ГОСУДАРСТВЕННЫЙ КОМИТЕТ СССР
ПО ГИДРОМЕТЕОРОЛОГИИ И КОНТРОЛЮ
ПРИРОДНОЙ СРЕДЫ

ТРУДЫ
ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ГЛАВНОЙ ГЕОФИЗИЧЕСКОЙ ОБСЕРВАТОРИИ
им. А. И. ВОЕЙКОВА

Выпуск

431

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В МЕТЕОРОЛОГИИ

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

ЛЕНИНГРАД ГИДРОМЕТЕОИЗДАТ 1980

УДК 551.5 : 311.17

Основное содержание сборника составляют вопросы исследования статистической структуры полей метеорологических элементов и применения данных о статистической структуре к различным теоретическим и прикладным задачам. Особое внимание уделено разработкам методов автоматического контроля метеорологической информации. Рассмотрены также вопросы рациональной методики осреднения метеорологических полей, в частности, при наличии периодических изменений.

Сборник рассчитан на научных работников, аспирантов и студентов метеорологических и родственных специальностей.

The investigation of the statistical structure of meteorological fields and the applications of the data on statistical structure to various theoretical and practical problems are the principle topics under consideration. Particular attention is paid to the methods of automatic quality control of meteorological information. Some problems of rational averaging methods of meteorological variables are also considered including the presence of periodical variations.

The publication is intended for graduate and post-graduate students and research workers in meteorology and adjacent sciences.

П $\frac{20807-030}{069(02)-80}$ 34-79(2) 1903040000

Л. С. Гандин, В. П. Тараканова

О ВЕРТИКАЛЬНОЙ СТАТИСТИЧЕСКОЙ СТРУКТУРЕ ОТНОСИТЕЛЬНОГО ГЕОПОТЕНЦИАЛА ИЗОБАРИЧЕСКИХ ПОВЕРХНОСТЕЙ

1. Сведения о статистических связях между значениями относительного геопотенциала различных слоев нужны для решения многих задач. В частности такого рода информация необходима для построения и реализации метода вертикального статистического контроля данных косвенного температурного зондирования атмосферы с метеорологических спутников [4].

Относительный геопотенциал изобарической поверхности $p = p_k$ над поверхностью $p = p_i$ равен, как известно, разности значений абсолютного геопотенциала этих поверхностей:

$$H_{i,\,k} = H_k - H_i. \tag{1}$$

Если имеется n изобарических поверхностей, то можно рассматривать $n(n-1)/2$ величин (1). Для наших целей удобно выделить среди них значения геопотенциала (толщины) слоев между соседними изобарическими поверхностями

$$h_i = H_{i,\,i+1} = H_{i+1} - H_i, \tag{2}$$

где имеется в виду, что поверхности нумеруются снизу вверх.

Легко показать, что, зная ковариации «элементарного» относительного геопотенциала h_i различных слоев, можно найти ковариации и для относительного геопотенциала слоев между несоседними поверхностями. ?

Будем ради простоты полагать все рассматриваемые метеорологические элементы отсчитываемыми от их средних в статистическом смысле значений (норм). Статистическое осреднение будем обозначать чертой сверху. Тогда, поскольку, например

$$H_{i,\,i+2} = h_i + h_{i+1},$$

то дисперсия $\overline{H^2}_{i,\,i+2}$ относительного геопотенциала $H_{i,\,i+2}$ связана ?

с дисперсиями и ковариациями «элементарного» относительного геопотенциала формулой

$$\overline{H^2_{i,\,i+2}} = \overline{h^2_i} + \overline{h^2_{i+1}} + 2\overline{h_i h_{i+1}},$$

ковариация величин $H_{i,\,i+2}$ и h_k — формулой

$$\overline{H_{i,\,i+2} h_k} = \overline{h_i h_k} + \overline{h_{i+1} h_k}$$

и т. п. Имея в виду использование подобного рода соотношений, достаточно обладать информацией лишь о дисперсиях $\overline{h^2_i}$ и ковариациях $\overline{h_i h_k}$ «элементарного» относительного геопотенциала h.

Для получения такой информации можно, разумеется, проводить обработку массового материала аэрологических наблюдений. Значительно экономнее, однако, применить косвенный путь, позволяющий использовать результаты уже выполненных обработок массового материала.

Может показаться, что для этой цели достаточно использовать вытекающие из определения (2) соотношения

$$\overline{h_i h_k} = \overline{H_{i+1} H_{k+1}} + \overline{H_i H_k} - \overline{H_{i+1} H_k} - \overline{H_i H_{k+1}} \qquad (3)$$

и, в частности,

$$\overline{h^2_i} = \overline{H^2_{i+1}} + \overline{H^2_i} - 2\overline{H_{i+1} H_i}\,, \qquad (4)$$

однако это не так: расчет дисперсий и ковариаций h по формулам (4) и (3) связан с настолько большой потерей точности, что его результаты совершенно не заслуживают доверия. Это не должно вызывать удивления, если учесть, что вычисления по формулам (3) и (4) представляют собой по существу определение вторых разностей эмпирически заданной ковариационной функции абсолютного геопотенциала.

Вместо точной формулы (2) к цели приводит использование приближенной, но выполняющейся с большой точностью формулы

$$h_i = B_i\,(T_i + T_{i+1}), \qquad (5)$$

связывающей относительный геопотенциал слоя между соседними главными изобарическими поверхностями с температурой T на этих поверхностях. В отличие от формул (1) и (2), применимых как к фактическим величинам, так и к отклонениям их от норм, формула (5) верна только для отклонений от нормы. Коэффициент

$$B_i = 3{,}37 \lg \frac{p_i}{p_{i+1}}$$

в этой формуле зависит при заданном наборе изобарических поверхностей только от номера i. Формула (5) выполняется точно при линейном изменении температуры с логарифмом давления. Практически ошибки формулы (5) могут быть существенны, если внутри слоя имеет место излом вертикального профиля температуры, например, если между поверхностями p_i и p_{i+1} расположена тропопауза.

Из (5) вытекают формулы

$$\overline{h_i^2} = B_i^2 \left(\overline{T_i^2} + \overline{T_{i+1}^2} + 2\overline{T_i T_{i+1}}\right), \tag{6}$$

$$\overline{h_i h_k} = B_i B_k \left(\overline{T_i T_k} + \overline{T_{i+1} T_{k+1}} + \overline{T_{i+1} T_k} + \overline{T_i T_{k+1}}\right), \tag{7}$$

с помощью которых легко вычислять дисперсии и ковариации h по известным дисперсиям и ковариациям температуры. Эти вычисления требуют, как правило, сложений, а не вычитаний, как в (3) и (4), и потому сопровождаются не потерей, а некоторым повышением точности.

2. В настоящее время накоплен обширный материал по вертикальной статистической структуре температуры. К сожалению, значительная часть этой информации имеется в виде ковариационных (или корреляционных) матриц, построенных для каждой отдельной станции. Такие данные практически невозможно использовать как вследствие их малой точности, что обусловлено значительными выборочными погрешностями, так и просто в силу отсутствия нужной степени универсальности.

Мы использовали данные о вертикальной структуре температуры из работ В. П. Болтенкова [1] (с учетом внесенных Р. Л. Каганом исправлений, см. [2]), О. Б. Мерцаловой [6], а также Л. И. Копровой и В. Г. Болдырева [5]. Кроме того, д-р К. Бергман из Национального метеорологического центра США любезно предоставил в наше распоряжение данные расчета вертикальной структуры температуры, произведенного им по текущей оперативной информации за осень (IX—XI) 1977 г. [8].

В табл. 1 приведена (выше главной диагонали) матрица коэффициентов корреляции температуры главных изобарических поверхностей, рассчитанная В. П. Болтенковым [1] по данным 60 станций в Канаде и США для зимнего сезона и исправленная Р. Л. Каганом [2]. По этой матрице и по значениям дисперсии температуры [2] были вычислены значения коэффициентов корреляции относительного геопотенциала ?

$$\frac{\overline{h_i h_k}}{\sqrt{\overline{h_i^2}\,\overline{h_k^2}}},$$

приведенные в табл. 1 ниже главной диагонали. Эти величины относятся не к уровням, а к слоям между соседними уровнями и потому расположены в табл. 1 посередине между строчками и между столбцами. Матрицы, расположенные таким образом, будем называть вложенными матрицами.

Сопоставление элементов основной и вложенной матриц показывает, что межуровенная корреляция «элементарного» относительного геопотенциала h ведет себя вполне аналогично корреляции температуры. Абсолютные значения коэффициентов корреляции относительного геопотенциала несколько выше, что является естественным следствием осреднения температуры в слое при расчете относительного геопотенциала.

Таблица 1

Матрицы коэффициентов корреляции (%) температуры T [2] и относительного геопотенциала h для зимнего сезона

p_k мбар	p_i мбар							
	1000	850	700	500	400	300	200	100
1000		67	57	47	45	34	—27	—14
850			74	68	57	29	—45	—45
	88							
700				76	67	44	—49	—46
	72	91						
500					94	53	—56	—70
	60	73	90					
400						67	—55	—70
	51	59	76	89				
300							—2	—46
	—16	—23	—21	—16	14			
200								51
	—41	—57	—64	—70	—60	62		
100								

Таблица 2

Коэффициенты корреляции (ниже главной диагонали, %), дисперсии и ковариации (дам^2) относительного геопотенциала для лета и зимы, рассчитанные по данным В. П. Болтенкова [1], [2]

Слой, мбар	Слой, мбар						
	850—1000	700—850	500—700	400—500	300—400	200—300	100—200
				Зима			
850—1000	8,9	8,7	10,5	5,4	4,7	—1,6	—10,7
700—850	88	11,0	14,7	7,2	6,0	—2,7	—16,3
500—700	72	91	23,8	13,1	11,5	—3,6	—27,1
400—500	60	73	90	8,8	8,2	—1,7	—18,0
300—400	51	59	76	89	9,5	1,5	—16,1
200—300	—16	—23	—21	—16	14	12,5	19,0
100—200	—41	—57	—64	—70	—60	62	73,0
				Лето			
850—1000	1,9	2,2	2,5	1,4	1,7	0,6	—2,6
700—850	90	3,1	3,9	2,1	2,4	1,0	—3,7
500—700	71	86	6,5	4,2	4,8	1,8	—6,8
400—500	57	66	91	3,3	3,9	1,5	—4,9
300—400	53	60	82	94	5,3	2,9	—5,8
200—300	15	21	27	32	46	7,2	6,6
100—200	—37	—40	—51	—57	—48	47	27,3

В табл. 2 представлены ковариационные и корреляционные матрицы относительного геопотенциала по тем же данным [1] и [2] для зимнего и летнего сезонов. Сопоставление этих матриц для двух сезонов позволяет сделать следующие выводы. Во-первых, летом в связи с повышением тропопаузы уровень перемены знака корреляции относительного геопотенциала расположен выше, чем зимой. Во-вторых, ковариации в летний период существенно меньше, чем зимой, что связано главным образом с меньшей изменчивостью (т. е. малыми дисперсиями) в летний период. Что же касается значений корреляционной функции, то их различия в разные сезоны гораздо меньше.

Пользуясь данными О. Б. Мерцаловой [6], можно рассмотреть изменения характеристик вертикальной статистической структуры относительного геопотенциала с широтой. В табл. 3 представлены результаты расчета коэффициентов корреляции, дисперсий и ковариации относительного геопотенциала для зимнего сезона по данным работы [6], относящимся к трем широтным зонам северного полушария: 25—40, 40—60 и 60—90°. Согласно данным табл. 3,

Таблица 3

Коэффициенты корреляции (ниже главной диагонали, %), дисперсии и ковариации (дам^2) относительного геопотенциала в трех широтных зонах, рассчитанные по данным О. Б. Мерцаловой [6]

Слой, мбар	Слой, мбар					
	850—1000	700—850	500—700	300—500	200—300	100—200
			Зона 25—40°			
850—1000	3,7	3,8	4,8	4,9	0,1	—5,2
700—850	86	5,3	7,8	7,6	0,2	—8,0
500—700	66	89	14,5	16,0	0,0	—16,8
300—500	51	67	84	24,8	5,0	—20,8
200—300	2	3	0	32	10,2	5,3
100—200	—46	—59	—75	—71	28	34,5
			Зона 40—60°			
850—1000	6,8	7,0	8,9	7,3	—1,4	—7,6
700—850	86	9,7	14,8	12,2	—2,5	—12,7
500—700	69	93	24,9	23,0	—3,9	—22,4
300—500	52	72	85	29,7	3,4	—17,8
200—300	—13	—21	—21	16	14,6	22,7
100—200	—35	—50	—54	—40	71	68,2
			Зона 60—90°			
850—1000	7,5	7,2	9,2	6,6	—1,4	—3,9
700—850	87	9,0	13,2	9,4	—2,1	—5,8
500—700	68	89	24,3	20,4	—4,0	—12,4
300—500	45	58	77	28,5	7,9	3,0
200—300	—10	—14	—17	30	23,3	46,6
100—200	—12	—17	—22	5	85	130,6

ковариации относительного геопотенциала в тропосфере низких широт значительно меньше, чем в умеренных широтах, что связано с меньшей изменчивостью температуры в низких широтах. Коэффициенты корреляции в низких широтах отличаются от их значений в других зонах главным образом за счет смены знака на большей высоте, что в свою очередь связано с большей высотой тропопаузы. Различия же коэффициентов в зонах 40—60 и 60—90° малы и едва ли существенно превосходят выборочную изменчивость рассчитанных характеристик.

Таблица 4

Коэффициенты корреляции (ниже главной диагонали, %), дисперсии и ковариации (дам^2) относительного геопотенциала некоторых слоев, рассчитанные по данным различных источников: [1] для зимы, [6] для зимы (зона 40—60°), [5] для зимы и [8] для осени

Слой, мбар	Источник	Слой, мбар			
		850—700	700—500	500—300	300—200
850—700	[1]	11,0	14,7	13,2	—2,7
	[6]	9,7	14,8	12,2	—2,5
	[5]	6,5	9,5	10,3	1,2
	[8]	6,8	9,6	7,3	0,6
700—500	[1]	91	23,8	24,5	—3,6
	[6]	93	24,9	23,0	—3,9
	[5]	87	18,5	22,1	—1,0
	[8]	90	16,7	14,2	0,8
500—300	[1]	68	85	34,7	—0,2
	[6]	72	85	29,7	3,4
	[5]	70	89	33,1	2,0
	[8]	51	64	29,5	9,9
300—200	[1]	—23	—21	—1	12,5
	[6]	—21	—21	16	14,6
	[5]	6	—3	11	10,2
	[8]	6	5	45	16,2

В табл. 4 приведены результаты расчета дисперсий, ковариаций и коэффициентов корреляции относительного геопотенциала по данным различных авторов. Сопоставление этих результатов между собой позволяет оценить степень надежности, с которой мы можем судить об этих характеристиках статистической структуры. Из табл. 4 видно, что согласование результатов, полученных по разным данным, удовлетворительное, причем в среднем оно несколько лучше для коэффициентов корреляции, чем для размерных величин — дисперсий и ковариаций. Оценивая это согласование, надо иметь в виду, что данные К. Бергмана [8] относятся к осени, и расчеты по ним дают поэтому некоторые закономерные различия по сравнению с другими приведенными в табл. 4 результатами, относящимися к зимнему сезону.

3. Выше упоминалось, что расчет характеристик вертикальной статистической структуры относительного геопотенциала по данным о характеристиках структуры абсолютного геопотенциала сопровождается потерей точности и потому не приводит к заслуживающим доверия результатам. Естественно, что обратная процедура — переход в расчетах вертикальной структуры от относительного геопотенциала к абсолютному — сопровождается возрастанием точности. Это наводит на мысль, что взамен непосредственного расчета ковариационных матриц абсолютного геопотенциала целесообразно определять их по данным о вертикальной структуре относительного геопотенциала. Такой путь аналогичен предложенному М. И. Юдиным [7] методу расчета горизонтальных ковариационных функций геопотенциала (или другого элемента), состоящему в вычислении конечных разностей, определении характеристик статистической структуры этих разностей и переходе от них к характеристикам структуры самого геопотенциала.

Нерегулярность сети станций вынудила М. И. Юдина определить разностные характеристики после интерполяции геопотенциала в точки регулярной сетки. Таким образом, результаты расчетов описывают структуру не истинных полей геопотенциала, а интерполированных. Это обстоятельство должно приводить, по крайней мере в принципе, к систематическим ошибкам в значениях корреляционной функции, а именно к завышению радиуса корреляции. При расчете же вертикальной структуры абсолютного геопотенциала по данным о структуре относительного геопотенциала причины такого рода погрешностей отсутствуют.

Наиболее простая и логичная процедура таких расчетов базируется на задании, наряду с дисперсиями и ковариациями относительного геопотенциала h, также дисперсий абсолютного геопотенциала H. Эта процедура сводится к рекуррентной последовательности явных формул. Именно сначала из формул (4) находим ковариации абсолютного геопотенциала соседних поверхностей:

$$\overline{H_iH_{i+1}}=\frac{1}{2}\left(\overline{H_i^2}+\overline{H_{i+1}^2}-\overline{h_i^2}\right), \tag{8}$$

затем, зная наряду с дисперсиями такие ковариации, из формул (5) при $k=i+1$ определяем ковариации

$$\overline{H_iH_{i+2}}=\overline{H_{i+1}H_{i+2}}+\overline{H_iH_{i+1}}-\overline{H_{i+1}^2}-\overline{h_ih_{i+1}},$$

далее из формул (5) при $k=i+2$ получаем

$$\overline{H_iH_{i+3}}=\overline{H_{i+1}H_{i+3}}+\overline{H_iH_{i+2}}-\overline{H_{i+1}H_{i+2}}-\overline{h_ih_{i+2}}$$

и т. д.

Пример расчетов по приведенным формулам представлен в табл. 5, в которой элементами основной матрицы являются ковариации абсолютного, а элементами вложенной матрицы —

ковариации относительного геопотенциала. Выше главной диагонали последней приведены ковариации $\overline{h_i h_j}$, рассчитанные по данным о структуре температуры (см. табл. 1). Эти ковариации вместе с заимствованными из работы [3] дисперсиями абсолютного геопотенциала $\overline{H_i^2}$ (главная диагональ основной матрицы) использованы для расчета ковариаций $\overline{H_i H_j}$, помещенных выше главной диагонали основной матрицы. Для сравнения ниже главной диагонали помещены результаты выполненного в [3] непосредственного расчета ковариаций $\overline{H_i H_j}$.

Таблица 5

Ковариации и дисперсии абсолютного и относительного геопотенциалов (дам^2)

p_k мбар	p_i мбар								
	850		700		500		300		200
850	99		106		121		142		129
		11,0		14,7		13,2		−2,7	
		19							
700	102		124		154		188		172
				23,8		24,6		−3,6	
		0		39					
500	124		146		207		266		246
						34,7		−0,2	
		16		15		51			
300	144		182		258		360		340
								12,5	
		−7		2		−12		33	
200	131		162		240		330		333

Сравнивая элементы основной матрицы в табл. 5 выше и ниже главной диагонали, легко убедиться, что различия между ними находятся в пределах выборочных погрешностей; наибольшее расхождение составляет около 5 %. Однако рассчитанные нами значения $\overline{H_i H_j}$ имеют то существенное преимущество, что они согласованы со статистической структурой относительного геопотенциала или, лучше сказать, со структурой температуры.

Элементами вложенной матрицы в табл. 5, помещенными ниже главной диагонали, являются значения дисперсий и ковариаций относительного геопотенциала, рассчитанные с помощью формул вида (3) и (4) по данным работы [3] о структуре абсолютного геопотенциала (т. е. по данным, приведенным в табл. 5 в основной матрице на главной диагонали и ниже ее). Можно убедиться, что они действительно имеют мало общего с результатами корректных расчетов, приведенными выше главной диагонали: различия имеют порядок 100 %.

Таким образом, подтверждается целесообразность расчета характеристик вертикальной статистической структуры абсолютного геопотенциала по данным о структуре относительного геопотенциала.

СПИСОК ЛИТЕРАТУРЫ

1. Болтенков В. П. Некоторые характеристики трехмерной макроструктуры температуры воздуха.— Труды ГГО, 1966, вып. 191, с. 47—57.
2. Гандин Л. С., Каган Р. Л. Статистические методы интерпретации метеорологических данных.— Л.: Гидрометеоиздат, 1976. 357 с.
3. ✓ Гандин Л. С., Кузнецова Т. И. О пространственной статистической структуре поля геопотенциала.— Труды ГГО, 1965, вып. 168, с. 84—93.
4. ✓ Гандин Л. С., Тараканова В. П. О вертикальном статистическом контроле данных косвенного зондирования с метеорологических спутников.— См. наст. сб.
5. Копрова Л. И., Болдырев В. Г. О статистических характеристиках вертикальной структуры полей температуры и влажности до больших высот.— Изв. АН СССР. Физика атмосферы и океана, 1970, т. 6, № 2, с. 154—167.
6. Мерцалова О. Б. Стандартизация вертикальных корреляционных связей температуры над северным полушарием.— Труды НИИАК, 1967, вып. 46, с. 129—148.
7. ✓ Юдин М. И. Некоторые закономерности структуры поля геопотенциала.— Труды ГГО, 1961, вып. 121, с. 3—18.
8. Bergman K. Personal communication, 1978.

Л. С. Гандин, В. П. Тараканова

О ВЕРТИКАЛЬНОМ СТАТИСТИЧЕСКОМ КОНТРОЛЕ ДАННЫХ КОСВЕННОГО ТЕМПЕРАТУРНОГО ЗОНДИРОВАНИЯ С МЕТЕОРОЛОГИЧЕСКИХ СПУТНИКОВ

1. Хорошо известно, что сеть аэрологических станций, производящих радиозондирование, распределена по земному шару крайне неравномерно. В хорошо обжитых районах суши эта сеть достаточно густая, а в пустынных областях и, особенно, на акваториях океанов и морей она очень редкая. Количество аэрологических станций в южном полушарии приблизительно в 10 раз меньше, чем в северном.

Эта неравномерность сети является серьезным препятствием к расширению областей, охватываемых моделями численного прогноза, и особенно к внедрению глобальных прогностических моделей, что, в свою очередь, ограничивает возможности повышения точности и увеличения заблаговременности прогнозов погоды.

Вместе с тем уже на протяжении около десяти лет наряду с данными радиозондирования распространяются в оперативном порядке результаты косвенного зондирования атмосферы с метеорологических спутников. Эта информация относится именно к акваториям океанов, т. е. к тем областям, где хуже всего обстоит дело с данными радиозондирования. Объем данных косвенного зондирования достаточно велик: количество пунктов, данные косвенного зондирования над которыми поступают в течение суток, имеет порядок тысячи над каждым полушарием.

Эта обширная информация остается, однако, до сих пор не включенной в общую систему автоматической обработки оперативной метеорологической информации, и при построении начальных полей для целей численного прогноза данные косвенного зондирования не используются. Так получилось потому, что, несмотря на весьма обнадеживающие результаты ряда модельных численных экспериментов и на положительные результаты некоторых экспериментов с фактическими данными косвенного зондирования (например, [8]), пока не удалось добиться устойчивого повышения качества объективного анализа и численного прогноза благодаря включению данных косвенного зондирования атмосферы.

Можно указать на две причины этого парадоксального положения. Одна из них состоит в том, что в процедурах усвоения спутниковой информации не учитывалось специфическое свойство этой информации, состоящее в коррелированности ошибок косвенного зондирования над различными пунктами. Наличие такой корреляции ошибок было предсказано сравнительно давно [2], а затем и обнаружено на фактическом материале [5]. Тогда же было показано, что неучет этого фактора может обесценить привлечение данных косвенного зондирования при объективном анализе [3]. Недавно Н. Филлипс [9] продемонстрировал, что включение данных косвенного зондирования без должного учета свойств ошибок этих данных может приводить не к улучшению, а к ухудшению результатов численного прогноза.

Вторая причина создавшегося положения состоит в следующем. Система автоматической оперативной обработки аэрологической информации включает в качестве важного этапа автоматический контроль. Процедуры контроля служат для выявления грубых ошибок, вкравшихся по тем или иным причинам в полученную прогностическим центром информацию, и последующего их исправления или, если не удается исправить, исключения ошибочных данных. Естественно, что данные косвенного зондирования тоже должны контролироваться, иначе отдельные грубые ошибки могут существенно исказить анализируемые поля. Однако в течение ряда лет этому вопросу не уделялось никакого внимания.

Между тем, как показывает исследование [6], не менее 8 % всех телеграмм с данными косвенного зондирования содержат грубые ошибки. Об этом свидетельствует опыт применения к данным о температуре и относительном геопотенциале, содержащимся, согласно прежнему коду, в телеграммах со спутниковой информацией, известной методики статического контроля.

К сожалению, в новом коде, действующем с июля 1977 г., предусмотрена передача данных только об относительном геопотенциале, а температура не передается. Это исключает возможность использования такого действенного средства, как статической контроль, и заставляет искать иных путей выявления и устранения грубых ошибок. Одним из таких путей может быть вертикальный статистический контроль.

Идея вертикального статистического контроля была предложена М. И. Юдиным и состоит в сравнении наблюденного значения контролируемого метеорологического элемента на каждом уровне с результатом его оптимальной интерполяции по данным других уровней. Применительно к обычной аэрологической информации об абсолютном геопотенциале эта идея была реализована Б. М. Ильиным и Л. В. Руховцом [7] (см. также [1]).

2. Данные косвенного зондирования передаются в виде относительного геопотенциала H_k различных изобарических поверхностей p_k над поверхностью $p_1 = 1000$ мбар. Можно было подвергать

вертикальному контролю непосредственно эти значения, однако для повышения чувствительности контроля естественно перейти к значениям относительного геопотенциала h_k слоев между соседними главными изобарическими поверхностями:

$$h_k = H_{k+1} - H_k, \tag{1}$$

и привлекать для контроля h_k данные об относительном геопотенциале соседних слоев. При контроле для каждого контролируемого слоя, кроме крайних в имеющемся наборе, будем привлекать данные двух примыкающих сверху и снизу слоев. С этой целью будем сравнивать наблюденное значение отклонения геопотенциала слоя от среднего значения (нормы) с результатом интерполяции данных двух соседних слоев по формуле

$$\hat{h}_k = a_k \tilde{h}_{k-1} + b_k \tilde{h}_{k+1}. \tag{2}$$

Здесь a_k и b_k — весовые множители,

$$\tilde{h}_i = h_i + \delta_i \tag{3}$$

представляет собой наблюденное значение отклонения относительного геопотенциала от нормы, h_i — истинное значение этого отклонения, δ_i — ошибка наблюдений. Заметим, что здесь и в дальнейшем вместо фактических значений рассматриваются отклонения от норм.

Веса a_k и b_k определим из требования минимума среднего квадрата разности между интерполированным значением $\hat{h}_k$ и наблюденным значением $\tilde{h}_k$:

$$\tilde{E}_k^2 = \overline{\left(a_k \tilde{h}_{k-1} + b_k \tilde{h}_{k+1} - \tilde{h}_k\right)^2}. \tag{4}$$

Величина (4) отличается от среднего квадрата ошибки интерполяции по формуле (2)

$$E_k^2 = \overline{\left(a_k \tilde{h}_{k-1} + b_k \tilde{h}_{k+1} - h_k\right)^2} \tag{5}$$

тем, что вместо истинного значения h_k в формуле (4) фигурирует наблюденное значение $\tilde{h}_k$, связанное с истинным формулой (3). Минимизация величины (4), соответствующая задаче контроля, не эквивалентна, вообще говоря, минимизации величины (5), если только не отсутствует корреляция между ошибками наблюдения δ.

Вводя ковариации измеренных величин

$$\tilde{m}_{ij} = \overline{\tilde{h}_i \tilde{h}_j}, \tag{6}$$

перепишем формулу (4) в виде

$$\tilde{E}_k^2 = a_k^2 \tilde{m}_{k-1,\,k-1} + 2a_k b_k \tilde{m}_{k-1,\,k+1} + b_k^2 \tilde{m}_{k+1,\,k+1} - 2a_k \tilde{m}_{k-1,\,k} - 2b_k \tilde{m}_{k+1,\,k} + \tilde{m}_{k,\,k}. \tag{7}$$

Минимизация (7) приводит к уравнениям для определения весов a_k и b_k:

$$\left.\begin{aligned} a_k\widetilde{m}_{k-1,\ k-1}+b_k\widetilde{m}_{k-1,\ k+1}=\widetilde{m}_{k-1,\ k}, \\ a_k\widetilde{m}_{k-1,\ k+1}+b_k\widetilde{m}_{k+1,\ k+1}=\widetilde{m}_{k+1,\ k}, \end{aligned}\right\} \tag{8}$$

пользуясь которыми можно формулу (7) записать в виде

$$\widetilde{E}_k^2=\widetilde{m}_{kk}-a_k\widetilde{m}_{k-1,\ k}-b_k\widetilde{m}_{k+1,\ k}. \tag{9}$$

Вычислив по формуле (9) величину $\widetilde{E}_k$, можно оценить допустимую невязку Δ_k, такую, что если

$$\left|\widetilde{h}_k-\hat{h}_k\right|<\Delta_k, \tag{10}$$

то вероятно наличие грубой ошибки. Именно, естественно принять, что

$$\Delta_k=A\widetilde{E}_k, \tag{11}$$

где константа A составляет 2,5—3,5.

Для самого верхнего и самого нижнего слоев интерполяция по формуле (2) невозможна и приходится пользоваться оптимальной экстраполяцией по данным одного слоя. Например, контроль относительного геопотенциала h_1 выполняется путем сравнения $\widetilde{h}_1$ с

$$\hat{h}_1=b_1\widetilde{h}_2, \tag{12}$$

где b_1 определяется из уравнения

$$b_1\widetilde{m}_{22}=\widetilde{m}_{12}, \tag{13}$$

а средний квадрат разности между $\widetilde{h}_1$ и $\hat{h}_1$ вычисляется по формуле

$$\widetilde{E}_1^2=\widetilde{m}_{11}-b_1\widetilde{m}_{12}. \tag{14}$$

Аналогично контролируется геопотенциал самого верхнего слоя.

3. Если ошибки измерений δ_i не коррелируют с истинными значениями измеряемых величин, то ковариация измеренных величин $\widetilde{m}_{ij}$ связана с ковариацией истинных значений

$$m_{ij}=\overline{h_ih_j} \tag{15}$$

и ковариацией ошибок измерений

$$n_{ij}=\overline{\delta_i\delta_j} \tag{16}$$

очевидным соотношением

$$\widetilde{m}_{ij}=m_{ij}+n_{ij}. \qquad (17)$$

Из приближенной формулы

$$h_i=B_i\,(T_i+T_{i+1}), \qquad (18)$$

связывающей относительный геопотенциал слоя между поверхностями p_i и p_{i+1} с температурой T на этих поверхностях (см. например, [1]), можно получить связь между дисперсиями ошибок геопотенциала h и температуры. В случае если ошибки температуры не коррелируют между собой, эта связь дается формулой

$$n_{ii}=B_i^2\,(d_i^2+d_{i+1}^2), \qquad (19)$$

где d_i^2 представляет собой дисперсию ошибок измерения температуры на поверхности p_i.

Таблица 1

Веса a_k, b_k и допустимые невязки Δ_k вертикального контроля относительного геопотенциала для зимнего сезона в умеренных широтах

Параметр	Вариант	Слой, мбар						
		1000—850	850—700	700—500	500—400	400—300	300—200	200—100
$a_k\cdot 10^2$	1	—	51	79	32	91	80	141
	2	—	49	79	32	83	57	113
	3	—	52	81	33	94	81	140
$b_k\cdot 10^2$	1	78	38	81	44	22	41	—
	2	73	36	66	38	17	60	—
	3	78	37	82	44	26	40	—
Δ_i гп. дам	1	4,0	2,8	3,9	2,8	4,0	5,3	17,9
	2	4,9	4,0	6,5	4,2	5,4	5,4	20,4
	3	4,2	2,8	4,0	2,9	3,7	5,4	18,6

В табл. 1 приведены значения весов a_k, b_k и допустимых невязок Δ_k вертикального контроля относительного геопотенциала h некоторых слоев, рассчитанные для трех вариантов. Вариант 1 моделирует случай аэрологических наблюдений. В этом случае для поверхности 1000 мбар было принято $d_1^2=5\,(°\mathrm{C})^2$, а для остальных поверхностей $d_2^2=d_3^2\ldots=1{,}5\,(°\mathrm{C})^2$. Ошибки геопотенциала различных слоев считались некоррелированными, т. е. $n_{ij}=0$ при $j\neq i$. Вариант 2 относится к случаю предположения, что ошибки косвенного зондирования не коррелировали между собой. Для них было принято $d_1^2=16\,(°\mathrm{C})^2$, $d_2^2=d_3^2=\ldots=6{,}25\,(°\mathrm{C})^2$. Наконец, вариант 3 отличается от варианта 2 предположением о коррелированности ошибок δ_i. Именно для коэффициентов корреляции между ошибками в соседних слоях и в слоях через один были приняты значения

$$\frac{n_{k-1,\,k}}{\sqrt{n_{k-1,\,k-1}n_{k,\,k}}}=\frac{n_{k+1,\,k}}{\sqrt{n_{k+1,\,k+1}n_{k,\,k}}}=0{,}6;$$

$$\frac{n_{k-1,\,k+1}}{\sqrt{n_{k-1,\,k-1}n_{k+1,\,k+1}}}=0{,}2.$$

Для всех вариантов были использованы ковариации истинных значений m_{ij}, рассчитанные в работе [4] по данным В. П. Болтенкова для зимы. Константа A в формуле (11) была принята равной 2,5.

Интерполяционные веса в табл. 1 заметно отличаются от тривиальных значений, равных 0,5 для промежуточных слоев и 1,0 для крайних. Эти отличия обусловлены особенностями вертикальной статистической структуры относительного геопотенциала. Вместе с тем изменения весов от одного варианта к другому сравнительно невелики. Что же касается допустимых невязок Δ, то они существенно различны для разных вариантов. Если бы ошибки косвенного зондирования не были коррелированы между собой (вариант 2), то возможности вертикального контроля таких данных были бы гораздо ниже возможностей аналогичного контроля данных радиозондирования (вариант 1). Но благодаря корреляции случайных ошибок косвенного зондирования возможности обнаружения грубых ошибок на их фоне весьма сильно возрастают. Они оказываются примерно такими же, как возможности контроля данных радиозондирования, несмотря на то, что случайные ошибки последнего гораздо меньше.

В связи с этим результатом интересно оценить степень надежности выполненных оценок, а также выяснить, как они зависят от широты места и времени года. Результаты соответствующих расчетов представлены в табл. 2. Все эти расчеты соответствуют варианту 3, моделируя тем самым контроль спутниковой информации с учетом как большой величины случайных ошибок такой информации, так и коррелированности этих ошибок. При расчетах использованы сведения о ковариациях m_{ij}, полученные в работе [4] по различным данным о ковариациях температуры изобарических поверхностей, а именно по данным О. Б. Мерцаловой (А), В. П. Болтенкова (Б), Л. И. Копровой и В. Г. Болдырева (В) и К. Бергмана (Г).

Из табл. 2 видно, что допустимые невязки Δ вертикального контроля существенно уменьшаются от зимы к лету и от высоких широт к низким. Соответственно возрастает чувствительность этого метода контроля, если судить о ней непосредственно по значениям Δ. С этой целью логичнее, однако, рассматривать отношения

Таблица 2

Веса, абсолютные и относительные допустимые невязки вертикального контроля относительного геопотенциала по данным косвенного зондирования

Слой, мбар	Параметр	А (зима)			Б		В (зима)	Г (лето)
		90—60°	60—40°	40—25°	зима	лето		
850—700	$b_k \cdot 10^2$	52	55	50	58	51	49	54
	Δ гп. дам	4,1	3,7	3,4	4,2	3,2	3,8	3,6
	$\chi \cdot 10^2$	130	113	135	124	158	139	129
700—500	$a_k \cdot 10^2$	106	103	95	90	77	85	115
	$b_k \cdot 10^2$	36	35	35	36	34	39	56
	Δ гп. дам	5,0	4,4	4,1	4,3	3,4	4,2	3,0
	$\chi \cdot 10^2$	96	83	98	83	110	80	68
500—300	$a_k \cdot 10^2$	91	98	105	105	112	116	82
	$b_k \cdot 10^2$	51	52	52	36	42	38	58
	Δ гп. дам	7,4	7,0	6,6	8,6	6,6	7,6	9,6
	$\chi \cdot 10^2$	125	115	118	133	136	120	159
300—200	$a_k \cdot 10^2$	27	36	51	39	47	28	53
	$b_k \cdot 10^2$	35	40	41	41	38	38	37
	Δ гп. дам	6,6	5,9	6,2	6,1	4,8	6,5	7,7
	$\chi \cdot 10^2$	125	134	163	197	141	170	170
200—100	$a_k \cdot 10^2$	185	143	68	140	96	114	53
	Δ гп. дам	17,5	16,2	16,0	18,6	13,6	14,8	17,5
	$\chi \cdot 10^2$	146	180	232	198	214	200	237

допустимой невязки к среднему квадратическому отклонению контролируемой величины от нормы

$$\chi_k = \frac{\Delta_k}{\sqrt{m_{kk}}}. \qquad (20)$$

Такие отношения несколько возрастают в среднем от зимы к лету и с уменьшением широты места.

Из табл. 2 видно также, что по данным различных авторов, относящимся к зиме для умеренных широт, получаются близкие значения допустимых невязок вертикального контроля относительного геопотенциала на основании данных косвенного зондирования. Что же касается интерполяционных весов a_k и b_k, то они мало меняются также при переходе от сезона к сезону и от одной широты к другой.

4. Рассмотрим реакцию описанного метода контроля на основные типы грубых ошибок. Наиболее часто грубые ошибки происходят вследствие искажения какого-либо одного значения H на канале связи. Такое искажение приведет к грубым ошибкам в значениях h для двух слоев, примыкающих к тому уровню, относительный геопотенциал которого над поверхностью 1000 мбар оказался искаженным. Поскольку значения h в каждом слое привлекаются

для контроля соседних слоев, появляются невязки контроля не в двух, а в четырех слоях. Именно пользуясь формулами (1)—(3), нетрудно показать, что искажение значения H_k на величину ζ_k

$$\widetilde{H}_k = H_k + \zeta_k \tag{21}$$

приведет к появлению следующих невязок

$$\widehat{\delta}_i = \widehat{h}_i - \widetilde{h}_i \tag{22}$$

вертикального контроля:

$$\left.\begin{aligned} \widehat{\delta}_{k-2} &= b_{k-2}\zeta_k, \\ \widehat{\delta}_{k-1} &= -(1+b_{k-1})\zeta_k, \\ \widehat{\delta}_k &= (1+a_k)\zeta_k, \\ \widehat{\delta}_{k+1} &= -a_{k+1}\zeta_k. \end{aligned}\right\} \tag{23}$$

Из формул (23) видно, что при переходе от каждого из этих слоев к следующему невязка меняет знак, причем невязки в слоях, примыкающих к уровню p_k, больше по величине, чем в соседних, и даже больше, чем искажение ζ_k.

В случаях когда уровень p_k, относительный геопотенциал H_k которого над поверхностью $p_1 = 1000$ мбар искажен на каналах связи, близок к самому верхнему или самому нижнему уровню, реакция вертикального контроля на это искажение соответственно изменяется. Если искажено значение относительного геопотенциала H_m самого последнего (верхнего) уровня p_m, то это сказывается лишь на двух невязках:

$$\left.\begin{aligned} \widehat{\delta}_{m-2} &= b_{m-2}\zeta_m, \\ \widehat{\delta}_{m-1} &= -\zeta_m. \end{aligned}\right\} \tag{24}$$

Искажение ζ_{m-1} на предпоследнем уровне приводит к трем невязкам:

$$\left.\begin{aligned} \widehat{\delta}_{m-3} &= b_{m-3}\zeta_{m-1}; \\ \widehat{\delta}_{m-2} &= -(1+b_{m-2})\zeta_{m-1}, \\ \widehat{\delta}_{m-1} &= (1+a_{m-1})\zeta_{m-1}, \end{aligned}\right\} \tag{25}$$

причем в последней формуле (25) вес a_{m-1} является экстраполяционным, а не интерполяционным. Аналогичный характер носит реакция контроля на искажение ζ_2 относительного геопотенциала $H_2 = h_1$:

$$\left.\begin{aligned} \widehat{\delta}_1 &= -(1+b_1)\zeta_2, \\ \widehat{\delta}_2 &= (1+a_2)\zeta_2, \\ \widehat{\delta}_3 &= -a_3\zeta_2, \end{aligned}\right\} \tag{26}$$

где b_1 представляет собой экстраполяционный вес при контроле h_1.

Вторым источником грубых ошибок является искажение одного из значений h в процессе расчета величин H_k при обработке данных косвенного зондирования. Такое искажение h может быть вызвано либо неправильно восстановленным по спутниковым данным значением температуры на одном из уровней, либо погрешностью, вкравшейся при расчете h по данным о температуре.

