ГЛАВНОЕ УПРАВЛЕНИЕ ГИДРОМЕТЕОРОЛОГИЧЕСКОЙ СЛУЖБЫ
ПРИ СОВЕТЕ МИНИСТРОВ СССР

ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ГЛАВНАЯ ГЕОФИЗИЧЕСКАЯ ОБСЕРВАТОРИЯ им. А. И. ВОЕЙКОВА

ТРУДЫ

ВЫПУСК 348

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В МЕТЕОРОЛОГИИ

Под редакцией
д-ра физ.-мат. наук Л. С. ГАНДИНА
д-ра физ.-мат. наук Р. Л. КАГАНА

ГИДРОМЕТЕОИЗДАТ
ЛЕНИНГРАД • 1975

УДК 551.5:311.17(061.6)

Сборник содержит работы по применению статистических методов интерпретации и обработки различных видов метеорологической информации. Представлен ряд результатов по проблеме четырехмерного анализа метеорологических полей. Рассмотрены вопросы точности определения статистических характеристик метеорологических полей, а также применения теории выбросов к анализу временных метеорологических рядов. В сборнике представлены также материалы по разработке и испытанию методов контроля метеорологической информации.

Сборник рассчитан на научных работников, аспирантов и студентов старших курсов, интересующихся применением статистических методов в метеорологии.

The publication includes papers dealing with the use of statistical methods for interpretation and treatment of different meteorological information. A number of results is presented concerning the problem of four — dimensional analysis of meteorological fields. There are considered problems of the accuracy of determining the statistical characteristics of meteorological fields, as well as application of the theory of excursions analysis of time meteorological series. The materials on the development and test of methods for checking meteorological information are given.

The publication is meant for researchers, post-graduates and senior students interested in the use of statistical methods in meteorology.

П $\frac{20807\text{-}139}{069(02)\text{-}75}$ 13-75(1)

Л. С. ГАНДИН, К. М. ЛУГИНА

ОБ УЧЕТЕ ДАННЫХ КОСВЕННОГО ЗОНДИРОВАНИЯ АТМОСФЕРЫ ПРИ ЧЕТЫРЕХМЕРНОМ АНАЛИЗЕ МЕТЕОРОЛОГИЧЕСКИХ ПОЛЕЙ

1. Как известно, сеть аэрологических станций, производящих зондирование атмосферы в заранее предписанные, единые физические моменты времени, распределена весьма неравномерно. Наряду с густой сетью таких станций в хорошо обжитых районах суши, существуют обширные области, главным образом акватории океанов, с весьма редкой сетью аэрологических станций. В южном полушарии количество таких станций в несколько раз меньше, чем в северном.

Перспективы расширения аэрологической сети весьма невелики. Не только организация новых кораблей погоды, производящих аэрологическое зондирование, но даже эксплуатация существующих кораблей погоды является весьма дорогостоящим мероприятием, и в настоящее время наблюдается тенденция не к увеличению, а, скорее, к сокращению флота кораблей погоды.

Вместе с тем задачи увеличения точности и заблаговременности прогнозов погоды требуют достаточно детальной информации об аэрологических полях над всем земным шаром или, по крайней мере, над полушарием, для которого выполняется прогноз. Такая информация настоятельно требуется также для достижения лучшего понимания закономерностей метеорологических процессов над океанами, в тропической и экваториальной зонах, в полярных районах и т. п., что в свою очередь необходимо для усовершенствования существующих прогностических методов.

Можно считать общепризнанным, что основным средством преодоления указанной трудности является привлечение наблюдений с быстро движущихся измерительных средств и, в первую очередь, с метеорологических спутников. Возможности использования спутниковой информации в дополнение к данным аэрологического зондирования базируются главным образом на так называемых методах косвенного зондирования атмосферы. Как показали работы последних лет, можно, используя решения обратных задач атмосферной оптики или основываясь на корреляционных связях, с удовлетворительной точностью восстанавливать вертикальные профили температуры и геопотенциала изобарических поверхностей по данным спутниковых измерений яркости уходящего излучения в различных спектральных интервалах.

Данные косвенного зондирования относятся к быстро следующим друг за другом и не фиксированным заранее срокам наблюдений, т. е. носят, как принято говорить, асиноптический характер. Такая асиноптическая информация уже сейчас распространяется в оперативном порядке. Удельный вес ее будет несомненно возрастать.

Для того чтобы учитывать (усваивать) асиноптическую информацию при объективном анализе метеорологических полей необходимо при таком анализе уметь использовать данные, относящиеся не только к разным точкам пространства, но и к разным моментам времени. Иначе говоря, необходим переход от применяемого в настоящее время чисто пространственного, трехмерного анализа метеорологических полей к анализу четырехмерному, т. е. пространственно-временному.

Проблема четырехмерного анализа была сформулирована уже несколько лет назад. С тех пор выполнен ряд исследований по четырехмерному анализу главным образом на уровне численных экспериментов. Результаты этих исследований представляют несомненный интерес. Однако основное внимание в них уделялось асиноптическому характеру данных косвенного зондирования.

Вместе с тем данные косвенного зондирования обладают еще по меньшей мере тремя особенностями, отличающими их от данных прямого аэрологического зондирования. Во-первых, косвенное зондирование дает, в отличие от прямого, осредненные по пространству величины, или, точнее говоря, масштаб такого осреднения существенно больше, чем для величин, полученных обычным аэрологическим зондированием. Во-вторых, аппаратура на спутнике работает в значительно более сложных условиях, чем при радиозондировании, а переход от спектральных яркостей к температуре и геопотенциалу является, разумеется, приближенным. Поэтому ошибки данных косвенного зондирования существенно больше, чем ошибки данных радиозондирования. Наконец, в третьих, если каждое новое радиозондирование производится с помощью нового экземпляра прибора, то все измерения на одном спутнике ведутся одним и тем же экземпляром аппаратуры. Поэтому в отличие от ошибок радиозондирования ошибки косвенного зондирования в различных точках должны тесно коррелировать между собой.

Указанные свойства ошибок косвенного зондирования не только следуют из умозрительных соображений. Они подтверждаются также результатами непосредственного анализа данных косенного зондирования, выполненного недавно В. П. Тараканоой [3].

2. Цель настоящей работы состоит в анализе возможностей етырехмерного анализа с учетом как асиноптического характера данных косвенного зондирования атмосферы, так и величины и степени коррелированности ошибок косвенного зондированияя. При этом предполагается, что четырехмерное усвоение инормации производится чисто статистическим путем, с помощью ространственно-временной оптимальной интерполяции.

Как известно (см., например, [1]), в случае, если ошибки аблюдений могут коррелировать между собой, но не коррелирурт с истинными значениями наблюдаемого метеорологического лемента f, уравнения метода оптимальной интерполяции для пределения весовых множителей P_i имеют вид

$$\sum_{j=1}^{n}(\mu_{ij}+\eta_i\eta_j\nu_{ij})P_j=\mu_{0i} \qquad (i=1,\ 2,\ \ldots,\ n), \tag{1}$$

де μ_{ij} — коэффициент корреляции между истинными значениями элемента f в i-том и j-том пунктах наблюдений; μ_{0i} — коэфициент корреляции между истинным значением элемента f i-том пункте наблюдений и искомым значением этого элемента; — число данных, используемых при интерполяции; η_i^2 — мера шибки наблюдений в i-том пункте, т. е. отношение среднего вадрата ошибки Δ_i^2 к дисперсии σ^2 элемента f; ν_{ij} — коэффициент корреляции между ошибками наблюдений в i-том и j-том унктах.

После того как веса P_i найдены путем решения системы уравений (1), нетрудно произвести интерполяцию по формуле

$$\hat{f}'_0=\sum_{i=1}^{n}P_i\tilde{f}'_i, \tag{2}$$

е величины со штрихом означают отклонения элемента f от реднего климатического значения (нормы) $\bar{f}$, знак $\sim$ относится наблюдаемым значениям элемента (в отличие от истинных), знак $\wedge$ — к результату интерполяции (также в отличие от стинного значения).

Независимо от этого можно после решения системы (1) оценить среднюю квадратическую ошибку интерполяции. С этой елью проще всего воспользоваться формулой

$$\varepsilon^2=1-\sum_{i=1}^{n}\mu_{0i}P_i, \tag{3}$$

де ε^2 — мера ошибки интерполяции, т. е. отношение среднего вадрата ошибки интерполяции к σ^2.

При этом предполагается, что величина σ является постоян ной. Было бы очень просто отказаться от этого упрощающег предположения, в чем, однако, нет нужды, поскольку данные на блюдения используются в сравнительно малой окрестности инте ресующей нас точки и в пределах такой окрестности величина σ меняется весьма мало.

Как известно, пространственные корреляционные функции ос новных метеорологических элементов можно с большой точностью считать однородными и изотропными по горизонтали или вдоль изобарических поверхностей, т. е. принять, что

$$\mu_{ij} = \mu(r_{ij}), \tag{4}$$

где r_{ij} — расстояние между i-тым и j-тым пунктами, а $\mu(r)$ — функция заданного вида.

Формула (4) верна в случае, если оба наблюдения относятся к одной и той же изобарической поверхности (что мы всегда бу дем предполагать) и если оба наблюдения выполнены в один и то же момент времени. Как показано в работе [2], с большой точностью выполняется более общая гипотеза о пространственно временной (точнее говоря, горизонтально-временной) однородно сти и изотропии. Именно, если расстояние между пунктами на блюдения равно r_{ij}, а интервал времени между ними составляе τ_{ij}, то можно воспользоваться формулой

$$\mu_{ij} = \mu\left(\sqrt{r_{ij}^2 + c^2 \tau_{ij}^2}\right), \tag{5}$$

где c — константа размерности скорости. Для приземного давле ния $c \simeq 35$ км/ч. В дальнейших расчетах будет использован именно это значение c.

Примем теперь, что часть исходных данных (например, дан ные пунктов с номерами $i=1, 2, \ldots, k$) представляет собой ре зультаты обычных радиозондовых наблюдений, а часть (с номера ми $i=k+1, k+2, \ldots, n$) — данные косвенного зондирования. Ошиб ки радиозондовых наблюдений будем, как это обычно делается считать белым шумом, т. е. предположим, что они не коррелиру ют ни между собой, ни с ошибками косвенного зондирования

$$\nu_{ij} = \begin{cases} 1 & \text{при } j = i; \\ 0 & \text{при } i = 1, 2, \ldots, k; \quad j = 1, 2, \ldots, n; \quad j \neq i; \\ 0 & \text{при } j = 1, 2, \ldots, k; \quad i = 1, 2, \ldots, n; \quad i \neq j; \end{cases} \tag{6}$$

Средний квадрат, а потому и меру ошибок радиозондировани будем считать одинаковыми для всех пунктов радиозондировани

$$\eta_1 = \eta_2 = \ldots = \eta_k = \eta. \tag{7}$$

Также одинаковой, но другой будем считать меру ошибок косвенного зондирования

$$\eta_{k+1}=\eta_{k+2}= \dots =\eta_n=\eta'. \tag{8}$$

Тогда можно записать систему (1) в следующем виде:

$$\begin{gathered}\sum_{j=1}^{n}\mu_{ij}P_j+\eta^2 P_i=\mu_{0i} \quad (i=1,\ 2,\ \dots,\ k),\\ \sum_{j=1}^{k}\mu_{ij}P_j+\sum_{j=k+1}^{n}\left(\mu_{ij}+\eta'^2\nu_{ij}\right)P_j=\mu_{0i}\\ (i=k+1,\ k+2,\ \dots,\ n).\end{gathered} \tag{9}$$

Например, при $k=2$, $n=5$ матрица коэффициентов системы (9) имеет вид (с учетом того, что $\mu_{ji}=\mu_{ij}$ и $\nu_{ji}=\nu_{ij}$).

$$\begin{pmatrix} 1+\eta^2 & \mu_{12} & \mu_{13} & \mu_{14} & \mu_{15}\\ \mu_{12} & 1+\eta^2 & \mu_{23} & \mu_{24} & \mu_{25}\\ \mu_{13} & \mu_{23} & 1+\eta'^2 & \mu_{34}+\eta'^2\nu_{34} & \mu_{35}+\eta'^2\nu_{35}\\ \mu_{14} & \mu_{24} & \mu_{34}+\eta'^2\nu_{34} & 1+\eta'^2 & \mu_{45}+\eta'^2\nu_{45}\\ \mu_{15} & \mu_{25} & \mu_{35}+\eta'^2\nu_{35} & \mu_{45}+\eta'^2\nu_{45} & 1+\eta'^2 \end{pmatrix}$$

Что касается коэффициентов ν_{ij} корреляции между ошибками косвенного зондирования, то они принимались зависящими лишь от расстояния между точками. При этом представляет интерес предельный случай, когда эти ошибки носят характер «черного шума», т. е. когда при всех i и j, превосходящих k,

$$\nu_{ij}=1 \quad (i,\ j=k+1,\ k+2,\ \dots,\ n). \tag{10}$$

3. Первая серия численных экспериментов относилась к случаю интерполяции по данным только асиноптических наблюдений. Предполагалось, что используется n таких наблюдений, расположенных на одной прямой AB и отстоящих друг от друга на расстояние r (рис. 1). Расстояние от этой прямой, моделирующей траекторию спутника, до интересующей нас точки 0 обозначено через ρ. При этом перпендикуляр, опущенный из точки 0 на прямую AB, пересекает ее в средней точке наблюдений. Такое предположение вполне естественно, поскольку можно считать, что имеется большое количество наблюдений на прямой AB, а для интерполяции в точку 0 используется лишь часть из них.

Считалось также, что наблюдения отстоят по времени от того момента, для которого нас интересует значение метеорологического элемента в точке 0, на некоторый интервал времени τ. Впрочем, поскольку в данном случае введение τ почти в точности

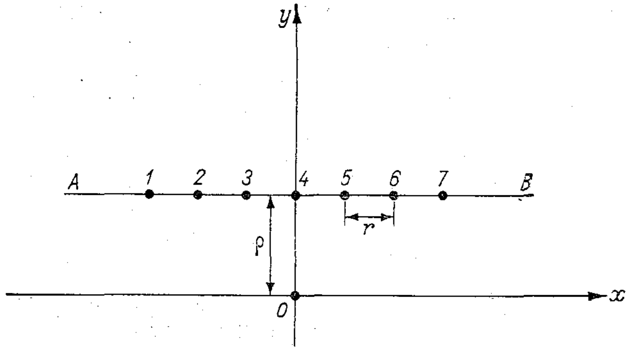


Рис. 1. Схема расположения пунктов асиноптических наблюдений.

эквивалентно соответствующему увеличению расстояния ρ, в дальнейшем анализируются результаты расчетов лишь при $\tau=0$.

Корреляционная функция анализируемого элемента задавалась формулой

$$\mu(r) = e^{-r/1,05}(1 + r/1,05) \tag{11}$$

(r в тысячах километров).

Как показано в работе [2], эта формула хорошо описывает, в частности, корреляционную функцию приземного давления. Корреляционная функция ошибок наблюдений считалась экспоненциальной

$$\nu(r) = e^{-ar}, \tag{12}$$

причём параметр a варьировался, так же как и мера ошибок наблюдения η'^2.

На рис. 2 показаны некоторые результаты вычислений средней квадратической относительной ошибки интерполяции ε для случая пяти влияющих точек ($n=5$) с расстоянием $r=200$ км между ними. Величина ε при различных значениях η'^2 и ρ представлена в функции от величины $\nu(0,2)$, т. е. от коэффициента корреляции между ошибками соседних наблюдений. Естественно, что с ростом $\nu(0,2)$ ошибка интерполяции возрастает, причем этот рост выражен тем сильней, чем больше η'. Величины ε, разумеется, увеличиваются с ростом расстояния от точки 0 до траектории AB. При этом относительное влияние коррелированности ошибок наблюдений уменьшается. В нашем случае это влияние наибольшее при $\rho=150$ км. Так, если мера ошибок наблюдения составляет 0,05, что является вполне разумной оценкой точности спутниковой информации, то при $\rho=150$ км переход от некоррелированных ошибок ($\nu=0$) к полностью коррелированным ($\nu=1$) соответствует увеличению ошибки интерполяции от 0,178 до 0,255, т. е. на 43%. Таким образом, наличие корреляции между ошибками асиноптических данных существенно уменьшает их информативность.

Это уменьшение наглядно проявляется при анализе влияния числа n учитываемых наблюдений на точность интерполяции. Так, в табл. 1 приведены результаты вычисления величины

$$\chi=\frac{\varepsilon(1)-\varepsilon(n)}{\varepsilon(1)}, \tag{13}$$

характеризующей относительное уменьшение ошибки интерполяции вследствие перехода от использования одного наблюдения к использованию n наблюдений. Такое уменьшение наиболее существенно при переходе от $n=1$ к $n=3$. При дальнейшем увеличении числа используемых наблюдений ε меняется в меньшей мере. При этом если в случае некоррелированных ошибок привлечение большого числа наблюдений приводит к существенному уменьшению ошибки интерполяции, то уже при $\nu(0,2)=0,7$ это повышение точности гораздо меньше. Еще меньше оно в случае наблюдений с полностью коррелированными ошибками.

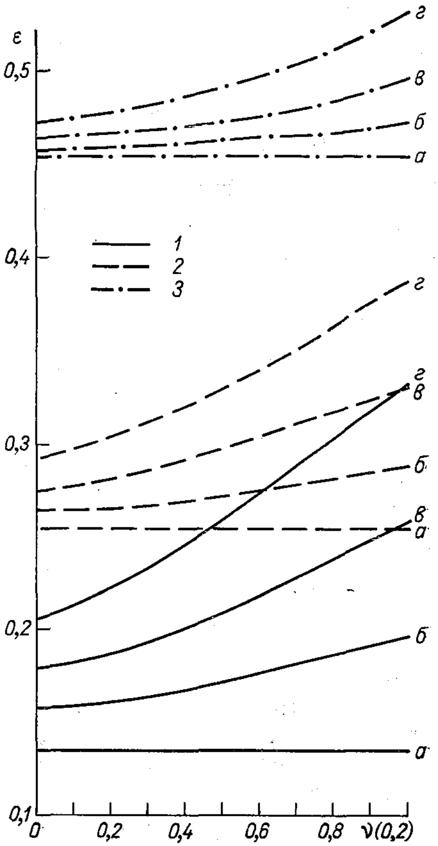


Рис. 2. Зависимость средней квадратической относительной ошибки информации от $\nu(0,2)$ и η'^2 ($n=5$, $r=200$ км) при различных значениях ρ.

1) $\rho=150$ км, *2)* $\rho=300$ км, *3)* $\rho=600$; *а)* $\eta'^2=0,0$; *б)* $\eta'^2=0,02$; *в)* $\eta'^2=0,05$; *г)* $\eta'^2=0,1$.

Сказанное означает, что в отличие от данных обычных наблюдений, для которых, как известно, увеличение числа влияющих станций в разумных пределах (до $n=6\sim8$) приводит к заметному уточнению анализа, применительно к асиноптической *информации это не так. Поскольку ошибки асиноптических данных* в разных пунктах существенно коррелированы, увеличение

Таблица 1

Относительное уменьшение (%) ошибки интерполяции при переходе от $n=1$ к $n=3$, $n=5$ и $n=9$; $\rho=300$ км; $r=200$ км; $\tau=0$

n	η'^2					
	0,01	0,02	0,07	0,10	0,15	0,20
			$\nu=0$			
3	5,1	8,2	17,5	20,5	23,4	25,1
5	5,6	9,3	20,9	24,8	29,0	31,5
9	5,6	9,3	21,4	25,6	30,6	33,7
			$\nu(0,2)=0,7$			
3	2,8	3,4	6,2	6,9	7,6	7,9
5	3,1	4,1	8,4	9,7	11,1	11,8
9	3,1	4,1	8,9	10,7	12,6	13,8

количества влияющих пунктов мало улучшает дело, и практически единственным путем, ведущим к уточнению анализа, является уменьшение ошибок асиноптической информации.

До сих пор предполагалось, что при вычислении интерполяционных весов каждый раз правильно учитывается наличие или отсутствие корреляции между ошибками наблюдений. Наряду с такими оценками представляет интерес выяснить, насколько ухудшаются результаты анализа, если между ошибками наблюдений существует корреляция, которая, однако, не принимается во внимание при расчете интерполяционных весов. Некоторые характерные результаты оценок такого рода приведены в табл. 2, где наряду со значениями ε при $\nu=0$ и при $\nu=1$ даны также величины ε', которые получаются при $\nu=1$, если используются веса P_i, рассчитанные в предложении, что $\nu=0$. Заметим, что в последнем случае интерполяция уже не является оптимальной, так что для расчета ε' нельзя пользоваться формулой (3), а необходимо использовать выражение

$$\varepsilon'^2=1-2\sum_{i=1}^{n}\mu_{0i}P_i+\sum_{i=1}^{n}\sum_{j=1}^{n}(\mu_{ij}+\eta'^2)P_iP_j. \tag{14}$$

Рассмотрение табл. 2 показывает, что в данном случае пренебрежение коррелированностью ошибок при расчете интерполяционных весов лишь незначительно ухудшает результаты анализа. Это особенно отчетливо видно из сопоставления данных в двух последних строках таблицы, в которых приведены относительные увеличения ε вследствие коррелированности ошибок наблюдения (предпоследняя строка) и дополнительные относительные увеличения вследствие неучета этого обстоятельства при расчете весов. Мы видим, что последние на порядок меньше первых.

Таблица 2

Средняя квадратическая относительная ошибка интерполяции при $\rho=150$ км, $r=200$ км, $\tau=0$, $n=5$ в случае некоррелированных ошибок наблюдений ($\varepsilon|_{\nu=0}$), в случае абсолютно коррелированных ошибок наблюдений ($\varepsilon|_{\nu=1}$) и в случае, когда ошибки наблюдений абсолютно коррелированы, но предполагаются некоррелированными в процессе вычисления интерполяционных весов (ε')

n	η'^2			
	0,01	0,02	0.05	0,10
$\varepsilon|_{\nu=0}$	0,148	0,157	0,178	0,205
$\varepsilon|_{\nu=1}$	0,167	0,194	0,255	0,327
ε'	0,170	0,199	0,266	0,346
$\frac{\varepsilon|_{\nu=1}-\varepsilon|_{\nu=0}}{\varepsilon|_{\nu=0}}\%$	12,8	23,6	43,3	59,5
$\frac{\varepsilon'-\varepsilon|_{\nu=1}}{\varepsilon|_{\nu=1}}\%$	1,8	2,6	4,3	5,8

Это означает, что имея дело только с асиноптической информацией, мы не получим больших ошибок анализа, если при расчете интерполяционных весов будем пренебрегать коррелированностью ошибок этой информации. Однако такое пренебрежение приведет к совершенно искаженным, завышенным оценкам точности анализа.

4. Перейдем теперь к результатам экспериментов, в которых предполагалось наличие как асиноптической информации, так и данных обычных, радиозондовых наблюдений.

Схема взаимного расположения пунктов наблюдений, использованная в этих экспериментах, показана на рис. 3. Предполагалось, что имеются 4 станции радиозондирования (точки с номерами $i=1, 2, 3, 4$), расположенные на одинаковом расстоянии h от интересующей нас точки 0 в направлениях x и y. Кроме того, используются асиноптические данные на траектории, параллельной оси x и отстоящей от точки 0 на расстояние ρ, причем эти данные относятся к моменту времени, отстоящему от рассматриваемого срока радиозондирования, для которого производится интерполяция, на интервал времени τ.

Эксперименты, о которых пойдет речь, относятся к случаю, когда используется 5 точек на траектории (точки с номерами $i=5$, 6, 7, 8, 9), на расстоянии $r=200$ км одна от другой (см. (рис. 3). В этих экспериментах расстояние ρ задавалось постоян-

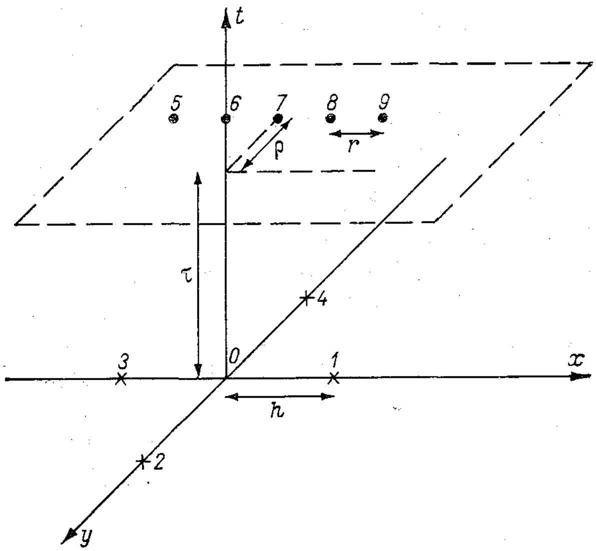


Рис. 3. Схема расположения пунктов наблюдений при использовании аэрологической и асиноптической информации.

1—4 — пункты аэрологической информации; *5—9* — пункты асиноптической информации; *0* — точка, в которой оценивается погрешность интерполяции.

ным, а именно $\rho=300$ км. Также постоянной принималась точность радиозондирования ($\eta^2=0{,}02$). Варьировались густота сети радиозондирования, т. е. расстояние h, сдвиг по времени τ, точность асиноптической информации, характеризуемая величиной η'^2, и степень коррелированности ошибок этой информации, описываемая коэффициентом корреляции $\nu(0{,}2)$ между ошибками в соседних точках траектории.

Рассмотрим некоторые результаты этих экспериментов, приведенные в табл. 3, а именно, суммы весов аэрологической информации, суммы весов асиноптической информации и средние квадратические относительные ошибки интерполяции. Эти результаты приведены в табл. 3 для двух значений h (200 и 800 км), трех значений τ (0; 3 и 6 ч), трех значений η'^2 (0,01; 0,1 и 0,2) и двух значений $\nu(0{,}2)$ (0 и 0,7).

При расстоянии h, равном 200 км, что соответствует расстоянию между соседними аэрологическими станциями около 280 км, вес аэрологической информации во всех случаях получается весьма большим, а вес асиноптических данных — почти пренебрежимо малым. Это вполне естественно, поскольку при таких расстоя-

Таблица 3

Сумма весов аэрологической ($\sum_{i=1}^{4} P_i$) и асиноптической ($\sum_{i=5}^{9} P_i$) информации и средняя квадратическая относительная ошибка интерполяции

τ ч	h км	η'^2	$\nu=0$			ν (0,2)=0,7			
			$\sum_{i=1}^{4} P_i$	$\sum_{i=5}^{9} P_i$	ε	$\sum_{i=1}^{4} P_i$	$\sum_{i=5}^{9} P_i$	ε	ε'
0	200	0,01	0,912	0,096	0,084	0,924	0,082	0,084	0,084
		0,1	0,962	0,050	0,086	0,987	0,021	0,086	0,087
		0,2	0,979	0,045	0,087	0,997	0,012	0,087	0,087
	800	0,01	0,034	0,988	0,165	0,146	0,880	0,176	0,177
		0,1	0,251	0,794	0,210	0,585	0,478	0,247	0,286
		0,2	0,427	0,628	0,234	0,762	0,304	0,269	0,298
3	200	0,01	0,936	0,070	0,085	0,944	0,062	0,085	0,085
		0,1	0,970	0,040	0,086	0,989	0,019	0,086	0,087
		0,2	0,982	0,028	0,087	0,998	0,011	0,087	0,087
	800	0,01	0,157	0,876	0,184	0,256	0,779	0,192	0,193
		0,1	0,323	0,726	0,220	0,618	0,436	0,251	0,285
		0,2	0,476	0,583	0,240	0,778	0,296	0,271	0,296
6	200	0,01	0,970	0,038	0,086	0,973	0,036	0,086	0,086
		0,1	0,982	0,026	0,086	0,994	0,014	0,087	0,087
		0,2	0,988	0,020	0,087	1,000	0,009	0,087	0,087
	800	0,01	0,406	0,645	0,219	0,477	0,575	0,221	0,222
		0,1	0,487	0,576	0,242	0,698	0,376	0,262	0,284
		0,2	0,588	0,511	0,255	0,820	0,264	0,274	0,292

ниях аэрологические данные позволяют получить весьма значительную информацию об искомой величине в точке 0, так что асиноптические наблюдения почти не дают дополнительной информации. Соотношение весов аэрологической и асиноптической информации закономерно меняется с изменением параметров наших экспериментов, а именно $\sum_{i=1}^{4} P_i$ растет, а $\sum_{i=5}^{9} P_i$ соответственно убывает с ростом ошибок асиноптических данных, с увеличением коррелированности этих ошибок и с ростом сдвига по времени τ. Эти эффекты, однако, выражены при $h=200$ км весьма слабо. Что касается точности интерполяции, характеризуемой ε, то она практически совсем не зависит от указанных параметров задачи.

Таким образом, в районах с густой аэрологической сетью, где достаточно высокая точность анализа может быть обеспечена путем использования только обычных наблюдений, асиноптические данные не могут заметным образом изменить результат анализа и увеличить его точность.

Существенно иная картина получается при редкой сети аэрологического зондирования, а именно при сети с расстоянием около 1100 км ($h=800$ км). В этом случае соотношение весов аэрологической и асиноптической информации сильно меняется в зависимости от параметров рассматриваемой задачи.

Так, если сдвиг асиноптической информации по времени отсутствует ($\tau=0$), а ошибки ее весьма малы ($\eta'^2=0{,}01$) и не коррелируют друг с другом ($\nu=0$), то ничтожно малым оказывается вес аэрологических данных, а асиноптические данные входят с суммарным весом, близким к единице. Увеличение η', ν и τ приводит к значительному уменьшению роли асиноптических данных и к соответственному возрастанию роли аэрологической информации. Так, при $\eta'^2=0{,}2$, $\nu(0{,}2)=0{,}7$ и $\tau=6$ ч суммарный вес аэрологических данных оказывается превосходящим вес асиноптических данных более чем в три раза.

Таким образом, в районах с редкой сетью аэрологических станций вклад асиноптической информации может быть весьма значительным и для корректного определения этого вклада необходимо надлежащим образом учитывать свойства асиноптической информации. Среди этих свойств, недостаточный учет которых может сильно исказить выводы, касающиеся соотношения вкладов разных видов информации, следует особенно отметить наличие корреляции между ошибками асиноптических данных. Так, при $\eta'^2=0{,}1$ и $\tau=0$ неучет указанного фактора приводит к выводу, что вес асиноптических данных должен более чем втрое превышать вес данных радиозондирования. В то же время при разумной оценке корреляции между ошибками спутниковой информации соотношение существенно иное — суммарный вес аэрологической информации даже несколько больше, чем асиноптической. Рассмотрение величин ε' в табл. 3, имеющих такой же смысл, как и ранее (см. табл. 2), показывает, что принятие неправильного соотношения между весами приводит также к заметному уменьшению точности анализа.

Увеличение ошибки анализа ε' вследствие пренебрежения коррелированностью ошибок асиноптической информации имеет тот же порядок величины, что и увеличение ε вследствие самого факта коррелированности этих ошибок.

Заслуживает внимания также тот факт, что приведенные оценки сравнительно слабо зависят от сдвига по времени τ. Так при $h=800$ км и $\eta'^2=0,1$ увеличение τ от 0 даже до 6 ч приводит (при коррелированных ошибках асиноптической информации) к росту суммарного веса аэрологических данных и уменьшению веса спутниковой информации примерно на 0,1. Ошибка анализа возрастает при этом только на 6%. При переходе же от $\tau=0$ к $\tau=3$ ч этот эффект значительно меньше.

Этот вывод представляется весьма существенным. Он означает, что при разработке методов четырехмерного анализа не столь уж существенно стремиться к тому, чтобы данные асиноптических наблюдений утилизировались в точности для тех моментов времени, к которым эти данные относятся; на это обстоятельство обращается основное внимание в современных численных экспериментах по четырехмерному усвоению данных. Гораздо важнее правильно учитывать статистические свойства ошибок асиноптической информации, а именно истинную среднюю величину этих ошибок и степень их коррелированности.

СПИСОК ЛИТЕРАТУРЫ

1. Гандин Л. С., Каган Р. Л., Полищук А. И. Об оценке информативности систем метеорологических наблюдений.—«Труды ГГО», 1972, вып. 286, с. 120—140.
2. Лугина К. М., Каган Р. Л. К вопросу о пространственно-временно́м анализе барического поля. — «Труды ГГО», 1974, вып. 336, с. 75—94.
3. Тараканова В. П. К вопросу о точности косвенного температурного зондирования атмосферы с искусственных спутников Земли. — «Метеор. и гидрол.», 1974, № 4, с. 76—78.

В. А. ШАХМЕЙСТЕР

ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ ПО СОВМЕСТНОМУ УЧЕТУ АЭРОЛОГИЧЕСКОЙ, СПУТНИКОВОЙ И ПРОГНОСТИЧЕСКОЙ ИНФОРМАЦИИ

Общеизвестно, что существующая в настоящее время сеть аэрологических станций весьма неравномерна. Поэтому точность объективного анализа в районах с редкой сетью довольно низкая, в то время как современные методы численных прогнозов требуют задания полей метеоэлементов по всему земному шару или по полушарию с достаточной степенью точности.

Поэтому в настоящее время назрела необходимость применения дополнительной информации при объективном анализе метеорологических полей с целью повышения точности анализа.

Рассмотрим вопрос о привлечении в качестве дополнительной информации данных, полученных с метеорологических спутников, и прогностической информации. В отличие от аэрологической эти два вида информации имеют свои особенности.

Во-первых, в отличие от аэрологических данных, спутниковые данные разбросаны непрерывно по пространству и времени, т. е. носят, как принято говорить, асиноптический характер. Поэтому при анализе асиноптических данных следует производить четырехмерное (пространственно-временное) усвоение данных наблюдений.

Во-вторых, существующие методы обработки спутниковых данных позволяют получать вертикальные профили температуры и геопотенциала по данным спектральных измерений уходящего излучения. Естественно, что ошибка измерения при таком косвенном зондировании атмосферы значительно больше, чем при прямом зондировании.

В-третьих, вследствие того, что все измерения на спутнике производятся одним и тем же экземпляром прибора, ошибки спутниковых данных коррелируют между собой. Результаты проверки всех этих свойств на фактическом материале полностью подтверждают высказанные выше предположения [1].

Перейдем теперь к особенностям прогностической информаии. Будем считать прогноз совершенным в том смысле, что икакой линейной комбинацией результатов прогноза его нелья улучшить. Из этого следует, что ошибки прогноза не коррелиуют с результатами прогноза и, следовательно, обладают отицательной корреляцией с истинными значениями метеоэлемена. Поэтому дисперсия прогноза должна быть занижена. В этом эстоит основное отличие ошибок прогноза от ошибок наблюдеий, которые не коррелируют с истинными значениями и завышат дисперсию метеоэлемента. При этом прогноз должен учитыаться только в той точке, куда производится интерполяция [2].

В настоящей работе основное внимание уделяется учету стаистических характеристик различных видов дополнительной инормации при четырехмерном анализе, а не их асиноптическому арактеру. Мы предполагали, что дополнительная информация сваивается в тот же момент времени, для которого производитя интерполяция. Это предположение не является ограничением, оскольку если бы асиноптические наблюдения отстояли от расматриваемого момента времени на интервал τ, то это было бы квивалентно увеличению расстояния от них до точки, куда проводится интерполяция, а это расстояние варьировалось в наих экспериментах [3]. Усвоение данных производилось чисто татистическим путем с помощью метода оптимальной интерполяии.

Итак, пусть в точку *0* производится интерполяция по данным аблюдений в n точках.

Тогда, если не делать никаких предположений относительно орреляции ошибок наблюдений как с истинным значением меоэлемента, так и между собой, уравнения метода оптимальной нтерполяции можно записать в виде

$$\sum_{j=0}^{n} (\overline{f'_i f'_j} + \overline{f'_i \delta_j} + \overline{\delta_i f'_j} + \overline{\delta_i \delta_j}) p_j = \overline{f'_i f'_0} + \overline{\delta_i f'_0}$$

$$(i = 0,\ 1,\ \ldots,\ n), \qquad (1)$$

це f'_i — отклонение истинного значения метеоэлемента от нормы i-той точке, δ_i — ошибка i-того наблюдения, p_j — весовой мноитель, черта сверху означает статистическое осреднение.

Выражение (1) представляет собой систему $n+1$ линейных авнений относительно весовых множителей p_j. Допустим теэрь, что при $i=0$ информация прогностическая, при $i=1, 2, \ldots, k$ — эрологическая, а при $i=k+1, k+2, \ldots, n$ — спутниковая.

Запишем систему (1) в виде:

$$\sum_{j=0}^{n} (\overline{f'_0 f'_j} + \overline{f'_0 \delta_j} + \overline{\delta_0^{(0)} f_j^0} + \overline{\delta_0^{(0)} f'_j}) p_j = \overline{f_0'^2} + \overline{\delta_0^{(0)} f'_0};$$

$$\sum_{j=0}^{n}\left(\overline{f_i' f_j'}+\overline{f_i'\delta_j}+\overline{\delta_i f_j'}+\overline{\delta_i\delta_j}\right)p_j=\overline{f_i' f_0'}+\overline{\delta_i f_0'}$$

$$(i=1,\ 2,\ \ldots,\ n), \qquad (2$$

где $\delta_0^{(0)}$ — ошибка прогноза.

Так как ошибки прогноза не коррелируют с результатами прогноза, то

$$\overline{\delta_0^{(0)}(f_j'+\delta_j^{(0)})}=0, \qquad (3$$

и, следовательно,

$$\overline{\delta_0^{(0)} f_j'}=-\overline{\delta_0^{(0)}\delta_j^{(0)}}. \qquad (4$$

Если учесть, что ошибки спутниковых данных коррелирую между собой, но не коррелируют ни с истинными значениями метеоэлемента, ни с ошибками аэрологических данных и прогноза, а ошибки аэрологических данных представляют собой белый шум, то система (2) примет вид:

$$\sum_{j=0}^{n}\left(\overline{f_0' f_j'}-\overline{\delta_0^{(0)}\delta_j^{(0)}}\right)p_j=\overline{f_0'^2}-\overline{\delta_0^{(0)2}};$$

$$\left(\overline{f_i' f_0'}-\overline{\delta_i^{(0)}\delta_0^{(0)}}\right)p_0+\sum_{j=1}^{n}\left(\overline{f_i' f_j'}\right)p_j+\overline{\delta_i^2}p_i=\overline{f_i' f_0'} \quad (i=1,\ 2,\ \ldots,\ k);$$

$$\left(\overline{f_i' f_0'}-\overline{\delta_i^{(0)}\delta_0^{(0)}}\right)p_0+\sum_{j=1}^{k}\left(\overline{f_i' f_j'}\right)p_j+\sum_{j=k+1}^{n}\left(\overline{f_i' f_j'}+\overline{\delta_i\delta_j}\right)p_j=\overline{f_i' f_0'}$$

$$(i=k+1,\ k+2,\ \ldots,\ n). \qquad (5$$

Будем считать, что дисперсия метеоэлемента постоянна и равна σ^2. Это предположение естественно, так как рассматриваема окрестность точки 0 сравнительно мала. Поделив все уравнени системы (5) на σ^2, получим:

$$\sum_{j=0}^{n}\left(\mu_{0j}-\eta_0^{(0)}\eta_j^{(0)}k_{0j}\right)p_j=1-\eta_0^{(0)2};$$

$$\left(\mu_{i0}-\eta_i^{(0)}\eta_0^{(0)}k_{i0}\right)p_0+\sum_{j=1}^{n}\mu_{ij}p_j+\eta_i^{(1)2}p_i=\mu_{0i} \quad (i=1,\ 2,\ \ldots,\ k);$$

$$\left(\mu_{i0}-\eta_i^{(0)}\eta_0^{(0)}k_{i0}\right)p_0+\sum_{j=1}^{k}\mu_{ij}p_j+\sum_{j=k+1}^{n}\left(\mu_{ij}+\eta_i^{(2)}\eta_j^{(2)}\nu_{ij}\right)p_j=\mu_{0i}$$

$$(i=k+1,\ k+2,\ \ldots,\ n), \qquad ($$

где μ_{ij} — коэффициент корреляции между истинными значениями элемента f в i-том и j-том пунктах наблюдений, k_{0j} — коэффициент корреляции между ошибками прогноза в 0-й и j-той точках, μ_{0i} — коэффициент корреляции между истинным значением элемента f в i-том пункте наблюдения и искомым значением этого элемента,

ν_{ij} — коэффициент корреляции между ошибками спутниковых наблюдений в i-том и j-том пунктах,

$\eta_i^{(0)2}$ — мера ошибки прогноза в i-том пункте,

$\eta_i^{(1)2}$ — мера ошибки аэрологических наблюдений,

$\eta_i^{(2)2}$ — мера ошибки спутниковых наблюдений.

Средний квадрат, а потому и меру ошибки наблюдений будем считать одинаковыми (для каждого вида информации). Поэтому

$$\eta_0^{(0)} = \eta_1^{(0)} = \eta_2^{(0)} = \ldots = \eta_n^{(0)} = \eta_0, \tag{7}$$

$$\eta_1^{(1)} = \eta_2^{(1)} = \ldots = \eta_k^{(1)} = \eta_1, \tag{8}$$

$$\eta_{k+1}^{(2)} = \eta_{k+2}^{(2)} = \ldots = \eta_n^{(2)} = \eta_2. \tag{9}$$

Таким образом, мера ошибки прогноза равна η_0^2, мера ошибки аэрологических данных равна η_1^2, а мера ошибки спутниковых данных — η_2^2.

Итак, систему (6) можно переписать в виде:

$$\sum_{j=0}^{n} \left(\mu_{0j} + \eta_0^2 k_{0j}\right) p_j = 1 - \eta_0^2;$$

$$\left(\mu_{i0} - \eta_0^2 k_{i0}\right) p_0 + \sum_{j=1}^{n} \mu_{ij} p_j + \eta_1^2 p_i = \mu_{0i} \quad (i = 1, 2, \ldots, k);$$

$$\left(\mu_{i0} - \eta_0^2 k_{i0}\right) p_0 + \sum_{j=1}^{k} \mu_{ij} p_j + \sum_{j=k+1}^{n} \left(\mu_{ij} + \eta_2^2 \nu_{ij}\right) p_j = \mu_{0i}$$
$$(i = k+1, k+2, \ldots, n). \tag{10}$$

Например, если имеются две аэрологические точки, две спутниковые и одна прогностическая, т. е. при $k=2$, $n=4$ мы получим систему уравнений с матрицей коэффициентов:

$$\begin{pmatrix} 1-\eta_0^2 & \mu_{01}-\eta_0^2 k_{01} & \mu_{02}-\eta_0^2 k_{02} & \mu_{03}-\eta_0^2 k_{03} & \mu_{04}-\eta_0^2 k_{04} \\ \mu_{10}-\eta_0^2 k_{10} & 1+\eta_1^2 & \mu_{12} & \mu_{13} & \mu_{14} \\ \mu_{20}-\eta_0^2 k_{20} & \mu_{21} & 1+\eta_1^2 & \mu_{23} & \mu_{24} \\ \mu_{30}-\eta_0^2 k_{30} & \mu_{31} & \mu_{32} & 1+\eta_2^2 & \mu_{34}+\eta_2^2 \nu_{34} \\ \mu_{40}-\eta_0^2 k_{40} & \mu_{41} & \mu_{42} & \mu_{43}+\eta_2^2 \nu_{43} & 1+\eta_2^2 \end{pmatrix} \tag{11}$$

Вектор правых частей системы запишется в виде

$$\left(1-\eta_0^2,\ \mu_{10},\ \mu_{20},\ \mu_{30},\ \mu_{40}\right). \qquad (12)$$

После решения системы (10) среднюю квадратическую ошибку интерполяции можно оценить по формуле

$$\varepsilon^{*2}=1-\sum_{i=1}^{n}\mu_{0i}p_i-\left(1-\eta_0^2\right)p_0, \qquad (13)$$

где ε^{*2} — мера ошибки интерполяции при наличии прогноза, p_i ($i=0,1, ..., n$) — решение системы (10).

В случае отсутствия какого-либо вида информации матрица системы метода оптимальной интерполяции может быть получена из (10) вычеркиванием соответствуюших строк и столбцов. Например, если нет прогностической информации, то в матрице (10) должны отсутствовать нулевая строка и нулевой столбец. Мера ошибки интерполяции при этом рассчитывается по формуле

$$\varepsilon^2=1-\sum_{i=1}^{n}\mu_{0i}\tilde{p}_i, \qquad (14)$$

где $\tilde{p}_i$ — соответствующие интерполяционные веса.

В дальнейшем всюду под ε^2 будем понимать меру ошибки интерполяции при отсутствии прогноза, а под ε^{*2} — при наличии прогноза.

Корреляционная функция истинных значений и корреляционные функции ошибок предполагались однородными и изотропными по горизонтали, т. е. предполагалось, что они являются функциями заданного вида от расстояния r между пунктами наблюдения.

Корреляционная функция μ анализируемого элемента задавалась формулой

$$\mu(r)=e^{-\frac{r}{1,05}}\left(1+\frac{r}{1,05}\right) \qquad (15)$$

(r — в тысячах километров [4]).

Корреляционная функция ошибок спутниковых наблюдений считалась равной

$$\nu(r)=e^{-ar}.$$

Параметр a был выбран таким образом, что $\nu\ (0,2)=0,7$. Корреляционная функция ошибок прогноза варьировалась в различных численных экспериментах. Все численные эксперименты проводились для фиксированного взаимного расположения станций, которое изображено на рис. 1. Точка 0, в которую производилась интерполяция, окружена четырьмя аэрологическими станциями, находящимися от нее на расстоянии h_1. На прямой *АБ*, которая расположена на расстоянии h_3 от точки 0 и моделирует траекторию спутника, находятся пять спутниковых точек на равных расстояниях h_2 одна от другой. При таком расположении станций была проведена серия численных экспериментов. Их целью было выяснить зависимость средней квадратической ошибки анализа

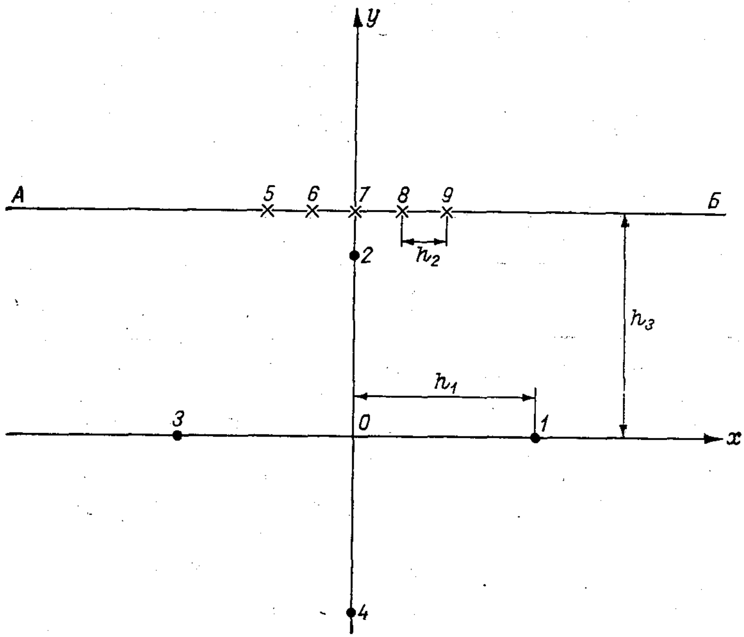


Рис. 1. Схема расположения станций.

1—4 — аэрологические станции; *5—9* — спутниковые точки; 0 — точка, куда производится интерполяция.

от различных параметров: от расстояния между станциями, от ошибки наблюдений, от корреляционных функций ошибок наблюдений, от количества и вида используемой информации, от правильности учета статистических характеристик различных видов информации при объективном анализе.

Перейдем к описанию результатов.

Первая серия численных экспериментов относится к выбору корреляционной функции ошибок прогноза. Эти функции в настоящее время изучены весьма слабо. Корреляционную функцию *К* ошибок прогноза следовало подобрать таким образом, чтобы матрица системы (10) для определения интерполяционных весов была бы положительно определена.

В работе [5] приведен пример корреляционной функции ошибок прогноза. Обозначим эту функцию через $K_1(r)$. При попытке использовать $K_1(r)$ в качестве K в системе (10) выяснилось, что при этом положительная определенность матрицы системы (10) нарушается.

Из рассмотрения матрицы системы (10), получающейся для четырех аэрологических точек и одной прогностической при $K=K_1(r)$, следует, что достаточным условием положительной определенности матрицы системы для определения интерполяционных весов является выполнение неравенства

$$\left(1-\eta_0^2\right) > \mu_{0i} - \eta_0^2 k_{0i} \quad (i=1,\ 2,\ 3,\ 4). \tag{16}$$

В проведенном численном эксперименте функция K подбиралась таким образом, чтобы неравенство (16) было выполнено,

Таблица 1

Влияние выбора корреляционной функции ошибок прогноза на точность интерполяции

N	n_1	n_2	n_3	K	$\sum_1^4 p_i$	$\sum_5^9 p_i$	p_0	ε	ε^*	Положительная определенность
1	4	0	1	$K_1(r)$	−0,969	—	2,046	0,306	0,604	нет
2	4	0	1	$K_1\left(\frac{r}{2}\right)$	—0,038	—	1,135	0,306	0,305	есть
3	4	0	1	$\mu(r)$	0,905	—	0,254	0,306	0,277	есть
4	4	5	1	$K_1(r)$	—0,696	0,965	0,732	0,222	0,398	нет
5	4	5	1	$K_1\left(\frac{r}{2}\right)$	0,412	0,644	—0,08	0,222	0,222	есть
6	4	5	1	$\mu(r)$	0,363	0,578	0,147	0,222	0,210	есть

и поэтому была обеспечена положительная определенность матрицы системы (10).

Результаты произведенных расчетов приведены в табл. 1, а также в табл. 2—5. В этих таблицах использованы следующие обозначения:

N — номер счета (номер строки),

n_1 — количество аэрологических точек,

n_2 — количество спутниковых точек,

n_3 — количество прогностических точек,

K — корреляционная функция ошибок прогноза,

$K_1\left(\frac{r}{2}\right)$ — корреляционная функция, получающаяся при растяжении $K_1(r)$ в два раза вдоль оси r.

$\sum_1^4 p_i$ — сумма весов аэрологической информации,

$\sum_5^9 p_i$ — сумма весов спутниковой информации,

p_0 — вес прогностической информации,

ε — относительная средняя квадратическая ошибка анализа при отсутствии прогностических данных,

ε^* — относительная средняя квадратическая ошибка анализа при наличии прогностической информации.

В табл. 1 приведены результаты расчета для $\eta_1^2=0{,}02$, $\eta_2^2=0{,}05$; $\eta_0^2=0{,}3$; $h_1=0{,}8$; $h_2=0{,}2$; $h_3=0{,}3$; $\nu(0{,}2)=0{,}7$; отмечено также наличие положительной определенности матрицы системы (10).

При расчетах в качестве корреляционной функции ошибок прогноза брались три функции: $K_1(r), K_1\left(\frac{r}{2}\right)$ и $\mu(r)$ (рис. 2).

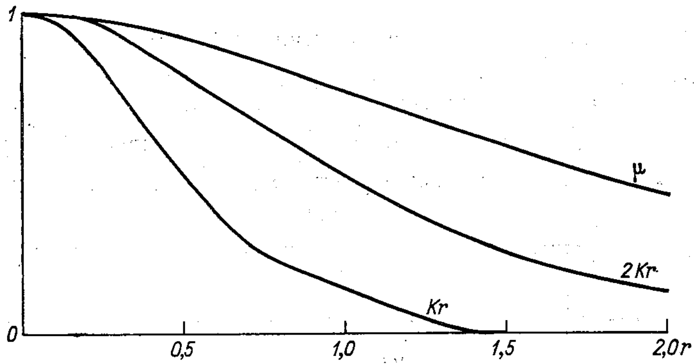


Рис. 2. Различные корреляционные функции ошибок прогноза

$Kr=K_1(r)$, $Kr=K_1\left(\frac{r}{2}\right)$, $\mu=\mu(r)$; r в тысячах километров.

