотрицательны. Для решения таких систем существуют очень эффективные компьютерные алгоритмы.

Если предположить, что случайное поле  $\xi = \xi(\bar{x})$  однородно и изотропно, то измерений для оценки функции одной переменной  $k(|\bar{x}_j - \bar{x}_m|)$  понадобится

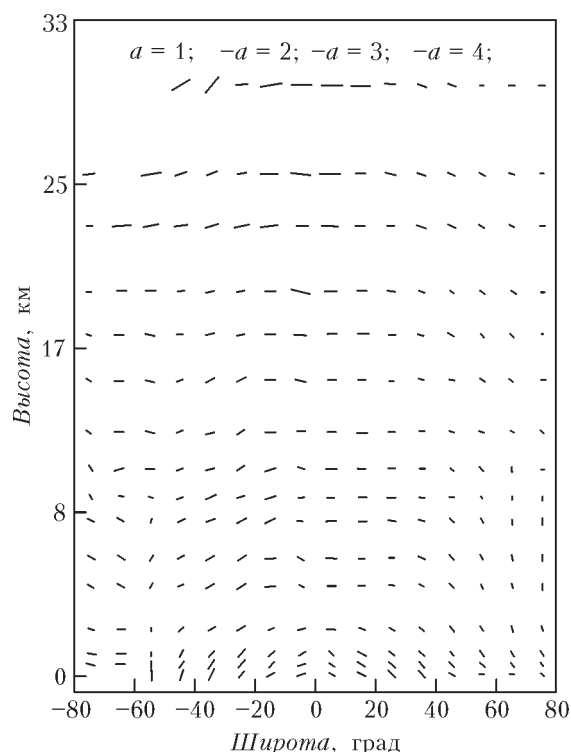


Рис.11. Анизотропия поля январских температур. Ориентация отрезка показывает направление наибольшей корреляции поля температуры в зависимости от широты и высоты (осреднение по долготе). Горизонтальное расположение отрезка означает, что корреляция сильнее вдоль параллели, вертикальное – вдоль меридиана. Чем длиннее отрезок, тем больше у корреляционной функции анизотропия  $a$ . Как видно, максимум  $a$  лежит в тропической стратосфере

ся во много раз меньше, чем для определения функции  $K$ , зависящей от двух векторных аргументов. Изотропность можно предполагать только по горизонтальным переменным. Простая статистическая гипотеза об однородности и изотропности по горизонтали использовалась в метеорологии около полувека. Однако накопление архивов метеоизмерений, рост вычислительных возможностей и разработка новых вычислительных алгоритмов привели к ее пересмотру. Оказалось, что и по горизонтальным переменным анизотропия, например температур, существенно проявляется в тропической зоне (рис.11). Алгоритм статистической интерполяции, учитывающий анизотропию интерполируемого поля, несколько сложнее изотропной версии.

### Заключение

*Пошли хлеб свой по водам, ибо спустя много дней ты найдешь его.*

Экклезиаст

Большие объемы метеорологической информации требуют адекватных вычислительных алгоритмов усвоения. Качество продукции определяется качеством собственно алгоритмов и точностью, ассортиментом и объемом измерений, а также мощностью компьютеров. В результате улучшается и численный прогноз погоды, и наше представление о климате планеты. В свою

очередь, знание полей средних и корреляционных функций метеорологических полей позволяет отлаживать численные климатические модели, довольно популярные в последние десятилетия. Пока корреляционные функции, генерируемые этими моделями, не окажутся близки к истинным корреляционным функциям, оцененным по архивам наблюдений, верить таким моделям не стоит.

В этой статье не было возможности объяснить, как именно используются данные ежедневных наблюдений<sup>24</sup> для прогноза погоды: какие эволюционные дифференциальные уравнения гидродинамического типа решаются, какие физические процессы они описывают, какие вычислительные алгоритмы используются. Для этого нужна отдельная книга и хорошая математическая подготовка читателя.

Первый численный прогноз погоды произвел шотландец Л.Ричардсон около века назад. Штат вычислителей был вооружен логарифмическими линейками. Вычислительная работа по прогнозу погоды на 6 часов для Европы продолжалась несколько лет и привела к чудовищно ошибочному прогнозу. Еще четверть века потребовала «работа над ошибками Ричардсона», в которой принимали участие и математики (Р.Курант, К.Фридрихс, Г.Леви), и «математики с метеорологическим уклоном» (Г.Россби, И.А.Кибель). Когда в Лос-Аламосе после изготовления атомной бомбы появилось немного свободного времени у ученых (и у их компьютеров), они занялись задачей прогноза погоды. Последние шестьдесят лет качество прогнозов неуклонно росло. Человечество начало пользоваться плодами своих многовековых вложений в метеорологию.

Но это уже другая история...

### Литература

1. О.А.Алдухов, В.А.Гордин. *Трехмерные корреляционные функции основных аэрологических величин*. – «Известия РАН», серия «Физика атмосферы и океана», 2001, 37(1).
2. О.А.Алдухов, В.А.Гордин. *Оценки анизотропии корреляционной структуры полей метеорологических величин по наблюдениям глобальной аэрологической сети*. – «Известия РАН», серия «Физика атмосферы и океана», 2005, 41(3).
3. А.И.Бедрицкий, Е.П.Борисенков, А.С.Коровченко, В.М.Пасецкий. *Очерки по истории Гидрометеорологической службы в России*, т.1. – С-П.: Гидрометеоиздат, 1997.
4. В.А.Гордин. «Кулоновский» алгоритм выбора влияющих станций. – «Метеорология и гидрология», 2003, №12.
5. В.А.Гордин *Как это посчитать? Обработка метеорологической информации на компьютере*. – М.: МЦНМО, 2005.
6. В.А.Гордин. *Математика, компьютер, прогноз погоды и другие сценарии математической физики*. – М.: Физматлит, 2010.
7. А.Х.Хргиан. *Очерки развития метеорологии*. – Л.: Гидрометеоиздат, 1959.

<sup>24</sup> Нужны не только собственно метеорологические наблюдения, но также наблюдения за альбедо поверхности, за океаном и реками, за свойствами листьев и корней растений, которые, как и почва, испаряют и поглощают влагу. Иначе не получится более точных прогнозов. «Чем дальше в лес, тем больше дров...»