Обозначая грубую ошибку в относительном геопотенциале h_k слоя между соседними главными изобарическими поверхностями p_k и p_{k+1} через ω, так что

$$\tilde{h}_k = h_k + \omega_k, \tag{27}$$

легко получить формулы для обусловленных этой ошибкой невязок вертикального контроля:

$$\left.\begin{aligned} \delta_{k-1} &= b_{k-1}\omega_k, \\ \delta_k &= -\omega_k, \\ \delta_{k+1} &= a_{k+1}\omega_k. \end{aligned}\right\} \tag{28}$$

Из (28) видно, что в данном случае тоже получаются невязки разных знаков, но не в четырех последовательных слоях, а в трех. Основное отличие рассматриваемого случая от предыдущего состоит, однако, не в этом факте, а в характере необходимого исправления. Если там нужно было исключить грубую ошибку ζ_k из только одного переданного значения относительного геопотенциала H_k, то здесь нужно исключить грубую ошибку ω_k из всех значений H_{k+1}, H_{k+2} и т. д.

Понятно, что если ошибка ω_k возникла в геопотенциале самого нижнего ($k=1$) или самого верхнего ($k=m-1$) слоя, то из трех равенств (28) остаются два: при $k=1$ пропадает первое из этих равенств, а при $k=m-1$ — последнее.

Можно упомянуть еще об искажениях, вызванных неисправностью измерительной, передающей или принимающей аппаратуры. Если такое искажение касается данных лишь одного спектрального канала, то естественно ожидать, что оно будет восприниматься как искажение одного из значений h. Если же будет искажена вся информация, то это приведет к большим невязкам для всех слоев.

В качестве примера использования рассмотренного метода контроля данных косвенного зондирования в табл. 3 приведены результаты его применения к некоторым из данных о температуре и геопотенциале, ранее подвергавшихся статическому контролю [6]. Во всех приведенных случаях с помощью статического контроля обнаружена и оценена ошибка (ζ_k) в одном из значений H. Это дало возможность рассчитать по формулам вида (23) ожидаемые значения ($\hat{\delta}(\zeta)$) невязок вертикального контроля. Как видно из табл. 3, фактические невязки $\hat{\delta}$ имеют значения, близкие к ожидаемым.

Таблица 3

Примеры вертикального контроля данных косвенного зондирования с одним ошибочным значением

Параметр	p мбар															p_k мбар	ζ
	1000		850		700		500		400		300		200		100		
H	0		131,4		283,0		532		671		877		1130		1572	400	−16,1
h		−4,6		−7,0		−14		−26		3		−15		−6			
$\hat{\delta}$						−12		23		−32		15					
$\hat{\delta}(\zeta)$						−13		23		−32		14					
H	0		142,0		306,1		587		750		959		1233		1655	500	8,6
h		6,0		5,5		17,9		−2		6		6		−16			
$\hat{\delta}$				4		−15		10		−6							
$\hat{\delta}(\zeta)$				3		−16		11		−8							
H	0		129,7		288,2		536		694		887		1146		1595	700	5,8
h		−6,3		−0,1		−15,2		−7		−10		−9		1			
$\hat{\delta}$		6		−9		10		−2									
$\hat{\delta}(\zeta)$		4		−8		10		−2									
H	0		130,3		282,5		534		699		879		1132		1564	400	9,4
h		−5,7		−6,4		−11,5		0		−23		−15		−16			
$\hat{\delta}$						7		−14		19		−11					
$\hat{\delta}(\zeta)$						8		−13		18		−8					
H	0		141,7		305,2		576		737		957		1230		1650	400	−10,1
h		5,7		4,9		7,8		−4		17		5		−28			
$\hat{\delta}$						−7		14		−20		−1					
$\hat{\delta}(\zeta)$						−9		15		−21		9					

СПИСОК ЛИТЕРАТУРЫ

1. Гандин Л. С., Каган Р. Л. Статистические методы интерпретации метеорологических данных.— Л.: Гидрометеоиздат, 1976.— 357 с.
2. Гандин Л. С., Каган Р. Л., Полищук А. И. Об оценке информационной значимости систем метеорологических наблюдений.— Труды ГГО, 1972, вып. 286, с. 120—140.
3. Гандин Л. С., Лугина К. М. Об учете данных косвенного зондирования атмосферы при четырехмерном анализе метеорологических полей.— Труды ГГО, 1975, вып. 348, с. 3—15.

4. Гандин Л. С., Тараканова В. П. О вертикальной статистической структуре относительного геопотенциала изобарических поверхностей.— См. наст. сб.

5. Тараканова В. П. К вопросу о точности температурного зондирования атмосферы с искусственных спутников Земли.— Метеорология и гидрология, 1974, № 4, с. 76—78.

6. Тараканова В. П. О повышении информативности сети станций южного полушария при использовании данных косвенного зондирования и прогноза. — Труды ГГО, 1977, вып. 397, с. 116—123.

7. Юдин М. И., Ильин Б. М., Руховец Л. В. Об одном способе контроля и исправления аэрологических телеграмм.— Метеорология и гидрология, 1964, № 5, с. 35—39.

8. Halem M. a. e. Preliminary report on DST-5 impact tests of temperature sounding data on Short-range forecasting.— GARP, Programme on Numerical Experimentation, 1976, N 11, p. 410—422.

9. Phillips N. A. The impact of synoptic observing and analysis systems on flow pattern forecasts.— Bull. Amer. Met. Soc., 1976, vol. 57, N 10, p. 1225—1240.

Ю. М. Либерман

АЛГОРИТМ КОМПЛЕКСНОГО СТАТИКО-ВРЕМЕННОГО КОНТРОЛЯ АЭРОЛОГИЧЕСКОЙ ИНФОРМАЦИИ

В настоящей работе описывается алгоритм контроля информации о геопотенциале H и температуре t, составными частями которого являются статический контроль (СК) и временно́й контроль (ВК). Алгоритм комплексного статико-временно́го контроля (СВК) предназначен для режимной обработки информации о значениях H и t на стандартных изобарических поверхностях, полученных по каналам связи в форме данных аэрологических телеграмм.

В основу алгоритма СВК положены следующие принципы.

1. Контроль телеграмм выполняется снизу вверх по слоям, границами которых являются стандартные изобарические уровни.

2. Обнаружению и по возможности исправлению подлежат ошибки, допущенные в процессе регистрации и обработки наблюдений на аэрологических станциях, а также возникшие при передаче данных по каналам связи и их обработке в региональном центре. Систематические ошибки, обусловленные неисправностью радиозонда или радиолокатора (так называемые ошибки радиозондирования) в данной схеме СВК не анализируются. Вообще говоря, комплекс СК и ВК позволяет обнаруживать ошибки радиозондирования. Соответствующий блок алгоритма логически весьма сложен; его реализация представляет собой ближайшую задачу совершенствования СВК.

3. Если диагноз ошибки, сформулированный блоком СК, однозначен, то принятое на основании СК решение окончательно. В такой ситуации нет необходимости в привлечении ВК. Исключением являются случаи восстановления значений обоих элементов — и H и t — на одном уровне и исправления относительного геопотенциала: в этих случаях сделанное исправление проверяется последующим ВК.

4. Если СК предлагает два или более вариантов исправления обнаруженной ошибки, то выбор одного или нескольких из этих вариантов производится на основании ВК.

5. При определенных условиях превышение по абсолютной величине критического значения Δ_i^{i+1} невязки СК ее фактическим значением δ_i^{i+1} является допустимым и необязательно служит признаком ошибки на уровнях i, $i+1$. Для таких случаев предусмотрена процедура реабилитации данных, признанных на предварительном этапе подозрительными.

6. В случаях когда однозначное решение не может быть принято из-за сложных комбинированных ошибок или недостатка данных, информация выдается из памяти ЭВМ для анализа человеком.

Блок СК включен в алгоритм СВК без существенных изменений по сравнению с ранее описанным вариантом [3]. В него лишь введена дополнительная проба на знак температуры, поскольку опыт показал, что весьма часто ошибка в значении температуры представляет собой искажение ее знака.

Временно́й контроль

Контроль по времени заключается в сравнении содержащегося в телеграмме значения метеорологического элемента f на проверяемом уровне с результатом интерполяции по данным двух ближайших сроков: предшествующего и последующего. Для интерполяции используется простейшая двухточечная линейная схема

$$f_n = a f_{n-j} + (1-a) f_{n+k} \quad (j=1,\ 2;\quad k=1,\ 2). \qquad (1)$$

Здесь индекс n означает момент времени, которому соответствует проверяемая телеграмма. Иначе говоря, n есть порядковый номер срока наблюдений в файле, содержащем N последовательных сроков, разделенных стандартными 12-часовыми интервалами. Для выполнения ВК срока n обязательно наличие хотя бы одного предшествующего и одного последующего сроков, отстоящих от проверяемого не более чем на ± 24 ч. Если данные за срок, ближайший к n, отсутствуют (например, телеграмма не поступила), то они заменяются данными более отдаленного предшествующего или последующего срока. Экстраполяция по времени не применяется, поэтому ВК крайних сроков файла $n=1$ и $n=N$ невозможен. Поскольку предельный сдвиг по времени не превосходит 24 ч, коэффициенты корреляции значений H или t в проверяемый срок со значениями в предшествующий и последующий сроки оказываются не ниже 0,8. При столь высокой корреляции применение для ВК какой-либо более сложной интерполяционной схемы, по-видимому, нецелесообразно.

Значения весового коэффициента a для каждого из четырех возможных вариантов интерполяции приведены в табл. 1. Там же указаны рассчитанные значения меры ошибки сопоставления результата интерполяции с контролируемым наблюдением $\widetilde{\varepsilon^2} = \widetilde{E^2}/\sigma^2$. Здесь $\widetilde{E^2}$ — средний квадрат ошибки сопоставления, σ^2 — дисперсия элемента (величину σ будем называть изменчивостью).

Таблица 1

Параметры временно́го контроля срока n по двум соседним срокам

Соседние сроки	a	$\widetilde{\varepsilon^2}$	$\widetilde{\varepsilon'}$
$n-1$ и $n+1$	1/2	0,038	0,2
$n-2$ и $n+1$	1/3	0,071	0,3
$n-1$ и $n+2$	2/3	0,071	0,3
$n-2$ и $n+2$	1/2	0,156	0,4

Величина $\widetilde{\varepsilon^2}$ найдена как сумма мер ошибок интерполяции и наблюдения:

$$\widetilde{\varepsilon^2}=\varepsilon^2+\eta^2. \tag{2}$$

Для расчета ε^2 использована формула

$$\varepsilon^2=1-2\,[\mu_{n,\,n-j}a+\mu_{n,\,n+k}(1-a)]+2\mu_{n-j,\,n+k}a\,(1-a)+$$
$$+(1+\eta^2)\,[a^2+(1-a)^2]\quad (j=1,\ 2;\quad k=1,\ 2), \tag{3}$$

где μ — значения автокорреляционной функции элемента для соответствующих интервалов времени [1]. Значение η^2 принято одинаковым для геопотенциала и температуры и равным 0,02. Временна́я автокорреляционная функция обоих элементов аппроксимирована зависимостью

$$\mu(\tau)=\left(1+\frac{\tau}{30}\right)e^{-\frac{\tau}{30}}, \tag{4}$$

где τ — время в часах.

Зависимость (4) представляет собой частный случай пространственно-временно́й корреляции геопотенциала в тропосфере умеренных широт [2].

На основе значений $\widetilde{\varepsilon^2}$, рассчитанных по формулам (2) и (3), мы приняли значения $\widetilde{\varepsilon'}=\sqrt{\widetilde{\varepsilon^2}}$ (табл. 1). Величина $\widetilde{\varepsilon'}$ позволяет определить допустимые невязки сопоставления

$$\Delta_i^\tau=K\widetilde{\varepsilon'}\sigma_i, \tag{5}$$

превышение которых по абсолютной величине сигнализирует о вероятной ошибке. Здесь индекс τ введен для отличия допустимой невязки ВК на уровне i от допустимой невязки СК в слое между уровнями i и $i+1$, для которой используется обозначение Δ_i^{i+1}. Эмпирический коэффициент K равен 3.

Допустимая невязка Δ_i^τ рассчитывается в процессе ВК для каждой контролируемой станции как функция широты φ. Из табл. 2,

Таблица 2

Зональное распределение изменчивости геопотенциала σ_H и температуры σ_t

p мбар	$\varphi°$								
	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90
	Изменчивость σ_H гп. дам								
100	6,7	6,4	10,0	15,8	20,3	22,6	24,2	25,2	25,2
150	5,7	6,0	9,9	16,7	21,2	23,2	23,6	23,3	23,3
200	4,5	5,5	9,8	17,6	22,1	23,8	22,9	21,2	21,2
250	4,0	4,9	8,8	16,0	21,1	23,5	22,6	20,4	20,4
300	3,3	4,2	7,8	14,3	20,0	23,2	22,3	20,2	20,2
400	2,4	3,5	6,2	10,8	16,3	19,4	18,9	17,9	17,9
500	2,2	2,6	4,5	8,4	13,2	15,8	15,5	15,4	15,4
700	1,7	2,2	3,2	5,6	8,5	10,7	10,7	10,6	10,6
850	1,4	1,7	2,8	4,5	6,9	8,8	8,7	8,4	8,4
1000	1,0	1,7	3,5	4,9	7,4	9,1	8,5	8,1	8,1
	Изменчивость σ_t °C								
100	3,4	3,2	3,4	3,9	4,3	4,5	4,8	5,4	5,4
150	2,8	2,9	3,3	4,0	4,7	5,0	5,1	5,4	5,4
200	1,9	2,5	3,1	4,1	5,1	5,6	5,3	5,4	5,4
250	1,8	2,4	3,3	4,5	5,2	5,4	5,1	5,0	5,0
300	1,8	2,4	3,4	5,0	5,3	5,2	4,8	4,6	4,6
400	1,9	2,2	3,4	5,0	6,1	6,4	6,0	5,5	5,5
500	1,6	1,8	2,9	4,6	6,0	6,8	6,5	6,1	6,1
700	1,3	2,0	2,9	4,7	6,4	6,9	6,8	6,5	6,5
850	1,1	1,8	3,3	5,2	6,7	7,1	7,4	6,9	6,9
1000	1,0	1,3	2,6	4,3	5,8	6,8	9,0	9,3	9,3

составленной по данным монографии [4] для осеннего сезона, видно, что изменчивости геопотенциала и температуры существенно зависят от широты. Пренебрежение этой зависимостью привело бы к недостоверности Δ_i^τ.

Алгоритм принятия решения

1. Контроль начинается с нижнего слоя телеграммы, ограниченного первым и вторым изобарическими уровнями.

1.1. При обнаружении ошибки в нижнем крайнем слое СК предлагает три варианта исправления, каждый из которых анализируется блоком ВК. Если ВК отверг два варианта из трех, то принимается единственный допущенный вариант. Если допущены два или три варианта или не допущен ни один вариант, а также если ВК невозможен, то решение должно приниматься человеком на основании информации, выданной на печать.

1.2. Если первый уровень неполон, т. е. содержит значение лишь одного элемента (либо H, либо t), то отсутствующее значение восстанавливается [3]. Для этого помимо полноты второго уровня требуется выполнение условия

$$|\delta_2^3| \leqslant \Delta_2^3, \tag{6}$$

которое предотвращает восстановление по данным соседнего уровня в случае их ошибочности.

1.3. До восстановления неполного крайнего уровня по полному соседнему уровню следовало бы проверить значение единственного элемента, имеющееся на неполном уровне. В частности, при восстановлении t_{1000} желательно быть уверенным в правильности H_{1000} и наоборот. Контроль H_{1000} возможен по приземным данным о давлении и температуре, содержащимся в телеграмме. Однако эти данные в свою очередь могут содержать ошибки и потому также нуждаются в проверке. Что касается t_{1000}, то контроль этой величины возможен лишь на базе ВК. В описываемом варианте СВК контроль H_{1000} и t_{1000} не предусмотрен.

2. Превышение допустимой невязки СК в двух смежных слоях говорит о сомнительности значений H_i или t_i на уровне i, разделяющем слои.

2.1. В большинстве случаев СК принимает однозначное решение об исправлении H_i или t_i по соотношению абсолютных величин невязок и их знаков.

2.2. Ситуация, в которой такое исправление не удается, свидетельствует об ошибочности значений обоих элементов. В таком случае ошибочные значения заменяются новыми, а именно рассчитанными по формулам

$$H_i = H_{i+1} - A_i^{i+1} - B_i^{i+1} t_{i+1} - B_i^{i+1} (H_{i+1} - H_{i-1} - A_i^{i+1} -$$

$$- A_{i-1}^i - B_i^{i+1} t_{i+1} - B_{i-1}^i t_{i-1}) / (B_i^{i+1} + B_{i-1}^i),$$

$$t_i = (H_{i+1} - H_{i-1} - A_i^{i+1} - A_{i-1}^i -$$

$$- B_i^{i+1} t_{i+1} - B_{i-1}^i t_{i-1}) / (B_i^{i+1} + B_{i-1}^i). \tag{7}$$

Значения параметров A и B приведены в работе [3].

Формулы (7) представляют собой решение системы второго порядка, которое весьма сильно меняется при малых изменениях коэффициентов системы. Эта неустойчивость требует обязательной проверки полученных исправленных значений с помощью ВК. Если ВК отверг хотя бы одно из значений H_i, t_i или если ВК невозможен, то фиксируется отказ от исправления данного уровня и от дальнейшего контроля вышележащих уровней.

2.3. Отсутствие данных H_i и t_i на промежуточном уровне телеграммы воспринимается по признаку 999 как ошибочность обоих этих значений. Обработка такого «пустого» уровня выполняется аналогичным образом по формулам (7) с последующим ВК.

3. Превышение допустимой невязки СК только в одном промежуточном слое указывает на возможную ошибку в вычислении относительного геопотенциала слоя, исправление которой заключается в изменении значений геопотенциала всех уровней, начиная с верхнего уровня данного слоя, на одну и ту же величину.

3.1. Исправление относительного геопотенциала производится только в тех случаях, когда не удается уменьшить невязку менее радикальным путем, а именно изменением H или t на одной из границ слоя. Для этого выполняются предварительные пробы, как описано в работе [3]. Если СК допускает изменение H или t на обеих границах слоя, то для уточнения диагноза привлекается ВК.

3.2. Если по условиям СК изменения H или t на границах слоя невозможны или ВК не допускает таких изменений, то исправляется относительный геопотенциал вычитанием фактической невязки из значений H верхней границы слоя и всех вышележащих уровней. Новые значения H подвергаются временно́му контролю. Если ВК дает отрицательный ответ на одном или нескольких уровнях, то контроль телеграммы прерывается и информация выдается на печать для субъективного анализа. В противном случае подтверждается исправление относительного геопотенциала и контроль продолжается.

4. Контроль заканчивается анализом верхнего крайнего слоя телеграммы.

4.1. При обнаружении ошибки в верхнем слое СК предлагает два варианта исправления. Оба они анализируются блоком ВК по аналогии с контролем нижнего слоя.

4.2. В случае неполноты верхнего крайнего уровня отсутствующее значение H или t восстанавливается по данным предпоследнего уровня. Эти данные должны быть признаны правильными на предшествующей стадии, а именно при контроле предпоследнего слоя.

5. Если невязки, превышающие допустимые значения, обнаружены в трех или более смежных слоях, то контроль телеграммы прекращается с выдачей информации на печать. В этих сложных случаях можно было бы продолжить анализ и попытаться поставить диагноз нескольких ошибок, содержащихся в телеграмме. Но такой дополнительный анализ затруднителен, а число подобных случаев сравнительно невелико. Поэтому целесообразно в таких случаях предоставить решение человеку.

6. Алгоритм принятия решения предусматривает послойную реабилитацию подозрительных данных при определенных условиях. Первое условие требует, чтобы в рассматриваемом слое превышение допустимой невязки СК составляло не более 20 % ее величины. Второе условие предполагает, что в соседних слоях фактическая невязка СК достаточно мала, а именно не превосходит 50 % допустимого значения. Неоднозначность результатов ВК рассматривается как признак, эквивалентный выполнению второго условия. Возможна реабилитация как для промежуточных, так и для крайних слоев.

Программная реализация СВК

Алгоритм СВК реализован на языке ФОРТРАН применительно к ЭВМ БЭСМ-6. Программа, содержащая 1300 фортранных операторов, состоит из трех подпрограмм.

Головная подпрограмма формирует рабочие массивы, заполняя их исходной информацией очередных телеграмм. Эта информация читается с магнитной ленты, где данные хранятся в виде макетов. Каждый макет соответствует одной станции. Три его первых слова содержат синоптический индекс станции, ее широту и долготу. Следующие 20 слов содержат значения H и t иа 10 уровнях от 1000 до 100 мбар. Таких групп по 20 слов в макете N, т. е. столько, сколько последовательных сроков наблюдений подлежат контролю. Размещение каждого срока в макете фиксировано; частичное или полное отсутствие данных обозначается признаком 999. Головная подпрограмма управляет сменой сроков данной станции и обеспечивает переход к следующей станции или окончание счета.

Подпрограмма временного контроля выполняет ВК одной станции на одном или нескольких смежных уровнях. Она имеет следующие шесть формальных параметров:

— порядковый номер проверяемого срока в файле;

— широта станции;

— порядковые номера нижнего и верхнего уровней, между которыми заключены проверяемые уровни (совпадение этих номеров означает, что ВК выполняется на единственном уровне);

— признак контроля H или t;

— значение коэффициента K, входящего в формулу (5).

Подпрограмма ВК обеспечивает выборку значений изменчивости по широте станции и номеру уровня из таблиц, хранящихся в памяти. Результатом ее работы являются признаки 0, 1 или 2, которые соответственно означают, что ВК дал положительный или отрицательный ответ или не дал никакого ответа. Последнее имеет место в случае невозможности ВК из-за недостатка данных.

Третья подпрограмма выполняет СК, устанавливает диагноз ошибки и вырабатывает решение о ее исправлении, вызывая при необходимости подпрограмму ВК. Эта третья подпрограмма наибольшая по объему: она содержит около 1000 операторов. Здесь же формируется подробная информация о результатах СВК, выдаваемая на печать. Такая информация содержит числовую и словесную части. В числовой форме выдаются значения H и t, содержащиеся в телеграмме, исправленные значения этих элементов и невязки СК и ВК до исправления и после него. В словесной форме печатаются диагнозы ошибок, сообщения о недостатке данных, их реабилитации, прекращении или продолжении контроля очередной телеграммы. Кроме того, печатаются таблицы, в которых содержатся сведения о восстановленных значениях H и t на крайних уровнях, а также о телеграммах с числом уровней менее 10. Несмотря на значительный объем программы, потребное машинное время невелико. Оно зависит от качества проверяемых телеграмм: чем меньше обнаруживается ошибок, тем быстрее заканчивается счет. В среднем контроль 100 телеграмм продолжается около 90 с включая трансляцию программы.

СПИСОК ЛИТЕРАТУРЫ

1. Гандин Л. С., Каган Р. Л. Статистические методы интерпретации метеорологических данных.— Л., Гидрометеоиздат, 1976.— 359 с.
2. Каган Р. Л., Лугина К. М. К вопросу о пространственно-временном анализе барического поля.— Труды ГГО, 1974, вып. 336, с. 75—94.
3. Либерман Ю. М. Алгоритм неоперативного статического контроля аэрологической информации.— Труды ГГО, 1978, вып. 412.
4. Oort A. H., Rasmusson E. M. Atmospheric circulation statistics.— Prof. Paper, 1971, N 5, NOAA, p. 95—97.

Ю. М. Либерман, В. П. Тараканова

ОПЫТ СТАТИКО-ВРЕМЕННОГО КОНТРОЛЯ АЭРОЛОГИЧЕСКОЙ ИНФОРМАЦИИ

В работе [1] описан алгоритм комплексного статико-временного контроля аэрологической информации в процессе ее режимной обработки. Отладка и испытания этого алгоритма позволили получить достаточно полную картину качества аэрологических телеграмм, передаваемых по линиям связи. Эту картину мы предполагаем охарактеризовать в отдельной работе. Одновременно удалось приобрести определенный опыт, полезный с точки зрения дальнейшего развития систем контроля данных. Его изложению посвящена настоящая работа.

Материалом для испытаний являются 3294 телеграммы со станций северного полушария, поступившие в Гидрометцентр СССР и записанные там на магнитную ленту. Эти телеграммы относятся к восьми последовательным срокам наблюдения начиная с 00 ч 12 сентября 1976 г.; интервал между сроками 12 ч. Телеграммы каждого срока записывались в порядке возрастания синоптических индексов в форме макетов, содержащих данные о геопотенциале H и температуре t, а также о дефиците точки росы и ветре на десяти стандартных изобарических поверхностях от 1000 до 100 мбар. Поэтому до выполнения контроля потребовалась предварительная перестройка файла исходных данных на ленте с помощью специальной программы. В результате мы получили постанционные макеты, содержащие значения H и t в порядке возрастания сроков (в дальнейшем сроки обозначаются номерами от 1 до 8).

Контроль выполнялся на ЭВМ БЭСМ-6 сеансами по 60—80 макетов. Счет в одном сеансе продолжался не более 2 мин, но увеличению числа макетов в сеансе препятствовал большой расход бумаги, обусловленный характером и объемом выводимой текстовой и цифровой информации. Затем эта информация многократно подвергалась тщательному субъективному анализу, на основании которого в алгоритм вносились исправления и усовершенствования. Характерно, что суммарный расход машинного времени составил около 120 мин, а всесторонний анализ результатов счета потребовал приблизительно 6 человеко-месяцев. Для такого анализа весьма полезным оказалось систематическое сопоставление результатов двух методов контроля: статического (СК) и статико-временного (СВК).

Перейдем к характеристике эффективности СВК. При этом выделим уровень 1000 мбар, поскольку из-за отсутствия в макете данных о давлении и температуре на уровне станции и высоте станции над уровнем моря мы часто не может однозначно судить о верности или ошибочности значений H_{1000} и t_{1000}. Сказанное относится, в частности, к случаям отсутствия в телеграмме значений H_{1000} и t_{1000}: такое отсутствие может быть оправданным (если приземное давление менее 1000 мбар) или ошибочным, но установить это не удается. Можно, однако, констатировать, что программа СВК восстановила отсутствующее значение t_{1000} в 1199 случаях, а восстановление H_{1000} произошло в 7 случаях. В дальнейшем в статистике ошибок уровень 1000 мбар не учитывается, за исключением особо оговоренных случаев.

Дефектных телеграмм, т. е. телеграмм с искаженными или отсутствующими значениями H и t оказалось 417, что составляет 13 % общего числа проконтролированных телеграмм. Искажение значения H обнаружено в 264 случаях, а значения t — в 201 случае; отсутствие значения H — в 52 случаях, а значения t — в 33 случаях. В 48 случаях искажены значения H и t на одном и том же уровне. В 17 случаях отсутствуют значения обоих элементов на промежуточном уровне.

В результате контроля исправлено и восстановлено 190 значений H и 108 значений t, в 6 случаях исправлены значения относительного геопотенциала.

Число неоднозначных исправлений, сделанных блоком статического контроля в рамках СВК, равно 59. Из них:

три варианта исправления в нижнем крайнем слое — 26;
два варианта исправления в верхнем крайнем слое — 23;
два варианта исправления на промежуточном уровне — 10.

В 21 случае, благодаря блоку временного контроля (ВК), удалось выбрать единственный приемлемый вариант из трех или двух, в 1 случае ВК позволил уменьшить число допустимых вариантов с 3 до 2.

Примеры однозначного выбора исправления приведены в табл. 1. Здесь и далее символом δ обозначены послойные невязки СК, комбинация 999 является признаком отсутствия данных, а ΔH означает изменение относительного геопотенциала нижнего слоя. Из примеров видно, что ВК не только отсеивает заведомо абсурдные варианты, но и делает выбор в случаях, когда внешне правдоподобными выглядят все предлагаемые варианты. Субъективный анализ показал, что в каждом из рассматриваемых случаев сделанный с помощью ВК выбор правилен.

Дефектная информация о значениях обоих элементов, и H и t, на одном и том же промежуточном уровне исправлена в 26 случаях. Об эффективности этих исправлений можно судить на основании табл. 2. В ее последней графе символы «+» и «—» соответствуют положительной и отрицательной реакции ВК при проверке исправления, предложенного блоком СК. Если ВК невозможен из-за отсутствия данных в соседние сроки или из-за того, что проверяемый срок является крайним, то в последней графе табл. 2 стоит символ 0.

Важным следствием обязательной проверки с помощью ВК является фильтрация ложных исправлений. Поясним это на примерах из табл. 2.

1. На станции 24 688 абсурдное значение $t_{700} = -414{,}9\,^\circ C$ обусловлено тем, что при его расчете по формулам (7) из работы [2] использовано искаженное значение геопотенциала вышележащего уровня $H_{500} = 854$ гп. дам.

2. На станции 44 231 исправленные значения $H_{300} = 913$ гп. дам и $t_{300} = -109{,}0\,^\circ C$ оказались ложными по аналогичной причине: в телеграмме ошибочны данные $H_{250} = 1026$ гп. дам и $t_{250} = -15{,}3\,^\circ C$.

3. На станции 99 906 значение $H_{200} = 1128$ гп. дам занижено на 100 гп. дам, а значение $t_{200} = -48{,}1\,^\circ C$ завышено на $10{,}0\,^\circ C$, поэтому неверен результат исправления $H_{150} = 1318$ гп. дам.

Предотвратить указанные ложные исправления в рамках СК трудно, поэтому дополнительное привлечение ВК представляется необходимым. Заметим, что восстановленные значения H и t на одном уровне иногда оказываются ложными не потому, что ошибочны используемые при расчете значения на соседних уровнях, а вследствие неустойчивости соответствующей системы уравнений [1]. И в этих случаях ВК своевременно сигнализирует о неприемлемости результатов восстановления (станции 33 008 и 42 809).

С другой стороны, ВК может привести к неверному суждению, если восстановление проверяемого срока выполнено правильно, а значение элемента в соседний срок ошибочно. Так произошло на станции 44 288, где ошибочно значение $t_{200} = -64{,}1\,^\circ C$ последующего срока 7. Подобные случаи легко обнаруживаются при просмотре выданных на печать результатов.

Ложные исправления возникают не только при обработке дефектных значений двух элементов на одном уровне. По-видимому, определенная малая доля ошибочных решений неизбежна даже при очень разветвленном и сложном алгоритме их принятия. В таких случаях ВК играет роль фильтра, разумеется, если ВК возможен. Для иллюстрации приведем примеры ложных исправлений в случаях комбинированных ошибок. Эти неверные исправления не были своевременно обнаружены только потому, что неполнота данных не позволила провести ВК (табл. 3).

1. На станции 57 036 верными были бы величины $t_{300} = -26{,}9\,^\circ C$ и $H_{250} = 1104$ дам. Вместо таких изменений было принято ошибочное решение о дефектности значений обоих элементов на уровне 250 мбар.

Таблица 1

Примеры однозначного выбора исправления крайнего слоя с помощью ВК

Станция	Исправленный уровень, мбар	Срок	H гп. дам	t °C	δ гп. дам	Исправление	Отклоненные варианты
4 340	500	1	541	—28,3	$\delta_{700}^{500} = -6,9$	$t_{500} = -24,6$	$H_{500} = 546$
		2	539	—10,5			
		3	522	—28,1			
8 509	150	4	1397	—51,5	$\delta_{200}^{150} = 510,6$	$H_{150} = 1395$	$t_{150} = 1158,4$
		5	1906	—54,9			
		7	1403	—62,7			
10 868	100	3	1643	—60,5	$\delta_{150}^{100} = 12,2$	$H_{100} = 1639$	$t_{100} = 34,8$
		4	1651	—55,3			
		5	1642	—58,5			
12 425	100	3	1643	—52,1	$\delta_{150}^{100} = -575,1$	$H_{100} = 1638$	$t_{100} = 1018,7$
		4	1063	—50,5			
		5	1643	—50,7			
15 420	1000	4	17	24,2	$\delta_{1000}^{850} = 58,7$	$H_{1000} = 20$	$t_{1000} = 259,1$ или $\Delta H = -59$
		5	—39	12,4			
		7	22	10,0			

Станция	Исправленный урсвень, мбар	Срок	H гп. дам	t °C	δ гп. дам	Исправление	Отклоненные варианты
25 428	500	1	540	—20,7	$\delta_{700}^{500} = -22,2$	$t_{500} = -20,1$	$H_{500} = 558$
		2	536	25,0			
		3	530	—22,1			
34 858	1000	4	19	15,4	$\delta_{1000}^{850} = 11,3$	$H_{1000} = 20$	$t_{1000} = 59,3$ или $\Delta H = -11$
		5	9	12,0			
		6	16	19,2			
35 394	100	3	1650	—53,3	$\delta_{150}^{100} = -52,0$	$t_{100} = -52,0$	$H_{100} = 1710$
		4	1658	35,6			
		5	1654	—55,5			
38 341	200	5	1213	—52,1	$\delta_{250}^{200} = -9,7$	$H_{200} = 1222$	$t_{200} = -77,4$
		6	1212	—47,9			
		7	1216	—50,5			
45 004	1000	1	5	999	$\delta_{1000}^{850} = 7,4$	$H_{1000} = 7$	$t_{1000} = 56,6$ или $\Delta H = -7$
		3	0	25,4			
		4	7	27,4			

Таблица 2

Восстановление значений H и/или t на одном уровне

Станция	Уровень, мбар	Проверяемый срок					Предшествующий срок			Последующий срок			ВК
			Данные телеграммы		Восстановленные значения								
		k	H	t	H	t	k	H	t	k	H	t	
1 001	250	1	958	26,0	1001	—64,7							0
20 891	400	2	158	32,0	695	—35,9	1	692	—39,7	3	699	—33,7	+
24 688	700	6	999	999	387	—414,9	5	2 2	—10,7	7	284	—11,1	—
28 698	500	8	999	999	565	—15,5	6	576	—15,5				0
29 698	700	8	317	81,0	307	1,6	7	309	1,2				0
33 008	700	7	999	999	305	31,2	6	305	—7,9	8	291	—8,5	—
38 606	300	1	999	999	946	—38,2				2	946	—38,9	0
40 650	250	7	1048	25,0	1099	—38,9	5	1098	—41,1	8	999	999	0
42 182	200	2	1015	20,2	1240	—53,4	1	1248	—52,5	3	1249	—50,1	—
42 809	500	7	568	999	574	—19,7	5	586	—2,1	8	583	—5,1	—
42 809	400	8	999	—40,7	755	—13,1	7	739	—19,1				0
42 867	300	1	999	999	968	—26,1							0
43 128	500	1	999	—50,5	577	—3,9							0
44 212	700	3	999	999	313	—0,1	2	311	—2,5	4	315	0,4	+
44 231	300	2	952	—42,1	913	—109,0	1	937	—43,1	3	942	—46,1	—
44 288	200	6	1113	—26,5	1213	—49,3	5	1218	—49,7	7	1206	—64,1	—
44 292	250	4	999	999	1199	—46,3	3	1208	—36,1	5	1208	—51,1	+
44 354	400	5	562	2,0	744	—24,7	3	743	—19,5	6	744	—25,9	0
47 401	150	4	999	30,0	1386	—54,4	3	1382	—51,7	6	1376	—52,7	+
47 412	500	4	339	—30,5	574	—13,5	3	568	—13,1	6	558	—15,7	+
47 909	500	4	129	12,0	590	—1,2	3	999	999	6	587	—4,7	0
47 945	250	3	999	999	1102	—37,6	1	1098	—39,9	4	1104	—37,1	+
54 374	400	4	999	999	722	—14,2	3	999	999	6	736	—22,5	0
54 662	500	4	535	—26,5	588	—3,3	3	999	999	6	583	—7,5	0
54 178	150	2	999	999	1436	—61,8	1	1441	—59,1	3	999	999	0
57 957	200	1	1356	6,0	1254	—46,7				2	999	999	0
58 847	150	2	999	999	1430	—63,2	1	1431	—60,1	3	1429	—64,1	+
59 287	200	2	999	999	1243	—48,1	1	1242	—49,3	4	1245	—50,7	+
71 909	400	1	999	999	690	—39,3				2	697	—37,1	0
72 572	300	4	999	999	950	—39,4	3	949	—41,7	6	948	—41,5	+
99 906	150	3	999	999	1318	—46,2	1	1393	—62,5	4	1361	—52,3	—

Таблица 3

Примеры ложных исправлений

Станция	Срок	Параметр	p мбар										Ложные исправления
			1000	850	700	500	400	300	250	200	150	100	
57 036	1	H	999	152	314	587	761	975	1004	1296	1480	1686	$H_{250} = 1123$ и
		t	999	13,4	—9,1	—2,3	—12,7	—16,9	—36,1	—47,3	—60,3	—71,1	$t_{250} = 28,6$
		δ		5,3	100,0	0,9	—3,5	—102,8	141,3	—0,7	—39,9		
58 424	3	H	12	151	315	558	999	999	999	999	999	999	$H_{500} = 598$ или
		t	20,6	16,0	20,2	8,0	999	999	999	999	999	999	$t_{500} = - 73,3$
		δ	1,3	1,7	—40,1								
78 367	1	H	13	156	320	592	764	974	1100	1247	1426	1247	$H_{150} = 1247$ и
		t	28,0	18,8	10,0	—6,3	—15,9	—31,7	—42,9	—54,3	—42,9	—54,3	$t_{150} = —491,5$
		δ	1,5	0,9	1,4	1,3	0,0	—0,1	0,8	—10,1	—445,3		
78 988	8	H	11	152	316	584	758	967	1093	1238	1415	1655	
		t	24,8	18,8	7,2	—6,1	—17,5	—32,5	—43,9	—56,3	—68,5	—71,1	$t_{400} = - 11,8$
		δ	1,3	0,9	—1,4	3,7	0,0	0,4	—0,2	—0,5	—1,1		

2. На станции 58 424 в действительности $t_{700} = 10{,}2\,°C$, $H_{500} = 588$ гп. дам, $t_{500} \approx -4{,}0\,°C$. Однако слой 700—500 мбар оказался в телеграмме самым верхним, и для него рекомендованы два неправильных варианта исправления.

3. Телеграмма со станции 78 367 сильно искажена на каналах связи: значение t_{250} повторено на месте t_{100}, а значения H_{200} и t_{200} повторены на месте H_{100} и t_{100}. В результате произошло ложное восстановление значений обоих элементов на уровне 150 мбар.

4. На станции 78 988 не следовало бы вносить никаких изменений. Однако, алгоритм предусматривает реабилитацию слоя 500—400 мбар лишь при превышении допустимой невязки, равной 3 гп. дам, не более чем на 20 %, тогда как фактически превышение равно 23 %.

Из всех видов ошибок, исправление которых возможно на основе рассматриваемого алгоритма СВК, наиболее трудны для диагноза ошибки, допущенные при вычислении относительного геопотенциала при обработке наблюдений на станциях. Два примера подобных ошибок приведены в табл. 4. На станции 11 520 исправление подтверждено последующим ВК. На станции 31 168 такая проверка невозможна из-за отсутствия данных в предшествующий и последующий сроки. В этом втором примере телеграмма содержит две изолированные ошибки: помимо занижения относительного геопотенциала слоя 500—400 мбар на 9 гп. дам искажено также значение H_{1000}. Выданные на печать три варианта исправления нижнего слоя 1000—850 мбар таковы: принять $H_{1000} = 14$ гп. дам, или $t_{1000} = 46{,}8\,°C$, или уменьшить относительный геопотенциал этого слоя на 9 гп. дам. На основании субъективного анализа наиболее вероятен первый вариант.

Из шести исправлений относительного геопотенциала, зафиксированных при испытаниях, четыре случая в результате субъективного анализа признаны верными, а 2 случая — вероятно верными. Тем не менее соответствующий блок алгоритма, по-видимому, может быть заметно усовершенствован по мере накопления опыта.

В испытанном алгоритме СВК предусмотрен отказ от исправления телеграммы при обнаружении недопустимых невязок в трех или более смежных слоях. Такая ситуация, свидетельствующая о значительных искажениях телеграммы вследствие комбинированных ошибок, встретилась в 45 случаях. В примерах, приведенных в табл. 5, ошибочные значения H и t даны курсивом. Анализ этих примеров допускает развитие алгоритма в двух направлениях. Во-первых, недопустимо большие невязки в смежных слоях часто образуют две пары. Их попарное исследование позволяет обнаружить и исправить две ошибки на разных уровнях. Так, легко видеть, что на станции 32 477 значения H_{400} и H_{250} занижены соответственно на 60 и 10 гп. дам. Во-вторых, одна из нескольких недопустимых невязок иногда лишь ненамного превосходит допуск. Такую невязку можно при определенных ограничениях считать приемлемой, тогда число недопустимых невязок уменьшится и диагноз упростится.