В первых трех строках табл. 1 приведены данные для случая, когда отсутствует спутниковая информация, в следующих трех — для случая, когда присутствуют все три вида информации. В 1-й и 4-й строках таблицы представлены результаты, полученные при $K=K_1(r)$; при этом положительная определенность нарушается и $\varepsilon^*>\varepsilon$. Во 2-й и 5-й строках таблицы приведены результаты расчета ε и ε^* при $K(r)=K_1\left(\frac{r}{2}\right)$. Неравенство (16) выполнено, положительная определенность не нарушается и введение прогностической информации ведет к уменьшению средней квадратической ошибки анализа.

При рассмотрении неравенства (16) видно, что оно заведомо выполнено при $K=\mu$. Результаты численных расчетов это подтвердили (строки 3-я и 6-я табл. 1). При этом положительная определенность матрицы системы (10) сохраняется, и добавление прогностической информации ведет к уменьшению средней квадратической ошибки анализа. Поэтому во всех наших последующих расчетах в качестве корреляционной функции ошибок прогноза была взята корреляционная функция истинных значений μ.

Во второй группе численных экспериментов варьировались параметры η_0^2; η_2^2; h_1; h_3; n_1; n_2; n_3 и рассматривалось влияние вариации этих параметров на точность интерполяции.

Все эксперименты проводились для $\eta_1^2=0{,}02$ и $\nu(0{,}2)=0{,}7$ и при $h_2=0{,}2$. Результаты соответствующих расчетов приведены в табл. 2, 3, 4.

Из табл. 2 видно, что при наличии только спутниковых данных введение прогностической информации в объективный анализ дает значительно больший эффект, чем когда имеется и спутниковая, и аэрологическая информация. Из сравнения между собой случаев 1, 2, 3 и 4, 5, 6 видно, что чем хуже спутниковая информация (т. е. чем больше η_2^2), тем больше относительное уменьшение средней квадратической ошибки анализа вследствие введения прогностической информации. Это уменьшение более сильное при отсутствии аэрологических данных. Например, при $\eta_2^2=0{,}02$ $\frac{\varepsilon-\varepsilon^{2*}}{\varepsilon}=7{,}9\%$, а при $\eta_2^2=0{,}1$ $\frac{\varepsilon-\varepsilon^{2*}}{\varepsilon}=12{,}2\%$. Заметим также что с увеличением η_2^2 суммы весов $\sum_5^9 p_i$ спутниковой информации уменьшаются, а с уменьшением — увеличиваются, а сами величины ε и ε^* естественным образом меняются с изменением η_2^2, а именно, с увеличением η_2^2 они увеличиваются, а с уменьшением η_2^2 — уменьшаются.

Таблица 2

Суммы весов аэрологической $\sum_1^4 p_i$, спутниковой $\sum_5^9 p_i$ и прогностической информации p_0, средние квадратические ошибки интерполяции ε и ε^*, а также относительное уменьшение средней квадратической ошибки интерполяции $\frac{\varepsilon-\varepsilon^*}{\varepsilon}$ вследствие включения прогностической информации при $\eta_0^2=0{,}3$; $h_1=0{,}8$, $h_3=0{,}3$ в зависимости от меры ошибки спутниковых данных η_2^2

N	n_1	n_2	n_3	η_2^2	$\sum_1^4 p_i$	$\sum_5^9 p_i$	p_0	ε^*	ε	$\frac{\varepsilon-\varepsilon^*}{\varepsilon}$ %
1	4	5	1	0,02	0,208	0,745	0,113	0,184	0,192	4,1
2	4	5	1	0,05	0,363	0,578	0,147	0,210	0,222	5,4
3	4	5	1	0,1	0,512	0,418	0,178	0,231	0,247	6,4
4	0	5	1	0,02	—	0,823	0,220	0,257	0,279	7,9
5	0	5	1	0,05	—	0,784	0,261	0,280	0,310	9,6
6	0	5	1	0,1	—	0,725	0,319	0,309	0,351	12,2

В табл. 3 приведено относительное уменьшение ошибки интерполяции вследствие введения прогностической информации при различных значениях меры ошибки прогностической информации η_0^2. В случаях 1, 2, 3 присутствуют все три вида информации, в случаях 4, 5, 6 — аэрологическая и прогностическая, в случаях 7, 8, 9 — спутниковая и прогностическая. Из сравнения между собой величин p_0 в различных строках табл. 3 видно, что с увеличением η_0^2 вес прогностической информации p_0 уменьшается, т. е. уменьшается влияние прогностической информации на результат

Таблица 3

Суммы весов аэрологической, спутниковой и прогностической информации, средние квадратические ошибки интерполяции ε и ε^*, а также относительное уменьшение средней квадратической ошибки интерполяции $\frac{\varepsilon-\varepsilon^*}{\varepsilon}$ вследствие включения прогностической информации при $\eta_2^2=0{,}05$; $h_1=0{,}8$; $h_3=0{,}3$ в зависимости от меры ошибки прогностической информации η_0^2

N	n_1	n_2	n_3	η_0^2	$\sum_1^4 p_i$	$\sum_5^9 p_i$	p_0	ε^*	ε	$\frac{\varepsilon-\varepsilon^*}{\varepsilon}$ %
1	4	5	1	0,1	0,281	0,446	0,341	0,185	0,222	16,6
2	4	5	1	0,3	0,363	0,578	0,147	0,210	0,222	5,4
3	4	5	1	0,5	0,386	0,614	0,094	0,216	0,222	2,7
4	4	0	1	0,1	0,599	—	0,508	0,225	0,306	26,4
5	4	0	1	0,3	0,905	—	0,256	0,277	0,306	9,4
6	4	0	1	0,5	1,008	—	0,171	0,292	0,306	4,5
7	0	5	1	0,1	—	0,515	0,515	0,227	0,310	26,7
8	0	5	1	0,3	—	0,784	0,261	0,280	0,310	9,6
9	0	5	1	0,5	—	0,875	0,175	0,296	0,310	4,5

интерполяции, а при уменьшении η_0^2 p_0 увеличивается, т. е. влияние прогностической информации увеличивается, а аэрологической и спутниковой уменьшается. Средняя квадратическая ошибка анализа ε^* при увеличении η_0^2 увеличивается, а при уменьшении η_0^2 — уменьшается.

Рассматривая величины $\frac{\varepsilon - \varepsilon^*}{\varepsilon}$ для различных видов информации (например, для случаев 1, 4, 7), мы видим, что наибольшее уменьшение ошибки анализа получается, когда имеется только спутниковая информация, а наименьшее — при наличии и спутниковой и аэрологической информации.

Таким образом, если имеется и спутниковая, и аэрологическая информация, то добавление лишь очень точной прогностической информации может привести к существенному улучшению анализа. В случаях, когда имеются только аэрологические данные, и в случаях, когда имеются только спутниковые данные, введение прогноза ведет к существенному улучшению анализа.

Таблица 4

Суммы весов аэрологической, спутниковой и прогностической информации, средние квадратические ошибки интерполяции ε и ε^*, относительное уменьшение средней квадратической ошибки интерполяции $\frac{\varepsilon-\varepsilon^*}{\varepsilon}$ вследствие включения прогностической информации при $\eta_2^2=0,05$ $\eta_0^2=0,3$ в зависимости от h_1 и h_3

N	n_1	n_2	n_3	h_1	h_3	$\sum_1^4 p_i$	$\sum_5^9 p_i$	p_0	ε^*	ε	$\frac{\varepsilon-\varepsilon^*}{\varepsilon}\%$
1	4	5	1	0,2	0,3	0,955	0,035	0,024	0,085	0,086	1,1
2	4	5	1	0,4	0,3	0,881	0,105	0,060	0,134	0,137	2,2
3	4	5	1	0,8	0,3	0,363	0,578	0,147	0,210	0,222	5,4
4	4	5	1	0,8	0,6	0,572	0,342	0,226	0,260	0,284	8,4
5	4	5	1	0,8	0,9	0,993	—0,093	0,253	0,276	0,304	9,2
6	4	0	1	0,2	0,3	0,991	—	0,025	0,086	0,087	1,1
7	4	0	1	0,4	0,3	0,988	—	0,068	0,142	0,146	2,9
8	4	0	1	0,8	0,3	0,905	—	0,256	0,277	0,306	9,4
9	0	5	1	0,8	0,3	—	0,784	0,261	0,280	0,310	9,6
10	0	5	1	0,8	0,6	—	0,580	0,500	0,387	0,480	19,3
11	0	5	1	0,8	0,9	—	0,420	0,679	0,451	0,623	27,5

В табл. 4 приведена зависимость средней квадратической ошибки интерполяции от расстояния h_1 между аэрологическими станциями и расстояния h_3 до спутниковой орбиты.

В строках таблицы с 1-й по 5-ю приведены результаты для случаев, когда присутствуют все три вида информации, в строках с 6-й по 8-ю — результаты для аэрологической и прогностической информации, с 9-й по 11-ю — для спутниковой и прогностической информации. Из сравнения строк таблицы между собой видно, что значения ε и ε^* растут с увеличением расстояния между станциями.

Рассмотрим 1, 2, 3-ю строки табл. 4. Из сравнения величин $\frac{\varepsilon-\varepsilon^*}{\varepsilon}$ видно, что в случае, когда аэрологические станции близки ($h_1=0,2$, $h_1=0,4$), введение прогностической информации бессмысленно. При этом вес прогностической информации пренебре кимо мал, а основная доля в общей сумме весов приходится на $\sum^{4} p_i$.

Сравнив результаты в 1-й, 2-й строках с результатами в 6-й, -й строках, можно сделать вывод, что когда аэрологические станции близки к точке, куда производится интерполяция, то не только добавление прогностической информации, но и добавление спутниковой информации ничего не дает в смысле уменьшения средней квадратической ошибки интерполяции. Строки 4 и 5-я табл. 4 относятся к случаю, когда и спутниковые точки и аэрологические станции далеко. При этом вес прогностической информации p_0 значителен, с увеличением h_3 он увеличивается, и влияние прогностической информации на уменьшение ε также значительное (порядка 9%).

Данные в 8-й строке, соответствующие случаю, когда станции далеки от точки, куда производится интерполяция, а спутниковой информации нет, показывают, что при введении прогностической информации средняя квадратическая ошибка анализа уменьшается довольно существенно.

Сравнивая 5-ю и 8-ю строки табл. 4, можно отметить, что случай, когда спутниковые точки дальше, чем аэрологические станции, практически эквивалентен случаю их отсутствия.

Если имеется только спутниковая информация (строки 9—11-я), то добавление прогностической информации ведет к значительному улучшению анализа, причем это улучшение тем больше, чем дальше спутниковые точки.

До сих пор считалось, что ошибки прогностической информации обладают отрицательной корреляцией с истинными значениями метеоэлемента, и это обстоятельство правильно учитывается при расчете интерполяционных весов.

Рассмотрим два случая. 1. В интересующей нас точке 0 добавляется не прогноз, а неточное наблюдение с мерой ошибки η_0^2, ошибки которого представляют собой белый шум. Такого рода подход изложен в [6]. Обозначим среднюю квадратическую ошибку такой интерполяции ε_1. 2. В точке 0 добавлен прогноз, но мы его считаем наблюдением, т. е. при расчете интерполяционных весов корреляция ошибок прогноза с истинными значениями не учитывается. Такая интерполяция не является оптимальной, и для расчета средней квадратической ошибки интерполяции следует воспользоваться полной формулой вида

$$\varepsilon_2^2 = 1 - 2\sum_{i=0}^{n} n_i \tilde{\tilde{p}}_i + \sum_{i=0}^{n}\sum_{j=0}^{n} m_{ij} \tilde{\tilde{p}}_i \tilde{\tilde{p}}_j, \tag{17}$$

где

$$\{n_i\}_{i=0,\ 1,\ \ldots,\ n} \tag{18}$$

— вектор правых частей системы (10),

$$\{m_{ij}\}_{j=0,1\ \dots\ ,\ n}^{i=0,1\ \dots\ ,\ n} \qquad (19$$

— матрица системы (10), а

$$\{\overset{\approx}{p}_i\}_{i=0,\ 1,\ \dots\ ,\ n} \qquad (20$$

— интерполяционные веса.

При условии, что корреляционная функция ошибок прогноз совпадает с корреляционной функцией инстинных значений, про ведя выкладки, аналогичные [7], где дано выражение для ε^2 пр переходе от n-го к $n+1$-му наблюдению, легко получить следую щие формулы для $\varepsilon_1^2, \varepsilon_2^2, \varepsilon^{*2}$:

$$\varepsilon_1^2 = \frac{\varepsilon^2\,\eta_0^2}{\varepsilon^2 + \eta_0^2}; \qquad (21$$

$$\varepsilon_2^2 = \frac{\varepsilon^2\,\eta_0^2}{\varepsilon^2 + \eta_0^2}\left(1 + \frac{2\,\varepsilon^2\,\eta_0^2}{\varepsilon^2 + \eta_0^2}\right) = \varepsilon_1^2\left(1 + 2\,\varepsilon_1^2\right); \qquad (22$$

$$\varepsilon_*^2 = \frac{\varepsilon^2\,\eta_0^2}{\varepsilon^2 + \eta_0^2 - \varepsilon^2\,\eta_0^2}. \qquad (23$$

Формулы (21)—(23) в точности совпадают с соответствующи ми формулами для одноточечного согласования, приведенным в [2], и таким образом, рассуждения, проведенные для однотс

Таблица

Влияние способа учета прогностической информации на точность интерполяции

N	n_1	n_2	n_3	ε	ε^*	ε_1	ε_2	$\frac{\varepsilon-\varepsilon^*}{\varepsilon}\%$	$\frac{\varepsilon-\varepsilon_2}{\varepsilon}\%$	$\frac{\varepsilon-\varepsilon_1}{\varepsilon}\%$	$\frac{\varepsilon^*-\varepsilon_1}{\varepsilon}\%$	$\frac{\varepsilon_2-\varepsilon^*}{\varepsilon}$
1	4	5	1	0,222	0,210	0,206	0,214	5,4	3,6	7,2	1,8	1,8
2	4	3	1	0,225	0,213	0,208	0,217	5,3	3,5	7,5	2,2	1,8
3	4	1	1	0,227	0,215	0,210	0,219	5,2	3,5	7,5	2,3	1,7
4	4	0	1	0,306	0,277	0,267	0,285	9,4	6,8	12,4	3,0	2,6
5	0	5	1	0,310	0,280	0,270	0,289	9,6	6,7	12,9	3,3	2,9
6	0	3	1	0,312	0,281	0,271	0,290	9,9	7,0	13,1	3,2	2,9
7	0	1	1	0,333	0,297	0,285	0,307	10,8	7,8	14,4	3,6	3,0

чечного согласования, сохраняются и для более общего случая

В табл. 5 приведены значения средней квадратической ошибк интерполяции при различных способах учета прогностическо информации, а также соответствующие им относительные умень ения ошибки интерполяции. Именно, величина $\frac{\varepsilon - \varepsilon^*}{\varepsilon}$ представяет собой относительное уменьшение ошибки анализа при праильном учете прогноза; $\frac{\varepsilon - \varepsilon_1}{\varepsilon}$ есть уменьшение средней квадраической ошибки анализа, если добавлено наблюдение; $\frac{\varepsilon - \varepsilon_2}{\varepsilon}$ — оответствующая величина для случая, когда мы считаем прогноз аблюдением.

Величины $\frac{\varepsilon^* - \varepsilon_1}{\varepsilon}$ и $\frac{\varepsilon_2 - \varepsilon^*}{\varepsilon}$ характеризуют искажения уменьения средней квадратической ошибки анализа вследствие неравильного введения прогноза. Все результаты, приведенные табл. 5, рассчитывались при $\eta_1^2=0{,}02$; $\eta_2^2=0{,}05$; $\eta_0^2=0{,}3$; $h_1=0{,}8$; $h_2=0{,}2$; $h_3=0{,}3$.

Из сравнения между собой величин $\frac{\varepsilon - \varepsilon^*}{\varepsilon}$ видно следующее.

Во-первых, добавление прогностической информации ведет ри надлежащем выборе корреляционной функции ошибок прооза), к увеличению точности интерполяции. При наличии спутниковой и аэрологической информации уменьшение ε менее ачительно, чем когда имеются только спутниковые точки или лько аэрологические станции, расположенные далеко друг от уга.

Во-вторых, если имеется только спутниковая информация, то шь увеличение количества спутниковых точек от одной до трех ачительно влияет на точность интерполяции, а дальнейшее увеечение количества спутниковых точек дает незначительный эфкт. Если же имеется существенно влияющая аэрологическая формация, то увеличение количества спутниковых точек от одй до пяти практически не влияет на повышение точности интерляции, а поэтому при объективном анализе с использованием ально существующей сети станций северного полушария можно пользовать только одну спутниковую точку в добавление аэрологическим.

В третьих, из сравнения величин ε^*, ε_1, ε_2 видно, что всюду $< \varepsilon^*$ и $\varepsilon_2 > \varepsilon^*$. Это означает, что если в точке 0 добавлено нечное наблюдение с мерой ошибки η_0^2, то средняя квадратическая ибка интерполяции получается меньше, чем при добавлении огноза, а если добавлен прогноз, который мы считаем наблюнием, то средняя квадратическая ошибка интерполяции завыется.

Следует отметить, что отклонения ε_1 и ε_2 от ε^* примерно одиковы, причем в слчае отсуствия одного из видов информации и отклонения несколько больше.

Из проделанной работы можно сделать следующие выводы.

1. При введении в объективный анализ прогностической инфорции необходимо надлежащим образом выбирать корреляционо функцию ошибок прогноза. При неправильном выборе этой нкции введение прогноза может привести не к улучшению,

а к ухудшению анализа. В частности, улучшение анализа зав домо обеспечено, если корреляционная функция ошибок прогноз совпадает с корреляционной функцией истинных значений.

2. Добавление к аэрологическим данным спутниковых и про ностических данных имеет смысл только в районах с редко сетью.

Спутниковую информацию следует добавлять только в то случае, если она лежит ближе, чем аэрологическая, к точке, куд производится интерполяция. Если же имеется только спутникова информация, то прогностическую информацию к ней следу добавлять и при близких спутниковых точках.

3. Неправильный учет статистических характеристик испол зуемой информации ведет к значительному увеличению теоретич ской средней квадратической ошибки анализа. Поэтому при об ективном анализе с использованием спутниковой и прогностич ской информации очень важно знать и правильно учитывать с ответствующие статистические характеристики используем информации.

СПИСОК ЛИТЕРАТУРЫ

1. Тараканова В. П. К вопросу о точности косвенного температурного зо дирования атмосферы с искусственных спутников Земли.—«Метеор и гидрол.», 1974, № 4, с. 76—78.
2. Гандин Л. С., Каган Р. Л. О построении системы объективного анали разнородных данных на основе метода оптимальной интерполяции и оп мального согласования. — «Метеорол. и гидрол.», 1974, № 5, с. 3-
3. Гандин Л. С., Лугина К. М. Об учете данных косвенного зонди вания атмосферы при четырехмерном анализе метеорологических полей. См. настоящий сборник.
4. Лугина К. М., Каган Р. Л. К вопросу о пространственно-временном а лизе барического поля.—«Труды ГГО», 1974, вып. 336, с. 75—94.
5. Kruger H. B. General and special approaches to the problem of objecti analysis of meteorological variables.— „Quart. J. R. Met. Soc.", 1968, № 403.
6. Машкович С. А. О повышении качества объективного анализа барическо поля над районами с редкой сетью аэрологических станций. — «Тру ММЦ», 1956, вып. 10, с. 31—39.
7. Гандин Л. С. Объективный анализ метеорологических полей. Л., Гидро теоиздат, 1963, 287 с.

В. А. ШАХМЕЙСТЕР

НЕКОТОРЫЕ ЭКСПЕРИМЕНТЫ ПО ЧЕТЫРЕХМЕРНОМУ АНАЛИЗУ

Как известно, в настоящее время все большее развитие получают наблюдения с метеорологических спутников. Использование этих наблюдений очень важно в районах с редкой сетью аэрологических станций. Следует отметить, что если до недавнего времени спутниковая информация носила лишь полукачественный характер, то сейчас, благодаря развитию методов косвенного зондирования атмосферы, имеется возможность восстанавливать вертикальные профили температуры и геопотенциала по данным спектральных измерений уходящего излучения.

Информация, получаемая со спутников, имеет свои особенности.

Эта информация носит непрерывный, или асиноптический, характер, т. е. она не привязана к определенной точке пространства и к фиксированному моменту времени. Поэтому при ее анализе необходимо производить четырехмерное (пространственно-временное) усвоение данных наблюдений. В отличие от методов трехмерного анализа экстраполяцию метеорологических элементов по времени разумнее проводить динамическим путем, т. е. основываясь на прогностической модели. Другой особенностью спутниковой информации, которую нужно учитывать при разработке схем четырехмерного усвоения, является коррелированность ошибок спутниковых измерений.

Коррелированность ошибок спутниковых данных является следствием того, что в отличие от радиозондовых измерений, где каждое новое измерение производится другим экземпляром прибора, спутниковые измерения производятся одним и тем же прибором непрерывно по времени.

Эти свойства ошибок спутниковых измерений подтверждаются и путем непосредственного анализа данных косвенного зондирования [6]. Все изложенные ниже численные эксперименты основаны на так называемой дискретной схеме усвоения данных наблюдений. При дискретной схеме усвоение данных наблюдения происходит в сравнительно редкие, дискретно расположенные моменты времени.

Наблюдения, произведенные в промежуточный момент, экстраполируются вперед по времени на основе прогностической модели к ближайшему из моментов, для которых происходит усвоение.

В наших расчетах и данные наблюдений и «истина» генерируются при помощи одной и той же прогностической модели. Мы как бы считаем, что используемая модель атмосферы абсолютно точна, поэтому при ее интегрировании мы получаем «истинные» поля. Данные наблюдений получаются искажением этих «истинных» полей. Такой подход хорош тем, что мы можем получить и «истину» и «наблюдения» с любым разрешением по пространству и времени. В то же время нельзя гарантировать, что результаты будут аналогичны при переходе к действительным данным наблюдений. Это является большим недостатком подобных экспериментов.

Целью наших экспериментов было, во-первых, выяснить, как можно повысить точность объективного анализа путем введения данных за предшествующие сроки. Во-вторых, мы хотели узнать, каково влияние коррелированности ошибок спутниковых измерений на результаты четырехмерного усвоения. В-третьих, интересно было посмотреть, как влияет учет коррелированности ошибок в схеме четырехмерного усвоения на его результаты.

Рассмотрим модель, которая использовалась для экстраполяции метеоэлементов по времени. Баротропное уравнение вихря для двумерного потока с учетом сжимаемости имеет вид

$$\frac{\partial \nabla^2 \Psi}{\partial t} - \frac{1}{L_0^2} \frac{\partial \Psi}{\partial t} = I(\nabla^2 \Psi + l, \ \Psi), \tag{1}$$

где Ψ — функция тока, ∇ — двумерный оператор Лапласа, L_0 — характерный масштаб длин, I — оператор Якоби, l — параметр Кориолиса.

Для расчета правой части (1) использовались конечно-разностные формулы Аракавы. Интегрирование по времени велось по методу Адамса — Бешфорта.

Область интегрирования можно описать следующим образом.

На карте полярной стереографической проекции с главным масштабом по широте 60° рассматривается квадратная сетка, образующая восьмиугольник. Шаг сетки 300 км. Сетка содержит 617 узлов.

Граничные условия таковы:

ψ — абсолютная постоянная на границе,

$\nabla^2 \psi$ — сохраняется на границе области.

Модель имеет три интегральных инварианта:

$$\overline{\nabla^2 \Psi} - \frac{1}{L_0^2} \overline{\Psi} = \text{const}, \tag{2}$$

$$\frac{1}{2}\overline{(\nabla\Psi)^2}+\frac{1}{2L_0^2}\overline{\Psi^2}=\mathrm{const}, \tag{3}$$

$$\overline{(\nabla^2\Psi+l)^2}+\frac{1}{L_0^2}\overline{(\nabla\Psi)^2}-\frac{2}{L_0^2}\overline{l\Psi}=\mathrm{const}, \tag{4}$$

где

$$\overline{(\quad)}=\frac{1}{s}\iint\limits_{(s)}(\quad)\,dxdy.$$

Подробное описание схемы интегрирования можно найти в [2].

Итак, пусть мы имеем N последовательных моментов времени:

$$t_{-(N-1)},\ t_{-(N-2)},\ \ldots,\ t_{-3},\ t_{-2},\ t_{-1},\ t_0.$$

В начальный момент $t_{-(N-1)}$ поле Ψ известно. С этим полем Ψ как с начальными данными уравнение (1) интегрируется численно вперед от момента $t_{-(N-1)}$ до момента t_0. Таким образом, теперь известны решения уравнения (1) в моменты времени

$$t_{-(N-1)},\ t_{-(N-2)},\ \ldots,\ t_{-1},\ t_0.$$

Эти решения мы будем считать аналогами истинного состояния атмосферы и обозначим:

$$\Psi_{\text{ист}}^{-(N-1)},\ \Psi_{\text{ист}}^{-(N-2)},\ \ldots,\ \Psi_{\text{ист}}^{0}.$$

Затем из двумерного массива в 617 точек сетки случайным образом выбираются 80, которые считаются станциями наблюдения. На станциях $\Psi_{\text{ист}}^{-n}$ ($n=0,1,\ \ldots,\ N-1$) изменяются на случайную малую величину, при этом получается новое множество — аналог данных наблюдения, которое мы обозначим $\Psi_{\text{набл}}^{-n}$.

Было проведено несколько экспериментов с различными способами моделирования ошибок наблюдения.

В первом эксперименте ошибки наблюдения считались независимыми и моделировались как малые случайные величины, распределенные по нормальному закону, с математическим ожиданием, равным нулю.

Средняя квадратическая ошибка наблюдения геопотенциальной высоты считалась равной 2 дкм. Во втором и третьем экспериментах для получения полей $\Psi_{\text{набл}}^{-n}$ были смоделированы случайные поля коррелированных ошибок. Корреляционная функция ошибок была выбрана равной $k\rho K_1(k\rho)$, где K_1 — функция Макдональда, ρ — расстояние между точками сетки, k — коэффициент, который был выбран равным 0,5. При этом коэффициент корреляции между значениями в двух соседних точках сетки оказался равным 0,78.

Способ построения случайных полей с заданной корреляционной функцией подробно описан в работе [3]. Идея его основан на том, что для всякой абсолютно интегрируемой и ограниченной корреляционной функции $B(r)$ и для любого однородног случайного поля существует линейное преобразование, посредст вом которого заданное случайное поле преобразуется в поле с за данной корреляционной функцией $B(r)$.

В нашем случае в качестве исходного бралось случайное не коррелированное поле, а линейное преобразование над ним осуществлялось с помощью решения уравнения Гельмгольца, гд правой частью служило исходное случайное поле.

Случайное поле с корреляционной функцией $k\rho K_1(k\rho)$ моделировалось нами таким образом, чтобы среднее по всему пол было равным нулю, а дисперсия была равна четырем.

После того как наблюдаемые поля были получены, по ним выполнялся объективный анализ во все точки сетки методом оптимальной интерполяции [5].

Для первого эксперимента объективный анализ проводился обычным способом.

Для второго эксперимента в алгоритме объективного анализа учитывалось, что ошибки наблюдения коррелируют между собой и не коррелируют с истинными значениями метеоэлементов в любой точке [4]. При этих предположениях уравнения метода оптимальной интерполяции имеют вид:

$$\sum_{j=1}^{k} (\mu_{ij} + \eta \nu_{ij}) p_j = \mu_{i0} \quad (i = 1, 2, \ldots, k), \tag{4a}$$

где μ_{ij} — коэффициент корреляции между истинными значениями элемента в i-том и j-том пунктах наблюдения; μ_{i0} — коэффициент корреляции между истинным значением элемента в i-том пункте наблюдения и искомым значением этого элемента; k — число данных, используемых при интерполяции; η — мера ошибки наблюдений; ν_{ij} — коэффициент корреляции между ошибками наблюдения в i-том и j-том пунктах.

В третьем эксперименте ошибки моделировались так же, как и во втором, но при объективном анализе коррелированност ошибок не учитывалась.

Обозначим поля, полученные в результате объективного анализа $\Psi^{-n}_{\text{анал}}$. После этого с $\Psi^{-n}_{\text{анал}}$ как с начальными данным уравнение (1) интегрируется от момента времени t_n до момент времени t_0. Полученное решение обозначим $\Psi^{-n}_{\text{пред}}$. Вместо $\Psi^{-0}_{\text{пре}}$ берется $\Psi^{0}_{\text{анал}}$. Синтезированное решение вычисляется по фор муле

$$\Psi_{\text{син}}(t_0) = \sum_{n=0}^{N-1} \omega_n \Psi^{-n}_{\text{пред}}, \tag{5}$$

где $\sum_{n=0}^{N-1} \omega_n = 1.$

Средняя квадратическая ошибка F равна

$$F = E\ \overline{(\Psi_{\text{ист}} - \Psi_{\text{син}})^2}, \tag{6}$$

где $(\overline{})$ означает осреднение по площади, а E — статистическое осреднение.

Требуется определить веса ω_n $(n=0,1, \ldots, N-1)$ так, чтобы F стала минимальной. Из (5) и (6) следует, что F имеет вид

$$F = \sum_{k=0}^{N-1} \sum_{n=0}^{N-1} \omega_n \omega_k p_{nk}, \tag{7}$$

где

$$p_{nk} = E\ \overline{\left(\Psi_{\text{ист}} - \Psi_{\text{пред}}^{-n}\right)\left(\Psi_{\text{ист}} - \Psi_{\text{пред}}^{-k}\right)}.$$

Используя метод множителей Лагранжа, находим ω_n, которые минимизируют квадратичную форму F, из уравнений

$$\begin{cases} \dfrac{\partial F}{\partial \omega_k} + \lambda = 0, \\ \sum\limits_{k=0}^{N-1} \omega_k = 1, \end{cases} \tag{8}$$

где λ — множитель Лагранжа.

Система (8) представляет собой систему $N+1$ линейных уравнений с $N+1$ неизвестным. Кроме того, мы рассматривали задачу минимизации

$$F_1 = \overline{(\Psi_{\text{ист}} - \Psi_{\text{син}})^2}, \tag{9}$$

которая отличается от предыдущей отсутствием статистического осреднения.

В наших экспериментах решались обе эти задачи. В качестве начальных данных были взяты два действительных поля геопотенциала за 1 и 10 августа 1971 г., сглаженные на границе. Расчеты с этими начальными данными производились независимо. Всего было произведено три серии экспериментов: в первой из них вводились некоррелированные ошибки наблюдения, во второй ошибки наблюдения были коррелированы и это учитывалось в построении полей $\Psi_{\text{анал}}$, а в третьей вводились коррелированные ошибки наблюдения, а при объективном анализе они считались некоррелированными. В дальнейшем будем называть эксперимент с исходными данными от 1 августа и с некоррелированными ошибками — экспериментом 1, с коррелированными ошибками — экспериментом 1а, а эксперимент, в котором ошибки моделируются коррелированными, но в дальнейших расчетах это не учитывается — экспериментом 1б. Соответствующие эксперименты с исходными данными от 10 августа 1971 г назовем экспериментами 2, 2а и 2б соответственно.

Таблица 1

Статистические характеристики полей ошибок наблюдений и средние квадраты истинной ошибки объективного анализа геопотенциальной высоты в экспериментах с исходной ситуацией 1

n	$\bar{a}_{1а}$	$\sigma^2_{1а}$	$k_{1а}$	σ^2_1	$\tilde{\varepsilon}^2_1$	$\tilde{\varepsilon}^2_{1а}$	$\tilde{\varepsilon}^2_{1б}$
12	0,032	4,4	0,74	4,0	10,220	9,93	10,02
11	0,390	4,8	0,70	5,3	9,585	12,59	12,59
10	0,08	3,8	0,78	4,4	8,309	9,04	8,96
9	0,12	3,9	0,73	5,2	8,820	9,91	9,77
8	—0,33	3,8	0,67	4,5	10,813	9,82	9,77
7	0,43	4,4	0,78	3,7	9,833	10,47	10,55
6	—0,004	5,7	0,71	4,7	12,318	11,91	11,88
5	—0,09	4,9	0,71	3,9	9,719	11,86	11,65
4	0,001	4,3	0,79	4,0	10,713	13,96	13,92
3	0,14	4,3	0,66	4,1	13,334	13,84	13,97
2	0,08	3,9	0,73	3,6	11,014	11,39	11,61
1	0,20	4,2	0,68	3,8	11,256	11,28	11,42
0	—0,09	3,6	0,73	4,3	10,002	11,26	11,54
Среднее	0,07	4,3	0,72	4,2	10,458	11,30	11,36
Среднее по 6 реализациям . . .	0,04	4,2	0,71	3,95	11,006	12,26	12,36

В табл. 1 и 2 представлены величины, которые подсчитывались при получении полей $\Psi^{-n}_{\text{набл}}$ и $\Psi^{-n}_{\text{анал}}$ в экспериментах 1, 1а, 1б и 2, 2а и 2б. В этих таблицах:

n — номер поля $\Psi^{-n}_{\text{анал}}$;

$\bar{a}_1{}^a$, $\bar{a}^{2a}$ — выборочные средние полей коррелированных ошибок (по 80 точкам, которые рассматриваются в качестве пунктов наблюдения) в экспериментах 1а и 2а соответственно;

$\sigma^2_{1а}$, $\sigma^2_{2а}$ — выборочные дисперсии полей ошибок в этих экспериментах;

$k_{1а}$, $k_{2а}$ — коэффициенты корреляции между значениями ошибок в двух соседних точках сетки в экспериментах 1а и 2а;

σ^2_1, σ^2_2 — выборочные дисперсии некоррелированных ошибок в экспериментах 1 и 2;

$\tilde{\varepsilon}^2$ — средний квадрат истинной ошибки анализа, геопотенциальной высоты, равный $\left(\frac{l}{g}\right)^2\overline{(\Psi_{\text{анал}}-\Psi_{\text{ист}})^2}$. Эта величина подсчитывалась во всех экспериментах 1, 1а, 1б, 2, 2а, 2б и обозначалась через $\tilde{\varepsilon}_1^2$, $\tilde{\varepsilon}_{1а}^2$, $\tilde{\varepsilon}_{1б}^2$, $\tilde{\varepsilon}_2^2$, $\tilde{\varepsilon}_{2а}^2$, $\tilde{\varepsilon}_{2б}^2$ соответственно.

Обозначим $E\left(\frac{l}{g}\right)^2 \overline{(\Psi_{анал} - \Psi_{ист})^2}$ через ε^2. Соответствующие величины для всех экспериментов обозначим: ε_1^2, $\varepsilon_{1а}^2$, $\varepsilon_{1б}^2$, ε_2^2, $\varepsilon_{2а}^2$, $\varepsilon_{2б}^2$.

Таблица 2

Статистические характеристики полей ошибок наблюдений и средние квадраты истинной ошибки объективного анализа геопотенциальной высоты в экспериментах с исходной ситуацией 2

n	$\bar{a}_{2а}$	$\sigma_{2а}^2$	$k_{2а}$	σ_2^2	$\tilde{\varepsilon}_2^2$	$\tilde{\varepsilon}_{2а}^2$	$\tilde{\varepsilon}_{2б}^2$
12	−0,05	4,5	0,71	3,9	8,73	12,50	12,50
11	−0,21	4,5	0,78	4,3	10,92	9,34	9,44
10	0,15	3,8	0,74	4,0	9,7	8,96	8,85
9	0,20	4,1	0,65	4,2	7,4	7,97	7,87
8	0,23	4,6	0,76	3,9	8,14	9,28	9,23
7	0,19	4,0	0,70	3,8	8,14	10,35	10,41
6	0,23	4,1	0,83	3,9	8,1	10,44	10,79
5	0,03	4,8	0,69	4,0	9,82	8,51	8,67
4	0,04	4,6	0,69	3,7	10,0	11,70	11,84
3	−0,42	3,4	0,74	3,6	11,21	10,37	10,36
2	−0,02	4,9	0,68	4,4	7,51	9,45	9,49
1	0,37	4,3	0,64	4,9	7,82	8,66	8,57
0	−0,14	3,8	0,77	3,7	6,7	8,42	8,45
Среднее	0,05	4,3	0,72	4,0	8,7	9,70	9,73
Среднее по 6 реализациям . . .	−0,02	4,3	0,70	4,0	8,8	9,51	9,56

Из сравнения этих величин для всех экспериментов видно, что выполняются следующие неравенства:

1) $\varepsilon_1^2 < \varepsilon_{1а}^2 < \varepsilon_{1б}^2$, $\varepsilon_2^2 < \varepsilon_{2а}^2 < \varepsilon_{2б}^2$,

2) $\varepsilon_2^2 < \varepsilon_1^2$, $\varepsilon_{2а}^2 < \varepsilon_{1а}^2$, $\varepsilon_{2б}^2 < \varepsilon_{1б}^2$.

Первая группа неравенств свидетельствуют о том, что, как следовало ожидать, средний квадрат ошибки объективного анализа при независимых ошибках наблюдения меньше, чем для анализа с зависимыми ошибками наблюдения, а анализ с зависимыми ошибками наблюдения, которые считаются независимыми, дает еще бóльшую истинную ошибку.

Вторая группа неравенств показывает, что средний квадрат ошибки анализа во всех экспериментах 2 меньше, чем в экспериментах 1.

Во всех экспериментах при интегрировании уравнения (1) для расчета полей $\Psi_{\text{ист}}$ и $\Psi_{\text{пред}}$ рассчитывались интегральные инварианты модели. Вследствие наличия временной конечно-разностной аппроксимации эти инварианты меняются со временем.

Из сравнения хода изменения инвариантов I_1, I_2, I_3 в экспериментах 1 и 2 следует, что в эксперименте 1 инварианты меняются значительно быстрее, чем в эксперименте 2. Из сравнения хода изменения инвариантов в экспериментах 1а и 2а и 1б и 2б также вытекает, что в экспериментах 1а и 1б инварианты изменяются быстрее, чем в экспериментах 2а и 2б соответственно.

Расчет матриц $\{p_{nk}\}_{n=0,\,k=0}^{N\quad N}$ производился следующим образом: сначала для каждого момента времени t_0, t_{-1}, ..., $t_{-(N-2)}$ подсчитывались матрицы $\left\{\tilde{p}_{nk}\right\}$, где

$$\tilde{p}_{nk} = \overline{\left(\Psi_{\text{ист}} - \Psi_{\text{пред}}^{-n}\right)\left(\Psi_{\text{ист}} - \Psi_{\text{пред}}^{-k}\right)}, \qquad (10)$$

и решалась задача (9), а затем проводилось статистическое осреднение по всем возможным реализациям (внутри одного эксперимента). Интегрирование уравнения (1) проводилось на трое суток вперед с шагом $\Delta t = 6$ ч. Для получения матриц $\{p_{nk}\}$ мы ограничились реализациями $\left\{\tilde{p}_{nk}\right\}$ для моментов t_0, t_{-1}, t_{-2}, t_{-3}, ..., t_{-6}. При этом полученные матрицы $\{p_{nk}\}$ имели восьмой порядок. Это означает, что при построении синтезированного решения $\Psi_{\text{син}}$ по формуле (5), при решении задачи (7) мы не привлекали данных «старше», чем за два дня. Это объясняется тем, что в рамках наших экспериментов мы были лишены возможности построения матриц $\{p_{nk}\}$ более высокого порядка, так как число реализаций матриц $\{p_{nk}\}$ для $N>7$ слишком мало. Матрицу $\{p_{nk}\}$, полученную в эксперименте 1, обозначим $P^{(1)}$, а матрицу $\{p_{nk}\}$, полученную в эксперименте $2 — P^{(2)}$. Кроме того, обозначим через $P^{(1+2)}$ матрицу с компонентами

$$p_{nk}^{(1+2)} = \frac{p_{nk}^{(1)} + p_{nk}^{(2)}}{2}. \qquad (11)$$

Таким образом $P^{(1+2)}$ есть средняя матрица $\{p_{nk}\}$ для экспериментов 1 и 2. Аналогично строились и матрицы $P^{(1а)}$, $P^{(2а)}$, $P^{(1а+2а)}$ в экспериментах 1а и 2а и матрицы $P^{(1б)}$, $P^{(2б)}$, $P^{(1б+2б)}$ в экспериментах 1б и 2б соответственно.

На рис. 1 представлены двумерные распределения компонент матриц $P^{(1)}$, $P^{(2)}$ и $P^{(1а)}$, $P^{(2а)}$. Заметим, что эти матрицы представляют собой матрицы ковариаций ошибок прогноза различной заблаговременности и по их двумерному распределению можно судить о нарушении предсказуемости в этих экспериментах. Из рисунка видно следующее.

Во-первых, поля $P^{(1)}$ и $P^{(1а)}$, а также $P^{(2)}$ и $P^{(2а)}$ очень похожи между собой. Поэтому мы не можем ожидать больших различий между результатами экспериментов с коррелированными и некоррелированными ошибками, что и будет показано дальше.

Во-вторых, поля $P^{(2)}$ и $P^{(2а)}$ значительно более гладкие, чем поля $P^{(1)}$ и $P^{(1а)}$ соответственно.

Этот факт хорошо согласуется с поведением интегральных инвариантов в экспериментах 1 и 2, 1а и 2а. Заметим, что хотя дву-

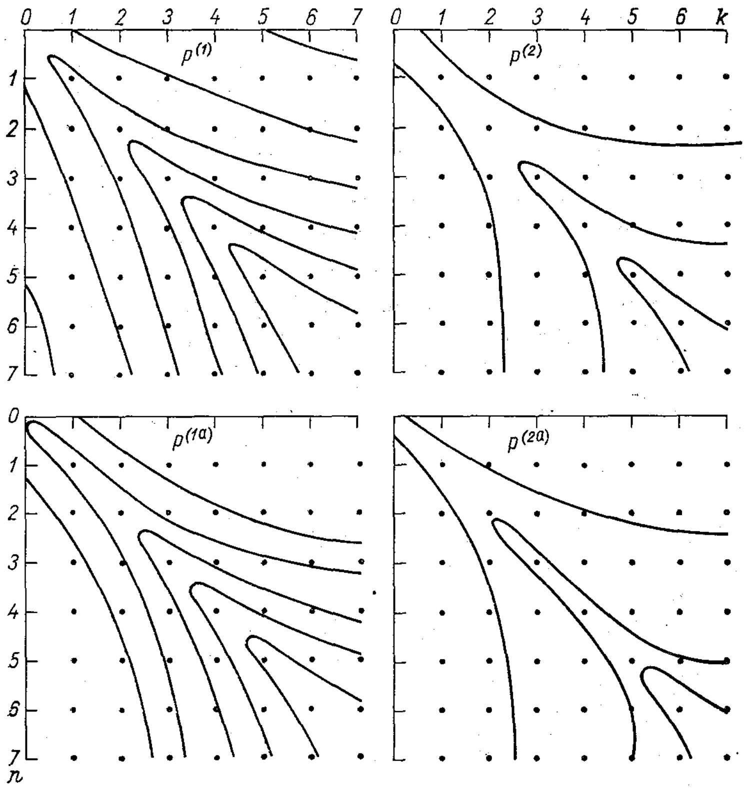


Рис. 1. Двумерные распределения матриц $P^{(1)}$, $P^{(2)}$, $P^{(1а)}$, $P^{(2а)}$.

мерные распределения для $P^{(1б)}$ и $P^{(2б)}$ здесь не приводятся, о них можно повторить все вышесказанное. Таким образом, мы видим, что в экспериментах 1, 1а, 1б предсказуемость нарушается быстрее, чем в экспериментах 2, 2а и 2б.

В табл. 3 представлены примеры весов ω_n в экспериментах 1, 1а, 1б, 2, 2а, и 2б для интервалов $\Delta t=6$ ч, соответствующие «возрасту» 0; 0,25; 0,5; ...; 1,75 дня. Из таблицы видно, что веса ω_n с увеличением n убывают. В последнем столбце таблицы помещены суммы первых пяти весов для всех экспериментов. Эта величина во всех экспериментах приближенно равна 0,9. Так как сумма всех весов равна единице, то можно ожидать, что основное уменьшение первоначальной ошибки анализа должно во всех экспериментах происходить за первый день.

Из сравнения сумм весов видно также, что сумма пяти весов для первых трех экспериментов меньше, чем для вторых трех. Это говорит о том, что в экспериментах серии 2 большую роль в построении $\Psi_{\text{син}}$ играют поля, возрастом не «старше» одного

Таблица 3

Примеры весов ω_n для всех проводимых экспериментов

№ эксперимента	Δt дни								$\sum\limits_{n=0}^{4} \omega_n$
	0	0,25	0,50	0,75	1,0	1,25	1,50	1,75	
1	0,535	0,136	0,117	0,008	0,062	0,007	0,049	0,085	0,858
1а	0,430	0,168	0,129	0,048	0,121	0,065	0,003	0,035	0,896
1б	0,477	0,118	0,142	0,042	0,124	0,055	0,002	0,040	0,903
2	0,424	0,232	0,166	0,094	0,052	0,010	0,021	0,001	0,969
2а	0,406	0,237	0,154	0,043	0,087	0,055	0,011	0,006	0,927
2б	0,420	0,240	0,152	0,035	0,083	0,058	0,013	0,002	0,930

дня, что связано с тем, что в экспериментах 2 предсказуемость нарушается медленнее и поэтому более «близкие» поля $\Psi_{\text{пред}}$ несут в себе большую информацию, необходимую для построения $\Psi_{\text{син}}$. Все эксперименты проводились в следующем порядке. Сначала решалась задача (9) — задача минимизации F_1 — среднего по площади квадратического отклонения синтезированного решения от истинного. При этом были рассчитаны матрицы $\{\tilde{p}_{nk}\}$. При проведении всех экспериментов рассчитывалась теоретическая средняя квадратическая ошибка четырехмерного анализа геопотенциальной высоты $\varepsilon_{\text{т}}^2$ и эмпирическая ε_u^2. Очевидно, что в случае задачи (9) $\varepsilon_{\text{т}}=\varepsilon_u$. Затем рассчитывались матрицы $\{p_{nk}\}$ и решалась задача (7) для каждого эксперимента со своей матрицей $\{p_{nk}\}$. Нетрудно видеть, что при этом

$$\sqrt{E(\varepsilon_u^2)} = \varepsilon_{\text{т}}. \qquad (12)$$

В дальнейшем вместо $\sqrt{E(\varepsilon_u^2)}$ будем писать $\tilde{E}$ (ε_u). После этого брались матрицы $P^{(1)}$, $(P^{(1а)}, P^{(1б)})$ и применялись к $\Psi_{\text{пред}}$, полученным из эксперимента 2 (2а, 2б) и наоборот. Затем были рас-

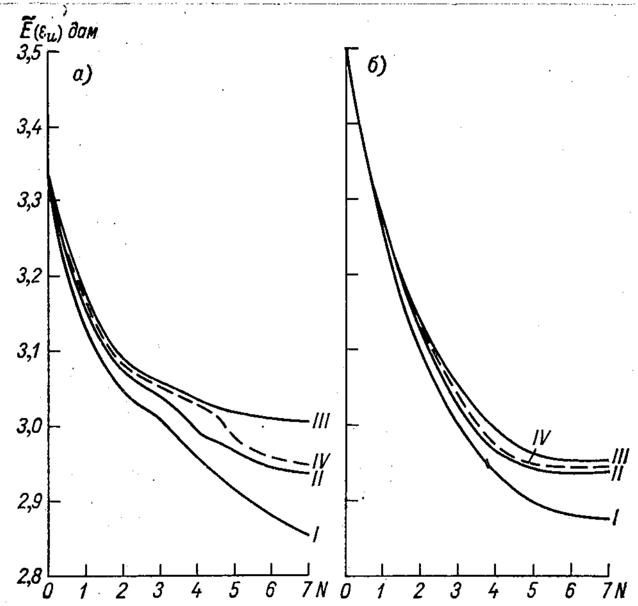


Рис. 2. $\tilde{E}(\varepsilon_u)$ в зависимости от N в экспериментах 1, 1а.

а — эксперимент 1, *б* — эксперимент 1а; *I* — $\tilde{E}(\varepsilon_u)$ при решении задачи (9), *II* — $\tilde{E}(\varepsilon_u)$ при использовании «своих» матриц $\{p_{nk}\}$, *III* — $\tilde{E}(\varepsilon_u)$ при использовании «чужих» матриц $\{p_{nk}\}$, *IV* — при использовании матриц $\{p_{nk}\}$, средних по двум экспериментам.

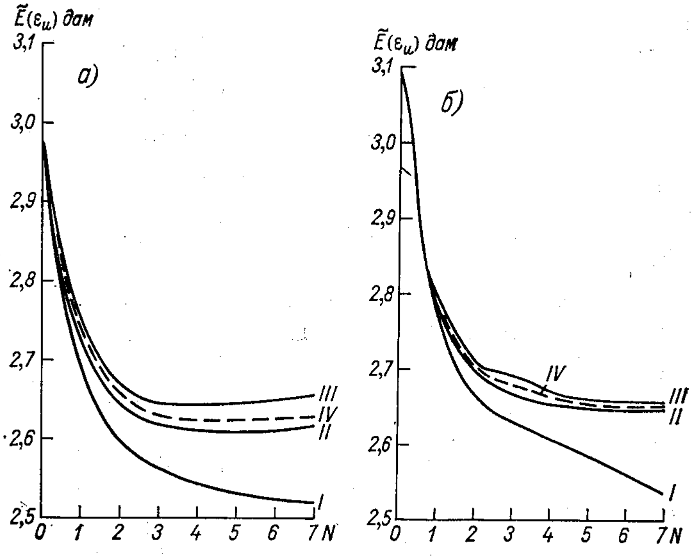


Рис. 3. $\tilde{E}(\varepsilon_u)$ в зависимости от N в экспериментах 2, 2а.
а — эксперимент 2, *б* — эксперимент 2а; усл. обозначения *I*—*IV* см. рис. 2.

считаны матрицы $P^{(1+2)}$, ($P^{(1а+2а)}$, $P^{(1б+2б)}$) и с ними решалась задача (7) для всех экспериментов. Результаты этих расчетов приведены на рис. 2, 3, 4, 5.