Таблица 4

Примеры исправления относительного геопотенциала

Срок	Параметр	Уровень, мбар									
		1000	850	700	500	400	300	250	200	150	100
		Станция 11520, уровни 7—10, $\Delta H=10$									
2	H	15	154	313	576	742	942	1067	1204	1384	1639
3	H	18	153	311	571	734	932	1041	1182	1365	1623
	t	999	8,2	—1,1	—18,1	—29,1	—47,1	—55,1	—57,1	—55,3	—52,9
	δ		0,7	0,2	0,4	0,1	—9,7	—0,3	0,3	—1,7	
—	$H_{\text{испр}}$	18	153	311	571	734	932	1051	1192	1375	1633
4	H	23	156	313	570	731	928	1046	1190	1376	1639
		Станция 31168, уровни 4—10, $\Delta H=9$									
7	H	5	146	298	540	699	893	1010	1155	1343	1608
	t	8,2	—3,5	—10,7	—25,9	—35,7	—50,9	—54,3	—50,3	—50,3	—49,5
	δ	9,2	1,1	—8,8	1,1	0,5	—0,9	1,2	0,4	0,3	
—	$H_{\text{испр}}$	14	146	298	549	708	902	1019	1164	1352	1617

Таблица 5

Примеры значительно искаженных телеграмм

Станция	Срок	Параметр	Уровень, мбар									
			1000	850	700	500	400	300	250	200	150	100
21 946	1	*H*	6	134	284	533	690	884	*504*	1144	*1345*	1603
		t	—0,5	—8,5	—13,3	—27,7	—36,3	—51,3	*23,0*	—48,1	—46,7	—47,7
		δ	0,2	0,4	0,7	—0,1	0,9	—518,4	470,2	11,9	—9,9	
32 477	1	*H*	—11	124	279	534	*634*	899	*1038*	1183	1378	1642
		t	999	4,2	—6,3	*—24,1*	—23,3	—30,9	—32,3	—38,7	—42,9	—52,1
		δ		0,6	0,8	—62,5	57,8	9,9	—9,8	—0,6	—3,6	
44 212	1	*H*	14	151	311	571	733	943	1011	*1127*	1384	999
		t	999	12,0	1,2	—18,3	—29,5	*—34,1*	*25,0*	—51,1	—51,9	—53,5
		δ		0,7	—0,2	—0,4	6,8	—75,6	—53,5	70,4		
44 212	4	*H*	18	157	315	577	741	*543*	1060	999	999	999
		t	999	11,0	0,4	*16,2*	—29,1	*24,0*	—51,5	999	999	999
		δ		0,3	—15,3	—9,8	—425,9	378,3				
44 288	2	*H*	2	999	310	573	740	*549*	1069	1217	1404	999
		t	999	999	—2,5	*11,0*	—23,9	—24,5	*—24,5*	—49,7	*—90,1*	999
		δ			—10,6	—6,8	—400,6	387,1	—5,7	15,9		
51 076	2	*H*	999	151	311	573	999	999	1016	999	999	999
		t	999	16,0	1,2	—17,1	*10,6*	999	*88,2*	999	999	999
		δ		0,6	0,4	250,1	—655,0	—419,3				
76 225	6	*H*	8	151	316	585	*75*	*96*	*508*	1230	1409	1653
		t	999	18,0	9,4	—8,1	—21,3	—37,1	—45,3	—54,5	—67,1	—71,1
		δ		1,9	—0,9	—678,4	—184,4	288,0	576,6	0,2	2,1	
99 906	7	*H*	24	161	321	588	*854*	957	*1089*	1223	1408	1662
		t	14,4	13,8	3,0	—12,3	—24,9	—39,7	—49,7	—53,7	—57,5	—59,1
		δ	0,1	0,5	2,2	100,2	—99,8	9,9	—10,2	1,8	—0,7	

Сформулируем еще несколько рекомендаций по совершенствованию алгоритма СВК.

1. Крайние уровни телеграммы, на которых одно из значений H или t отсутствует, названы в работе [1] неполными. Их восстановление по данным соседнего полного уровня иногда оказывается ложным из-за ошибочности единственного значения, сохранившегося на неполном крайнем уровне. Особенно часто неверно восстанавливается значение t_{1000} из-за ошибки в H_{1000}. В свою очередь ошибка в значении H_{1000} может возникнуть либо при передаче по линиям связи, либо при расчете на станции. В последнем случае ее можно было бы обнаружить и исправить, если предусмотреть в алгоритме повторение расчета H_{1000}, выполненного на станции. Для этого необходимо включить в макет приземные значения давления и температуры и высоту станции над уровнем моря. Следует, однако, иметь в виду, что эта информация также нуждается в контроле.

2. Алгоритм СВК допускает реабилитацию слоя, невязка которого несколько превышает допустимую. По-видимому, в определенных ситуациях целесообразна и обратная процедура, а именно признание невязки недопустимой, несмотря на то, что принятый предел не превышен. Например, на станции 99 903 в срок 1 встретилась такая ситуация:

Уровень, мбар	850		700		500		400
H дам	153		316		585		755
t °С	16,0		—0,5		—7,1		—18,5
δ дам		2,9		4,1		0,4	
Δ дам		3		4		3	

Хотя допустимая невязка Δ_{850}^{700} не превышена, но очевидна необходимость совместного анализа фактических невязок δ_{850}^{700} и δ_{700}^{500}: он показывает, что значение t_{700} занижено на 9 °С.

3. Весьма полезно проверять знак температуры. В ходе испытаний СВК искажение знака обнаружено в 68 случаях, причем знак минус отсутствовал в 53 случаях и был ошибочно присвоен в 15 случаях. По-видимому, пробу знака целесообразно выполнять при любой попытке исправления температуры.

4. Велики возможности уточнения временно́го контроля в отношении оптимизации его параметров и критериев на разных широтах и высотах. В принятом варианте временной контроль зачастую оказывается излишне жестким или, наоборот, чрезмерно мягким.

5. Желательно исключить применение цифровой комбинации 999 как признака отсутствия данных, чтобы этот признак, воспринятый как значение геопотенциала или температуры, не вызывал бы появление еще одной недопустимой невязки СК, что иногда ведет к напрасному отказу от контроля телеграммы.

В заключение приведем пример ошибки, обусловленной неисправностью радиозонда или радиолокатора. Телеграммы со станции 1001, приведенные в табл. 6, позволяют уверенно заключить,

Таблица 6

Пример ошибки радиозондирования на станции 1001 в срок 5

Срок	Параметр	Уровень, мбар									
		1000	850	700	500	400	300	250	200	150	100
4	H	—1	122	275	528	687	880	998	1145	1334	1602
	t	999	—0,7	—9,3	—24,1	—36,3	—50,5	—50,5	—48,3	—48,1	—49,1
5	H	—20	110	261	511	669	868	995	1150	1348	1623
	t	999	—3,9	—11,5	—27,3	—36,5	—35,5	—34,9	—36,7	—39,5	—43,9
6	H	—7	123	276	529	686	999	999	999	999	999
	t	999	—1,1	—8,5	—26,7	—39,1	999	999	999	999	999
7	H	7	137	290	544	703	898	1015	1159	1348	1614
	t	0,8	—3,1	—8,1	—24,1	—35,1	—49,1	—57,1	—48,7	—49,1	—49,5

Таблица 7

Данные ВК геопотенциала (дам) станции 1001 в срок 5

Уровень, мбар	Срок			Невязка ВК	
	4	5	6	фактическая	допустимая
1000	—1	—20	—7	—11,8	6,4
850	122	110	123	—12,5	6,7
700	275	261	276	—13,9	8,5
500	528	511	529	—17,5	12,3
400	687	669	686	—17,5	14,3

Таблица 8

Данные ВК температуры (°С) станции 1001 в срок 5

Уровень, мбар	Срок			Невязка ВК	
	4	5	7	фактическая	допустимая
300	−50,5	−35,5	−49,1	14,5	5,5
250	−50,5	−34,9	−57,1	17,8	6,0
200	−48,3	−36,7	−48,7	11,7	6,5
150	−48,1	−39,5	−49,1	8,9	6,5

что в срок 5 температура завышена на всех уровнях, начиная с 300 мбар, а геопотенциал занижен на всех уровнях до 300 мбар. Тем не менее требования СК в этот срок не нарушены, а потому ошибку можно обнаружить лишь на основе ВК. Данные ВК приведены в табл. 7 и 8. Ошибки подобного рода вызываются не только неисправностью измерительной аппаратуры, но также погрешностями, допущенными при выдержке исправного радиозонда перед запуском. Разработка блока СВК, выявляющего и исправляющего такие «ошибки радиозондирования», представляет собой первоочередную задачу.

Авторы признательны сотруднику Гидрометцентра СССР Е. А. Локтионовой, подготовившей файл аэрологических телеграмм для испытания алгоритма СВК.

СПИСОК ЛИТЕРАТУРЫ

1. Либерман Ю. М. Алгоритм комплексного статико-временно́го контроля аэрологической информации.— См. наст. сб.

Л. П. Клягина

ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ КОМПЛЕКСНОГО КОНТРОЛЯ ПОЛЯ ПРИЗЕМНОГО ДАВЛЕНИЯ

Во ВНИИГМИ—МЦД с 1974 г. начаты работы по созданию централизованного архива гидрометеорологических данных на машинных носителях. Вполне естественно, что данные могут иметь ошибки, которые необходимо не только выявлять, но и исправлять, чтобы при расчетах вторичных характеристик не возникало дополнительных ошибок.

Ввиду того, что при создании банка данных перерабатывается массовый материал наблюдений по большому числу станций, появляется возможность использования комплексного контроля, идея которого была выдвинута Л. С. Гандиным [3] в связи с контролем аэрологических наблюдений и нашла свое отражение в ряде последующих работ [2, 4]. В этих работах предлагается использовать сочетание нескольких видов контроля. Так, например, для приземных полей можно использовать контроль пространственной совместимости информации в комплексе с контролем ее временно́й совместимости. Такого рода контроль целесообразно применять к рядам, составленным из срочных или суточных значений элемента. Если ряд состоит из месячных значений, то временной контроль использовать нецелесообразно, ввиду малой связности месячных значений во времени.

Для использования предлагаемых видов контроля необходимо знать допустимую невязку, величина которой определяет ошибочна информация или нет. В настоящей работе при использовании срочных данных приземного поля давления путем статистической обработки относительных ошибок сопоставления информации с расчетными значениями находится критерий, величина которого дает основание считать сомнительным значение давления на контролируемой станции.

Исходные данные для расчетов взяты из телеграмм, получаемых СЗ УГКС через каждые 3 ч на 505 станциях северного полушария за 3 и 4 января 1977 г.

Контролю подвергались все станции в порядке их расположения. Сначала применялся пространственный контроль, а затем временной.

Пространственный контроль основывается на оптимальной интерполяции данных [4]. Данные окружающих станций интерполируются на станцию, ряд которой проверяется, и затем сопоставляется наблюденное значение ($f_н$) метеорологического элемента с результатом интерполяции ($f_р$). Если разность $|f_р - f_н|$ превышает по абсолютной величине некоторую допустимую невязку θ, то наблюденное значение считается сомнительным. Алгоритм пространственного контроля реализован в виде программы контроля поля приземного давления, за основу которой взята базовая программа объективного анализа метеорологических полей, разработанная Ю. М. Либерманом [6].

Основное отличие программы контроля от базовой программы анализа состоит в том, что оптимальная интерполяция производится не в узел регулярной сетки, а на контролируемую станцию, имеющую свои координаты.

Блок поиска станций, влияющих на контролируемую станцию, основан на принципе метода машинной карты, предложенной С. Л. Белоусовым [2]. В этом блоке узел географической сетки заменен на контролируемую станцию, которая автоматически исключается из числа влияющих. Так как в нашем случае используется довольно густая сеть станций, то количество влияющих ограничено шестью станциями. При поиске учитывается требование симметричности станций относительно контролируемой.

Значение давления на контролируемой станции определяется по формуле

$$f_р = \bar{f}_0 + \sigma_0 \sum_{i=1}^{n} p_i \frac{f_i - \bar{f}_i}{\sigma_i}, \tag{1}$$

где p_i — интерполяционный вес, σ — среднее квадратическое отклонение элемента, n — число станций, влияющих на контролируемую, черта сверху означает осреднение.

Система уравнений для определения интерполяционных весов может быть записана в виде

$$\sum_{j=1}^{n} p_i \mu_{ij} + \eta^2 p_i = \mu_{0i}, \quad i = 1, 2, \ldots, n, \tag{2}$$

где μ_{ij} и μ_{0i} — значения автокорреляционной функции рассматриваемого элемента, которую в нашем случае будем считать однородной и изотропной, т. е. зависящей только от расстояния. Меру ошибки наблюдений η^2 считаем постоянной для этого района и согласно оценкам [5] примем равной 0,01, что соответствует средней квадратической ошибке исходных данных около 1 мбар при изменении дисперсии поля давления в рассматриваемом районе от 4 до 6 мбар.

Пространственную корреляционную функцию поля приземного давления можно считать также постоянной для всего района. В этом случае пространственная корреляционная функция для января [7] имеет следующий вид:

$$\mu(\rho)=\left(1+\frac{\rho}{1,05}\right)e^{-\rho/1,05}, \tag{3}$$

где ρ — в тыс. км.

Решение системы уравнений вида (2) для определения интерполяционных весов выполняется методом исключения с выбором главного элемента по столбцу. Веса, найденные решением системы (2), служат для интерполяции отклонений от норм в формуле (1). С целью повышения степени однородности и изотропности поля давления при интерполяции значения отклонений давления (f_i) от норм ($\overline{f_i}$) нормируются по изменчивости элемента.

Поэтому для выполнения интерполяции необходимо знать нормы и изменчивости на контролируемых станциях. В качестве норм и изменчивостей используют средние многолетние средние месячные поля приземного давления в узлах географической сетки с шагом по широте и долготе, равным 10°, и заданных вдоль параллелей в направлении от экватора и к востоку от Грпнвича. Значения $\overline{f}$ и σ в узлах этой сетки заимствованы из работы [1].

В формулу (1) входят нормы и изменчивости, полученные путем аппроксимации их полиномом

$$f(\varphi, \lambda)=A\lambda^2+B\varphi\lambda+C\varphi^2+D\lambda+E\varphi+F, \tag{4}$$

т. е. предполагается квадратическая зависимость полей $\overline{f}$ и σ от φ и λ. Если известны значения $\overline{f}$ и σ в узлах регулярной сетки с шагом $\Delta\varphi=\Delta\lambda=10°$, то, используя метод наименьших квадратов, легко получить выражения для коэффициентов полинома. Эта процедура подробно изложена в работе [6].

Мерой точности оптимальной интерполяции служит величина

$$\varepsilon_{\text{ин}}^2=1-\sum_{i=1}^{n} p_i\mu_{0i}, \tag{5}$$

названная мерой ошибки интерполяции. В данном случае нам необходимо оценить меру ошибки сопоставления фактического и проинтерполированного значений, которая представляет собой

$$\varepsilon_{\text{соп}}^2=\varepsilon_{\text{ин}}^2+\eta_i^2. \tag{6}$$

Тогда средняя квадратическая ошибка сопоставления имеет следующий вид:

$$E_{\text{соп}}=\sigma_0\varepsilon_{\text{соп}}. \tag{7}$$

В качестве допустимой невязки горизонтального контроля применяют величину $0=KE_{\text{соп}}$, где K — критерий сомнительности.

Нам необходимо было выяснить, какое значение K должно быть использовано в случае горизонтального контроля срочного поля приземного давления. С этой целью произведен горизонтальный контроль на 505 станциях за восемь сроков 3 и 4 января 1977 г. в районе от 20° в. д. до 90° з. д. и от 85° до 30° с. ш. Получены ошибки между расчетным и фактическим материалом и нормированы на среднюю квадратическую ошибку сопоставления (в отличие от [2], где разность нормировалась на $E_{\text{ин}}$). Таким образом, получен ряд относительных ошибок, который затем подвергается статистической обработке.

На рис. 1 представлено фактическое распределение ошибок, которое с большой точностью подчиняется нормальному закону. Всякое теоретическое распределение характеризуется своими основными параметрами: средним значением и дисперсией. В нашем случае среднее значение близко к нулю, а среднее квадратическое отклонение равно 1,6. Входя с этими параметрами в функцию распределения случайной величины при нормальном законе, получаем теоретическую кривую (рис. 1).

Из рисунка видно, что полученная теоретическая кривая не совсем точно описывает фактическое распределение. Вполне естественно, что можно изменить вид кривой, уменьшив величину σ, но взять ее совсем малой нельзя, так как получим, что почти вся информация будет находиться в зоне грубых ошибок. Поэтому, уменьшая σ, будем смотреть, чтобы теоретическая кривая как можно лучше описывала фактическую. Такая картина получается в случае $\sigma = 1{,}3$ (рис. 1).

Таким образом, можно считать информацию сомнительной, если абсолютная разность между расчетным и истинным значением будет превышать $4{,}0E_{\text{соп}}$, где значение $K = 4$ получено как 3σ. Однако ошибка горизонтального контроля может превышать допустимую и тогда, когда ошибочным является значение не на контролируемой станции, а на одной из соседних, использованной в качестве влияющей. Поэтому прежде чем исключить сомнительную величину (или исправить ее) необходимо выяснить, данные какой именно станции ошибочны. Указанное обстоятельство и создает необходимость комплексного контроля. В этом случае наряду с горизонтальным контролем целесообразно обратиться к временно́му.

Временно́й контроль предполагает интерполяцию метеорологического элемента во времени и затем также сравнение проинтерполированного значения с измеренным. Допустимая невязка будет зависеть как от временной изменчивости контролируемого элемента, так и от способа интерполяции. Исследования показали [4], что применение наиболее простой схемы линейной интерполяции по данным за два срока дает хорошие результаты:

$$f_{\text{р}}(t) = \frac{1}{2}\left[\widetilde{f}(t+\tau) + \widetilde{f}(t-\tau)\right], \qquad (8)$$

где τ — интервал между сроками; $f_{\text{р}}(t)$ — результат интерполяции, а $\widetilde{f}(t \pm \tau)$ — наблюденное значение давления на станции соответственно за предыдущий и последующий сроки. Это значение представляет собой сумму истинного значения давления $f(t+\tau)$, $f(t-\tau)$ и ошибки наблюдения его $\Delta(t+\tau)$ и $\Delta(t-\tau)$.

Это необходимо иметь в виду при определении меры ошибки сопоставления $\varepsilon^2_{\text{соп}}$, которая связана со средним квадратом ошибки формулой (7). Используя соотношение (7) и формулу (8) получаем выражение для линейной интерполяции по двум точкам для

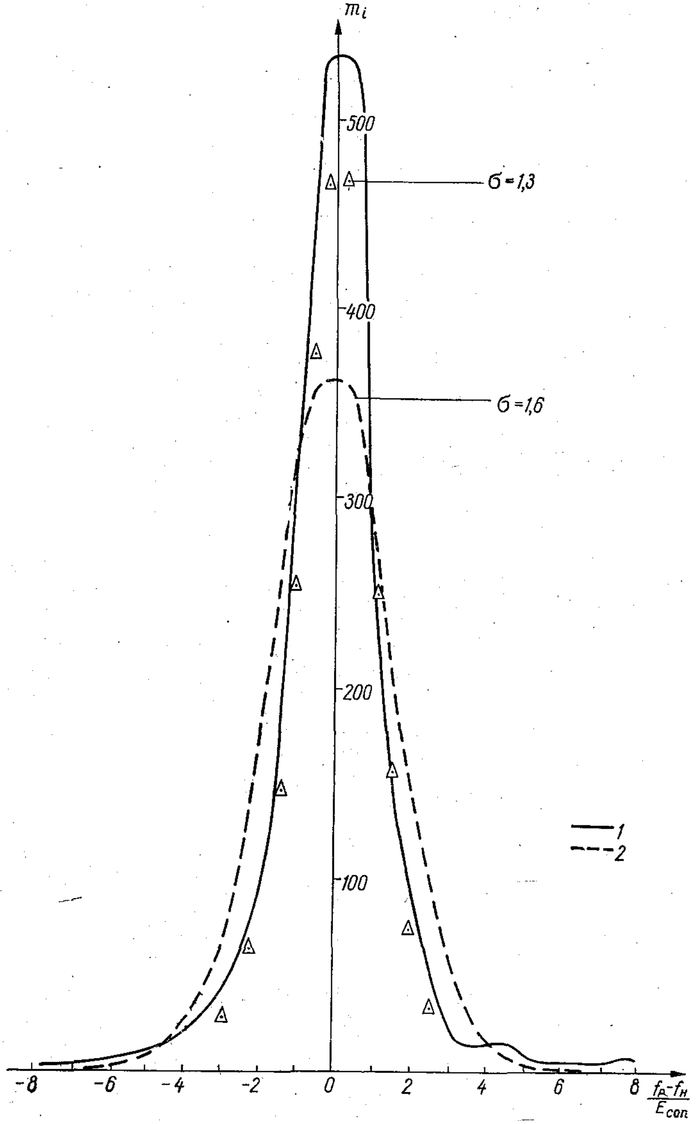

Рис. 1. Фактическое (*1*) и теоретическое (*2*) распределение относительных ошибок при горизонтальном контроле.

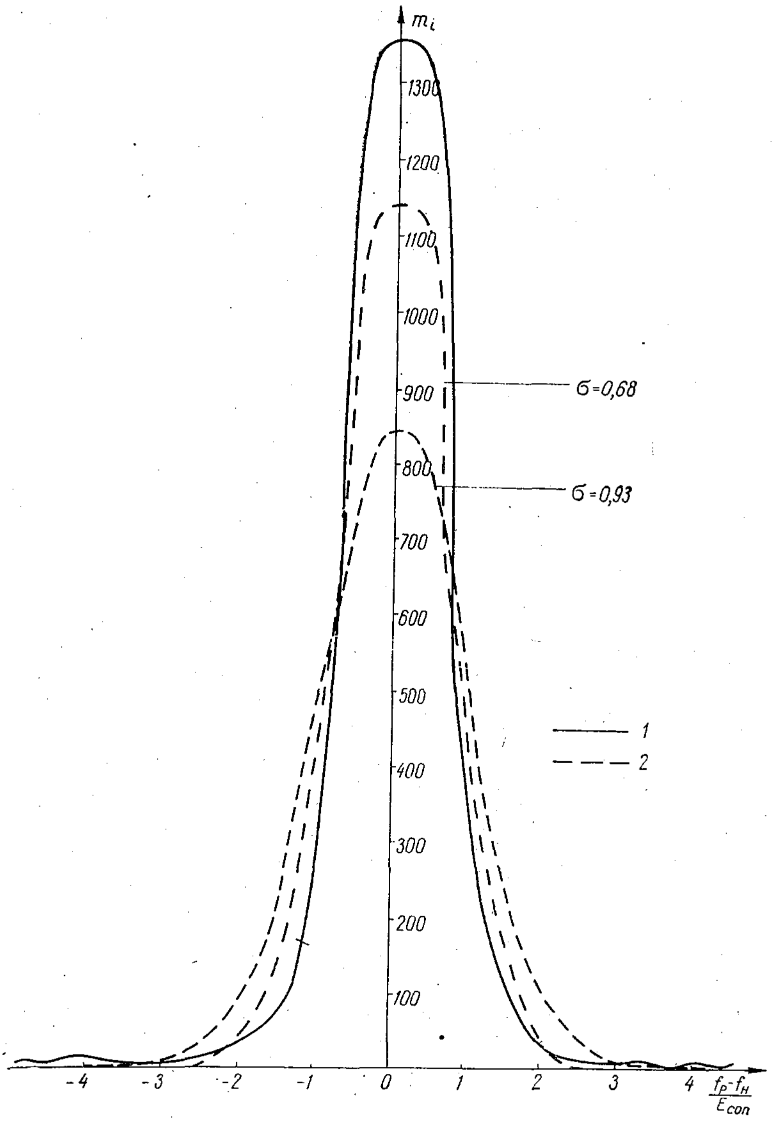

Рис. 2. Фактическое (*1*) и теоретическое (2) распределение относительных ошибок при временном контроле.

каждого конкретного случая. Если нет пропусков наблюдений и имеются значения давления для двух соседних сроков, то формула имеет следующий вид:

$$\varepsilon^2_{\text{соп}}(\tau)=\frac{3}{2}(1+\eta^2)-2\mu(\tau)+\frac{1}{2}\mu(2\tau). \qquad (9)$$

Если имеется наблюдение за предыдущий и нет последующего срока, то для расчетов привлекается значение давления за следующий срок. В этом случае формула имеет вид

$$\varepsilon^2_{\text{соп}}(\tau) = \frac{14}{9}(1+\eta^2) - \frac{4}{3}\mu(\tau) - \frac{2}{3}\mu(2\tau) + \frac{4}{9}\mu(3\tau). \qquad (10)$$

Если пропущены последующие два срока, то формула изменяется,

$$\varepsilon^2_{\text{соп}}(\tau) = \frac{38}{25}(1+\eta^2) + \frac{12}{25}\mu(5\tau) - \frac{20}{25}\mu(3\tau) - \frac{30}{25}\mu(2\tau). \qquad (11)$$

В этих формулах $\mu(\tau)$ — временная корреляционная функция, которая, согласно работе [7] для рассматриваемого района имеет следующий вид:

$$\mu(\tau) = r_0'(\tau)\left(1 + \frac{\tau}{30}\right)e^{-\frac{\tau}{30}},$$

где τ в часах.

Аналогично горизонтальному контролю проводился временной контроль поля приземного давления для указанного выше района по 10 срокам наблюдения. В результате получен ряд относительных ошибок, который также подвергался статистической обработке. Фактическая и теоретическая кривые представлены на рис. 2. На основании анализа кривых распределения относительных ошибок можно прийти к выводу, что сомнительной в данном случае можно считать информацию, если разность между расчетным и исходным значениями превышает $2E_{\text{соп}}$, где значение коэффициента $K = 2$ получено как произведение 3σ (при $\sigma = 0{,}7$).

В дальнейшем предполагается, используя полученные критерии отбраковки для горизонтального и временного контроля, локализовать ошибки наблюдения, установить их причины, исправить их, а если это невозможно, то отбраковать ошибочную информацию.

СПИСОК ЛИТЕРАТУРЫ

1. Атлас климатических характеристик температуры, плотности и давления воздуха, ветра и геопотенциала в тропосфере и нижней стратосфере северного полушария.— 1975, вып. 1, Гидрометцентр СССР.— 210 с.
2. Белоусов С. Л., Гандин Л. С., Машкович С. А. Обработка оперативной метеорологической информации с помощью электронных вычислительных машин.— Л., Гидрометеоиздат, 1968.— 192 с.
3. Гандии Л. С. и др. О статистическом контроле аэрологических телеграмм.— Труды ГГО, 1964, вып. 151, с. 3—16.
4. Гандин Л. С., Каган Р. Л. Статистические методы интерпретации метеорологических данных.— Л., Гидрометеоиздат, 1976.— 359 с.
5. Гандин Л. С., Мелешко В. П., Мещерская А. В. О применении универсальных цифровых машин для исследования статистической структуры метеорологических полей.—Труды ГГО, 1963, вып. 143, с. 113—129.
6. Либерман Ю. М. Эксперимент по использованию прогностической информации в объективном анализе.— Труды ГГО, 1977, вып. 397, с. 99—106.
7. Лугина К. М., Каган Р. Л. К вопросу о пространственно-временном анализе барического поля.— Труды ГГО, 1974, вып. 336, с. 75—94.

Л. П. Клягина, Э. М. Скворцова

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МЕТОДА ОПТИМАЛЬНОЙ ИНТЕРПОЛЯЦИИ ДЛЯ КОНТРОЛЯ СРЕДНЕЙ МЕСЯЧНОЙ ТЕМПЕРАТУРЫ ВОЗДУХА

Ряды метеорологических наблюдений, используемые для климатологических обобщений, составляются на основе данных первичных наблюдений, осредненных за определенные промежутки времени (сутки, пентаду, декаду, месяц, год и т. д.).

Присутствие ошибок в климатологических рядах искажает результаты статистической обработки, поэтому необходимо проводить их контроль.

До недавнего времени контроль климатологической информации проводился вручную, субъективным путем. С внедрением ЭВМ в практику климатологической обработки появилась возможность автоматического контроля, существенное преимущество которого заключается в его объективности. К настоящему времени разработаны различные методы контроля климатологической информации. Достаточно полный обзор таких методов приведен в работе [4].

В настоящей статье для контроля средней месячной температуры воздуха использован пространственный (горизонтальный) контроль, где использован метод оптимальной интерполяции [3]. Алгоритм пространственного контроля реализован в виде программы, за основу которой взята программа горизонтального контроля поля приземного давления [6].

Интерполируемое значение средней месячной температуры воздуха вычисляется по формуле

$$t_0=\bar{t}_0+\sigma_0\sum_{i=1}^{n}p_i\frac{t'_i}{\sigma_i}, \tag{1}$$

где $\bar{t}_0$ — норма средней месячной температуры воздуха; $t'_i = t_i-\bar{t}_i$ — отклонение средней месячной температуры воздуха от нормы; σ_i — среднее квадратическое отклонение средних месячных температур; n — число влияющих станций, данные которых тесно коррелируют с данными контролируемой станции; p_i — интерполяционные веса.

Применять оптимальную интерполяцию для значений самого элемента неудобно, так как нельзя считать поле средней месячной температуры воздуха статистически однородным и изотропным. Поэтому в расчетах участвуют не сами значения метеорологических величин, а их отклонения от нормы.

Значения интерполяционных весов определяются путем решения системы линейных уравнений

$$\sum_{i=1}^{n} \mu_{ij} p_i + \eta^2 p_i = \mu_{0i} \quad (j=1,\ 2,\ \ldots,\ 6), \tag{2}$$

где μ_{ij} — корреляционная функция, зависящая только от расстояния между станциями; $\eta^2 = \frac{\Delta^2}{\sigma^2}$ — мера ошибки определения средней месячной температуры воздуха.

В наших расчетах η принимается одинаковой для каждой из n станций.

Данные о статистической структуре полей средней месячной температуры позволяют не только выполнять интерполяцию на конкретную станцию, но и оценить среднюю квадратическую разность между проинтерполированным и наблюденным значениями. Она равна средней квадратической ошибке сопоставления

$$E_{\text{соп}} = \sigma_0 \left[\left(1 - \sum_{i=1}^{n} p_i \mu_{0i} \right) + \eta_i^2 \right]^{1/2}, \tag{3}$$

которая зависит в основном от густоты сети станций.

Мерой ошибки сопоставления служит величина

$$\varepsilon_{\text{соп}} = \frac{E_{\text{соп}}}{\sigma_0}. \tag{4}$$

В решении системы уравнений (2) участвуют не более шести станций, расположенных на расстоянии не менее 12 км и не более 1700 км. Последнее ограничение связано с тем, что на расстоянии более 1700 км корреляционная функция температуры воздуха [8], используемая в настоящей работе, затухает, т. е. приближается к нулю.

В работе использованы значения норм температуры воздуха и ее изменчивости в узлах географической сетки с шагом 10° ($\Delta\varphi = \Delta\lambda = 10°$), заданные вдоль параллелей в направлении от экватора и к востоку от Гринвича [1]. Значения средней месячной температуры воздуха взяты из [5].

Для использования метода оптимальной интерполяции в целях контроля средней месячной температуры воздуха необходимо знание статистической структуры этого элемента. Достаточно полно пространственная корреляция средней месячной температуры воздуха исследована в работах [7, 8].

В данной работе использована корреляционная функция вида

$$r'(\rho)=r'(0)\,e^{\left(-\frac{\rho}{\rho_0}\right)^n}, \tag{5}$$

где $r'(\rho)$ — эмпирически определенное значение корреляционной функции; $r'(0)$ — значение эмпирической корреляционной функции, полученное путем экстраполяции на нулевое расстояние; ρ — расстояние; ρ_0 — радиус корреляции; n — структурный параметр, принимаемый, согласно работе [8], для зимы равным 1,49, а для лета 1,67.

С помощью метода оптимальной интерполяции по данным о температуре воздуха на окружающих станциях вычисляется значение температуры на проверяемой станции. Затем находится разность между рассчитанным значением температуры $t_р$ и наблюденным $t_н$. Если эта разность оказывается большой, то можно высказать предположение о наличии ошибки. Большая величина Δt может быть обусловлена как погрешностью в наблюдениях и их обработке, так и ошибкой интерполяции. Последняя зависит от расстояния между проверяемой и окружающими станциями, от взаимного их расположения, а также от характера затухания корреляционной функции.

Величина средней квадратической ошибки сопоставления $E_{соп}$ позволяет судить о том, какой в среднем должна быть ошибка сопоставления. Если сопоставить невязку Δt и теоретическую ошибку сопоставления, то можно установить, существуют ли ошибки в исходной информации. В качестве допустимой невязки горизонтального контроля применяют величину $KE_{соп}$, где K — константа.

В работе [2] при контроле оперативной информации для выбора значения параметра K была проведена серия специальных численных экспериментов, где последовательно задавались разные значения K. В нашем случае имеется в распоряжении архив средней месячной погодичной температуры воздуха, поэтому стало возможным определить параметр K на основании закона распределения относительной величины $\frac{\Delta t}{E_{соп}}$. Для этой цели использованы значения средней месячной температуры воздуха на 158 станциях, расположенных на территории СССР в районе между 61 и 49° с. ш. и 37 и 57° в. д. за период 1951—1970 гг. для января и июля.

Вычисления проводились для значений $\eta^2_{I}=0{,}002$ и $\eta^2_{VII}=0{,}008$, что соответствует ошибке определения средней месячной температуры воздуха 0,1 и 0,2 °С.

Результаты расчетов показывают, что для рассматриваемой территории интерполяция может быть проведена с ошибкой $\varepsilon<0{,}14$ для лета и $\varepsilon<0{,}10$ для зимы, что не превышает погрешности определения средних месячных значений температуры воздуха.

На рис. 1 представлены для января и июля эмпирические кривые распределения относительных ошибок. Из рисунка видно, что эти кривые подчиняются нормальному закону распределения, что дает возможность по статистикам, полученным на основании этого ряда, построить теоретическую кривую распределения. На рисунке область грубых ошибок лежит при $K > 3$. Причем параметр K меняется от 4 зимой до 5 летом. Если учесть, что зимой в среднем для всей рассматриваемой территории средняя квадратическая ошибка сопоставления температуры воздуха $E_{\text{соп}} \approx 0{,}4 \div 0{,}5$, а летом $E_{\text{соп}} \approx 0{,}3$, то допустимые разности между интерполированным и наблюденным значениями будут составлять около 1,5 °С. Но для каждой станции рассчитываются свои значения средней квадратической ошибки, поэтому и допустимые разности будут колебаться, однако в любом случае они не должны превышать $|\Delta t| < K E_{\text{соп}}$.

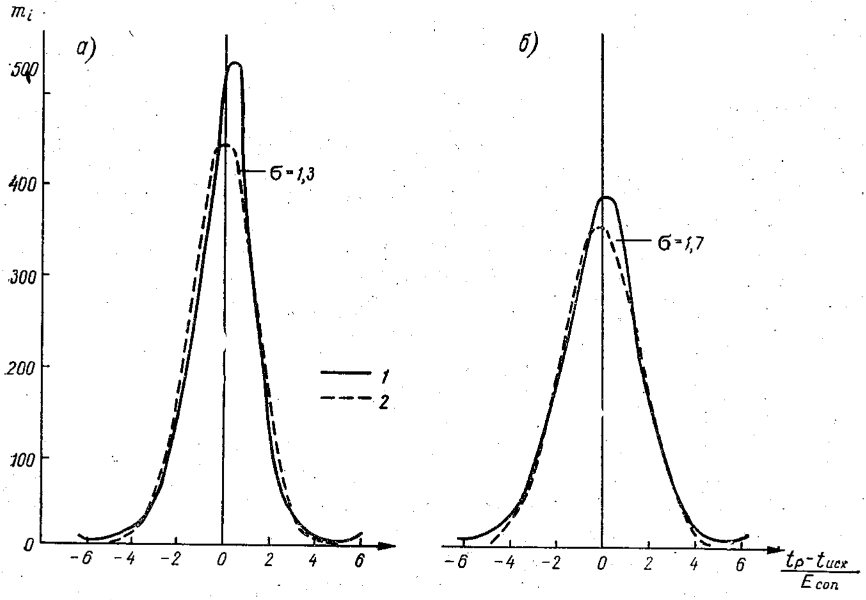

Рис. 1. Эмпирическая (*1*) и теоретическая (*2*) кривые распределения ошибок интерполяции для января (*а*) и июля (*б*).

Контроль средней месячной температуры воздуха на ЕТС показал, что количество ошибочных данных составляет около 2 % общего числа случаев.

Причем для проверки метода в исследуемые ряды искусственно вносились ошибки, которые были также обнаружены.

Систематические ошибки отмечены на двух станциях (Починки за период 1959—1970 гг. и Фурманово за период 1955—1970 гг.) для января и на двух станциях (Аиучино и Туймазы за весь период наблюдений) для июля.

На значения средней месячной температуры воздуха большое влияние оказывает подстилающая поверхность. Использованный метод оптимальной интерполяции, очевидно, недостаточно учитывает этот факт, что приводит к отбраковке правильных данных при контроле.

Например, на ст. Туймазы во все годы значения средней месячной температуры выделяются как ошибочные. Анализ этих температур на ст. Туймазы и влияющих на нее станциях показал, что невязки возникают вследствие того, что при интерполяции большой вклад в результат вносят данные станций, находящихся на возвышенности (Бугульма, Аксаково) и поэтому имеющих более низкие температуры.

Таким образом, при контроле средней месячной температуры воздуха методом оптимальной интерполяции необходимо совершенствование метода подбора влияющих станций с учетом характера подстилающей поверхности.

СПИСОК ЛИТЕРАТУРЫ

1. Атлас климатических характеристик температуры, плотности и давления воздуха, ветра и геопотенциала в тропосфере и нижней стратосфере северного полушария/Под ред. Д. И. Стехновского.— М., Гидрометцентр СССР, ВНИИГМИ—МЦД, 1975.— 210 с.
2. Белоусов С. Л., Гандин Л. С., Машкович С. Л. Обработка оперативной метеорологической информации с помощью ЭВМ.— Л.: Гидрометеоиздат, 1968.— 282 с.
3. Гандин Л. С. Объективный анализ метеорологических полей.— Л.: Гидрометеоиздат, 1963.— 287 с.
4. Гандин Л. С., Каган Р. Л. Статистические методы интерпретации метеорологических данных.— Л.: Гидрометеоиздат, 1976.— 232 с.
5. Справочник по климату СССР, ч. П, вып. 8—13, 18, 28, 29.— Л.: Гидрометеоиздат, 1965—1967.
6. Клягина Л. П. Определение параметров комплексного контроля поля приземного давления.— См. наст. сб.
7. Лугина К. М., Малашенко Л. Я. Пространственная корреляция аномалий температуры воздуха и использование ее при рационализации сети станций.— Труды ГГО, 1972, вып. 286, с. 26—38.
8. Лугина К. М., Тараканова В. П. Пространственная структура поля средней месячной температуры.— В кн.: Исследование статистической структуры метеорологических полей. (Материалы международного симпозиума специалистов гидрометслужб социалистических стран.) Том. 1. М.: Гидрометеоиздат, 1975, с. 73—79.

Л. С. Гандин, В. П. Тараканова

К МЕТОДИКЕ СОГЛАСОВАНИЯ НАБЛЮДЕНИЙ С ДАННЫМИ ЧИСЛЕННОГО ПРОГНОЗА

1. Во многих применяемых в настоящее время методах объективного анализа метеорологических полей в качестве исходной информации наряду с данными наблюдений используются результаты численного прогноза на рассматриваемый срок. Так, в методе полиномиальной аппроксимации данные численного прогноза привлекаются при определении коэффициентов полинома методом наименьших квадратов [13]. В широко распространенном методе последовательных коррекций [11, 12] прогностическая информация привлекается для построения предварительного поля, а данные наблюдений используются для внесения исправлений в него. В последние годы разработан также ряд методов объективного анализа, базирующихся на оптимальной интерполяции отклонений данных наблюдений от результатов прогноза на рассматриваемый срок. Таков метод, предложенный Х. Кругером [14] и применяющийся в метеорологической службе Канады, а также методы совместного анализа геопотенциала и ветра, разработанные Т. Шлаттером [18] и Е. Резерфордом [17]. Аналогичный подход использовался в численных экспериментах по четырехмерному анализу и по планированию систем наблюдений, выполненных в работах К. Мияcoда и О. Талаграна [15], [19], Л. Бенгтссона и Н. Густавссона [9], [10], Л. Бенгтссона [8] и др.

Опыт показал, что оптимальная интерполяция отклонений наблюденных значений от прогностических обеспечивает достаточно высокое качество объективного анализа. Вместе с тем при таком подходе не учитываются специфические свойства прогностических полей, отличающие их как от наблюденных, так и от истинных полей. Поэтому имеется опасность, что в районах с редкой сетью наблюдений, где прогностическая информация существенно влияет на результат анализа, точность последнего может оказаться недостаточно высокой. С этой точки зрения заслуживает предпочтения подход, состоящий в согласовании отклонений наблюденных и прогностических значений от климатических норм, т. е. подход, при котором наряду с наблюдениями и прогнозом используется также климатическая информация.

Весьма простой метод согласования такого рода был предложен в работе [3] (см. также [2] и [4]). Он базируется на двух предположениях. Первое из них состоит в том, что прогноз статистически совершенен, а именно не может быть улучшен путем прямого или обратного сглаживания прогностических аномалий. Иначе говоря, если попытаться заменить прогностическое значение $f'_{0п}$ отклонения элемента f от нормы в некоторой точке «исправленным» значением

$$f'_{0п}=kf'_{0п}+\sum_{i=1}^{n} c_i f'_{iп}, \tag{1}$$

где $f'_{iп}$ — прогностические аномалии в других точках, то наименьшее среднее квадратическое отклонение значения $f_{0п}$ от истинного получится при весах $k=1$ и $c_i=0$, т. е. когда «исправленное» прогностическое значение равно исходному. Заметим, что указанным свойствам действительно обладает статистический прогноз, базирующийся на принципе наименьших квадратов, при условии, что прогностические значения во всех точках получены по одной и той же исходной информации. При этом условии обладает подобным свойством и результат оптимальной интерполяции.