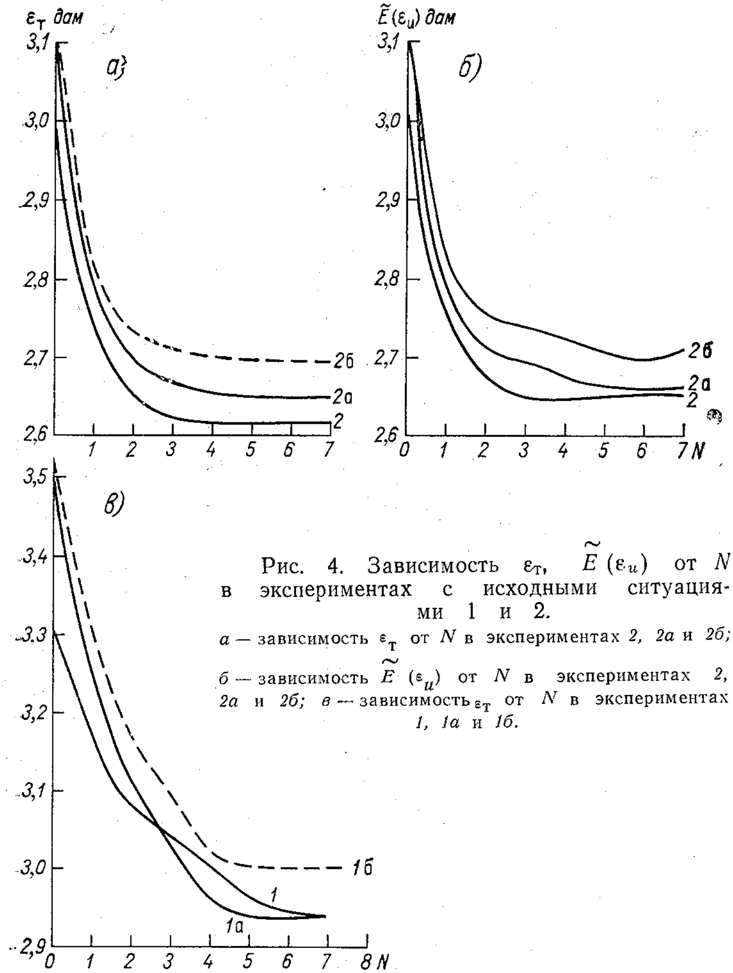


Рис. 4. Зависимость $\varepsilon_т$, $\tilde{E}(\varepsilon_u)$ от N в экспериментах с исходными ситуациями 1 и 2.

а — зависимость $\varepsilon_т$ от N в экспериментах *2, 2а* и *2б*; *б* — зависимость $\tilde{E}(\varepsilon_u)$ от N в экспериментах *2, 2а* и *2б*; *в* — зависимость $\varepsilon_т$ от N в экспериментах *1, 1а* и *1б*.

На рис. 2 и 3 изображены кривые изменения $\tilde{E}(\varepsilon_u)$ в зависимости от N — числа $\Psi^{-n}_{пред}$, участвующих в построении $\Psi_{син}$ в экспериментах 1, 1а, 2, 2а. Из сравнения этих рисунков видно, что взаимное расположение и вид кривых *I, II, III, IV* аналогичны во всех экспериментах, хотя для каждого эксперимента имеются свои особенности.

Из рис. 3, например, видно, что с увеличением N величина $\tilde{E}$ (ε_u) убывает, причем наилучший результат, как и следовало ожидать, дает решение задачи (9) кривая *I*).

При решении задачи (7) с использованием «своих» матриц $\{p_{nk}\}$ $\tilde{E}(\varepsilon_u)$ возрастает (кривая *II* лежит выше кривой *I*), затем при замене «своих» матриц «чужими» значительно возрастает (кривая *III*), но почти возвращается к прежнему уровню (кривая *II*) при использовании средних матриц $\{p_{nk}\}$ для двух экспериментов. Заметим, что характер и взаимное расположение этих кривых в экспериментах 2а и 2б аналогичны.

Из рис. 4 *а* видно, что для экспериментов 2, 2а и 2б начальное неравенство $\varepsilon_2 < \varepsilon_{2а} < \varepsilon_{2б}$ сохраняется и для теоретических ошибок четырехмерного анализа. На рис. 4 *б* изображены кривые изменения $\tilde{E}$ (ε_u) в экспериментах 2, 2а, 2б при использовании «чужих» матриц $\{p_{nk}\}$. Взаимное расположение кривых *2, 2а* и *2б.* одинаково на рис. 4 *а* и *б*.

Из графиков, приведенных на рис. 4 *а,* видно, что различия в окончательном уровне уменьшения ошибки между кривыми *2а* и *2б* больше, чем различия между первоначальными ошибками объективного анализа в экспериментах 2а и 2б. Таким образом, отсутствие учета коррелированности ошибок наблюдения в схеме четырехмерного анализа ухудшает его результаты.

Из рис. 4 *в* видно, что кривая *1 а* пересекает кривую *1,* и, таким образом, для ошибок четырехмерного анализа в экспериментах 1 и 1а первоначальное соотношение $\varepsilon_1 < \varepsilon_{1а}$ не сохраняется.

Из сравнения рис. 4 *а* и *в* видно, что как первоначальная величина ошибки объективного анализа, так и окончательный уровень ее уменьшения в экспериментах 2, 2а и 2б значительно ниже, чем в экспериментах 1, 1а и 1б. Кроме того, кривые на рис. 4 *а* раньше начинают оставаться на одном уровне, чем на рис. 4 *в*. Все это связано с тем, что, как уже отмечалось выше, в экспериментах 2, 2а и 2б предсказуемость нарушается медленнее, чем в экспериментах 1, 1а и 1б, а сумма первых пяти весов ω_n больше, чем в экспериментах 1, 1а и 1б.

Сравним теперь $\tilde{E}$ (ε_u) и $\varepsilon_т$ для различных экспериментов.

На рис. 5 *а* и *б* теоретическая $\varepsilon_т$ и истинные $\tilde{E}$ (ε_u) кривые ошибок анализа в экспериментах 1а и 2а лежат довольно далеко друг от друга (кривые *I* и *II*), но они значительно приближаются друг к другу при использовании матриц $P^{(1а+2а)}$ (кривые *III* и *IV*). Таким образом, увеличение числа реализаций для статистического осреднения при расчете матриц $\{p_{nk}\}$ ведет к сближению истинной и теоретической ошибок анализа.

Кроме того, из анализа всех теоретических кривых следует, что окончательный уровень уменьшения ошибки равен получен-ному после 1,5 дня. Таким образом, в связи с нарушением предсказуемости данные «старше» чем за 1,5 дня можно не привлекать для построения синтезированного решения $\Psi_{\text{син}}$.

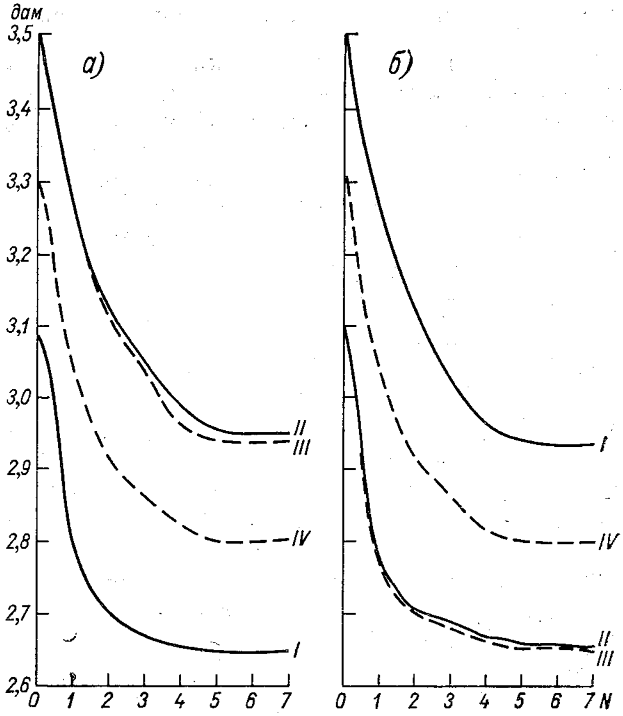


Рис. 5. Зависимость $\varepsilon_{\text{т}}$ и $\widetilde{E}$ (ε_u) от N в экспериментах 1а и 2а.

а — эксперимент 1а с коррелированными ошибками: *I* — $\varepsilon_{\text{т}}$ для матриц $P^{(2a)}$, *II* — $\widetilde{E}$ (ε_u) для матриц $P^{(2a)}$, *III* — $\widetilde{E}(\varepsilon_u)$ для матриц $P^{(1a+2a)}$, *IV* — $\varepsilon_{\text{т}}$ для матриц $P^{(1a+2a)}$; *б*—эксперимент 2а с коррелированными ошибками: *I* — $\varepsilon_{\text{т}}$ для матриц $P^{(1a)}$, *II* — $\widetilde{E}$ (ε_u) для матриц $P^{(1a)}$, *III* — $\widetilde{E}$ (ε_u) для матриц $P^{(1a+2a)}$; *IV* — $\varepsilon_{\text{т}}$ для матриц $P^{(1a+2a)}$.

Из табл. 4, в которой представлено процентное отношение окончательного уменьшения теоретической ошибки четырехмерного анализа ε_0—ε_N к первоначальной ошибке ε_0 объективного анализа для всех экспериментов, можно сделать следующие выводы:

Во-первых, при наличии коррелированных ошибок наблюдения роль четырехмерного усвоения повышается в том смысле, что отношение $\frac{\varepsilon_0 - \varepsilon_N}{\varepsilon_0}$ больше для коррелированных ошибок, чем для некоррелированных. В то же время сама величина ε_N для некоррелированных ошибок меньше, чем для коррелированных.

Во-вторых, учет коррелированности ошибок в схеме анализа ведет к некоторому улучшению в окончательном уровне уменьшения ошибки анализа. Например, ошибка уменьшается на 5,9% в эксперименте 1а и на 14,5% в эксперименте 1б.

Таблица 4

Относительное уменьшение ошибки в результате четырехмерного анализа

Ошибки	Исходная ситуация	
	1	2
Некоррелированные	11,4	12,1
Коррелированные	15,9	14,1
Коррелированные, но считаются некоррелированными	14,5	12,8

Заметим, что аналогичные эксперименты, но только для некоррелированных ошибок наблюдения проводились Миякода и Талаграном [1].

Результаты, полученные нами в экспериментах с некоррелированными ошибками, очень похожи на результаты [1]. Наше рассмотрение несколько шире, чем в [1], там не производился анализ различных путей построения матриц $\{p_{nk}\}$, сравнение истинной и теоретической ошибок анализа и не изучалось влияние коррелированности ошибок спутниковых измерений на результаты четырехмерного усвоения.

СПИСОК ЛИТЕРАТУРЫ

Mijakoda K. and Talagrand O. The assimilation of past data in dynamical analysis, Part I. Tellus, 1971, vol. 22, N 4—5, p. 310—317.

Гандин Л. С., Каган Р. Л., Мелешко В. Н., Руховец Л. В. Упрощенная модель для численных экспериментов по термическому режиму атмосферы. — «Труды ГГО», 1970, вып. 256, с. 72—97.

Глуховский А. Б. О статистическом моделировании метеорологических полей. — «Изв. АН СССР, Физика атмосферы и океана.» 1969, № 7, т. 5. с. 724—729.

Гандин Л. С., Каган Р. Л., Полищук А. И. Об оценке информативности систем метеорологических наблюдений. — «Труды ГГО», 1972, вып. 286, с. 120—140.

Гандин Л. С. Объективный анализ метеорологических полей. Л., Гидрометеоиздат. 1966. 268 с.

Тараканова В. П. К вопросу о точности косвенного температурного зондирования атмосферы с искусственных спутников Земли. — «Метеорол. и гидрол.», 1974. № 4, с. 76—78.

В. П. ТАРАКАНОВ

О ЗАКОНОМЕРНОСТЯХ ОШИБОК ОПРЕДЕЛЕНИЯ ТЕМПЕРАТУРЫ И ГЕОПОТЕНЦИАЛА ПО ДАННЫМ КОСВЕННОГО ЗОНДИРОВАНИЯ АТМОСФЕРЫ

В оперативной практике анализа и прогноза метеорологич ских полей используются не только данные радиозондов, но и да ные температурного спутникового зондирования атмосферы. П следние играют большую роль при анализе в районах с редк сетью аэрологических станций. Поскольку наблюдения со сп ников непрерывны во времени, то при анализе необходимо пр изводить пространственно-временное усвоение этих данных. В р ботах [5, 6] представлены различные схемы четырехмерно объективного анализа асинхронных метеорологических данн Однако в этих схемах не учитываются их особенности, основн из которых является коррелированность ошибок спутников данных, снижающая их информативность [2]. Главной причин этого является отсутствие данных о структуре ошибок спутни вого зондирования.

В работе автора [7] была сделана попытка количествен оценки структуры ошибок данных косвенного температурн зондирования атмосферы. Полученные при этом выводы носи предварительный характер, поскольку расчеты были сделаны небольшому объему исходных данных. В настоящей статье р сматривается пространственная структура поля ошибок спут ковых данных о температуре и геопотенциале изобарических верхностей, полученная на бóльшем материале. Использовали результаты косвенного зондирования атмосферы с американск метеорологических спутников серии «Нимбус» за период с дек ря 1971 г. по май 1972 г. Для исследования взят район, хоро освещенный аэрологическими данными, расположенный меж 10° з. д. и 180° в. д. и 40 и 80° с. ш. Данные были взяты из те грамм, ежедневно поступающих в Гидрометцентр СССР и люб ю предоставленных нам Я. М. Хейфецем и В. А. Пастуховым. Расчеты проводились для двух сезонов — зима (декабрь — февраль) и весна (март — май). Для зимы было отобрано 405, для весны — 363 вертикальных профиля температуры и геопотенциала. Методика обработки данных изложена в работе [7].

В табл. 1 приведены значения средних квадратических ошибок σ_δ спутникового зондирования. Как видно из таблицы, в ходе средних квадратических ошибок наблюдаются сезонные различия: весной их значения больше, чем зимой. Значения σ_{δ_T} для весны оказались гораздо меньшими (до уровня 300 мбар), чем полученные в работе [7]. Это, вероятно, можно объяснить тем, что в [7] для расчета использовались данные зондирований не

Таблица 1

Средние квадратические ошибки спутникового зондирования на различных изобарических поверхностях

Ошибка	Уровень, мбар				
	850	700	500	300	200
			Зима		
σ_{δ_T} °C	1,8	1,8	2,3	1,9	2,6
σ_{δ_H} дам	1,7	1,7	2,8	4,3	5,1
			Весна		
σ_{δ_T} °C	2,3	2,7	2,5	2,3	3,0
σ_{δ_H} дам	2,9	3,6	5,3	7,4	8,5

олько за весенний, но и за летний период. Отсюда можно сделать вывод, что значения σ_{δ_T} летом могут быть еще больше, чем весной.

Значения средних квадратических ошибок температуры в общем увеличиваются с высотой. Небольшой минимум, который прослеживается на уровне 300 мбар (а весной также и на уровне 500 мбар), возможно, связан с недостаточным объемом данных. Следует, однако, иметь в виду, что такой минимум в умеренных широтах прослеживается и в средних квадратических отклонениях самой температуры, приведенных в работе И. В. Ханевской [8]. Сравнение данных [8] с данными табл. 1 показывает, что величина σ_{δ_T} для зимнего периода достигает 25—30% от средней квадратической изменчивости температуры на нижних и 40% на верхних уровнях. Для весеннего периода относительная погрешность спутниковых данных о температуре $\sigma_{\delta_T}/\sigma_T$ оказывается еще больше.

Зависимость погрешности спутниковых данных о геопотенциале от сезона и от уровня в атмосфере более четко выражена, чем для соответствующих данных о температуре. Для весеннего пе риода значения σ_{δ_H} в 1,5—2 раза больше, чем для зимнего пе риода. Особенно резко прослеживается увеличение σ_{δ_H} с высо той. Так, на изобарической поверхности 200 мбар σ_{δ_H} в 3 раз больше, чем на поверхности 850 мбар. Такой сильный рост оши бок обусловлен накоплением ошибок в определении относитель ного геопотенциала различных слоев, связанных с погрешностя ми определения температуры.

Сравнение величин σ_{δ_H} с величинами σ_H, приведенным в работе Л. Г. Заставенко [3], показывает, что относительна погрешность определения геопотенциала $\sigma_{\delta_H}/\sigma_H$ составляет дл зимнего периода около 15% для нижних уровней и 25—30% дл верхних уровней тропосферы. Для весны эти значения оказыва ются гораздо бо́льшими.

При рассмотрении данных табл. 1 следует иметь в виду, чт приведенные в ней значения σ_δ завышены по сравнению с ис тинными погрешностями спутниковых наблюдений за счет оши бок интерполяции соответствующих полей по данным обычны аэрологических наблюдений, с которыми сравниваются данны спутниковых наблюдений. Выделить влияние ошибок интерполя ции оказывается затруднительным, однако некоторое представ ление об этом можно получить, если учесть что для данного райо на средняя квадратическая ошибка интерполяции составляе около 10% от среднего квадратического отклонения для поля гео потенциала и около 15—20% от среднего квадратического откло нения для температуры [1, 4]. В связи с этим можно ориенти ровочно считать, что приведенные в табл. 1 значения средни квадратических ошибок завышены для нижних уровней атмосфе ры примерно на 20%. Для более высоких уровней влияние оши бок интерполяции уменьшается и завышение для поверхност 200 мб составляет около 10%. Указанные оценки относятс к зимнему периоду, для весны ошибки интерполяции нескольк меньше, чем для зимы. Это связано с тем, что весной меньш естественная изменчивость полей геопотенциала и температуры Учитывая, что значения σ_δ для этого сезона больше, чем для зи мы, можно ожидать, что вклад ошибок интерполяции для этог сезона существенно меньше.

Сказанное следует иметь в виду и при анализе данных о про странственной корреляции ошибок спутниковых данных, приве денных на рис. 1 и 2. Как видно из рисунков, корреляци ошибок спутниковых данных прослеживается до значительны расстояний. Учитывая при этом имеющийся объем данных, раз личия между пространственной корреляцией ошибок температур и геопотенциала, по-видимому, можно считать несущественными

При сопоставлении корреляционных функций для различны сезонов обращает на себя внимание существенно более высока корреляция для весеннего сезона (рис. 1 *б*, 2 *б*), для которог коэффициенты корреляции для первой градации расстояни среднее расстояние около 250 км) достигают 0,7—0,8, а для имнего периода они, как правило, не превышают 0,5. Возможно, то эти различия связаны не столько с различиями в корреляции стинных ошибок спутниковых данных, сколько в существенном

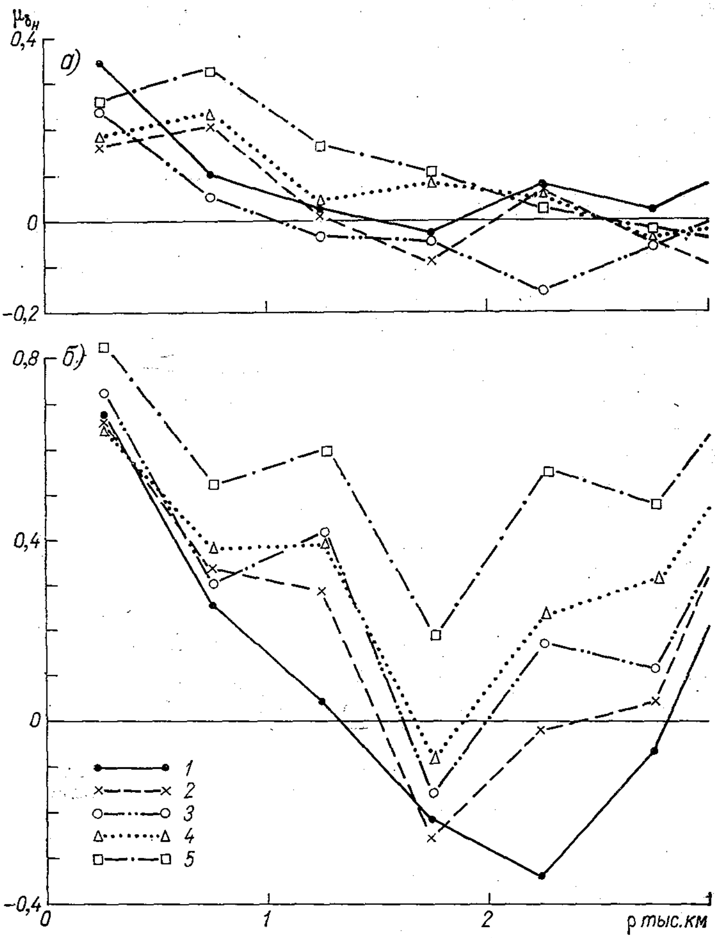


Рис. 1. Пространственная корреляция ошибок спутникового зондирования. Поле геопотенциала.

а — зима (XII—II), *б* — весна (III—V); *1* — 850 мбар, *2* — 700 мбар, *3* — 500 мбар, *4* — 300 мбар, *5* — 200 мбар.

лиянии упомянутых выше ошибок интерполяции для зимнего езона.

Более высокая пространственная корреляция ошибок для уров- 200 мбар (в весенний период) также, вероятно, определяется лиянием этого фактора. Во всяком случае значения пространственной корреляции ошибок 0,7—0,8 для расстояний 300—500 км между пунктами наблюдений представляются вполне реальными.

При рассмотрении рис. 1 и 2 можно видеть, что пространственная корреляция достигает минимума на расстояниях около 2000 км, далее наблюдается резкий рост корреляционной функ-

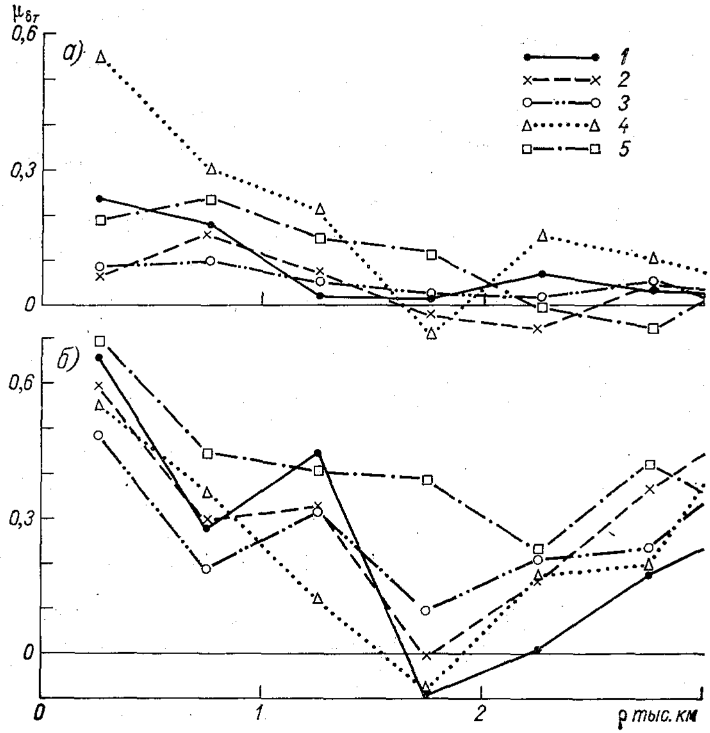


Рис. 2. Пространственная корреляция ошибок спутникового зондирования. Поле температуры.

Усл. обозначения см. рис. 1.

ции, достигающей на расстояниях порядка 3500 км значений сравнимых со значениями, полученными для малых расстояний. Вероятно, такой ход корреляционной функции связан с характерными размерами барических образований в умеренных и высоких широтах и с зависимостью ошибок спутниковых данных от синоптической ситуации. Этот вопрос заслуживает дополнительного исследования, поскольку наличие такой зависимости открывало бы дополнительные возможности для уточнения данных спутникового зондирования.

Из данных табл. 2 следует, что межуровенная корреляция ибок температуры прослеживается лишь в нижней тропосфере, я ошибок спутниковых данных о геопотенциале отмечается су-ственная положительная корреляция во всей толще тропосфе-. Наиболее значительна корреляция для смежных уровней,

Таблица 2

журовенная корреляция ошибок данных спутникового зондирования

овень. мбар	μ_{δ_T}				μ_{δ_H}			
	700	500	300	200	700	500	300	200
				Зима				
850	0,24	0,22	0,02	—0,03	0,38	0,33	0,09	0,06
700		0,46	0,07	—0,05		0,48	0,30	0,19
500			0,09	—0,16			0,59	0,46
300				0,05				0,57
				Весна				
850	0,20	0,20	0,02	0,003	0,45	0,28	0,18	0,14
700		0,50	—0,05	—0,02		0,52	0,32	0,29
500			—0,06	—0,18			0,57	0,43
300				0,02				0,56

я которых коэффициент корреляции составляет около 0,6, при-м, учитывая влияние ошибок интерполяции эти величины сле-ует считать еще заниженными.

СПИСОК ЛИТЕРАТУРЫ

Белоусов С. Л., Гандин Л. С., Машкович С. А. Обработка оперативной метеорологической информации с помощью электронных вычислительных машин. Л., Гидрометеоиздат, 1968. 282 с.

Гандин Л. С. Проблемы усвоения данных метеорологических наблюдений. — «Метеорол. и гидрол». 1971, № 3, с. 15—21.

Заставенко Л. Г. Барическое поле тропосферы северного полушария. — «Труды НИИАК», 1972, вып. 86, 237 с.

Либерман Ю. М., Лугина К. М. Опыт объективного анализа среднемесячных аэрологических полей.—«Труды ГГО», 1970, вып. 267, с. 63—78.

5. Машкович С. А., Вейль И. Г. Численные эксперименты по четырехме ному объективному анализу на основе спектральной прогностической мод ли. — «Метеорол. и гидрол.», 1972, № 3, с. 3—15.
6. Мусаелян Ш. А., Хорошилов А. Ф. О пространственно-временном ан лизе асинхронных метеорологических данных.—«Метеорол. и гидрол 1973, № 2, с. 36—46.
7. Тараканова В. П. К вопросу о точности косвенного температурного зо дирования атмосферы с искусственных спутников Земли.—«Метеоро и гидрол.», 1974, № 4, с. 76—78.
8. Ханевская И. В. Температурный режим свободной атмосферы над севе ным полушарием. Л., Гидрометеоиздат, 1968. 300 с.

Ю. М. ЛИБЕРМАН

РАСПРЕДЕЛЕНИЕ АБСОЛЮТНЫХ ОШИБОК АНАЛИЗА ГЕОПОТЕНЦИАЛА В ЮЖНОМ ПОЛУШАРИИ

Реализация планов Всемирной службы погоды, Проекта исследований глобальных атмосферных процессов и других международных программ требует интенсивного изучения атмосферы южного полушария. В связи с этим в последнее десятилетие резко возросла потребность в разработке специальной методики численного анализа метеорологических полей в этой части Земного шара. Два обстоятельства сильно затрудняют решение этой задачи. Во-первых, сеть станций в южном полушарии значительно реже, чем в северном. В частности, аэрологическая сеть насчитывает здесь лишь немногим более 100 станций, тогда как в северном полушарии — свыше 700. Во-вторых, преобладание облачности в зоне умеренных широт южного полушария препятствует получению спутниковой метеорологической информации в должном объеме.

Естественно в первую очередь рассмотреть вопрос о фактической точности анализа, которую обеспечивает существующая сеть станций. Применительно к геопотенциалу эта задача исследовалась в работе [4] на базе оптимальной интерполяции [1]. В качестве оценки точности анализа использовалась мера ошибки оптимальной интерполяции

$$\varepsilon = \frac{E}{m_H(0)}, \qquad (1)$$

представляющая средний в статистическом смысле квадрат абсолютной ошибки анализа E, отнесенный к дисперсии геопотенциала $m_H(0)$. Нормирование на дисперсию обусловливает зависимость характеристики ε лишь от густоты наблюдательной сети, то есть от числа и взаимного расположения станций, данные которых интерполируются в узел регулярной сетки. Практичес для вычисления ε использовалась формула

$$\varepsilon = 1 - \sum_{i=1}^{n} p_i \mu_{0i},$$

где n — число станций, μ_{0i} — автокорреляционные коэффициен отвечающие паре «узел — станция», p_i — оптимальные интерпо ционные веса.

Для целей численных прогнозов существенно знание не то ко поля самого метеорологического элемента, но и полей его ди ференциальных характеристик: производных по горизонтальн координатам, лапласиана, якобиана и т. п. Точность вычислен этих характеристик по данным сети также поддается априорн оценке [3]. Так, средний квадрат ошибки определения пер конечной разности, отнесенный к дисперсии и умноженный квадрат шага дифференцирования r, т. е. величина

$$D = \frac{r^2 \overline{\beta^2}}{m_H(0)}$$

находится как простая комбинация значений ε, а именно:

$$D = \varepsilon_1 + \varepsilon_2 + \varepsilon_{12}.$$

В формулах (3), (4) черта означает осреднение в статис ческом смысле, индексы 1, 2 соответствуют точкам, по данн анализа в которых вычисляется конечная разность, а значе ε_{12} описывает корреляционную связь между ошибками анал в этих точках [2].

Географическое распределение величин ε, D характериз ошибки анализа геопотенциала и его первой конечной разнос отнесенные к дисперсии элемента. Оценка соответствующих солютных ошибок требует задания поля дисперсии. Отсутст фактических данных о дисперсии геопотенциала в южном по шарии вынудило нас ограничиться в работе [4] расчетами д северного полушария. В настоящее время мы имеем необходим данные по южному полушарию, любезно предоставленные в р поряжение ГГО сотрудником Национального центра исследо ний атмосферы США Р. Л. Дженни (R. L. Jenne). Это позвол привести оценки абсолютных ошибок анализа.

В нашем распоряжении имеется карта среднего квадрати ского отклонения приземного давления σ_p для июля. Эта ка построена Дженни по наблюдениям за 62 дня (июль 1957 и июль 1958 г.). Очевидно, рассчитанная по столь малому ря изменчивость должна заметно отличаться от средней многол ней изменчивости в обычном климатологическом смысле. оценке Дженни, это отличие составляет примерно 40% в стор занижения. Можно, однако, думать, что географическое расп деление изменчивости на полушарии отражено верно. По аналогичной причине мы сочли возможным воспользоваться данными об изменчивости приземного давления σ_p, а не данными об изменчивости геопотенциала σ_H, полагая связь между σ_p и σ_H прямо пропорциональной.

Рис. 1. Распределение абсолютных ошибок анализа геопотенциала (в условных единицах).

На рис. 1 приведено распределение величины

$$K=\sqrt{\varepsilon}\,\sigma_p. \qquad (5)$$

Эта величина численно равна средней квадратической ошибке интерполяции приземного давления (в миллибарах) по наблюдениям существующей в южном полушарии аэрологической сети. Разумеется, она не совпадает с ошибкой интерполяции геопотенциала поверхности 500 мб, однако, как уже сказано, може приближенно считаться пропорциональной ей. Распределение зависит, с одной стороны, от особенностей сети, с другой сто роны,— от характера поля σ_P. Из 110 станций 14 находятс в Антарктиде, 96 — на других материках. Обширные океаниче ские акватории почти полностью лишены аэрологических наблю дений. В Антарктиде сочетание «густой» (конечно, по масштаба южного полушария) сети и относительно малой изменчиво сти обусловливает сравнительно высокую точность анализа Здесь значения K лишь в 2—3 раза больше, чем в Австралии

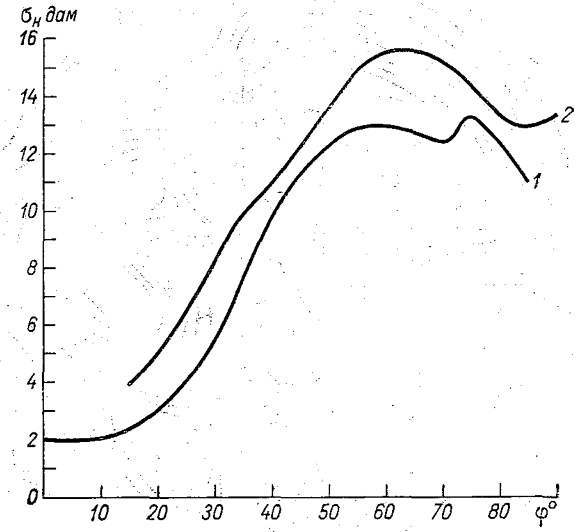


Рис. 2. Средний зональный профиль изменчивости геопотенциала поверхности 500 мбар для зимнего сезона.

1 — южное полушарие, 2 — северное полушарие.

Южной Африке и Южной Америке. Заметим, что в северном по лушарии Центральный полярный бассейн относится к районам наименьшей точности. Разумеется, «высокая» точность анализа в Антарктиде очень далека от требований существующих схем численного прогноза. Много хуже обстоит дело на океанах, зани мающих подавляющую часть площади полушария. Рис. 1 дает также представление о распределении абсолютной средней квад ратической ошибки вычисления первой конечной разности

$$\sqrt{B} = \sigma_H \sqrt{D}. \quad (6$$

Данные, представленные на рис. 1, могут оказаться полез ными при планировании новых наблюдательных систем, вклю

іающих метеорологические спутники, океанские буйковые стан-
ции и автоматические трансозонды.

Обратимся теперь к вопросу о том, какую погрешность внес-
ю бы в анализ поля геопотенциала использование осредненного
дoль широтных кругов зонального профиля изменчивости вместо
рактически использованной карты. Как известно, такой упрощен-
ый способ задания σ_H достаточно широко используется в опе-
ративных схемах анализа. Мы воспользовались полученными от
Дженни данными для расчета зимнего зонального профиля
изменчивости геопотенциала поверхности 500 мбар для южного
олушария. Рассчитанный нами профиль приведен на рис. 2. Его
тличие от аналогичного профиля для северного полушария [5]
бусловлено не только климатологическими причинами, но и ма-
ым объемом выборки. Это не позволяет сравнивать абсолютные
начения σ_H на одинаковых широтах двух полушарий, однако не
иешает сопоставить зависимость от широты. Наиболее сущест-
енным отличием профиля южного полушария оказывается бли-
ость значений σ_H на широтах 50, 60, и 70°; в северном полуша-
рии выделяется максимум на широте 60°.

При сопоставлении сглаженного зонального профиля изменчи-
ости с картой ее для южного полушария выявилось, что разли-
ия между ними наиболее велики на параллелях 50, 60 и 70°,
доль которых в поле дисперсии локализованы замкнутые экст-
ремумы. Выполненные оценки показали, что по указанной причине
редние квадратические ошибки $\sqrt{E}$ и $\sqrt{B}$ оказываются здесь за-
иженными на 15—20%, а в некоторых случаях фиктивное зани-
кение ошибок достигает 30%. Это обстоятельство явно указы-
ает на нецелесообразность использования зонального профиля
изменчивости при оценках точности анализа.

СПИСОК ЛИТЕРАТУРЫ

1. Гандин Л. С. Объективный анализ метеорологических полей. Л., Гидрометеоиздат, 1963. 287 с.
2. Гандин Л. С., Ильин Б. М., Либерман Ю. М., Юдин М. И. О точности определения конечных разностей при анализе метеорологических полей.—«Труды ГГО», 1965, вып. 168, с. 113—122.
3. Либерман Ю. М. О точности определения конечных разностей при анализе поля геопотенциала по данным существующей аэрологической сети. — «Изв. АН СССР. Физика атмосферы и океана», 1965, т. 1, № 11, с. 1130—1137.
4. Либерман Ю. М. О точности анализа поля геопотенциала над северным и южным полушариями. — «Труды ГГО», 1968, вып. 228, с. 41—48.
5. Заставенко Л. Г. Барическое поле тропосферы северного полушария. — «Труды НИИАК», 1972, вып. 86. 236 с.

Р. Л. КАГА

О ТОЧНОСТИ ВЫБОРОЧНЫХ СТРУКТУРНЫХ ФУНКЦИЙ

1. При решении задач статистической метеорологии одним из наиболее употребительных характеристик пространственно изменчивости поля являются структурные функции. В связ с этим представляет практический интерес оценка точности, к торая может быть достигнута при получении пространственны структурных функций по ограниченному объему данных.

Некоторые оценки такого рода будут получены ниже. Пр этом будет использоваться подход, примененный для оценки то ности пространственных ковариационных и корреляционных фун ций в работах [1, 2], продолжением которых является настояща статья. Этот подход сводится к определению пространственно корреляции между различными выборочными моментами и к уч ту этой связности при оценке точности, которая достигаетс при осреднении индивидуальных моментов. Учитывая, что с ответствующая методика подробно изложена в [1, 2], мы по во можности сократим выкладки и уделим основное внимание ана лизу полученных результатов.

2. Будем рассматривать случайное поле величины f (x, y) Значение этой величины в точке i с координатами x_i и y_i буде обозначать f_i. Полагаем, что средние значения величины f в всех точках поля равны нулю. На общность наших выводов эт не повлияет.

Будем обозначать чертой сверху осреднение по статистическо му ансамблю. Тогда для любой точки имеем

$$\overline{f_i} = 0. \qquad (1$$

Пространственная изменчивость величины f между точкам i и k характеризуется значением структурной функции

$$b_{ik} = \overline{(f_i - f_k)^2}. \qquad (2$$

Для описания связи между значениями f и f_k могут использоваться также ковариационный (R_{ik}) и корреляционный (r_{ik}) моменты:

$$R_{ik}=\overline{f_i f_k}; \qquad r_{ik}=\frac{\overline{f_i f_k}}{\sqrt{\overline{f_i^2}\,\overline{f_k^2}}}=\frac{R_{ik}}{\sqrt{R_{ii}R_{kk}}}. \tag{3}$$

Значения структурной функции могут быть выражены через ковариации с помощью формулы

$$b_{ik}=R_{ii}+R_{kk}-2R_{ik}. \tag{4}$$

Для однородного поля

$$R_{ii}=R_{kk}=\sigma^2 \quad \text{и} \quad b_{ik}=2\sigma^2(1-r_{ik}), \tag{5}$$

где σ^2 — дисперсия поля.

Практически расчет моментов ведется лишь по ограниченному объему данных, вследствие чего полученные значения их будут отличаться от истинных. Пусть мы располагаем n независимыми реализациями поля f. Тогда для каждой пары точек можно получить оценку величин b. Так, для пары точек i и k

$$b^*_{ik}=\frac{1}{n}\sum_{j=1}^{n}\left(f_i^{(j)}-f_k^{(j)}\right)^2. \tag{6}$$

Здесь b^*_{ik} — выборочное значение структурной функции. Верхний индекс при величине f означает номер реализации, так что суммирование производится по всем реализациям.

Для пары точек s и t имеем аналогично

$$b^*_{st}=\frac{1}{n}\sum_{j=1}^{n}\left(f_s^{(j)}-f_t^{(j)}\right)^2. \tag{6а}$$

Нетрудно видеть, что величины b^*_{ik} и b^*_{st} являются несмещенными оценками соответствующих значений структурной функции, так что

$$\overline{b^*_{ik}}=b_{ik}; \quad \overline{b^*_{st}}=b_{st}. \tag{7}$$

Статистическая связь между этими оценками характеризуется их ковариацией

$$m_b(i,\ k;\ s,\ t)=\overline{b^*_{ik}b^*_{st}}-b_{ik}\,b_{st}. \tag{8}$$

Учитывая (3), (6) и (6а), получаем:

$$\overline{b_{ik}^{*} b_{st}^{*}} = \frac{1}{n^2} \sum_{j=1}^{m} \sum_{g=1}^{m} \overline{\left(f_i^{(j)} - f_k^{(j)}\right)^2 \left(f_s^{(g)} - f_t^{(g)}\right)^2};$$

$$b_{ik} b_{st} = (R_{ii} + R_{kk} - 2R_{ik})(R_{ss} + R_{tt} - 2R_{st}). \tag{9}$$

Таким образом, произведение $b_{ik}b_{st}$ выражается через вторые моменты поля, а в $\overline{b_{ik}^{*}\, b_{st}^{*}}$ входят четвертые моменты. Полагаем, что величина f имеет распределение, близкое к нормальному, так что четвертые моменты ее связаны со вторыми формулой

$$\overline{f_i^{(j)} f_k^{(j)} f_s^{(g)} f_t^{(g)}} = R_{ik} R_{st} + \delta_{jg}(R_{is}R_{kt} + R_{it}R_{ks}), \tag{10}$$

где

$$\delta_{jg} = \begin{cases} 1 & \text{при } j = g \\ 0 & \text{при } j \neq g \end{cases} \text{— символ Кронекера.}$$

Подстановка (9) и (10) в (8) дает

$$m_b(i, k; s, t) = \frac{2}{n}(R_{it} + R_{ks} - R_{is} - R_{kt})^2 =$$

$$= \frac{1}{2n}(b_{it} + b_{ks} - b_{is} - b_{kt})^2. \tag{11}$$

Частными случаями (11) являются формулы, определяющие дисперсии значений структурной функции:

$$\sigma_{b_{ik}}^2 = m_b(i, k; i, k) = \frac{2}{n} b_{ik}^2 = \frac{2}{n}(R_{ii} + R_{kk} - 2R_{ik})^2;$$

$$\sigma_{b_{st}}^2 = m_b(s, t; s, t) = \frac{2}{n} b_{st}^2 = \frac{2}{n}(R_{ss} + R_{tt} - 2R_{st})^2. \tag{12}$$

Учитывая (11) и (12), получаем коэффициент корреляции между значениями b_{ik}^{*} и b_{st}^{*}

$$\mu_b(i, k; s, t) = \frac{m_b(i, k; s, t)}{\sigma_{b_{ik}} \sigma_{b_{st}}} = \frac{(b_{it} + b_{ks} - b_{is} - b_{kt})^2}{4 b_{ik} b_{st}} =$$

$$= \frac{(R_{it} + R_{ks} - R_{is} - R_{kt})^2}{(R_{it} + R_{kk} - 2R_{ik})(R_{ss} + R_{tt} - 2R_{st})}. \tag{13}$$

Заметим, что этот коэффициент корреляции не зависит от числа реализаций, использованных при расчете, и определяется лишь характером связи между значениями поля в пунктах задания исходных данных. Для случая однородного и изотропного случайного поля эта связь, а следовательно, и величина μ_b определяется взаимным расположением точек i, k, s, t, расстоянием между ними и заданием пространственной корреляционной функции поля f.

В этом случае подстановка (5) в (12) и (13) дает:

$$\sigma^2_{b_{ik}} = \frac{8\sigma^4}{n}(1 - r_{ik})^2;$$

$$\sigma^2_{b_{st}} = \frac{8\sigma^4}{n}(1 - r_{st})^2;$$

$$\mu_b(i,\ k;\ s,\ t) = \frac{(r_{it} + r_{ks} - r_{is} - r_{kt})^2}{4(1 - r_{ik})(1 - r_{st})}. \tag{14}$$

3. Аналогично тому, как это делалось в [1], ограничимся рассмотрением трех вариантов расположения точек, при которых расстояние между точками i и k и между точками s и t одина-

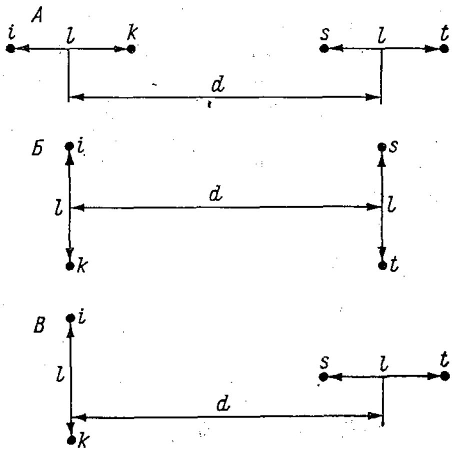


Рис. 1. Схема расположения точек задания исходных данных в вариантах А, Б, В.

ково. Обозначим его l, а расстояние между точками, делящими соответствующие отрезки пополам, обозначим d (см. рис. 1).

Нетрудно видеть, что для случая, когда все точки расположены вдоль одной прямой (вариант А, соответствующий «продольной» корреляции значений структурной функции)

$$\mu_b^{\text{А}}(l, \quad = \frac{1}{4}\left[\frac{2r(d) - r(d + l) - r(|d - l|)}{1 - r(l)}\right]^2. \tag{15а}$$

Для варианта Б, соответствующего «поперечной» корреляции выборочной структурной функции, получаем

$$\mu_b^{\text{Б}}(l,\ d) = \left[\frac{r(d) - r(\sqrt{l^2 + d^2})}{1 - r(l)}\right]^2. \tag{15б}$$

Конкретные значения $\mu_b^{\text{А}}$ и $\mu_b^{\text{Б}}$ существенно зависят от статистической структуры поля.

Для варианта В, независимо от типа статистической структуры, корреляция между значениями выборочной структурной функции отсутствует, так что

$$\mu_b^{\text{В}}(l,\ d) = 0. \tag{15в}$$

Это означает, что выборочные структурные функции в двух взаимноперпендикулярных направлениях являются независимыми и существенно дополняют друг друга. Для ковариационных и корреляционных функций такой независимости не отмечается, так что совместное использование этих функций, вычисленных для разных направлений, гораздо менее эффективно.

Заметим также, что, независимо от структуры поля, корреляция значений структурной функции не может быть отрицательной. Как следует из [1], корреляция ковариационных моментов может быть отрицательной для поля, пространственная корреляционная функция которого имеет отрицательную область, а для выборочных корреляционных моментов корреляция может оказаться отрицательной и при существенно положительных корреляционных функциях поля f.

Конкретные расчеты по формулам (15) были выполнены для корреляционных функций вида:

$$\begin{aligned} r_1(l) &= \exp\{-(l/l_0)^2\}; \\ r_2(l) &= (1 + l/l_0)\exp\{-l/l_0\}; \\ r_3(l) &= \exp\{-l/l_0\}. \end{aligned} \tag{16}$$

Некоторые результаты приводятся в табл. 1, из которой видно, что корреляция значений выборочных структурных функций является наиболее высокой для случайных полей, структура которых описывается функцией r_1, а функция r_3 соответствует наименьшим значениям μ_b. Корреляция в продольном направлении ($\mu_b^{\text{А}}$) меньше корреляции в поперечном направлении ($\mu_b^{\text{Б}}$).

Для корреляционных функций (16) при фиксированном d отмечается увеличение μ_b с увеличением расстояния l между точками задания исходных данных (т. е. с уменьшением $r(l)$). Исключением в этом отношении является случай, когда эти точки

Таблица 1

Коэффициент корреляции между выборочными значениями структурной функции

$r(d)$	$r(l)$							
	Схема А				Схема Б			
	0,9	0,8	0,7	0,5	0,9	0,8	0,7	0,5
	$r_3(l)=\exp\{-l/l_0\}$							
1,0	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,9	0,00	0,23	0,43	0,63	0,15	0,32	0,49	0,65
0,8	0,00	0,01	0,08	0,33	0,04	0,11	0,23	0,40
0,7	0,00	0,01	0,02	0,11	0,01	0,04	0,10	0,23
0,6	0,00	0,01	0,02	0,00	0,00	0,02	0,04	0,13
0,5	0,00	0,00	0,01	0,06	0,00	0,01	0,02	0,06
0,4	0,00	0,00	0,01	0,04	0,00	0,00	0,01	0,03
0,3	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,01
0,2	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00
	$r_2(l)=(1+l/l_0)\exp\{-l/l_0\}$							
1,0	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,9	0,19	0,33	0,43	0,59	0,55	0,62	0,66	0,72
0,8	0,02	0,05	0,12	0,28	0,33	0,39	0,44	0,51
0,7	0,00	0,00	0,01	0,08	0,20	0,24	0,28	0,32
0,6	0,02	0,01	0,01	0,00	0,12	0,15	0,17	0,23
0,5	0,03	0,04	0,04	0,02	0,06	0,08	0,10	0,14
0,4	0,04	0,05	0,06	0,07	0,03	0,04	0,05	0,08
0,3	0,03	0,04	0,06	0,09	0,01	0,02	0,02	0,04
0,2	0,02	0,03	0,04	0,08	0,00	0,01	0,01	0,01
	$r_1(l)=\exp\{-(l/l_0)^2\}$							
1,0	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,9	0,52	0,53	0,55	0,59	0,81	0,81	0,81	0,81
0,8	0,21	0,23	0,24	0,29	0,64	0,64	0,64	0,64
0,7	0,05	0,06	0,07	0,10	0,49	0,49	0,49	0,49
0,6	0,00	0,00	0,00	0,02	0,36	0,36	0,36	0,36
0,5	0,03	0,02	0,02	0,00	0,25	0,25	0,25	0,25
0,4	0,10	0,09	0,08	0,05	0,16	0,16	0,16	0,16
0,3	0,17	0,16	0,15	0,13	0,09	0,09	0,09	0,09
0,2	0,20	0,20	0,20	0,19	0,04	0,04	0,04	0,04

расположены в соответствии со схемой Б, а структура поля за дается функцией r_1. В этом случае

$$\mu_b^{Б}(l,\ d) = r_1^2(d). \qquad (17$$

Заметим, что, как следует из [1, 2], корреляция ковариаци (μ_R) и коэффициентов корреляции (μ_r) для этого случая такж равна $r_1^2\ (d)$. Для других корреляционных функций и при дру гом расположении точек значения $\mu_b(l,\ d)$, как правило, сущес венно меньше соответствующих значений $\mu_r(l,\ d)$ и $\mu_R(l,\ d)$

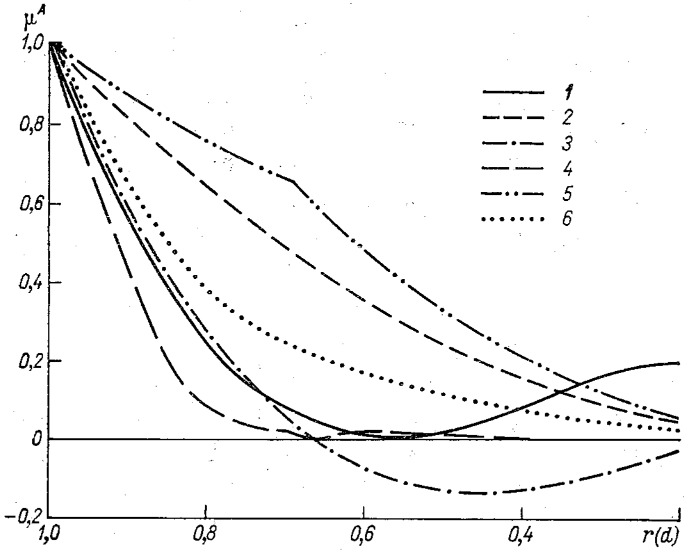


Рис. 2. Зависимость корреляции выборочных моментов от расстояния между парами станций (схема А) при корреляционных функциях
$r_1(1-\mu_b(l,\ d);\ 2-\mu_R(l,\ d);\ 3-\mu_r(l,\ d))$ и $r_3(4-\mu_b(l,\ d);\ 5-\mu_R(l,\ d);\ 6-\mu_r(l,\ d))$.

Это видно, в частности, из рис. 2, на котором приводится зависимость коэффициентов корреляции между различными выборочными моментами от d для схемы А. При таком расположении точек для фиксированных l корреляция μ_b при увеличении d убывает до нуля, после чего имеет вторичный максимум. Для расположения точек по схеме Б характерно монотонное убывание μ_b с увеличением расстояния d. Следует иметь в виду, что такой характер зависимости $\mu_b(l,\ d)$ в значительной мере определяется спецификой рассмотренных корреляционных функций (16). Для случайных функций другого вида эта зависимость мо-ет измениться. Так, например, в случае «закона первой сте- ени»

$$r_1(l) = 1 - l/l_0 \tag{18}$$

одстановка в (15а) дает

$$\mu_b^\wedge(l,\ d) = \begin{cases} \left(\dfrac{l-d}{l}\right)^2 & \text{при } d \leqslant l \leqslant l_0, \\ 0 & \text{при } l < d \leqslant l_0. \end{cases} \tag{19}$$

4. Данные о пространственной связности отдельных значений ыборочной структурной функции могут использоваться для опре- еления точности расчета структурной функции.

Пусть оценка структурной функции для некоторой градации асстояний, срединное значение которой равно l, получается пу- м осреднения N моментов, соответствующих парам пунктов, асстояния между которыми лежат в пределах данной града- ии:

$$\hat{b}(l) = \frac{1}{N} \sum_{p=1}^{N} b^*_{i_p k_p}. \tag{20}$$

Дисперсия величины $\hat{b}(l)$ определяется формулой

$$\sigma^2_{\hat{b}(l)} = \sigma^2_{b(l)}\, \hat{\mu}_b(l), \tag{21}$$

е, в соответствии с (12),

$$\sigma^2_{b(l)} = \frac{2}{n}\, b^2(l) \tag{22}$$

ть дисперсия значения структурной функции для одной пары чек, а

$$\hat{\mu}_b(l) = \frac{1}{N^2} \sum_{p=1}^{N} \sum_{q=1}^{N} \mu_b(i_p,\ k_p;\ i_q,\ k_q) \tag{23}$$

ть среднее из коэффициентов корреляции между всеми значе- иями структурной функции, попавшими в данную градацию рас- ояния.