Второе предположение состоит в том, что корреляционная функция ошибок прогноза совпадает с корреляционной функцией истинных значений элемента f. Из этих двух предположений следуют простые формулы одноточечного согласования интерполированных и прогностических значений, т. е такого, что согласованное значение отклонения f' получится в виде линейной комбинации результата оптимальной интерполяции в рассматриваемую точку и прогностического значения f' в ней.

Эксперименты, выполненные Ю. М. Либерманом [6], показали, что применение данного метода может заметно увеличить точность анализа в районах, где данных наблюдений мало, а ошибки прогноза сравнительно невелики. Однако из самых общих соображений следует, что корреляционная функция ошибок прогноза должна не совпадать с корреляционной функцией истинных значений, а затухать существенно быстрее с ростом расстояния между точками. Этот вывод подтверждается всеми эмпирическими данными [9, 14, 16] и означает, что второе из перечисленных выше предположений недостаточно корректно. Вместе с тем, если просто отказаться от него и воспользоваться эмпирически определенной корреляционной функцией ошибок прогноза, то имеется реальная опасность нарушения положительной определенности матриц при определении весов для согласования наблюдений и прогноза, а потому получения вообще абсурдных результатов.

С целью преодоления этой трудности Р. Л. Каган [5] предложил подход, используемый и в настоящей работе. Его идея заключается в моделировании корреляционной функции ошибок численного прогноза корреляционной функцией ошибок статистической экстраполяции во времени. Последняя состоит в статистически оптимальной (в смысле принципа наименьших квадратов) оценке прогностического значения в момент t в каждой рассматриваемой точке по данным наблюдений в момент $t-\tau$ в системе точек, расположенных заданным образом по отношению к рассматриваемой.

Заметим, что поскольку при такой процедуре прогностические значения в разных точках получаются при использовании различной информации, она не является статистически совершенной в указанном выше смысле. Вместо этого выполняется менее сильное ограничение: если в формуле (1) заранее считать все веса c_i равными нулю, то оптимальное значение веса k окажется равным единице.

Выполненные Р. Л. Каганом [5] расчеты показали, что одноточечная схема статистического прогноза, сводящаяся к чисто временной экстраполяции по данным в той же точке, недостаточна для моделирования корреляционных функций и дисперсий ошибок прогноза, близких к тем, которые имеют место в случае численного прогноза. Поэтому Каган рассмотрел более общую, пятиточечную схему, в которой используется информация в момент $t-\tau$ как в самой точке, так и в еще четырех точках, ее окружающих и расположенных симметрично по отнощению к ней. Он показал, что надлежащим выбором параметров этой схемы можно достичь хорошего согласования статистических характеристик ошибок прогноза по ней с известными из эмпирических данных характеристиками ошибок численных прогнозов.

Однако практическая реализация согласования на основе такой схемы достаточно проста лишь в случае, если ограничиться одноточечным согласованием. Такое ограничение, естественно, связано с понижением точности анализа, степень которого заранее неизвестна, и желательно этим ограничением не пользоваться. В принципе на основе пятиточечной схемы Кагана можно выполнять и непосредственное согласование наблюдений и прогноза, однако процедура такого согласования требует вычисления большого количества пространственно-временных расстояний и потому слишком громоздка.

Этого недостатка можно избежать, если исходить из схемы статистического прогноза, инвариантной относительно любого поворота координат. Такой путь и используется в настоящей работе.

2. Пусть статистический прогноз выполняется по данным во всех точках окружности радиуса r, окружающих рассматриваемую точку, т. е. по формуле

$$f_п(t,\ 0)=\frac{\bar{a}}{2\pi}\int\limits_0^{2\pi} f(t-\tau,\ r,\ \vartheta)\,d\vartheta, \tag{2}$$

где введены полярные координаты r, ϑ с полюсом в рассматриваемой точке (при $r=0$ координата ϑ становится неопределенной и поэтому опускается), f и $f_п$ — истинное и прогностическое значения отклонения метеорологического элемента от нормы, a — экстраполяционный вес.

Дисперсия ошибок такого прогноза

$$E_п^2=\overline{[f_п(t,\ 0)-f(t,\ 0)]^2}, \qquad (3)$$

где черта означает статистическое осреднение, равна, согласно (2),

$$E_п^2=\sigma^2-2M_{п0}+\sigma_п^2, \qquad (4)$$

где

$$\sigma^2=\overline{f^2(t,\ 0)} \qquad (5)$$

— дисперсия истинного значения элемента f,

$$M_{п0}=\overline{f_п(t,\ 0)\,f(t,\ 0)}=\frac{a}{2\pi}\int\limits_0^{2\pi}\overline{f(t-\tau,\ r,\ \vartheta)\,f(t,\ 0)}\,d\vartheta \qquad (6)$$

— ковариация истинного и прогностического значений и

$$\sigma_п^2=\overline{f_п^2(t,\ 0)}=\frac{a^2}{4\pi^2}\int\limits_0^{2\pi}\int\limits_0^{2\pi}\overline{f(t-\tau,\ r,\ \vartheta)\,f(t-\tau,\ r,\ \vartheta')}\,d\vartheta\,d\vartheta' \qquad (7)$$

— дисперсия прогностического значения.

Разделим (4) на σ^2 и введем обозначения

$$\varepsilon_п^2=E_п^2/\sigma^2, \qquad (8)$$

$$S_п^2=\sigma_п^2/\sigma^2, \qquad (9)$$

$$\mu_{п0}=M_{п0}/\sigma^2. \qquad (10)$$

Тогда получим

$$\varepsilon_п^2=1-2\mu_{п0}+S_п^2, \qquad (11)$$

где

$$\mu_{п0}=\frac{a}{2\pi\sigma^2}\int\limits_0^{2\pi}\overline{f(t-\tau,\ r,\ \vartheta)\,f(t,\ 0)}\,d\vartheta \qquad (12)$$

и

$$S_п^2=\frac{a^2}{4\pi^2\sigma^2}\int\limits_0^{2\pi}\int\limits_0^{2\pi}\overline{f(t-\tau,\ r,\ \vartheta)\,f(t-\tau,\ r,\ \vartheta')}\,d\vartheta\,d\vartheta'. \qquad (13)$$

Предположим, что пространственно-временна́я корреляционная функция μ элемента f однородна по горизонтали и стационарна по времени, т. е. что

$$\frac{1}{\sigma^2}\overline{f(\mathbf{r},\ t)\,f(\mathbf{r}+\boldsymbol{\rho},\ t+\tau)}=\mu(\rho,\ |\tau|). \qquad (14)$$

Подставляя (14) в (12) и (13), легко показать, что

$$\mu_{\text{п}0} = a\mu(\tau,\ r) \tag{15}$$

и что величину $S^2_{\text{п}}$ можно представить в виде

$$S^2_{\text{п}} = a^2 I_1(r), \tag{16}$$

где

$$I_1(r) = \frac{1}{\pi} \int\limits_0^{\pi} \mu\left(2r \sin\frac{\vartheta}{2},\ 0\right) d\vartheta. \tag{17}$$

Подставляя (15) и (16) в (11), получим

$$\varepsilon^2_{\text{п}} = 1 - 2a\mu(r,\ \tau) + a^2 I_1(r). \tag{18}$$

Все предыдущее было верно для любого значения весового множителя a. Выберем его теперь так, чтобы мера ошибки прогноза $\varepsilon^2_{\text{п}}$ была минимальной. Тогда получим

$$a(r,\ \tau) = \mu(r,\ \tau)/I_1(r), \tag{19}$$

$$S^2_{\text{п}}(r,\ \tau) = \mu_{\text{п}0}(r,\ \tau) = a(r,\ \tau)\,\mu(r,\ \tau), \tag{20}$$

$$\varepsilon^2_{\text{п}}(r,\ \tau) = 1 - \mu_{\text{п}0}(r,\ \tau). \tag{21}$$

3. Пусть для получения согласованного с прогнозом значения в некоторой точке используется, в дополнение к данным наблюдений в n точках, прогностическое значение только в интересующей нас точке, т. е. согласование производится по формуле

$$\hat{f}_0 = \sum_{i=0}^{n} a_i \tilde{f}_i + a_{\text{п}} f_{\text{п}0}, \tag{22}$$

где $\tilde{f}_i$ — отклонения наблюденных значений от норм, $\hat{f}_0$ — отклонение согласованного значения от нормы, a_i $(i = 1,\ 2,\ \ldots,\ n)$ и $a_{\text{п}}$ — весовые множители.

Рассуждая аналогично предыдущему, для меры ошибки согласования, производимого по формуле (22), нетрудно получить следующую формулу:

$$\hat{\varepsilon}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \tilde{\mu}_{ij} + a^2_{\text{п}} S^2_{\text{п}0} + 2 \sum_{i=1}^{n} a_i a_{\text{п}} \varkappa_{i0} - 2 \sum_{i=1}^{n} a_i \mu_{i0} - 2a_{\text{п}}\mu_{\text{п}0} + 1, \tag{23}$$

где

$$\tilde{\mu}_{ij} = \frac{1}{\sigma^2} \overline{\tilde{f}_i \tilde{f}_j}, \tag{24}$$

$$\mu_{i0} = \frac{1}{\sigma^2} \overline{f_i f_0}, \tag{25}$$

$$\varkappa_{i0} = \frac{1}{\sigma^2} \overline{f_i f_{\text{п}0}}. \tag{26}$$

Величины μ_{i0} представляют собой значения корреляционной функции $\mu(\rho, 0)$ (14) для расстояний ρ_{i0} от соответствующего пункта наблюдений до интересующей нас точки. Величины $\tilde{\mu}_{ij}$ отличаются от μ_{i0} тем, что они описывают связь между наблюденными значениями. В частности, если ошибки наблюдений δ_i и δ_j в разных пунктах некоррелированы, то

$$\tilde{\mu}_{ij}=\begin{cases}\mu_{ij} & \text{при } j\neq i,\\ 1+\eta_i^2 & \text{при } j=i,\end{cases}$$

где $\eta^2_i=\overline{\delta}^2_i/\sigma^2$ — мера ошибок наблюдений. Наконец, величины $\varkappa_{i0}$ характеризуют статистическую связь между значениями элемента f в пунктах наблюдений и прогностическим значением в интересующей нас точке.

В частном случае $a_\text{п}=0$ формула (23) описывает меру ошибки обычной интерполяции с произвольными весами a_i.

Значения весов a_i и $a_\text{п}$ для случая оптимального согласования, получатся минимизацией величины $\widehat{\varepsilon^2}$ относительно этих весов, что приводит к системе линейных уравнений

$$\sum_{j=1}^{n} a_j\tilde{\mu}_{ij}+a_0\varkappa_{i0}=\mu_{i0}\qquad (i=1,\ 2,\ \ldots,\ n),$$

$$\sum_{i=1}^{n} a_i\varkappa_{i0}+a_\text{п}S^2_{\text{п}0}=\mu_{\text{п}0}\qquad (27)$$

для их определения.

Рассмотрение уравнений (27) показывает, что, переходя от интерполяции наблюдений к согласованию их с прогнозом по формуле (22), нужно дополнительно знать, помимо рассмотренных выше величин $\mu_{\text{п}0}$ и $S^2_{\text{п}0}$, еще только значения $\varkappa_{i0}$, определенные формулой (26). Подставляя в нее $f_{\text{п}0}=f_\text{п}(t,\ 0)$ из (2) и пользуясь предположением (14), можно получить формулу

$$\varkappa(\rho;\ r,\ \tau)=a(r,\ \tau)\,I_2(\rho;\ r,\ \tau),\qquad (28)$$

где

$$I_2(\rho;\ r,\ \tau)=\frac{1}{\pi}\int\limits_0^{\pi}\mu\left(\sqrt{\rho^2+r^2+2\rho r\cos\vartheta},\ \tau\right)d\vartheta,\qquad (29)$$

а ρ — расстояние от пункта наблюдения до интересующей нас точки.

4. Рассмотренная выше схема статистического прогноза зависит от двух параметров — его заблаговременности τ и радиуса окружности r. Естественно подбирать эти параметры, исходя из требования, чтобы статистические характеристики этого прогноза наиболее близко соответствовали характеристикам применяемых численных прогнозов.

Одной из таких характеристик является мера ошибок прогноза $\varepsilon^2_{\text{п}}$, пропорциональная среднему квадрату прогностической ошибки. Задание величины $\varepsilon^2_{\text{п}}$ (21) по результатам оценок оправдываемости численных прогнозов дает одно соотношение между параметрами r и τ. Для получения недостающего второго соотношения проще всего, в принципе, было бы использовать какую-либо характеристику связи фактических и прогностических значений в разных точках, например радиус корреляции функции $\varkappa(\rho)$ (26).

Однако никаких данных об этой функции для численных прогнозов пока не имеется. Вместо этого существует, как уже упоминалось, некоторая информация о корреляционной функции ошибок численных прогнозов

$$\nu(\rho)=\frac{1}{E^2_{\text{п}}}\overline{[f_{\text{п}}(\mathbf{r})-f(\mathbf{r})]\,[f_{\text{п}}(\mathbf{r}+\boldsymbol{\rho})-f(\boldsymbol{\rho})]}. \qquad (30)$$

Поэтому с целью подбора параметров схемы статистического прогноза необходимо рассчитать корреляционную функцию ошибок прогноза по этой схеме. Преобразуя для нее (30) с помощью соотношений (2), (8), (14), (25) и (26), получим формулу

$$\nu(\rho;\ r,\ \tau)=\frac{\lambda\,(\rho;\ r,\ \tau)+\mu\,(\rho,\ 0)-2\varkappa\,(\rho;\ r,\ \tau)}{\varepsilon^2_{\text{п}}\,(r,\ \tau)}, \qquad (31)$$

где величины μ, $\varkappa$ и $\varepsilon^2_{\text{п}}$ были введены выше, а функция

$$\lambda\,(\rho)=\frac{1}{\sigma^2}\overline{f_{\text{п}}(\mathbf{r})\,f_{\text{п}}(\mathbf{r}+\boldsymbol{\rho})} \qquad (32)$$

пропорциональна корреляционной функции прогностических значений элемента f и может быть рассчитана по получающейся после подстановки (2) в (32) формуле

$$\lambda\,(\rho;\ r,\ \tau)=a^2\,(r,\ \tau)\,I_3\,(\rho;\ r,\ \tau), \qquad (33)$$

где

$$I_3\,(\rho;\ r,\ \tau)=\frac{1}{2\pi^2}\int\limits_0^{\pi}\int\limits_0^{2\pi}\times$$

$$\times\,\mu\left(\sqrt{\rho^2+4r^2\sin^2\frac{\vartheta'-\vartheta}{2}-2\rho r\,(\cos\theta'-\cos\vartheta)},\ 0\right)d\vartheta'\,d\vartheta. \qquad (34)$$

5. На основании приведенных формул были с помощью ЭВМ выполнены расчеты функций $\varkappa(\rho)$, $\lambda(\rho)$ и $\nu(\rho)$, а также констант схемы статистического прогноза a, $\mu_{\text{п}0}=S^2_{\text{п}0}$ и $\varepsilon_{\text{п}}$ при различных значениях параметров τ и r. В этих расчетах использовалась предложенная в работе [7] (см. также [1]) аппроксимация пространственно-временной корреляционной функции приземного давления

$$\mu\,(\rho,\ \tau)=e^{-C\sqrt{\rho^2+V^2\tau^2}}\left(1+C\sqrt{\rho^2+V^2\tau^2}\right), \qquad (35)$$

содержащая два параметра C и V, так что результаты расчетов зависят также от этих параметров. Вычисление интегралов I_1, I_2 и I_3 по формулам (17), (29) и (34) производилось путем численных квадратур, точность которых проверялась удвоением числа точек деления интервалов интегрирования. Параметр C в формуле (35) в соответствии с результатами работ [1] и [7] был принят равным 10^{-3} 1/км, а для параметра V использовалось два значения: 35 км/ч для зимних условий [7] и 24 км/ч для летних [1]. Рассмотрим некоторые результаты этих расчетов.

В табл. 1 и 2 приведены рассчитанные значения трех величин: $\varepsilon^2_{\text{п}}$, $\nu(400)$ и $\nu(1000)$. Первая из них представляет собой меру ошибок статистического прогноза, а вторая и третья — коэффициенты корреляции между этими ошибками на расстояниях 400 и 1000 км соответственно. Указанные значения приведены для интервала изменения параметров схемы статистического прогноза 500 км $\leqslant r \leqslant$ 1000 км и 6 ч $\leqslant \tau \leqslant$ 12 ч. Таблица 1 относится к случаю $V = 35$ км/ч, а табл. 2 — к случаю $V = 24$ км/ч.

Согласно данным по изменчивости геопотенциала и оценкам оправдываемости численных прогнозов, мера ошибок численного прогноза этого элемента на сутки меняется примерно в пределах от 0,10 до 0,20. Далее, имеющиеся сведения о корреляционных функциях ошибок численного прогноза геопотенциала на сутки позволяют заключить, что значение $\nu(400)$ должно находиться в пределах приблизительно от 0,60 до 0,70, а значение $\nu(1000)$ — в пределах от 0,14 до 0,28.

Совместный учет всех этих ограничений заставляет выбрать сравнительно узкие области значений параметров статистического прогноза r и τ в качестве тех, при которых статистический прогноз достаточно успешно с рассматриваемой точки зрения моделирует собой динамический прогноз. В табл. 1 и 2 эти значения даны курсивом.

На рис. 1 представлены наибольшие и наименьшие корреляционные функции ошибок статистического прогноза среди рассчитанных при указанных значениях параметров. На нем приведены также корреляционные функции ошибок численных прогнозов геопотенциала, заимствованные из работ Х. Кругера [14], Л. Бенгтссона и Н. Густавссона [9] и Д. Питерсена [16]. Мы видим, что рассчитанные кривые действительно лежат в пределах разброса эмпирической информации о корреляционных функциях ошибок численного прогноза.

Важно отметить, что значительный разброс этой информации сравнительно мало сказывается на параметрах согласования наблюдений с прогнозом по формуле (22): в то время как функция $\nu(\rho)$ варьирует значительно, константа $S_{\text{п}0}$ и функция $\varkappa(\rho)$, от которых зависят весовые множители a_i и $a_{\text{п}0}$, изменяются сравнительно в узких пределах. Особенно малы изменения функции

$$\chi(\rho;\ r,\ \tau) = \frac{\varkappa(\rho;\ r,\ \tau)}{S_{\text{п}0}(r,\ \tau)}, \tag{36}$$

Таблица 1

Мера ошибки $\varepsilon_п^2$ и коэффициенты корреляции $\nu(400)$ и $\nu(1000)$ статистического прогноза при различных значениях r и τ ($V=35$ км/ч)

r	τ						
	6	7	8	9	10	11	12
$\varepsilon_п^2 \cdot 10^3$							
0,5	58	68	79	91	104	118	133
0,6	74	83	93	104	116	130	144
0,7	93	*101*	*111*	*121*	132	145	158
0,8	116	*124*	*132*	142	152	164	176
0,9	142	150	158	165	176	187	198
1,0	172	178	186	194	203	213	224
$\nu(400) \cdot 10^3$							
0,5	630	674	710	739	764	784	800
0,6	630	666	697	725	749	770	787
0,7	648	*675*	*700*	*723*	744	763	780
0,8	674	*693*	*712*	730	747	764	778
0,9	701	715	729	743	757	770	782
1,0	727	737	748	758	769	779	789
$\nu(1000) \cdot 10^3$							
0,5	257	305	346	380	409	434	455
0,6	188	238	283	323	359	390	417
0,7	137	*184*	*228*	*270*	308	343	374
0,8	109	*130*	*191*	231	268	303	336
0,9	104	139	174	210	244	277	309
1,0	117	146	175	206	236	266	294

Таблица 2

Мера ошибки $\varepsilon_п^2$ и коэффициенты корреляции $\nu(400)$ и $\nu(1000)$ статистического прогноза при различных значениях r и τ ($V=24$ км/ч)

r	τ						
	6	7	8	9	10	11	12
$\varepsilon_п^2 \cdot 10^3$							
0,5	44	49	54	60	67	74	82
0,6	60	65	70	75	81	88	95
0,7	80	84	89	94	*100*	*106*	*113*
0,8	104	108	112	117	*122*	*128*	134
0,9	132	135	139	144	148	154	160
1,0	162	166	169	173	178	182	188

r	τ						
	6	7	8	9	10	11	12
				$\nu(400)\cdot10^3$			
0,5	528	568	605	638	668	694	717
0,6	557	584	611	637	661	683	704
0,7	598	616	634	653	*671*	*689*	*705*
0,8	640	652	664	677	*690*	*703*	716
0,9	678	686	694	704	713	722	732
1,0	711	717	723	729	736	743	750
				$\nu(1000)\cdot10^3$			
0,5	147	190	230	266	299	328	354
0,6	86	123	160	197	231	263	293
0,7	48	80	112	145	*177*	*208*	*238*
0,8	36	62	88	116	*144*	*173*	200
0,9	45	65	87	110	134	158	182
1,0	70	86	103	122	141	162	182

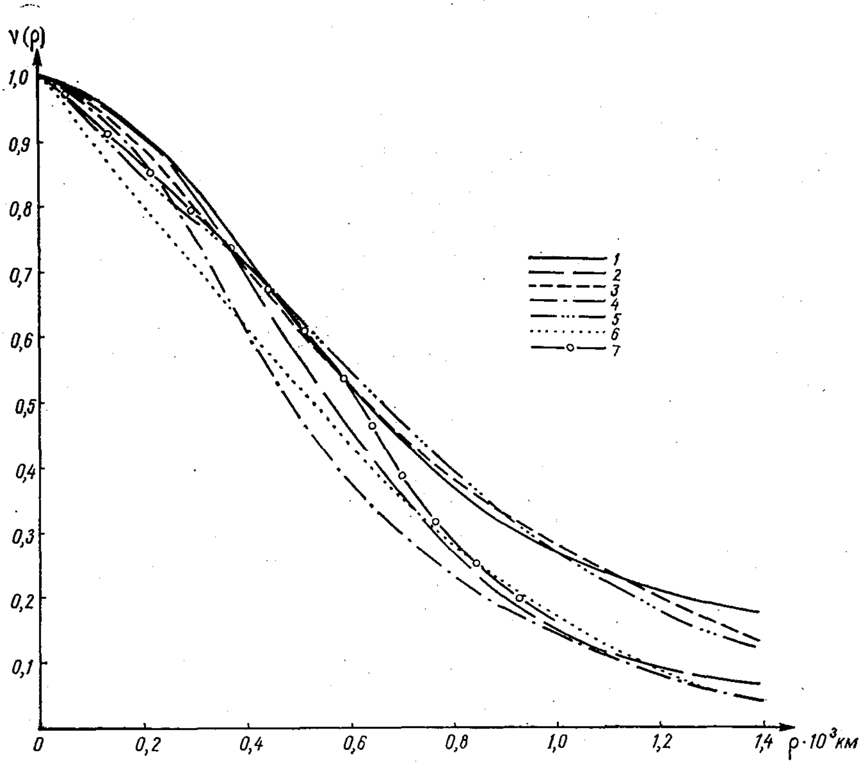

Рис. 1. Корреляционные функции γ(ρ) ошибок статистического и численного прогноза.

1 — r=700 км, τ=9 ч, V=35 км/ч; *2* — r=800 км, τ=10 ч, V=24 км/ч; *3* — поверхность 500 мбар [9]; *4* — поверхность 500 мбар [14]; *6* — поверхность 1000 мбар [17]; *6* — поверхность 500 мбар [17]; *7* — поверхность 300 мбар [16].

представляющую собой, как легко убедиться на основании формул (9) и (7), взаимную корреляционную функцию фактических и прогностических значений согласуемого элемента.

Так, все зависимости $\chi(\rho)$ при рассмотренных выше десяти наборах параметров r, τ и V, укладываются между двумя кривыми $\chi_{\min}(\rho)$ и $\chi_{\max}(\rho)$, приведенными на рис. 2.

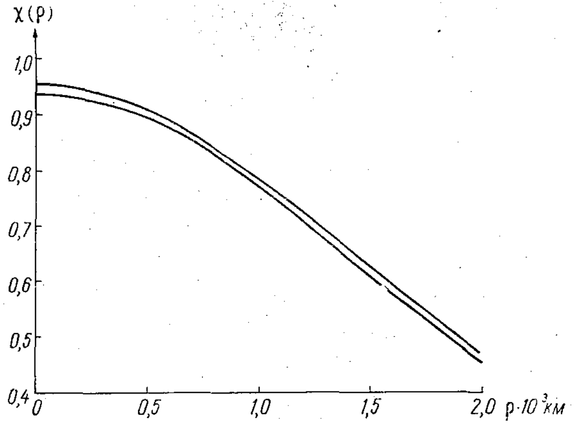

Рис. 2. Взаимная корреляционная функция $\chi(\rho)$ фактических и прогностических значений согласуемого элемента.

Имея это в виду, целесообразно несколько преобразовать систему уравнений (27) для определения весовых множителей, а именно воспользовавшись формулами (36) и (20), привести эту систему к виду

$$\sum_{j=1}^{n} a_j \tilde{\mu}_{ij} + b\chi_{i0} = \mu_{i0} \quad (i=1,\ 2,\ \ldots,\ n), \qquad (37)$$

$$\sum_{i=1}^{n} a_i \chi_{i\text{п}} + b = S_{\text{п}0},$$

где обозначено

$$b = a_{\text{п}} / S_{\text{п}0}. \qquad (38)$$

СПИСОК ЛИТЕРАТУРЫ

1. Антоненко В. С., Лугина К. М. Пространственно-временна́я структура поля давления для летнего сезона.— Труды ГГО, 1977, вып. 397, с. 70—77.
2. Гандин Л. С. Четырехмерный анализ метеорологических полей.— Л.: Гидрометеоиздат, 1976.— 61 с.
3. Гандин Л. С., Каган Р. Л. О построении системы объективного анализа разнородных данных на основе метода оптимальной интерполяции

и оптимального согласования.— Метеорология и гидрология, 1974, № 5, с. 3—10.
4. Гандин Л. С., Каган Р. Л. Статистические методы интерпретации метеорологических данных.— Л.: Гидрометеоиздат, 1976.— 359 с.
5. Каган Р. Л. Некоторые возможности использования данных о пространственно-временной структуре барического поля.— Труды ГГО, 1976, вып. 374, с. 3—21.
6. Либерман Ю. М. Эксперимент по использованию прогностической информации в объективном анализе.— Труды ГГО, 1977, вып. 397, с. 99—106.
7. Лугина К. М., Каган Р. Л. К вопросу о пространственно-временно́м анализе барического поля.— Труды ГГО, 1974, вып. 336, с. 75—94.
8. Bengtsson L. Impact on enhancement of the Arctic observing system.— Rep. 13th session of the Joint Organizing Committee. Annex E, App. C, GARP, WMO, Geneve, 1977.
9. Bengtsson L., Gustavsson N. An experiment in the assimilation of data in dynamical analysis.— Tellus, 1971, vol. 23, N 4—5, p. 328—336.
10. Bengtsson L., Gustavsson N. Assimilation of non-synoptic observations.— Tellus, 1972, vol. 24, N 5, p. 383—399.
11. Bergthorsson P., Döös B. R. Numerical weather map analysis.— Tellus, 1955, vol. 7, N 5. Перевод в сб. «Численные методы прогноза погоды». Под ред. Л. С. Гандина и А. С. Дубова.— Гидрометеоиздат, 1960.
12. Cressman G. P. An operational objective analysis system.— Mon. Wea. Rev., 1959, vol. 87, N 10, p. 367—374.
13. Johnson D. H. Preliminary research in objective analysis.— Tellus, 1957, vol. 9, N 3.
14. Kruger H. B. General and special approaches to the problem of objective analysis of meteorological variables.— Quart. J. Roy. Met. Soc., 1969, vol. 95, N 403, p. 21—39.
15. Miyakoda K., Talagrand O. The assimilation of past data in dynamical analysis. I.— Tellus, 1971, vol. 23, N 4—5, p. 310—317.
16. Petersen D. P. A comparison of the performance of quasi-optimal and conventional objective analysis schemes.— J. Appl. Met. 1973, vol. 12, N 7, p. 1093—1101.
17. Rutherford I. D. Data assimilation by statistical interpolation of forecast error fields.— J. Atmos. Sci., 1971, vol. 29, N 5, p. 809—815.
18. Schlatter T. W. Some experiments with a multivariate statistical objective analysis scheme.— Mon. Wea. Rev., 1975, vol. 103, p. 246—257.
19. Talagrand O., Miyakoda K. The assimilation of past data in dynamical analysis. II.— Tellus, 1971, vol. 23, N 4—5, p. 318—327.

Е. И. Хлебникова

О РАСЧЕТЕ ПРОДОЛЖИТЕЛЬНОСТЕЙ ВЫБРОСОВ РЯДОВ СРЕДНЕЙ СУТОЧНОЙ ТЕМПЕРАТУРЫ ВОЗДУХА

При изучении выбросов случайных процессов [5] в первую очередь представляет интерес такая случайная величина, как продолжительность выброса. Если средняя продолжительность выбросов при некоторых предположениях находится сравнительно легко, то строгий аналитический расчет распределения продолжительностей выбросов даже для гауссовских процессов представляет значительные математические трудности. Для процессов с дискретным временем расчет распределения продолжительностей выбросов сводится к вычислению интегралов большой кратности и применительно к сколько-нибудь длинным выбросам оказывается практически невозможным. Поэтому и для процессов с дискретным временем существенна разработка различных приближенных способов расчета распределения продолжительностей выбросов.

Один из возможных способов приближенного расчета распределения продолжительностей выбросов случайной последовательности $X(t)$ через уровень C основан на предположении о том, что последовательность $Y(t)$, принимающая значения x_+, если $X(t) > C$ и x_-, если $X(t) < C$ (последовательность импульсов) является марковской цепью какого-нибудь порядка [3, 4, 6]. Представляет интерес выяснение того, в какой мере это предположение может быть использовано для случайных процессов конкретной вероятностной структуры, в частности, применительно к гауссовским процессам с заданной корреляционной функцией.

Пусть $X(t)$ — стационарная случайная последовательность. Введем обозначения:

$$p_+ = P\{X(t_i) > C\}$$

— вероятность пребывания $X(t)$ выше рассматриваемого уровня C;

$$p_{++} = P\{X(t_i) > C / X(t_{i-1}) > C\}$$

— вероятность пребывания $X(t)$ в данный момент времени выше уровня C при условии, что в предыдущий момент времени процесс также был выше уровня C.

Аналогичный смысл имеют обозначения:

$$p_{-++}=P\{X(t_i)>C/X(t_{i-1})>C,\ X(t_{i-2})<C\},$$
$$p_{+++}=P\{X(t_i)>C/X(t_{i-1})>C,\ X(t_{i-2})>C\},$$
$$p_{-+++}=P\{X(t_i)>C/X(t_{i-1})>C,\ X(t_{i-2})>C,\ X(t_{i-3})<C\},$$
$$p_{++++}=P\{X(t_i)>C/X(t_{i-1})>C,\ \mathrm{X}(t_{i-2})>C,\ X(t_{i-3})>C\}.$$

В случае когда последовательность $Y(t)$ может рассматриваться как односвязная марковская цепь, вероятность того, что длина выброса равна k, находится по формуле [3]

$$p_1(k)=(1-p_{++})\,p_{++}^{k-1}. \qquad (1)$$

В предположении двухсвязности $Y(t)$ выражение для вероятности выброса заданной длины k имеет вид [3]:

$$p_2(1)=1-p_{-++},$$
$$p_2(k)=p_{-++}(1-p_{+++})\,p_{+++}^{k-2}, \quad k>1. \qquad (2)$$

Нетрудно показать, что если исходить из трехсвязности последовательности импульсов $Y(t)$, то для определения распределения длительностей выбросов можно воспользоваться формулами:

$$p_3(1)=1-p_{-++},$$
$$p_3(2)=p_{-++}(1-p_{-+++}),$$
$$p_3(k)=p_{-++}p_{-+++}(1-p_{++++})\,p_{++++}^{k-3}, \quad k>2. \qquad (3)$$

По формулам (1)—(3) были произведены расчеты распределения продолжительностей выбросов через различные уровни для гауссовской последовательности $X(t)$ в предположении, что корреляционная функция ее $r(\tau)$ имеет вид

$$r(\tau)=\exp(-\alpha\tau). \qquad (4)$$

Коэффициент корреляции $q=\exp(-\alpha)$ между смежными членами последовательности при расчетах варьировался.

Для проверки возможности применения формул (1)—(3) к процессам заданной структуры распределения продолжительностей выбросов были рассчитаны также путем непосредственного моделирования временных рядов на ЭВМ [2].

Рассмотрение гауссовского процесса с корреляционной функцией вида (4) определялось как тем, что именно таким образом удовлетворительно описывается вероятностная структура средней суточной температуры воздуха, представляющая интерес для многих приложений, так и тем, что в силу марковости процесса $X(t)$ сравнение распределений продолжительностей выбросов, рассчитанных в предположении марковости $Y(t)$, с истинным позволяет наглядно продемонстрировать потерю свойства марковости последовательности при укрупнении состояний.

Из априорных соображений ясно, что чем выше относительный уровень $c\,(c=(C-m)/\sigma$, где m и σ^2 — соответственно среднее и дисперсия $X(t))$ и чем меньше коэффициент корреляции q, тем ближе к истинным должны быть распределения продолжительностей выбросов, рассчитанные по формулам (1)—(3). Однако, как показали проведенные оценки, даже для $q=0{,}9$ и $c=0$ распределение продолжительностей выбросов вполне удовлетворительно может быть получено с помощью формулы (3).

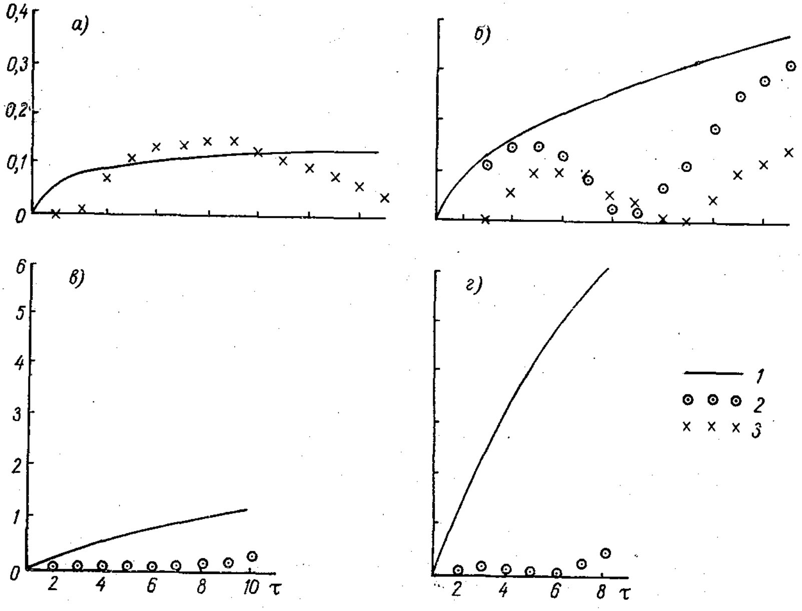

Рис. 1. Точность расчета распределения продолжительностей выбросов по марковской модели ($q=0{,}9$).

а) $c=0$, б) $c=1$, в) $c=2$, г) $c=3$; 1) $\frac{\sigma_F(\tau)}{F(\tau)}$, 2) $\frac{|F_2(\tau)-F(\tau)|}{F(\tau)}$, 3) $\frac{|F_3(\tau)-F(\tau)|}{F(\tau)}$.

Представление о точности, который можно достигнуть путем использования описанного подхода, дает рис. 1. Здесь $F_i(\tau) = \sum_{j\geqslant k} p_i(j)$ $(i=1, 2, 3)$ — вероятности выбросов длительности большей или равной τ, рассчитанные по формулам (1)—(3), $F(\tau)$ — вероятности выбросов длительности не меньше τ, полученные с помощью метода статистического моделирования по 360 реализациям, длина каждой из которых 1000 членов. Через $\sigma_F(\tau)$ обозначены средние квадратические отклонения вероятностей выбросов длительности не меньше τ, рассчитываемых по 60 реализациям, длина которых равна 30 членам. По величинам $\sigma_F(\tau)$, полученным также с помощью метода статистического моделирования, можно судить, в частности, с какой погрешностью за

Таблица 1

Распределения длительностей выбросов последовательности с $q=0{,}75$ (в тысячных долях единицы)

τ	$c=0$				$c=1$				$c=2$				$c=3$			
	$F_1(\tau)$	$F_2(\tau)$	$F_3(\tau)$	$F(\tau)$	$F_1(\tau)$	$F_2(\tau)$	$F_3(\tau)$	$F(\tau)$	$F_1(\tau)$	$F_2(\tau)$	$F_3(\tau)$	$F(\tau)$	$F_1(\tau)$	$F_2(\tau)$	$F_3(\tau)$	$F(\tau)$
1	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
2	770	673	673	674	570	507	507	505	374	342	342	341	216	204	204	192
3	593	537	508	506	325	313	299	290	140	145	141	141	046	053	052	049
4	457	429	411	397	185	193	190	180	053	061	062	061	009	014	014	005
5	352	343	333	320	105	119	121	114	020	025	027	029	001	004	004	000
6	271	274	270	258	060	073	077	073	008	009	012	012				
7	209	219	219	211	034	045	050	048	003	003	005	006				
8	161	175	177	173	019	028	032	030								
9	124	140	143	141	011	017	021	020								
10	096	112	116	115	006	010	014	013								
11	074	090	094	094	003	006	009	009								
12	057	072	076	078	002	003	006	006								
13	044	058	062	064	001	002	004	004								
14	034	047	050	053												
15	026	038	040	044												
16	020	031	032	036												
17	015	025	026	030												
18	011	020	021	024												
19	008	016	017	020												
20	006	013	014	017												
	σ_1	σ_2	σ_3	$\bar{\tau}$	σ_1	σ_2	σ_3	$\bar{\tau}$	σ_1	σ_2	σ_3	$\bar{\tau}$	σ_1	σ_2	σ_3	$\bar{\tau}$
	3,81	4,33	4,50	4,35	1,75	1,96	2,02	2,32	0,97	1,06	1,08	1,60	0,59	0,62	0,62	1,28

Таблица 2

Значения параметров p_{-++}, p_{-+++}, p_{++++} для стационарного гауссовского процесса с корреляционной функцией вида $r(\tau)=q^{\tau}$

c	q						
	0,65	0,70	0,75	0,80	0,85	0,90	0,95
Параметр p_{-++}							
0,0	0,657	0,666	0,673	0,681	0,688	0,695	0,701
0,5	0,557	0,575	0,592	0,610	0,628	0,647	0,668
1,0	0,454	0,480	0,507	0,535	0,565	0,597	0,634
1,5	0,355	0,388	0,423	0,460	0,501	0,545	0,599
2,0	0,266	0,302	0,342	0,387	0,436	0,493	0,563
2,5	0,189	0,226	0,269	0,318	0,375	0,442	0,527
3,0	0,129	0,163	0,204	0,254	0,316	0,391	0,492
Параметр p_{-+++}							
0,0	0,724	0,740	0,754	0,768	0,782	0,794	0,806
0,5	0,624	0,650	0,676	0,701	0,726	0,750	0,777
1,0	0,517	0,553	0,590	0,627	0,664	0,703	0,746
1,5	0,411	0,500	0,500	0,548	0,599	0,653	0,712
2,0	0,312	0,360	0,413	0,470	0,532	0,601	0,678
2,5	0,225	0,273	0,329	0,392	0,464	0,546	0,643
3,0	0,156	0,201	0,255	0,315	0,398	0,491	0,607
Параметр p_{++++}							
0,0	0,760	0,784	0,810	0,837	0,865	0,895	0,931
0,5	0,655	0,690	0,727	0,765	0,806	0,850	0,901
1,0	0,543	0,587	0,635	0,685	0,739	0,798	0,867
1,5	0,431	0,482	0,538	0,600	0,667	0,741	0,830
2,0	0,327	0,382	0,443	0,513	0,592	0,682	0,790
2,5	0,234	0,288	0,351	0,425	0,513	0,618	0,747
3,0	0,163	0,212	0,273	0,348	0,441	0,557	0,698

счет ограниченности выборки могут быть рассчитаны распределения средней суточной температуры воздуха по 60-летним рядам наблюдений в течение какого-либо месяца. При этом предполагается, что систематическая ошибка за счет ограниченности длины реализации устранена [1].

Из рис. 1 видно, что использование трехсвязной марковской модели при расчете распределения продолжительностей выбросов для положительных уровней в целом дает лучшие результаты, чем непосредственная обработка данных наблюдений в течение месяца за 60-летний период. Для высоких уровней ($c=2$, 3) вполне удовлетворительные результаты можно получить, используя двусвязную марковскую модель.

С уменьшением коэффициента корреляции q увеличивается степень близости значений $F_2(\tau)$ и $F_3(\tau)$ друг к другу и к «истинным» вероятностям выбросов заданной длины. В качестве примера в табл. 1 приведены полученные различными способами распределения длительностей выбросов последовательности с коэффициентом корреляции между смежными членами $q = 0{,}75$. Из этой таблицы хорошо видно, как меняются значения рассчитываемых вероятностей с изменением гипотезы о связности последовательности $Y(t)$.

Как можно видеть из формулы (3), в предположении трехсвязности последовательности импульсов распределение длительностей выбросов через заданный уровень определяется тремя параметрами: p_{-++}, p_{-+++} и p_{++++}. Значения этих параметров для гауссовского стационарного процесса с экспоненциальной корреляционной функцией для уровней $c = 0$, 1, 2, 3 и некоторых коэффициентов корреляции q приведены в табл. 2. Поскольку, как следует из вышеизложенного, применение трехсвязной марковской модели для расчета распределения продолжительностей выбросов случайной последовательности такой структуры дает вполне удовлетворительные результаты, используя табл. 2—4, для заданного набора c, q легко получить вероятности выбросов любой длины.

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л., Федорченко Е. И. О расчете статистических характеристик выбросов случайной функции.— Труды ГГО, 1970, вып. 268, с. 146—172.
2. Каган Р. Л., Канашкин В. К., Федорченко Е. И. О расчете характеристик временных рядов методом статистического моделирования.— Труды ГГО, 1972, вып. 286, с. 71—82.
3. Левин Б. Р., Фомин Я. А. Применение временно́й дискретизации случайного процесса для нахождения распределения длительности его выбросов.— Радиотехника, 1965, т. 20, № 10, с. 1—8.
4. Румянцев В. А., Бовыкин И. В. Основы теории случайных выбросов гидрологических рядов.— Обзоры ВНИИГМИ—МЦД. Гидрология суши. Обнинск, 1977.— 34 с.
5. Тихонов В. И. Выбросы случайных процессов.— М.: Наука, 1970.— 392 с.
6. Gabriel K. P., Neumann J. A Markov chain model for daily rainfall occurrence at Tel Aviv.— Quart. J. Roy. Met. Soc., 1962, vol. 88, N 375, p. 90—95.

Р. Л. Каган, Е. Е. Сибир

К ВОПРОСУ О ВЫБОРЕ ИНТЕРВАЛА ОСРЕДНЕНИЯ ПЕРИОДИЧЕСКИ КОРРЕЛИРОВАННЫХ ВЕЛИЧИН

1. Для большинства метеорологических элементов характерна нестационарность за счет наличия четко выраженного суточного и годового хода. Эта нестационарность проявляется как в изменчивости средних величин, так и в изменчивости других статистических параметров, например дисперсии и характеристик временной корреляции. С этой точки зрения метеорологические ряды могут рассматриваться как периодически коррелированные процессы, статистические характеристики которых в каждом периоде зависят от фазы, т. е. от интервала времени между рассматриваемым моментом и началом отсчета периода.

В метеорологической практике с целью исключения нестационарности широко пользуются осредненными величинами. Так, для исключения влияния суточного хода можно рассматривать осредненные за сутки значения метеорологических элементов. Для исключения влияния годового хода естественно рассматривать их средние годовые значения. Полученные после такого осреднения ряды во многих случаях можно считать стационарными (для средних годовых величин — при отсутствии годового хода, а для средних суточных величин — в периоды, когда не сказывается сезонный ход).

В связи с практической важностью процедуры осреднения, разным аспектам ее посвящено большое количество работ. В них детально исследуется вопрос о точности осреднения, выполняемого по ограниченному числу данных на периоде осреднения при использовании разных способов осреднения, и о том, как при этом изменяются характеристики осредняемой величины. Такого рода исследования очень существенны при анализе рядов наблюдений за длинные периоды времени, в течение которых могут меняться как сроки наблюдений, так и способы их осреднения.

Вместе с тем следует иметь в виду, что и в идеальном случае, когда измерения производятся столь часто, что осреднение по периоду может производиться точно, внутренняя нестационарность процесса в течение периода оказывает влияние на его статистические характеристики. Это влияние проявляется в том, что дисперсии, корреляционные функции и другие характеристики могут существенно зависеть от выбора начала периода.

Очевидно, например, что корреляция между средними для смежных суток будет выше, если конец первых суток и начало последующих характеризуются повышенной дисперсией. При выборе за начало суток часов с пониженными значениями дисперсии корреляция между средними значениями смежных суток уменьшится. В равной мере можно ожидать, что при выборе начала суток в часы с повышенной дисперсией дисперсия осредненных величин окажется меньшей, чем при другом выборе начала метеорологических суток. Аналогично должно обстоять дело и при годовом осреднении.

Величина этого эффекта должна зависеть от статистических характеристик осредняемого элемента, в первую очередь от степени и характера его нестационарности и временно́й связности. Очевидно, для разных элементов они могут быть различными. В климатологической практике этот эффект не учитывается, начало периода осреднения определяется соображениями практического удобства и обычно устанавливается единым для всех метеорологических элементов (например, 19 часов или полночь при суточном осреднении, 1 января или 1 июля при годовом осреднении).

В связи с этим возникает вопрос, насколько это обстоятельство практически существенно. В частности, не может ли привести смена начала периода осреднения (даже в предположении, что осреднение выполняется точно) к нарушению однородности ряда осредненных величин. Из соображений интуитивного характера это представляется маловероятным, однако для уверенного ответа на этот вопрос необходимо иметь количественные оценки. Ниже будут приведены некоторые оценки, полученные для простейшего случая периодически коррелированных процессов, для которых вся нестационарность определяется изменением нормы и дисперсии на периоде осреднения.

2. Пусть мы имеем процесс $f(t)$. Процесс этот полагаем нестационарным с математическим ожиданием (нормой) $\overline{f}(t)$ и ковариационной функцией

$$R(t_1,\ t_2)=\overline{f(t_1)\,f(t_2)}-\overline{f(t_2)}\cdot\overline{f(t_2)}. \tag{1}$$

Здесь черта сверху означает осреднение, понимаемое в статистическом смысле.

Дисперсия для различных моментов времени t_1 и t_2 составляет

$$\sigma^2(t_1)=R(t_1,\ t_1),\quad \sigma^2(t_2)=R(t_2,\ t_2). \tag{2}$$

Для простоты будем считать, что нестационарность ковариации определяется изменением нормы и дисперсии, а нормированные значения $(f(t)-\overline{f}(t))/\sigma(t)$ являются стационарными, так что корреляционная функция r величины f зависит лишь от разности аргументов. В этом случае

$$R(t_1, t_2)=\sigma(t_1)\sigma(t_2)r(|t_2-t_1|). \tag{3}$$

Рассмотрим скользящую среднюю с периодом осреднения T:

$$F(t)=\frac{1}{T}\int\limits_{t}^{t+T} f(t)\,dt. \tag{4}$$

Среднее значение и ковариационная функция величины F равны соответственно

$$\overline{F}(t)=\frac{1}{T}\int\limits_{t}^{t+T} \overline{f}(t)\,dt, \tag{5}$$

$$R_F(t_1, t_2)=\frac{1}{T^2}\int\limits_{t_1}^{t_1+T}\int\limits_{t_2}^{t_2+T} R(x, y)\,dx\,dy. \tag{6}$$

Если подставить (3) в (6), то получим

$$R_F(t_1, t_2)=\frac{1}{T^2}\int\limits_{t_1}^{t_1+T}\int\limits_{t_2}^{t_2+T} \sigma(x)\sigma(y)r(|x-y|)\,dx\,dy. \tag{7}$$

В общем случае нестационарный процесс остается нестационарным и после сглаживания. Рассмотрим случай периодически коррелированного процесса с периодом, равным периоду осреднения T. Для такого процесса

$$\overline{f}(t+T)=\overline{f}(t),\quad \sigma(t+T)=\sigma(t) \tag{8}$$

и осреднение приводит к постоянству средней величины

$$\overline{F}(t)=\frac{1}{T}\int\limits_{0}^{T} \overline{f}(t)\,dt=\text{const}. \tag{9}$$

Однако для ковариационной функции нестационарность по-прежнему имеет место, и величина F также является периодически коррелированной с периодом T. В частности, для дисперсии скользящих средних по периоду и для ковариации средних по смежным периодам имеем соответственно

$$\sigma_F^2(t)=R_F(t, t)=\frac{1}{T^2}\int\limits_{t}^{t+T}\int\limits_{t}^{t+T} \sigma(x)\sigma(y)r(|x-y|)\,dx\,dy, \tag{10}$$

$$R_F(t, t+T)=\frac{1}{T^2}\int\limits_{t}^{t+T}\int\limits_{t}^{t+T} \sigma(x)\sigma(y+T)r(|x-y-T|)\,dx\,dy. \tag{11}$$

Для оценки степени нестационарности этих характеристик необходимо задание конкретного вида функций $r(\tau)$ и $\sigma^2(t)$. Интегрирование в формулах (10) и (11) производится, вообще говоря, численно, однако в некоторых случаях квадратуры в них берутся и в результате получаются сравнительно простые аналитические выражения.

Рассмотрим, например, случай, когда ход средних квадратических отклонений исходной величины f может быть описан простой гармоникой

$$\sigma(t)=\sigma_0\left[1+A\cos\left(2\pi\frac{t}{T}\right)\right], \tag{12}$$

а корреляционная функция определяется экспоненциальной зависимостью

$$r(t_1,\ t_2)=\exp\left(-\frac{|\tau|}{T_0}\right), \tag{13}$$

где $\tau=t_1-t_2$ — сдвиг времени, а T_0 — масштаб корреляции. В этом случае подстановка в (10) и (11) дает

$$\mu_0(t)=\mu_{0c}+\alpha B-2\alpha\left(\frac{T_0}{T}\right)^2\left(1-e^{-T/T_0}\right)\cos\left(2\pi\frac{t}{T}\right)\times$$
$$\times\left[2+A\cos\left(2\pi\frac{t}{T}\right)\right], \tag{14}$$

$$\mu_T(t)=\mu_{Tc}\left[(1+\alpha^2-\alpha A)+\alpha\cos\left(2\pi\frac{t}{T}\right)\left(2+A\cos\left(2\pi\frac{t}{T}\right)\right)\right]. \tag{15}$$

Здесь

$$\alpha=\frac{AT^2}{T^2+4\pi^2T_0^2},$$

$$B=A\frac{T_0}{T}+2\left(\frac{T_0}{T}\right)^2\left(1-e^{-T/T_0}\right)(A-\alpha),$$

$$\mu_\tau(t)=\frac{R_F(t,\ t+\tau)}{\sigma_0^2}$$

— относительная ковариация осредненных значений, μ_{0c} и μ_{Tc} — значения относительной ковариации для стационарного случая ($A=0$), которые определяются формулами

$$\mu_{0c}=2\left(\frac{T_0}{T}\right)^2\left(e^{-T/T_0}-1+\frac{T}{T_0}\right), \tag{14a}$$

$$\mu_{Tc}=\left(\frac{T_0}{T}\right)^2\left(1-e^{-T/T_0}\right)^2. \tag{15a}$$

Из (14) видно, что дисперсия средней является периодической величиной. Минимумы ее совпадают с максимумами дисперсии исходной величины f. Следовательно, наибольшее сглаживание ее достигается при выборе в качестве начала периода осреднения моментов, для которых дисперсия величины f максимальна. Максимум дисперсии осредненной величины соответствует выбору в качестве начала периода осреднения моментов минимальной дисперсии величины f.

Из (14) следует, что разность между наибольшим и наименьшим значениями величины μ_0 равна

$$\Delta\mu_0 = \frac{8T_0^2(1-e^{-T/T_0})}{T^2+4\pi^2T_0^2}A. \tag{16}$$

Она пропорциональна относительной амплитуде изменения среднего квадратического отклонения A величины f. Нетрудно видеть, что величина $\Delta\mu_0$ мала как при больших, так и при малых значениях масштаба корреляции T_0.

В самом деле при $T_0 \gg T$

$$\Delta\mu_0 \approx \frac{2}{\pi^2}\frac{T}{T_0}A, \tag{16а}$$

а при $T_0 \ll T$

$$\Delta\mu_0 \approx 8\left(\frac{T_0}{T}\right)^2 A. \tag{16б}$$

Наибольших значений величина $\Delta\mu_0$ достигает при масштабах корреляции T_0, сравнимых с величиной периода осреднения. Максимальна она при $T_0 \sim 0{,}4T$, когда $\Delta\mu_0 \approx 0{,}18A$ (заметим, что при этом $\mu_{0c} \approx 0{,}5$).

Из (15) следует, что ковариация для смежных периодов μ_{Tc} также является периодической величиной, находящейся в противофазе с μ_{0c}, так что ее максимумы и минимумы совпадают с максимумами и минимумами дисперсии осредняемой величины f. Разность между наибольшим и наименьшим ее значениями составляет

$$\Delta\mu_T = 4\alpha\mu_{Tc} = \frac{4T_0^2(1-e^{-T/T_0})}{T^2+4\pi^2T_0^2}A. \tag{17}$$

Эта разность достигает максимальных значений, близких к $0{,}075A$, также при $T_0 \sim 0{,}4T$ (при этом $\mu_{Tc} \approx 0{,}13$).

3. Оценки зависимости ковариации осредненных значений от сдвига момента начала периода осреднения были произведены также для корреляционных функций вида

$$r(\tau) = \left(1+\frac{|\tau|}{T_0}\right)\exp\left(-\frac{|\tau|}{T_0}\right) \tag{18}$$

и

$$r(\tau) = \exp\left[-\frac{1}{2}\left(\frac{\tau}{T_0}\right)^2\right]. \tag{19}$$

Эти зависимости оказываются аналогичными уже описанным для корреляционной функции (13). Это видно, например, из рис. 1, на котором приводятся некоторые оценки, полученные для корреляционной функции (18). По осям ординат этого рисунка отло-

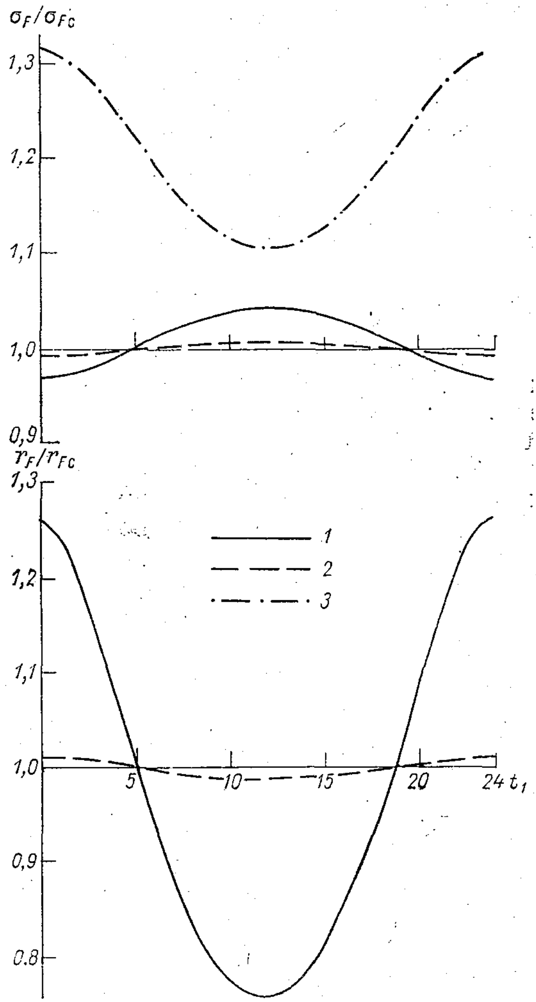

Рис. 1. Зависимость от начала периода осреднения среднего квадратического отклонения σ_F/σ_{F_c} и корреляции для смежных периодов r_F/r_{F_c} значений осредненных величин при корреляционной функции $r(\tau) = \left(1 + \frac{|\tau|}{T_0}\right) \times$ $\times \exp\left(-\frac{|\tau|}{T_0}\right)$ для $T_0=0{,}2T$ (*1*) и $T_0=T$ (2), а также при $r(\tau) = \cos\left(\frac{7\tau}{T}\right) \exp\left(-\frac{|\tau|}{T}\right)$(*3*).

Момент $t=0$ соответствует максимальной дисперсии исходной величины.

жены величины $\sigma_F(t)/\sigma_{Fc} = \sqrt{\frac{\mu_0(t)}{\mu_{0c}}}$ (рис. 1 *а*) и $r_F(t)/r_{Fc} = \sqrt{\frac{\mu_T(t)}{\mu_{Tc}} \Big/ \frac{\mu_{0T}}{\mu_{0c}}}$ (рис. 1 *б*), характеризующие превышение средним квадратическим отклонением осредненных величин σ_F и коэффициентом корреляции средних по смежным периодам r_F их значений для стационарного случая (σ_{Fc} и r_{Fc} соответственно).

Из рис. 1 видно, что при совмещении начала периода осреднения с максимумом дисперсии осредняемой величины f дисперсия осредненных значений будет минимальной, а корреляция осредненных по смежным периодам значений — максимальной. Дисперсия осредненной величины будет максимальной при выборе в качестве начала периода осреднения момента минимума дисперсии исходной величины. Корреляция средних для смежных периодов будет в этом случае минимальной.

Имея это в виду, мы ограничимся приведением в табл. 1 разностей между максимальным и минимальным значениями среднего квадратического отклонения ($\Delta\sigma_F$) и коэффициента корреляции (Δr_F) средних значений. Эти разности удобно приводить в долях относительной амплитуды A среднего квадратического отклонения исходной величины f, поскольку они приближенно пропорциональны ей.

Таблица 1

Разности между максимальными и минимальными значениями σ_F и r_F на периоде (в % от амплитуды A) в зависимости от корреляционной функции и масштаба корреляции T_0

T_0/T	$r(\tau) = \exp\left(-\frac{\lvert\tau\rvert}{T_0}\right)$		$r(\tau) = \left(1 + \frac{\lvert\tau\rvert}{T_0}\right) \times \exp\left(-\frac{\lvert\tau\rvert}{T_0}\right)$		$r(\tau) = \exp\left[-\frac{1}{2}\left(\frac{\tau}{T_0}\right)^2\right]$	
	$\Delta\sigma_F/A$	$\Delta r_F/A$	$\Delta\sigma_F/A$	$\Delta r_F/A$	$\Delta\sigma_F/A$	$\Delta r_F/A$
0,05	3	9	6	14	3	8
0,10	7	16	12	24	7	16
0,20	11	23	14	26	17	21
0,40	11	22	10	17	15	26
0,60	10	19	6	10	10	17
0,80	8	17	4	7	6	10
1,0	7	14	3	4	4	5
2,0	4	8	1	1	1	1
3,0	3	6	—	—	1	0
5,0	2	4	0	0	0	0
10,0	1	2	0	0	0	0

Для сравнения в табл. 2 приведены значения σ_{Fc}/σ_0 и r_{Fc} для стационарного случая.

Из табл. 1 видно, что при аналогичном ходе зависимости раз-

Таблица 2

Среднее квадратическое отклонение средних по периоду σ_{Fc}, корреляция средней по смежным периодам r_{Fc} в стационарном случае и корреляция исходных величин на интервале, равном периоду осреднения $r(T)$ (в %)

T_0/T	$r(\tau)=\exp\left(-\frac{\lvert\tau\rvert}{T_0}\right)$			$r(\tau)=\left(1+\frac{\lvert\tau\rvert}{T_0}\right)\times$ $\times\exp\left(-\frac{\lvert\tau\rvert}{T_0}\right)$			$r(\tau)=$ $=\exp\left[-\frac{1}{2}\left(\frac{\tau}{T_0}\right)^2\right]$		
	σ_{Fc}/σ_0	r_{Fc}	$r(T)$	σ_{Fc}/σ_0	r_{Fc}	$r(T)$	σ_{Fc}/σ_0	r_{Fc}	$r(T)$
0,05	32	3	0	43	4	0	35	2	0
0,10	43	5	0	58	9	0	48	4	0
0,20	57	12	1	75	20	4	65	10	0
0,40	71	27	8	89	44	29	83	23	4
0,60	78	38	19	94	60	50	90	40	25
0,80	83	47	29	96	71	65	94	54	46
1,0	86	54	37	97	78	74	96	65	61
2,0	92	73	61	99	92	91	99	89	88
3,0	95	81	72	—	—	96	99	95	95
5,0	97	88	82	100	98	98	100	98	98
10,0	98	94	90	100	99	99	100	99	99

ностей Δ от масштаба корреляции имеют место значительные различия для разных видов корреляционной функции. Это вполне естественно, поскольку одному и тому же значению T_0 соответствует различная степень связности в зависимости от аппроксимации корреляционной функции. Более сопоставимые результаты получаются при рассмотрении зависимости разностей Δ от значений корреляционных функций для интервалов времени, равных длине интервала осреднения $r(T)$.

Как видно из рис. 2, на котором представлены соответствующие зависимости для всех трех рассмотренных корреляционных функций, они для сравнительно связных рядов ($r(\tau) > 0{,}1$) мало отличаются друг от друга и могут быть приближенно описаны простыми формулами:

$$\frac{\Delta\sigma_F}{A\sigma_0} \approx 0{,}14-0{,}19r(T)+0{,}05r^2(T), \tag{20}$$

$$\frac{\Delta r_F}{A} \approx 0{,}26-0{,}38r(T)+0{,}12r^2(T). \tag{21}$$

Для процессов малой связности (при $r(T) < 0{,}1$) расхождения оказываются более существенными и такой единой аппроксимации подобрать не удается. Количественные оценки в этом случае следует вести раздельно для каждой конкретной корреляционной функции.

Заметим, что указанные закономерности характерны для корреляционных функций, которые монотонно убывают с увеличением сдвига времени.

В случае отсутствия монотонности, а тем более для знакопеременных корреляционных функций может отмечаться их нарушение, в частности совпадение максимумов и минимумов дисперсий осредненных и исходных значений величины f. Так, например, обстоит дело с приведенной на рис. 1 зависимостью $\Delta\sigma_F/\sigma_0$ для корреляционной функции $r(\tau)=\cos\left(\frac{7\tau}{T}\right)\exp\left(-\frac{|\tau|}{T}\right)$. Принципиальную возможность таких случаев следует иметь в виду, хотя практически они, по-видимому, встречаются редко.

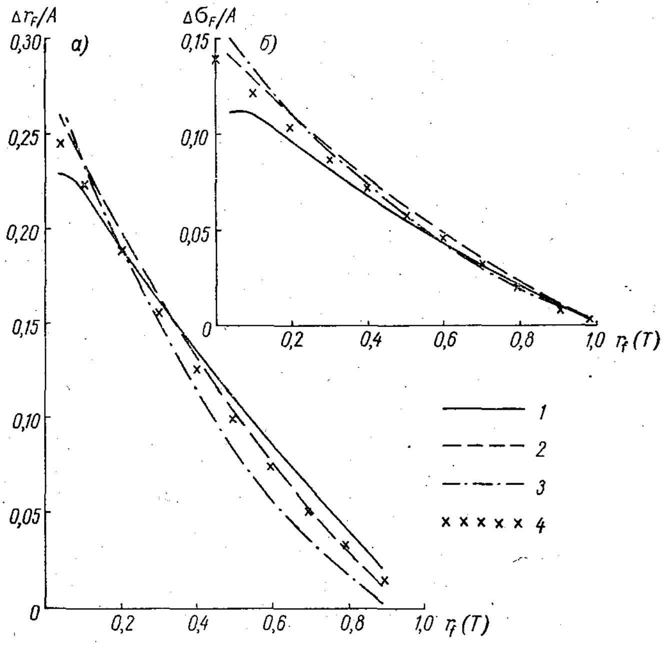

Рис. 2. Зависимость разностей между максимальными и минимальными значениями r_F (*а*) и σ_F (*б*) на периоде (в % от амплитуды A) от значений корреляционных функций для интервалов времени, равных периоду осреднения.

1) $r(\tau)=\exp\left(-\frac{|\tau|}{T_0}\right)$; *2*) $r(\tau)=\left(1+\frac{|\tau|}{T_0}\right)\exp\left(-\frac{|\tau|}{T_0}\right)$, *3*) $r(\tau)=\exp\left[-\frac{1}{2}\left(\frac{\tau}{T_0}\right)^2\right]$, *4*) формулы (20) и (21).

4. Применительно к конкретным метеорологическим процессам использованная выше статистическая модель периодически коррелированного процесса является, конечно, существенно упрощенной. Известно, например (см. [1, 5]), что для таких характеристик, как температура и влажность воздуха, предположение о стационарности корреляционной функции в течение суток является довольно грубым приближением, равно как и описание хода среднего квадратического отклонения простой гармоникой. Тем не менее для

ориентировочных оценок использование этой модели представляется возможным.

Суточный ход среднего квадратического отклонения температуры воздуха для летних и зимних месяцев существенно различен. Для зимних месяцев, для которых он вообще выражен сравнительно слабо, характерно наличие максимума в ночные часы, для летних месяцев — в дневные часы [4, 5]. Относительные амплитуды суточного хода оказываются различными для разных районов, однако для территории СССР характерные значения их составляют 0,01—0,15 для зимних и 0,15—0,20 для летних месяцев [3]. Лишь в отдельных районах они достигают значений, равных примерно 0,3. Для относительной влажности в зимние месяцы суточный ход среднего квадратического отклонения практически отсутствует, а летом и особенно в переходные сезоны наблюдается четко выраженный максимум его в дневные часы. В качестве характерных значений относительной амплитуды в эти сезоны могут быть, по-видимому, приняты значения $A = 0{,}2 \div 0{,}3$.

Для зимних месяцев временны́е корреляционные функции температуры и относительной влажности воздуха могут считаться стационарными. Убывание корреляции трудно описать одним аналитическим выражением, однако, если пренебречь сравнительно малыми сдвигами времени τ, можно воспользоваться простой экспоненциальной зависимостью (13) при масштабах корреляции T_0 примерно 2—3 суток.

Для летних месяцев масштаб корреляции меньше и составляет около двух суток для температуры и около одних суток для относительной влажности. Для этих месяцев предположение о стационарности временно́й корреляции является гораздо более грубым.

Исходя из сказанного, можно ожидать, что при определении в качестве начала метеорологических суток какого-нибудь из дневных часов (например, 13 или 15) средние суточные значения температуры воздуха и относительной влажности летом оказались бы более устойчивыми и более тесно коррелирующими друг с другом, чем для календарных суток. Зимой, наоборот, более устойчивые характеристики дает осреднение по календарным суткам.

Количественно эти различия сравнительно невелики. Из табл. 1 следует, что различие между максимальным и минимальным значениями σ_F должны составлять около 3—4 % от амплитуды A для температуры воздуха в течение года и для относительной влажности зимой, а для относительной влажности в переходные сезоны и летом — около $0{,}07A$. Различия между максимальными и минимальными значениями межсуточной корреляции примерно вдвое больше.

Учитывая приведенные выше характерные значения A для рассматриваемых элементов, приходим к выводу, что различия лежат в пределах 1 % от средних значений σ_0 среднего квадратического отклонения для температуры воздуха и в пределах 2 % — для относительной влажности.

В абсолютных цифрах эти различия составляют около 0,05 °С для температуры и около 0,3—0,4 % для относительной влажности.

Для межсуточной корреляции эти различия могут достигать около 0,02 для температуры воздуха (при средних значениях порядка 0,7) и около 0,05 для относительной влажности (при средних значениях порядка 0,5).

В качестве примера в табл. 3 приводятся оценки, полученные для июля по данным ГМС Барабинск (относительная влажность) и ГМО Ленинград (температура воздуха). При этом для ГМС Барабинск расчеты выполнялись по приведенным выше формулам в предположении справедливости формул (12) и (13) (исходные данные были взяты из работы [2]). Для ГМО Ленинград вместо задания аналитических формул использовались полученные Е. И. Федорченко [5] фактические данные о суточном ходе σ и $r(\tau)$ для температуры воздуха, а интегрирование в (10) и (11) производилось численно.

Очевидно, что полученные различия в значениях σ_F практического значения не имеют. Более существенными могут оказаться различия в межсуточной корреляции, поскольку они сопоставимы с реально достижимой точностью оценки этой корреляции по фактическим данным (погрешность определения σ_F по фактическим данным оказывается значительно больше различий за счет нестационарности), однако и ими в большинстве случаев можно пренебречь.

Исключение могут составить некоторые задачи, в которых требуется одновременное задание средних квадратических отклонений и корреляции осредненных величин. К числу таких задач относится определение характеристик выбросов через высокие уровни. Известно, что среднее число выбросов некоторого параметра через заданный уровень при прочих равных условиях увеличивается при увеличении среднего квадратического отклонения этого параметра и при уменьшении корреляции между значениями параметра за смежные сроки. Одновременное изменение этих характеристик в противоположном направлении может привести к заметному изменению числа выбросов по сравнению со стационарным случаем. Оценки показывают, что для температуры и относительной влажности для высоких уровней (в 2—3 раза превышающих среднее квадратическое отклонение) различия в среднем числе выбросов средних суточных значений в зависимости от выбора начала периода осреднения могут в неблагоприятных условиях достигать 10—20 % от числа выбросов в стационарном случае.

5. При годовом осреднении будем для простоты считать исходными средние месячные величины. Для большинства метеорологических элементов характерна большая нестационарность этих величин по дисперсии при сравнительно небольшой связности ряда.

Так, для температуры воздуха дисперсия имеет четко выраженный годовой ход с максимумом в зимние и минимумом в летние

Таблица 3

Средние квадратические отклонения средних величин и корреляция их для смежных периодов при суточном (за июль) и годовом осреднении

Осреднение	Станция	Величина	Параметр					
			σ_0	A	T_0	σ_c	r_c	σ_{max}
Суточное	Ленинград	Температура воздуха	3,1 °C	0,18	2,2 сут	2,9 °C	0,73	2,9 °C
	Барабннск	Относительная влажность	12,5 %	0,28	0,8 сут	10,4 %	0,47	7,5 %
Годовое	9 НИСП	Температура воздуха	2,1 °C	0,4	0,1	0,9 °C	0,05	0,8 °C
	НИСП *I*	Температура поверхности моря	1,2 °C	0,4	0,4	0,8 °C	0,27	0,8 °C

Осреднение	Станция	Величина	Параметр				
			σ_{min}	$\Delta\sigma/\sigma_c$	r_{max}	r_{min}	$\Delta r/r_c$
Суточное	Ленинград	Температура воздуха	2,8 °C	0,01	0,78	0,76	0,02
	Барабннск	Относительная влажность	7,3 %	0,02	0,50	0,45	0,10
Годовое	9 НИСП	Температура воздуха	0,8 °C	0,03	0,06	0,02	0,76
	НИСП *I*	Температура поверхности моря	0,8 °C	0,06	0,30	0,22	0,33

месяцы, а относительная амплитуда A среднего квадратического отклонения имеет порядок 0,5. Поскольку, однако, корреляция средней месячной температуры быстро затухает, влияние нестационарности дисперсии также оказывается сравнительно небольшим.

Ряды средних месячных значений величины f можно в первом приближении рассматривать как сумму периодически коррелированного случайного процесса f_0 с корреляционной функцией $r_{f_0}(\tau)$ и дисперсией $\sigma^2_{f_0}(t)$ и некоррелированного белого шума δ с дисперсией $\sigma^2_\delta(t)$:

$$f(t)=f_0(t)+\delta(t). \tag{22}$$

Ковариационная функция величин f представляет собой сумму ковариационных функций их составляющих. В частности, для дисперсии имеем

$$\sigma^2(t)=\sigma^2_{f_0}(t)+\sigma^2_\delta(t). \tag{23}$$

Полагая фиксированной долю дисперсии шумовой составляющей в полной дисперсии $\eta^2=\sigma^2_\delta/\sigma^2$, получаем корреляционную функцию процесса

$$r(\tau)=\frac{1}{1+\eta^2}r_{f_0}(\tau) \text{ при } \tau>0, \tag{24}$$

или

$$r(\tau)=r'(0)\cdot r_{f_0}(\tau) \text{ при } \tau>0, \tag{25}$$

где $r'(0)=\dfrac{1}{1+\eta^2}$ — значение корреляционной функции, полученное путем экстраполяции на $\tau=0$.

Составляющие дисперсии в (23) могут быть в этом случае представлены в виде

$$\sigma^2_{f_0}=\sigma^2\cdot r'(0), \quad \sigma^2_\delta=\sigma^2(1-r'(0)). \tag{26}$$

Рассмотрим теперь осредненные по 12 месяца́м значения f:

$$F=F_0+D, \tag{27}$$

где

$$F_0=\frac{1}{T}\int_0^T f_0\,dt; \quad D=\frac{1}{12}\sum_{i=1}^{12}\delta_i. \tag{28}$$

Составляющие F_0 и D величины F, подобно f_0 и δ, являются некоррелированными, так что

$$\sigma^2_F(t)=\sigma^2_{F_0}(t)+\sigma^2_D, \tag{29}$$

где

$$\sigma^2_D=\frac{1-r'(0)}{144}\sum_{i=1}^{12}\sigma^2(t_i), \tag{30}$$

а $\sigma^2_{F_0}$ оценивается по данным о $\sigma^2_{f_0}(t)$ и $r_{f_0}(\tau)$ в соответствии с формулой (10).

В случае когда годовой ход величины $\sigma^2_{f_0}$ описывается формулой (12), имеем

$$\sigma^2_D = \frac{1 - r'(0)}{12}\left(1 + \frac{A^2}{2}\right)\sigma^2_0. \tag{31}$$

Коэффициент корреляции между смежными средними годовыми значениями составляет

$$r_F(t,\ t+T) = \frac{\sigma^2_{F_0}(t)}{\sigma^2_{F_0}(t) + \sigma^2_D} r_{F_0}(t,\ t+T), \tag{32}$$

причем корреляционная функция r_{F_0} определяется в соответствии с формулой (11).

Рассмотрим в качестве примера ряды средней месячной температуры воздуха и поверхности моря по данным кораблей погоды в Северной Атлантике. Характеристики временно́й структуры этих величин приведены в статье Е. Крауса и Р. Моррисона [6]. Из приведенных данных следует, что для всех девяти кораблей погоды для температуры воздуха можно принять в среднем $\sigma_0 = 2{,}1$ °C, $A = 0{,}4$, $r'(0) = 0{,}5$. Таким образом, на долю некоррелированной составляющей изменчивости приходится половина дисперсии средней месячной температуры воздуха.

Будем полагать корреляционную функцию экспоненциально убывающей по формуле (13) с масштабом корреляции $T_0 = 0{,}1$ года. В этом случае оценки по приведенным формулам показывают, что при годовом осреднении вклад некоррелированной составляющей уменьшается лишь до $^1/_3$ от полной дисперсии средней годовой температуры воздуха.

Из табл. 3, в которой приведены некоторые оценки влияния выбора начала периода осреднения, видно, что для среднего квадратического отклонения возможные различия за этот счет невелики и лежат в пределах 5 % от его значений для стационарного случая. Различия в межгодовой корреляции, будучи сами по себе невелики (примерно 0,04), сопоставимы с величиной самого коэффициента корреляции для стационарного случая, так что они были бы весьма существенными, если бы можно было быть уверенным в пригодности принятой аппроксимации для сдвигов времени порядка года и более. К сожалению, исходные данные по структуре не дают оснований для такой уверенности.

Гораздо большей временно́й связностью характеризуются средние месячные значения температуры поверхности моря. Так, осреднение по всем девяти кораблям дает при экспоненциальной аппроксимации приближенно $r'(0) = 0{,}7$ и $T_0 = 0{,}15$ года, а для корабля I, расположенного между Исландией и Ирландией, даже $r'(0) = 0{,}9$ и $T_0 = 0{,}4$ года. В связи с этим годовое осреднение оказывается более эффективным с точки зрения сглаживания некоррелированной составляющей. Вклад ее в дисперсию в среднем для девяти кораблей уменьшается от 30 до 12 %, а для корабля I — от 10 до 2 %.

Поскольку годовой ход дисперсии температуры поверхности моря обратен соответствующему годовому ходу для температуры воздуха (она минимальна в зимние и максимальна в летние месяцы), влияние нестационарности на эти элементы также противоположно. Дисперсия средних годовых значений температуры поверхности моря максимальна для календарных годов и минимальна для годов, начало которых приурочивается к летним месяцам. Межгодовая корреляция при осреднении по календарным годам максимальна. Соответствующие различия следовало бы, по-видимому, признать существенными с той же оговоркой о надежности исходных статистических данных.

Из изложенного можно сделать следующие выводы.

1. Статистические характеристики осредненных по времени периодически коррелированных процессов даже при точном выполнении осреднения зависят от выбора начала периода осреднения.

2. Эта зависимость тем больше, чем больше амплитуда среднего квадратического отклонения осредняемой величины. Зависимость мала при большой и малой временно́й связности ряда и максимальна при масштабах корреляции, сравнимых с периодом осреднения.

3. При монотонном убывании корреляции со сдвигом времени в пределах периода выбор в качестве начала периода момента максимальной дисперсии обеспечивает минимум дисперсии осредняемой величины и максимум корреляции ее значений для смежных периодов. При достаточно высокой связности ряда ($r(T) > > 0{,}1$) возможная изменчивость этих характеристик мало зависит от вида корреляционной функции.

4. Выполненные для конкретных метеорологических элементов оценки показывают, что этот эффект для них практически мало существен как при суточном осреднении (в силу большой связности метеорологических элементов в пределах суток), так и при годовом осреднении (в силу малой связности их в течение года). Возможные различия статистических характеристик осредненных величин в зависимости от выбора начала года или суток малы по сравнению с погрешностями расчета этих характеристик по фактическим данным. С этой точки зрения начало периода осреднения может выбираться произвольно, конечно, при условии, что само осреднение выполняется точно.

СПИСОК ЛИТЕРАТУРЫ

1. Жуковский Е. Е. и др. Исследование статистических характеристик относительной влажности воздуха.— Сборник трудов по агроном. физике, 1969, вып. 20, с. 5—28.

2. Марченко А. С., Помозова Л. И., Чубенко М. А. Временная статистическая структура метеорологических процессов.— Труды НИИАК, 1968, вып. 54 (4), с. 35—46.
3. Сибир Е. Е. Об амплитуде суточного хода параметров температуры воздуха.— См. наст. сб.
4. Тихомирова Л. В. Опыт исследования статистических закономерностей распределения температуры воздуха в отдельные часы суток.— Труды НИИАК, 1969, вып. 51, с. 92—107.
5. Федорченко Е. И. О временной структуре ежечасных наблюдений за температурой воздуха в Ленинграде.— Труды ГГО, 1975, вып. 348, с. 112—122.
6. Kraus E. B., Morrison R. E. Local interaction between the sea and the air at monthly and annual time scales.— Quart. Journ. Roy. Meteorol. Soc., 1966, vol. 92, N 391, p. 114—127.

Р. Л. Каган, Е. Е. Сибир

О ТОЧНОСТИ ОСРЕДНЕНИЯ ПЕРИОДИЧЕСКИ КОРРЕЛИРОВАННЫХ ВЕЛИЧИН

1. В работе [6] показано, что статистическая нестационарность метеорологических процессов в определенной мере сказывается на характеристиках их осредненных значений даже в том случае, если осреднение производится точно. В действительности метеорологические измерения производятся обычно дискретно во времени, что делает невозможным точное осреднение. В этих условиях средняя за период (например, за сутки) заменяется осредненным значением из имеющихся данных наблюдений (например, их средней арифметической). Очевидно, такая замена приводит как к систематическим, так и к случайным ошибкам в определении средней величины. В климатологической практике вопросу о систематических погрешностях осреднения уделяется значительное внимание. Разработаны способы их устранения как за счет рационального выбора сроков наблюдения, так и за счет использования подходящих способов осреднения (см., например, [1, 10]). Меньше внимания уделялось оценке случайных ошибок, что связано с трудоемкостью соответствующей обработки эмпирических данных, а также с трудностью обобщения полученных результатов.

Задача оценки точности осреднения существенно облегчается, если воспользоваться для этой цели подходом, основанным на использовании данных о временно́й структуре осредняемой величины. В этом случае необязательна непосредственная обработка экспериментальных данных, достаточно использовать априорную информацию о временно́й изменчивости, заданную, например, в виде корреляционной функции.

Такой подход к оценке точности одномерного осреднения был ранее применен в работах [2, 7] для определения точности получения средней путем арифметического осреднения данных равномерно расположенных на интервале осреднения отсчетов.

Оценки такого рода, выполнявшиеся для стационарных процессов, показали, что точность осреднения при заданной изменчивости осредняемых величин и точности их измерения зависит от числа сроков наблюдений на периоде, от расположения этих сроков относительно периода и от способа осреднения. Для нестационарных процессов должно, очевидно, сказаться влияние дополнительного фактора — вида и степени нестационарности. Важно представлять, насколько это влияние существенно по сравнению с влиянием упомянутых ранее факторов.

Пусть необходимо оценить истинную среднюю процесса $f(t)$ за период $(0, T)$

$$F_0=\frac{1}{T}\int_0^T f(t)\,dt. \qquad (1)$$

Для этой цели мы располагаем n измерениями на периоде осреднения в моменты $t_1, t_2, \ldots, t_n$. Обозначим $f_i=f(t_i)$ и учтем также, что данные измерений содержат ошибки, которые мы обозначим $\delta_1, \delta_2, \ldots, \delta_n$. Тогда оценка величины F_0 по данным измерений дается формулой

$$F=\sum_{i=1}^{n} p_i(f_i+\delta_i), \qquad (2)$$

где p_i — ве́са осреднения.