Из формулы (23) видно, что вследствие положительной связ- ости между отдельными значениями выборочной пространствен- ой структурной функции, точность ее с увеличением числа пар растет медленнее, чем это имело бы место при их независи- ости. При увеличении числа пунктов, более или менее равно- ерно расположенных на фиксированной территории, величина

$\mu_b(l)$ стремится не к нулю, а некоторой конечной величине. Пр таком предельном переходе выражение (23) преобразуетс в кратный интеграл (подробнее см. [1]), вычисление которо определяет максимально возможную точность расчета структу ной функции при использовании данных наблюдений за велич ной f на рассматриваемой территории.

Оценки величины $\hat{\mu}_b$, полученные таким образом для ква ратов со стороной $L = l_0$ при экспоненциальной корреляционно функции r_3, приводятся в табл. 2. Там же для сравнения прив дятся полученные в [1, 2] значениях средних по территории коэфф циентов корреляции ковариаций и коэффициентов корреляци $\hat{\mu}_R$ и $\hat{\mu}_r$ соответственно [1].

Таблица

Зависимость $\hat{\mu}_b$, $\hat{\mu}_R$ и $\hat{\mu}_r$ от l для квадратов со стороной $L = l_0$

$r(l) = \exp\{-l/l_0\}$

$\hat{\mu}$	l/l_0					
	0	0,2	0,4	0,6	0,8	1,0
$\hat{\mu}_b$	0,00	0,02	0,07	0,13	0,20	0,25
$\hat{\mu}_R$	0,39	0,48	0,56	0,58	0,58	0,56
$\hat{\mu}_r$	0,20	0,23	0,30	0,35	0,40	0,43

Из табл. 2 видно, что осреднение значений структурной фун ции, полученных для различных пар станций на рассматриваемс площади, тем более уточняет структурную функцию, чем меньш значения аргумента l. Как уже указывалось ранее, уточнені значительно больше, чем соответствующее уточнение, достигаемс при осреднении ковариационных и корреляционных моментов.

При увеличении размеров квадрата L значения $\hat{\mu}_b$ уменьш ются, что соответствует уменьшению ошибки выборочной стру турной функции.

Для экспоненциальной корреляционной функции r_3 завис мость $\hat{\mu}_b$ от аргумента l и от стороны квадрата L удовлетвор тельно аппроксимируется формулой

$$\hat{\mu}_b(l, L) = 0{,}28 \frac{l^{1,6}}{L^{1,85}}. \qquad (2$$

[1] Заметим, что в [1] значения $\hat{\mu}_R(l)$ для этого случая вследствие ошиб в расчетах были приведены неточно. Здесь эта ошибка устранена.

то видно из рис. 3, на котором приводятся значения $\hat{\mu}_b(l, L)$, ассчитанные по этой формуле (линии) и путем приближенного ычисления интегрального аналога формулы (23) (точки).

Подстановка (22) и (24) в (21) дает для корреляционной ункции r_3

$$\frac{\sigma_{\hat{b}}}{b}=0{,}75\,\frac{l^{0.8}}{n^{1/2}L^{0{,}92}}. \qquad (25)$$

Как показано в [1], оценками, типа полученных выше для вадрата, можно приближенно пользоваться для областей осредне- ия другой формы, лишь бы они были не слишком вытянутыми. этом случае входящий в (24) и (25) параметр L определяется ак сторона квадрата, имеющего ту же площадь S, что и область адания исходных данных, т. е.

$$L=\sqrt{S}. \qquad (26)$$

Значения погрешности (25) являются заниженными, посколь- у при получении их предполагалось, что используются данные аблюдений во всех точках бласти. Оценки показывают, днако, что при задании 50— 0 пунктов, равномерно рас- оложенных по территории, асхождениями между сумма- и (23) и интегралами, полу- ающимися при предельном ереходе, практически можно ренебречь.

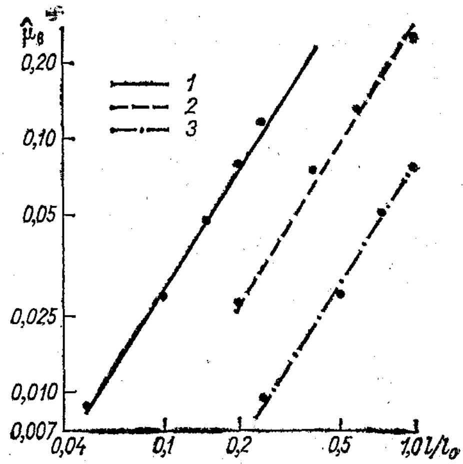


Рис. 3. Зависимость $\hat{\mu}_b(l, L)$ в случае экспоненциальной корреляционной функции r_3.
$1-L=0{,}5l_0$; $2-L=l_0$; $3-L=2l_0$.

Для полей, структура кото- ых описывается корреляцион- ыми функциями r_1 и r_2, полу- ить такие аппроксимацион- ые формулы оказывается за- руднительным. Для того что- ы дать представление о ха- актере зависимости точности труктурных функций от их ргумента и от площади обла- ти осреднения, в табл. 3 при- одятся значения $\hat{\mu}_b$ для раз- ичных корреляционных функ- ий. С целью большей сопоставимости результатов, они даны не ля фиксированных значений аргумента, а для фиксированных начений корреляционной (а значит, и структурной) функции.

Из табл. 3 и из данных, приведенных в [1], следует, что из ассмотренных типов случайных полей при одном и том же объеме выборки характеристики статистической структуры опред ляются точнее всего для полей, описываемых корреляционнс функцией r_3. С наименьшей точностью характеристики структ ры определяются для полей, описываемых функцией r_1.

Таблица

Зависимость $\hat{\mu}_b$ от l для квадратов со стороной L

r (L)	r	r (l)						
		0,9	0,8	0,7	0,6	0,5	0,4	0,3
0,3	r_1	0,23	0,28	0,31	0,34	0,37	0,39	0,42
	r_2	0,09	0,13	0,18	0,23	0,26	0,31	0,36
	r_3	0,01	0,02	0,04	0,07	0,12	0,17	0,24
0,1	r_1	0,15	0,17	0,20	0,21	0,24	0,26	0,30
	r_2	0,05	0,06	0,08	0,09	0,12	0,15	0,19
	r_3	0,00	0,01	0,01	0,02	0,03	0,05	0,08

В заключение отметим, что наиболее существенной особе ностью полученных оценок погрешностей выборочных структу ных функций представляется их более быстрое уменьшение пр привлечении дополнительных данных по сравнению с ковари ционными и корреляционными функциями. Если учесть, чт структурные функции имеют и некоторые другие практически преимущества по сравнению с последними (например, меньшу чувствительность к временной нестационарности поля), то в н которых случаях может оказаться предпочтительным пользоват ся структурными функциями, как более надежно определяемым на эмпирическом материале.

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л. О точности расчета пространственных корреляционных фун ций. — «Труды ГГО», 1973, вып. 308, с. 3—19.
2. Каган Р. Л. О точности расчета пространственных корреляционных функ ций. II.—«Труды ГГО», 1974, вып. 336, с. 3—19.

Р. Л. КАГАН, Е. И. ФЕДОРЧЕНКО

К РАСЧЕТУ ВЕРОЯТНОСТИ ВЫБРОСА НОРМАЛЬНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

1. В ряде метеорологических задач приходится вычислять вероятность выброса значений метеорологического элемента через заданный уровень. Аналитически вероятность $p(C)$ выброса вверх через уровень C случайной последовательности f на промежутке $[t_1, t_2]$ может быть найдена по формуле (см., например, [1])

$$p(C)=\int_{-\infty}^{C}\int_{C}^{\infty}\varphi(x, y)\,dx dy, \tag{1}$$

где $\varphi(x, y)$ — совместная плотность распределения значений последовательности f в моменты времени t_1 и t_2.

Для нормальных случайных величин интегральные двумерные функции распределения затабулированы [2], что позволяет, исходя из формулы (1), с помощью таблиц вручную оценивать значения вероятностей выбросов нормальных последовательностей. Практически, однако, при проведении многократных расчетов возникает необходимость реализации алгоритма вычисления вероятности выброса на ЭВМ.

В [3] для расчета на ЭВМ вероятности выброса нормальной последовательности использовалась формула (1), приведенная к виду

$$p(C, r)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c_1}e^{-\frac{x^2}{2}}\left[1-\Phi\left(\frac{c_2-rx}{\sqrt{1-r^2}}\right)\right]dx. \tag{2}$$

Здесь $c_1=\frac{C-m_1}{\sigma_1}$, $c_2=\frac{C-m_2}{\sigma_2}$ — нормированные уклонения уровня C от математических ожиданий m_1 и m_2 величины f для моментов времени t_1 и t_2 соответственно, а σ_1 и σ_2 — соответств ющие средние квадратические отклонения; r — коэффициент ко реляции между значениями величины f в эти моменты времени,

$$\Phi(u)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-\frac{z^2}{2}}dz \qquad ($$

— интеграл вероятностей. Двукратное интегрирование в форму (2) выполняется численно.

Более эффективный алгоритм расчета вероятности выбро нормальной последовательности, требующий лишь однократно численного интегрирования, может быть построен при использ вании формулы [4]:

$$p(C,r)=\Phi(c_1)\left[1-\Phi(c_2)\right]-\frac{1}{2\pi}\int_0^r\frac{1}{\sqrt{1-x^2}}\exp\left(-\frac{c_1^2-2xc_1c_2+c_2^2}{2(1-x^2)}\right)dx. \qquad ($$

В частности, для случая стационарной последовательност для которой $m_1=m_2=m$; $\sigma_1=\sigma_2=\sigma$; $c_1=c_2=c$, формула (приводится к виду

$$p(C,\ r)=\frac{1}{2\pi}\int_r^1\frac{1}{\sqrt{1-x^2}}\exp\left(-\frac{c^2}{1+x}\right)dx. \qquad ($$

Подынтегральная функция в (4) имеет особенность при $r=$ что затрудняет непосредственное ее использование при высок коэффициентах корреляции. Вместе с тем для оценки вероятнос выброса последовательности с высокой временно́й связность оказывается возможным получить упрощенные формулы, осн ванные на асимптотическом разложении функции $p(C,\ r)$ в о рестности $r=1$, расчеты по которым можно производить да без помощи вычислительной машины.

2. Рассмотрим сначала случай нормальной стационарной п следовательности. Анализ формулы (5) показывает, что п $r\to 1$ вероятность $p(C,\ r)$ пропорциональна $\sqrt{1-r}$. Поэто удобно рассматривать величину $y\ (c,\ r)$, определяемую форм лой

$$y(c,\ r)=(1-r)^{-\frac{1}{2}}p(C,\ r)=\frac{(1-r)^{-\frac{1}{2}}}{2\pi}\int_r^1\frac{1}{\sqrt{1-x^2}}\exp\left(-\frac{c^2}{1+x}\right)dx. \qquad ($$

Представим $y(c,\ r)$ в виде разложения в ряд Тейлора в окр тности $r=1$. Для нахождения производных этой величины воспользуемся тождеством

$$2(1-r)\frac{\partial y(c,\ r)}{\partial r}-y(c,\ r)=-F(c,\ r), \tag{7}$$

где

$$F(c,\ r)=\frac{1}{\pi}(1+r)^{-\frac{1}{2}}\exp\left(-\frac{c^2}{1+r}\right). \tag{8}$$

Функция $F(c,\ r)$ не имеет особенностей при $r>-1$. Можно доказать, что для нее справедливо равенство

$$\frac{\partial^n F(c,\ r)}{\partial r^n}=\frac{1}{2^{2n}}\frac{\partial^{2n}F(c,\ r)}{\partial c^{2n}}. \tag{9}$$

Последовательно дифференцируя (7) по r и учитывая, что производные от $y(c,\ r)$ конечны, получаем

$$\left.\frac{\partial^n y(c,\ r)}{\partial r^n}\right|_{r=1}=-\frac{1}{1+2n}\left.\frac{\partial^n F(c,\ r)}{\partial r^n}\right|_{r=1} \tag{10}$$

или, согласно (9),

$$\left.\frac{\partial^n y(c,\ r)}{\partial r^n}\right|_{r=1}=-\frac{1}{(1+2n)\,2^{2n}}\frac{\partial^{2n}F(c,\ 1)}{\partial c^{2n}}=$$

$$=-\frac{1}{\sqrt{\pi}\,(1+2n)\cdot 2^{2n}}\frac{\partial^{2n}}{\partial c^{2n}}\left[\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{c^2}{2}\right)\right]=$$

$$=-\frac{\sqrt{2}}{2\pi(1+2n)\cdot 2^{2n}}\exp\left(-\frac{c^2}{2}\right)H_{2n}(c), \tag{11}$$

где $H_{2n}(c)$ — полиномы Чебышева — Эрмиата [5].

С учетом (11) разложение $y(c,\ r)$, в ряд Тейлора в окрестности $r=1$ будет иметь вид

$$y(c,\ r)=\frac{\sqrt{2}}{2\pi}\exp\left(-\frac{c^2}{2}\right)\sum_{i=0}^{\infty}(-1)^i\frac{(1-r)^i}{2^{2i}\,i!\,(1+2i)}H_{2i}(c). \tag{12}$$

Подстановка (12) в (6) дает выражение для расчета вероятности выброса нормальной стационарной последовательности

$$p(C,\ r)=\frac{\sqrt{2(1-r)}}{2\pi}\exp\left(-\frac{c^2}{2}\right)\sum_{i=0}^{\infty}(-1)^i\frac{(1-r)^i}{2^{2i}\,i!\,(1+2i)}H_{2i}(c). \tag{13}$$

Практически полиномы $H_{2i}(c)$ удобно находить по рекуррентной формуле

$$H_{2i+2}(c)=(c^2-4i-1)\,H_{2i}(c)-2i(2i-1)\,H_{2i-2}(c), \tag{14}$$

учитывая, что

$$H_0(c)=1, \quad H_2(c)=c^2-1. \qquad (1$$

Входящий в (13) ряд сходится при $-1<r\leqslant 1$. Скорость сх димости тем больше, чем меньше относительный уровень c. О нако даже при очень высоких значениях $c=4$ использование членов разложения для любых положительных коэффициентов ко реляции обеспечивает расчет вероятности с погрешностью, не пр вышающей 10^{-7}.

Наиболее удобно, однако, использовать формулу (13) пр высоких значениях коэффициентов корреляции, когда в ней мо но ограничиться несколькими первыми членами разложения. В многих случаях оказывается достаточным использование фо мул

$$p_0(C,\ r)=\frac{\sqrt{2(1-r)}}{2\pi}\exp\left(-\frac{c^2}{2}\right) \qquad (13$$

или

$$p_1(C,\ r)=\frac{\sqrt{2(1-r)}}{2\pi}\exp\left(-\frac{c^2}{2}\right)\left[1-\frac{1}{12}(c^2-1)(1-r)\right], \qquad (13$$

получающихся при удержании в (13) одного или двух член разложения.

Формула (13а) при $r\geqslant 0{,}5$ и $c\leqslant 1{,}5$ определяет вероятнос выброса с относительной погрешностью, не превышающей 5 что в ряде случаев достаточно для практических приложени Для более высоких уровней точность этой формулы сравнительн невысока. Как видно из рис. 1, характеризующего погрешнос двучленной формулы (13б), последняя при $r\geqslant 0{,}8$ обеспечива точность в пределах 5% даже для очень высоких уровн $c=4$.

Заметим, что в указанном диапазоне изменения аргумент бо́льшая точность обеспечивается приведенными в [6] аппро симационными формулами:

$$p=\frac{1}{2\pi}\arccos r\exp\left\{-\frac{c^2}{2}-0{,}095(1-r)\,c^{1{,}85}\right\}, \qquad (16$$

$$p=\frac{1}{2\pi}\arccos r\exp\left\{-(0{,}58-0{,}08r)\,c^2\right\}. \qquad (16$$

В большинстве случаев, однако, предпочтительнее исполь вание формул (13а) или (13б) как из-за простоты их непосредс венного применения, так и благодаря возможности их уточнен путем привлечения следующих членов разложения.

3. Для случая нестационарной нормальной последовательн сти, характеризующейся не очень большой вариацией матем

ического ожидания и среднего квадратического отклонения от члена к члену, естественно представить вероятность выброса

$$p(C,\ r) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \times$$

$$\times \int\limits_{-\infty}^{C}\int\limits_{C}^{\infty} \exp\left\{-\frac{(x-m_1)^2-2r(x-m_1)(y-m_2)+(y-m_2)^2}{2\sigma_1\sigma_2(1-r^2)}\right\} dxdy =$$

$$= \frac{1}{2\pi\sqrt{1-r^2}} \int\limits_{-\infty}^{c_1}\int\limits_{c_2}^{\infty} \exp\left\{-\frac{u^2-2ruv+v^2}{2(1-r^2)}\right\} dudv = f(c_1,\ c_2) \qquad (17)$$

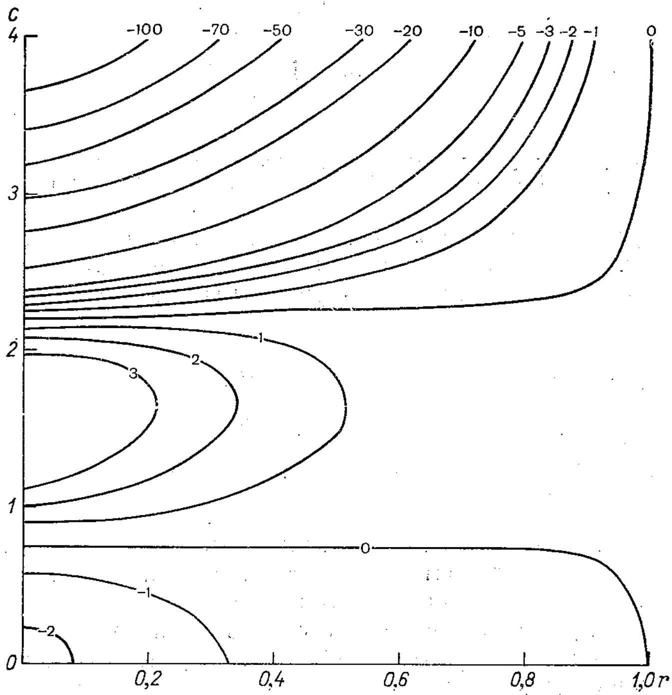


Рис. 1. Относительная погрешность (%) определения вероятности выброса по формуле (13 б).

в виде разложения в ряд

$$f(c_1,\ c_2) = f(c_1,\ c_1) + \left.\frac{\partial f(c_1,\ c_2)}{\partial c_2}\right|_{c_2=c_1}(c_2-c_1) +$$

$$+ \frac{1}{2}\left.\frac{\partial^2 f(c_1,\ c_2)}{\partial c_2^2}\right|_{c_2=c_1}(c_2-c_1)^2 + \ldots \qquad (18)$$

Дифференцируя $f(c_1, c_2)$ по второму аргументу, получаем:

$$A_1 = \left.\frac{\partial f(c_1, c_2)}{\partial c_2}\right|_{c_2=c_1} = -\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{c_1^2}{2}\right) \Phi\left(c_1 \sqrt{\frac{1-r}{1+r}}\right);$$

$$A_2 = \left.\frac{\partial^2 f(c_1, c_2)}{\partial c_2^2}\right|_{c_2=c_1} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{c_1^2}{2}\right) \times$$

$$\times \left[c_1 \Phi\left(c_1 \sqrt{\frac{1-r}{1+r}}\right) + \frac{r}{\sqrt{2\pi(1-r^2)}} \exp\left(-\frac{c_1^2}{2}\frac{1-r}{1+r}\right)\right];$$

$$A_3 = \left.\frac{\partial^3 f(c_1, c_2)}{\partial c_2^3}\right|_{c_2=c_1} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{c_1^2}{2}\right) \times$$

$$\times \left[-(c_1^2-1)\, \Phi\left(c_1 \sqrt{\frac{1-r}{1+r}}\right) + \right.$$

$$\left. + \left(c_1 - \frac{2rc_1}{\sqrt{2\pi(1-r^2)}}\frac{1}{1+r}\right) \exp\left(-\frac{c_1^2}{2}\frac{1-r}{1+r}\right)\right]. \qquad (19$$

Аналогично можно получить выражение и для следующи производных. Практически во многих случаях достаточно удер жания в (18) первых членов. В результате могут использоватьс формулы

$$p_2(C, r) = f(c_1, c_1) + A_1(c_2 - c_1) \qquad (20$$

или

$$p_3(C, r) = f(c_1, c_1) + A_1(c_2 - c_1) + \frac{1}{2} A_2(c_2 - c_1)^2. \qquad (21$$

Здесь $f(c_1, c_1)$ — вероятность выброса через уровень C стацио нарной последовательности с математическим ожиданием m и средним квадратическим отклонением σ_1.

Заметим, что при не очень больших значениях уровня c_1 и до статочно больших r в формуле (19) для A_1 можно использовать приближенное выражение

$$\Phi\left(c_1 \sqrt{\frac{1-r}{1+r}}\right) \approx \frac{1}{2}\left(1 + \frac{c_1}{\sqrt{\pi}} \sqrt{1-r}\right). \qquad (22$$

Учитывая его, можно использовать приближенную формулу

$$p_4(C, r) = f(c_1, c_1) - \frac{1}{2\sqrt{2\pi}} \times$$

$$\times \exp\left(-\frac{c_1^2}{2}\right)\left(1 + \frac{c_1}{\sqrt{\pi}} \sqrt{1-r}\right)(c_2 - c_1). \qquad (23$$

Точность формул (20), (21) и (23) зависит от степени нестационарности исследуемого процесса. При $|\Delta c|=|c_2-c_2|<0{,}1$ эти формулы обеспечивают достаточную точность расчета вероятности выброса даже для очень высоких уровней c_1. Это хорошо видно, например, из табл. 1, в которой приводятся оценки вероятности выброса для уровня $c_1=3$, полученные по точной формуле (4) и по формулам (20), (21) и (23). При $\Delta c=0{,}05$ отклонения значений, полученных по формуле (20) и даже по формуле (23),

Таблица 1

Вероятность выброса нестационарной последовательности для $c_1=3$ при расчете по точной формуле (p) и приближенным формулам (p_2, p_3 и p_4)

r	Δc												
	0	0,05				0,1				0,2			
	$10^6\,p$	$10^6\,p$	$10^6\,p_2$	$10^6\,p_3$	$10^6\,p_4$	$10^6\,p$	$10^6\,p_2$	$10^6\,p_3$	$10^6\,p_4$	$10^6\,p$	$10^6\,p_2$	$10^6\,p_3$	$10^6\,p_4$
0	1348	1143	1127	1143	1050	967	906	972	751	686	463	728	155
0,1	1345	1140	1124	1141	1056	964	903	970	768	684	462	727	190
0,2	1338	1134	1118	1135	1060	959	898	965	781	680	458	723	224
0,3	1326	1123	1108	1124	1058	950	889	955	791	673	452	715	255
0,4	1301	1104	1088	1104	1048	932	872	937	792	659	440	701	280
0,5	1268	1071	1056	1072	1025	903	843	908	781	636	419	678	294
0,6	1210	1019	1003	1020	989	856	797	861	751	598	383	640	292
0,7	1120	937	921	937	906	781	723	786	693	537	326	580	266
0,8	978	807	791	807	783	663	605	668	588	441	232	484	199
0,9	799	588	572	588	569	464	405	470	399	281	71	329	59

от полученных по точной формуле, лежат в пределах нескольких процентов. Погрешность расчетов по формуле (21) находится в пределах 0,1%. При $\Delta c=0{,}1$ погрешность формул (20) и (23) может превышать 10%, однако погрешность формулы (21) лежит, как правило, в пределах 1%. При $\Delta c=0{,}2$ погрешность даже этой формулы может превысить 10%. Очевидно, при такой нестационарности ряда необходимо использовать более точные формулы.

Исходя из сказанного, можно решать вопрос о возможности использования приближенных способов учета нестационарности ряда при расчете вероятностей выбросов. Рассмотрим в качестве примера ряды средней суточной температуры воздуха. На рис. 2 приводятся значения Δc, характеризующие нестационарность этих рядов для ГМС Москва — Сельскохозяйственная академия. Они вычислены по формуле

$$\Delta c = \frac{(m_1 - m_2) + c_1(\sigma_1 - \sigma_2)}{\sigma_2}. \qquad (24$$

Значения m и σ для этой станции получены авторами в [7].

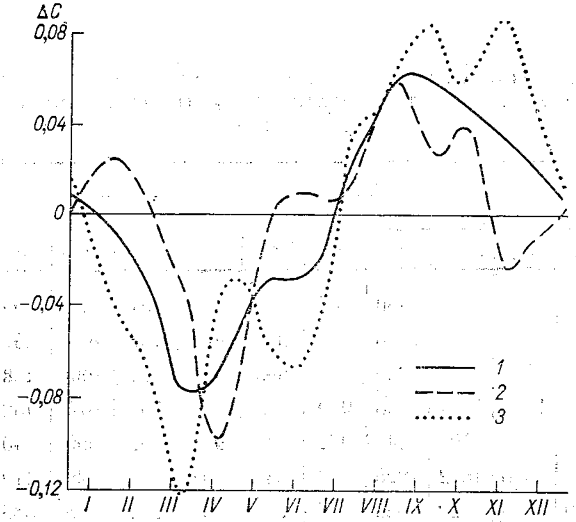


Рис. 2. Мера нестационарности Δc (сутки$^{-1}$) средней суточной температуры воздуха, ГМС Москва — с.-х. академия.

1) $c_1=0$; 2) $c_1=4$; 3) $c_1=-4$.

Из рис. 2 видно, что в этом случае нестационарность в течение года практически укладывается в пределы $|\Delta c| < 0{,}1$, что делает возможным использование приведенных выше упрощенных формул. В случае рассмотрения этих рядов не в суточной, а в какой-либо другой дискретности, мера нестационарности Δc, естественно, должна быть увеличена (например, для декадной дискретности — в 10 раз). В таких случаях необходим более тщательный учет нестационарности ряда.

СПИСОК ЛИТЕРАТУРЫ

1. Свешников А. А. Прикладные методы теории случайных функций. М., «Наука», 1968, 463 с.
2. Смирнов Н. В., Большев Л. Н. Таблицы для вычисления функции двумерного нормального распределения. М., Изд. АН СССР, 1962. 204 с.

Каган Р. Л., Федорченко Е. И. О расчете статистических характеристик выбросов случайной функции. — «Труды ГГО», 1970, вып. 268, с. 146—172.

Крамер Г., Лидбеттер М. Стационарные случайные процессы. М., «Мир», 1969. 398 с.

Кендалл М., Стьюарт А. Теория распределений. М., «Наука», 1966, 587 с.

Анискин Л. В., Персии С. М. О погрешности измерения экстремальных значений случайного процеса.—«Труды ГГО», 1972, вып. 292, с. 12—25.

Каган Р. Л., Федорченко Е. И. О восстановлении годового хода моментов метеорологических рядов. См. наст. сборник.

Р. Л. КАГАН, Е. И. ФЕДОРЧЕНКО

О ВЛИЯНИИ ДИСКРЕТНОСТИ ИЗМЕРЕНИЙ НА ТОЧНОСТЬ ОПРЕДЕЛЕНИЯ ЧИСЛА ВЫБРОСОВ СЛУЧАЙНОГО ПРОЦЕССА

1. Введение

Одной из важных статистических характеристик случайного процесса является среднее число выбросов его через тот или иной уровень на интервале заданной длины. Применительно к метеорологическим рядам эта характеристика позволяет судить о повторяемости и о продолжительности периодов, в течение которых интересующий нас метеорологический элемент превышает какие-либо существенные для практики пороговые значения. Практически число выбросов метеорологических процессов приходится в большинстве случаев определять путем обработки рядов наблюдений, выполненных в дискретные моменты времени. Необходимо уметь оценивать влияние дискретности отсчетов на статистические характеристики выбросов временных рядов.

Некоторые оценки влияния дискретности отсчетов на точность определения среднего числа выбросов процесса были выполнены авторами в [1]. При этом считалось, что дискретизация приводит к тому, что вместо числа выбросов процесса определяется число выбросов соответствующей последовательности. Поскольку при определении выбросов случайной последовательности предполагается, что в промежутке времени между измерениями значения процесса изменяются монотонно, это приводит к занижению числа выбросов, тем большему, чем больше интервал времени между измерениями. С этой точки зрения может создаться впечатление, что для правильной оценки числа выбросов выгодно производить возможно более частые измерения.

В действительности дело обстоит так лишь в идеализированном случае отсутствия ошибок наблюдений. В реальных условиях в процессе обработки фиксируются как действительные выбросы измеряемой величины, так и фиктивные «выбросы» за счет ошибок измерений. Удельный вес последних тем больше ем чаще производятся измерения. В пределе, при очень частых змерениях, все зафиксированные выбросы могут оказаться фик-ивными[1]. Завышение числа выбросов за счет ошибок измерения принципе может быть уменьшено путем надлежащего сглажи-ания данных наблюдений в процессе измерений или обработки. обычных условиях, когда это не предусмотрено, единственным утем уменьшения числа фиктивных выбросов оказывается ди-кретизация отсчетов.

При увеличении интервала между измерениями число фиктив-ых выбросов за счет ошибок наблюдения уменьшается, но уве-ичивается число пропущенных реальных выбросов. Очевидно, ожно ставить вопрос о выборе оптимальной в некотором смыс-е дискретности отсчетов аналогично тому, как это делается при ценке точности конечно-разностного представления производ-ых (см., например, [3]).

В настоящей статье рассматривается постановка этой задачи получаются некоторые численные оценки. Наряду с качествен-ым анализом влияния дискретизации отсчетов показаны пути олучеиия простых количественных критериев выбора оптималь-ой дискретности измерений. Мы ограничимся оценками для нор-ального стационарного процесса, для которого расчеты выпол-яются наиболее просто. Ясно, однако, что аналогичные резуль-аты могут быть получены и для случайных процессов другого ида.

2. Постановка задачи

Пусть производятся измерения нормального стационарного роцесса $f(t)$, для которого полагаем известными математиче-кое ожидание m, среднее квадратическое отклонение σ_f и кор-еляционную функцию $r_f(\tau)$. Указанные характеристики полно-тью определяют статистические свойства процесса. В частности, спользуя их, можно определить среднее число выбросов через ровень C на интервале времени T по формуле

$$\overline{N}_c(0) = \frac{T}{2\pi} \sqrt{-r_f''(0)}\, e^{-\frac{c^2}{2}}, \tag{1}$$

де $c = \dfrac{C - m}{\sigma_f}$ — относительное отклонение уровня C от матема-

[1] К аналогичному с формальной точки зрения завышению числа выбросов южет приводить учет микромасштабных пульсаций измеряемой величины, воз-ействие которых в рамках рассматриваемой задачи (например, при расчете рочности строительных конструкций) является пренебрежимо малым. Так, работе В. Н. Иванова и Р. Л. Стратоновича [2] показано, что учет высоко-астотных пульсаций приводит к завышению числа выбросов скорости ветра на несколько порядков.

тического ожидания, а $r''_f(0)$ представляет собой вторую про-изводную[1] корреляционной функции $r_f(\tau)$ при $\tau=0$.

Поскольку процесс считается стационарным, среднее число выбросов вверх и вниз через каждый уровень совпадает. Для определенности можно считать, что речь идет о числе выбросов вверх. Для нормального процесса знак относительного отклонения не играет роли. Мы ограничимся рассмотрением случая $C>m$ $(c>0)$ и будем говорить о высоких и низких уровнях, имея в виду абсолютные значения этого отклонения. Полученные результаты будут справедливы и для уровней $C<m$. Так уровням C и $C_1=2m-C$ соответствуют относительные отклонения c и $c_1=-c$, так что статистические характеристики числа выбросов для них совпадают.

В ходе измерений вместо истинных значений f получаются величины

$$\tilde{f}(t)=f(t)+\delta(t). \tag{2}$$

Полагаем, что ошибка измерений $\delta(t)$ имеет нормальное распределение с нулевым средним значением и не коррелирует с величиной $f(t)$. Величина $\tilde{f}(t)$, очевидно, также имеет нормальное распределение, и среднее число выбросов ее может быть определено по формуле (1) после замены в ней истинных статистических характеристик процесса характеристиками этой величины.

Рассмотрим для простоты случай, когда корреляция ошибки затухает со временем по тому же закону, что и корреляция измеряемой величины, а различие состоит лишь в масштабах корреляции. В этом случае

$$r_\delta(\tau)=r_f(k\tau), \tag{3}$$

где $k=\frac{T_f}{T_\delta}$, T_f и T_δ — масштабы корреляции величин f и δ соответственно.

Нетрудно видеть, что при сделанных предположениях

$$\sigma^2_{\tilde{f}}=(1+\eta^2)\,\sigma^2_f, \tag{4}$$

$$r_{\tilde{f}}(\tau)=\frac{r_f(\tau)+\eta^2 r_f(k\tau)}{1+\eta^2}, \tag{5}$$

где $\eta^2=\frac{\sigma^2_\delta}{\sigma^2_f}$ — мера ошибок наблюдения.

---

[1] Предполагается, что эта производная существует, так что процесс $f(t)$ является дифференцируемым.

Из (5) следует, что

$$r''_{\tilde f}(0) = \frac{1+\eta^2 k^2}{1+\eta^2}\, r''_f(0). \tag{6}$$

Масштаб корреляции ошибок измерения T_δ естественно счи-
ть меньшим масштаба корреляции измеряемой величины, так
о $k>1$ и $r''_{\tilde f}(0) > r''_f(0)$. Обращаясь к формуле (1), видим, что
о должно приводить к увеличению числа выбросов. В предель-
ом случае некоррелированных ошибок измерений $k \to \infty$ и чис-
) выбросов безгранично возрастает.

Другим фактором, также приводящим к увеличению числа
ыбросов, является завышение дисперсии и, следовательно, за-
іжение входящего в формулу (1) относительного уровня c:

$$\tilde c = \frac{C-m}{\sigma_{\tilde f}} = \frac{c}{\sqrt{1+\eta^2}}. \tag{7}$$

Таблица 1

Зависимость увеличения (M_1) среднего числа выбросов случайного эцесса через уровень $C=m$ ($c=0$) от ошибок измерений и масштаба их корреляции

η^2	k									
	1	2	5	10	15	20	30	50	100	∞
)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
,0001	1,00	1,00	1,00	1,00	1,01	1,02	1,04	1,11	1,41	∞
,001	1,00	1,00	1,01	1,05	1,11	1,16	1,38	1,58	3,32	∞
,005	1,00	1,01	1,06	1,22	1,45	1,73	2,34	3,63	7,12	∞
,01	1,00	1,02	1,11	1,41	1,81	2,22	3,15	5,07	10,0	∞
,02	1,00	1,04	1,21	1,71	2,32	2,97	4,32	7,02	14,0	∞
,05	1,00	1,07	1,46	2,39	3,41	4,47	6,55	10,9	21,8	∞
,1	1,00	1,13	1,78	3,16	4,62	6,11	9,08	15,1	30,1	∞

Подстановка в формулу (1) выражений (6) и (7) дает

$$\overline{\tilde N}_c(0) = \overline{N}_c(0) \sqrt{\frac{1+\eta^2 k^2}{1+\eta^2}}\, e^{\frac{1}{2} c^2 \frac{\eta^2}{1+\eta^2}} \tag{8}$$

ли

$$\overline{\tilde N}_c(0) = \overline{N}_c(0)\, M_1(k, \eta)\, M_2(c, \eta), \tag{9}$$

е множители $M_1(k, \eta) = \sqrt{\frac{1+k^2\eta^2}{1+\eta^2}}$ и $M_2(c, \eta) = e^{\frac{1}{2} c^2 \frac{\eta^2}{1+\eta^2}}$ харак-
ризуют влияние первого и второго факторов. Представление об

их вкладе при некоторых значениях, входящих в них параметро дают соответственно табл. 1 и 2.

Из этих таблиц видно, что при небольших мерах ошибок и мерения ($\eta^2<0{,}01$) завышение числа выбросов процесса практ чески не зависит от высоты уровня C. Оно сколько-нибудь заме но лишь в тех случаях, когда масштаб корреляции ошибок и мерения значительно (по крайней мере на порядок) мены масштаба корреляции измеряемой величины. При больших зн чениях меры ошибок высота уровня сказывается лишь при зн чительном удалении его от математического ожидания ($c>1{,}5$

Таблица

Зависимость увеличения (M_2) среднего числа выбросов процесса при масштабе корреляции ошибок $T_\delta = T_f$ от относительного уровня $c=\frac{C-m}{\sigma}$ и от меры ошибок измерения процесса η^2

η^2	c								
	0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0
0	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,0001	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,001	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,01	1,01
0,005	1,00	1,00	1,00	1,01	1,01	1,02	1,02	1,03	1,04
0,01	1,00	1,00	1,00	1,01	1,02	1,03	1,04	1,06	1,08
0,02	1,00	1,00	1,01	1,02	1,04	1,06	1,09	1,13	1,17
0,05	1,00	1,00	1,02	1,05	1,10	1,16	1,24	1,34	1,46
0,1	1,00	1,01	1,05	1,11	1,20	1,32	1,49	1,74	2,08

Однако и в этом случае влияние мелкомасштабных шумов ок зывается, как правило, более существенным, чем влияние сниж ния относительного уровня, характеризуемое множителем M_2.

В случае, когда измерения процесса производятся через и тервалы времени Δ, судить о числе выбросов процесса оказыв ется возможным лишь на основании числа выбросов соответс вующей последовательности. В частности, среднее число выбр сов в течение периода времени T может быть (см., наприме [4]) определено по формуле

$$\tilde{\tilde{N}}_c(\Delta)=\frac{T}{\Delta\sqrt{2\pi}}\int\limits_{-\infty}^{\tilde{c}} e^{-\frac{x^2}{2}}\left[1-\Phi\left(\frac{\tilde{c}-r_1x}{\sqrt{1-r_1^2}}\right)\right]dx, \qquad (1$$

где $r_1=r_{\tilde{f}}(\Delta)$ — коэффициент корреляции между данными изм рения величины f, полученными через интервал времени Δ

$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$ — интеграл вероятностей.

В отличие от формул (8) и (9), справедливых при любой ррреляционной функции величины f, соотношение между истинным средним числом выбросов и средним числом выбросов $_c(\Delta)$, фиксируемых при дискретных отсчетах через интервал Δ, соответствии с формулами (10) и (5), требует задания значений корреляционной функции $r_f(\Delta)$ и $r_f(k\Delta)$ и потому зависит от ида этой функции.

Мы ограничимся рассмотрением случайных процессов, корелляционные функции которых имеют вид

$$r_f(\tau)=\left(1+\frac{\tau}{T_f}\right)e^{-\tau/T_f} \tag{11}$$

ли

$$r_f(\tau)=e^{-\frac{1}{2}(\tau/T_f)^2}. \tag{12}$$

В обоих случаях $r''_f(0)=-\frac{1}{T_f^2}$, так что, согласно (1), олучаем для периода времени T истинное значение среднего исла выбросов процесса

$$\overline{N}_c(0)=\frac{1}{2\pi}\frac{T}{T_f}e^{-\frac{c^2}{2}}. \tag{13}$$

С другой стороны, для этого же периода последовательность змерений, интервал времени между членами которой составляг Δ, будет в среднем иметь количество выбросов, определяемое ормулой (10).

3. Оценка систематических ошибок в определении числа выбросов

Приведем некоторые результаты сопоставления величин $\overline{N}_c(0)$ $\overline{\overline{N}}_c(\Delta)$, полученные таким образом для различных дискретнотей измерений и при различном уровне ошибок измерения η_2.

Таблицы 3 и 4 иллюстрируют влияние дискретности отсчетов ля случайных процессов, корреляционные функции которых писываются соответственно формулами (11) и (12). В этих табицах приводятся значения отношения $\nu_c(\Delta)=\overline{\overline{N}}_c(\Delta)/\overline{N}_c(0)$ для ровня $c=0$ (т. е. $C=m$) при фиксированном значении $\eta^2=0{,}01$.

Из табл. 3 и 4 видно, что при малых интервалах между изерениями имеет место систематическое завышение числа выброов последовательности по сравнению с истинным числом выброов для процесса. Это завышение тем больше, чем больше пааметр k, т. е. чем меньше масштаб корреляции ошибок измеений. При больших интервалах Δ полученное в результате измерений число выбросов меньше истинного числа выбросов пр цесса. Для каждого значения k существует такой интервал ди кретизации Δ_0, для которого среднее число выбросов, получен ное путем обработки данных измерений, совпадает со средни числом выбросов процесса. Будем называть этот интервал опти мальным интервалом дискретизации. Значения этого интервал для разных k приводятся в предпоследних строках табл. 3 и Сравнение их показывает, что оптимальный интервал дискрет

Таблица

Отношение среднего числа выбросов последовательности за уровень $c=0$ к истинному среднему числу выбросов процесса.

$$(r(\tau)=(1+\tau/T_f)\,e^{-\frac{\tau}{T_f}};\quad \eta^2=0{,}01)$$

$\frac{\Delta}{T_f}$	$r(\Delta)$	k					
		1	2	5	10	20	∞
0	1,000	1,00	1,02	1,11	1,41	2,22	∞
0,1	0,995	0,96	0,97	1,05	1,20	1,44	1,71
0,2	0,982	0,94	0,95	1,00	1,08	1,15	1,17
0,3	0,963	0,91	0,92	0,96	1,00	1,02	1,02
0,4	0,938	0,88	0,89	0,92	0,94	0,95	0,95
0,5	0,910	0,86	0,86	0,88	0,90	0,90	0,90
0,6	0,878	0,83	0,84	0,86	0,86	0,86	0,86
0,7	0,844	0,81	0,82	0,83	0,83	0,83	0,83
0,8	0,809	0,79	0,79	0,80	0,80	0,80	0,80
0,9	0,772	0,76	0,77	0,78	0,78	0,78	0,78
1,0	0,736	0,74	0,75	0,75	0,76	0,76	0,76
$(\Delta_0/T_f)_{c=0}$		0	0,03	0,21	0,30	0,32	0,33
$(\Delta_0/T_f)_{c=3}$		0,12	0,16	0,25	0,31	0,33	0,33

зации зависит не только от масштаба корреляции ошибок изме рения, но и от корреляционной функции измеряемых величин.

Приведенные данные подтверждают, что в реальных условия при наличии ошибок измерений очень большая частота измерени не только не обязательна, но и не желательна, поскольку она при водит к систематическим ошибкам при определении числа выбро сов процесса. Эти систематические ошибки тем больше, чем боль ше ошибки измерения; соответственно должен возрастать и опти мальный интервал Δ_0. Представление о зависимости оптимально го интервала от меры ошибок наблюдения дают рис. 1 и 2, отно сящиеся к процессам с корреляционными функциями (11) и (12) соответственно.

Приведенные выше данные характеризуют число выбросо через уровень $c=0$, т. е. через математическое ожидание (при менительно к метеорологическим рядам можно говорить о числе ревышений нормы). Для других уровней упомянутые закономерости также имеют место, хотя и отмечаются некоторые колиественные различия. Основное различие состоит в том, что для ысоких уровней при малых интервалах Δ между измерениями вышение числа выбросов оказывается еще бóльшим, чем для $=0$, а при больших Δ пропускается большее число выбросов, чем ля $c=0$. Это связано с тем, что, как уже упоминалось (см. форулу (8)), для непрерывного процесса, а следовательно, и при змерениях с достаточно малым интервалом Δ наличие ошибок змерения сказывается тем больше, чем выше уровень. С другой ороны, чем выше уровень, тем меньше средняя продолжитель-

Таблица 4

Отношение среднего числа выбросов последовательности за уровень $c=0$ к истинному среднему числу выбросов процесса.

$$\left(r(\tau)=e^{-\frac{1}{2}(\tau/T_f)^2};\quad \eta^2=0{,}01\right)$$

$\frac{\Delta}{T_f}$	$r(\Delta)$	k					
		1	2	5	10	20	∞
0	1,000	1,00	1,02	1,11	1,41	2,22	∞
0,1	0,995	1,00	1,01	1,10	1,33	1,64	1,72
0,2	0,980	1,00	1,01	1,09	1,19	1,22	1,22
0,3	0,956	0,99	1,01	1,06	1,09	1,10	1,10
0,4	0,923	0,99	1,00	1,04	1,04	1,04	1,04
0,5	0,883	0,98	0,99	1,01	1,02	1,02	1,02
0,6	0,835	0,97	0,98	0,99	0,99	0,99	0,99
0,7	0,783	0,96	0,97	0,98	0,98	0,98	0,98
0,8	0,726	0,95	0,96	0,96	0,96	0,96	0,96
0,9	0,667	0,93	0,94	0,94	0,94	0,94	0,94
1,0	0,607	0,92	0,92	0,93	0,93	0,93	0,93
$(\Delta_0/T_f)_{c=0}$		0	0,38	0,57	0,57	0,57	0,57
$(\Delta_0/T_f)_{c=3}$		0,32	0,36	0,43	0,44	0,44	0,44

ость выброса, что приводит к увеличению вероятности пропуска ыброса процесса при больших интервалах времени между изерениями.

В результате оптимальный интервал дискретизации для выких уровней может существенно отличаться от оптимального нтервала для уровня $c=0$. В качестве примера в последних строах табл. 3 и 4 приведены значения Δ_0/T_f для уровня $c=3$.

При организации наблюдений, разумеется, необходимо устаавливать какую-то единую дискретность отсчетов. С практичеой точки зрения представляется наиболее целесообразным исодить из оптимальной дискретности для уровня $c=0$ при задан-

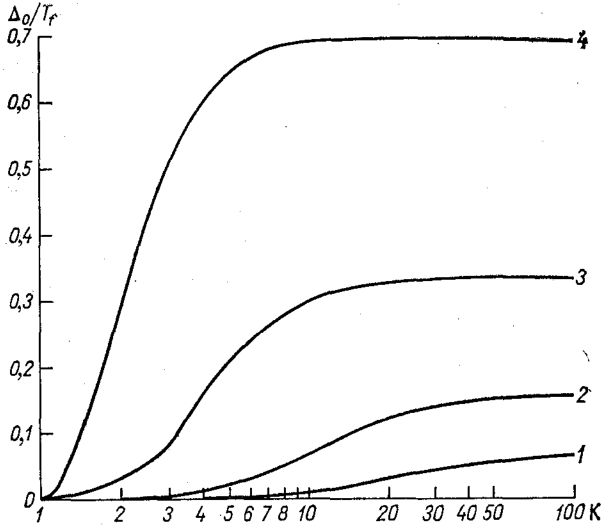


Рис. 1. Зависимость оптимального интервала между измерениями от меры ошибок и масштаба их корреляции. $c=0$, корреляционная функция (11).
1) $\eta^2=0{,}0001$, *2*) $\eta^2=0{,}001$, *3*) $\eta^2=0{,}01$, *4*) $\eta^2=0{,}1$.

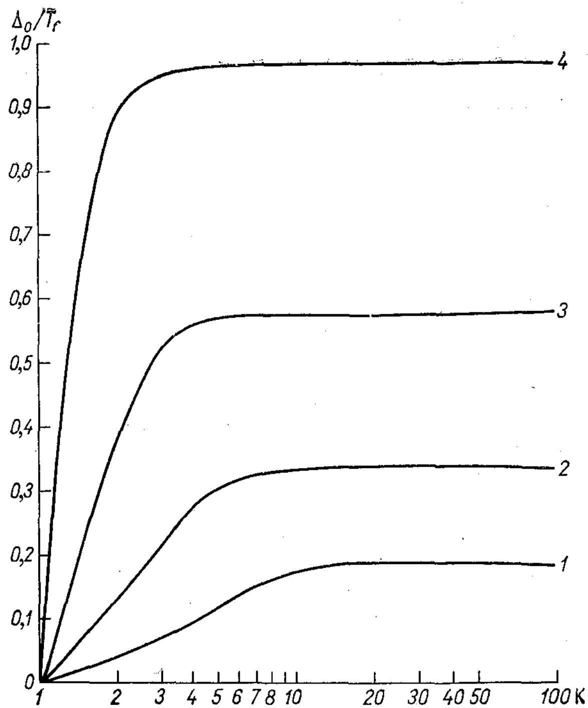


Рис. 2. Зависимость оптимального интервала между измерениями от меры ошибок и масштаба их корреляции. $c=0$, корреляционная функция (12).
Усл. обозначения см. рис. 1.

й точности измерений. Как показано в [1], во многих случаях казывается целесообразным непосредственный расчет числа ыбросов производить лишь для уровней, достаточно близких норме, а среднее число выбросов для высоких уровней полуть расчетным путем, используя эмпирические данные, полученые для низких уровней.

Как было показано выше, влияние дискретности отсчетов ильно зависит от степени коррелированности ошибок измереия, которая в свою очередь, определяется используемой системой мерений. Особый интерес представляет случай некоррели-ованных ошибок измерений ($k=\infty$). При определенных услоиях такими ошибками характеризуются данные стандартных этеорологических наблюдений (например, за температурой и давлением воздуха) после первичной обработки, позволяющей исключить систематические ошибки наблюдений.

Для случая некоррелируемых ошибок также могут быть почены таблицы типа табл. 3 и 4 для различных значений меры шибок наблюдений. Однако, учитывая сильную зависимость реельтатов от характеристик временно́й структуры, представляся более полезным нахождение общих формул, которые позволяя хотя бы приближенно получать нужные оценки для различых условий. Такие приближенные формулы удается получить, итывая, что при метеорологических измерениях наибольший нтерес представляет обработка сравнительно точных измереий, для которых $\eta^2 \leqslant 0{,}01$, а связность метеорологических рядов является довольно высокой.

Как показано в [5], при достаточно высоких коэффициентах рреляции между смежными членами стационарной нормальной оследовательности, можно вместо формулы (10) использовать риближенную формулу

$$\bar{\bar{N}}_c(\Delta) = \frac{1}{2\pi}\frac{T}{\Delta} e^{-\frac{c^2}{2}} \sqrt{2(1-\tilde{r}_1)} \left\{1 - \frac{1}{12}(c^2-1)(1-\tilde{r}_1)\right\}. \qquad (14)$$

Представим ее в несколько видоизмененной форме

$$\bar{\bar{N}}_c(\Delta) = \frac{1}{2\pi}\frac{T}{\Delta} e^{-\frac{c^2}{2}} \sqrt{2(1-\tilde{r}_1) - \frac{c^2-1}{3}(1-\tilde{r}_1)^2}. \qquad (15)$$

Из сравнения (15) и (13) следует, что

$$\nu_c(\Delta) = \frac{\bar{\bar{N}}_c(\Delta)}{\bar{N}_c(0)} = \frac{T_f}{\Delta}\sqrt{2(1-\tilde{r}_1) - \frac{c^2-1}{3}(1-\tilde{r}_1)^2}. \qquad (16)$$

Будем представлять корреляционную функцию процесса f при алых значениях сдвига времени τ в виде разложения

$$r_f(\tau) = 1 - \frac{1}{2}\left(\frac{\tau}{T_f}\right)^2 + a\left(\frac{\tau}{T_f}\right)^3 + b\left(\frac{\tau}{T_f}\right)^4 + \ldots \qquad (17)$$

Линейный член в (17) отсутствует, поскольку мы полага процесс дифференцируемым. В случае, если процесс дифференц руем дважды, в (17) отсутствует также кубический член, т. $a=0$. Так обстоит дело, например, в случае корреляционной фун ции (12).