Ошибка осреднения составляет

$$\Delta=F-F_0=\sum_{i=1}^{n} p_i(f_i+\delta_i)-\frac{1}{T}\int_0^T f(t)\,dt. \qquad (3)$$

Как уже указывалось, оценка систематической погрешности осреднения $\overline{\Delta}$ (черта сверху означает статистическое осреднение) принципиальных трудностей не представляет. Будем считать, что эта погрешность устранена, так что

$$\overline{\Delta}=0, \qquad (4)$$

и будем характеризовать точность осреднения дисперсией величины Δ:

$$E^2=\overline{\Delta^2}=\sum_{i=1}^{n}\sum_{j=1}^{n} p_i p_j R_{ij}+\sum_{i=1}^{n} p_i^2\overline{\delta_i^2}+\frac{1}{T^2}\int_0^T\int_0^T R(t_1,\ t_2,)\,dt_1\,dt_2-$$

$$-\frac{2}{T}\sum_{i=1}^{n} p_i\int_0^T R(t_i,\ t)\,dt, \qquad (5)$$

где

$$R_{ij}=R(t_i,\ t_j)=\overline{f(t_i)\,f(t_j)}-\overline{f(t_i)}\cdot\overline{f(t_j)} \qquad (6)$$

— ковариация значений величин f в моменты времени t_i и t_j.

При выводе (5) предполагалось, что ошибки измерения случайны в том смысле, что их средняя величина равна нулю, а значения ошибок в разные моменты времени не зависят друг от друга и от значений измеряемой величины.

Интегрирование в (5) может быть выполнено точно лишь при некоторых простых типах ковариационных функций. В общем случае оно должно выполняться численно.

Полагаем, что нестационарность ковариационной функции определяется лишь нестационарностью средней величины $\overline{f}(t)$ и дисперсии $\sigma^2(t)$, а корреляционная функция r является стационарной и зависит лишь от сдвига по времени, так что

$$R(t_i,\ t_j)=\sigma(t_i)\,\sigma(t_j)\,r(\tau), \tag{7}$$

где $\tau=|t_i-t_j|$ — сдвиг по времени.

Будем считать также величину f периодически коррелированной с периодом, равным периоду осреднения T, так что

$$\sigma(t+T)=\sigma(t), \quad \overline{f}(t+T)=\overline{f}(T). \tag{8}$$

Пусть, кроме того, измерения производятся через равные интервалы времени $h=\frac{T}{n}$, а истинная средняя по периоду аппроксимируется средней арифметической из данных измерений. При этом

$$t_i=t_1+\frac{T}{n}(i-1); \quad p_i=\frac{1}{n}. \tag{9}$$

Такая упрощенная постановка задачи оправдывается тем, что она близка к реальности в случае определения средних суточных значений метеорологических элементов. В то же время она позволяет достаточно четко выявить влияние нестационарности на точность осреднения.

Подстановка (7)—(9) в (5) дает

$$E^2=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i\sigma_j r_{ij}+\frac{\eta^2}{n^2}\sum_{i=1}^{n}\sigma_i^2+\frac{1}{T^2}\int_0^T\int_0^T\sigma(t_1)\,\sigma(t_2)\times$$

$$\times r(|t_1-t_2|)dt_1\,dt_2-\frac{2}{Tn}\sum_{i=1}^{n}\sigma_i\int_0^T\sigma(t)\,r(|t-t_i|)\,dt. \tag{10}$$

Здесь использованы обозначения $r_{ij}=r(h|i-j|)$; $\sigma_i=$ $=\sigma(t_1+h(i-1))$; $\eta^2=\frac{\overline{\delta_i^2}}{\sigma_i^2}$ — мера ошибок измерения, которая принята одинаковой для всех сроков наблюдения t_i.

2. По приведенным формулам были выполнены расчеты для случая, когда ход среднего квадратического отклонения σ описывается гармоникой с периодом T:

$$\sigma(t)=\sigma_0\left[1+A\cos\frac{2\pi}{T}(t-t_0)\right]. \tag{11}$$

Здесь σ_0 — среднее значение σ, $A<1$ — его относительная амплитуда, t_0 — момент времени, в который σ достигает максимума.

Корреляционная функция аппроксимировалась экспоненциальной зависимостью

$$r(\tau)=\exp\left(-\frac{\tau}{T_0}\right), \tag{12}$$

где T_0 — масштаб корреляции.

Расчеты производились для различных значений параметров A, T/T_0, η, числа n сроков измерений на периоде, а также при различных сочетаниях моментов времени t_0 (максимум дисперсии) и t_1 (момент первого отсчета). Приведем некоторые результаты, иллюстрирующие зависимость точности осреднения от этих параметров.

При фиксированном числе измерений n средняя арифметическая характеризует истинную среднюю наиболее точно в том случае, если

$$t_i=\frac{T}{n}\left(i-\frac{1}{2}\right) \quad (i=1,\ 2,\ \ldots,\ n), \tag{13}$$

т. е. когда каждое из наблюдений приходится на середину интервалов длиной $h=\frac{T}{n}$, на которые разбит весь период осреднения. При смещении t_1 и всех остальных моментов измерения относительно этих наиболее выгодных моментов погрешности осреднения существенно увеличиваются. Особенно велики они при $t_1>\frac{T}{n}$ (или $t_1<0$), когда крайние сроки наблюдений выходят за пределы периода осреднения. Такое расположение моментов отсчета, конечно, крайне нежелательно и мы будем впредь полагать, что $0\leqslant t_1\leqslant\frac{T}{n}$ т. е. что все моменты наблюдений расположены в пределах периода осреднения.

Указанная закономерность, отмечавшаяся ранее [4] в стационарном случае ($A=0$), остается справедливой и для нестационарного процесса. Более того, она проявляется тем резче, чем больше степень нестационарности. Как следует из рис. 1, при неблагоприятном выборе моментов измерения точность осреднения при большем числе измерений может оказаться меньше, чем при осреднении данных меньшего числа оптимально размешенных отсчетов.

В табл. 1 представлены оценки относительных средних квадратических ошибок осреднения $\varepsilon=E/\sigma_0$ для случая $\eta^2=0{,}001$ и $t_0=0$ (начало периода осреднения соответствует максимуму дисперсии). При этом для каждого числа измерений n в табл. 1 приводятся данные лишь для двух размещений их на периоде: ε_1 — для оптимального ($t_1=T/2n$) и ε_2 — для наиболее неблагоприятного размещения ($t_1=0$).

Из табл. 1 видно, что как в стационарном, так и в нестационарном варианте ошибка осреднения тем больше, чем меньше число измерений на периоде и чем меньше связность осредняемой величины (т. е. чем больше T_0/T). Погрешность осреднения при оптимальном выборе сроков отсчета (ε_1) практически мало зависит от A и близка к погрешности осреднения для стационарного процесса. Исключение составляет случай $n = 1$, когда средняя по периоду заменяется данными измерений в момент времени, совпадающий с его центром. Этот случай большого практического интереса не представляет.

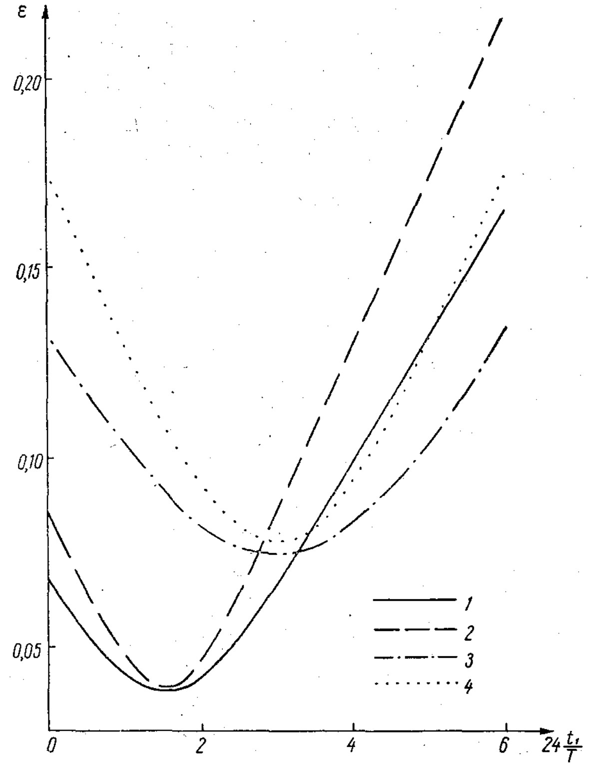

Рис. 1. Зависимость средней квадратической относительной ошибки осреднения $\varepsilon = E/\sigma_0$ от выбора момента первого отсчета ($t_0=0$, $T_0=2T$, $\eta^2=0{,}001$, корреляционная функция (12)).

1) $n=8$, $A=0$; 2) $n=8$, $A=0{,}4$; 3) $n=4$, $A=0$; 4) $n=4$, $A=0{,}4$.

При неблагоприятном размещении сроков наблюдений погрешность осреднения оказывается тем большей, чем больше степень нестационарности. Погрешность осреднения при этом возрастает примерно пропорционально параметру нестационарности A. Это значит, что правильный выбор сроков наблюдений тем важнее, чем больше нестационарность осредняемых величин. Отметим также существенное увеличение влияния нестационарности с ростом T_0.

Таблица 1

Средние квадратические (в % от σ_0) ошибки определения средней по периоду по данным n наблюдений ($t_0=0$, $\eta^2=0{,}001$); корреляционная функция (12)

n	T_0/T									
	0,1		0,5		1,0		2,0		5,0	
	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2
					$A=0$					
12	11	12	5	7	4	6	3	5	2	3
8	16	18	7	11	5	9	4	7	3	4
4	31	35	14	22	10	17	7	13	5	9
3	40	45	19	29	14	23	10	18	6	12
2	55	62	29	43	20	35	15	26	9	18
1	88	97	55	82	40	67	29	51	19	34
					$A=0{,}2$					
12	11	13	5	8	4	7	3	5	2	4
8	16	19	7	12	5	10	4	8	3	5
4	31	37	15	24	10	20	8	15	5	10
3	40	47	19	33	14	26	10	20	6	14
2	55	65	29	49	21	40	15	30	10	20
1	74	113	80	120	36	78	29	61	24	43
					$A=0{,}4$					
12	11	14	5	9	4	8	3	6	2	4
8	17	21	8	14	5	11	4	9	6	6
4	32	39	15	28	11	22	8	17	5	12
3	41	50	20	37	14	30	10	23	7	16
2	55	70	30	55	21	45	15	35	10	23
1	61	131	45	109	41	92	40	75	40	58

При несовпадении начала периода осреднения с максимумом дисперсии (при $t_0 \neq 0$) соотношение между точностью различных вариантов осреднения остается таким же. При оптимальном размещении моментов измерения точность осреднения оказывается близкой к точности осреднения в стационарном случае. Однако при других вариантах размещения влияние нестационарности в зависимости от сдвига максимума дисперсии t_0 оказывается различным. Это иллюстрируется табл. 2, в которой приводятся оценки точности осреднения для случаев $t_0=0$ (вариант I), $t_0=T/4$ (вариант И) и $t_0=T/2$ (вариант III) при $A=0{,}2$. Там же для сравнения приводятся оценки, соответствующие стационарному случаю, когда $A=0$ (вариант IV).

Из табл. 2 видно, что в варианте II, при котором начало периода приходится на момент перехода дисперсии через среднее значение, нестационарность практически не сказывается на точности

Таблица 2

Средние квадратические (в % от σ_0) ошибки определения средней по периоду при различном размещении его начала относительно максимума дисперсии ($A=0{,}2$, $\eta^2=0{,}001$); корреляционная функция (12)

n	ε	T_0/T															
		0,1				0,5				1,0				2,0			
		I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
12	ε_1	11	11	11	11	5	5	5	5	4	4	4	4	3	3	3	3
	ε_2	13	12	12	12	8	7	7	7	7	6	5	6	5	4	4	5
8	ε_1	16	16	16	16	7	7	7	7	5	5	5	5	4	4	4	4
	ε_2	19	18	18	18	12	11	10	11	10	9	8	9	8	7	6	7
4	ε_1	31	31	31	31	15	15	15	14	10	10	10	10	8	7	7	7
	ε_2	37	35	34	35	25	21	20	22	20	17	15	17	15	13	12	13
3	ε_1	40	40	40	40	19	19	19	19	14	14	14	14	10	10	10	10
	ε_2	47	45	44	45	33	28	26	29	26	22	20	23	20	17	15	18
2	ε_1	55	56	55	55	29	29	29	29	21	21	21	20	15	15	15	15
	ε_2	65	62	61	62	49	41	39	43	40	32	31	35	30	24	23	27

осреднения. В варианте III, при котором начало периода осреднения совпадает с минимумом дисперсии, нестационарность приводит к увеличению точности осреднения при неблагоприятном размещении сроков наблюдений. Однако уточнение это сравнительно невелико и, конечно, такое размещение остается нежелательным. При характерных для суточного хода температуры и влажности воздуха значениях параметров $A = 0{,}2$ и $T_0 = 2T$ даже в варианте III неблагоприятное размещение моментов отсчета приводит к увеличению средней квадратической ошибки примерно в 1,5 раза по сравнению с ошибками для оптимального размещения (в варианте I ошибка увеличивается примерно в 2 раза).

Представленные оценки получены при задании меры ошибки наблюдений $\eta^2 = 0{,}001$. Практически это, вероятно, максимальная точность, на которую можно рассчитывать при метеорологических измерениях. При увеличении ошибок измерений точность осреднения, естественно, уменьшается как в стационарном, так и в нестационарном случае. Некоторые оценки точности осреднения по данным восьми сроков измерений для различных значений η^2 приведены в табл. 3 (в этой таблице приняты те же обозначения вариантов, что и в табл. 2, а в качестве единицы времени принимается $T/24$, что соответствует, например, 1 ч при суточном осреднении). Из табл. 3 видно, что при увеличении ошибок влияние нестационарности уменьшается, однако это проявляется в сколько-нибудь заметной мере лишь при очень больших значениях η^2.

Таблица 3

Средние квадратические (в тысячных долях от σ_0) ошибки определения средней по периоду при различной точности данных измерений ($T_0=2T$, $n=8$, $A=0{,}2$); корреляционная функция (12)

$24\frac{t_1}{T}$	η^2											
	0			0,001			0,01			0,1		
	I	III	IV	I	III	IV	I	III	IV	I	III	IV
0	76	57	67	77	59	67	84	68	75	136	127	130
0,5	58	47	53	59	48	53	68	59	63	127	122	123
1,0	43	39	42	44	41	42	56	53	54	121	120	119
1,5	37	37	38	38	38	38	51	51	51	119	119	118
2,0	43	39	42	44	41	42	56	53	54	121	120	119
2,5	58	47	53	59	48	53	68	59	63	127	122	123
3,0	76	57	67	77	59	67	84	68	75	136	127	130
3,5	95	69	81	96	70	82	101	78	89	148	132	138
4,0	115	83	98	115	83	99	120	90	105	161	140	149

Таким образом, влияние нестационарности существенно повышает погрешность осреднения при неблагоприятном размещении моментов измерения на периоде осреднения, однако практически не сказывается при оптимальном их размещении согласно (13). В этом случае точность осреднения можно оценивать без учета нестационарности.

Как следует из [5], в стационарном случае точность осреднения процессов, характеризующихся экспоненциально убывающей корреляционной функцией (12), при таком размещении моментов измерения осредняемых величин приближенно описывается формулой

$$\varepsilon_1 \approx \sqrt{\frac{1}{6}\frac{T}{T_0}\frac{1}{n^2}+\frac{\eta^2}{n}}, \tag{14}$$

погрешность которой при $\frac{T}{T_0 n} < 2$ не превышает нескольких процентов. Очевидно, эта формула вполне пригодна для оценки точности осреднения и в нестационарном случае при условии оптимального выбора моментов измерения.

Аналогичные оценки точности осреднения при различной связности и различном числе и точности наблюдений были выполнены также для процессов, корреляционная функция которых имеет вид

$$r(\tau)=\left(1+\frac{\tau}{T_0}\right)\exp\left(-\frac{\tau}{T_0}\right). \tag{15}$$

Мы ограничимся лишь представлением табл. 4, которая аналогична приведенной ранее табл. 1. Из нее видно, что отмеченные выше закономерности имеют место и для таких процессов. При этом зависимость точности осреднения от выбора t_1 оказывается еще большей, чем для процессов, описывающихся корреляционной функцией (12), так что оптимальный выбор сроков в этом случае оказывается еще более важным.

3. Приведенные оценки позволяют судить о характере влияния нестационарности на точность определения средних величин по дискретным данным и о соотношении его с влиянием других факторов. Однако они получены при специфическом задании временно́й структуры осредняемого процесса формулами (11) и (12) (или (15)), так что эти оценки в количественном отношении могут отличаться от фактических при осреднении метеорологических элементов с существенно отличной временно́й структурой. При наличии надежных данных о временно́й структуре не представляет труда выполнить соответствующие расчеты по приведенным выше формулам. Более того, при этом не обязательны принятые нами для упрощения расчетов предположения о стационарности корреляционных функций и о гармоническом ходе среднего квадратического отклонения осредняемых величин.

Практически, однако, принципиальная возможность оценивать точность осреднения по формуле (5) по фактическим данным о структуре процесса без использования упрощающих предположений о них осложняется сравнительно небольшой надежностью экспериментальных статистических данных, которая может приводить к значительным ошибкам в результатах расчетов.

Таблица 4

Средние квадратические (в % от σ_0) ошибки определения средней по периоду по данным n наблюдений ($t_0=0$, $\eta^2=0{,}001$); корреляционная функция (15)

n	T_0/T									
	0,1		0,5		1,0		2,0		5,0	
	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2
					$A=0$					
12	2	6	1	5	1	3	1	2	1	1
8	4	10	1	7	1	5	1	3	1	2
4	13	23	2	14	2	9	2	6	2	3
3	21	32	3	19	2	12	2	7	2	4
2	37	52	6	28	3	19	2	11	2	5
1	75	95	22	60	10	38	5	21	3	10
					$A=0{,}2$					
12	2	7	1	6	1	4	1	2	1	1
8	4	11	1	8	1	6	1	3	1	2
4	13	26	2	17	2	11	2	7	2	3
3	21	36	4	22	2	15	2	9	2	4
2	37	57	7	34	4	22	3	13	2	6
1	64	109	22	70	19	47	20	31	20	23
					$A=0{,}4$					
12	2	9	1	6	1	4	1	3	1	1
8	4	13	1	10	1	6	1	4	1	2
4	14	29	2	19	2	13	2	8	2	4
3	22	41	4	26	2	17	2	10	2	5
2	37	63	9	40	4	26	3	15	2	7
1	56	124	35	84	19	61	12	48	6	42

В этом смысле показательна работа [12], в которой оценивается точность получения средних суточных значений температуры воздуха по наблюдениям за ограниченное число сроков в Варшаве. При расчетах использована формула, аналогичная формуле (5), а суточный ход статистических характеристик задавался по фактическим данным при отказе от упрощающих предположений. Однако правильная по существу постановка задачи была в значительной мере сведена на нет тем обстоятельством, что при расчете корреляционных функций использовались данные за несколько полных лет без исключения влияния годового хода. Это привело к значительному завышению временно́й корреляции, которая в [12] для сдвига времени, равного суткам, превышает 0,9 вместо значений порядка 0,70—0,75. Более того, представленные в [12] ковариационные матрицы имеют отрицательные собственные числа и, следовательно, не являются положительно определенными. Поэтому полученные в [12] значения ошибок средних суточных температур представляются сомнительными.

Опыт показывает, что необходима существенная предварительная обработка фактических данных о временной структуре, используемых при численных оценках по формуле (5). Наиболее важным является обеспечение положительной определенности ковариационных матриц, что сравнительно просто достигается в предположении стационарности корреляционной функции, а при отказе от него является довольно затруднительным.

Определенные трудности представляет и выбор метода численного интегрирования в (5) в тех случаях, когда детальность измерений (определяемая их числом на периоде) близка к максимальной детальности задания корреляционных функций. В этих условиях погрешности осреднения, полученные по формуле (5), могут определяться не столько фактической его точностью, сколько отличием весов осреднения от весов, принятых при численном интегрировании корреляционной функции [3].

В силу сказанного в выполнявшихся нами расчетах при аналитическом задании корреляционной функции (12) интегрирование велось методом Симпсона при делении периода на 48 интервалов. Оценки показывают, что при таком шаге интегрирования указанный эффект на точности вычислений не сказывается.

Приведем в качестве примера выполненные по приведенной методике оценки точности вычисления средних суточных значений температуры воздуха на ст. Ленинград, ГМО. В качестве исходных использовались корреляционные функции и ежечасные значения дисперсии температуры воздуха, полученные в [9] (влияние годового хода в них исключено). При этом для устранения возможных нарушений положительной определенности получающейся корреляционной матрицы взяты значения корреляционной функции, осредненные по суточному периоду в предположении ее стационарности. Это потребовало существенного сглаживания корреляционной функции для июля. В результате для расчета использовались данные о структуре, нредставленные в табл. 5 (средние квадратические отклонения) и табл. 6 (корреляционные функции). Значения меры ошибок измерения, согласно [9], были приняты 0,001 для января и 0,004 для июля, что соответствует абсолютным значениям средних квадратических ошибок примерно 0,2 °C.

Из табл. 5 видно, что среднее квадратическое отклонение температуры в январе имеет для данной станции максимум в начале суток и минимум в послеполуденные часы, однако в количественном отношении суточный ход его невелик, амплитуда суточного хода составляет лишь около 0,04 от среднего значения σ. Заметно больше (порядка 0,2 от среднего значения) оказывается амплитуда суточного хода для июля. При этом, в отличие от января, максимум дисперсии в этом месяце приходится на послеполуденные часы, а минимум — на начало суток.

Полученные оценки погрешностей осреднения при разном числе сроков наблюдений в сутки и при разном их размещении приведены в табл. 7. При этом имеется в виду, что осреднение ведется по календарным суткам (от 0 до 24 ч московского времени)

Таблица 5

Суточный ход (°C) среднего квадратического отклонения температуры воздуха. Ленинград, ГМО

t ч	1	2	3	4	5	6	7	8	9	10	11	12
Январь	6,5	6,5	6,5	6,5	6,6	6,6	6,6	6,6	6,6	6,5	6,4	6,2
Июль	2,7	2,7	2,7	2,7	2,7	2,6	2,8	3,1	3,4	3,6	3,7	3,8
***t* ч**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
Январь	6,1	6,1	6,1	6,2	6,3	6,4	6,5	6,5	6,6	6,6	6,6	6,6
Июль	3,8	3,9	3,8	3,8	3,6	3,4	3,3	3,1	2,8	2,7	2,7	2,7

Таблица 6

Временная корреляционная функцня (в тысячных долях едииицы) температуры воздуха. Ленинград, ГМО

τ ч	1	2	3	4	5	6	7	8	9	10	11	12
Январь	998	994	987	978	967	956	944	932	920	908	896	884
Июль	977	950	915	882	847	816	785	760	738	720	707	695
τ ч	13	14	15	16	17	18	19	20	21	22	23	24
Январь	873	862	851	840	829	818	808	797	787	777	767	757
Июль	687	678	672	665	660	656	653	650	649	648	645	638

методом среднего арифметического. Там же для сравнения приведены оценки для стационарного варианта, в котором использованы средние за сутки значения среднего квадратического отклонения, равные 6,4 °C для января и 3,3 °C для июля.

Таблица 7

Средние квадратические ошибки (°C) определения средней суточной температуры воздуха по данным n сроков наблюдений через $24/n$ часов. Ленинград, ГМО

Первый срок, ч	Январь				Июль			
	n							
	3	4	8	12	3	4	8	12
	Стационарный вариант							
1	0,59	0,40	0,12	0,06	0,49	0,34	0,14	0,08
2	0,41	0,22	0,12	0,20	0,42	0,29	0,14	0,17
3	0,24	0,12	0,29	0,38	0,35	0,25	0,22	0,24
4	0,15	0,22	0,47	0,56	0,35	0,29	0,32	0,36
5	0,24	0,40	0,66	0,75	0,35	0,34	0,42	0,46
6	0,41	0,58	0,84	0,93	0,42	0,44	0,52	0,56
E_0	0,15	0,12	0,08	0,06	0,35	0,29	0,13	0,08
E_1	0,77	0,58	0,29	0,20	0,58	0,44	0,22	0,17
E_2	0,12	0,10	0,07	0,06	0,12	0,10	0,07	0,06
E_3	0,09	0,06	0,02	0,01	0,33	0,23	0,11	0,05
	Нестационарный вариант							
1	0,60	0,40	0,12	0,06	0,45	0,31	0,14	0,08
2	0,41	0,23	0,12	0,20	0,40	0,29	0,14	0,16
3	0,24	0,12	0,29	0,38	0,35	0,25	0,19	0,21
4	0,15	0,23	0,48	0,57	0,34	0,28	0,27	0,31
5	0,24	0,40	0,67	0,76	0,35	0,32	0,35	0,38
6	0,41	0,58	0,85	0,94	0,39	0,40	0,43	0,47
E_0	0,15	0,12	0,08	0,06	0,34	0,25	0,13	0,08
E_1	0,77	0,58	0,29	0,20	0,49	0,39	0,19	0,16

Наряду со значениями погрешностей $E(t_1)$, характеризующими точность осреднения при использовании в качестве первого срока наблюдений момента времени t_1(1, 2, ..., 6 ч), приведены оценки средних квадратических погрешностей осреднения E_0 при оптимальном размешении сроков, когда $t_1 = T/2n$ (в нашем случае, например, при $n = 8$ это соответствует $t_1 = 01$ ч 30 мин), и E_1 при «неблагоприятном» размещении[1], когда $t_1 = T/n$. Для стационар-

---

[1] В действительности, конечно, случай, когда $t_1 > T/n$ (например, $t_1 \geqslant 6$ ч при $n=8$), является еще более неблагоприятным, поскольку при этом в осреднении участвуют на равных правах наблюдения, выходящие за пределы периода осреднения, однако мы уже условились ранее такие случаи не рассматривать.

ного варианта приведены также оценки $E_2=\frac{\sigma\eta}{n}$, характеризующие погрешность осреднения за счет случайных ошибок измерений, и $E_3=\sqrt{E_0^2-E_2^2}$, определяющие гипотетическую точность осреднения при оптимальном выборе сроков и при отсутствии ошибок измерения.

Представленные в табл. 7 оценки как по порядку величины, так и по характеру зависимости от t_1 неплохо согласуются с оценками точности определения средней суточной температуры, полученными в [11] путем непосредственной обработки фактических данных наблюдений в Вансдорфе (ГДР).

Из табл. 7 следует, что влияние нестационарности в январе очень невелико. Это является естественным следствием малого суточного хода дисперсии в этом месяце. Для июля нестационарность дисперсии в течение суток является существенной и приводит к заметному уменьшению погрешностей осреднения при неблагоприятном выборе сроков.

При сравнении погрешностей осреднения для января и июля следует иметь в виду, что различия их при одном и том же выборе сроков определяются двумя факторами. Во-первых, температура воздуха в зимние месяцы характеризуется гораздо большей временно́й связностью, чем в летние (см. табл. 6). Это обеспечивает при прочих равных условиях меньшую относительную (в долях σ) погрешность осреднения. Во-вторых, дисперсия температуры воздуха в зимние месяцы в несколько раз больше дисперсии летней температуры (см. табл. 5). Это может приводить к тому, что при определенных условиях ошибка осреднения для января превышает соответствующую ошибку осреднения для июля. Именно такое сочетание факторов имеет место при неблагоприятном расположении сроков измерения. В этих условиях средняя квадратическая погрешность (E_1) осреднения трехсрочных, четырехсрочных и даже восьмисрочных наблюдений для января оказывается в 1,5 раза больше, чем для июля. При оптимальном же выборе сроков соответствующая погрешность (E_0) оказывается в январе примерно вдвое меньше, чем для июля.

Следует отметить, что ошибки измерения вносят в абсолютном выражении для обоих месяцев одинаковый вклад в погрешность осреднения (E_2). Поскольку, однако, этот вклад существенно больше погрешностей E_3, которые имели бы место при осреднении истинных значений температуры для января, полная погрешность осреднения при оптимальном выборе сроков в значительной мере определяется именно ошибками измерения. Для июля дело обстоит так лишь при большом числе сроков ($n\geqslant 12$), при меньшем числе сроков ошибки измерения сказываются сравнительно мало.

Рассмотрение данного примера подтверждает полученные ранее выводы о том, что даже значительная нестационарность периодически коррелированного процесса мало сказывается на точности осреднения его по периоду при условии использования данных наблюдений в правильно подобранные сроки или при отнесении наблюдений в заранее заданные сроки к специально подобранному периоду. Важно, что при этом достигается и гораздо бо́льшая точность осреднения. Так, средняя арифметическая из восьми наблюдений температуры за сроки 00, 03, 06, 09, 12, 15, 18 и 21 ч московского времени на ст. Ленинград, ГМО характеризует истинную среднюю за календарные сутки от 00 до 24 ч или за метеорологические сутки от 21 до 21 ч (см. [8]) с погрешностью, значительно большей (в июле в 1,5 раза, а в январе даже в 3,5 раза), чем за период с 22 ч 30 мин.

Разумеется, подбор сроков измерений или периода осреднения практически производится с учетом многих соображений, из которых обеспечение минимума случайной ошибки осреднения, рассматриваемой нами в данном случае, является отнюдь не самым важным. Часто приходится учитывать такие соображения, как возможность правильного описания суточного хода с фиксацией экстремумов в нем, а иногда даже экономическую возможность обеспечения наблюдений в те или иные часы суток. Большое значение имеет единообразие и сравнимость данных наблюдений в различных географических районах и в разные периоды времени. Ради такого единообразия нередко предпочтительнее производить наблюдения в единые сроки и при изменении метеорологических суток от района к району.

В этих условиях может оказаться целесообразным среднее значение определять не методом среднего арифметического, а путем использования более точных квадратурных формул. Так, уже использование простой формулы трапеций может дать значительный выигрыш в точности средней величины. Важно, что при этом не требуется никаких изменений в организации наблюдений, а вносятся лишь небольшие изменения в процессе обработки, что существенных трудностей не представляет. В перспективе, по мере накопления данных о статистической структуре метеорологических процессов, можно думать и об оптимальном осреднении с учетом этих данных [4], что должно обеспечить дальнейшее повышение точности используемых средних величин.

СПИСОК ЛИТЕРАТУРЫ

1. Вильд Г. О температуре воздуха в Российской Империи.— СПБ, 1882.— 393 с.
2. Гандин Л. С., Каган Р. Л. О точности определения средней высоты снежного покрова по дискретным данным.— Труды ГГО, 1962, вып. 130, с. 3—10.
3. Гандин Л. С., Каган Р. Л. К вопросу об аппроксимации характеристик статистической структуры.— Труды ГГО, 1966, вып. 191, с. 18—21.
4. Гандин Л. С., Соловейчик Р. Э. Об оптимальных формулах численных квадратур для стационарных случайных функций.— Зап. Ленинградского горного института, т. 43, вып. 3, 1964, с. 48—58.
5. Каган Р. Л. О точности определения средней по площади по данным точечных измерений.— Труды ГГО, 1965, вып. 175, с. 117—131.

6. Каган Р. Л., Сибир Е. Е. К вопросу о выборе интервала осреднения периодически коррелированных величин.— См. наст. сб.
7. Лайхтман Д. Л., Каган Р. Л. Некоторые вопросы рационализации снегосъемок.— Труды ГГО, 1960, вып. 108, с. 3—18.
8. Наставление гидрометеорологическим станциям и постам, вып. 3, ч. 1. Л.: Гидрометеоиздат, 1969.— 308 с.
9. Федорченко Е. И. О временной структуре ежечасных наблюдений за температурой воздуха в Ленинграде.— Труды ГГО, 1975, вып. 348, с. 112—122.
10. Conrad V., Pollak L. W. Methods in climatology.— Cambridge, Massachusetts: Harvard University Press, 1950.— 459 p.
11. Junghans H. Der Terminfehler bei Lufttemperatur, relativer Luftfeuchte, Dampfdruck und Luftdruck.— Zeitschr. für Met., 1969, Bd 21, H. 3—4, S. 113—117.
12. Koziel S. O średnej dobowej i optymalnych terminach obserwacji temperatury powietrza.— Wiad. meteorol. i gosp. wodnej, 1976, vol. 3, N 2, p. 77—85.

Р. Л. Каган, Е. И. Хлебникова

К ВОПРОСУ О СХОДИМОСТИ РАСПРЕДЕЛЕНИЯ К НОРМАЛЬНОМУ ДЛЯ СУММ СВЯЗНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

1. В метеорологии, наряду со значениями метеорологических величин, полученными в те или иные моменты времени, приходится постоянно иметь дело и с осредненными по времени значениями. При этом осреднение может производиться как непрерывно, путем использования каких-либо интеграторов, так и дискретно, по данным в те или иные моменты времени. Поэтому важно представлять себе, какими вероятностными характеристиками обладают величины, полученные путем непрерывного или дискретного осреднения, а также ошибки замены непрерывных средних их дискретными аналогами.

Подобная задача решается просто для случая, когда осреднению подвергаются нормально распределенные величины. В этом случае распределения средних величин и отклонений от средних также являются нормальными и полностью определяются соответствующими дисперсиями, получение которых не вызывает больших затруднений.

В случае же когда распределение осредняемых величин отклоняется от нормального, знания дисперсии недостаточно, необходимо иметь представление о старших моментах исследуемых величин, которые в общем случае оценить затруднительно.

В соответствии с центральной предельной теоремой распределение средних величин асимптотически стремится к нормальному при произвольном распределении исходных величин. Можно думать поэтому, что при достаточном осреднении средние величины могут практически считаться нормальными и в тех случаях, когда исходные величины существенно отклоняются от этого распределения. Тем более это должно иметь место, если мы интересуемся характеристиками не самих средних величин, а ошибками их определения, распределение которых должно сходиться к нормальному с увеличением объема осредняемой выборки гораздо быстрее.

В настоящей статье указанные закономерности рассматриваются на примере специфически заданных стационарных случайных процессов, для которых выводятся соответствующие аналитические формулы. Полученные оценки представляют и самостоятельный интерес, поскольку такие процессы хорошо описывают изменение во времени составляющих кинетической энергии движения в свободной атмосфере или скоростного напора ветра.

Хотя можно думать, что полученные результаты в качественном отношении справедливы и для процессов с другой временно́й структурой, соответствующие количественные оценки для них могут быть совершенно иными. Поскольку аналитически получить их не всегда возможно, часто предпочтительнее использовать метод статистического моделирования. Этот метод, однако, также имеет некоторые ограничения, прежде всего в отношении точного получения нужных характеристик по выборкам, которые можно смоделировать и обработать в пределах разумного машинного времени. В связи с этим в настоящей статье выполнены также некоторые оценки точности расчета характеристик осредненных величин и ошибок их определения методом статистического моделирования.

2. Рассмотрим случайный процесс $y(t) = x^2(t)$, где $x(t)$ — стационарный гауссовский случайный процесс с нулевым математическим ожиданием, дисперсией σ^2 и корреляционной функцией

$$r(\tau) = \exp(-\tau/T_0), \tag{1}$$

где T_0 — масштаб корреляции.

Наряду с этим случайным процессом рассмотрим его интегральное среднее (за период T) значение

$$F(t) = \frac{1}{T} \int\limits_t^{t+T} y(s)\, ds$$

и его дискретный аналог

$$G(t) = \frac{1}{n} \sum_{i=1}^{n} y\left(t + \Delta\left(t - \frac{1}{2}\right)\right),$$

где $n = T/\Delta$ — число отсчетов, Δ — интервал между отсчетами.

Оценки показывают (см., например, [4]), что выбор отсчетов, осуществленный при получении среднего арифметического в виде $G(t)$, дает наименьшую ошибку

$$\delta(t) = G(t) - F(t) \tag{2}$$

замены величины F ее дискретным аналогом.

Нетрудно получить моменты распределения величин F, G и δ. Выпишем формулы для определения первых трех моментов, при этом в силу стационарности процесса

$$\overline{F(t)} = \overline{G(t)} = \sigma^2, \tag{3}$$

$$\overline{\delta(t)} = 0, \tag{4}$$

где черта сверху означает операцию математического ожидания:

$$\sigma_F^2 = \overline{\left(F(t) - \overline{F(t)}\right)^2} = \sigma^4 \left[\frac{2}{\beta} - \frac{1 - \exp(-2\beta)}{\beta^2}\right], \tag{5}$$

$$\sigma_G^2 = \overline{\left(G(t) - \overline{G(t)}\right)^2} = \sigma^4 \left[\frac{2}{n^2}\left(\frac{n(1+p^2)}{1-p^2} - \frac{2p^2(1-p^{2n})}{(1-p^2)^2}\right)\right], \tag{6}$$

$$\sigma_\delta^2 = \overline{\delta^2(t)} = \sigma^4 \left[\frac{4dp(1-p^{2n})}{\beta^2(1-p^2)} - \frac{2}{\beta} - \frac{1-p^{2n}}{\beta^2} + \right.$$

$$\left. + \frac{2d(1+p^2)}{\beta(1-p^2)} - \frac{4d^2p^2(1-p^{2n})}{\beta^2(1-p^2)^2}\right]. \tag{7}$$

Здесь $\beta = T/T_0$ — безразмерный интервал осреднения, $d = \Delta/T_0$ — безразмерный интервал между отсчетами, $p = \exp(-\Delta/T_0)$ — коэффициент корреляции между смежными членами исходного ряда x.

Из формул (5)—(7) следует, что с увеличением длины реализации при прочих равных условиях дисперсия как средних величин F и G, так и их разности быстро уменьшается. При больших значениях длины реализации дисперсия величины F обратно пропорциональна T $\left(\sigma_F^2 \to \frac{2\sigma^4}{\beta}\right)$. Дисперсия разности σ_δ^2 при этом также убывает обратно пропорционально T при фиксированном интервале между отсчетами. Однако в случае если число отсчетов n на интервале зафиксировано, значение σ_δ^2 стремится к постоянной $\frac{2\sigma^4}{n}$, совпадающей с асимптотическим значением σ_G^2 при больших интервалах между отсчетами Δ (или при малых интервалах корреляции T_0).

При малых значениях интервалов T и Δ удобно использовать получающиеся путем разложения в ряд формулы:

$$\sigma_F^2 = 2\sigma^4 \left[1 - \frac{2}{3}\beta + \frac{1}{3}\beta^2 + \ldots\right], \tag{5a}$$

$$\sigma_G^2 = 2\sigma^4 \left[1 - \frac{2}{3}\,\frac{n^2-1}{n}\,d + \frac{1}{3}(n^2-1)\,d^2 + \ldots\right], \tag{6a}$$

$$\sigma_\delta^2 = \frac{2\sigma^4}{3}\,\frac{d}{n}\left[1 - \frac{3d^2}{20} + \ldots\right]. \tag{7a}$$

Из этих формул следует, что при уменьшении интервала осреднения (соответственно при уменьшении числа отсчетов при фиксированном интервале между ними) дисперсия осредненных значений F и G стремится к дисперсии неосредненных величин y, т. е. к величине $2\sigma^4$. Дисперсия же ошибки осреднения δ быстро уменьшается как за счет уменьшения интервала осреднения, так и за счет увеличения числа измерений на нем. Из (7а) видно, что при фиксированном d дисперсия ошибки убывает приблизительно пропорционально росту числа отсчетов n. Это соответствует отмеченной в [3] закономерности, характерной для любых рядов с экспоненциально убывающей корреляцией. В случае фиксированного интервала осреднения T вместо формулы (7а) можно записать

$$\sigma_\delta^2 = \frac{\sigma^4\beta}{n^2}\left[1 - \frac{3\beta^2}{20n^2} + \ldots\right], \tag{7б}$$

т. е. в этом случае линейно убывает с увеличением числа отсчетов не дисперсия, как это было бы при отсутствии корреляции, а среднее квадратическое отклонение ошибки.

Более подробно описывать зависимости (5)—(7) нецелесообразно, учитывая, что они детально исследованы и отчасти затабулированы в работах [2, 3, 5, 6]. Результаты этих работ можно использовать с учетом того, что $\sigma_y^2 = 2\sigma_x^4$, а коэффициент корреляции между двумя смежными членами последовательности y равен p^2.

Моменты третьего порядка величин F, G и δ определяются формулами:

$$m_F = \overline{\left(F(t) - \overline{F(t)}\right)^3} = \frac{12\sigma^6}{\beta^2}\left[1 - \frac{1}{\beta} + \left(1 + \frac{1}{\beta}\right)\exp(-2\beta)\right], \tag{8}$$

$$m_G = \overline{\left(G(t) - \overline{G(t)}\right)^3} = \frac{8\sigma^6}{n^3}\left[n\,\frac{1 + 4p^2 + p^4}{(1-p^2)^2} + \frac{6np^{2(n+1)}}{(1-p^2)^2} - \right.$$
$$\left. - \frac{6p^2(1+p^2)(1-p^{2n})}{(1-p^2)^3}\right], \tag{9}$$

$$m_\delta = \overline{\delta^3(t)} = \sigma^6\left[-\frac{24d(1+p^2)}{\beta^2(1-p^2)} + \frac{8d^2}{\beta^2} + \frac{12(1-p^{2n})}{\beta^3} + \frac{24}{\beta^2} + \right.$$
$$\left. + \frac{48d^2p^2(1-p^{2n})}{\beta^3(1-p^2)^2} - \frac{48dp(1-p^{2n})}{\beta^3(1-p^2)}\right]. \tag{10}$$

Из формул (8) и (9) видно, что с увеличением интервала осреднения T третьи моменты величин F и G убывают обратно пропорционально квадрату T. Следовательно, с учетом формул (5) и (6) убывание асимметрии этих величин при больших интервалах осреднения происходит обратно пропорционально $\sqrt{T}$. При увеличении T для фиксированного числа отсчетов n третий момент ошибки осреднения δ стремится к величине $\frac{8\sigma^6}{3}$, соответствующей случаю осреднения некоррелированных величин. Асимметрия ошибки осреднения в этом случае стремится к величине $\sqrt{\frac{8}{n}}$.

При малых значениях интервала осреднения, а также величины d могут быть получены разложения:

$$m_F=\frac{8\sigma^6}{d^2}\left(1-d+\frac{3}{5}d^2+\ldots\right), \tag{11}$$

$$m_G=8\sigma^6\left(1-nd+\frac{d}{n}+\frac{3n^2d^2}{5}-d^2+\frac{2d^2}{5n^2}+\ldots\right), \tag{12}$$

$$m_\delta=\frac{6d^2\sigma^6}{5n^2}\left(1-\frac{5}{9}nd+\frac{10}{27}n^2d^2+\ldots\right). \tag{13}$$

Наибольший практический интерес представляет рассмотрение не самих третьих моментов, а соответствующих коэффициентов асимметрии $A=\frac{m}{\sigma^3}$. Расчет их может быть выполнен по формулам (5)—(10). Некоторые результаты расчетов приведены в табл. 1.

В табл. 1 представлена зависимость A_G от d и T. Первая строка таблицы, соответствующая $d=0$, характеризует зависимость коэффициента асимметрии величины F, обозначаемого далее A_F, от интервала осреднения. Эта величина до значений T порядка T_0 мало отличается от значений коэффициента асимметрии для неосредненных значений y, равного $2\sqrt{2}$. При больших интервалах осреднения A_F убывает приблизительно обратно пропорционально $\sqrt{T}$ и асимптотически приближается к величине $\sqrt{\frac{18}{\beta}}$.

При оценке асимметрии средней арифметической величины A_G следует иметь в виду, что при малых интервалах d между отсчетами она, естественно, должна быть близка к A_F. При больших интервалах между отсчетами, превышающих интервал корреляции исходных величин T_0 (при этом корреляции между осредняемыми величинами меньше e^{-2}) асимметрия средней арифметической величины должна приближаться к асимметрии суммы независимых значений y. Последняя, как нетрудно показать, определяется формулой

$$A_G\big|_{r(1)=0}=\sqrt{\frac{8}{n}}=\sqrt{\frac{8d}{\beta}}. \tag{14}$$

Из табл. 1 видно, что до значений безразмерного интервала d порядка 1 значения A_G действительно близки к A_F и лишь слегка уменьшаются с увеличением d. В окрестности $d=1$ величина A_G достигает минимального значения, после чего начинает увеличиваться, асимптотически приближаясь к значениям, определяемым формулой (14).

В табл. 1 представлены также аналогичные зависимости для коэффициента асимметрии ошибки осреднения δ. Из этой таблицы видно, что при не слишком больших интервалах между отсчетами ($d<1$) коэффициент асимметрии при фиксированном T линейно

Таблица 1

Коэффициенты асимметрии средних величин и ошибок осреднения в зависимости от интервала осреднения и дискретности отсчетов на нем

d	$\beta = T/T_0$									
	0,1	0,2	0,5	1,0	2,0	5,0	10	20	50	100
					Средние величины					
0	2,83	2,82	2,79	2,68	2,41	1,78	1,30	0,94	0,60	0,42
0,005	2,83	2,82	2,79	2,68	2,41	1,78	1,30	0,94	0,60	0,42
0,010	2,83	2,82	2,79	2,68	2,41	1,78	1,30	0,94	0,60	0,42
0,020	2,83	2,82	2,79	2,68	2,41	1,78	1,30	0,94	0,60	0,42
0,025	2,83	2,82	2,79	2,68	2,41	1,78	1,30	0,94	0,60	0,42
0,050	2,83	2,82	2,79	2,68	2,41	1,78	1,30	0,94	0,60	0,42
0,100	2,83	2,82	2,78	2,68	2,41	1,77	1,30	0,93	0,60	0,42
0,200	—	2,83	—	2,67	2,40	1,76	1,29	0,93	0,59	0,42
0,250	—	—	2,77	2,66	2,39	1,76	1,29	0,93	0,59	0,42
0,500	—	—	2,83	2,63	2,33	1,71	1,26	0,90	0,58	0,41
1,000	—	—	—	2,83	2,32	1,65	1,20	0,86	0,56	0,39
2,000	—	—	—	—	2,83	—	1,44	0,94	0,60	0,42
5,000	—	—	—	—	—	2,83	2,00	1,41	0,89	0,63
10,00	—	—	—	—	—	—	2,83	2,00	1,26	0,89
20,00	—	—	—	—	—	—	—	2,83	—	1,26
50,00	—	—	—	—	—	—	—	—	2,83	2,00
100,0	—	—	—	—	—	—	—	—	—	2,83
					Ошибки осреднения					
0,005	0,03	0,02	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00
0,010	0,06	0,04	0,02	0,01	0,01	0,00	0,00	0,00	0,00	0,00
0,020	0,13	0,09	0,05	0,03	0,02	0,01	0,01	0,00	0,00	0,00
0,025	0,17	0,11	0,03	0,04	0,02	0,01	0,01	0,01	0,00	0,00
0,050	0,33	0,22	0,12	0,08	0,05	0,02	0,02	0,01	0,01	0,00
0,100	0,66	0,44	0,25	0,15	0,09	0,05	0,03	0,02	0,01	0,01
0,200	—	0,89	—	0,30	0,18	0,10	0,07	0,05	0,03	0,02
0,250	—	—	0,62	0,38	0,23	0,12	0,08	0,06	0,04	0,02
0,500	—	—	1,25	0,76	0,45	0,25	0,16	0,11	0,07	0,05
1,000	—	—	—	1,52	0,91	0,49	0,33	0,23	0,14	0,10
2,000	—	—	—	—	1,78	—	0,66	0,45	0,28	0,20
5,000	—	—	—	—	—	2,17	1,49	1,04	0,65	0,46
10,00	—	—	—	—	—	—	2,44	1,72	1,08	0,77
20,00	—	—	—	—	—	—	—	2,63	—	1,17
50,00	—	—	—	—	—	—	—	—	2,75	1,94
100,0	—	—	—	—	—	—	—	—	—	2,79

возрастает с ростом *d* или, что то же самое, линейно убывает с ростом числа отсчетов *n*. Такая зависимость может быть объяснена на основе подхода, аналогичного использованному в [3] для оценки дисперсии ошибки осреднения, а именно можно рассматривать ошибку осреднения как среднюю арифметическую из ошибок, с которыми каждый из отсчетов представляет среднее значение на части интервала T/n, на центр которого он приходится. Как показано в [3], при экспоненциально-убывающей корреляционной функции эти элементарные ошибки являются практически некоррелированными уже для смежных интервалов. Поэтому, обозначая через $A_1(d)$ асимметрию элементарной ошибки, зависимость A_δ при фиксированном T от d и n можно приближенно описывать формулой

$$A_\delta = A_1(d)/\sqrt{n}. \qquad (15)$$

Величина $A_1(d)$ получается путем подстановки в формулы (7) и (10) значения $n = 1$, что дает

$$A_1(d) = 2\sqrt{2}\,\frac{1 - \frac{3}{d} + \frac{3}{d^2} - \frac{6p}{d^2} + \frac{3(1-p^2)}{2d^3}}{\left(1 - \frac{1}{d} + \frac{2p}{d} - \frac{1-p^2}{2d^2}\right)^{3/2}}. \qquad (16)$$

Зависимость $A_1(d)$ графически представлена на рис. 1, из которого видно, что при малых интервалах A_1 возрастает с увеличением d примерно пропорционально $\sqrt{d}$ или обратно пропорционально $\sqrt{n}$, что в сочетании с зависимостью (15) приводит к упоминавшемуся выше линейному убыванию асимметрии с ростом n.

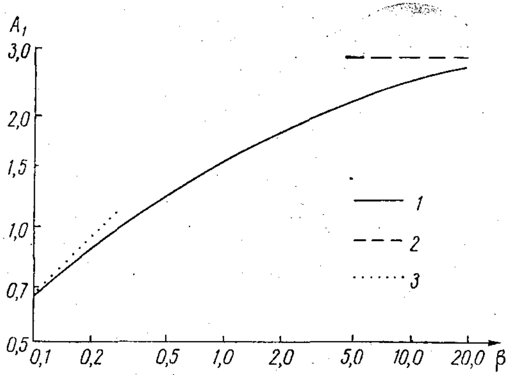

Рис. 1. Коэффициент асимметрии ошибки представления средней одним отсчетом на середине интервала осреднения.

1 — точная зависимость $A_1(\beta)$, 2 — предельное значение $A = \sqrt{8}$ при больших β, 3 — зависимость $A_1(\beta)$ при малых β.

Заметим, что при больших d рост $A_1(d)$ оказывается более медленным и при очень больших d эта величина асимптотически приближается к коэффициенту асимметрии точечных значений y, равному $\sqrt{8}$.

Из вышеизложенного следует, что при достаточно большом числе отсчетов n коэффициент асимметрии ошибки осреднения оказывается близким к 0 и при таких значениях интервала осреднения T, когда асимметрия искомой средней A_F еще существенно отлична от 0. Аналогично должно обстоять дело и с коэффициентом эксцесса и с кумулянтами более высокого порядка. В результате распределение ошибок осреднения быстро приближается к нормальному.

При малом числе отсчетов это распределение может существенно отличаться от нормального. Однако при этом сами ошибки осреднения являются очень большими и сравнимыми по порядку с изменчивостью осредняемой величины. Для многих практических задач такое осреднение большого интереса не представляет.

3. При выполнении оценок методом статистического моделирования реальные непрерывные процессы заменяются формируемыми в ЭВМ их дискретными аналогами. Точность этой имитации (в смысле соответствия статистической структуры ряда заданной модели), число отсчетов и объемы моделируемых выборок могут быть в принципе сколь угодно велики. Практически, однако, они ограничиваются мощностью ЭВМ и имеющимся машинным временем. Будем считать, что интересующий нас процесс моделируется точно и ограничимся оценкой влияния дискретности ряда и объемов выборки на характеристики точности осредненных величин, получаемых методом статистического моделирования.

При расчетах этим методом интегральная средняя F заменяется «эталонной» средней

$$F_э = \frac{\Delta_э}{T} \sum_{i=1}^{n} y\left(t + \Delta_э\left(t - \frac{1}{2}\right)\right), \tag{17}$$

где $\Delta_э$ — минимальный интервал между отсчетами в моделируемом ряде.

Средние арифметические значения $G(t)$ при различной дискретности Δ получаются путем использования прореженного «эталонного» ряда (при этом, конечно, значения Δ могут быть лишь кратными значению $\Delta_э$). В результате обработки моделированных рядов вместо характеристик ошибок представления истинной средней получаются характеристики величины

$$\delta_1(t) = G(t) - F_э(t). \tag{18}$$

Очевидно, что при этом может допускаться погрешность, тем большая, чем ближе Δ к $\Delta_э$. Полезно оценить возможные за этот счет искажения характеристик точности осреднения. Это представляет интерес и с точки зрения обработки эмпирических данных так называемым методом прореживания, при котором точность осреднения дискретных данных оценивается аналогичным образом путем прореживания исходных, возможно более частых, наблюдений. При этом, как правило, принимаемый в качестве эталонного ряд приходится брать через интервалы времени, значительно превышающие интервалы $\Delta_э$, которые можно реализовать на ЭВМ.

При принятых предположениях о временно́й структуре ряда

$$\delta_1(t) = 0. \tag{19}$$

После простых (хотя и громоздких) преобразований можно получить второй и третий моменты распределения δ_1:

$$\sigma_{\delta_1}^2 = \frac{2\sigma^4}{\beta}\left[\frac{d\,(1+p^2)}{1-p^2} - \frac{d_э\,(1+q^2)}{1-q^2}\right] - \\ - \frac{4\sigma^4\,(1-\exp(-2\beta))}{\beta^2}\left[\frac{dp}{1-p^2} - \frac{d_э q}{1-q^2}\right]^2, \tag{20}$$

$$m_{\delta_1} = 8\sigma^6\left[\frac{d^2}{\beta^2} + \frac{2d_э^2}{\beta^2} + \frac{12d_э^2 q^2}{\beta^2\,(1-q^2)^2} - \frac{3d_э d\,(1+q^2)\,(1+p^2)}{\beta^2\,(1-q^2)\,(1-p^2)} + \right. \\ \left. + \frac{6d_э\,(1+q^2)\,(1-\exp(-2\beta))}{\beta\,(1-q^2)}\left(\frac{d_э q}{\beta\,(1-q^2)} - \frac{dp}{\beta\,(1-p^2)}\right)^2\right], \tag{21}$$

где $d_э = \Delta_э/T_0$ — безразмерный интервал между членами эталонного ряда, $q = \exp(-d_э)$.

Анализ формул (20) и (21) показывает, что моменты величины δ_1 систематически меньше моментов истинной ошибки осреднения δ. Это занижение тем больше, чем ближе Δ к $\Delta_э$. По формулам (20) и (21) были выполнены оценки для некоторых сочетаний параметров β, d и $d_э$. Заметим, что при этих расчетах в наиболее интересном диапазоне изменения параметров при значениях d и $d_э$, малых и близких между собой, приходится преодолевать определенные вычислительные трудности, связанные с тем, что входящие в расчет и сопоставляемые между собой величины малы по абсолютной величине и находятся на пределе точности стандартных операций ЭВМ. В самом деле, из (20) и (21) нетрудно получить выражения:

$$\sigma_{\delta_1}^2 \approx \frac{2\sigma^4}{3\beta}\left(d^2 - d_э^2\right)\left[1 - \frac{1-\exp(-2\beta)}{24\beta}\left(d^2 - d_э^2\right) - \frac{d^2 + d_э^2}{15}\right], \tag{22}$$

$$m_{\delta_1} \approx \frac{8d^4}{15\beta^2}\sigma^6\left[1 - \frac{5d_э^2}{d^2} + \frac{4d_э^4}{d^4} + \frac{5\,(1-\exp(-2\beta))}{8\beta}\left(1 - \frac{d_э^2}{d^2}\right)^2\right]. \tag{23}$$

Из этих выражений видно, что при d и $d_э$ порядка 10^{-3} и 10^{-4} величины $\sigma_{\delta_1}^3/\sigma^6$ и m_{δ_1}/σ^6 имеют соответственно порядок 10^{-9} и 10^{-12}. Поэтому расчеты при малых d и $d_э$ оказалось необходимым вести, используя на ЭВМ величины с двойной точностью.

Абсолютные значения ошибок определения σ_δ и A_δ существенно зависят от выбора интервала осреднения T и интервала между отсчетами. Однако соответствующие относительные ошибки от них мало зависят. Это хорошо видно из рис. 2, на котором представлена зависимость относительных ошибок определения σ_δ и A_δ от отношения $\Delta/\Delta_э$ для двух произвольно выбранных сочетаний T и Δ.

Из рис. 2 следует, что для определения коэффициента асимметрии с систематической погрешностью менее 10 % необходимо, чтобы интервал между отсчетами по крайней мере в 5—6 раз превышал интервал отсчетов для эталонной последовательности. Для ценки с той же точностью среднего квадратического отклонения ошибки определения σ_δ достаточно, чтобы эталонная последовательность была гуще оцениваемой в 3 раза. Использование более редких эталонных последовательностей приводит к сильному занижению оцениваемых моментов.

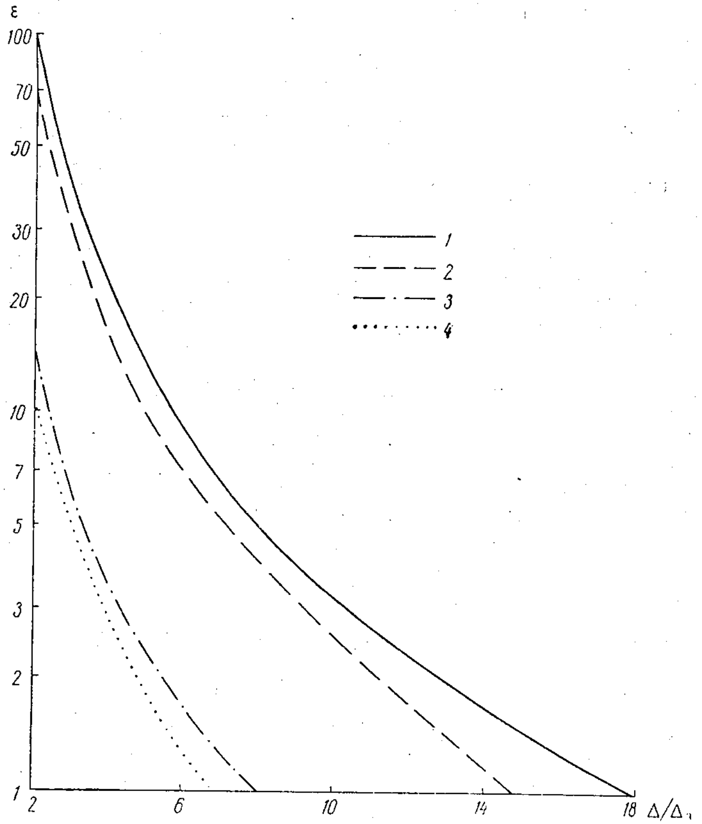

Рис. 2. Относительное значение (в %) характеристик ошибок осреднения в зависимости от степени прореживания ряда.

1) ε_{A_δ} при $T=5T_0$ $(\Delta=T_0)$, 2) ε_{A_δ} при $T=T_0$ $(\Delta=0{,}01T_0)$, 3) $\varepsilon_{\sigma_\delta}$ при $T=5T_0$ $(\Delta=T_0)$, 4) $\varepsilon_{\sigma_\delta}$ при $T=T_0$ $(\Delta=0{,}01T_0)$.

Требование выбора малого интервала между членами эталонного ряда приводит к необходимости моделирования реализаций, состоящих из очень большого числа членов. Наряду с этим приходится учитывать наличие случайных выборочных погрешностей, для уменьшения которых необходимо моделирование достаточно большого числа реализаций процесса. В совокупности эти требования приводят к очень большим объемам выборок, что в сочетании с необходимостью использования сложных процедур моделирования рядов и их обработки может потребовать значительного машинного времени. Так, например, для выполнения опытов по п лучению методом статистического моделирования характеристи точности осреднения процесса $y(t)$ по выборке из 500 реализаци длиной в 1575 членов каждая (такое число членов бралось дл удобства прореживания) оказалось необходимым затрачивать н ЭВМ БЭСМ-6 около 30 мин для каждой дискретности эталонно последовательности.

Тем не менее, как видно из табл. 2, в которой наряду с пол ченными аналитическими оценками σ_δ и A_δ приводятся соответс вующие величины, полученные методом статистического моделиро вания (σ_{δ_1} и A_{δ_1}), видно, что точность последних оказывается вы сокой лишь для σ_{δ_1}. Оценка коэффициента асимметрии ошибо осреднения методом статистического моделирования дает относи тельную погрешность порядка 10 % лишь при больших значе ниях A_δ. При малых значениях A_δ она дает представление лиш о порядке оцениваемой величины. Более точная ее оценка требуе моделирования выборок большего объема.

Таблица

Характеристики ошибок осреднения (в сотых долях единицы) на реализации длины $1575d_э$ при различном числе отсчетов n и при различном шаге эталонной последовательности $d_э = —\ln q$

n	$d/d_э$	q											
		0,99				0,95				0,90			
		σ_{δ_1}	σ_δ	A_{δ_1}	A_δ	σ_{δ_1}	σ_δ	A_{δ_1}	A_δ	σ_{δ_1}	σ_δ	A_{δ_1}	A_δ
1	1575	136	137	218	258	143	141	307	276	133	141	290	28
3	525	755	773	116	123	84	80	137	152	80	81	140	15
5	315	53	52	96	79	61	61	91	112	58	62	104	12
7	225	42	41	43	58	55	51	127	90	56	52	143	10
9	175	34	33	65	45	45	44	83	76	45	46	84	8
15	105	21	21	28	27	33	33	67	51	34	35	72	6
21	75	16	15	17	19	27	27	44	37	29	29	57	5
35	45	9	9	15	12	18	18	28	23	22	21	34	3
45	35	7	7	5	9	16	15	16	18	18	18	30	2

СПИСОК ЛИТЕРАТУРЫ

1. Алексеев Г. А. Методы оценки случайных погрешностей гидрометеоро логической информации.— Л.: Гидрометеоиздат, 1975.— 96 с.
2. Виленкин С. Я. Статистические методы исследования систем автомати ческого регулирования.— М.: Советское радио, 1967.— 200 с.
3. Каган Р. Л. О точности определения средней по площади по данны точечных измерений.— Труды ГГО, 1965, вып. 175, с. 117—131.
4. Каган Р. Л., Сибир Е. Е. О точности осреднения периодически кор релированных величин.— См. наст. сб.
5. Полищук А. И., Каган Р. Л. О статистической структуре осреднен ных значений метеорологических элементов.— Труды ГГО, 1966, вып. 191 с. 92—121.
6. Федорченко Е. И. О влиянии связности метеорологических рядов н точность выборочных моментов.— Труды ГГО, 1974, вып. 336, с. 25—47.

Е. Е. Сибир

ОБ АМПЛИТУДЕ СУТОЧНОГО ХОДА СТАТИСТИЧЕСКИХ ПАРАМЕТРОВ ТЕМПЕРАТУРЫ ВОЗДУХА

Данные об амплитуде суточного хода дисперсии температуры воздуха представляют большой интерес как для теории, так и для практики. Они необходимы, например, при расчете экстремальных величин для получения осредненных по периоду значений и др.

Суточный ход нормы исследован в литературе достаточно широко (см., например, [2]), мы хотели рассмотреть его связь с общей изменчивостью температуры воздуха, характеризующейся ее средним квадратическим отклонением. Сведения о дисперсии температуры воздуха за конкретные сроки содержатся в ряде работ [4, 5, 6]. Представляется интересным получить также данные об амплитуде суточного хода дисперсии температуры воздуха.

Для этой цели нами использовались средние за месяц значения дисперсии температуры воздуха за сроки 01, 07, 13 и 19 ч для 42 станций на территории Советского Союза (рис. 1), полученные Н. А. Лепехиной и Е. И. Федорченко [3] по данным наблюдений за температурой воздуха за период 1936—1965 гг. с учетом годового хода, и средние ежечасные значения температуры воздуха для всех месяцев года для этих же станций, приведенные в Справочниках по климату СССР (для некоторых станций данные о средних ежечасных значениях температуры отсутствуют).

Были вычислены относительные амплитуды суточного хода нормы и дисперсии в виде

$$A_m = \frac{m_{\max} - m_{\min}}{2\bar{\sigma}}, \qquad (1)$$

$$A_D = \frac{D_{\max} - D_{\min}}{2\bar{D}}, \qquad (2)$$

где

$$\bar{D} = \frac{D_{01} + D_{07} + D_{13} + D_{19}}{4}.$$

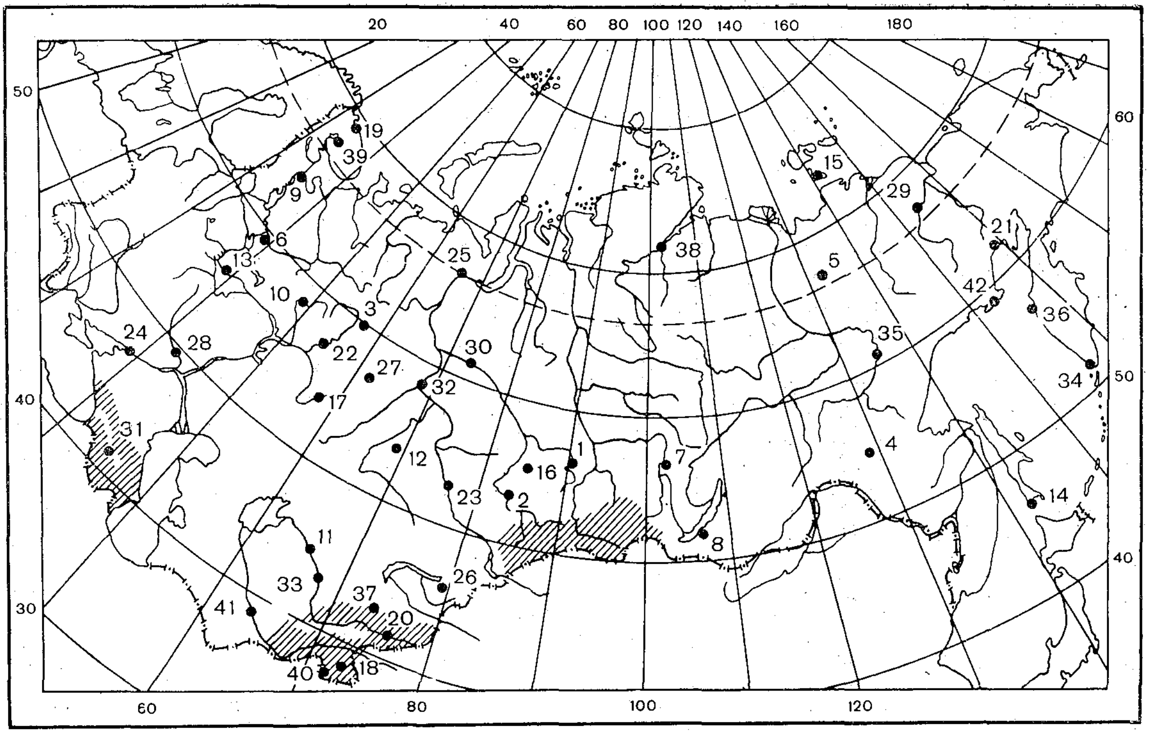

Рис. 1. Схема расположения станций (номера станций такие же, как в табл. 2).

Таблица 1

Относительные амплитуды дисперсии температуры воздуха по ежечасным и четырехсрочным наблюдениям. Ленинград, ГМО

Число сроков	Месяц			
	I	IV	VII	X
24	0,08	0,27	0,40	0,10
4	0,08	0,25	0,35	0,06

Таблица 2

Амплитуды суточного хода нормы и дисперсии температуры воздуха (в %)

№ пп	Станция	Январь		Апрель		Июль		Октябрь	
		A_m	A_D	A_m	A_D	A_m	A_D	A_m	A_D
1	Ачинск	—	18	—	16	—	34	—	20
2	Барнаул	30	26	79	23	130	39	63	35
3	Бисер	13	12	78	17	87	28	27	5
4	Бомнак	68	10	115	21	138	42	92	18
5	Верхоянск	—	—	—	44	—	37	—	13
6	Вологда	8	12	69	10	110	41	42	7
7	Илимск	27	11	119	37	154	41	59	23
8	Кабанск	51	20	113	34	145	59	106	34
9	Кемь	4	6	53	14	52	41	31	10
10	Киров	6	6	66	21	78	32	33	5
11	Кзыл-Орда	39	20	110	31	206	33	135	33
12	Кокчетав	—	14	—	37	—	42	—	31
13	Кольчугино	—	12	—	19	—	35	—	30
14	Корсаков	44	27	74	23	66	14	77	19
15	Котельный	2	1	40	23	27	36	5	2
16	Мариинск	—	17	—	16	—	37	—	23
17	Мелеуз	—	17	—	33	—	39	—	23
18	Мургат	94	14	165	6	180	40	199	7
19	Мурманск	3	1	56	25	58	39	21	8
20	Нарын	60	21	120	42	181	52	173	29
21	Наяхан	—	4	—	37	—	48	—	12
22	Ножовка	—	31	—	13	—	34	—	8
23	Павлодар	—	14	—	39	—	36	—	47
24	Ростов-иа-Дону	17	12	96	38	151	36	77	16
25	Салехард	3	3	41	23	68	24	19	10
26	Сарканд	—	12	—	24	—	27	—	22
27	Свердловск	17	32	74	32	96	40	39	16
28	Серафимович	14	15	74	38	112	34	58	26
29	Среднеколымск	4	10	100	16	64	36	34	15
30	Сургут	7	15	53	19	77	21	23	6
31	Тбилиси	79	16	103	58	162	50	108	27
32	Тобольск	14	8	62	17	90	34	41	12
33	Туркестан	—	17	—	33	—	13	—	29

№ пп	Станция	Январь		Апрель		Июль		Октябрь	
		A_m	A_D	A_m	A_D	A_m	A_D	A_m	A_D
34	Усть-Большерецк	—	20	—	48	—	9	—	33
35	Усть-Мая	—	—	—	49	—	30	—	18
36	Усть-Хайрюзово	—	6	—	40	—	29	—	15
37	Фрунзе	49	3	86	37	154	26	99	38
38	Хатанга	1	2	43	18	58	38	12	1
39	Хибины	4	3	56	43	72	28	16	15
40	Хорог	95	37	137	37	199	19	209	18
41	Чарджоу	6	25	181	42	603	31	175	42
42	Ямск	—	9	—	54	—	51	—	23

Здесь A_m и A_D — амплитуды суточного хода нормы и дисперсии; $m_{\max}$ и $m_{\min}$ — максимальное и минимальное за сутки значения нормы; $D_{\max}$ и $D_{\min}$ — максимальное и минимальное за сутки значения дисперсии; $\bar{D}$ — среднее за сутки значение дисперсии; $\bar{\sigma} = \sqrt{\bar{D}}$ — среднее за сутки значение среднего квадратического отклонения; D_{01}, D_{07}, D_{13} и D_{19} — средние значения дисперсии за 01, 07, 13 и 19 ч.

Амплитуды нормы и дисперсии определялись для всех месяцев года. При этом максимальные и минимальные за сутки значения нормы выбирались из 24 средних ежечасных значений, а дисперсии — из четырех имеющихся у нас срочных значений.

В табл. 1 приведены амплитуды дисперсии для ст. Ленинград, ГМО, полученные при выборе максимальных и минимальных за сутки значений дисперсии из 24 и 4 сроков наблюдений для четырех центральных месяцев сезонов. Как и следовало ожидать, по четырем срокам мы получаем заниженные значения амплитуд, особенно для месяцев с хорошо выраженным суточным ходом дисперсии, но и по четырем срокам мы можем оценить порядок величины амплитуд суточного хода дисперсии.

В табл. 2 представлены амплитуды нормы и дисперсии для четырех центральных месяцев сезонов для 42 указанных станций. Амплитуды нормы для всей территории СССР минимальны зимой и максимальны летом и, как правило, увеличиваются с увеличением широты места. Амплитуды дисперсии также минимальны зимой и увеличиваются весной и летом.

Для северных районов амплитуды нормы и дисперсии зимой близки к нулю, так как в это время года светлая часть суток или совсем отсутствует, или кратковременна. В другие сезоны года амплитуды имеют различный ход в зависимости от местоположения и соответственно условий циркуляции в конкретных районах. Однако почти везде амплитуды возрастают летом и достигают 0,5—0,7 для нормы и 0,3—0,4 для дисперсии (рис. 2 *а*).

Для южных районов амплитуды нормы летом возрастают очень сильно, а амплитуды дисперсии меняются мало в течение года.

Это объясняется большим суточным ходом температуры в этих районах летом и устойчивостью погодных условий (рис. 2 *б*).

При изучении влияния суточного хода температуры иногда бывает удобнее пользоваться не амплитудой суточного хода диспер-

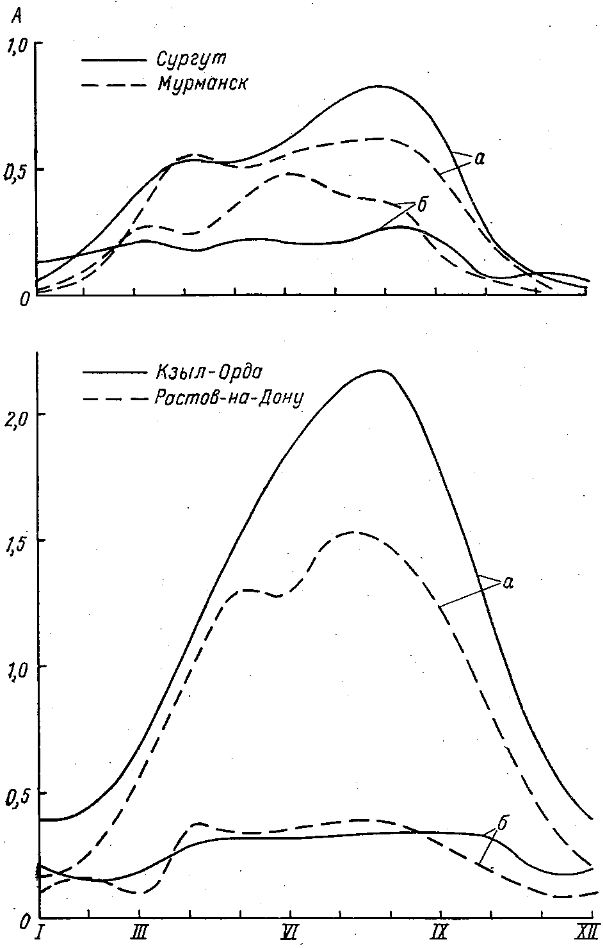

Рис. 2. Годовой ход относительной амплитуды суточного хода нормы (*а*) и дисперсии (*б*) температуры воздуха для четырех станций.

сии A_D, а аналогичной ей величиной $A_\sigma = \frac{\sigma_{\max} - \sigma_{\min}}{2\bar{\sigma}}$ (см. например, [1]), характеризующей изменчивость среднего квадратического отклонения. Оценки показывают, что $A_\sigma \approx \frac{1}{2} A_D$.

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л., Сибир Е. Е. К вопросу о выборе интервала осреднения периодически коррелированных величин.— См. наст. сб.
2. Климат СССР, Вып. 1—6.— Л.: Гидрометеоиздат, 1958—1963.
3. Лепехина Н. А., Федорченко Е. И. О суточном ходе дисперсии температуры воздуха.— Труды ГГО, 1973, вып. 308, с. 99—109.
4. Мамонтов Н. В. Краткая характеристика среднего квадратического отклонения температуры воздуха на территории СССР по ежедневным наблюдениям в отдельные часы суток.— Труды ЗСРНИГМИ, 1975, вып. 16 с. 3—26.
5. Марченко А. С., Помозова Л. И., Чубеико М. А. Временная статистическая структура метеорологических процессов.— Труды НИИАК вып. 54(4), с. 35—46.
6. Тихомирова Л. В. Опыт исследования статистических закономерностей распределения температуры воздуха в отдельные часы суток.— Труды НИИАК, 1969, вып. 51, с. 92—107.

Н. Г. Андронова

СТАТИСТИЧЕСКАЯ ОЦЕНКА ЧУВСТВИТЕЛЬНОСТИ ОДНОЙ МОДЕЛИ АТМОСФЕРНОГО ПЕРЕНОСА К ВОЗМУЩЕНИЯМ ЕЕ ПАРАМЕТРОВ

При работе с моделями атмосферных процессов типа моделей общей циркуляции атмосферы (ОЦА) или моделей фотохимии некоторых атмосферных газов (O_3, NO_x и пр.) возникает задача оценки их внутренней изменчивости от различного рода возмущений. Такая задача сравнительно легко решается для простых моделей, допускающих прямые оценки возможных изменений переменных на выходе модели при вариации тех или иных исходных данных. Однако для многопараметрических, многомерных моделей, в которых могут одновременно действовать различные источники внутренней неустойчивости, определяющие так называемый собственный шум модели, задача существенно усложняется.

Для решения сформулированной проблемы в [2—5] предлагается использовать известный метод проверки статистической значимости гипотез. Такой метод позволяет оценить реакцию сложной модели на введенное определенным способом малое возмущение, давая возможность указать вероятность различения сигнала (возмущения), поданного на вход модели, на ее выходе на фоне собственного шума. Под шумом модели здесь понимается рост и эволюция случайных малых ошибок (СМО), которые оцениваются с помощью определенных статистических приемов при различных реализациях модели.

С целью анализа реакции модели на СМО в рассмотрение вводятся следующие величины:

$$r=\frac{|\Delta|}{\sigma}, \qquad (1)$$

$$\Delta=X-Y, \qquad (2)$$

$$\sigma=\left\{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2\right\}^{1/2}, \qquad (3)$$

$$\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i. \qquad (4)$$

Здесь Y — невозмущенная реализация; X — реализация при введении возмущения в начальные или граничные условия, или при расчете модели с возмущенными параметрами и пр.; x_i — реализации, построенные введением СМО в начальные данные или краевые условия, или параметры модели; N — число временных серий. Величина σ определяет внутреннюю изменчивость (шум) модели, величина Δ — сигнал (возмущение), вводимый в модель.

Предполагается, что моделируются стационарные случайные процессы и что величины, используемые для расчета r, независимы и взяты из нормального распределения с единичной дисперсией. Тогда можно связать отношение r с отношением, определяющим t — распределение Стьюдента с k степенями свободы:

$$t=\frac{(\overline{X}-\overline{Y})}{(s/k)^{1/2}}\left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{1/2}. \qquad (5)$$

Здесь

$$\overline{X}=\frac{1}{N_1}\sum_{i=1}^{N_1} X_i; \quad \overline{Y}_i=\frac{1}{N_2}\sum_{i=1}^{N_2} Y_i;$$

$$s=\sum_{i=1}^{N}(x_i-\overline{x})^2; \quad \overline{x}=\frac{1}{N}\sum_{i=1}^{N} x_i,$$

где Y_i — контрольная реализация; X_i, x_i — возмущенные реализации; N_1, N_2, N — число соответствующих реализаций; $k=N-1$ — число степеней свободы.

Видно, что, если r следует (1), то связь между t и r определится функциональной зависимостью вида

$$t=\sqrt{2}r \ (N_1=1, \ N_2=1, \ N=5).$$

Придавая различные значения величине $r>1$, что является необходимым условием различимости сигнала на фоне шума, и пользуясь таблицами t-распределения для заданного числа степеней свободы, можно получить величины уровней значимости, которые являются мерой недостатка надежности того, что сигнал, посланный в модель, проявляется на ее выходе на фоне собственного шума.

В терминах математической статистики это означает, что, задавая «критические» значения r, мы определим уровни значимости, при которых так называемая нуль-гипотеза, заключающаяся в предположении о статистической значимости разности математических ожиданий теоретической и экспериментальной совокупностей, будет отвергнута, а принята соответственно альтернативная гипотеза.

В настоящей работе предложенный метод использован для оценки «внутренней изменчивости» одномерной численной среднезональной модели переноса и фотодиссоциации N_2O в стратосфере

]. В этой модели поведение концентрации N_2O со временем опи-
вается уравнением

$$\frac{\partial \rho C}{\partial \tau}+\frac{\partial}{\partial z}\left(v_z \rho C - K_z \frac{\partial \rho C}{\partial z}\right)+\Lambda \rho C=0, \qquad (6)$$

е τ — время; z — вертикальная координата; ρ — плотность воз-
уха; C — концентрация N_2O; v_z — скорость вертикального пере-
са; K_z — вертикальный коэффициент макротурбулентной диффу-
и; Λ — параметр стока (распада примеси). Вертикальные про-
или K_z и Λ представлены на рис. 1.

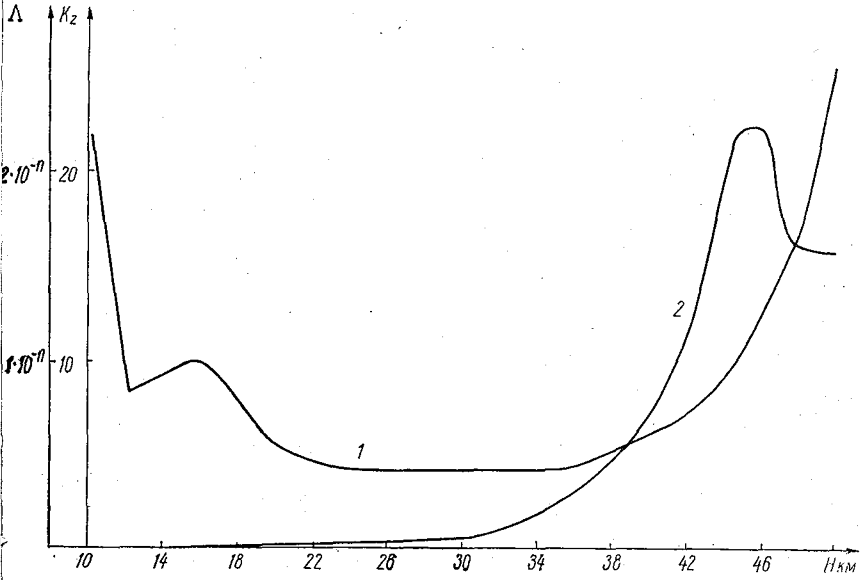

Рис. 1. Профили средних годовых параметров модели.