При некоррелированных ошибках наблюдений, из (5) след ет, что

$$\tilde{r}_1 = \frac{r_f(\Delta)}{1+\eta^2}. \tag{18}$$

Учитывая (17) и (18) и полагая меру ошибок наблюдений малой, получаем приближенное равенство

$$1-\tilde{r}_1 = \eta^2 + \frac{1}{2}\left(\frac{\Delta}{T_f}\right)^2 - a\left(\frac{\Delta}{T_f}\right)^3 - b\left(\frac{\Delta}{T_f}\right)^4, \tag{19}$$

подстановка которого в (16) дает

$$\nu_c(\Delta) = \sqrt{1-2\left(\frac{T_f}{\Delta}\right)^2\left\{a\left(\frac{\Delta}{T_f}\right)^3+\left(b+\frac{c^2-1}{24}\right)\left(\frac{\Delta}{T_f}\right)^4-\eta^2\right\}}. \tag{20}$$

Оптимальный выбор дискретности Δ соответствует обращени в нуль фигурной скобки в подкоренном выражении, что прив дит к условию

$$a\left(\frac{\Delta_0}{T_f}\right)^3+\left(b+\frac{c^2-1}{24}\right)\left(\frac{\Delta_0}{T_f}\right)^4=\eta^2. \tag{21}$$

Для процессов, имеющих лишь первую производную, вторы членом левой части уравнения (21) можно пренебречь по сравн нию с первым. Это означает, что для таких процессов при неко релированных ошибках измерений оптимальный интервал вр мени между измерениями мало зависит от уровня c и при задан ной мере ошибок η^2 может приближенно определяться по фо муле

$$\Delta_0 = T_f\sqrt[3]{\eta^2/a}. \tag{22}$$

В частности, для процессов с корреляционной функцией (11 $a={}^1/_3$, так что

$$\Delta_0 = T_f\sqrt[3]{3\eta^2}. \tag{23}$$

Для процессов, дифференцируемых дважды, $a=0$ и для опре деления оптимальной дискретности получаем формулу

$$\Delta_0 = T_f\sqrt[4]{\frac{\eta^2}{b+\frac{c^2-1}{24}}}. \tag{24}$$

В эту формулу существенно входит высота уровня, для которого определяется число выбросов. Оптимальный интервал оказывается тем меньше, чем больше уровень c.

Например, для процессов с корреляционной функцией (12) $a=0$, и $b=1/8$, так что

$$\Delta_0=T_f\sqrt[4]{\frac{12\,\eta^2}{1+\frac{c^2}{2}}}. \tag{25}$$

Из данных табл. 5, в которой приводятся оценки оптимальных интервалов дискретизации, полученные с помощью приближенных формул (23) и (25) и точные значения оптимальных интервалов, видно, что при $c=0$ формулы (23) и (25) дают хорошее согласование с фактическими данными во всем рассмотренном диапазоне изменений η^2. Для высоких уровней c удовлетворительные результаты получаются лишь для $\eta^2\leqslant 0{,}01$. Очевидно, что эти ограничения относятся также и к возможности использования приближенной формулы (20) для оценки систематической погрешности эмпирически определенного числа выбросов.

Таблица 5

Расчетные (Δ_p) и фактические (Δ_0) значения оптимального интервала при некоррелированных ошибках измерений

η^2	Корреляционная функция (11)			Корреляционная функция (12)			
	$\frac{\Delta_p}{T_f}$	Δ_0/T_f		Δ_p/T_f		Δ_0/T_f	
		$c=0$	$c=3$	$c=0$	$c=3$	$c=0$	$c=3$
0,0001	0,07	0,07	0,07	0,19	0,12	0,19	0,13
0,001	0,14	0,14	0,15	0,33	0,23	0,33	0,23
0,005	0,25	0,26	0,26	0,50	0,32	0,49	0,36
0,01	0,31	0,33	0,33	0,59	0,38	0,58	0,44
0,02	0,39	0,41	0,44	0,70	0,46	0,67	0,56
0,05	0,53	0,55	0,67	0,88	0,57	0,83	0,80
0,1	0,67	0,69	0,93	1,05	0,68	0,96	$>1{,}00$

В качестве примера оценим целесообразную дискретность данных для расчета числа выбросов температуры и давления воздуха на станциях Северо-Запада Европейской территории СССР. При этом ограничимся зимним периодом, для которого изменения не только давления, но и температуры воздуха могут приближенно описываться как нормальный стационарный процесс.

По данным работы [6], временна́я корреляция давления воздуха, приведенного к уровню моря, для этого района в январе описывается формулой (11). При этом $T_f=30$ ч, а $\eta^2=0{,}001$. При дисперсии давления порядка 200 мбар² это соответствует средней квадратической ошибке 0,45 мбар, что значительно превышает инструментальную погрешность наблюдений, которая обычно (см. например, [7]) полагается равной около 0,2 мбар. Это превышение, по-видимому, связано с дополнительными погрешностями возникающими в ходе приведения давления к уровню моря.

По данным, полученным одним из авторов [8], временна́я корреляция температуры воздуха в январе в Ленинграде для сдвигов по времени τ до 6 ч также удовлетворительно описывается формулой (11) при $T_f=16$ ч. Полагая среднюю квадратическую погрешность стандартных измерений температуры воздуха в психрометрической будке равной 0,2°С и учитывая, что для января дисперсия температуры составляет около 40 (°С)², получаем снова оценку меры ошибок наблюдений[1] $\eta^2=0{,}001$.

Как уже указывалось, для корреляционной функции вида (11) в формуле (20) можно в первом приближении пренебречь членом в фигурных скобках, содержащим шаг измерения Δ в четвертой степени, а следовательно, и зависимостью от уровня, выбросы через который нас интересуют. В результате формула (20) принимает вид

$$\nu=\sqrt{1+2\eta^2\left(\frac{T_f}{\Delta}\right)^2-\frac{2}{3}\frac{\Delta}{T_f}}. \tag{26}$$

Подстановка в нее конкретных значений параметров статистической структуры дает для давления на уровне моря

$$\nu_p\approx\sqrt{1+\frac{1{,}8}{\Delta^2}-0{,}022\,\Delta}, \tag{26а}$$

а для температуры воздуха

$$\nu_\Theta\approx\sqrt{1+\frac{0{,}512}{\Delta^2}-0{,}042\,\Delta}. \tag{26б}$$

Интервал времени Δ между измерениями в этих формулах отсчитывается в часах.

Зависимость отношения ν от Δ, полученная таким образом, представлена в табл. 6.

Из табл. 6 следует, что при выборе $\Delta=1$ ч, т. е. при использовании данных ежечасных наблюдений, число выбросов, полученных при обработке фактических данных, будет сильно завы-

[1] Заметим, что для других сезонов дисперсия температуры существенно меньше, соответственно увеличивается мера ошибок наблюдений. Меняется также и характер временно́й корреляции. Выполнение оценок для этих сезонов представляется, однако, нецелесообразным прежде всего из-за существенной нестационарности температурных рядов для этих сезонов, делающей невозможным непосредственное применение полученных выше формул.

щено (на 21% для температуры, а для давления даже на 67%). Очевидно, что полученные таким образом результаты требуют последующей корректировки. Возможно, разумеется, вводить в процессе расчета числа выбросов какие-то процедуры фильтрации данных, например, путем неучета выбросов малой длительности, которые при большой частоте наблюдений возникают главным образом за счет ошибок измерений. Однако разработка методики подобной фильтрации требует проведения специальных исследований.

Таблица 6

Отношение среднего числа выбросов, полученного по данным наблюдений, к истинному среднему числу выбросов. Январь, Северо-Запад ЕТС

часы	1	2	3	4	5	6
Давление	1,67	1,18	1,06	1,01	0,98	0,95
Температура воздуха	1,22	1,02	0,97	0,93	0,90	0,87

Наиболее простым путем уменьшения систематических ошибок является в настоящее время использование данных наблюдений, полученных через бо́льшие интервалы времени. Оптимальные с этой точки зрения интервалы, в соответствии с формулой (23), составляют 4,3 ч для определения выбросов давления на уровне моря и 2,3 ч для выбросов температуры воздуха.

Конечно, задание таких дробных интервалов практически неприемлемо, однако из табл. 6 видно, что использование трехчасового интервала дает гораздо более точные результаты, чем использование ежечасных данных. Это тем более существенно, что в настоящее время наблюдения на основной сети метеорологических станций СССР выполняются 8 раз в сутки, а данные ежечасных наблюдений, которые обычно рекомендуются для обработки такого рода, имеются лишь для ограниченного числа станций.

4. Оценка случайных ошибок числа выбросов за счет дискретности отсчетов

Выбор оптимальной дискретности данных в соответствии с описанной методикой позволяет исключить систематические ошибки в определении среднего числа выбросов. Следует, однако, иметь в виду, что он не обеспечивает получения правильных значений других статистических характеристик случайных процессов. В частности, при дискретизации отсчетов, в том числе и при оптимальной в нашем смысле дискретности, имеет место искажение распределения выбросов по длительностям. Оценка этих искажений и введение соответствующих поправок являются затруднительными и для случая отсутствия ошибок наблюдений. При наличии этих ошибок соответствующие оценки еще более осложняются.

Сложным является и вопрос об оценке случайных погрешносте определения числа выбросов, возникающих за счет дискретиза ции. Между тем представление хотя бы о порядке этих погрешно стей иметь очень важно, поскольку от него зависит точность по лученных в ходе обработки данных о среднем числе выбросов равно как и объем подлежащих обработке данных.

Хотя аналитический расчет этих характеристик в принцип возможен, практически он требует преодоления очень больши вычислительных трудностей. Представляется поэтому, что дл этой цели наиболее подходящим является использование метод статистического моделирования. Этот метод в настоящее время до вольно хорошо разработан. Результаты применения его к расче ту некоторых характеристик выбросов приводятся в работе [9] Ниже мы рассмотрим пример, иллюстрирующий возможности ме тода статистического моделирования применительно к оценке слу чайных погрешностей числа выбросов процесса за счет дискрет ности измерений.

Авторы получали на ЭВМ М-220 реализации нормальных ста ционарных случайных последовательностей с нулевым математи ческим ожиданием, единичной дисперсией и с корреляционной функцией (11). Для этой цели использовалась приведенная в кни ге В. В. Быкова [10] рекуррентная формула

$$f_i = a_0 x_i + a_1 x_{i-1} - b_1 f_{i-1} - b_2 f_{i-2}. \qquad (27)$$

Здесь x_i — псевдослучайные числа, нормально распределенны с нулевым математическим ожиданием и единичной дисперсией а коэффициенты a_i и b_i определяются формулами[1]:

$$a_0 = \sqrt{\frac{A_1 \pm \sqrt{A_1^2 - 4A_0}}{2}}; \quad a_1 = \frac{A_0}{a_0};$$

$$b_1 = -2\rho; \quad b_2 = \rho^2, \qquad (28)$$

где Δ — интервал последовательности;

$$\rho = e^{-\Delta/T_f};$$

$$A_0 = \rho^3\left(1 + \frac{\Delta}{T_f}\right) - \rho\left(1 - \frac{\Delta}{T_f}\right);$$

$$A_1 = 1 - 4\rho^2 \frac{\Delta}{T_f} - \rho^4.$$

Чтобы избежать искажения начала реализаций, первые два члена каждой из них находились по формулам:

$$f_1 = x_1;$$

$$f_2 = a_0 x_2 + r(\Delta) x_1 + \sqrt{1 - a_0^2 - (r(\Delta))^2}\, x_0. \qquad (29)$$

[1] В приводимых в [10] формулах для нахождения коэффициентов a_i и b_i содержится несколько опечаток. Здесь они устранены.

Легко убедиться путем последовательного расчета коэффици-нтов корреляции между разными членами f_i реализаций, что для оделируемых в соответствии с этим алгоритмом последователь-остей они в точности совпадают с соответствующими значения-и корреляционной функции (11). Возможные фактические раз-ичия связаны лишь с несовершенством используемых стандарт-ых программ выдачи псевдослучайных чисел, и, разумеется, ограниченным объемом фактически моделируемых объемов дан-ых.

Оценка ошибок в определении числа выбросов за счет дис-ретности отсчетов производилась следующим образом.

а. Моделировались «эталонные» реализации последовательно-ти, соответствующие столь частому измерению процесса, кото-ое практически исключало бы пропуски выбросов процесса. актически в качестве эталонной дискретности принимался в боль-инстве случаев интервал $\Delta_э=0{,}01T_f$. Для данной корреляцион-ой функции она соответствует коэффициенту корреляции между межными членами последовательности $r(\Delta)=0{,}99995$, что значи-ельно превышает корреляцию между последовательными отсче-ами любых реально наблюдаемых метеорологических элементов.

б. Определялось число выбросов $N_э$ для каждой эталонной еализации. Систематическое занижение числа выбросов для вы-ранной дискретности при отсутствии ошибок составляет, соглас-о формуле (26), около 0,3%. На вопросе о случайной погрешно-ти числа выбросов эталонной последовательности мы остановимся иже.

в. Производилось прореживание эталонных реализаций путем тбора каждого второго, четвертого и т. д. их члена.

г. На полученные ранее величины f_i накладывались «ошибки змерений» δ_i, которые могут моделироваться надлежащим обра-ом. В данном случае для простоты использовались псевдослу-айные числа $\delta_i=x_i\eta$, что соответствует некоррелированным ошиб-ам измерения при мере ошибок η^2.

д. Для каждой реализации производился расчет числа выбро-ов N полученной последовательности «измерений» $\tilde{f}_i=f_i+\delta_i$ сравнение его с числом выбросов эталонной последовательности. азность $\Delta N=N-N_э$ и считается ошибкой определения числа вы-осов процесса.

е. Путем осреднения по всем реализациям находились как ха-актеристики (средние арифметические и дисперсии) числа вы-осов N последовательностей при различной длине реализации различной дискретности их, так и соответствующие характери-ики ошибок в определении числа выбросов.

В ходе расчетов в большинстве случаев моделировалось по 0 независимых реализаций процесса. Лишь для длинных реали-ций ($T>5T_f$) с целью экономии машинного времени модели-овалось по 200 реализаций.

В табл. 7 приводятся статистические характеристики числа вы-осов через уровни $c=0$, $c=1$, $c=2$ для моделированных по приведенному алгоритму эталонных выборок. Для уровня $c=$(наряду со средним числом выбросов $\bar{N}_э$ на реализации T и сред ним квадратическим отклонением числа выбросов σ_N для эталон ной выборки, приводятся теоретические оценки среднего числ выбросов для процесса $\bar{N}(0)$, полученные по формуле (13), и дл последовательности с эталонной дискретностью $\bar{N}$ $(\Delta_э)$, получен ные по формуле (10). Сравнение их показывает, что занижени числа выбросов на эталонной выборке за счет дискретности е значительно меньше выборочных погрешностей, которые при ис пользованном числе реализаций составляют несколько процентс от среднего числа выбросов.

Таблица

Статистические характеристики числа выбросов эталонных выборок

$\frac{T}{T_f}$	Число реализаций	$c=0$				$c=1$		$c=2$	
		$\bar{N}(0)$	$\bar{N}(\Delta_э)$	$\bar{N}_э$	σ_N	$\bar{N}_э$	σ_N	$\bar{N}_э$	σ_N
1	500	0,159	0,159	0,156	0,38	0,104	0,34	0,018	0,1[illegible]
2	500	0,317	0,317	0,332	0,52	0,184	0,44	0,036	0,2[illegible]
3	500	0,477	0,476	0,486	0,64	0,300	0,53	0,070	0,2[illegible]
5	500	0,796	0,793	0,780	0,74	0,478	0,68	0,094	0,3[illegible]
7	200	1,114	1,110	1,110	0,83	—	—	—	—
10	200	1,592	1,586	1,520	1,02	—	—	—	—
20	200	3,183	3,162	3,090	1,23	—	—	—	—

Из табл. 8 видно, что дисперсия числа выбросов последова тельности уменьшается с увеличением интервала между ее чл нами, однако это уменьшение сравнительно невелико. Это да основание считать, что значения σ_N, полученные для эталонно реализации, могут быть с точностью до выборочной погрешност приписаны и моделируемому непрерывному процессу.

В табл. 8 приводятся данные о дисперсии числа выбросов лиш до уровня $c=2$. Для более высоких уровней число выбросов с щественно уменьшается и получение надежных оценок методо статистического моделирования требует использования гораз бо́льших объемов данных. Практической необходимости в этом н было, поскольку известно, что для этих уровней распределени числа выбросов приближается к закону Пуассона, следовательн

$$\sigma_N \approx \sqrt{\bar{N}_c(\Delta)} \quad (c \geqslant 2). \qquad (3$$

Среднее число выбросов последовательности с корреляционной функцией (11) на реализации длины T при интервале Δ между членами последовательности и отсутствии ошибок измерений, в соответствии с формулами (13), (15) и (26), определяется приближенным выражением

$$\overline{N_c}(\Delta) = \frac{1}{2\pi}\frac{T}{T_f}e^{-\frac{c^2}{2}}\left(1 - \frac{1}{3}\frac{\Delta}{T_f}\right). \tag{31}$$

Таблица 8

Среднее квадратическое отклонение числа выбросов последовательности на реализации длины T при различных интервалах между членами

$\frac{T}{T_f}$	Δ/T_f					
	0,01	0,02	0,05	0,10	0,20	0,50
			$c=0$			
1	0,38	0,37	0,37	0,37	0,35	0,33
2	0,52	0,52	0,51	0,50	0,49	0,45
3	0,64	0,63	0,62	0,61	0,58	0,53
5	0,74	0,74	0,73	0,71	0,68	0,65
7	0,84	0,84	0,82	0,82	0,81	0,74
10	1,02	1,02	0,98	0,97	0,94	0,86
20	1,23	1,23	1,2)	1,18	1,14	0,96
			$c=1$			
1	0,34	0,34	0,33	0,37	0,35	0,33
2	0,44	0,44	0,42	0,41	0,40	0,36
3	0,53	0,53	0,52	0,51	0,48	0,44
5	0,68	0,67	0,66	0,65	0,63	0,59
			$c=2$			
1	0,13	0,13	0,13	0,13	0,13	0,11
2	0,20	0,20	0,20	0,18	0,18	0,18
3	0,28	0,28	0,28	0,28	0,28	0,24
5	0,30	0,30	0,30	0,30	0,30	0,28

Для не очень длинных реализаций пропуск выброса является сравнительно редким событием. Поэтому при оценке средней квадратической ошибки определения числа выбросов за счет дискретности изменений при отсутствии ошибок измерений также естественно исходить из закона Пуассона. В этом случае из (31) следует, что $\sigma_{\Delta N} \approx \sqrt{\Delta \bar{N}} \approx \sqrt{\bar{N} \cdot \frac{1}{3} \frac{\Delta}{T_f}}$, так что

$$\sigma_{\Delta N} \approx \sqrt{\frac{1}{6\pi}\left(\frac{T}{T_f}\right)\left(\frac{\Delta}{T_f}\right)e^{-\frac{c^2}{4}}}. \qquad (32)$$

Из табл. 9, в которой приведены оценки величины $\sigma_{\Delta N}$, полученные методом статистического моделирования и путем расчет

Таблица

Средняя квадратическая ошибка определения числа выбросов через уровень $c=0$ за счет дискретности измерений

$r(\tau)=(1+\tau/T_f)\, e^{-\tau/T_f}$; $\eta^2=0$

$\frac{T}{T_f}$	$\bar{N}$	Δ/T_f								
		Расчет методом статистического моделирования					Расчет по формуле (32)			
		0,02	0,05	0,10	0,20	0,50	0,01	0,02	0,20	0,50
1	0,159	0,04	0,04	0,09	0,13	0,18	0,02	0,03	0,07	0,16
2	0,317	0,04	0,11	0,13	0,17	0,25	0,03	0,05	0,10	0,23
3	0,477	0,06	0,12	0,13	0,20	0,31	0,04	0,06	0,13	0,28
5	0,796	0,05	0,12	0,16	0,27	0,38	0,05	0,07	0,16	0,36
7	1,114	0,00	0,10	0,17	0,23	0,34	0,06	0,09	0,19	0,43
10	1,592	0,00	0,14	0,21	0,33	0,48	0,07	0,10	0,23	0,50
20	3,183	—	0,23	0,27	0,45	0,69	0,10	0,14	0,32	0,52

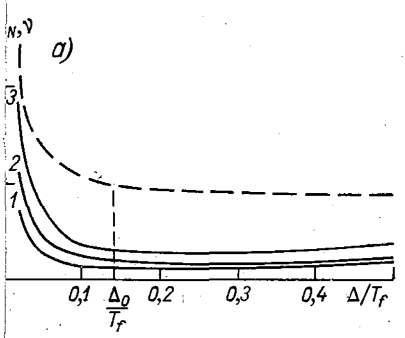
по формуле (32), видно, что использование последней дает удовлетворительные результаты уже при $c=0$. Для высоких уровне согласование должно быть значительно лучшим. Приведенны в табл. 9 оценки для эталонной последовательности $\left(\frac{\Delta}{T_f}=0{,}01\right.$ показывают, что средняя квадратическая ошибка числа выбросов для нее также составляет несколько процентов от истинного числа выбросов моделируемого процесса.

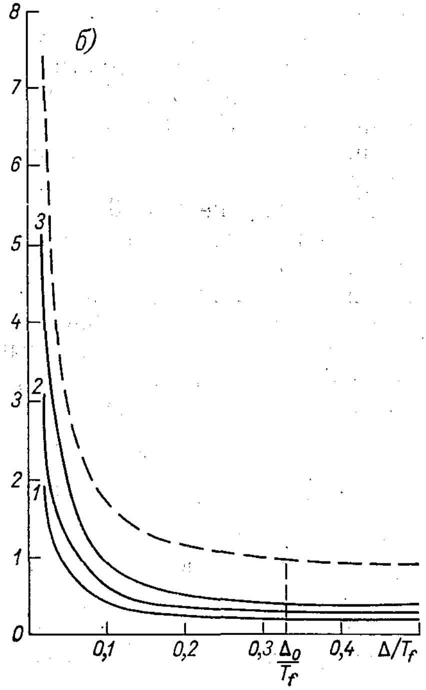
В случае наличия погрешностей измерения могут иметь мест как пропуски фактически существующих выбросов процесса, та и регистрация фиктивных выбросов. Ошибки в определении числа выбросов могут иметь разный знак и распределение их может н подчиняться закону Пуассона. Некоторые оценки для этого слу чая приводятся на рис. 3, на котором представлена зависимост от дискретности измерений отношения измеренного среднего чис ла выбросов к среднему числу выбросов для эталонной последо вательности и средней квадратической ошибки определения числа

ıбросов на реализации. Полученные для разной длины реали-
ıции T зависимости $\nu(\Delta)$ мало различаются между собой и прак-
ıчески совпадают с теоретической, которая представлена на
ıс. 3.

Полученные данные недо-
ıтаточны для суждения о ха-
ıактере распределения оши-
ıок числа выбросов и для на-
ıежного определения зависи-
ıости величины их от точ-
ıости измерений. Тем не ме-
ıее, некоторые закономер-
ıости они позволяют выявить.

Наличие ошибок наблюде-
ıий приводит при большой





ıис. 3. Зависимость от дискретности измерения Δ средней квадратической ıшибки $\sigma_{\Delta N}$ определения числа выбросов на реализации длиной T (сплошные ıинии) и отношения ν среднего числа выбросов последовательности с дискрет-ıостью Δ к числу выбросов эталонной последовательности (штриховые линии). $c=0$, корреляционная функция (11).

а) $\eta^2=0{,}001$, *б*) $\eta^2=0{,}01$; *1*) $T=T_f$, *2*) $T=2T_f$, *3*) $T=5T_f$.

ıустоте измерений не только к большим систематическим, но и к ıольшим случайным ошибкам. Последние при малых интервалах ıискретности Δ убывают с увеличением Δ. При некоторой дискрет-ıости случайные ошибки достигают минимальных значений, после ıего рост интервала Δ приводит уже к увеличению уровня ошибок. ıискретность, при которой достигается минимум случайных оши-ıок, может быть также названа оптимальной дискретностью изме-ıений. Она не совпадает с оптимальной дискретностью, опреде-ıенной нами выше как дискретность, при которой число выбросов ıоследовательности является несмещенной оценкой числа выбро-ıов процесса. Согласно полученным результатам, оптимальная дискретность в смысле минимума случайной ошибки числа в[ы]бросов соответствует измерениям через интервалы, бо́льшие, ч[ем] Δ_0. Так, для рассматриваемого примера (см. также табл. 5) $\Delta_0 = 0,14T_f$ при $\eta^2=0,001$ и $\Delta_0=0,33T_f$ при $\eta^2=0,01$. Из рис. 3 сл[е]дует, что минимальная ошибка достигается при $\eta^2=0,001$ для и[н]тервалов около $0,2T_f$, а при $\eta^2=0,01$ минимум в рассмотренн[ом] диапазоне интервалов до $\Delta=0,5T_f$ не обнаруживается. Вмес[те] с тем, поскольку в области минимума ошибки сравнительно ма[ло] зависят от интервала Δ, представляется, что использование опт[и]мального интервала Δ_0 может практически считаться обеспеч[и]вающим и минимум случайных ошибок в определении числа в[ы]бросов.

Из сравнения рис. 3 с табл. 9 видно, что для оптимальной ди[с]кретности измерений уровень ошибок в определении числа вы[бросов] повышается по сравнению с ошибками для случая обр[а]ботки точных измерений той же дискретности. Для рассмотре[н]ного примера это повышение составляет около 30%. Уточнен[ие] этих оценок и более детальный анализ влияния дискретности и[з]мерения на статистические характеристики выбросов случайн[ых] процессов требует дальнейших исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л., Федорченко Е. И. О применении теории выбросов к и[с]следованию температурных рядов. — «Труды ГГО», 1970, вып. 267, с. 86—99.
2. Иванов В. Н., Стратонович Р. Л. О числе серий выбросов турбулен[т]ных пульсаций скорости. — «Метеорол. и гидрол.», 1972, № 11, с. 37—42.
3. Каган Р. Л. О построении оптимальных конечно-разностных и квадрат[у]рных формул для однородных и изотропных случайных полей. — Из[в.] АН СССР, сер. физики атмосферы и океана, 1967, т. 3, № 6, с. 592—60[…]
4. Свешников А. А. Прикладные методы теории случайных функций — [М.,] «Наука», 1968. 463 с.
5. Каган Р. Л., Федорченко Е. И. К расчету вероятности выбро[сов] нормальной последовательности.— См. наст. сборник.
6. Лугина К. М., Каган Р. Л. К вопросу о пространственно-временном а[на]лизе барического поля. — «Труды ГГО», 1974, вып. 336, с. 75—94.
7. Наблюдения на гидрометеорологической сети СССР. Определение понят[ий] гидрометеорологических элементов и оценка точности наблюдений. Под р[ед.] О. А. Городецкого. Л., Гидрометеоиздат, 1970. 91 с.
8. Федорченко Е. И. О временно́й структуре ежечасных наблюдений за те[м]пературой воздуха в Ленинграде. — См. наст. сборник.
9. Каган Р. Л., Канашкин В. К., Федорченко Е. И. О расчете хара[к]теристик временных рядов методом статистического моделирования. — «Тр[у]ды ГГО», 1972, вып. 286, с. 71—82.
10. Быков В. В. Цифровое моделирование в статистической радиотехнике. М., «Советское радио», 1971. 328 с.

Р. Л. КАГАН, Е. И. ФЕДОРЧЕНКО

О ВОССТАНОВЛЕНИИ ГОДОВОГО ХОДА МОМЕНТОВ МЕТЕОРОЛОГИЧЕСКИХ РЯДОВ

При исследовании структуры метеорологических рядов важное значение имеет оценка моментов их распределений (среднего значения, дисперсии, асимметрии, эксцесса). Для ряда задач существенно уметь определять моменты распределений для конкретных дней месяца. Укажем, например, на задачу расчета вероятностей выбросов метеорологических рядов, при решении которой, как показано в [1], для переходных месяцев необходимо учитывать внутримесячный ход моментов распределения. С определением ежедневных значений моментов связана и оценка средних за месяц значений моментов.

Получение моментов распределений на каждый день месяца эмпирическим способом очень трудоемко. Кроме того, значения моментов для конкретных дней месяца, полученные непосредственной обработкой данных наблюдений, имеют небольшую точность из-за ограниченности выборки, и поэтому результаты такого расчета затруднительно использовать без дополнительного сглаживания их.

Между тем, для метеорологических элементов, распределения которых приводятся в климатологических справочниках, можно попытаться решить обратную задачу, а именно: по данным о повторяемостях значений метеорологических элементов в месячных совокупностях восстановить значения моментов метеорологических элементов на каждый день месяца.

При этом необходимо иметь в виду, что по повторяемостям значений метеорологических элементов по данным наблюдений за месяц можно найти непосредственно лишь средние за месяц значения начальных моментов. Поскольку значения моментов в течение месяца изменяются, то для получения по начальным моментам значений центральных моментов необходимы дополнительные предположения о характере изменения моментов распределения во времени.

Пусть $m_k(t)$ и $\mu_k(t)$ — значения соответственно начального и центрального моментов k-того порядка в момент времени t (t — время в месяцах, отсчитываемое от начала года). Черта сверху в дальнейших обозначениях означает осреднение по промежутку $[i-1, i]$, где i — номер месяца. Например $\overline{m_k} = \int\limits_{i-1}^{i} m_k(t)\,dt$ — среднее за i-й месяц значение начального момента k-того порядка. Значения $\overline{m_k}$ можно найти непосредственным расчетом по повторяемостям значений метеорологических элементов, приводимым в климатологических справочниках. Для нахождения среднего за i-й месяц значения центрального момента k-того порядка $\overline{\mu_k}$ необходимо знание не только средних значений начальных моментов $\overline{m_j}$ ($j=1, 2, \dots, k$), но и хода $m_1(t)$ и $\mu_j(t)$ ($j=2, \dots, k-1$) в течение рассматриваемого месяца.

Действительно, $\overline{\mu_k}$ ($k=2, 3, 4$) выражается через $\overline{m_k}$ следующим образом:

$$\overline{\mu_2} = \overline{m_2} - \overline{m_1^2}; \tag{1}$$

$$\overline{\mu_3} = \overline{m_3} - 3\,\overline{\mu_2 m_1} - \overline{m_1^3}; \tag{2}$$

$$\overline{\mu_4} = \overline{m_4} - 4\,\overline{\mu_3 m_1} - 6\,\overline{\mu_2 m_1^2} - \overline{m_1^4}. \tag{3}$$

Например, чтобы определить $\overline{\mu_2}$, нужно знать не только $\overline{m_2}$, но и $\overline{m_1^2}$. Для нахождения последнего необходимо задание хода $m_1(t)$ в течение месяца.

Таким образом, чтобы найти средние за месяц значения центральных моментов $\overline{\mu_k}$ можно по среднему значению $\overline{m_1}$ восстановить ход $m_1(t)$, после чего по формуле (1) найти $\overline{\mu_2}$ и восстановить $\mu_2(t)$, затем по формуле (2) найти $\overline{\mu_3}$ и восстановить $\mu_3(t)$ и т. д., т. е. для нахождения $\overline{\mu_k}$ достаточно задать алгоритм восстановления (способ интерполяции) внутримесячного хода моментов по данным об их средних значениях. При такой процедуре получения моментов задача нахождения среднего значения центрального момента k-того порядка эквивалентна задаче нахождения ежедневных значений момента ($k-1$)-го порядка.

Остановимся теперь на способах восстановления внутримесячного хода моментов.

В [1] был рассмотрен способ восстановления первых двух моментов распределения, который сводится к кусочной параболической интерполяции моментов для каждого месяца по данным о средних значениях этих моментов для рассматриваемого, предыдущего и последующего месяцев. Испытание этой методики на материале средней суточной температуры воздуха дало весьма удовлетворительные результаты. Вместе с тем подобная параболическая интерполяция, обеспечивая совпадение фактических средних за месяц значений моментов с расчетными, не предусматривает совпадения значений моментов на концах месяцев, в реультате чего моменты могут претерпевать скачки при переходе от одного месяца к другому. Это обстоятельство оказывается существенным при исследовании старших моментов распределения, для которых использование параболической интерполяции может приводить к заметным искажениям. Поэтому при рассмотрении 3-го и 4-го моментов желательно использовать такие способы восстановления их хода, которые обеспечивали бы непрерывность значений моментов в течение года.

Ниже будут рассмотрены два способа такого восстановления, основанные на применении рядов Фурье и метода сплайнов.

Опишем сначала алгоритм восстановления внутримесячного хода моментов по данным об их средних значениях, основанный на использовании тригонометрической интерполяции.

Пусть $X(t)$ — значение момента какого-либо порядка в момент времени t (t — время в месяцах, отсчитываемое от начала года). Будем искать $X(t)$ в виде

$$X(t)=a_0+2\sum_{k=1}^{5}\left(a_k\cos\frac{\pi k}{6}(t-6{,}5)+b_k\sin\frac{\pi k}{6}(t-6{,}5)\right)+ \\ +a_6\cos(\pi(t-6{,}5)). \tag{4}$$

Коэффициенты a_k ($k=0, 1, \ldots, 6$) и b_k ($k=1, \ldots, 5$) можно найти из условий:

$$\int_{i-1}^{i} X(t)\,dt=f_i \qquad (i=1,\ 2,\ \ldots,\ 12), \tag{5}$$

где f_i — средние значения моментов в i-й месяц, которые предполагаются известными.

Решая систему (5), получаем:

$$a_0=\frac{1}{12}\sum_{i=1}^{12}f_i;$$

$$a_k=\frac{\pi k}{12\sin(\pi k/12)}\cdot\frac{1}{12}\sum_{i=1}^{12}f_i\cos\left(\frac{\pi k}{6}(i-7)\right)\quad (k=1,\ \ldots,\ 6);$$

$$b_k=\frac{\pi k}{12\sin(\pi k/12)}\cdot\frac{1}{12}\sum_{i=1}^{12}f_i\sin\left(\frac{\pi k}{6}(i-7)\right)\quad (k=1,\ \ldots,\ 5). \tag{6}$$

Остановимся теперь на использовании для восстановления внутримесячного хода моментов сплайновой (кусочно-полиномиальной) интерполяции [2, 3].

Пусть на отрезке $[a,\ b]$ задана сетка $a=x_0<x_1<\ldots<x_n=b$, в узлах которой заданы значения g_i ($i=0, \ldots, n$) функции $g(x)$, определенной на отрезке $[a,\ b]$. Под сплайн-функцией, интерполирующей функцию $g(x)$, обычно понимается функция $h(x)$, на

промежутке $[x_{i-1}, x_i]$ $(i=1, 2, ..., n)$ представляющая собой полином некоторой степени $m<n-1$, непрерывная на участке $[a, b]$ вместе со своими производными до $(m-1)$-го порядка включительно и такая, что $h(x_i)=g_i (i=0, 1, ..., n)$. Для периодических с периодом $(b-a)$ функций естественно условие $h^{(j)}(a)=h^{(j)}(b)$ $(j=1, ..., m-1)$. Для непериодических функций для замыкания системы условий старшие производные в точках a, b обычно полагают равными нулю.

Наиболее употребительными являются сплайны — полиномы 3-й степени.

Интерполяция кубическими сплайнами оказывается эффективной на практике благодаря своему свойству минимума кривизны, которое заключается в том, что кубический сплайн $h(x)$ минимизирует значение $\int\limits_a^b (h''(x))^2 dx$ в классе функций, интерполирующих функцию $g(x)$ и имеющих на $[a, b]$ суммируемые с квадратом вторые производные [2].

Для задания кубического сплайна $h(x)$, интерполирующего периодическую с периодом $(b-a)$ функцию $g(x)$ на $[a, b]$, нужно решить систему из $4n$ уравнений, неизвестными которой являются коэффициенты полиномов 3-й степени на отрезках $[x_{i-1}, x_i]$. Причем n уравнений этой системы представляют собой условия $h(x_i)=g_i$, остальные $3n$ уравнений — условия непрерывности функции $h(x)$ и ее первых двух производных в узлах x_i $(i=1, ..., n)$.

Поскольку нам необходимо восстановить внутримесячный ход моментов по их средним значениям, а не по значениям в точках, то для этой цели метод сплайнов приходится использовать в несколько измененном виде.

Пусть $X(t)$, как и раньше,— момент распределения исследуемого ряда какого-либо порядка в момент времени t. С помощью узлов x_i $(i=1, ..., 12)$ разделим год на 12 промежутков и будем искать $X(t)$ в виде кусочно-кубической функции, непрерывной вместе со своими первыми и вторыми производными и удовлетворяющей условиям (5). В качестве узлов x_i выберем середины месяцев[1], т. е. будем искать $X(t)$ в виде:

$$X(t)=\begin{cases} A_{i-1}+B_{i-1}\left(t-i+\frac{3}{2}\right)+C_{i-1}\left(t-i+\frac{3}{2}\right)^2+ \\ +D_{i-1}\left(t-i+\frac{3}{2}\right)^3, \quad t\in\left[i-1,\ i-\frac{1}{2}\right], \\ A_i+B_i\left(t-i+\frac{1}{2}\right)+C_i\left(t-i+\frac{1}{2}\right)^2+ \\ +D_i\left(t-i+\frac{1}{2}\right)^3, \quad t\in\left[i-\frac{1}{2},\ i\right]. \end{cases} \tag{7}$$

[1] При выборе в качестве узлов x_i $(i=1, ..., 12)$ границ месяцев сплайн, удовлетворяющий условиям (5), построить не удается.

Коэффициенты A_i, B_i, C_i и D_i ($i=1$, ..., 12) находятся из усло-
ий (5) и условий непрерывности $X(t)$ вместе с первыми двумя оизводными в узлах x_i ($i=1$, ..., 12), которые имеют следующий д:

$$
\begin{cases}
\dfrac{A_{i-1}}{2}+\dfrac{3B_{i-1}}{8}+\dfrac{7C_{i-1}}{24}+\dfrac{15D_{i-1}}{64}+\dfrac{A_i}{2}+\dfrac{B_i}{8}+\dfrac{C_i}{24}+\dfrac{D_i}{64}=f_i; \\
A_i=A_{i-1}+B_{i-1}+C_{i-1}+D_{i-1}; \\
B_i=B_{i-1}+2C_{i-1}+3D_{i-1}; \\
2C_i=2C_{i-1}+6D_{i-1}. \qquad (i=1,\ \dots,\ 12).
\end{cases}
\tag{8}
$$

Решение системы (8) из 48 уравнений можно свести к решею системы уравнений 12-го порядка относительно коэффициенв C_i:

$$
\frac{1}{24}C_{i-2}+\frac{19}{6}C_{i-1}+\frac{115}{12}C_i+\frac{19}{6}C_{+1}+\frac{1}{24}C_{i+2}=
=8f_{i-1}-16f_i+8f_{i+1},
\tag{9}
$$

коэффициенты B_i, D_i, A_i найти последовательно по формулам:

$$
B_i=f_{i+1}-f_i+\frac{1}{192}C_{i-1}-\frac{115}{192}C_i-\frac{77}{192}C_{i+1}-\frac{1}{192}C_{i+2};
\tag{10}
$$

$$
D_i=\frac{1}{3}C_{i+1}-\frac{1}{3}C_i;
\tag{11}
$$

$$
A_i=f_i+\frac{1}{8}B_{i-1}+\frac{5}{24}C_{i-1}+\frac{17}{64}D_{i-1}-
-\frac{1}{8}B_i-\frac{1}{24}C_i-\frac{1}{64}D_i.
\tag{12}
$$

Указанные алгоритмы были реализованы на ЭВМ М-220 и исытаны на фактическом материале. В качестве примера ниже риведем некоторые результаты расчетов, выполненных на матеиале наблюдений за температурой воздуха на ГМС Москва — ельскохозяйственная академия.

Для этой станции в качестве исходных использовались данные повторяемостях средних суточных температур воздуха, привеенные для каждого месяца года в [4], которые были получены утем обработки рядов наблюдений за 1881—1960 гг.

По повторяемостям были определены значения начальных моентов, по которым в соответствии с формулами (1)—(3) и с писанной выше процедурой были последовательно вычислены ак средние за месяц значения центральных моментов, так и ежедневные их значения. При этом для восстановления внутримесячого хода моментов использовались тригонометрический, сплайовый и описанный в [1] параболический способы интерполяции.

По ежедневным значениям центральных моментов вычислялис значения на каждый день коэффициентов асимметрии $A(t)$ и экс цесса $E(t)$, по которым получались также средние месячные зна чения этих характеристик. При расчете дисперсии $\sigma^2(t)$ с цель исключения влияния группировки данных вводилась поправк Шеппарда, но при получении асимметрии и эксцесса использова

Таблица

Фактические и расчетные средние за месяц значения моментов (ГМС Москва — Сельскохозяйственная академия)

Характеристика	Месяц											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
$\overline{\theta}_1$	—10,6	—9,5	—4,6	4,2	12,2	16,3	18,5	16,4	10,7	4,3	—2,1	—7,
$\overline{\theta}$	—10,2	—9,5	—4,5	4,2	12,1	16,3	18,5	16,4	10,7	4,1	—2,1	—7,
$\overline{\sigma_1^2}$	58,1	43,9	26,0	15,6	22,6	16,6	11,5	12,3	12,6	16,8	24,3	46,
$\overline{\sigma_2^2}$	58,9	43,3	25,0	16,9	22,4	17,3	11,3	11,4	12,6	17,1	24,7	49,
$\overline{\sigma_3^2}$	58,9	43,3	25,1	16,7	22,5	17,3	11,3	11,3	12,6	17,1	24,7	49,
$\overline{\sigma_4^2}$	58,9	43,2	25,6	18,4	22,3	17,3	11,3	11,7	13,0	17,1	25,0	49,5
$\overline{A}_1$	—0,51	—0,25	—0,52	—0,02	—0,08	—0,08	0,12	0,28	0,11	—0,18	—0,64	—0,
$\overline{A}_2$	—0,53	—0,26	—0,40	0,26	—0,07	—0,10	0,03	0,25	0,18	—0,22	—0,46	—0,
$\overline{A}_3$	—0,53	—0,27	—0,37	0,25	—0,08	—0,06	0,03	0,23	0,23	—0,23	—0,48	—0,
$\overline{A}_4$	—0,53	—0,27	—0,42	0,31	—0,04	—0,14	0,03	0,24	0,14	—0,15	—0,49	—0,
$\overline{E}_1$	—0,3	—0,4	—0,2	—0,2	—0,6	—0,6	—0,3	—0,2	—0,3	—0,2	0,4	0,
$\overline{E}_2$	—0,1	—0,4	—0,6	—0,6	—0,5	—0,6	—0,3	—0,2	—0,5	0,2	—0,0	0,
$\overline{E'_2}$	—0,1	—0,4	—0,6	—0,6	—0,5	—0,6	—0,3	—0,2	—0,4	0,2	0,2	0,
$\overline{E'_3}$	—0,1	—0,4	—0,6	—0,6	—0,4	—0,7	—0,3	—0,0	—0,5	0,2	0,2	0,
$\overline{E'_4}$	—0,1	—0,4	—0,6	—0,4	—0,6	—0,5	—0,3	—0,3	—0,2	—0,0	0,3	0,

лись значения дисперсии и 4-го центрального момента, вычислен ные без учета поправок на группировку, поскольку, как показан в [5], эти поправки значительно уменьшают точность определе ния асимметрии и эксцесса.

С целью сопоставления полученных таким образом моментны характеристик с фактическими значениями параллельно был вы полнен расчет последних по данным ежедневных наблюдений н той же станции за 1879—1972 гг., любезно предоставленным нам А. И. Неушкиным. При обработке фактических данных было со чтено целесообразным использование возможно более полны рядов наблюдений, что давало основания рассчитывать на полу

ение статистически более надежных результатов. С другой сто-
оны, следует, конечно, иметь в виду, что несовпадение периодов ложет привести к известным различиям в оценках моментов.

Полученные в результате расчетов средние значения моментов ля различных месяцев года приводятся в табл. 1. При этом ндексом 1 помечены значения моментов, полученные путем не-осредственного расчета, а индексами 2, 3 и 4 — значения момен-ов, вычисленные с помощью сплайновой, тригонометрической параболической, описанной в [1], интерполяции соответственно. Іоскольку средняя месячная температура рассчитывается непо-редственно по распределениям, приводимым в Справочнике, таблице приводятся только эти значения средних месячных тем-ератур, обозначенные через $\overline{\theta}$, и значения средних месячных емператур, полученные по фактическим данным, которые обо-начены через $\overline{\theta_1}$. При вычислении средних значений эксцесса $\overline{E}'_2$, $'_3$ и $\overline{E}'_4$ использовалась несколько видоизмененная методика, которой будет сказано иже.

Рассмотрение табл. 1 оказывает, что оценки редних значений моментов, олученные упомянутыми ремя способами, оказыва-отся очень близкими друг другу и сравнительно ма-о отличаются от фактиче-ких. При оценке расхож-ений между расчетными фактическими значениями оментов следует иметь в иду выборочные погреш-ости этих характеристик, риближенные значения которых представлены в табл. 2. При-еденные в таблице значения ошибок получены по методике, писанной в [6], в предположении, что ряд близок к нормальному экспоненциально-убывающей корреляционной функцией, и что оэффициент корреляции между смежными членами ряда равен ,8. Учитывая эти погрешности, согласование между расчетными фактическими средними значениями моментов можно считать довлетворительным.

Таблица 2

Средние квадратические ошибки определения средних месячных значений моментов распределения средней суточной температуры

Число лет	Характеристика			
	$\sigma_{\overline{\theta}}$	$\sigma_{\overline{\sigma^2}}$	$\sigma_{\overline{A}}$	$\sigma_{\overline{E}}$
25	$0{,}10\sqrt{\overline{\sigma^2}}$	$0{,}11\,\overline{\sigma^2}$	0,15	0,27
90	$0{,}06\sqrt{\overline{\sigma^2}}$	$0{,}06\,\overline{\sigma^2}$	0,08	0,15

Различия между указанными способами интерполяции выяв-яются при изучении внутримесячного хода моментов, получен-ого при их использовании. Рассмотрение рис. 1—4, на которых редставлен годовой ход моментов средней суточной температу-ы, показывает, что полученные непосредственным расчетом для тдельных дней года значения моментов не являются достаточно адежными даже при использовании рядов наблюдений за 90 лет, следствие чего изменения значений моментов ото дня ко дню являются малоупорядоченными. При необходимости применени таких данных целесообразно было бы производить их сглажива ние, например, путем скользящего осреднения. С этой точки зре ния определенные преимущества представляет использовани проинтерполированных моментов, поскольку при интерполяци

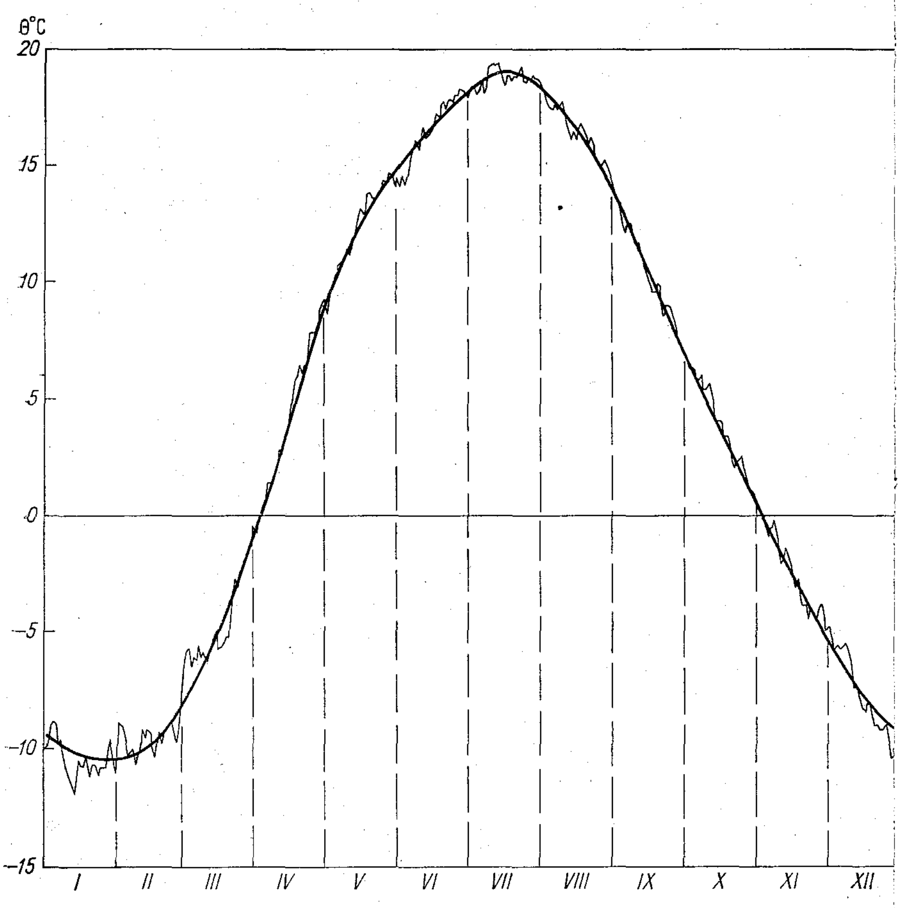


Рис. 1. Годовой ход среднего значения средней суточной температуры во духа (ГМС Москва — с.-х. академия).

Усл. обозначения см. рис. 2.

ход моментов автоматически получается сглаженным, несущес венные флуктуации моментов ото дня ко дню исключаются, и то же время общий ход моментов в течение года улавливаетс достаточно хорошо.

Кривые годового хода средней суточной температуры, восста новленного путем использования разных способов интерполяци практически совпадают, поэтому на рис. 1 мы ограничилис

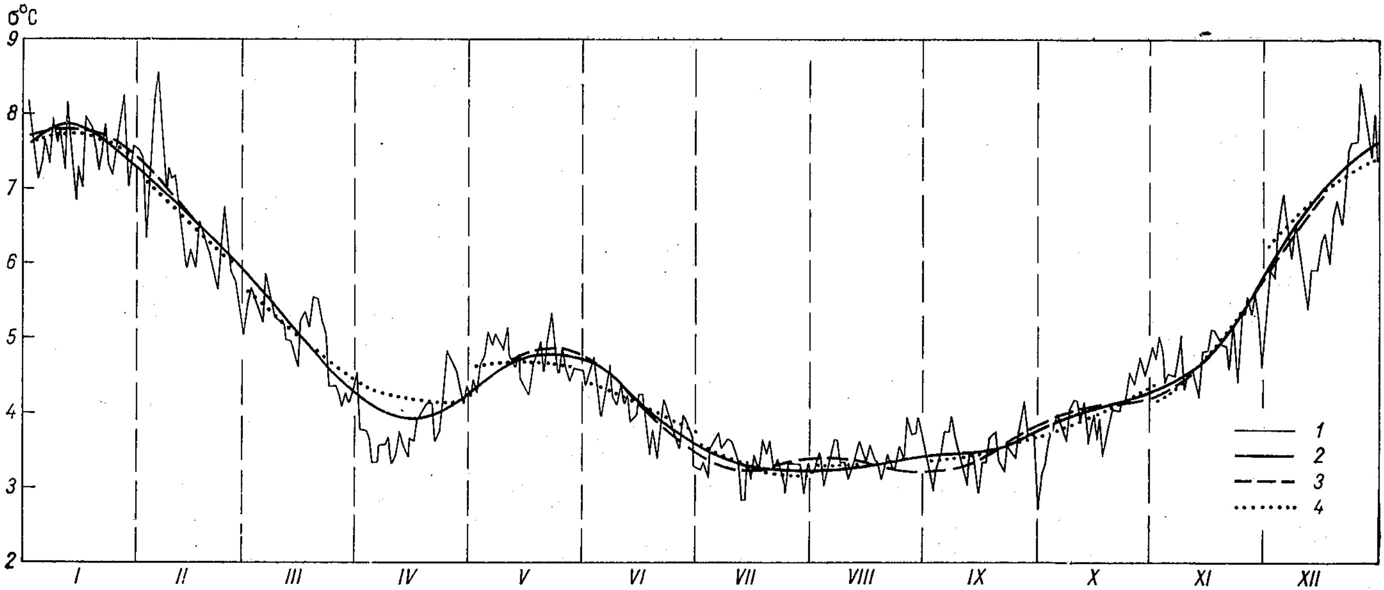
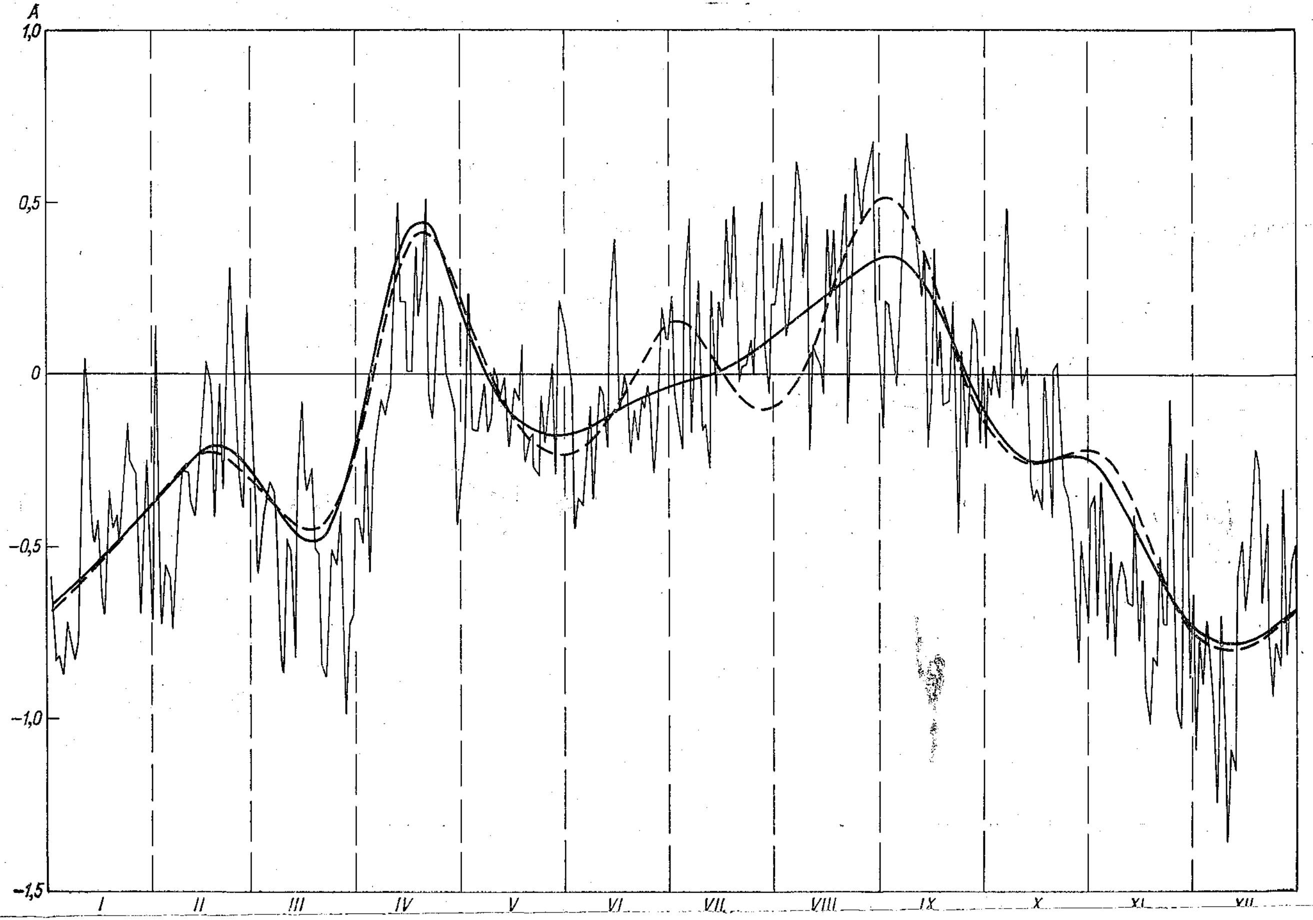


Рис. 2. Годовой ход среднего квадратического отклонения средней суточной температуры в здуха (ГМС Москва — с.-х. академия).

1 — фактические данные, *2* — восстановленные путем сплайновой интерполяции, *3* — восстановленные путем тригонометрической интерполяции, *4* — восстановленные путем интерполяции параболой.



A
1,0
0,5
0
-0,5
-1,0
-1,5
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII

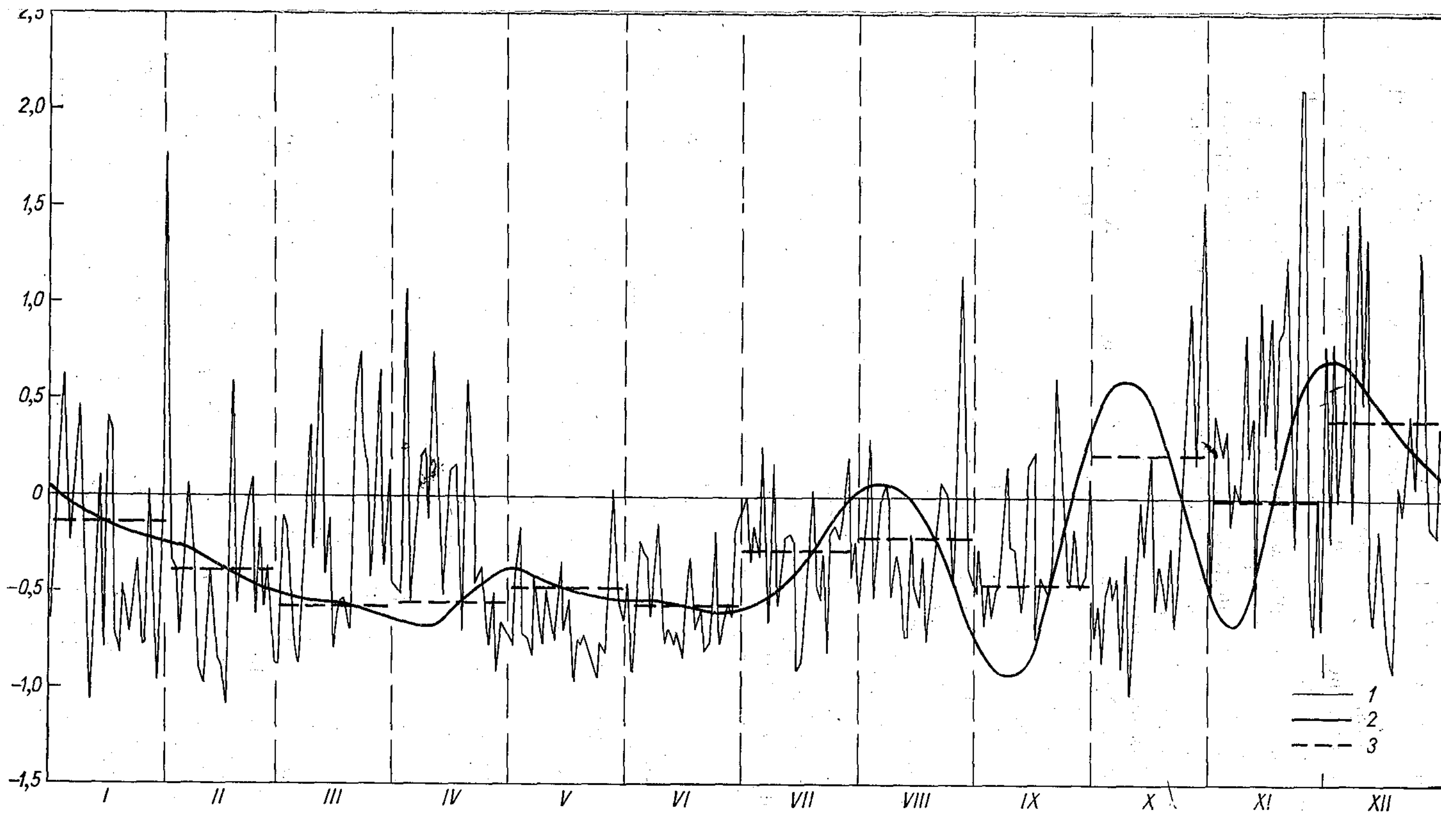


Рис. 4. Годовой ход коэффициента эксцесса средней суточной температуры воздуха (ГМС Москва — с.-х. академия). 1 — фактические данные, 2 — $E_2(t)$ (восстановленные путем сплайновой интерполяции), 3 — $\overline{E_2(t)}$.

проведением одной линии. Кривые годового хода среднего кв ратического отклонения температуры, полученного расчетным с собом, также сравнительно мало различаются между собой. Н более существенно то, что использование параболической инт поляции приводит к разрыву значений дисперсии при переходе одного месяца к другому. Для старших моментов разрывы яв ются еще более значительными и настолько искажают реальн ход моментов, что использование этого способа интерполяции д них представляется нецелесообразным. Соответствующие кри на рис. 3 и 4 не приводятся.

Наибольший интерес с точки зрения выбора рациональн способа интерполяции представляет рис. 3, на котором изобра годовой ход коэффициента асимметрии. Из этого рисунка вид что в течение большей части года кривые годового хода коэф циента асимметрии, восстановленного с помощью тригономет ческой и сплайновой интерполяции, хорошо согласуются как м ду собой, так и с фактическими данными. Исключение составля летние месяцы (июнь — август), для которых тригонометричес интерполяции дает два максимума в ходе коэффициента асимм рии, в то время как метод сплайнов описывает сравнительно р номерный рост асимметрии в течение лета с одним максимум в конце его. Судя по имеющимся данным, метод сплайнов обес чивает более реалистическое описание хода коэффициента ас метрии. Это, по-видимому, связано с упомянутым нами свойст минимальной кривизны сплайнов.

Из рассмотрения рис. 4, на котором приведен годовой ход з чений эксцесса, видно, что ход последних носит особенно неу рядоченный характер. Описание этого хода любым из указанн выше способов позволяет уловить лишь самые общие законом ности. Как видно из рис. 4, даже использование метода сплай дает довольно большие расхождения с фактическими данны Эти расхождения, по-видимому, связаны не только с возможн несовершенством принятого способа интерполяции, но и с у упомянутой ненадежностью индивидуальных моментов. В то время осредненные за месяц значения проинтерполированных ментов дают некоторое представление о ходе фактических д ных (см. линии *3* на рис. 4). В силу этого было сочтено целесо разным при восстановлении годового хода эксцесса исходить постоянства его в течение месяца. Рассчитанные таким обра значения несколько отличаются от тех, которые получаю осреднением восстановленных путем интерполяции сплайн ежедневных значений эксцесса, однако различия эти сравнит но невелики. Именно такая упрощенная методика использовал при расчете значений эксцесса $\bar{E}_2'$, $\bar{E}_3'$ и $\bar{E}_4'$, приведен в табл. 1.

Авторам представляется, что приведенные результаты п тверждают целесообразность интерполяции для восстановле годового хода моментов распределения. Наиболее перспективн с этой точки зрения, является, по-видимому, использование ме а сплайнов. В то же время следует иметь в виду, что точность акого восстановления при заданном объеме исходных данных меньшается с увеличением порядка момента, вследствие чего ри расчете моментов температурных рядов приходится отказыаться от детального восстановления хода эксцесса внутри месяа. Возможно, что такое восстановление окажется целесообразым для метеорологических элементов, распределения которых ущественно отличаются от нормального.

СПИСОК ЛИТЕРАТУРЫ

. Каган Р. Л., Федорченко Е. И. О расчете статистических характеристик выбросов случайной функции. — «Труды ГГО», 1970, вып. 268, с. 146—172.

. Алберг Дж., Нильсон Э., Уолш Дж. Теория сплайнов и ее приложения. М., «Мир», 1972. 316 с.

. Марчук Г. И. Методы вычислительной математики. Новосибирск, «Наука», 1973. 352 с.

. Справочник по климату СССР. Часть II. Л., Гидрометеоиздат, 1965.

. Федорченко Е. И. О влиянии группировки данных на точность выборочных моментов. — «Труды ГГО», 1974, вып. 336, с. 159—165.

. Федорченко Е. И. О влиянии связности метеорологических рядов на точность выборочных моментов. — «Труды ГГО», 1974, вып. 336, с. 25—47.

Е. И. ФЕДОРЧЕНК

О ВРЕМЕННОЙ СТРУКТУРЕ ЕЖЕЧАСНЫХ НАБЛЮДЕНИ ЗА ТЕМПЕРАТУРОЙ ВОЗДУХА В ЛЕНИНГРАДЕ

1. При изучении рядов наблюдений за температурой в различ ные часы суток наибольшее внимание уделялось суточному ход средней температуры воздуха. Суточному ходу других момент распределения температуры, а также характеристик временнó корреляции ее посвящено сравнительно мало работ. Из таких р бот следует упомянуть в первую очередь статью Л. В. Тихомир вой [1], в которой рассматривается суточный ход дисперси асимметрии и эксцесса. Некоторые данные по этому вопро приводятся в [2]. Исследование суточного хода дисперсии темп ратуры воздуха на материале ежечасных наблюдений на станци Ленинград, ГМО, проводится в работе [3], в которой с помощь статистических тестов показано, что имеющиеся данные о суто ном ходе дисперсии противоречат гипотезе стационарности ряд В работе [4] рассматривается также вопрос о суточном ходе вр меннóй корреляционной функции температуры для этой же ста ции и указывается на значимость его для сдвигов по времени 3 ч и более.

К сожалению, существующие оценки получены на основани обработки рядов наблюдений лишь за один день каждого месяц Поэтому из результатов, приводимых в работе, для 1 мая и 1 се тября не вполне ясно, реальны ли сделанные выводы о суточно ходе и в какой мере их можно распространить на другое врем года.

Целью настоящей работы было попытаться рассмотреть во можно детальнее суточный ход первых четырех моментов распр деления температуры и временнóй корреляции ее. В качестве и ходных данных брались ежечасные наблюдения за температуро воздуха (по термографу) на ст. Ленинград, ГМО, для четыр центральных месяцев сезонов (январь, апрель, июль, октябрь) 1944—1965 гг. Для того чтобы на результатах возможно меньш азывалось влияние нестационарности ряда за счет годового хода, в то же время, чтобы полученные оценки были достаточно усйчивы, расчеты производились для каждой пентады месяца отдельности (при этом к пентаде добавлялось одно наблюдение следующей; наблюдения за 31-е число в январе, июле, октябре брасывались).

По этим данным для каждого срока наблюдений i ($i=1, 2, \ldots,$ 24) k-того дня j-той пентады рассматриваемого месяца ($k=1, 2,$ 4, 5; $j=1, 2, \ldots, 6$) вычислялось среднее $m(i, j, k)$ в этот момент емени, дисперсия $\sigma^2(i, j, k)$, коэффициенты асимметрии $A(i, j, k)$ эксцесса $E(i, j, k)$. Кроме того, для сдвигов по времени $\tau=$ 1, 2, ..., 120 ч вычислялись значения коэффициентов корреляции $i, i+\tau, j, k$) между температурой в i-тый срок наблюдений того дня j-той пентады и температурой в момент времени через часов от указанного.

По индивидуальным значениям одноточечных и корреляционых моментов для каждого срока наблюдений i ($i=1, \ldots, 24$) пяти ток пентады рассчитывались средние за пентаду значения:

$$m(i, j)=\frac{1}{5}\sum_{k=1}^{5} m(i, j, k);$$

$$\sigma^2(i, j)=\frac{1}{5}\sum_{k=1}^{5} \sigma^2(i, j, k);$$

$$A(i, j)=\frac{1}{5}\sum_{k=1}^{5} A(i, j, k);$$

$$E(i, j)=\frac{1}{5}\sum_{k=1}^{5} E(i, j, k) \quad (j=1, 2, \ldots, 6);$$

$$r'(i, i+\tau, j)=\frac{1}{5}\sum_{k=1}^{5} r(i, i+\tau, j, k) \quad (\tau=1, 2, \ldots, 120). \qquad (1)$$

Значения $m(i, j)$, $\sigma^2(i, j)$, $A(i, j)$, $E(i, j)$, $r'(i, i+\tau, j)$ осреднялись о каждому из аргументов и получались как средние за месяц начения моментов для каждого срока наблюдений:

$$m_1(i)=\frac{1}{6}\sum_{j=1}^{6} m(i, j);$$

$$\sigma_1^2(i)=\frac{1}{6}\sum_{j=1}^{6} \sigma^2(i, j);$$

$$A_1(i) = \frac{1}{6} \sum_{j=1}^{6} A(i, j);$$

$$E_1(i) = \frac{1}{6} \sum_{j=1}^{6} E(i, j);$$

$$r'_1(i, i+\tau) = \frac{1}{6} \sum_{j=1}^{6} r'(i, i+\tau, j), \qquad (2)$$

так и средние за пентаду значения:

$$m_2(j) = \frac{1}{24} \sum_{i=1}^{24} m(i, j);$$

$$\sigma_2^2(j) = \frac{1}{24} \sum_{i=1}^{24} \sigma^2(i, j);$$

$$A_2(j) = \frac{1}{24} \sum_{i=1}^{24} A(i, j); \qquad (3)$$

$$E_2(j) = \frac{1}{24} \sum_{i=1}^{24} E(i, j);$$

$$r'_2(\tau, j) = \frac{1}{24} \sum_{i=1}^{24} r'(i, i+\tau, j).$$

При высокой временно́й связности целесообразно пользоваться не непосредственным осреднением коэффициентов корреляции, а осреднением величин z, полученных из коэффициентов корреляции путем преобразования Фишера [5]. Как показал опыт, это позволяет существенно ослабить влияние отдельных ошибок в исходных данных. Поэтому наряду со значениями $r'(i, i+\tau, j)$, $r'_1(i, i+\tau)$ и $r'_2(\tau, j)$, нами вычислялись значения $r(i, i+\tau, j)$, $r_1(i, i+\tau)$ и $r_2(\tau, j)$, полученные обратным преобразованием средних значений z-преобразований коэффициентов корреляции.

2. Остановимся сначала на одноточечных моментах. Полученные результаты показали, что при расчете моментов в апреле для выявления суточного хода как среднего, так и дисперсии и асимметрии существенно учитывать нестационарность за счет годового хода. Лишь для эксцесса, вследствие малой точности его расчета из-за ограниченности выборки, не удается получить значимых результатов для отдельных пятидневок.

В табл. 1 для каждой пентады апреля приводятся средние за пентаду значения средней температуры, дисперсии и коэффициента асимметрии на каждый срок наблюдений. Из этой таблицы дно, что в апреле суточный ход моментов для разных пентад цественно различен. Рассмотрим, например, суточный ход диэрсии. В [1, 2] было показано, что для дисперсии можно вылить два типа суточного хода: зимний, с наибольшим значением сперсии в предутренние часы и наименьшим в дневные, и летй, с максимумом дисперсии в дневные часы. Из табл. 1 видно,) в Ленинграде в апреле происходит смена типов суточного да дисперсии: в первой пятидневке максимум дисперсии прихотся на ночные часы, а минимум — на дневные, в последней пядневке виден явно выраженный максимум в дневные часы.

Претерпевают изменения в апреле и закономерности в суточм ходе асимметрии. В ночные и утренние часы коэффициент имметрии от начала месяца растет, достигая максимума в пяй пятидневке, после чего к концу месяца начинает уменьшаться. ри этом в первой половине месяца коэффициент асимметрии ночные и утренние часы отрицателен, а во второй половине — ложителен. В дневные и вечерние часы наибольших значений эффициент асимметрии достигает в четвертой пятидневке, прим отрицательные значения коэффициента асимметрии в эти сы наблюдаются лишь в первой пятидневке. Заметим, что максмум в весенний переходный месяц в период после схода снежно- покрова, по-видимому, характерен и для годового хода коэфициента асимметрии средней суточной температуры (см. [6]).

В октябре нестационарность за счет годового хода сказываетна изменении значений моментов от пентады к пентаде в меньэй степени, чем в апреле. Эту нестационарность необходимо итывать при определении суточного хода среднего значения. Заэтны различия от пятидневки к пятидневке и в значениях диерсии для каждого срока наблюдений (значения дисперсии стут от начала месяца к концу, что согласуется с известными нее особенностями годового хода дисперсии [2]), однако различя эти невелики и для многих задач ими можно пренебречь. Заономерности суточного хода дисперсии от пентады к пентаде октябре не меняются. При определении суточного хода моменв в январе и июле нестационарностью за счет годового хода течение месяца можно пренебречь.

В табл. 2 представлены средние за месяц значения моментов аспределения температуры на каждый срок наблюдений для янаря, апреля, июля и октября. Ход $m_1(i)$, естественно, аналогичен уточному ходу температуры, приводимому в Справочнике [7]. уточный ход дисперсии в январе и июле отвечает закономерэстям, указанным в [1, 2]. Зимой максимум дисперсии наблюается в предутренние часы, а минимум в дневные, летом наборот. В октябре суточный ход дисперсии выражен очень слабо. апреле, как уже указывалось, суточный ход дисперсии в течеие месяца претерпевает большие изменения, но в среднем он ринадлежит к летнему типу.

Рассмотрение значений асимметрии показывает, что в среднем аиболее асимметрично распределение температуры в январе

Таблица 1

Суточный ход средних за пентаду значений моментов распределения температуры воздуха (ГМС Ленинград, ГМО; аирель)

Срок, часы	Среднее						Дисперсия						Асимметрия					
	I	II	III	IV	V	VI	I	II	III	IV	V	VI	I	II	III	IV	V	VI
1	—1,8	0,4	0,9	2,1	2,8	4,8	23,7	14,1	10,5	9,5	11,9	14,2	—0,8	—0,4	—0,0	0,3	0,4	0,3
2	—1,7	0,1	0,7	1,9	2,6	4,6	23,3	14,1	10,3	9,2	12,4	13,6	—0,8	—0,5	—0,1	0,3	0,4	0,3
3	—1,3	—0,0	0,5	1,6	2,4	4,3	21,8	14,3	10,2	8,7	12,2	12,8	—0,8	—0,6	—0,1	0,3	0,4	0,2
4	—0,8	—0,2	0,4	1,6	2,2	4,2	19,9	15,1	10,2	8,6	11,8	12,0	—0,9	—0,6	—0,1	0,2	0,3	0,3
5	—0,1	—0,4	0,3	1,5	2,2	4,1	18,3	15,0	10,3	8,6	12,2	11,8	—1,0	—0,6	—0,1	0,3	0,3	0,3
6	0,4	—0,5	0,3	1,6	2,6	4,5	18,1	14,7	10,0	8,2	11,6	11,9	—1,2	—0,7	—0,2	0,3	0,5	0,3
7	1,0	—0,1	0,8	2,4	3,5	5,4	18,1	13,3	8,8	7,9	11,6	12,9	—1,1	—0,8	—0,2	0,3	0,7	0,1
8	1,6	0,6	1,5	3,3	4,5	6,4	16,6	11,4	8,7	9,3	13,9	16,6	—0,8	—0,8	0,1	0,5	0,8	0,1
9	2,0	1,3	2,4	4,2	5,5	7,6	15,0	10,1	8,7	11,2	16,5	21,4	—0,3	—0,6	0,4	0,6	0,7	0,2
10	2,2	2,2	3,3	5,1	6,4	8,6	14,3	10,4	10,2	13,5	19,7	25,6	—0,0	—0,3	0,6	0,6	0,7	0,2
11	2,2	2,8	3,9	5,8	7,1	9,4	15,1	11,5	11,5	15,7	21,6	29,2	0,1	—0,0	0,8	0,5	0,6	0,3
12	2,1	3,4	4,4	6,4	7,6	10,1	16,8	12,0	13,0	17,7	23,7	30,1	0,0	0,2	0,8	0,5	0,5	0,3
13	1,9	4,0	4,8	7,0	8,1	10,5	17,5	13,3	13,3	19,3	24,2	33,0	—0,1	0,4	0,8	0,6	0,4	0,3
14	1,5	4,1	4,9	7,1	8,4	10,5	18,2	14,3	14,5	19,7	25,3	32,0	—0,3	0,4	0,7	0,6	0,3	0,3
15	1,1	4,1	5,0	7,1	8,4	10,5	20,0	14,0	15,1	19,6	24,3	31,8	—0,5	0,4	0,7	0,7	0,3	0,3
16	0,7	4,0	5,0	7,0	8,3	10,5	20,8	13,9	15,2	19,5	24,1	30,8	—0,7	0,4	0,7	0,7	0,3	0,3
17	0,3	3,7	4,7	6,6	8,0	10,1	21,6	13,4	14,3	18,1	22,1	29,4	—0,8	0,4	0,6	0,8	0,3	0,3
18	—0,2	3,2	4,3	6,2	7,4	9,4	23,4	12,4	13,4	17,3	19,6	27,5	—0,9	0,4	0,6	0,9	0,3	0,4
19	—0,6	2,4	3,6	5,3	6,5	8,4	24,3	11,3	12,4	14,9	16,9	24,6	—0,9	0,4	0,5	0,9	0,4	0,3
20	—1,0	1,9	3,0	4,4	5,4	7,4	25,7	11,2	12,0	12,6	14,1	21,3	—0,9	0,3	0,4	0,8	0,3	0,3
21	—1,3	1,5	2,5	3,7	4,7	6,7	27,8	11,2	11,3	11,6	13,3	19,1	—0,8	0,3	0,3	0,7	0,4	0,3
22	—1,8	1,1	2,1	2,2	4,2	6,1	27,7	10,5	11,3	10,8	12,9	18,4	—0,7	0,0	0,3	0,7	0,4	0,4
23	—1,9	0,8	1,8	2,7	3,9	5,6	27,8	10,5	11,2	9,5	13,3	16,4	—0,7	—0,1	0,2	0,5	0,5	0,4
24	—1,8	0,5	1,6	2,3	3,5	5,3	25,7	10,7	11,2	8,9	12,7	15,2	—0,7	—0,2	0,1	0,3	0,5	0,3
Среднее	0,1	1,7	2,6	4,1	5,2	7,2	21,0	12,6	11,5	12,8	16,6	21,2	—0,6	—0,1	0,3	0,5	0,4	0,3

Суточный ход средних за месяц значений моментов распределения температуры воздуха (ГМС, Ленинград, ГМО)

Срок, часы	Среднее				Дисперсия				Асимметрия				Эксцесс			
	I	IV	VII	X	I	IV	VII	X	I	IV	VII	X	I	IV	VII	X
1	—7,0	1,5	14,9	5,1	42,6	14,0	7,5	13,2	—0,6	—0,0	0,1	—0,2	—0,4	—0,2	—0,2	—0,7
2	—7,1	1,4	14,5	4,9	42,6	13,8	7,4	13,2	—0,6	—0,1	0,1	—0,2	—0,4	—0,1	—0,3	—0,6
3	—7,1	1,3	14,2	4,8	42,6	13,3	7,5	13,3	—0,5	—0,1	0,0	—0,2	—0,4	—0,1	—0,3	—0,6
4	—7,2	1,2	14,1	4,7	42,8	12,9	7,5	13,3	—0,6	—0,2	0,0	—0,2	—0,4	—0,1	—0,3	—0,7
5	—7,2	1,2	14,4	4,6	43,3	12,7	7,2	13,3	—0,6	—0,1	0,1	—0,2	—0,4	0,0	—0,4	—0,6
6	—7,2	1,5	15,2	4,5	43,3	12,4	6,8	13,4	—0,6	—0,2	0,1	—0,2	—0,3	0,1	—0,5	—0,6
7	—7,1	2,2	16,5	4,5	43,2	12,1	7,9	13,4	—0,6	—0,2	0,1	—0,2	—0,3	0,2	—0,6	—0,6
8	—7,1	3,0	17,4	4,7	43,0	12,4	9,6	12,7	—0,6	—0,0	0,1	—0,2	—0,3	0,2	—0,6	—0,7
9	—7,1	3,8	18,3	5,1	43,1	14,0	11,4	11,7	—0,6	0,2	0,1	—0,2	—0,2	0,1	—0,5	—0,8
10	—7,0	4,6	19,0	5,7	42,5	15,6	13,0	11,0	—0,6	0,3	0,1	—0,2	—0,2	0,1	—0,5	—0,8
11	—6,7	5,2	19,6	6,3	40,8	17,4	13,9	10,9	—0,6	0,4	0,1	—0,1	—0,2	—0,0	—0,5	—0,8
12	—6,5	5,7	20,0	6,8	38,7	18,9	14,5	11,3	—0,6	0,4	0,1	—0,1	—0,1	—0,1	—0,5	—0,8
13	—6,2	6,1	20,4	7,1	37,0	20,1	14,8	11,8	—0,6	0,4	0,1	—0,1	—0,1	—0,1	—0,5	—0,7
14	—6,2	6,1	20,5	7,2	36,7	20,7	15,0	11,9	—0,6	0,3	0,2	—0,1	—0,2	—0,3	—0,5	—0,7
15	—6,3	6,0	20,5	7,2	37,1	20,8	14,6	12,0	—0,6	0,3	0,2	—0,1	—0,2	—0,2	—0,5	—0,6
16	—6,5	5,9	20,4	7,0	38,7	20,7	14,1	12,1	—0,6	0,3	0,2	—0,1	—0,2	—0,2	—0,5	—0,7
17	—6,7	5,6	20,0	6,5	40,0	19,8	13,0	11,9	—0,6	0,3	0,3	—0,1	—0,3	—0,1	—0,3	—0,7
18	—6,8	5,1	19,6	6,1	40,8	18,9	11,9	12,2	—0,6	0,3	0,3	—0,1	—0,3	—0,0	—0,3	—0,7
19	—6,8	4,3	19,1	5,9	41,8	17,7	11,0	12,4	—0,6	0,3	0,3	—0,1	—0,4	—0,0	—0,3	—0,7
20	—6,9	3,5	18,4	5,6	42,4	16,2	9,7	12,6	—0,6	0,2	0,3	—0,1	—0,4	—0,2	—0,3	—0,7
21	—7,0	3,0	17,2	5,5	43,1	15,7	8,1	12,7	—0,5	0,2	0,3	—0,1	—0,5	—0,3	—0,2	—0,7
22	—7,0	2,3	16,4	5,3	43,2	15,3	7,5	12,8	—0,5	0,2	0,3	—0,2	—0,5	—0,2	—0,1	—0,7
23	—7,0	2,2	15,8	5,2	43,4	14,8	7,2	13,0	—0,5	0,1	0,2	—0,2	—0,5	—0,3	—0,2	—0,7
24	—7,0	1,9	15,3	5,0	43,5	14,1	7,2	13,3	—0,6	0,1	0,2	—0,2	—0,4	—0,3	—0,2	—0,7
Среднее	—7,7	3,1	17,7	4,8	41,6	16,0	10,8	12,5	—0,6	0,1	0,2	—0,1	—0,3	—0,1	—0,4	—0,7

(хотя, как видно из табл. 1, в апреле в отдельные пятиднев значения коэффициента асимметрии достигают еще больших абсолютной величине значений, чем в январе). Суточный х асимметрии в январе отсутствует. В июле и октябре асиммет ность распределений незначительна. В июле в вечернее время блюдается небольшая положительная асимметрия, в осталы время суток распределение можно считать симметричным. В тябре суточный ход асимметрии отсутствует.

Средние за месяц значения коэффициента эксцесса д всех четырех месяцев практически в каждый срок наблюден отрицательны. Наибольшие по абсолютной величине значения эффициента эксцесса наблюдаются в октябре, при этом в течен суток они не меняются. В июле значения коэффициента эксце днем по абсолютной величине несколько больше, чем в н ные часы, а в январе — наоборот. Для апреля низкая точнос расчета эксцесса (средняя квадратическая ошибка определен среднего на каждый срок наблюдений значения коэффициен эксцесса приблизительно равна 0,25) не позволяет судить о з чимости полученных результатов.

3. Перейдем теперь к рассмотрению результатов, полученн для корреляционных моментов.

На рис. 1 для января, апреля, июля и октября представле зависимость от сдвига по времени τ средней за месяц коррел ционной функции

$$r(\tau)=\frac{1}{6}\sum_{j=1}^{6} r_2(\tau,\ j),$$

полученной осреднением корреляционных моментов как по с кам наблюдений, так и по пятидневкам. Из этого рисунка вид что чем теплее месяц, тем быстрее в среднем за месяц убыва корреляционная функция в течение суток.

Для апреля и июля даже осредненные корреляционные фун ции не изменяются монотонно, а имеют четко выраженные ма симумы при сдвигах времени τ, кратных суткам. Для октября э максимумы, отражающие влияние суточного хода температуры временную корреляцию ее, выражены гораздо слабее, а для я варя совсем не прослеживаются.

Однако при рассмотрении результатов, представленных рис. 1, следует иметь в виду, что в действительности условия ст ционарности как в течение месяца, так и в течение суток, мож считать выполняющимися лишь для января. Для остальных м сяцев имеет место нестационарность в течение суток, особен существенная для апреля и июля.

В табл. 3 для апреля и июля для каждого срока наблюден $i=1, 2, \ldots, 24$ приводятся средние за пентаду значения коэффиц ентов корреляции $r(i,\ i+1,\ j)$ между температурой в i-т

($i+1$)-й срок наблюдений, т. е. для сдвига по времени $\tau=1$ =1, 2, ..., 6). По методике [8] нами была найдена средняя квадратическая погрешность определения $r(i, i+1, j)$, которая для преля оказалась равной примерно 0,003, а для июля — 0,005. читывая эту погрешность, легко видеть, что условие стационарости коэффициента корреляции в течение суток не выполняется,

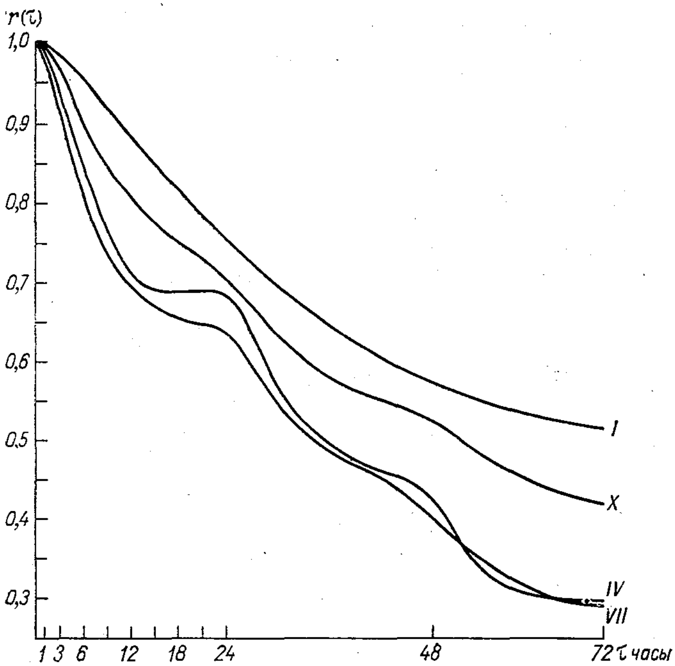


Рис. 1. Средние за месяц корреляционные функции температуры воздуха (ГМС Ленинград, ГМО).

суточный ход коэффициента корреляции существен. Действильно, несмотря на связность значений $r(i, i+1, j)$ для различх i, разность между максимальным и минимальным значениями $i, i+1, j$) внутри каждой пятидневки, например, для июля боe чем в 10 раз превышает среднюю квадратическую погрешсть определения коэффициента корреляции. В суточном ходе $i, i+1, j$) просматривается два минимума: первый, наиболее ыраженный,— в утренние часы, он хорошо связывается со временем восхода солнца (смещен по отношению ко времени восхоа примерно на 2—3 часа), второй — в часы перед заходом солн. Наибольших значений межчасовая корреляция температуры оздуха достигает в ночные часы.

Что же касается нестационарности коэффициента корреляции счет годового хода, то, как видно из табл. 3, в ночные и дневные часы разброс коэффициентов корреляции, полученных д различных пятидневок, лежит в пределах возможных погрешн стей за счет ограниченности выборки, а в часы, близкие к те когда в суточном ходе коэффициента корреляции наблюдает

Таблица

Средние за пентаду значения межчасовой корреляции температуры воздуха на ГМС Ленинград, ГМО (в тысячных долях секунды)

Срок, часы	Апрель						Июль					
	I	II	III	IV	V	VI	I	II	III	IV	V	V
1	990	992	993	990	990	995	986	990	991	992	992	98
2	989	996	995	990	988	991	979	989	986	989	990	99
3	983	995	994	987	990	994	985	989	987	991	994	99
4	976	994	995	993	993	993	964	979	984	990	987	99
5	985	997	994	987	989	985	928	936	940	955	969	97
6	987	990	979	952	959	962	927	939	936	940	931	93
7	983	984	973	957	964	977	955	974	977	975	978	97
8	979	979	963	973	977	980	976	977	982	982	986	98
9	983	977	976	983	989	984	983	984	977	990	987	98
10	974	978	983	985	988	987	973	978	980	989	988	98
11	978	975	986	989	987	993	981	971	980	985	986	98
12	984	979	979	986	990	990	986	981	970	981	988	98
13	987	989	985	989	985	981	974	980	980	976	989	96
14	985	987	991	988	986	987	959	975	970	983	984	97
15	991	991	992	989	988	991	968	958	977	973	975	97
16	991	985	991	989	988	990	974	966	978	982	968	96
17	987	989	989	989	987	990	981	971	956	986	972	98
18	991	972	982	986	969	989	977	972	962	985	975	97
19	994	988	982	981	986	987	974	974	981	988	986	98
20	996	989	986	984	987	988	963	968	962	975	969	97
21	989	981	990	984	991	991	964	973	976	985	978	97
22	994	990	992	974	991	991	970	970	985	987	988	98
23	986	991	992	970	991	996	983	983	981	989	987	98
24	988	992	991	983	990	990	979	987	986	991	988	98

минимум, разброс коэффициентов корреляции может превыша выборочные погрешности из-за смешения минимума в суточн ходе коэффициента корреляции (за счет сдвига времени восхо и захода солнца от пятидневки к пятидневке). Отсюда следуе что с целью уменьшения влияния выборочных погрешностей мо

Суточный ход средних за месяц коэффициентов корреляции r_1 $(i, i+\tau)$ температуры воздуха на ГМС Ленинград, ГМО (в тысячных долях единицы)

Срок i, часы	Январь						Апрель						Июль						Октябрь					
	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=6$	$\tau=12$	$\tau=24$	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=6$	$\tau=12$	$\tau=24$	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=6$	$\tau=12$	$\tau=24$	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=6$	$\tau=12$	$\tau=24$
1	997	992	985	953	888	721	992	977	959	871	619	645	990	972	956	741	533	623	994	984	972	928	785	671
2	997	991	983	951	887	714	992	976	955	821	599	666	987	973	950	642	510	628	995	984	970	935	766	653
3	997	991	982	950	886	715	991	973	952	772	585	670	988	967	891	591	490	629	994	983	971	931	752	639
4	997	991	982	952	883	715	991	973	926	727	586	676	983	906	750	584	496	636	994	984	974	909	764	639
5	997	990	982	953	878	718	990	945	877	688	594	685	950	814	724	634	534	655	994	985	974	868	780	632
6	996	990	981	954	875	727	972	914	845	692	630	697	934	865	820	764	663	699	995	984	962	829	795	626
7	997	989	982	951	875	733	973	921	866	734	714	702	972	938	914	852	760	684	992	971	928	790	795	622
8	997	991	985	956	880	743	975	936	894	787	770	697	981	959	936	864	790	641	987	952	906	800	818	648
9	997	992	984	957	881	751	982	951	918	817	790	691	985	961	937	858	805	622	983	947	909	838	836	682
10	997	991	983	956	881	762	982	955	919	851	782	692	982	958	930	859	804	621	983	955	923	895	847	708
11	997	991	985	955	877	769	985	956	929	888	778	697	982	959	938	871	790	640	986	964	948	913	846	738
12	997	992	985	949	877	783	985	964	943	912	762	708	982	960	931	866	773	634	988	974	962	928	820	745
13	997	991	982	946	870	787	985	969	957	912	729	707	978	949	924	866	742	624	990	979	969	907	797	746
14	997	989	979	947	869	792	987	975	961	898	728	701	975	950	922	877	691	609	992	980	965	895	787	756
15	996	989	978	950	875	798	990	978	959	872	713	698	972	940	913	873	673	608	990	973	946	884	770	765
16	996	988	980	954	882	795	989	973	949	851	703	704	975	942	920	839	669	621	988	963	941	887	775	773
17	996	990	984	955	887	793	989	963	931	836	693	714	975	944	938	810	716	621	987	969	953	900	787	771
18	997	992	984	953	890	780	985	954	927	828	726	709	975	966	930	780	793	641	991	980	966	907	798	754
19	997	991	984	952	891	772	984	960	931	832	773	704	982	946	888	751	820	630	994	983	969	907	802	735
20	997	991	983	951	890	763	988	966	940	865	788	686	968	914	858	743	782	642	994	982	965	903	822	725
21	997	991	983	953	890	758	987	966	935	869	758	666	980	934	898	811	717	651	993	979	960	904	843	613
22	997	991	981	953	891	742	989	965	938	874	723	657	980	952	920	865	651	638	993	979	963	915	853	717
23	997	990	981	953	891	729	988	965	944	874	695	640	984	958	939	889	590	620	993	981	967	923	846	710
24	996	990	982	954	890	724	989	971	950	883	674	643	986	970	949	878	558	615	994	984	972	931	832	701
Среднее	997	991	983	953	883	755	987	965	938	846	713	687	979	950	915	815	696	638	992	977	960	900	808	707

но производить осреднение коэффициентов корреляции, получен ных для разных пентад, но при этом нужно иметь в виду, что ми нимум в суточном ходе осредненных за месяц коэффициентов кор реляции будет сглаженным по сравнению с минимумами в отдель ных пятидневках.

К аналогичным выводам можно прийти, изучая суточный хо коэффициентов корреляции для других сдвигов по времени.

В табл. 4 на каждый срок наблюдений приводятся средние за месяц значения коэффициентов корреляции для различных сдви гов по времени. Из этой таблицы видно, что в январе суточног хода коэффициента корреляции практически нет, а в остальны месяцы при сдвигах во времени от 1 до 6 ч он хорошо просле живается.

Таким образом, приведенные результаты подтверждают еде ланный в [4] вывод о существенной нестационарности временно корреляции температуры воздуха. К сожалению, имеющихся дан ных недостаточно для построения сколько-нибудь надежной ана литической модели суточного хода характеристик ее временно структуры. Очевидна необходимость детального изучения это го вопроса на бóльшем материале с использованием данных на блюдений на станциях, расположенных в разных климатических зонах СССР.

СПИСОК ЛИТЕРАТУРЫ

1. Тихомирова Л. В. Опыт исследования статистических закономерностей распределения температуры воздуха в отдельные часы суток. — «Труды НИИАК», 1969, вып. 51, с. 92—107.
2. Лепехина Н. А., Федорченко Е. И. О суточном ходе дисперсии температуры воздуха. — «Труды ГГО», 1972, вып. 308, с. 99—109.
3. Киселева Т. Л., Чудновский А. Ф. Статистическое исследование суточного хода температуры воздуха. — «Бюлл. науч.-техн. информации по агрономической физике», 1968, № 11, с. 17—38.
4. Киселева Т. Л. Статистический анализ корреляционных функций. — «Бюлл. научн.-техн. информации по агрономической физике», 1968, № 11, с. 39—45.
5. Кендалл М., Стьюарт А. Теория распределений. М., «Наука», 1966. 587 с.
6. Каган Р. Л., Федорченко Е. И. О восстановлении годового хода моментов метеорологических рядов. — См. наст. сборник.
7. Справочник по климату СССР. Ч. II. Л., Гидрометеоиздат, 1965.
8. Каган Р. Л. О точности расчета пространственных корреляционных функций. — «Труды ГГО», 1973, вып. 308, с. 3—19.

Л. С. ГАНДИН, А. Е. ПРИГОДИЧ

О СТАТИСТИЧЕСКОМ КОНТРОЛЕ ВЕРТИКАЛЬНЫХ ПРОФИЛЕЙ ГЕОПОТЕНЦИАЛА

1. В работе [4] был предложен метод контроля вертикальных профилей геопотенциала, базирующийся на оптимальной интерполяции на контролируемый уровень по данным других уровней с последующим сравнением интерполированного и наблюденного значений. Опыт применения этого метода в рамках оперативных прогнозов, выполняемых совместно ГГО и СЗУГМС, показал, что метод обеспечивает сравнительно эффективный контроль информации о геопотенциале, а также позволяет исправлять грубо ошибочные значения и восстанавливать единичные пропуски данных. Вместе с тем анализ показывает, что статистический метод вертикального контроля геопотенциала может быть усовершенствован в нескольких направлениях.

а) В качестве норм $\bar{H}_k$ на каждом уровне p_k в работе [4] рекомендуется брать значения геопотенциала, осредненные по всем контролируемым станциям для рассматриваемого срока. Тем самым пренебрегается географической изменчивостью $\bar{H}_k$ и в частности их зависимостью от широты места. Пока контролю подвергаются данные в сравнительно небольшой области, такие единые нормы можно считать приемлемыми. Однако при переходе к большей области контроля принятие единых норм приводит к заметным погрешностям. Вместо таких норм целесообразно в процессе контроля пользоваться полями норм, взятыми из аэроклиматических данных.

б) Интерполяция, согласно [4], проводится по данным всех имеющихся уровней. При этом веса, с которыми входят значения H_i на уровнях, отдаленных от контролируемого, малы. Легко видеть, что на величины таких весов сильно влияют выборочные погрешности, с которыми известны компоненты матрицы автоковариаций геопотенциала

$$m_{ij} = \overline{H'_i H'_j} \tag{1}$$

(здесь черта сверху означает статистическое осреднение, а штрих— отклонение от среднего). Как правило, учет этих весов не влияе ни на результат, ни, следовательно, на качество контроля. Однак в отдельных случаях такой учет может искажать результаты конт роля. Целесообразнее поэтому ограничиваться интерполяцией п нескольким уровням, ближайшим к контролируемому, подобн тому как при объективном анализе в точку регулярной сетки про изводится интерполяция по данным лишь нескольких, близки к этой точке станций.

в) За неимением другой информации в работе [4] для опре деления интерполяционных весов a_{ik} была использована матриц ковариаций (1) не самого геопотенциала, а его межсуточных из менений. Значения a_{ik} так же, как и допустимые невязки контро ля Δ_k, определенные эмпирическим путем, считаются постоянны ми как по области контроля, так и во времени. Между тем известно, что ковариации m_{ij} и особенно дисперсии

$$\sigma_i^2 = m_{ii} \tag{2}$$

имеют заметный широтный и сезонный ход. Можно ожидать, что веса и допустимые невязки вертикального контроля также буду меняться с широтой и от сезона к сезону. За время, прошедшее после выполнения работы [4], накоплен обширный материал по межуровенной корреляции геопотенциала. Такие данные имеют ся, в частности, в работе [2]. Используя их, можно не только вы числить веса p_{ik}, но и оценить теоретическим путем допустимые невязки Δ_k для различных широтных поясов и сезонов года.

2. Рассмотрим подробнее реализацию приведенных соображе ний на примере схемы контроля данных 9 уровней p_k, а именно, 1000, 850, 700, 500, 400, 300, 200, 150 и 100 мбар при использова нии двух ближайших уровней для контроля данных каждого уровня.

Приведенная в работе [1] информация о значениях σ_i и о ко эффициентах межуровенной корреляции

$$\mu_{ij} = \frac{m_{ij}}{\sigma_i \sigma_j} \tag{3}$$

не содержит данных, касающихся уровней 400 и 150 мбар.

Кроме того, профили σ для разных сезонов и поясов имеют ряд особенностей, обусловленных, скорее всего, не реальными за кономерностями, а выборочными погрешностями. Поэтому эти профили были сглажены следующим образом. Отношения

$$\chi_{ij} = \frac{\sigma_i}{\sigma_j} \tag{4}$$

считались для широтных поясов 40—60 и 25—40° независимыми ни от широты, ни от сезона, а для пояса 0—25° независимыми от сезона. Значения σ на уровне 500 мбар (т. е. σ_4) брались непосредственно из [2], а значения на других уровнях p_i надились путем умножения σ_4 на χ_{i4}. В качестве последних принимались величины этих отношений, осредненные по обоим езонам и (кроме пояса 0—25°) по широтным поясам. Эти осредненные значения χ_{i4} приведены в табл. 1 наряду с исходными сглаженными профилями σ.

Таблица 1

Исходные и сглаженные вертикальные профили σ, профили отношения σ/σ_4 и значения меры ошибок наблюдений η^2

Пояс. °ш	Сезон	Поверхность p_k мбар (№)									η^2
		1000 (1)	850 (2)	700 (3)	500 (4)	400 (5)	300 (6)	200 (7)	150 (8)	100 (9)	
		Исходные значения σ_k дам									
—60	з	7,21	6,96	8,16	11,63		16,25	15,39		14,14	
	л	4,06	4,63	5,30	7,53		11,65	13,42		11,54	
—40	з	5,42	5,46	6,56	9,49		13,77	14,40		9,91	
	л	3,28	3,41	3,60	4,57		7,40	9,58		6,90	
—25	з	2,35	1,85	1,97	2,89		4,08	5,52		7,37	
	л	2,67	1,99	1,99	2,65		3,19	4,32		6,71	
		Сглаженные значения σ_k/σ_4									
—60		0,66	0,65	0,74	1,00	1,20	1,41	1,45	1,48	1,50	
—25		0,66	0,65	0,74	1,00	1,15	1,31	1,60	2,08	2,55	
		Сглаженные значения σ_k дам									
—60	з	7,6	7,5	8,6	11,6	13,9	16,3	16,8	17,2	17,4	0,02
	л	5,0	4,9	5,5	7,5	9,0	10,6	10,9	11,1	11,2	0,05
—40	з	6,3	6,2	7,0	9,5	11,4	13,4	13,8	14,1	14,2	0,03
	л	3,0	3,0	3,4	4,6	5,5	6,5	6,7	6,8	6,9	0,13
—25	з	1,9	1,9	2,1	2,9	3,3	3,8	4,6	6,0	7,4	0,32
	л	1,8	1,8	2,0	2,7	3,1	3,5	4,3	5,6	6,9	0,37

Помимо дисперсий σ_k^2 самого геопотенциала, для определения есов a_{ik} надо знать дисперсии δ_k^2 случайных ошибок его определения. Мы приняли меру ошибок наблюдения η_k^2, т. е. отношение их дисперсий

$$\eta_k^2 = \frac{\delta_k^2}{\sigma_k^2}, \tag{5}$$

зависящей от высоты. В соответствии с существующими оценками (см., например, [1]), для зимнего сезона в широтном поясе

40—60° было принято значение $\eta^2=0{,}02$, что соответствует средней квадратической ошибке δ_4 на уровне 500 мбар около 2 дам. В других поясах, а также для лета значения η^2 определялись в пред положении, что δ_k не зависит от широты и времени года, так что различия в η_k определяются полностью сезонными и широтными ходами σ_k.

Понятно, что распределения μ_{ij}, приведенные в [2], также было бы желательно сгладить, однако сделать это непросто. Вместе с тем можно ожидать, что хотя веса p_{ik} сильно зависят от коэффициентов μ_{ij}, результаты интерполяции, а потому и контроля мало изменятся с изменением μ_{ij}. По этим причинам величины μ_{ij} были взяты непосредственно из [2].

Они были дополнены значениями μ_{35}, μ_{45}, μ_{56}, μ_{57}, μ_{68}, μ_{78} и μ_{89} (т. е. коэффициентами, описывающими корреляцию геопотенциала поверхностей 400 и 150 мбар и нескольких ближайших к ним в принятой схеме), полученными путем интерполяции.

Коэффициенты μ_{35}, μ_{57} и μ_{68} определялись путем простой линейной интерполяции по формулам:

$$\begin{aligned}
\mu_{35} &= \frac{1}{2}(\mu_{34}+\mu_{36});\\
\mu_{57} &= \frac{1}{2}(\mu_{47}+\mu_{67});\\
\mu_{68} &= \frac{1}{2}(\mu_{67}+\mu_{69}).
\end{aligned} \tag{6}$$

Что касается коэффициентов μ_{45}, μ_{56}, μ_{78} и μ_{89}, описывающих связи уровней 400 и 150 мбар с соседними, то для определения этих коэффициентов пришлось использовать более сложный способ интерполяции. Именно, функция $\mu(x, y)$ (x и y — значения давления) аппроксимировалась симметричным полиномом третьей степени, обращающимся в единицу при $x=y$:

$$\mu(x, y)=1-A(x-y)^2-B(x-y)^2(x+y). \tag{7}$$

Коэффициенты A и B в формуле (7) вычислялись для определения $\mu_{i-1,\,i}$ и $\mu_{i,\,i+1}$ методом наименьших квадратов по данным о трех величинах: $\mu_{i-1,\,i+1}$, $\mu_{i-2,\,i-1}$ и $\mu_{i+1,\,i+2}$. Это привело к следующим формулам:

$$\begin{aligned}
\mu_{45} &= 0{,}48+0{,}16\,\mu_{46}+0{,}29\,\mu_{67}+0{,}07\,\mu_{34};\\
\mu_{56} &= 0{,}66+0{,}26\,\mu_{46}+0{,}12\,\mu_{67}-0{,}04\,\mu_{34};\\
\mu_{78} &= 0{,}71+0{,}16\,\mu_{79}+0{,}06\,\mu_{89}+0{,}07\,\mu_{67};\\
\mu_{89} &= 0{,}66+0{,}26\,\mu_{79}+0{,}12\,\mu_{89}-0{,}04\,\mu_{67}.
\end{aligned} \tag{8}$$

3. Как уже упоминалось, интерполяция в рассматриваемой схеме контроля производится по данным двух ближайших уровней. Для контроля данных всех поверхностей, кроме крайних

000 и 100 мбар), такими влияющими уровнями были поверх-
сти, расположенные непосредственно под и над контролируемой.
ри этом интерполяционные веса определялись по формулам:

$$a_{k-1,\,k}=\sigma_k\,b_{k-1,\,k};\qquad a_{k+1,\,k}=\sigma_k\,b_{k+1,\,k},\tag{9}$$

е $b_{k-1,k}$ и $b_{k+1,k}$ вычислены путем решения системы урав-
ний

$$\left.\begin{aligned}(1+\eta^2)\,b_{k-1,\,k}+\mu_{k-1,\,k+1}\,b_{k+1,\,k}&=\mu_{k-1,\,k};\\ \mu_{k-1,\,k+1}\,b_{k-1,\,k}+(1+\eta^2)\,b_{k+1,\,k}&=\mu_{k,\,k+1}.\end{aligned}\right\}\tag{10}$$

ера ошибки сопоставления проинтерполированного и наблюден-
го на уровне p_k значений определена по формуле

$$\varepsilon_k^2=1+\eta^2-b_{k-1,\,k}\,\mu_{k-1,\,k}-b_{k+1,\,k}\,\mu_{k,\,k+1},\tag{11}$$

средняя квадратическая ошибка такого сопоставления — по
рмуле

$$E_k=\sigma_k\varepsilon_k.\tag{12}$$

Для крайних уровней обе влияющие поверхности находятся
одну сторону от контролируемой. Соответственно изменены
рмулы (10) и (11).

Некоторые результаты этих расчетов сведены в табл. 2. В этой
блице применена упрощенная индексация, а именно, индекс 1
ответствует нижнему, а индекс 2 — верхнему из влияющих уров-
й. Значения E, умноженные на коэффициент $N=2{,}5$, приведен-
е в табл. 2, целесообразно, по-видимому, принимать в качестве
пустимых невязок. При нормальном законе распределения слу-
йных ошибок сопоставления это соответствует вероятности око-
99% того факта, что правильные данные не будут поставлены
д сомнение. Разумеется, можно выбрать и другое значение N.

Как видно из табл. 2, интерполяционные веса a_1 и a_2 для кон-
оля данных всех уровней, кроме крайних, мало меняются от
зона к сезону и с широтой места, и их можно заменить сред-
ми значениями $\bar{a}_1$ и $\bar{a}_2$, также приведенными в табл. 2. Что ка-
ется уровней 1000 и 100 мбар, для контроля которых произво-
тся не интерполяция, а экстраполяция, то здесь картина иная.
мой в высоких широтах вес ближайшего уровня превосходит
иницу, а вес следующего уровня отрицателен. При переходе
летнему сезону и к низким широтам различие между весами
еньшается. Это объясняется уменьшением дисперсии σ^2, кото-
е вызывает увеличение меры ошибок наблюдений η^2 (см. фор-
лу (5)).

Впрочем, для крайних уровней можно без существенного умень-
ния эффективности контроля выполнять его по данным лишь
ного, ближайшего уровня. В этом случае, как видно из табл. 2,

Таблица

Параметры девятиуровенной схемы статистического контроля профилей геопотенциала

Пояс, °ш	Сезон	Параметр	Поверхность, мбар										
			1000	1000	850	700	500	400	300	200	150	100	10
40—60	з	$a_1 \cdot 10^2$	90	125	51	57	53	46	80	51	57	−43	
	л		77	77	41	48	51	53	73	47	57	−11	
25—40	з		85	107	40	63	42	54	57	50	59	−2	
	л		65	52	30	53	56	57	62	41	52	12	
0—25	з		49	46	35	49	34	50	52	51	63	27	
	л		54	40	30	48	56	48	52	48	59	42	
		$\bar{a}_1 \cdot 10^2$	—	—	38	53	49	51	63	48	58	--	
40—60	з	$a_2 \cdot 10^2$		−32	46	44	53	50	32	52	48	135	93
	л			0	50	44	51	46	35	54	45	101	90
25—40	з			−23	54	35	58	46	40	50	40	90	91
	л			28	52	37	46	42	41	54	44	71	84
0—25	з			6	37	23	50	41	36	35	33	70	83
	л			18	44	32	41	38	32	36	33	66	90
		$\bar{a}_2 \cdot 10^2$		—	47	36	51	44	36	46	40	—	88
40—60	з	$\varepsilon \cdot 10^2$	47	44	31	20	25	24	23	25	21	37	40
	л		65	67	49	40	33	32	32	32	31	46	47
25—40	з		56	54	34	31	29	24	27	28	31	46	45
	л		82	75	63	54	49	47	50	49	48	62	60
0—25	з		101	101	89	91	81	72	72	72	73	85	75
	л		99	98	85	80	77	75	76	75	74	83	85
40—60	з	2,5 E дам	8,9	8,3	5,8	5,2	7,2	6,3	9,4	10,6	9,9	15,9	17
	л		8,4	8,4	6,0	5,4	6,2	7,1	8,5	8,6	8,5	12,8	13
25—40	з		8,9	8,6	5,2	5,5	7,0	6,8	8,9	9,5	10,8	16,5	16
	л		6,1	5,6	4,7	4,6	5,6	6,5	8,1	8,2	8,1	10,6	10
0—25	з		5,0	4,8	4,2	4,8	5,9	5,9	10,9	8,3	10,9	15,1	16
	л		4,5	4,4	3,8	4,0	5,2	5,8	10,4	8,1	10,4	14,2	14

можно и для контроля геопотенциала поверхности 100 мбар и пользовать не зависящий от сезона и широтного пояса весов множитель $\bar{a}_2$.

В отличие от весов допустимые невязки, как показыва табл. 2, существенно зависят от времени года и особенно от ш ротного пояса. Они уменьшаются от зимы к лету и от высок широт к низким. Из этого, однако, отнюдь не следует, что возмо сти вертикального статистического контроля геопотенциала льше летом и в низких широтах, чем зимой и в высоких широх. Как видно из табл. 2, значения меры ошибки сопоставления летом больше, чем зимой, а в низких широтах они приближася к единице. Последнее означает, что с помощью данного споба контроля удается в низких широтах отбраковать лишь данie, мало вероятные вообще, безотносительно к значениям на седних уровнях.

Все эти закономерности находят простое объяснение. Увеличее ε летом и в низких широтах обусловлено увеличением меры

Таблица 3

ıтрица величин $a_{ik} \cdot 10^2$ десятиуровенной схемы интерполяции по трем влияющим уровням. Широтный пояс 40—60°, зима

ıнтролируе-ıй уровень, мбар	Влияющие уровни, мбар									
	1000	850	700	500	400	300	250	200	150	100
1000		111	9	—24						
850	50		48	—2						
700	3	53		41						
500		—13	59		54					
400			—28	93		36				
300				—25	86		42			
250					—3	58		47		
200						—24	90		34	
150							—3	64		39
100							—28	—18	135	

шибок наблюдения η^2, которое, в свою очередь, вызвано уменьением дисперсии σ^2. При переходе же от ε к E решающим окаывается уменьшение σ, входящего множителем в формулы виа (12).

Так или иначе, целесообразно не пользоваться единым для сех сезонов и широтных зон набором допустимых невязок контоля геопотенциала изобарических поверхностей, а варьировать х в зависимости от времени года и в особенности от широтного ояса. Следует указать, что внесение такого усовершенствования алгоритм контроля не вызывает сколько-нибудь серьезных заруднений.

4. В рассмотренной схеме для контроля привлекались данные е более двух уровней. Расчеты показывают, что добавление еще дного или нескольких уровней не должно повысить точности онтроля. Так, в табл. 3 представлена матрица интерполяционных есов для случая, когда используются три влияющих уровня (как

правило, два ниже и один выше контролируемой поверхности Уже простое рассмотрение этой матрицы показывает, что вес на более удаленного уровня, как правило, либо отрицателен, либ весьма мал. Вычисление теоретических ошибок также показывае что эффект от добавления третьего влияющего уровня, как пр вило, мал. Однако для некоторых поверхностей учет дополнител ного уровня может несколько улучшить результаты контроля.

Что же касается использования при контроле еще большеı числа уровней, то опыты по применению метода в различных ег вариантах позволяют с несомненностью утверждать, что использс вание большого числа влияющих уровней не приводит к улучше нию результатов.

Таблица

Эмпирические и теоретические ошибки сопоставления (дам) интерполированных и наблюденных значений геопотенциала

Поверхность, мбар	Месяц	$\overline{\delta H}$	$\lvert\overline{\delta H}\rvert$	$0{,}8\ E$	$\lvert\delta H_{max}\rvert$	$2{,}5\ E$
1000	I	—0,1	1,5	2,7	5,2	8,3
	VII	—0,4	1,4	2,7	4,5	8,4
850	I	—0,1	0,8	1,9	3,1	5,8
	VII	0,2	0,7	1,9	3,0	6,0
700	I	0,4	1,0	1,6	3,9	5,2
	VII	0,1	0,9	1,7	3,1	5,4
500	I	—0,2	1,3	2,3	4,7	7,2
	VII	—0,2	1,0	2,0	4,3	6,2
300	I	—0,4	2,6	3,0	7,6	9,4
	VII	0,4	1,9	2,7	7,3	8,5

Представляет интерес рассмотреть наряду с теоретическим также и эмпирические значения ошибок сопоставления интерпо лированных и наблюденных значений геопотенциала. Соответст вующие данные для умеренных широт представлены в табл. 4 в виде средних с учетом знака, средних абсолютных, а также мак симальных по величине ошибок сопоставления. Для сравнения в табл. 4 приведены величины $2{,}5E$ из табл. 2, а также значения $0{,}8E$, которые при нормальном распределении ошибок соответст вуют их среднему абсолютному значению. Мы видим, что как средние, так и максимальные эмпирические ошибки заметно мень ше их теоретических оценок. Поэтому можно надеяться, что чув ствительность рассмотренного метода даже выше, чем это выте кает из данных табл. 2.

Некоторые результаты, касающиеся эффективности контроля потенциала с помощью описанного метода, приведены в рабо- [3].

В заключение укажем, что аналогичным путем могут строиться годы статистического контроля вертикальных профилей других теорологических элементов.

СПИСОК ЛИТЕРАТУРЫ

Гандин Л. С. Объективный анализ метеорологических полей. Л., Гидрометеоиздат, 1963. 288 с.

Мерцалова О. Б. Стандартизация вертикальных корреляционных связей давления. — «Труды НИИАК», 1972, вып. 79, с. 37—53.

Пригодич А. Е. Об одном методе контроля вертикальных профилей геопотенциала главных изобарических поверхностей. — См. наст. сборник.

Юдин М. И., Ильин Б. М., Руховец Л. В. Об одном способе контроля и исправления аэрологических телеграмм. — «Метеорол. и гидрол.», 1964, № 5, с. 35—39.

А. Е. ПРИГОДИ

ОБ ОДНОМ МЕТОДЕ КОНТРОЛЯ ВЕРТИКАЛЬНЫХ ПРОФИЛЕЙ ГЕОПОТЕНЦИАЛА ГЛАВНЫХ ИЗОБАРИЧЕСКИХ ПОВЕРХНОСТЕЙ

1. В статье [8] был предложен статистический способ контроля правильности данных геопотенциала стандартных изобарических поверхностей, основанный на применении метода оптимальной интерполяции по вертикали; будем называть его далее вертикальным контролем (ВК). Некоторые вопросы реализации ВК рассматривались также в работах [1—5]. В Ленинградском бюро погоды накоплен значительный опыт его применения, тем более ценный, что этот способ не получил такого широкого распространения как, например, горизонтальный контроль (ГК) данных геопотенциала [1]. В данной работе мы обсудим некоторые вопросы практической реализации ВК и полученные при этой реализации результаты. Затем будет рассмотрен усовершенствованный метод контроля, названный условно вертикально-потенциальным контролем (ВПК), а также результаты его применения в качестве единственного способа контроля и исправления данных о вертикальном распределении геопотенциала. Подобная ситуация возникает по существу при последовательном соединении нескольких способов контроля, что широко применяется в настоящее время, и при применении комплексного горизонтально-вертикального метода контроля [5] в районах, слабо освещенных наблюдениями. ВПК так же как и ВК, может быть использован для дополнительного контроля геопотенциала в кустовых пунктах обработки данных радиозондирования атмосферы [10], а также данных спутникового термического зондирования атмосферы.

Напомним, что ВК основывается на сравнении величин невязок δ_i, рассчитанных по данным проверяемой станции со статистически найденными допустимыми величинами невязок Δ_i, причем

$$\delta_i = H'_i - H'_{iH}, \qquad (1)$$

е H'_i— отклонение от нормы i-того уровня, определенное по данм, переданным со станции; H'_{iH} — проинтерполированное отклоние от нормы, определенное по данным соседних уровней этой станции.

Применительно к особенностям ВК ошибки в данных геопотенала удобно разделить на два класса: простая ошибка (в сводке еется только одно ошибочное значение геопотенциала) и сложя ошибка (имеются ошибки в данных геопотенциала более чем ного уровня).

Рассмотрим зависимость между величиной простой ошибки реакцией ВК (приводимые ниже величины ошибок характерны я уровней 850—500 мбар). Мелкие ошибки размером менее -4 дам пропускаются контролем. Диапазон ошибок 3—8 дам вывает появление невязки больше допустимой только на том уров, где имеется ошибка. Ошибки 5—15 дам вызывают появление $>\Delta_i$ сразу на двух-трех уровнях и возникает проблема опреления уровня, содержащего ошибку, ибо интерполяционные авнения допускают несколько вариантов исправлений. Ошибки, льшие 15—20 дам, вызывают появление невязок, превышающих итические, на трех — шести уровнях, но проблема выбора уров с ошибкой решается обычно без затруднений.

В случае наличия в данных сложной ошибки общую законорность вывести затруднительно. Наиболее неприятным является риант, когда имеем две ошибки на уровнях, близких к крайним, учайно согласованные по величине; при такой ситуации может ть пропущена крупная ошибка.

Таким образом, ВК требует наличия определенного алгоритма инятия решения (АПР). В программе для ЭВМ, составленной М. Ильиным [8], в основе АПР лежит способ автоматического следовательного перебора попыток исправления сомнительных овней в направлении от нижних уровней к верхним. Можно полгать, что такой АПР не сможет полностью решить задачу разления ошибок в данных на простые и сложные, а также задачу ределения уровня в случае простой ошибки порядка 10 дам.

2. Оценки контроля данных геопотенциала шести уровней при мощи ВК проводились по двум выборкам. Первая содержала выше 20 тыс. телеграмм за 1964—1966 гг. При этом были взяты едующие значения предельных невязок Δ_i и предельных велин допустимой ошибки исправления геопотенциала $\overline{\Delta_i}$:

ровень, мбар	1000	850	700	500	300	200
дам	9	4	6	8	9	27
дам	5	4	5	5	6	12

начения Δ_i существенно увеличены по сравнению с первоначальым вариантом [8], что объясняется «необходимостью сохранеия информации в малоосвещенных районах» [7]. Исправления еопотенциала оказались верными (в соответствии с допусками

$\overline{\Delta_i}$) в 81% случаев; неверные исправления наиболее часто про водились на уровне 850 мбар. Восстановление отсутствующих з чений было правильным в 94% случаев. Правильно выполнена браковка данных в 58% случаев. Число невыявленных ошиб геопотенциала составило около 1% [7]; отметим, что последн величину трудно оценить точно.

Вторая выборка содержала 21 тыс. телеграмм за все меся 1971 г. Каждый случай содержал в среднем 200 станций. Из эт го количества ВК исправлял и восстанавливал данные на 20 ста

Таблиц

Число исправлений простых ошибок на отдельных уровнях вертикальным контролем

	Уровень, мбар						
	1000	850	700	500	300	200	Все
Верные исправления	165	209	154	183	178	346	12
Неверные исправления	63	88	47	54	62	180	4
Исправлен уровень $i+1$	33	0	1	3	0	—	
Исправлен уровень $i-1$	—	26	41	0	0	19	

циях (10%) и отбраковывал на трех (1,5%). Однако разобра удалось данные только 2040 сводок, подвергшихся воздейств контроля. Допуски были взяты следующими:

Уровень, мбар	1000	850	700	500	300	2
Δ_i дам	9	4	7	8	10	
$\overline{\Delta_i}$ дам	3,5	2,5	3,5	4,5	5,5	

Небольшое увеличение Δ_i и уменьшение $\overline{\Delta_i}$ объясняется накопл нием соответствующего эмпирического материала.

Приведем выводы об эффективности работы ВК в отношен сложных ошибок. Всего обнаружено (разными способами) 294 т кие ошибки. Рассматриваемый вариант метода ВК, примененн изолированно от других методов контроля, правильно отбракові вает 76% сложных ошибок, пропускает 24% и еще отбраковыва простые ошибки в количестве, равном 30% от числа сложн ошибок.

В табл. 1 приведены результаты исправления данных, которы АПР контроля идентифицировал как содержащие простые оши ки. Из данных таблицы следует, что ВК исправил правильно 71 случаев простых ошибок. Эта величина значительно хуже, чем п лученная в работе [8]. Объяснить это можно следующими обсто тельствами. При применении ВК возникает проблема выбор ровня, содержащего ошибку (см. п. 1); эта причина снизила спешность исправления на 7%. На 6% ухудшился результат иза проскальзывания сложных ошибок, принимавшихся за простые, ошибочного исправления правильных данных. Оказали влияние акже применение существенно больших величин Δ_i, применение есов теплого периода для других сезонов года и недостаточно очное задание норм геопотенциала.

Из ряда соображений следовало бы ожидать наибольшую точость исправления простых ошибок на уровне 850 мб, но данные абл. 1, как и результаты оценки первой выборки, не подтверждают этого. Именно на этом уровне имеет место наибольшее число шибочных исправлений вследствие принятия неправильного реіения об уровне с ошибкой. Причина состоит в том, что ошибка H_{1000}, равная 5—11 дам, может вызывать невязку $|\delta_{1000}| < \Delta_{1000}$ одновременно $|\delta_{850}| > \Delta_{850}$. Подобное явление имеет место и на райних верхних уровнях при ошибке H_{200} в 10—30 дам. Таким обазом, причиной явилась рассогласованность (несоответствие) веичин Δ_{1000} и Δ_{850} (Δ_{300} и Δ_{200}).

При сравнении данных третьей и четвертой строк в табл. 1 особенно, если учесть, что первая величина в третьей строке меет своим происхождением другую причину) видно большое реобладание количества исправлений данных более низких уровей. Это связано с недостатком используемого АПР (именно направлением перебора уровней снизу вверх). Итак, были обаружены недостатки ВК, не носящие принципиального характеа, и можно было надеяться на их частичное или полное устраіение.

Для целей контроля оперативных данных о геопотенциале меодом оптимальной интерполяции требуется знание норм геопоенциала. Анализ ряда упрощенных способов задания норм показал, что такие упрощения могут быть причиной неправильного контроля данных. В итоге было решено задавать нормы над всей ерриторией полушария в узлах географической сетки с шагами ю широте 10° и по долготе 20°. Такая величина шагов обеспечивает довлетворительное описание климатических центров действия атюсферы. При этом для описания одного поля норм требуется 190 ячеек памяти ЭВМ. В одной ячейке вместе с нормой геопотенциала должна содержаться и норма его изменчивости. Предвариельные оценки показали, что программа интерполяции норм обеспечивает расчет $\overline{H_i}$ в нижней тропосфере с точностью около 1 дам, а в верхней — около 3 дам. В отдельных точках на поверхностях 300 и 200 мбар из-за нелинейности поля норм погрешности определения их достигают 5—7 дам. Положение существенно облегчает тот факт, что ошибки интерполяции норм обычно плавно меняются от уровня к уровню.

3. Рассмотрим вопрос о применении закономерностей распределения потенциальной температуры Θ в атмосфере для целей контроля вертикального профиля геопотенциала.

Эта идея возникла, с одной стороны, из того факта, что в слу чаях, когда АПР вертикального контроля делает заключение об и правлении геопотенциала уровня, не содержащего ошибку, прои ходят серьезные искажения распределения толщин слоев, а сл довательно, и Θ. С другой стороны, С. Л. Белоусов высказал мыс о возможности использования закономерностей распределения при согласовании полей объективного анализа на разных уровня что важно для схем прогноза, чувствительных к вертикальн структуре исходного поля Θ.

Пользуясь значениями геопотенциала изобарических повер ностей, можно рассчитать для средней точки слоя среднее давл ние и среднюю температуру слоя, а значит и среднюю потенц альную температуру слоя Θ. Известно, что

$$\Theta = -\frac{g}{R}\zeta^{-\lambda}\frac{\partial H}{\partial \ln p}, \qquad (\ $$

где $\zeta = \frac{p}{1000}$, $\lambda = 0{,}286$.

Отсюда получаем формулу

$$\Theta_i = k_i \Delta H_i, \qquad (\ $$

где

$$k_i = \frac{g}{R}\left(\frac{\sqrt{p_i p_{i+1}}}{1000}\right)^{-0{,}286} \cdot \frac{1}{\ln p_i - \ln p_{i+1}}, \quad \Delta H_i = H_{i+1} - H_i \text{ (дам)}.$$

При расчете использовались следующие значения коэффицие тов k_i:

Уровень, мбар	1000	850	700	500	300	2
k_i	2 152	1,896	1,180	0,8775	1,2605	

Известно, что в слоях воздуха достаточно большой толщин вертикальный градиент Θ имеет ограниченные пределы измен ния, выход за которые может быть квалифицирован как следс вие ошибок в исходных данных. «Потенциальный контроль» данны геопотенциала (ПК) заключается в проверке того, удовлетворя ют ли данные на станции допустимым значениям четырех па раметров: H_i, Θ_i, $\delta\Theta_i$, $\delta^2\Theta_i$, где $\delta\Theta_i$ и $\delta^2\Theta_i$ — первая и вторая ра ности потенциальной температуры. Первый этап здесь представ ляет собой так называемый контроль по допустимым значения геопотенциала; практически оказалось удобным слить его с соб ственно потенциальным контролем. Определение верхней и ниж ней границ допустимых значений H_i произведено на основани материалов аэроклиматических справочников, атласов, карт [6 9 и др.]. Определение границ Θ_i, $\delta\Theta_i$, $\delta^2\Theta_i$ первоначально был сделано на материалах выборки из 650 верных зондов и зате уточнялось на материалах 10 тыс. сводок по территории Еврази сти Северной Америки, Африки, Атлантики севернее 30—5° с. ш. Если получались значения параметров, близкие к пре-льным, то данные о геопотенциале контролировались другими етодами. Принятые значения верхних и нижних границ пара-етров H_i, Θ_i, $\Delta\Theta_i$, $\Delta^2\Theta_i$ (для всего года) приведены в табл. 2. увствительность ПК к случайным изолированным ошибкам на ровнях 850, 700, 500 мбар несколько превосходит чувствитель-ость ВК и, по-видимому, близка к чувствительности статиче-кого контроля (СК). Так удавалось обнаружить ошибки $|\delta H_{850}| = 2$ дам, $|\delta H_{700}| = 3$ дам, $|\delta H_{1000}|$ и $|\delta H_{500}|$, равные 4 дам. есколько ниже чувствительность ПК к обнаружению ошибок ипа просчета слоя.

Таблица 2

раничные значения геопотенциала H_i, потенциальной температуры Θ_i и ее конечных разностей $\Delta\Theta_i$, $\Delta^2\Theta_i$ (за год)

омер уровня i	1	2	3	4	5	6
ровень, мбар	1000	850	700	500	300	200
$H_{i\,min}$ дам	—60	77	225	455	785	1035
$H_{i\,max}$ дам	60	180	334	605	995	1285
$\Theta_{i\,min}$ К	228	246	261	279	301	
$\Theta_{i\,max}$ К	323	326	334	345	363	
$\Delta\Theta_{i\,min}$ К		—0,5	0,5	2,4	3	
$\Delta\Theta_{i\,max}$ К		26	21	30	51	
$\Delta^2\Theta_{i\,min}$ К		—14,3	—10,6	—10		
$\Delta^2\Theta_{i\,max}$ К		13,5	21	36		

Вопрос о диагнозе небольших ошибок (порядка 2—2,5 дам) аэрологических телеграммах, насколько нам известно, до сих ор не рассматривался. ПК помогает обнаружить и сгладить или аже исправить такие погрешности.

В табл. 3 для примера приведен анализ данных одного радио-онда. Обращает на себя внимание значение $T°C$ на уровне 00 мбар, однако просто отвергнуть его нельзя. В строке ΔH_i—ΔH_{iT} приведены разности между значениями толщины слоев, приведенных в телеграмме и рассчитанных по значениям темпе-атуры на верхней и нижней границе слоев (ΔH_{iT}). Малые вели-ины разностей свидетельствуют о том, что данные удовлетво-ряют статическому контролю. Однако ход знака разностей районе уровня 700 мбар относится к случаю неблагоприятного очетания мелких погрешностей. Действительно, ПК сигнализи-рует о неблагополучии в данных геопотенциала в районе уровней 700, 500 мбар, так как значения $\delta\Theta_2$ и $\delta^2\Theta_1$ выше критических.

ВК дает повышенное значение невязки на уровне 700 мбар. Так
образом, при использовании ВПК будет принято решение о сгл
живании данных этого зонда.

4. Рассмотрим алгоритм принятия решения в случае сочет
ния вертикального контроля с потенциальным (ВПК).

Для контролируемой станции рассчитываются значения $\bar{I}$
H'_i, H'_{iH}, Θ_i, $\delta\Theta_i$, $\delta^2\Theta_i$. Далее идет поуровенная «проверка» б
ошибочности данных вертикальным и потенциальным контрол
ми по отдельности. Каждый контроль отмечает уровни, призн
ваемые сомнительными, и делает положительное или отрицател
ное заключение о данных, которое мы далее будем обознача
названием контроля с прибавлением знака плюс (данные верны
или минус (данные сомнительны).

Таблица

Анализ данных радиозондирования; станция 72816, 00 ч, 2 февраля 1974 г.

Уровень P_i мбар	1000		850		700		500		300		20
Высота H_i дам	—27,6		91,9		230,0		470		815		1082
Температура T_i °C	≈—20		—30,3		—23,1		—39,1		—44,7		—51
$\Delta H_i = H_{i+1} - H_i$ дам		119,5		138,1		240		345		267	
ΔH_{iT}		118		140		238		345,5		266,5	
$\Delta H_i - \Delta H_{iT}$		1,5		—2		2		—0,5		0,5	
Θ_i		257,2		261,8		283,2		302,7		336,6	
$\delta\Theta_i$			4,6		21,4		19,5		33,9		
$\delta^2\Theta_i$					15,8		—1,9		14,4		

Принят следующий АПР. При сочетании $ВК_+$, $ПК_+$ делае
вывод, что данные верны. При сочетании $ВК_-$, $ПК_\pm$ начинаем «по
пытки исправления» для уровней, отмеченных вертикальным конт
ролем как сомнительные. Предусмотрен и второй этап попыто
исправления — по уровням, отмеченным потенциальным контро
лем как сомнительные, или даже (если $ПК_+$) по всем уровням
При сочетании $ВК_+$, $ПК_-$ начинаем попытки исправления перебо
ром уровней, отмеченных потенциальным контролем, а при необ
ходимости — всех уровней. Последний вариант работы ВПК во
зникает при небольших ошибках или значениях геопотенциала
близких к экстремальным, и представляет собой по существу сгла
живание вертикального профиля геопотенциала.

Если попытки исправления не удались, то данные отбраковываются.

5. Рассмотрим результаты применения ВПК.

Вначале приведем значения использованных допустимых невязок вертикального контроля, входящего в состав ВПК

Уровень, мбар	1000	850	700	500	300	200
Δ_i дам	5,8	3,5	4,5	6	8	14

Как видим, удалось значительно уменьшить их по сравнению с применявшимися ранее (п. 2); при этом были использованы результаты работы [3].

Опыты показали, что для ВПК характерна значительная чувствительность к точности задания норм. Так, попытка использовать нормы геопотенциала за апрель для контроля данных за март привела к частым необоснованным отбраковкам данных уровня 200 мбар в высоких широтах. Применение норм января для контроля данных за декабрь и февраль не вызывает столь крупных нежелательных последствий, однако и здесь имеем увеличение числа необоснованных воздействий ВПК на данные. Можно сделать вывод, что в данной постановке задачи для ВПК необходимо применение своих норм геопотенциала за каждый месяц.

Вертикально-потенциальному контролю были подвергнуты 52 случая оперативных данных геопотенциала шести уровней (свыше 10 тыс. сводок) за январь, февраль, апрель, июль, октябрь, декабрь 1971 г. Было решено сделать общую оценку работы контроля без подразделения итогов по отдельным уровням. При этом ставилась оценка плюс, если исправление было верным, в соответствии с допусками $\overline{\Delta}_i$, или были отбракованы данные, содержащие сложную ошибку; в противном случае ставилась оценка минус; ставилась оценка плюс — минус, если была существенно уменьшена только часть сложной ошибки. Отдельно подсчитывалось число случаев сглаживания данных геопотенциала (признаки его: имеется реакция только у ПК; данные геопотенциала исправлены на величину 0,5—3 дам).

Непосредственному воздействию контроля подверглись данные 1375 телеграмм (13,5% общего количества). Из них признано верными 84,5% воздействий, неверными 12,5%; сглаживание было в 3% случаев. Для этих же случаев при ВК признано верными 67,7%, неверными — 32,3% воздействий на данные. Количество ошибок вследствие неточного определения уровня с ошибкой уменьшилось более чем в три раза по сравнению с ВК — от 55 до 15, но одновременно оно увеличилось до 40 (0,4% всех данных) за счет путаницы уровней при мелких ошибках (2—4 дам), которые ранее пропускались вертикальным контролем. Относительно малое улучшение достигнуто на уровне 200 мб. Помимо упоминавшихся ранее, укажем еще отдельные недостатки ВПК (в скобках указан % от общего числа сводок):

— отбраковка при наличии одной ошибки в данных (<0,2%);
— исправление данных при отсутствии ошибки (<0,1%);

— пропуск случаев, когда две «согласованные» ошибки вызы вают при работе ВПК появление третьей ошибки (<0,1%).

Ближайшими задачами развития метода контроля являются по нашему мнению, объединение ВПК с горизонтальным контро лем; использование норм геопотенциала и изменчивости для уточ нения допустимых невязок, параметров и весов; исследование ра боты ВПК в комплексе контролей СК—ВПК—ГК.

В заключение автору приятно поблагодарить Л. С. Гандин и С. Л. Белоусова за помощь и поддержку, оказанные при вы полнении этой работы.

СПИСОК ЛИТЕРАТУРЫ

1. Белоусов С. Л., Гандин Л. С., Машкович С. А. Обработка опера тивной метеорологической информации с помощью электронных вычисли тельных машин. Л., Гидрометеоиздат, 1968, 282 с.
2. Гандин Л. С. Об автоматическом контроле текущей метеорологической ин формации. — «Метеорол. и гидрол.», 1969, № 3, с. 3—13.
3. Гандин Л. С., Пригодич А. Е. О статистическом контроле вертикальны профилей геопотенциала. — См. наст. сб.
4. Гандин Л. С., Смирнова Л. В. Об автоматическом контроле оператив ной метеорологической информации. — «Труды ГГО», 1972, вып. 308, с. 60—73.
5. Гандин Л. С., Тарасюк В. В. О комплексном контроле аэрологической информации. — «Метеорол. и гидрол.», 1971, № 5, с. 3—9.
6. Заставенко Л. Г. Барическое поле тропосферы Северного полушария. — «Труды НИИАК», 1972, вып. 86, 239 с.
7. Орлова Л. С. К вопросу об оперативном испытании численных схем прог нозов карт барической топографии и приземного поля давления. — В Сб работ Ленинградской ГМО, 1966, вып. 3, Изд. Северо-Западного УГМС с. 284—298.
8. Юдин М. И., Ильин Б. М., Руховец Л. В. Об одном способе контроля и исправления аэрологических телеграмм. — «Метеоролог. и гидрол.», 1964 № 5, с. 35—39.
9. Аэроклиматический атлас Северного полушария. Т. 1. Л., Гидрометеоиздат 1961, 298 с.
10. Наставление гидрометеорологическим станциям и постам. Вып. 4, ч. 1Па Л., Гидрометеоиздат, 1973. 256 с.

Я. В. ПАРФИНЕВИЧ

ТРЕХКОМПОНЕНТНАЯ МОДЕЛЬ КОМПЛЕКСНОГО КОНТРОЛЯ ГЕОПОТЕНЦИАЛА И ТЕМПЕРАТУРЫ ИЗОБАРИЧЕСКИХ ПОВЕРХНОСТЕЙ

1. Введение

Проведенные под руководством Л. С. Гандина исследования 4, 5] показали, что к аэрологической информации о геопотен-иале h и температуре t можно применять три разных метода онтроля: горизонтальный и вертикальный статистические [2, 9] етоды и статический контроль h и t [2, 8]. Эти способы контро-я выявляют все типы ошибок, содержащиеся в аэрологических елеграммах. Если рассматривать применение в комплексе гори-онтального, вертикального статистического и статического мето-ов контроля по отношению к двум элементам h и t, то получа-тся пятикомпонентный контроль. Модель комплексного контроля ожно упростить, учитывая, что информация о качестве данных, олученная с помощью результатов вертикального статистическо-о контроля, в принципе не вносит много нового по отношению горизонтальному контролю h и t и статическому контролю этих лементов. С другой стороны, статический контроль обладает ядом преимуществ перед вертикальным статистическим конт-олем.

По этим соображениям мы ограничились трехкомпонентной мо-елью контроля.

Построению такой модели предшествовало сравнение чувстви-ельности трех компонент комплексного контроля.

Порог чувствительности статического контроля можно харак-еризовать множителем N_c, так как значения допустимых невязок ыражаются через $N_c\Delta_{i\,\text{доп}}^{i+1}$, где $\Delta_{i\,\text{доп}}^{i+1}$ — некие эталонные эмпириче-ки найденные [7, 8] значения допустимых невязок в слое между ровнями i и $i+1$.

Порог чувствительности горизонтального контроля геопотен-иала и температуры определяют множители N_h и N_t соответственно, так как допустимые невязки горизонтальной оптимально интерполяции удобно принять равными: $N_h\sqrt{\varepsilon+\eta}\sigma_h$, $N_t\sqrt{\varepsilon+\eta}\sigma_t$ где ε и η — мера ошибки оптимальной интерполяции и мера ошибок наблюдения, σ — средние квадратические отклонения соответствующих элементов.

Таблица

Классификация основных типов изолированных ошибок

Тип ошибки	Признаки ошибки по компонентам комплексного контроля
I. Ошибки геопотенциала или температуры, возникающие либо в каналах связи, либо в процессе кодирования или раскодирования аэрологической информации	Невязки горизонтального контроля h или превышают допустимые; невязки статического контроля в слоях, соседних с данным уровнем, в случае ошибок h выше допустимых, а в случае ошибки t — не меньше допустимых невязок при уменьшенном приблизительно вдвое, пороге чувствительности
II. Ошибки зондирования, возникающие в результате неисправности или неправильной регулировки зонда.	Начиная с какого-то уровня невязки горизонтального контроля h или t превышают допустимые значения. Невязки статического контроля меньше допустимых.
Ошибки местоположения как результат неверного восприятия индекса станции	Этот вид ошибок с точки зрения признаков аналогичен ошибкам зондирования, но невязки ошибок горизонтального контроля превышают допустимые начиная с самого нижнего уровня
III. Ошибки неправильного вычисления относительного геопотенциала	Невязки статического контроля в данном случае превышают допустимые значения; невязки горизонтального контроля h начиная с уровня, ограничивающего сверху данный слой, превышают допустимые значения. Невязки горизонтального контроля t меньше допустимых
IV. Ошибки неправильного вычисления высоты 1000 мбар или перенос ошибки вычисления относительного геопотенциала слоев, расположенных ниже самого нижнего рассматриваемого контролем уровня	Начиная с самого нижнего расчетного уровня невязки горизонтального контроля h превышают допустимые значения. Невязки горизонтального контроля t и статического контроля меньше допустимых

Статический контроль по сравнению с горизонтальным контролем h и t хуже реагирует на ошибки в значениях t, чем на ошибки в значениях h. Следовательно, ошибки t выявляются в основном горизонтальным контролем t и подтверждаются и уточняются результатами статического контроля. Ошибки в значениях h выявляются главным образом статическим контролем подтверждаются и локализуются горизонтальным контролем h.. роги чувствительности компонент комплексного контроля приты следующими: $N_c=0{,}75$ (по отношению к h) и $N_c=0{,}5$ (по ношению к t), $N_h=4$, $N_t=5$.

По результатам многокомплексного контроля представляется зможным определить причины возникновения ошибок. Это пожено в основу алгоритма принятия решения и исправления ибочной информации. Чтобы поставить оптимальный диагноз, обходимо располагать классификацией ошибок, содержащихся аэрологических телеграммах. В соответствии с сочетанием ошик в сводке можно разделить ошибки телеграммы на изолироннные[1], неизолированные и смешанные. Для каждого из расатриваемых трех видов контроля с точки зрения признаков, торые им свойственны, выделим четыре основных типа изоливанных ошибок (табл. 1).

2. Алгоритм комплексного контроля и исправления ошибок

Все станции, сообщения которых имеются в совокупности реданной информации, обрабатываются в две итерации последовательно двумя блоками: блок I — получение невязок по испольемым методам контроля, блок II — принятие решения о пральности информации и в случае ошибок исследование возможсти исправления на данной итерации. Далее будет показано, о хотя блок II исправляет ошибки главным образом по результам статического контроля, внесение в сводку изменения оизводится так, чтобы введенное исправление не противоречило ловию горизонтальной согласованности данного элемента. Поольку при горизонтальной интерполяции иногда допускаются убые погрешности, связанные с присутствием в данных влияюих станций ошибок, для устранения которых принимаемые ряде случаев меры неэффективны, заключение о правильности правления ложной информации сделать затруднительно. Поэто- оказалось целесообразным разделить совокупность всей инрмации на три класса: 1) правильная, 2) ложная, но уверенно правляемая и 3) неисправимая (правильной будем также счить информацию с внесенными исправлениями на этой итерации) сомнительная. Уточним, что на первой итерации исправляются авным образом изолированные и иногда смешанные ошибки; на орой итерации рассматриваются только сомнительные станции, ичем в горизонтальном контроле в качестве влияющей испольется информация, признанная правильной на первой итерации.

В блоке I для данной станции по всем уровням получаются вязки трех видов контроля. Первым выполняется статиче-

[1] Вообще говоря, понятие изолированности ошибки связано с применяемой тодикой контроля. По-разному, например, определяются изолированные ибки вертикального статистического и статического контроля. В нашем слуе изолированность определяется результатами статического контроля.

ский контроль. Определяются случаи с невязками, превышаюш ми допустимые, запоминается вектор невязок. В цикле по урс ням в начале производится горизонтальный контроль h с доп стимой невязкой, определяемой значением множителя $N_h=4$, а з тем горизонтальный контроль t при значении множителя $N_t=$ Горизонтальный контроль выполняется по-разному, в зависимос от невязок статического контроля соседних слоев (процеду конкретизации). Если обе эти невязки или одна из них прево ходит допустимые значения, а невязки горизонтального контро меньше допустимой при $N_h=4$, тогда понижается порог чувств тельности до значения множителя $N_h=2{,}2$. Если даже при знач нии $N_h=2{,}2$, при которой допустимая невязка близка к ошиб оптимальной интерполяции h, невязка горизонтального контро геопотенциала меньше допустимой, то аналогичная процедура вы выполняется и для температуры, причем порог чувствительност горизонтального контроля t понижается до $N_t=3{,}2$. Однако ино да, несмотря на процедуру конкретизации, имеет место такое н соответствие невязок статического и горизонтального контр ля, при котором статический контроль указывает на наличи ошибки, а горизонтальный контроль этого не подтверждает. В т ких случаях при условии достаточного количества влияющи станций (не меньше двух), в зависимости от величины невязк статического контроля можно принять решение о наличи ошибки. Отметим, что дальнейшее уменьшение допустимых нев зок горизонтального контроля до ошибки оптимальной интерп ляции является нецелесообразным, поскольку в этом случае р зультаты контроля в принципе не вносят новой информаци В построенной схеме уверенность в наличии ошибки всегда по тверждается независимо от значения множителя N путем поочс редного исключения по одной из влияющих станций и проверкі действительно ли есть ошибка.

После выполнения статического и горизонтального контрс ля запоминаются следующие характеристики:

1) вектор невязок статического контроля: $\delta=\{\delta_i^{i+1}\}_{i=1,\ldots,n-1}$,

где n — количество уровней;

2) вектор невязок (превосходящих допустимые) горизонталь ного контроля h: $\Delta h=\{\Delta h_i\}$, где

$$\Delta h_i=\begin{cases}0, & \text{если } |\Delta h_i|<\Delta h_{i\,\text{доп}},\\ \Delta h_i, & \text{если } |\Delta h_i|\geqslant\Delta h_{i\,\text{доп}};\end{cases}$$

3) вектор невязок (превосходящих допустимые) горизонталь ного контроля t: $\Delta t=\{\Delta t_i\}$, где

$$\Delta t_i=\begin{cases}0, & \text{если } |\Delta t_i|<\Delta t_{i\,\text{доп}},\\ \Delta t_i, & \text{если } |\Delta t_i|\geqslant\Delta t_{i\,\text{доп}}.\end{cases}$$

Для данной станции по значениям указанных векторов нетруд-о найти следующие интегральные характеристики:

$$\sum_h = \sum_{i=1}^{n} \delta_i, \quad \text{где } \delta_i = \begin{cases} 0 & \text{при } \Delta h_i = 0, \\ 1 & \text{при } \Delta h_i \neq 0, \end{cases}$$

аналогично для горизонтального контроля t и для статиче-кого контроля — Σ_t, Σ_c.

сли $\sum_h \neq 0 \vee \sum_t \neq 0 \vee \sum_c \neq 0$, то вызывается блок II.

Имея в виду специфику ошибок, встречающихся в аэрологиче-ких телеграммах, и особенности используемых методов контроля, ожно исходить из следующего:

1. Будем полагать, что при указанных порогах чувствитель-ости наличие ошибки всегда скажется на результатах хотя бы дного из видов контроля, или, по-другому, что вероятность слу-айного сочетания грубых ошибок, которые компенсируются при ыполнении отдельных видов контроля, пренебрежимо мала.

2. Исходя из того, что невязки горизонтального контроля раз-ыты (как полученные с помощью статистической интерполяции), удем классифицировать ошибки по результатам всех видов кон-роля, но количественные исправления будем вводить в основном о данным статического контроля.

3. Ошибки зондирования бракуются без исправления.

На первой итерации с помощью алгоритма принятия решения олученный для данной станции комплект невязок рассматрива-тся с точки зрения возможности наличия в телеграммах сле-ующих ошибок:

1. Ошибки в значениях геопотенциала:

а) на самом нижнем уровне $i=1$; если невязка статиче-кого контроля в самом нижнем слое превышает допустимое зна-ение, т. е. $|\delta_1^2| > N_c \Delta^2_{1\,доп}$ (в описываемой модели для статического онтроля относительного геопотенциала принят множитель $N_c = 0{,}75$), причем выполняются условия:

изолированности

$$\Delta t_1 = 0 \wedge (\Delta h_2 = 0 \wedge \Delta t_2 = 0 \vee |\delta_2^3| < \Delta_2^3),$$

корректности внесения исправления

$$|\Delta h_1 + \delta| < |\Delta h_1|,$$

где δ — предполагаемая ошибка, $\delta = \delta_1^2$, тогда в значение h_1 вно-ится исправление

$$h_1 := h_1 + \delta;$$

б) на промежуточных уровнях $i = 2, 3, \ldots, n-1$; если невязка ста-ического контроля в нижнем по отношению к данному уровню слое больше допустимой, т. е. $|\delta_{i-1}^{i}| > N_c \Delta_{i-1\,доп}^{i}$ и невязк статического контроля в верхнем соседнем слое больше пол вины (удобно, ради исключения пропусков ошибок, принять хо бы в одном из соседних с данным уровнем слоев $N_c<0{,}75$) доп стимой величины, т. е. $|\delta_{i}^{i+1}| > 0{,}5\, \Delta_{i\,доп}^{i+1}$, или, наоборот, $|\delta_{i-1}^{i}| > 0{,}5 \times \Delta_{i-1\,доп}^{i} \wedge |\delta_{i}^{i+1}| > N_c \Delta_{i\,доп}^{i+1}$, причем выполняются условия:

изолпрованности

$$\left|\delta_{i-1}^{i} + \delta_{i}^{i+1}\right| < N_c \left|\Delta_{i-1\,доп}^{i} + \Delta_{i\,доп}^{i+1}\right|,$$

корректности внесения исправления

$$|\Delta h_i + \delta| < \begin{cases} |\Delta h_i| & \text{при } \Delta h_i \neq 0, \\ 2\sqrt{\varepsilon + \eta}\,\sigma_h^i & \text{при } \Delta h_i = 0, \end{cases}$$

где

$$\delta = \left(\delta_{i}^{i+1} \cdot \Delta_{i-1\,доп}^{i} - \delta_{i-1}^{i} \cdot \Delta_{i\,доп}^{i+1}\right) / \left(\Delta_{i\,доп}^{i+1} + \Delta_{i-1\,доп}^{i}\right),$$

тогда в значение h вносится исправление $h_i := h_i + \delta$. В описывае мую процедуру 1б) целесообразно включить исправление ошибо температуры для тех случаев, когда погрешность сказалась лиш на невязках статического контроля, применяя более жестки критерий изолированности и определения именно ошибки темпе ратуры:

$$\left|\delta_{i-1}^{i}/B_{i-1} - \delta_{i}^{i+1}/B_i\right| < \left|\Delta_{i-1\,доп}^{i} + \Delta_{i\,доп}^{i+1}\right|.$$

Здесь B_i, а в последующем и A_i — коэффициенты статическог контроля (см. [2, 8]);

в) на верхнем уровне $i=n$; если невязка статического конт роля в самом верхнем слое превышает допустимое значение, т. е $|\delta_{n-1}^{n}| > N_c \Delta_{n-1\,доп}^{n}$, причем выполняются условия:

изолированности

$$\Delta t_n = 0 \wedge (\Delta h_n = 0 \wedge \Delta t_n = 0 \vee |\delta_{n-1}^{n}| < \Delta_{n-1\,доп}^{n},$$

корректности внесения исправления

$$|\Delta h_n + \delta| < |\Delta h_n|,$$

где $\delta = -\delta_{n-1}^{n}$, тогда в значение h_n вносится исправление

$$h_n := h_n + \delta.$$

2. Ошибки в значениях температуры:

а) на самом нижнем уровне, $i=1$; если невязка горизонтально го контроля температуры на самом нижнем уровне превышает до пустимое значение, т. е. $\Delta t_1 \neq 0$, и абсолютное значение невязк

тического контроля превосходит допустимую величину, т. е. $< N_c \Delta^2_{1доп}$ (подтверждение наличия ошибки температуры тическим контролем производится при $N_c=0{,}5$), причем полняются условия:
изолированности

$$\Delta h_1 = 0 \wedge (\Delta h_2 = 0 \wedge \Delta t_2 = 0 \vee |\delta_2^3| < \Delta^3_{2\,доп}),$$

корректности внесения исправления

$$|\Delta t_1 + \delta| < |\Delta t_1|,$$

е $\delta = \delta_1^2/B_1$, тогда в значение t_1 вносится исправление $t_1 := t_1 + \delta$;

б) на промежуточных уровнях, $i=2, 3, ..., n-1$; если невязка ризонтального контроля температуры на i-том уровне превыша- допустимое значение, т. е. $\Delta t_i \neq 0$, и абсолютные значения невя- к статического контроля в соседних слоях превышают допу- имые (множитель $N_c=0{,}5$) величины, т. е.

$$|\delta^i_{i-1}| > N_c \Delta^i_{i-1\,доп} \wedge |\delta_i^{i+1}| > N_c \Delta^{i+1}_{i\,доп},$$

ричем выполняются условия:
изолированности

$$|\delta^i_{i-1}/B_{i-1} - \delta_i^{i+1}/B_i| < N_c |\Delta^i_{i-1\,доп}/B_{i-1} + \Delta^{i+1}_{i\,доп}/B_i|,$$

корректности внесения исправления

$$|\Delta t_i + \delta| < |\Delta t_i|,$$

де

$$\delta = (\delta^i_{i-1} \Delta i^{i+1}_{доп} + \delta_i^{i+1} \Delta^i_{i-1\,доп}) / (B_{i-1} \Delta^{i+1}_{i\,доп} + B_i \Delta^i_{i-1\,доп}),$$

огда в значение t_i вносится исправление $t_i := t_i + \delta$;

в) на верхнем уровне, $i=n$; если невязка горизонтального онтроля температуры на n-м уровне превышает допустимое начение, т. е. $\Delta t_n \neq 0$, и абсолютное значение невязки стати- еского контроля в верхнем слое превышает допустимую величи- у, т. е. $|\delta^n_{n-1}| > N_c \Delta^n_{n-1\,доп}$, причем выполняются условия:
изолированности

$$\Delta h_n = 0 \wedge (\Delta t_{n-1} = 0 \wedge \Delta h_{n-1} = 0 \vee |\delta^{n-1}_{n-2}| < \Delta^{n-1}_{n-2\,доп},$$

корректности внесения исправления

$$|\Delta t_n + \delta| < |\Delta t_n|,$$

де $\delta = \delta^n_{n-1}/B_{n-1}$, тогда в значение t_n вносится исправление $t_n := t_n + \delta$.

Исправлению с помощью описанных алгоритмов 1 и 2 подвергаются лишь ошибки типа I (по основной классификаци Они составляют, как правило, большинство ошибок, так что пос вызова процедур 1 и 2 целесообразно пересчитать заново ко плект невязок (это, естественно, не требует выполнения гориз тальной интерполяции). Прежде чем перейти к обнаружен и исправлению ошибок просчета (т. е. ошибок типа П1 и IV, с табл. 1), следует исключить из дальнейшего рассмотрения в можность наличия ошибки зондирования. Исследование хар терных признаков погрешности зондирования привело к следу щему условию:

$$\sum_t \geqslant 3 \wedge \left(\sum_h + \sum_t\right) \geqslant 7 \wedge \Delta h_4 \neq 0 \wedge |\Delta h_5| > |\Delta h_4|, \qquad ($$

где значения 3 и 7, найденные эмпирически, соответствуют шест уровенной (850, 700, 500, 400, 300 и 200 мбар) модели контрол Если условие (1) выполняется, то телеграмма считается на да ной итерации неисправимой, если нет, то комплект невяз в дальнейшем рассматривается с точки зрения возможности н личия в телеграмме следующих ошибок.

3. Ошибки неправильного вычисления относительного геоп тенциала. Происходит проверка в цикле по слоям — нет ли невя ки статического контроля больше допустимой. Предположим, ч в слое $(i, i+1)$ невязка δ_i^{i+1} превышает допустимое значение, то да, если выполняется хотя бы одно из следующих условий:

$$\text{а)} \bigwedge_{j \geqslant i+1} \left|\Delta h_j - \delta_i^{i+1}\right| < 2\sqrt{\varepsilon + \eta}\,\sigma_h^i,$$

$$\text{б)} \quad \left|\delta_i^{i+1}\right| > 2 \cdot \Delta_{i\ \text{доп}}^{i+1} \wedge \left|\delta_{i-1}^{i}\right| < \Delta_{i-1\ \text{доп}}^{i} \wedge \left|\delta_{i+1}^{i+2}\right| < \Delta_{i+1\ \text{доп}}^{i+2},$$

считается, что имеет место ошибка неправильного вычисления о носительного геопотенциала и вносится исправление:

$$\bigwedge_{j \geqslant i+1} h_j := h_j - \delta_i^{i+1},$$

если же эти условия не выполняются, телеграмма на первой ите рации считается неисправимой.

4. Ошибки неправильного вычисления высоты 1000 мбар ил перенос ошибки вычисления относительного геопотенциала слоев расположенных ниже самого нижнего рассматриваемого в дан ной модели контроля уровня.

Если $\Sigma_h = n$ (т. е. невязки горизонтального контроля геопотенциала превышают допустимые на всех уровнях), тогда, если вы полняется условие

$$\frac{\sum\limits_{i=1}^{n} \left(1/\sigma_h^i\right) \left(\Delta h_i - \widetilde{\Delta h}\right)}{\sum\limits_{i=1}^{n} \left(1/\sigma_h^i\right)} < \frac{\sum\limits_{i=1}^{n} \left(1/\sigma_h^i\right) N_{h_{\text{кр}}} \sqrt{\varepsilon + \eta}\,\sigma_h^i}{\sum\limits_{i=1}^{n} \left(1/\sigma_h^i\right)},$$

$$\widetilde{\Delta} h = \left(\sum_{i=1}^{n} \left(1/\sigma_h^i\right) \Delta h_i \right) \Big/ \sum_{i=1}^{n} \left(1/\sigma_h^i\right),$$

и после преобразования

$$\sum_{i=1}^{n} \left(\Delta h_i - \widetilde{\Delta} h\right) / \sigma_h^i < N_{h\,\text{кр}}\, n \sqrt{\varepsilon + \eta}, \tag{2}$$

итается, что имеет место ошибка, и вносится исправление $=h_i - \Delta h_i$, где $i=1, \ldots, n$. В противоположном случае на дан-м этапе телеграмма считается неисправимой.

Таким образом, на первой итерации исправляются изолиро-нные ошибки всех типов (за исключением ошибок зондирова-я) и в ряде случаев смешанные ошибки. Запоминаются номера исправимых станций, т. е. станций с сомнительной информа-ей. На второй итерации рассматриваются лишь сомнительные анции, но при горизонтальном контроле выбираются в качест-влияющих те станции, которые признаны правильными на пер-й итерации. Кроме упомянутых процедур 1, 2, 3, 4 исправляют-те изолированные ошибки, исправление которых на первой ерации было невозможным из-за наличия ошибки в данных ияющих станций (не выполнялось условие корректности внесе-я исправления при втором просмотре выполняется).

5. Процедура исправления неизолированных ошибок.

В основу алгоритма исправления неизолированных ошибок по-жены следующие соображения. Система сообщенных телеграм-й n значений h и t может заключать в себе $2n$ ошибочных све-ний. Обозначим вектор ошибок $h: \delta x = \{x_1, \ldots, x_n\}$, а вектор ибок $t - \delta y = \{y_1, \ldots, y_n\}$, далее будем искать истинные значе-я $\hat{h}$ и $\hat{t}$ в виде: $\hat{h} = h - \delta x$ и $\hat{t} = t - \delta y$. Задача состоит в отыска-и векторов δx и δy. Для этого воспользуемся уравнением атики, расписанным для $(n-1)$ слоев. Получается система $(n-1)$ уравнения следующего вида:

$$h_{i+1} - h_i - A_i - B_i(t_i + t_{i+1}) = \delta_i^{i+1}. \tag{3}$$

ходя из того, что

$$\left|\hat{h}_{i+1} - \hat{h}_i - A_i - B_i\left(\hat{t}_i + \hat{t}_{i+1}\right)\right| = \left|d_i^{i+1}\right| < \Delta_{i\,\text{доп}}^{i+1},$$

стему (3) можно переписать с точностью до допустимых невя-к (полагая, что в среднем невязки статического контроля тинных значений $\hat{h}$ и $\hat{t}$ равны нулю $d_i^{i+1}=0$) следующим обра-м:

$$x_{i+1} - x_i - B_i(y_i + y_{i+1}) = \delta_i^{i+1}, \quad i = 1, \ldots, n-1. \tag{4}$$

Если количество ошибок равняется количеству невязок, п восходящих допустимые значения (это можно рассматривать і предельный случай исправляемости), то система (4) имеет точі решение при условии, что известна локализация (диагноз) оі бок. Локализацию ошибочных элементов будем считать известі по данным горизонтального контроля h и t, имея в виду, что дежность ее обеспечивается ранее описанной процедурой конк тизации. Решение системы линейных алгебраических уравнеі (4) удобно находить численными итерационными методами.

Рассмотрим i-тый уровень. Возможны три случая: $x_i \neq 0 \wedge y_i =$ $x_i \neq 0 \wedge y_i \neq 0$; $x_i = 0 \wedge y_i \neq 0$ (случай правильной информации $x_i =$ $\wedge y_i = 0$ на i-том уровне, естественно, исправлению не подвер ется), для каждого из которых выпишем соответствующие реі ния относительно x_i или y_i. Эти величины получим решени системы двух уравнений с двумя (или одним) неизвестными. Д i-того уровня соответствующие уравнения системы (4) мож записать в виде:

$$x_i - B_{i-1}\, y_i = \delta^i_{i-1};$$
$$-x_i - B_i\, y_i = \delta^{i+1}_i.$$

Учитывая, что невязки истинных значений $\hat{h}$ и $\hat{t}$ по модулю пр порциональны допустимым невязкам статического контро соответствующих слоев (что, как показано в [4], обобща и уточняет решение), получим:

в случае $x_i \neq 0 \wedge y_i = 0$

$$x_i = \left(\delta^i_{i-1}\, \Delta^{i+1}_{i\ \text{доп}} - \delta^{i+1}_i\, \Delta^{i+1}_{i-1\ \text{доп}}\right) / \left(\Delta^i_{i-1\ \text{доп}} + \Delta^{i+1}_{i\ \text{доп}}\right),$$
$$y_i = 0;$$

в случае $x_i \neq 0 \wedge y_i \neq 0$

$$x_i = \left(\delta^i_{i-1} B_i - \delta^{i+1}_i B_{i-1}\right) / (B_i + B_{i-1}),$$
$$y_i = -\left(\delta^i_{i-1} + \delta^{i+1}_i\right) / (B_i + B_{i-1})$$

(полагается $d^{iH}_i = d^i_{i-1} = 0$);

в случае $x_i = 0 \wedge y_i \neq 0$

$$x_i = 0,$$
$$y_i = -\left(\delta^i_{i-1}\, \Delta^{i+1}_{i\ \text{доп}} + \delta^{i+1}_i\, \Delta^i_{i-1\ \text{доп}}\right) / \left(B_i\, \Delta^i_{i-1\ \text{доп}} + B^i_{i-1}\, \Delta^{i+1}_{i\ \text{доп}}\right).$$

Как нетрудно заметить, формулы (6), (7), (8) справедлив для $i \neq 1$ и $i \neq n$. Для определения ошибок крайних слоев на был принят следующий способ.

Предположим, что $x_i \neq 0 \vee y_i \neq 0$, где $i = 1$ или $i = n$, и буде искать решение x_i, y_i в виде:

для самого нижнего слоя

$$x_1 = \alpha_1 \Delta h_1 \delta_1^2; \quad y_1 = \alpha_1 \Delta t_1 \delta_1^2,$$
$$x_2 = \alpha_1 \Delta h_2 \delta_1^2; \quad y_2 = \alpha_1 \Delta t_2 \delta_1^2; \tag{9}$$

для верхнего слоя

$$x_n = \alpha_n \Delta h_n \delta_{n-1}^n; \quad y_n = \alpha_n \Delta t_n \delta_{n-1}^n;$$
$$x_{n-1} = \alpha_n \Delta h_{n-1} \delta_{n-1}^n; \quad y_{n-1} = \alpha_n \Delta t_{n-1} \delta_{n-1}^n, \tag{10}$$

де α_i — некий коэффициент, отыскание которого гарантирует выолнение следующих равенств: $\hat{h} = h_i - x_i$, $\hat{t}_i = t_i - y_i$ ($\hat{h}$, $\hat{t}$ — истиные значения элементов).

Описанный способ есть не что иное как разброс невязок статического контроля крайних слоев пропорционально невязкам оризонтального контроля температуры и геопотенциала крайних ровней. Подстановка h_i и t_i с учетом (9) и (10) в уравнение статики для самого нижнего слоя дает

$$\delta_1^2 = d_1^2 + \left[\alpha_1 \Delta h_2 \delta_1^2 - \alpha_1 \Delta h_1 \delta_1^2 - B_1(\alpha_1 \Delta t_2 \delta_1^2 + \alpha_1 \Delta t_1 \delta_1^2)\right].$$

Іолагая, что $\hat{\delta}_1^2 = \delta_1^2(\hat{h}_1, \hat{h}_2, \hat{t}_1, \hat{t}_2)$, получаем

$$\alpha_1 = \frac{1}{\Delta h_2 - \Delta h_1 - B_1(\Delta t_1 + \Delta t_2)}, \tag{11}$$

ı для самого верхнего уровня —

$$\alpha_n = \frac{1}{\Delta h_n - \Delta h_{n-1} - B_{n-1}(\Delta t_n + \Delta t_{n-1})}. \tag{12}$$

Совокупность формул (6)—(12) окончательно определяет реıения задачи. Можно показать, что описанный способ является обобщением метода исправления изолированных ошибок. Итерационное решение задачи исправления неизолированных ошибок производится следующим образом. В начале каждой итерации по формулам (9)—(12) определяются значения ошибок x_i или y_i краевых уровней, затем, используя формулы (6)—(8), можно определить значения ошибок для i-того промеждуточного слоя, делая а каждой итерации сдвиг по слоям. В качестве первого приближения применяются значения $h_i^{(0)}$, $t_i^{(0)}$, $\delta_i^{i+1(0)}$, рассчитанные по данным телеграмм. В дальнейшем $t_i^{(\nu)} = t_i^{(\nu-1)} - y_i^{(\nu)}$, $h_i^{(\nu)} = h_i^{(\nu-1)} - x_i^{(\nu)}$ ı, если $\delta_i^{i+1(\nu)} < \Delta_{i\text{доп}}^{i+1}$ для всех $i = 1, \ldots, n-1$, счет прекращается. Продолжительность расчетов определяет наперед заданное максимальное количество итераций ($\nu_{\max} \simeq n-1$).

3. Результаты численных экспериментов

Программа комплексного контроля была реализована на ЭВМ М-222 в квазиоперативном режиме. На вход программы поступа ют в неупорядоченном виде аэрологические телеграммы, переупа кованные согласно программе первичной обработки С. Л. Бело усова [2]. На выходе получается информация о качестве теле грамм. В машинной памяти (МОЗУ 1) накапливалась информа ция о 360 зондах. Для каждого зонда выполнялся поиск и подбо

Таблица

Статистическое распределение ошибок

Тип	I						II	III	IV	Неизолированные ошибки		Смешанные ошибки
Элемент	h			t						нет данных	прочие	
Уровень	$i=n$	$i=2,\ldots,n-1$	$i=1$	$i=n$	$i=2,\ldots,n-1$	$i=1$						
% ошибок	12	26	2	6	4	8	6	4	2	18	8	4
Σ %	40			18			12			26		4

влияющих станций. Затем вызывался блок получения невязок и, если данные зондирования вызывали сомнение хотя бы по одной компоненте контроля на любом уровне, вызывался блок принятия решения.

Оптимизация программы заключалась, во-первых, в подборе соответствующих эмпирически определяемых значений и, во-вторых, в совершенствовании логики и последовательности внесения исправлений. Так, например, оказалось целесообразным в начале исправлять изолированные ошибки в значениях геопотенциала, а затем — в значениях температуры. Это явилось следствием того, что ошибки t хуже контролируются статическим контролем, и выполнение соответствующего критерия, более жесткого, чем использованный раньше в программе К. А. Семендяева [8], иногда давало сомнительные результаты.

Программа была испытана на материале 900 аэрологических телеграмм, поступивших с территории, по которой составлялись оперативные прогнозы в Северо-Западном УГМС (по схеме Б. М. Ильина — Л. В. Руховца) в июне 1973 г. Анализ материала показал, что примерно в 15% всех телеграмм содержалась по крайней мере одна ошибка. Статистическое распределение выявленных ошибок по данным 900 зондов представлено в табл. 2.

зультаты испытаний оказались вполне удовлетворительными. ік правило, хорошо исправлялись все изолированные ошибки, том числе и ошибки на крайних уровнях. В ряде случаев оказась, что зонды с кажущимися ошибками зондирования на первой ерации, на второй, когда в качестве влияющей привлекалась авильная информация, были признаны правильными. Вообще воря, браковать по горизонтальному контролю крайне неэффеківно. Об эффективности схем комплексного и последовательного этодов контроля можно судить по следующему сопоставлению:

1) количество ошибочных телеграмм, исправляемых комплекым контролем на первой итерации, составляет 8,6% всех телеамм или 57% всех ошибочных телеграмм (телеграмм хотя с одй ошибкой);

2) количество ошибочных телеграмм, исправляемых схемой омплексного контроля на второй итерации, равно 4,9% всех теэграмм или 33% всех ошибочных телеграмм;

3) количество неисправимых с помощью комплексного контоля ошибочных телеграмм составляет 1,5% всех телеграмм или)% всех ошибочных телеграмм;

4) количество ошибочных телеграмм, надежно исправляемых эи последовательном способе контроля, составляет 5,1% всех телеграмм или 34% всех ошибочных телеграмм;

5) количество ошибочных телеграмм, исправление которых ри последовательном способе контроля сомнительно (т. е. допукаются случаи не только пропуска, но и генерирования ошибок), эставляет 6,4% всех телеграмм или 36% всех ошибочных телеэамм;

6) количество неисправимых с помощью последовательной схеы контроля ошибочных телеграмм составляет 4,5% всех телеэамм или 30% всех ошибочных телеграмм;

7) количество данных, ложно забракованных горизонтальным энтролем геопотенциала, составляет 2% всех телеграмм, а колиэство данных, ложно забракованных горизонтальным контролем эмпературы,—5% всех телеграмм.

Как показано в работе М. О. Кричак [6], критерий отбраковки есть функция той доли информации, которой мы готовы покертвовать. Отсюда следует, что, жертвуя одинаковой долей инормации, в случае комплексного контроля можно повышать увствительность контроля. При принятых нами порогах чувствиэльности $N_c = \{0{,}75 \text{ или } 0{,}5\}$, $N_h=4$, $N_t=5$ количество забракоанной информации уменьшилось приблизительно в три раза по тношению к последовательной системе статического контроя и исправления, горизонтального контроля h и горизонтального онтроля t.

Не останавливаясь на очевидных преимуществах комплексноо контроля над последовательными системами контроля, обсуим некоторые вопросы, существенные для реализуемой схемы.

Исправление неизолированных ошибок наиболее надежно приенительно к ошибкам h и t на одном уровне (назовем их полуизолированными). Тогда из уравнения статики можно най ошибки в h и t как решение линейной системы двух уравнен с двумя неизвестными.

Вопрос, насколько эффективен и оптимален обший итерацис ный способ исправления неизолированных ошибок вместе с в становлением отсутствующей информации на крайних уровня применительно к оперативным условиям, остается неясным. И правление по итерационному способу в принципе менее коррек но, чем обычный способ исправления изолированных и полуизол рованных ошибок. С другой стороны, неизолированные ошиб встречаются в оперативных условиях весьма редко, а восстана ливать или исправлять такие ошибки целесообразнее примен тельно к редкой сети, где ценность информации возрастает.

Что касается ошибок зондирования, вполне мыслимы случа когда критерий (1) не выполняется, а в то же время имеет ошибка зондирования, например, на последних двух уровнях. П этому необходимо выделить случай ошибок зондирования, ск зывающихся менее чем на трех уровнях. Предполагается вмес (1) применять более жесткое условие:

$$\sum_t \geqslant 3 \wedge \left(\sum_h + \sum_t\right) \geqslant 7 \vee \Delta h_n \neq 0 \wedge \Delta t_n \neq 0 \wedge \left|\delta^n_{n-1}\right| < \Delta^n_{n-1\ \text{доп}},$$

где n — самый верхний уровень.

Упомянутая проблема тесно связана с самой реализацией про граммы. На вычислительной машине с бо́льшим объектом опер тивной памяти можно было бы производить расщепление на лож ную и правильную информацию не только по станциям, но и п уровням. Это обеспечило бы бо́льшую точность, но реализаци такого алгоритма на языке АЛГОЛ требует использования гр моздких процедур упаковки и распаковки.

СПИСОК ЛИТЕРАТУРЫ

1. Андерсон Т. Введение в многомерный статический анализ. М., 196 500 с.
2. Белоусов С. Л., Гандин Л. С., Машкович С. А. Обработка оператив ной метеорологической информации с помощью электронных вычислитель ных машин. Л., Гидрометеоиздат, 1968, 282 с.
3. Гандин Л. С. Об автоматическом контроле текущей метеорологической ин формации. — «Метеорол. и гидрол.», 1969, № 3, с. 3—13.
4. Гандин Л. С., Тарасюк В. В. О комплексном контроле аэрологической информации. — «Метеорол. и гидрол.», 1971, № 5, с. 3—9.
5. Гандин Л. С., Смирнова Л. В. Об автоматическом контроле оператив ной метеорологической информации. — «Труды ГГО», 1973, вып. 308, с. 60— 72.

Кричак М. О. К вопросу об автоматическом контроле данных измерений ветра в свободной атмосфере.—«Труды ГГО», 1970, вып. 267, с. 117—135.

Рубцов И. В. Автоматическая подготовка исходных данных для численного анализа и прогноза погоды. — «Труды ГГО», вып. 124, стр. 30—38.

Семендяев К. А. и др. Автоматическая раскодировка аэрологических телеграмм. — «Труды ГМЦ», 1967, вып. I, с. 72—79.

Юдин М. И., Ильин Б. М., Руховец Л. В. Об одном способе контроля и исправления аэрологических телеграмм. — «Метеорол. и гидрол.», 1964, № 5, с. 35—39.

Л. С. ГАНДИН

О КОРРЕЛЯЦИОННЫХ ФУНКЦИЯХ КАЧЕСТВЕННЫХ ПРИЗНАКОВ

1. Пусть метеорологический элемент x может принимать только два значения: x_1 и x_2, а метеорологический элемент y — только два значения: y_1 и y_2. Пусть далее по данным одновременных измерений элементов x и y получена следующая таблица совместной повторяемости (сопряженности) значений этих элементов:

y	x		Σ
	x_1	x_2	
y_1	p_{11}	p_{21}	p_{01}
y_2	p_{12}	p_{22}	p_{02}
Σ	p_{10}	p_{12}	1

Так что, например, p_{12} есть повторяемость совпадения значений $x=x_1$ и $y=y_2$, а p_{20} — повторяемость значения $x=x_2$ независимо от значения y.

Тогда для средних значений $\bar{x}$ и $\bar{y}$ элементов x и y легко получить формулы:

$$\bar{x}=p_{10}x_1+p_{20}x_2; \qquad \bar{y}=p_{01}y_1+p_{02}y_2, \tag{1}$$

для дисперсий $\sigma_x^2=\overline{x^2}-\bar{x}^2$ и $\sigma_y^2=\overline{y^2}-\bar{y}^2$ — формулы

$$\sigma_x^2=p_{10}\,p_{20}(x_2-x_1)^2; \qquad \sigma_y^2=p_{01}\,p_{02}(y_2-y_1)^2, \tag{2}$$

для ковариации $\mu=\overline{xy}-\bar{x}\bar{y}$ — формулу

$$\mu=(p_{11}\,p_{22}-p_{12}\,p_{21})\,(x_2-x_1)\,(y_2-y_1) \tag{3}$$

, наконец, для коэффициента корреляции $r = \frac{\mu}{\sigma_x \sigma_y}$ между эле-
ентами x и y — формулу

$$r = \frac{p_{11} p_{22} - p_{12} p_{21}}{\sqrt{p_{10} p_{20} p_{01} p_{02}}}. \tag{4}$$

В отличие от формул (1)—(3) выражение (4) не содержит ни
дной из величин x_1, x_2, y_1 и y_2. Таким образом, коэффициент кор-
еляции между переменными, каждая из которых может прини-
ать лишь два значения, не зависит от этих значений, а зависит
ишь от их совместных повторяемостей.

Разумеется, это свойство коэффициента корреляции наруша-
тся, если хоть одна из переменных может принимать более двух
начений.

При заданных величинах p_{10}, p_{20}, p_{01} и p_{02}, описывающих пов-
оряемости значений каждой из переменных x и y независимо от
начений другой из этих переменных, совместные повторяемости
$_{11}$, p_{12}, p_{21} и p_{22} связаны между собой четырьмя очевидными соот-
ошениями:

$$_{11} + p_{12} = p_{10}; \quad p_{21} + p_{22} = p_{20}; \quad p_{11} + p_{21} = p_{01}; \quad p_{12} + p_{22} = p_{02}, \tag{5}$$

з которых лишь три независимых. Поэтому при таком условии
юбая из величин p_{11}, p_{12}, p_{21}, p_{22}, r может быть выражена через
юбую другую из них. Например, в силу (4) и (5) имеют место
ормулы:

$$\left.\begin{aligned} p_{11} &= p_{10} p_{01} + r\sqrt{p_{10} p_{20} p_{01} p_{02}}; \\ p_{12} &= p_{10} p_{02} - r\sqrt{p_{10} p_{20} p_{01} p_{02}}; \\ p_{21} &= p_{20} p_{01} - r\sqrt{p_{10} p_{20} p_{01} p_{02}}; \\ p_{22} &= p_{20} p_{02} + r\sqrt{p_{10} p_{20} p_{01} p_{02}}, \end{aligned}\right\} \tag{6}$$

ыражающие совместные повторяемости через коэффициент кор-
еляции. В частности, при отсутствии корреляции ($r=0$) из (6)
олучаются простые соотношения:

$$p_{11} = p_{10} p_{01}; \quad p_{12} = p_{10} p_{02}; \quad p_{21} = p_{20} p_{01}; \quad p_{22} = p_{20} p_{02}, \tag{7}$$

видетельствующие о независимости переменных x и y.

Из формул (6) также следует, что коэффициент корреляции r
ожет обратиться в единицу только, если $p_{01} = p_{10}$ и, следова-
ельно, $p_{02} = p_{20}$. Именно, при $r=1$ выполняются равенства:

$$p_{11} = p_{10} = p_{01}; \quad p_{22} = p_{20} = p_{02}; \quad p_{12} = p_{21} = 0. \tag{8}$$

Аналогичным образом, при $r=-1$ выполняются равенства:

$$p_{12} = p_{10} = p_{02}; \quad p_{21} = p_{20} = p_{01}; \quad p_{11} = p_{22} = 0. \tag{9}$$

Если не выполняется ни одно из равенств $p_{01}=p_{10}$ $(p_{02}=p_{20}$ и $p_{02}=p_{10}(p_{01}=p_{20})$, то ни при каких значениях совместных повторяемостей p_{11}, p_{12}, p_{21} и p_{22} коэффициент корреляции r не может обращаться ни в $+1$, ни в -1.

2. Имеются три ситуации, при которых метеорологический элемент целесообразно рассматривать в качестве дихотомической переменной, т. е. переменной, могущей принимать всего два значения.

Первая ситуация возникает, когда в силу крайней неточности имеющейся информации о метеорологическом элементе — это обычно относится, прежде всего, к прогностической информации — целесообразно разделять все возможные значения элемента всего на две градации. Чаще всего в качестве таких градаций применяются градации «выше нормы» и «ниже нормы». Используются и другие способы разделения всего интервала значений элемента на две градации.

В такой ситуации величину r следует рассматривать в качестве приближенной оценки истинного коэффициента корреляции между двумя непрерывными переменными x и y. Эта оценка, как мы видели, не зависит от того, какие количественные значения x_1, x_2, y_1 и y_2 приписываются градациями обеих переменных. Но она, разумеется, зависит от того, каким образом производится разделение на градации.

Вторая ситуация обусловливается тем или иным конкретным приложением информации о метеорологическом элементе, а именно, таким, при котором возникает определенная опасность, если этот элемент выходит в ту или иную сторону за некоторое «пороговое» значение. В этих условиях целесообразно разделить все возможные значения элемента на две градации — выше и ниже порогового значения. Например, условия заморозков для растений характеризуют тем, что температура воздуха (или температура поверхности почвы) ниже некоторого значения $T_{кр}$, и вводят две градации $T<T_{кр}$ (наличие заморозка) и $T\geqslant T_{кр}$ (отсутствие заморозка). Штормовые условия, например для судоходства, характеризуют превышением некоторого критического значения $v_{кр}$ скоростью ветра v. В этом случае градация $v>v_{кр}$ соответствует наличию шторма, а градация $v\leqslant v_{кр}$ — его отсутствию.

В обеих рассмотренных ситуациях целесообразность разделения всех возможных значений метеорологического элемента никак не связана со свойствами самого элемента, а обусловлена либо точностью информации о нем, либо потребностями приложений. В отличие от этого, в третьей ситуации мы имеем дело с метеорологическим элементом, дифференциальная функция распределения которого имеет особенность (обращается в бесконечность) в нуле, т. е. с элементом, характеризующим явление, которое может не существовать. К числу таких явлений относятся, например, осадки, грозы, туманы. Часто о таком явлении действительно неизвестно ничего, кроме того, имеет оно место или г. Но и в тех случаях, когда мы обладаем более детальной ин-рмацией (например, измерениями интенсивности осадков), ределенный интерес представляет статистика рассматриваемой ременной как переменной дихотомической.

Во всех перечисленных ситуациях статистическую связь меж-значениями двух переменных удобно характеризовать коэф-щиентом r, который называют коэффициентом качественной рреляции.

3. Коэффициент качественной корреляции естественно привле-ть к рассмотрению и в случае рассмотрения связи значений ух дихотомических метеорологических элементов (или значений ного и того же элемента) в различных точках или в разные мо-нты времени. В этом смысле целесообразно говорить о функ-ях качественной корреляции — как автокорреляции, так и вза-ной корреляции. Для таких функций можно, по аналогии обычными корреляционными функциями, вводить гипотезы об нородности и изотропии, предполагая коэффициент качественной рреляции зависящим лишь от расстояния между точками. В бо-е узком смысле предположения об однородности и изотропии ли стационарности по времени) для функции качественной ав-корреляции означают также выполнение равенств:

$$p_{01}=p_{10};\quad p_{02}=p_{20};\quad p_{21}=p_{12}, \tag{10}$$

)и которых формула (4) приобретает более простой вид

$$r=\frac{p_{11}p_{22}-p_{12}^2}{p_{10}p_{20}} \tag{11}$$

соответственно упрощаются формулы (5).

Разумеется, вместо величины r можно для характеристики ространственной или временной связности поля дихотомическо-) элемента пользоваться и иными характеристиками, например, юбой из величин p_{11}, p_{12} или p_{22}. Иногда вводят и другие меры вязности. Так, полагая $x_1=y_1=0$, а $x_2=y_2=1$, получим для струк-урной функции $b=\overline{(x-y)^2}$ простое выражение $b=p_{12}+p_{21}$, кото-ое еще более упрощается при условиях (10), а именно $b=2p_{12}$. ля ковариации μ (см. формулу (3)) получим

$$\mu=p_{11}p_{22}-p_{12}p_{21},$$

ли при условиях (10)

$$\mu=p_{11}p_{22}-p_{12}^2.$$

Io сравнению со всеми этими характеристиками функция качест-енной корреляции обладает тем преимуществом, что она, во-пер-ых, не зависит от значений, присваиваемых градациям рассмат-иваемой переменной и, во-вторых, в известной мере аналогична широко распространенным корреляционным функциям непреры ных метеорологических элементов.

4. Приведем пример функции качественной автокорреляци В работе [1] исследована совместная повторяемость гроз в пун тах, расположенных на различных расстояниях друг от друга. І второй строке табл. 1 приведены полученные по данным этой р боты значения среднего числа A дней за месяц, в которые ос ществляется гроза в обоих пунктах, отстоящих друг от друга расстоянии ρ. Значению $\rho=0$ соответствует среднее число дней с грозой за месяц в одном пункте. Рассчитаем по этим да

Таблица

Функция качественной корреляции осуществления грозы в течение д

ρ км	0	45	75	125	175	250	350	450	600	73
A дни/месяц	8,8	5,3	3,9	2,8	2,3	1,9	1,5	1,3	1,2	1,
r	1	0,45	0,22	0,05	−0,03	−0,09	−0,16	−0,19	−0,20	−0,

ным функцию качественной корреляции гроз, понимая под x_1 ос ществление грозы в течение дня, а под x_2 — ее отсутствие. И формул (5), (10) и (11) легко получить формулу

$$r=\frac{p_{11}-p_{10}^2}{p_{10}(1-p_{10})}, \qquad (1$$

выражающую в однородном и изотропном случае коэффициен корреляции через вероятность p_{10} осуществления события в одном пункте и вероятность p_{11} осуществления его в обои пунктах. В нашем случае

$$p_{11}=\frac{A}{N}; \quad p_{10}=\frac{A_0}{N}, \qquad (1$$

где N — число дней в месяце. С учетом (13) можно переписат формулу (12) в виде

$$r=\frac{N}{A_0(N-A_0)}A-\frac{A_0}{N-A_0}, \qquad (14$$

связывающем r с A линейной зависимостью (величины N и A_0 о расстояния ρ, разумеется, не зависят). В нашем примере $N=31$ $A_0=8{,}8$ (в днях за месяц), так что формула (14) дает

$$r=0{,}16A-0{,}40, \qquad (15$$

где коэффициенты выписаны с точностью до двух значащи цифр. По формуле (15) рассчитана функция качественной корре ляции $r(\rho)$, значения которой приведены в нижней строке табл. 1

ы видим, в частности, что эта функция переходит через нуль и расстоянии ρ около 160 км, что соответствует среднему ра-усу грозоопасного очага.

Интересно, что при больших расстояниях r принимает не очень лые отрицательные значения.

Заметим, что полученную функцию $r(\rho)$ нельзя смешивать корреляционной функцией $\mu(\rho)$ числа дней с грозой [2]. По-едняя характеризует связь между числом дней с грозой в пунк-х, отстящих друг от друга на расстоянии ρ, независимо от того, уществляется гроза в обоих пунктах в один и тот же день или разные дни. Поэтому функция $\mu(\rho)$ затухает с ростом ρ гораздо едленнее, чем только что рассмотренная функция $r(\rho)$.

В заключение считаю приятным долгом выразить признатель-ость Л. Л. Брагинской за помощь при использовании получен-ых ею данных.

СПИСОК ЛИТЕРАТУРЫ

Брагинская Л. Л., Степаненко Г. А. О характеристиках совместной повторяемости гроз в различных пунктах. — «Труды ГГО», 1974, вып. 336, с. 68—74.

Лугина К. М., Масанова М. Д. Пространственная изменчивость числа дней с грозой. — «Труды ГГО», 1973, вып. 308, с. 145—158.

Л. С. ГАНДИН, Л. Л. БРАГИНСКА

ОБ ОЦЕНКЕ ЭКОНОМИЧЕСКОЙ ЭФФЕКТИВНОСТИ РАБО ПО УСОВЕРШЕНСТВОВАНИЮ МЕТОДОВ ОБЪЕКТИВНОГ АНАЛИЗА МЕТЕОРОЛОГИЧЕСКИХ ПОЛЕЙ

1. В последнее время большое внимание уделяется оценк экономической эффективности научных работ вообще и рабо в области метеорологии в частности. При постановке едва ли н каждой работы требуют оценить в денежных единицах эффект о внедрения ее результатов. Это требование предполагает, что ре зультаты намеченной работы известны заранее, хотя для боль шинства научных исследований, а именно, для наиболее важны из них, это, разумеется, не так. Но даже по уже выполненно научной работе оценить корректно ее экономическую эффектив ность весьма непросто. Методы, разработанные для этой цели не учитывают специфики исследований по метеорологии и много образия направлений этих исследований, базируются на ряде п меньшей мере спорных положений, а главное, предполагают зна ние величин, которые в действительности неизвестны.

Вместе с тем ясно, что наличие хотя бы грубых порядковы оценок эффективности научных работ в области метеорологи весьма желательно. Знание таких оценок было бы полезно, ра зумеется, не для подавления той или иной тематики — занятия которым, поскольку оно не требует никакой квалификации, под час злоупотребляют, а для выдвижения проблем, разработка ко торых сулит значительный экономический эффект в ближайшем или более отдаленном будущем.

Научные работы по большинству естественных наук можно подразделить на познавательные (их часто называют фундамен тальными) и прикладные. Грубо говоря, познавательные работы направлены на изучение существующих закономерностей, а при кладные — на приложение этих закономерностей к решению кон кретных хозяйственных задач, выходящих обычно за рамки дан ной науки.

Специфической особенностью метеорологии является наличие, и, точнее говоря, большой удельный вес еще одного класса бот, которые мы будем называть информационными. Это — раты, так или иначе связанные с получением, обработкой и обобщением метеорологической информации. К этой категории относится широкий класс работ, начиная от разработки аппаратуры я метеорологических измерений и кончая выдачей метеорологических прогнозов и климатологической информации[1].

Работы информационного класса развиваются, разумеется, тесном взаимодействии с познавательными и прикладными исследованиями. Разработки как средств измерения, так и методов работки, климатологического обобщения и прогноза базируются на результатах познавательных работ и вызываются в значительной мере потребностями как познавательных, так и прикладных работ. Тем не менее выделение информационных работ отдельную категорию целесообразно по ряду причин и, в частности, потому, что возможности и пути оценки экономической эффективности таких работ существенно иные, чем для исследований познавательного и прикладного направлений.

2. Если отвлечься от малосущественных частностей, то существующая (см. например [4]) методика оценки экономической эффективности научных работ сводится к следующему. Затраты на выполнение научной работы могут состоять из единовременных затрат E_0 и постоянных затрат. Если обозначить через dE/dt постоянные затраты, отнесенные к единице времени, то получим формулу

$$E = E_0 + \frac{dE}{dt} t, \tag{1}$$

где t — время, обычно исчисляемое в годах. Аналогичная формула

$$G = G_0 + \frac{dG}{dt} t. \tag{2}$$

может быть написана для выигрыша, получаемого после внедрения результатов научной работы: он состоит из единовременного выигрыша G_0 и постоянного выигрыша, «скорость» которого, т. е. выигрыш за единицу времени, составляет dG/dt.

Знания четырех величин E_0, dE/dt, G_0 и dG/dt, фигурирующих в формулах (1) и (2), достаточно для оценки экономической эффективности данной научной работы. Эту оценку можно производить разными путями, причем различия между ними непринципиальны. Можно, например, оценить «время окупаемости» работы t_0,

[1] Заметим, что действия, направленные на выдачу климатологической информации потребителю, иногда рассматривают как работы по прикладной климатологии. Это, разумеется, не так. Прикладными являются работы, в которых исследуется влияние метеорологических факторов на те или иные народнохозяйственные объекты или процессы.

для чего следует приравнять величины E и G и из полученного р
венства найти t_0:

$$t_0 = \frac{E_0 - G_0}{\frac{dG}{dt} - \frac{dE}{dt}}.$$

Можно вместо этого задать на основе тех или иных соображени
(или без таковых) нормативный период окупаемости[1] $t_н$ и оцени
при $t=t_н$ отношение выигрыша к затратам

$$\chi = \left.\frac{G}{E}\right|_{t=t_н}, \qquad (4)$$

которым чаще всего принято характеризовать эффективность н
учных работ.

Затраты на выполнение научной работы всегда носят единовр
менный характер, постоянная часть их равна нулю. Что касаетс
затрат на внедрение результатов работ, то они наряду с единовр
менной частью могут содержать также постоянные затраты, таки
например, как затраты на амортизацию внедренной аппаратур
или на оплату машинного времени, необходимого для реализаци
внедренной процедуры обработки данных наблюдений. Часто, о
нако, постоянные затраты при внедрении пренебрежимо малы ил
вообще отсутствуют, так что $dE/dt=0$.

Что касается выигрыша G, то он, в противоположность затр
там, чаще всего носит постоянный характер, т. е. единовременны
выигрыш отсутствует $(G_0=0)$.

При этих условиях формула (3) принимает совсем простой ви

$$t_0 = \frac{E_0}{dG/dt}, \qquad (5)$$

а выражение (4) для χ может быть переписано в виде

$$\chi = \frac{t_н}{t_0}, \qquad (6)$$

т. е. отношение выигрыша к затратам при нормативном период
окупаемости равно отношению последнего к фактическому перио
ду окупаемости рассматриваемой работы.

3. Трудности оценки экономической эффективности научных ра
бот относятся, конечно, не к выводу рассмотренных формул, кото
рый вполне элементарен. Столь же элементарна и непринципиаль
на детализация подобных формул, выполненная, например, в упо
мянутой выше работе [4]. Трудности, как практические, так и прин

[1] Часто принимают, например, что $\frac{1}{t_н}=0{,}12$ год$^{-1}$, что соответствует нор
мативному периоду окупаемости $t_н=8{,}3$ года.

пиальные, возникают при попытке корректно оценить величины затрат и выигрышей, входящие в правые части формул)—(5).

Часто думают, что оценки затрат на выполнение научных работ удностей не вызывают. В качестве таких оценок берут суммы, траченные на выполнение соответствующей научной темы. Они лючают деньги, фактически выплаченные исполнителям (заратная плата, командировочные средства) и затраты на приобрение и эксплуатацию оборудования. Иногда в эти суммы включают дополнительные, накладные расходы. Однако такой путь корректен по крайней мере по двум принципиальным причинам.

Во-первых, любая научная работа проводится не на голом сте, а на основе результатов других, выполненных ранее работ, особенности работ познавательного направления. В этом отношеи познавательные работы аналогичны производству средств оизводства. Экономический эффект познавательных работ проляется через внедрение выполненных на их основе работ информационного и прикладного направлений. Отсюда вытекает, что число затрат на выполнение данной работы нужно включать кже некоторую долю затрат на те работы — преимущественно знавательные,— без которых данная работа была бы невозможй. На выполнение каких именно работ и какую именно долю трат — на эти вопросы ответить чрезвычайно трудно. Ясно одно: бестоимость любой научной работы выше затрат на выполнение енно этой работы.

Во-вторых, верно ли, что заработная плата и другие перечисенные суммы представляют собой те потери, которые несет обество вследствие выполнения данной научной работы? Иначе воря, верно ли, что если бы данная работа не выполнялась, то работная плата была бы сэкономлена? Разумеется, это неверно: полнители все равно получали бы заработную плату, все равно асходовали бы командировочные средства и т. п. Но они в этом учае выполняли бы другую научную работу, внедрение результатов которой или результатов последующих, базирующихся на ей работ дало бы определенный экономический эффект. Именно от эффект — эффект от другой возможной работы, которой мог ы заниматься научный коллектив, если бы он не выполнял расматриваемой работы — и измеряет реальные потери на выполнение данной работы. Они гораздо больше или по крайней мере олжны быть гораздо больше, чем заработная плата. В противом случае работник не имеет морального права заниматься научными исследованиями и должен переключиться на другой вид еятельности.

Из сказанного следует, что, определяя задачи тем или иным аучным коллективам, планируя их деятельность, необходимо осоенно тщательно учитывать сложившуюся специализацию таких оллективов, круг их научных интересов, имеющиеся заделы. Это еобходимо для того, чтобы задачи, поручаемые научным коллективам, решались наиболее квалифицированно, с наименьшими з тратами и в кратчайшие сроки. Неоправданные частые перекл чения тематики не только вредны для развития науки и практи но и влекут за собой экономический ущерб, поддающийся колич ственной оценке.

Что же касается затрат на выполнение научных работ, то сказанного следует, во всяком случае, что, приравнивая их зарнл те исполнителей и другим подобным суммам, мы неизбежно п уменьшаем реальные затраты и тем самым преувеличиваем эко мическую эффективность научных работ.

4. Возможности и пути оценки выигрыша G, достигаемо в результате выполнения научного исследования, существен различны для трех перечисленных выше групп работ. Наибол трудно, разумеется, выполнить такую оценку для познавател ных работ. Напротив, перспективы экономического выигрыша внедрения результатов прикладных работ сравнительно легко по даются хотя бы грубой количественной оценке.

Что же касается работ информационного направления, то мо но указать лишь один, узкий класс таких работ, для которых в личину G оценить не очень сложно.

Это такие работы, направленные на усовершенствование м тодов получения или обработки информации, в результате кот рых состав, объем, точность и другие параметры выдаваемой и формации не меняются, а изменяются лишь способы получен или обработки данных. К этому классу относятся многие работ по автоматизации измерений, передаче и обработке данных. В игрыш от внедрения таких работ состоит в экономии людских р сурсов, затрачиваемых в случае отсутствия автоматизации. Э экономия сравнительно легко поддается оценке. Для корректно выполнения ее надо учитывать два обстоятельства. Во-первы оценка должна производиться для достаточно длительного пери да времени, следующего за моментом перехода от ручных проц дур к автоматическим. Во-вторых, необходимо учитывать, ч экономия людских ресурсов отнюдь не равна экономии на зар ботной плате, а, как уже говорилось, гораздо больше. Не принима во внимание этих обстоятельств, можно сильно занизить эконом ческий эффект автоматизации той или иной процедуры или даж прийти к выводу об ее убыточности в тех случаях, когда он в действительности прибыльна.

Однако подавляющее большинство работ информационного н правления не принадлежит к только что рассмотренному кла су. Внедрение большинства работ приводит к изменению пар метров информации: она оказывается либо более широкой, либ более точной, либо более своевременной. Такое усовершенств вание параметров информации позволяет уменьшить потери потр бителей, пользующихся метеорологическими данными. Это умен шение потерь потребителей и составляет выигрыш от внедрени соответствующей информационной работы.

К сожалению, все это крайне трудно поддается сколько-ни-дь обоснованным количественным оценкам[1]. Попытки таких ценок применительно к отдельным потребителям метеорологиче-ой информации предпринимались, но они в лучшем случае по-оляют представить себе порядок величины уменьшения потерь. ри этом надо иметь в виду, что для правильной оценки выигры-а нужно учесть потери всех возможных потребителей данной ин-эрмации, как тех, которым такая информация нужна сейчас, к и тех, которым она потребуется в будущем. Иначе мы неиз-жно занизим экономический эффект рассматриваемой работы.

Существует, однако, прием, с помощью которого можно иног-1 обойти все эти трудности. Этот прием можно назвать спосо-эм оцениваемого эквивалента. Он состоит в следующем.

Зададимся вопросом, не существует ли пути, с помощью ко-рого можно было бы достичь того же усовершенствования пара-этров информации, какого удается достигнуть на основании зультатов рассматриваемой работы. Если этот другой путь суще-вует, то попытаемся оценить затраты, связанные с его реали-цией. Этих затрат удается избежать, потому что выполнена ссматриваемая работа. Поэтому они характеризуют экономи-еский эффект этой работы.

5. Способ оцениваемого эквивалента применим, в частности, вычислению экономического эффекта работ по усовершенство-анию методов объективного анализа метеорологических полей. дея таких оценок весьма проста и состоит в следующем [1]. совершенствование методики анализа поля того или иного эле-ента приводит, естественно, к уточнению результатов этого ана-иза. Можно ли достичь такого же уточнения, применяя старую, е усовершенствованную методику анализа? Можно, но для этого ришлось бы соответственно расширить сеть станций, на которых роизводятся наблюдения. Затраты на функционирование допол-ительных станций, установки которых удается избежать благо-аря усовершенствованию методики анализа, и представляет собой арактеристику экономического эффекта такого усовершенст-эвания.

Рассмотрим конкретный пример.

Как известно, применяемая в настоящее время в Гидрометцент-е СССР методика объективного анализа полей геопотенциала ад северным полушарием [1] базируется на использовании ис-одных данных только о геопотенциале. В районах с густой сетью танций такая методика обеспечивает требуемую точность ана-иза, в то время как в областях с редкой сетью анализ недоста-эчно точен.

И. Клуге [2] разработал усовершенствованную методику объек-ивного анализа, включающую учет данных о ветре. Он показал, то внедрение этой методики позволит снизить среднюю квадра-

[1] Приходится подчеркнуть, что речь идет именно о сколько-нибудь обосно-анных оценках в отличие от оценок никак не обоснованных.

тическую ошибку анализа на 17%. С другой стороны, по оценк С. А. Машковича [3], такое же уменьшение ошибки анализа мог бы быть достигнуто при применении существующей методики ан лиза, если в дополнение к действующим станциям северного г лушария установить в нужных местах еще 25 станций.

Затраты по содержанию одной аэрологической станции в хор шо обжитых районах суши составляют около 50 тыс. руб. в г В пустынных районах и тем более на океанах (корабли погоды эта стоимость гораздо выше. Кроме того, существуют еще един временные затраты на установку станции. Пренебрегая всем этими дополнительными затратами, получим, что внедрение ус вершенствованной методики, предложенной в работе [2], да экономию

$$\frac{dG}{dt}=50\cdot 25=1250 \text{ тыс. руб. в год.}$$

Что касается затрат, связанных с выполнением данной работы, т нетрудно оценить следующие суммы:

1) стипендия, выплаченная исполнителю (И. Клуге проходи аспирантуру в СССР), составила

$$E_1=2{,}4 \text{ тыс. руб.};$$

2) часть заработной платы руководителя, соответствующа доле времени, которая была потрачена на руководство аспира том

$$E_2=0{,}6 \text{ тыс. руб.};$$

3) стоимость машинного времени, израсходованного в процесс выполнения работы:

$$E_3=3{,}0 \text{ тыс. руб.};$$

4) затраты, необходимые для внедрения усовершенствованно методики в оперативную практику:

$$E_4=0{,}8 \text{ тыс. руб.}$$

Суммирование дает

$$E=E_1+E_2+E_3+E_4=6{,}8 \text{ тыс. руб.}$$

Отсюда по формуле (5) находим время окупаемости

$$t_0=5{,}45\cdot 10^{-3} \text{ года}=2{,}0 \text{ дня},$$

а по формуле (6) — отношение выигрыша к затратам

$$\chi=1520,$$

т. е. каждый затраченный рубль дает выигрыш свыше 1500 руб лей.

Несомненно, что эта величина завышена по сравнению с дей ствительностью. Одна из причин завышения кроется в самой сущ ности способа оцениваемого эквивалента. Второй путь, фигури ующий в этом способе (в нашем случае это — организация до-олнительных станций), потому и не используется в действитель-ости, что он требует слишком больших затрат. Однако в приве-енных оценках это завышение нивелируется тем обстоятельством, го мы пренебрегли единовременными затратами на установку ганций, а также исходили из стоимости содержания стан-ий на суше, в то время как в действительности большинство ополнительных станций пришлось бы организовывать на аква-ориях океанов.

Более существенна другая причина завышения оценок. Она остоит в том, что в действительности, как об этом подробно го-орилось выше, затраты на выполнение научной работы гораздо ольше, чем об этом можно судить по непосредственно израсходо-анным суммам. Однако именно такого рода суммы всегда исполь-уют в процессе оценки экономической эффективности различных абот, а коэффициент χ получается при этом гораздо мень-им. Поэтому на основании приведенного примера можно сде-ать заключение, что экономическая эффективность научных абот, направленных на усовершенствование методов объективного нализа, действительно значительно выше, чем для большинства ругих работ.

Укажем в заключение, что аналогичным способом можно оце-ивать эффективность работ по усовершенствованию методов рогноза погоды, основываясь на том, что точность прогноза за-исит, среди прочих причин, от точности анализа исходного поля, последняя в свою очередь — от количества станций.

СПИСОК ЛИТЕРАТУРЫ

1. Гандин Л. С. Современные проблемы объективного анализа метеорологических полей. — В кн.: Труды V Всесоюзного метеорологического съезда. 1972, т. 2, с. 106—117.
2. Клуге И. Об использовании данных о ветре при объективном анализе высотного барического поля. — «Труды ГГО», 1970, вып. 267, с. 32—51.
3. Машкович С. А. Некоторые вопросы планирования аэрологических станций в свете задач объективного анализа аэрологических наблюдений. — В кн.: Труды Симпозиума по численным методам прогноза погоды. Гидрометеоиздат., Л., 1964, с. 215—223.
4. Экономические проблемы повышения эффективности научных разработок. Л., 1972, 338 с.

СОДЕРЖАНИЕ

Труды ГГО, вып. 348

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В МЕТЕОРОЛОГИИ

Редактор Е. И. Ильиных
Техн. редактор Н. Ф. Грачева
Корректор Г. С. Макарова

Сдано в набор 24/II 1975 г. Подписано к печати 10/VII 1975 г. М-31718. Формат $60\times90^{1}/_{16}$. Бумага тип. № 1. Печ. л. 11. Уч.-изд. л. 11,38. Тираж 650 экз. Индекс МЛ-116. Заказ № 193. Цена 80 коп.
Гидрометеоиздат. 199053, Ленинград, 2-я линия, д 23.

Сортавальская книжная типография Управления по делам издательств, полиграфии и книжной торговли Совета Министров Карельской АССР. Сортавала, Карельская, 42.