1 — коэффициент турбулентной диффузии K_z, 2 — параметр стока Λ.

С целью сравнения реакции модели на различного рода возму-
ения уравнение (6) приведено к безразмерному виду введением
арактерных пространственного и временно́го масштабов: $\chi=$
$=10^3$ м — вертикальный масштаб; $T=6{,}37\cdot10^6$ с — масштаб вре-
ени.

Модель реализуется методом сеток с шагом 2 км по высоте и
5 суток по времени и учитывает среднее сезонное изменение вхо-
ящих в нее параметров. Граничные условия на величину концен-
рации N_2O заданы на уровнях 10 и 50 км ($C|_{z=10}=C_0$, $C|_{z=50}=$
$=0$). Начальное распределение концентрации газа задается про-
илем по высоте.

Реализации модели проводились при условии стационарности
ременных серий, что обеспечивалось предположением о постоян-
тве во времени входящих в нее параметров.

Реализация организовывалась следующим образом. В нулев момент времени ($\tau = 0$) рассматривалась модель с постоянны средними годовыми параметрами, рассчитываемая далее до дес того года (240 шагов по времени). Первая (невозмущенная) ре лизация (X_0) была названа контрольной. Последняя реализаци (X_N) использовалась для организации сигнала (Δ), посылаемо в модель. Остальные реализации ($X_1, \ldots, X_{N-1}$) отличались от ко трольной тем, что на последнем шаге первого года в модель ра личными способами вводилось малое возмущение (СМО). Дл удобства проведения эксперимента СМО брались из равномерно распределения на [—1, 1], генерируемого датчиком случайных ч сел (UR), и были различны для различных реализаций. Реализ ции $X_1, \ldots, X_{N-1}$ использовались для оценки внутреннего шум модели.

Всего проводилось две совокупности реализаций, с $N = 7$ $N = 11$.

В работе рассматривались следующие аспекты оценки стат стической значимости задаваемого модельного возмущения:

1) как выбрать способ вычисления шума модели;

2) каким способом организовать сигнал, подаваемый на вх модели;

3) с какого момента времени проводимые реализации можн считать независимыми;

4) каков оптимальный объем выборки рассчитываемых реал заций.

Известно, что рассматриваемая простая диффузионная модел весьма устойчива к вводимым малым возмущения. Поэтому в ц лях отработки и дальнейшего развития методики, предложенно в [2—5], СМО вводились на всех уровнях модели не только в ра считываемую величину концентрации C, но и в параметры моде K_z и Λ, причем вводить СМО в эти параметры необходимо был на протяжении всей реализации.

Были опробованы следующие способы введения СМО:

а) на последнем шаге первого года в распределение концентра ции N_2O по формулам:

$$\widetilde{C} = C\,(1 + 0{,}07\mathrm{UR})$$ (7

и

$$\widetilde{C} = C\,(1 + 0{,}01\mathrm{UR});$$ (8

б) в распределение величины K_z в виде

$$\widetilde{K}_z = K_z\,(1 + 0{,}01\mathrm{UR})$$ (9

и в распределение величины Λ в виде

$$\widetilde{\Lambda} = \Lambda\,(1 + 0{,}01\mathrm{UR})$$ (10

и

$$\widetilde{\Lambda} = \Lambda\left(1 + \mathrm{UR}\Big/\sum_{i=0}^{n}\Lambda_i\right),$$ (11

причем в K_z и Λ возмущение вводилось на всем протяжении времени реализаций, начиная с первого года.

Для организации сигнала (Δ), получаемого на выходе модели, начиная с первого шага нулевого года ($\tau = 0$) на протяжении всей реализации вводились возмущения:

1) в граничные условия путем задания на нижней границе (10 км) искусственного выброса N_2O, имевшего вид:

$$\widehat{C}_0 = 0{,}1C_0, \tag{12}$$

$$\widehat{C}_0 = 0{,}5C_0; \tag{13}$$

2) в профиль параметра K_z в виде:

$$\widehat{K}_z = 0{,}1K_z, \tag{14}$$

$$\widehat{K}_z = 0{,}5K_z; \tag{15}$$

3) в профиль параметра Λ в виде:

$$\widehat{\Lambda} = 0{,}05\Lambda, \tag{16}$$

$$\widehat{\Lambda} = 0{,}1\Lambda. \tag{17}$$

В статистической теории оценка независимости временных серий производится с помощью построения корреляционной функции и вычисления ее уровней значимости. Однако в сложных моделях, как, например, модель ОЦА, такое построение представляет известную трудность и требует большой затраты машинного времени.

В [3] предложен метод, позволяющий судить о степени взаимной зависимости временных реализаций по изменению со временем величины среднего квадратического отклонения мгновенных разностей возмущенной и невозмущенной реализации (Δ_{rms}), которая при предположении стационарности и эргодичности испытаний пропорциональна величине среднего квадратического отклонения, вычисленного по одной реализации.

В настоящей работе аналогично [4] рассматривалась величина δ, вычисляемая по формуле

$$\delta(t) = \left\{ \frac{1}{n} \sum_{i=1}^{n} [X_0^i(t) - X_1^i(t)]^2 \gamma_i \right\}^{1/2}, \tag{18}$$

где n — число уровней, на которых вводилась СМО, X_0 — контрольная реализация, X_1 — возмущенная реализация, γ — весовая функция, которая при равномерной сетке модели равна 1.

График величины δ, названный в [4] кривой предсказуемости, характеризует рост и эволюцию СМО, введенной в модель. Для сложной модели типа модели ОЦА НЦИА (GCM NCAR) в силу ее высокой внутренней изменчивости форма кривой δ указывает на рост СМО до определенного момента времени, после которого график δ выходит на участок горизонтального плато, что соот ветствует установившемуся уровню шума модели.

Кривые предсказуемости, рассчитанные по вышеописанной мо дели для переменной *C* (концентрации N_2O) и параметров *K* (коэффициента турбулентной диффузии) и Λ (коэффициента фото диссоциации газа), приведены на рис. 2. Рассматриваемая модель как уже отмечалось, относится к простейшим диффузионным мо

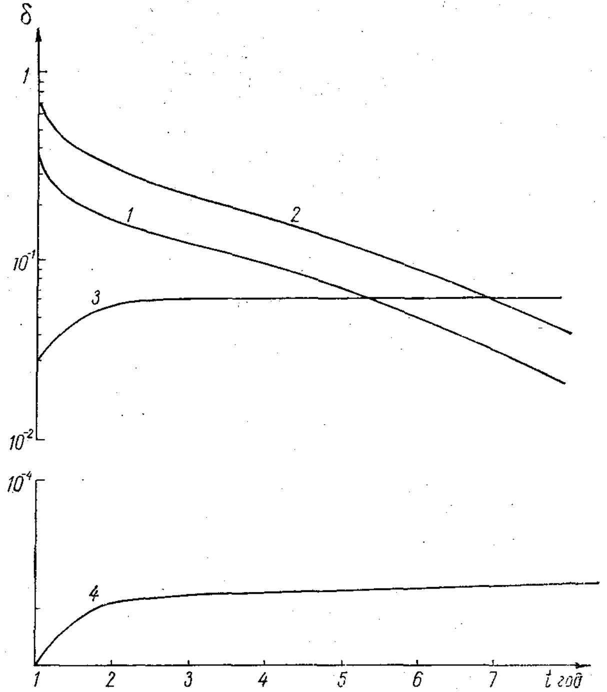

Рис. 2. Кривые предсказуемости δ, полученные для концентрации N_2O и параметров K_z и Λ.

1 — СМО введены в профиль N_2O по (8), *2* — СМО введены в профиль N_2O по (7), *3* — СМО введены в параметр K_z по (9), *4* — СМО введены в параметр Λ по (10).

делям и устойчива к малым возмущениям. Поэтому кривая предсказуемости для переменной $C(\delta^C)$ непрерывно убывает, т. е СМО, введенные в начальное распределение концентрации N_2O сглаживаются с течением времени.

Момент выхода графика функции $\delta^C(\tau)$ на стационарный режим убывания можно считать моментом независимости временных реализаций, начиная с которого эргодическая гипотеза подтверждается. На рис. 2 (кривые *1* и *2*) этот момент соответствует второму году после введения СМО ($\tau = 2$). С целью проверки последнего утверждения была вычислена корреляционная функция временных реализаций X_0 и X_1 с шагом 1 год. О ее поведении мо-

Таблица 1

Корреляционная функция $\mu(\tau)$ временных реализаций X

τ годы	$\mu(\tau)$
1	1
2	$3{,}0\cdot10^{-4}$
3	$1{,}0\cdot10^{-4}$
4	$3{,}7\cdot10^{-5}$
5	$1{,}4\cdot10^{-5}$
6	$5{,}9\cdot10^{-6}$

кно судить по табл. 1, из которой видно, что связь между сече-иями случайных процессов мала, начиная уже с $\tau = 2$, и с этого омента нужно производить осреднение рассматриваемых величин.

Графики «кривых предсказуемости» для параметров K_z и Λ рис. 2) возрастают в течение первого года после введения СМО, затем выходят на горизонтальный участок. Такое поведение кри-ых указывает на то, что модели требуется около года для пере-аспределения концентрации N_2O за счет турбулентного обмена и отодиссоциации, если возмущение введено в параметры, опреде-яющие скорость указанных процессов. Далее модель переходит новое устойчивое состояние. В этом случае момент $\tau = 2$ также ожно считать моментом статистической независимости временных еализаций.

Как указывалось выше, под шумом модели понимается рост и волюция СМО, введенной в модель определенным способом. Шум одели, определяемый величиной σ, оценивался по формуле (3), и ходящие в нее величины осреднялись по годовому промежутку, ачиная со второго года ($\tau = 2$). Результаты расчетов приведены а рис. 3 и 4. Из рис. 3 видно, к каким изменениям в поведении высотой σ^C приводят различные способы задания СМО. При меньшении СМО (задание СМО по (8)) σ^C также убывает (ср. кривые *1* и *2* на рис. 3). Увеличение количества степеней свободы риводит при одинаковых СМО к незначительным изменениям ормы кривых (ср. кривые *2* и *3* на рис. 3). Ход σ^C с высотой мо-кет быть понят, если учесть, что граничные условия сохранялись еизменными на протяжении всей реализации. Поэтому максимум ^{C} достигается на средних уровнях.

Как видно из рис. 3 (кривая *4*), величина среднего квадрати-еского отклонения, полученного введением СМО в параметр $\Lambda(\sigma^\Lambda)$, меньше, чем при введении ее непосредственно в распреде-ение концентрации $N_2O(\sigma^C)$:

$$\sigma^C_{\max}=0{,}15, \quad \sigma^\Lambda_{\max}=1{,}5\cdot10^{-5}.$$

тсюда можно заключить, что внутренняя изменчивость модели тносительно параметра Λ существенно ниже, чем относительно онцентрации N_2O.

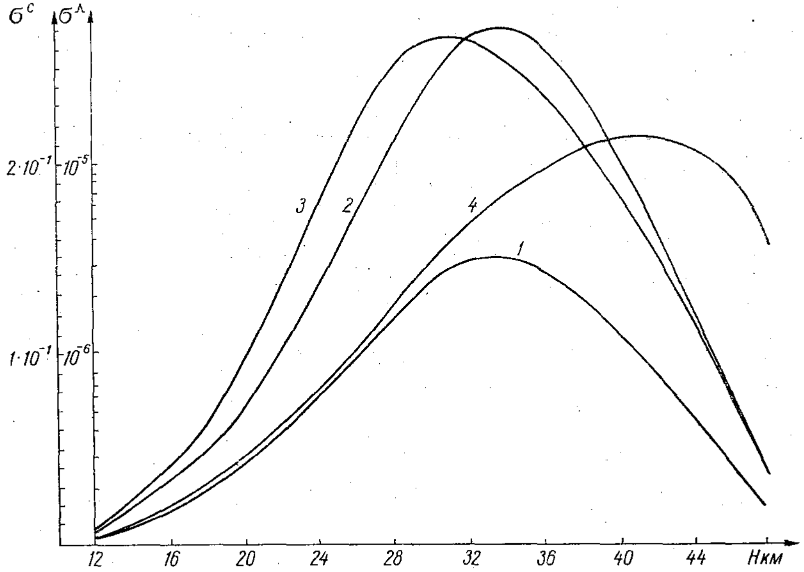

Рис. 3. Распределение по высотам величин σ^C и σ^Λ.

1 — СМО введены в профиль N_2O по (8), *2* — СМО введены в профиль N_2O по (7), *3* — σ^C, рассчитанная для 9 степеней свободы, *4* — СМО введены в профиль Λ по (10).

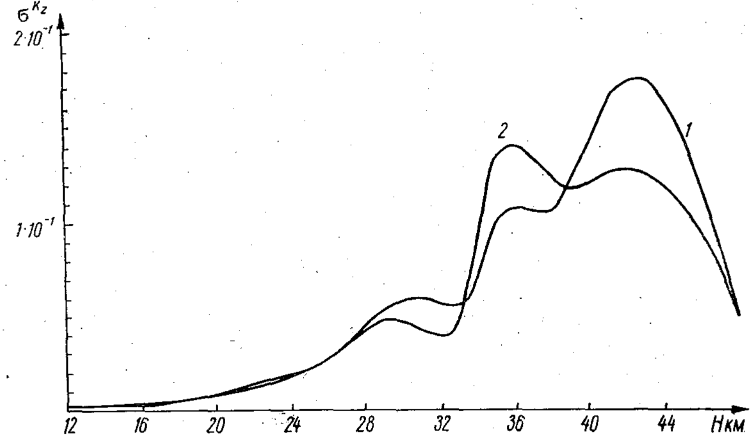

Рис. 4. Распределение по высоте величины σ^{K_z}.

1 — по расчетам для 4 степеней свободы, *2* — по расчетам для 9 степеней свободы.

Величина σ^{Λ} на нижних уровнях мала, поскольку Λ (рис. 1, кривая *2*) на этих уровнях близка к нулю. Далее σ^{Λ} монотонно возрастает (в силу малой изменчивости Λ), достигая максимума около 40 км, а затем уменьшается. Для понимания этого факта необходимо учесть взаимосвязь процессов фотодиссоциации и турбулентной диффузии: выше 40 км коэффициент K_z (рис. 1, кривая *1*) резко возрастает (усиливается турбулентный перенос) и таким образом ограничивается роль фотодиссоциации в установлении профиля концентрации.

Величина σ^{Λ} рассчитана для 9 степеней свободы. Задание СМО по формуле (11) приводит к нереально большим вариациям Λ, и в результате модель выходит из равновесного состояния.

Величина σ^{K_z} (рис. 4) остается малой до уровня $H = 24$ км, так как на этих высотах интенсивность турбулентного перемешивания велика, и профиль концентрации быстро устанавливается при введении СМО в K_z. На уровнях от 24 до 42 км профиль K_z практически постоянен и имеет минимальное значение (рис. 2, кривая *1*), и на этом участке наблюдается эффект неравномерного накопления ошибок, связанный, по-видимому, с неравномерным заданием СМО, которые генерирует датчик случайных чисел UR. Выше 42 км значение K_z резко возрастает и σ^{K_z} соответственно уменьшается.

Изменчивость модели относительно параметра $K_z(\sigma^{K_z}_{\max} = 0{,}14$ для 4 степеней свободы и $\sigma^{K_z}_{\max} = 0{,}17$ для 9 степеней свободы) сравнима с изменчивостью относительно величины концентрации N_2O.

Изменение числа степеней свободы при расчете величины σ^{K_z} приводит к заметным изменениям ее на тех уровнях, где происходит неравномерное накопление ошибок, однако в целом характер поведения профилей подобен (рис. 4, кривые *1* и *2*).

Экспериментальный коэффициент r был вычислен по формуле (1). Графики распределения этого коэффициента по высоте для различных экспериментов представлены на рис. 5 и 6. Из этих рисунков следует:

1. Поведение коэффициента r, образованного введением возмущения в C, K_z и Λ, воспроизводится моделью принципиально различным образом. При организации сигнала введением возмущения в нижнее граничное условие ($H = 10$ км) значение r^C (рис. 5, кривые *1*—*3*) монотонно убывает с высотой. Смысл такого поведения состоит, видимо, в том, что сигнал передается в модели снизу вверх турбулентной диффузией, постепенно затухая. С увеличением числа степеней свободы существенных изменений в поведении коэффициента r^C не наблюдается (ср. кривые *1* и *3* на рис. 5).

Кривая изменения r^{Λ} с высотой (рис. 5) при введении возмущения (16) и (17) в величину Λ была получена для 9 степеней свободы. С увеличением сигнала с 5 до 10 % коэффициент r^{Λ}

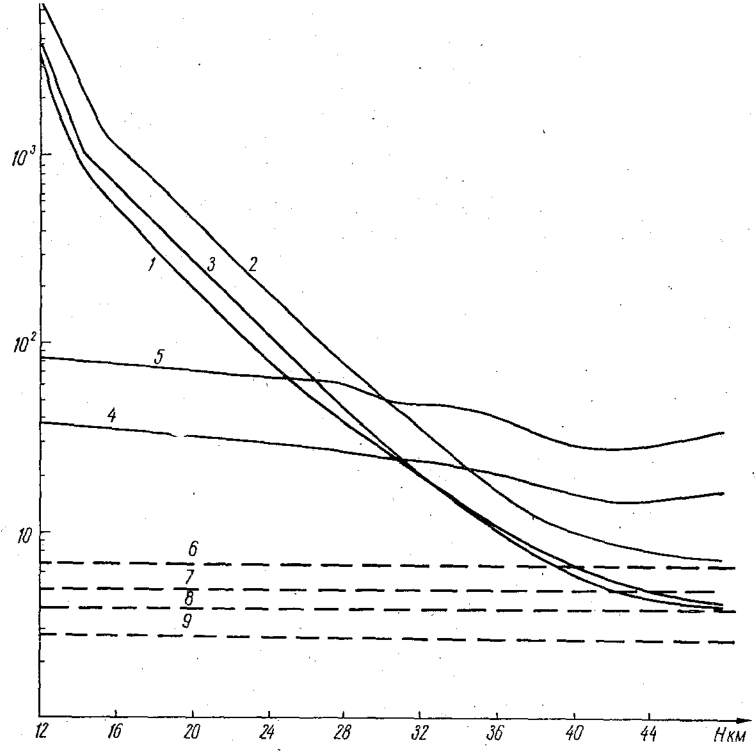

Рис. 5. Распределение по высоте величин r^{C} и r^{Λ}.

1 — возмущение граничного условия по (12), *2* — то же по (13), *3* — то же для 9 степеней свободы; *4* — возмущение параметра Λ по (16), *5* — то же (по 17); *6* — 99 %-ная вероятность проявления сигнала при 4 степенях свободы, *7* — то же при 9 степенях свободы, *8* — 95 %-ная вероятность проявления сигнала при 4 степенях свободы, *9* — то же при 9 степенях свободы.

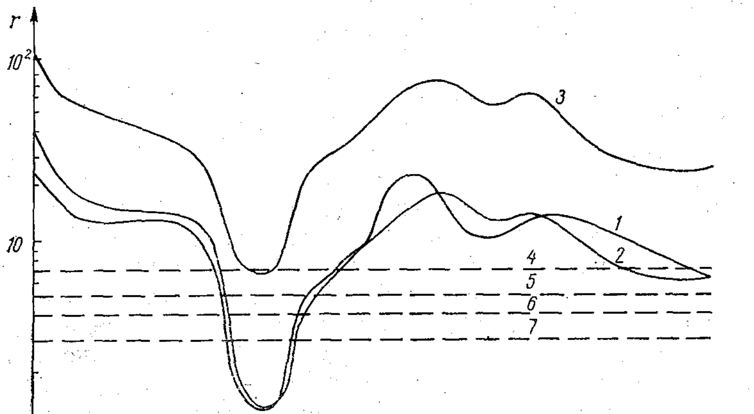

Рис. 6. Распределение по высоте коэффициента r^{Kz}.

1 — возмущение введено по (14), *2* — то же для 9 степеней свободы, *3* — то же по (15), *4* — 99 %-ная вероятность проявления сигнала при 4 степенях свободы, *5* — то же при 9 степенях свободы, *6* — 95 %-ная вероятность проявления сигнала при 4 степенях свободы, *7* — то же при 9 степенях свободы.

(ср. кривые *4* и *5* на рис. 5) на всех высотах соответственно возрастает. Наблюдается незначительное уменьшение коэффициента r^{Δ} в сторону больших высот, связанное с возрастанием интенсивности распада N_2O с высотой.

Профиль коэффициента r^{K_z} при введении возмущения в параметр K_z (рис. 6) имеет резкий минимум на уровне 24 км, где значение K_z достигает минимума. Поведение коэффициента r^{K_z} выше 24 км обусловливается низкой скоростью турбулентного перемешивания (что приводит к нарастанию введенного возмущения) и характером изменения с высотой σ^{K_z}.

Увеличение числа степеней свободы приводит к незначительным изменениям r^{K_z} (ср. кривые *1* и *2* на рис. 6), а увеличение сигнала ведет к практически одинаковому росту r^{K_z} на всех высотах (рис. 6, кривые *1* и *3*).

2. На рис. 5 (кривые *6—9*) и на рис. 6 (кривые *4—7*) указаны значения r, соответствующие 95 и 99 %-ной вероятности того, что сигнал, посланный в модель, проявится на фоне шума. Эти значения получены из таблицы уровней значимости (табл. 2), построенной на основании t-распределения Стьюдента для четырех и девяти степеней свободы. Они указывают, с какой вероятностью при заданном числе степеней свободы и соответствующих значениях r принятая 0-гипотеза отвергается. Например, сигнал, посланный в модель, проявится на фоне шума с 95 %-ной вероятностью, начиная с $r = 4$ как при четырех, так и при девяти степенях свободы. Для 99 %-ной вероятности сигнал будет значим, начиная с $r = 7$ для четырех степеней свободы и с $r = 5$ для девяти степеней свободы.

Таблица 2

Значимые уровни (%) для $t = \sqrt{2}\, r$ при разном числе степеней свободы

r	Число степеней свободы	
	4	9
1	51,69	49,57
2	22,86	18,93
3	10,02	6,20
4	4,69	1,94
5	2,38	0,62
6	1,31	0,21
7	0,77	0,08
8	0,47	0,03
9	0,31	0,01

Из рис. 5 видно, что возмущение нижнего граничного условия проявляется с высокой надежностью практически на всех уровнях.

Так, при 4 степенях свободы сигнал до уровня 38 км проявляется с вероятностью выше 99 %, при 9 степенях свободы — до 44 км и в обоих случаях (4 и 9 степеней свободы) проявляется на всех уровнях с вероятностью выше 95 %.

При рассмотренных возмущениях параметра Λ (5 и 10 % от величины Λ) сигнал статистически значимо проявляется на всех уровнях модели, поскольку шум модели относительно параметра Λ чрезвычайно низок.

Сигнал, получаемый 50 %-ным возмущением коэффициента K_z для 9 степеней свободы, оказывается также статистически значимым на всех уровнях модели. Вероятности проявления сигнала при 10 %-ном возмущении K_z приведены в табл. 3.

Таблица 3

Высоты H статистической значимости сигнала при возмущении параметра K_z

Число степеней свободы	Доверительная версятность, %	H км	
		ниже	выше
4	95	22	26
	99	22	28
9	95	23	26
	99	22	27

Из таблицы видно, что при увеличении числа степеней свободы сигнал становится значимым с заданной вероятностью для большего числа уровней.

В заключение можно сделать следующие выводы:

1) скорость роста и процесс эволюции СМО различны для различных рассмотренных переменных;

2) осреднение рассматриваемых величин нужно производить с момента приблизительной статистической независимости временны́х реализаций, начинающихся из одного и того же начального состояния, т. е. с момента, когда модель приспособится к заданному возмущению и достигнет устойчивого состояния;

3) необходимо тщательное планирование проводимых экспериментов;

4) для оценки уровней значимости модели желательно максимально возможное расширение объема выборки (увеличение числа степеней свободы).

СПИСОК ЛИТЕРАТУРЫ

1. Кароль И. Л., Киселев А. А. Модельные расчеты сезонного содержания соединений азота в стратосфере умеренных широт.— Труды ГГО, 1977, вып. 394, с. 56—65.

2. Chervin R. M. Variability of GCM climate simulation and statistical significance in GCM climate change experiments.— In: Report at Joint DMG/AMS Int. Conf. on Simulation of Large Scale Atm. Processes. Hamburg, F. R. Germany, 30 Aug.— 4 Sept. 1976.
3. Chervin R. M., Schneider S. H. A study of the response of NCAR GCM climatological statistics of random perturbations: estimating noise levels.— J. Atmos. Sci., 1976, vol. 33, N 3, p. 391—404.
4. Chervin R. M., Schneider S. H. On determining the statistical significance of climate experiments with General Circulation Models.— J. Atmos. Sci., 1976, vol. 33, N 3, p. 405—412.
5. Chervin R. M., Washington W. M., Schneider S. H. Testing the statistical significance of the response of the NCAR General Circulation Model to North Pacific Ocean surface temperature anomalies.— J. Atmos. Sci., 1976, vol. 33, N 3, p. 413—423.

Б. Т. Курбанов

О ПРОСТРАНСТВЕННОЙ СТАТИСТИЧЕСКОЙ СТРУКТУРЕ ТЕМПЕРАТУРЫ И ВЫСОТЫ ТРОПОПАУЗЫ

Исследование пространственной статистической структуры метеорологических полей представляет интерес с точки зрения решения целого ряда практических задач, например задачи объективного анализа этих полей. К настоящему времени выполнено большое количество исследований статистической структуры метеополей, прежде всего температуры и геопотенциала изобарических поверхностей как тропосферных, так и стратосферных.

Однако до настоящего времени неизвестны сведения о пространственной структуре тропопаузы. Между тем статистические сведения о высоте и температуре тропопаузы необходимы прежде всего для нужд современной авиации. Известно, например, что переход на сверхзвуковую скорость целесообразно производить именно на уровне тропопаузы.

Кроме того, в некоторых прогностических моделях (см., например, [11]) поверхность тропопаузы выбирается как координатная поверхность.

Целью настоящей работы является расчет пространственных корреляционных функций высоты и температуры полярной тропопаузы.

Исходным материалом для расчета служили данные радиозондирования над Европейской территорией СССР за 1970—1975 гг. Расчеты производились по данным за серединный месяц каждого сезона (январь, апрель, июль и октябрь). Для уменьшения статистической связи между ситуациями они выбирались отстоящими друг от друга на трое суток. Таким образом, в каждом месяце, за исключением апреля, для расчета использовалось 66 ситуаций (в апреле — 60). Для расчета пространственной корреляционной функции температуры были использованы данные 33 радиозондирующих станций, для расчета же корреляционной функции высоты тропопаузы — данные 37 станций.

На рис. 1 показаны станции, данные с которых были использованы. Кружками обведены четыре станции, данные с которых не использовались при расчете корреляционной функции температуры. Тропопауза рассматривалась как полярная, если ее высота не превышала 14 км. Если над одним пунктом одновременно наблюдались несколько тропопауз, к расчетам привлекались сведения о той тропопаузе, которая подтверждалась данными о тропопаузе на соседних станциях. В тех случаях, когда этим путем не удавалось уверенно выявить «истинную» тропопаузу, данные этого пункта к расчетам не привлекались.

Корреляционная функция рассчитывалась по программе, описанной в [4]. По этой программе производится расчет корреляционных моментов для каждой пары станций. Затем корреляционные моменты осредняются в рамках выбранной градации расстояний. Эти средние значения относятся к середине градации. Кроме значений корреляционной функции, выдаются на печать средние значения за рассматриваемый период и средние квадратические отклонения σ для каждого пункта и осредненные по всей территории.

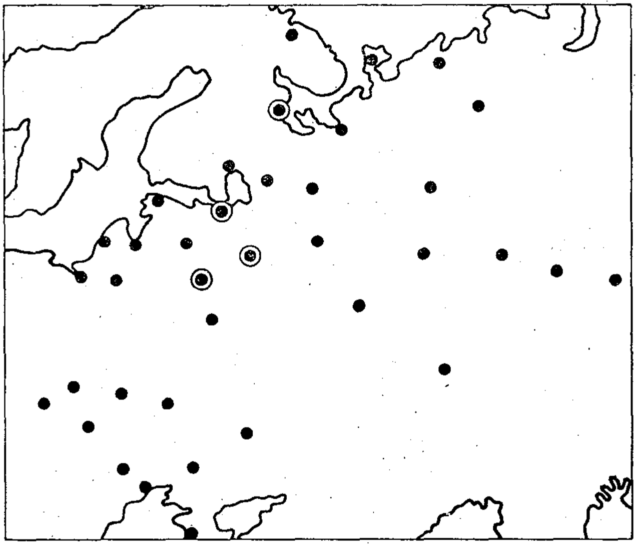

Рис. 1. Расположение станций, данные которых использовались для расчетов.

Для выяснения однородности полей температуры и высоты (геопотенциала) тропопаузы были построены карты средних квадратических отклонений температуры и геопотенциала для каждого месяца.

Анализ показал, что поле средних квадратических отклонений температуры σ_T наиболее однородно в июле. В январе и октябре значения σ_T несколько повышаются к западу и понижаются к северу рассматриваемой территории. В апреле наблюдается повышение σ_T в центральной и восточной частях, а понижение — на севере и юге.

Анализ полей средних квадратических отклонений высоты тропопаузы σ_h показал, что это поле, так же как и поле σ_T наиболее однородно в июле. В остальные месяцы величина σ_h плавно меняется по территории, несколько понижаясь к юго-западу. В апреле и июле σ_h повышается к востоку, в октябре — к северу.

В табл. 1 приведены средние, максимальные и минимальные значения средних квадратических отклонений температуры и высоты для рассматриваемой территории. Из таблицы видно, что наибольшие $\sigma_{T\text{ср}}$ наблюдаются в январе, а наименьшие — в июле, что объясняется циркуляционными особенностями сезонов.

Таблица 1

Значения средних квадратических отклонений температуры и высоты тропопаузы, полученные автором и в работах [1] и [8]

Месяц	σ_T °C			σ_h гп. дам			[8]		[1]
	ср	макс	мин	ср	макс	мин	σ_T °C	σ_h гп. дам	σ_T °C
I	5,9	7,3	4,5	116	139	86	5,8	113	5,5
IV	5,2	6,8	4,0	110	133	79	—	—	4,9
VII	4,5	6,2	3,7	95	115	73	4,5	80	3,8
X	5,5	6,6	4,0	130	167	98	—	—	4,2

Анализ табл. 1 показывает, что наименьшие значения $\sigma_{h\text{ср}}$, так же как и $\sigma_{T\text{ср}}$, наблюдаются в июле. Обращает на себя внимание тот факт, что наибольшие $\sigma_{h\text{ср}}$ наблюдаются в октябре, а не в январе. По-видимому, это объясняется тем, что для широт, охватывающих большую часть рассматриваемой территории, высота тропопаузы от лета к зиме понижается быстро и практически линейно. Об этом свидетельствуют данные З. М. Маховера, приводимые в [6] и [7].

Одной из причин больших значений $\sigma_{h\text{ср}}$ в октябре является длина периода осреднения (месяц), в течение которого высота тропопаузы постоянно понижается. Для других серединных месяцев подобное явление не наблюдается или выражено слабо.

В табл. 1 приведены также средние квадратические отклонения температуры и высоты тропопаузы, полученные З. М. Маховером [8] по данным за два года для Свердловска. Как следует из таблицы, наши данные хорошо согласуются с данными З. М. Маховера.

В табл. 1 для сравнения приведены также средние квадратические отклонения температуры поверхности 200 мбар, полученные В. П. Болтенковым [1]. На этом уровне σ_T достигает своего максимального значения в свободной атмосфере. Сравнение σ_T тропопаузы и изобарической поверхности 200 мбар показывает, что

σ_T тропопаузы больше максимального σ_T изобарических поверхностей в тропосфере и нижней стратосфере. По данным Л. С. Гандина и Т. И. Кузнецовой [3], для зимы значения σ_h высот изобарических поверхностей 300 и 200 мбар, между которыми чаще всего расположена тропопауза, составили соответственно 19,0 и 18,2 дам, а по данным С. М. Олевской [9], для этих же уровней — соответственно 18,0 и 16,7 дам.

Сравнение данных табл. 1 с вышеприведенными показало, что σ_h высоты тропопаузы значительно (на порядок) больше σ_h высоты изобарических поверхностей. Таким образом, подтверждается вывод З. М. Маховера [8] о том, что изменения высоты тропопаузы превосходят изменения высоты изобарических поверхностей, а изменения температуры достигают по величине максимальной изменчивости температуры изобарических поверхностей.

Для проверки гипотезы об однородности и изотропности полей температуры и высоты (геопотенциала) тропопаузы были построены карты изокоррелят для ст. Вологда, расположенной примерно в центре выбранной территории. Анализ карт показал, что поле температуры до расстояний примерно 1000 км достаточно изотропно. Анизотропность поля высоты выражена слабо в октябре и апреле. В целом поля высот тропопаузы можно принять изотропными до расстояний примерно 700—900 км.

Ограниченный объем данных, используемый при выполнении данной работы, не позволяет в должной мере исследовать, насколько оправдано применение гипотезы об однородности и изотропности полей температуры и высоты тропопаузы. Поэтому выводы об однородности и изотропности следует считать ориентировочными.

В табл. 2 приведены значения корреляционной функции температуры $m(\rho)$, а в табл. 3 — аналогичные значения корреляционной функции высоты тропопаузы. В этих таблицах даны средние значения расстояний ρ и число пар станций N, попадающих в заданную градацию. Значения коэффициентов корреляции закономерно уменьшаются с расстоянием и хорошо аппроксимируются аналитическими функциями.

Для температуры:

$$m(\rho) = m(0)\, e^{-1,22\rho^{1,51}}$$

зимой,

$$m(\rho) = m(0)\, e^{-1,425\rho^{1,522}}$$

весной,

$$m(\rho) = m(0)\, e^{-2,01\rho^{1,599}}$$

летом,

$$m(\rho) = m(0)\, e^{-0,807\rho^{1,255}} \cos(1,227\rho)$$

осенью.

Таблица 2

Значения корреляционной функции температуры тропопаузы по месяцам

$\rho \cdot 10^{-3}$ км	N	Месяц			
		I	IV	VII	X
0,20	12	0,895	0,817	0,756	0,826
0,30	18	0,797	0,756	0,688	0.719
0,40	20	0,715	0,663	0,564	0,681
0,50	15	0,654	0,531	0,431	0,567
0,60	27	0,535	0,504	0,370	0,475
0,70	25	0,487	0,465	0,260	0,385
0,80	32	0,395	0,360	0,181	0,262
0,925	42	0,320	0,278	0,166	0,226
1,10	53	0,254	0,165	0,101	0,086
1,30	51	0,140	0,135	0,078	—0,007
1,50	54	0,098	0,065	0,085	—0,087
1,70	48	0,038	0,033	0,039	—0,144
1,90	31	0,032	0,008	—0,023	—0,181
2,25	70	—0,024	0,042	0,028	—0,128
2,75	30	—0,016	0,069	0,048	—0,105

Таблица 3

Значения корреляционной функции высоты тропопаузы по месяцам

$\rho \cdot 10^{-3}$ км	N	Месяц			
		I	IV	VII	X
0,20	15	0,790	0,776	0,732	0,810
0,30	26	0,757	0,747	0,696	0,730
0,40	28	0,677	0,682	0,610	0,683
0,50	23	0,633	0,579	0,514	0,593
0,60	37	0,553	0,534	0,409	0,526
0,70	32	0,482	0,467	0,302	0,468
0,80	41	0,449	0,397	0,265	0,343
0,90	30	0,341	0,328	0,196	0,304
1,00	38	0,310	0,249	0,167	0,251
1,20	118	0,196	0,151	0,084	0,127
1,45	61	0,078	0,051	0,019	0,004
1,65	61	—0,012	—0,008	0,018	—0,066
1,85	44	—0,070	—0,078	—0,026	—0,114
2,10	52	—0,115	—0,100	—0,021	—0,151
2,40	37	—0,112	—0,103	—0,021	—0,174
2,80	23	—0,159	—0,082	—0,042	—0,167

Для высоты:

$$m(\rho)=m(0)\,e^{-0{,}496\rho^{0{,}95}}\cos(0{,}970\rho)$$

зимой,

$$m(\rho)=m(0)\,e^{-0{,}605\rho^{1{,}15}}\cos(0{,}982\rho)$$

весной,

$$m(\rho)=m(0)\,e^{-1{,}64\rho^{1{,}63}}$$

летом,

$$m(\rho)=m(0)\,e^{-0{,}582\rho^{1{,}065}}\cos(1{,}061\rho)$$

осенью.

Значения $m(0)$ получены путем интерполяции корреляционных функций на нуль и приведены в табл. 4.

Таблица 4

Некоторые параметры корреляционных функций температуры и высоты (геопотенциала) тропопаузы и изобарических поверхностей 300 и 200 мбар

Месяц	$m(0)$	$\rho_0\cdot10^3$ км	η^2	δ °C	[1, 2]	
					$\rho_{0,300}\cdot10^3$ км	$\rho_{0,200}\cdot10^3$ км
			Температура			
I	0,96	0,85	0,04	1,2	1,30	1,18
IV	0,94	0,75	0,06	1,2	—	—
VII	0,90	0,60	0,11	1,4	1,06	0,72
X	0,96	0,71	0,04	1,0	—	—

Месяц	$m(0)$	$\rho\cdot10^3$ км	η^2	δ гп. дам	[10]		[5]	
					$\rho_{0,300}\cdot10^3$ км	$\rho_{0,200}\cdot10^3$ км	$\rho_{0,300}\cdot10^3$ км	$\rho_{0,200}\cdot10^3$ км
			Высота (геопотенциал)					
I	0,90	0,90	0,11	37	1,20	1,36	1,24	1,40
IV	0,89	0,84	0,12	36	—	—	—	—
VII	0,86	0,64	0,16	36	1,34	1,62	—	—
X	0,92	0,82	0,09	37	—	—	—	—

По данным значений корреляционных функций температуры и высоты построены графики нормированных корреляционных функций для каждого месяца.

На рис. 2 приведены графики нормированных корреляционных функций температуры тропопаузы. Как следует из рисунка, наиболее быстро затухание корреляционной функции наблюдается в июле. Корреляционные функции за апрель и октябрь занимают промежуточное положение между январскими и июльскими функциями. Совершенно аналогично ведут себя и корреляционные функции высоты тропопаузы (рис. 3). Ослабление корреляционных связей в июле по сравнению с январем объясняется, по-видимому, сезонными изменениями масштаба циркуляции.

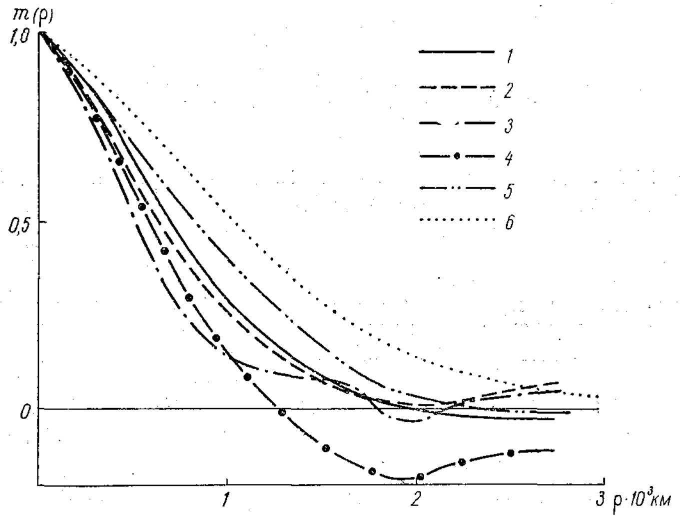

Рис. 2. Нормированные корреляционные функции температуры тропопаузы и изобарической поверхности 200 мбар.

1 — для января, 2 — для апреля, 3 — для июля, 4 — для октября, 5 — по данным [9] для лета, 6 — по данным [6] для зимы.

На рис. 2 приведены также корреляционные функции температуры изобарической поверхности 300 мбар за зимний и летний сезоны по данным В. П. Болтенкова [1, 2], а на рис. 3 приведены корреляционные функции высоты поверхности 200 мбар по данным М. С. Татарской [10] и В. М. Мартемьянова [5]. Обращает на себя внимание более быстрое затухание корреляционной функции температуры и высоты тропопаузы по сравнению с соответствующими корреляционными функциями изобарических поверхностей.

Как следствие указанного выше факта, радиусы корреляции ρ_0 (расстояние, на котором корреляция убывает в e раз), приводимые в табл. 4, оказались значительно меньше радиусов корреляции соответствующих функций изобарических поверхностей 300 и 200 мбар, рассчитанных по данным В. П. Болтенкова [6, 9], М. С. Татарской [10] и В. И. Мартемьянова [11].

Переход через нуль корреляционной функции высоты изобарических поверхностей 300 и 200 мбар зимой по данным работ [2], [3], [5] и [9] происходит на расстояниях соответственно 2,16 и 2,55; 2,08, и 2,26; 2,06 и 2,16; 2,05 и 2,40 тыс. км. По данным настоящей работы, переход корреляционной функции высоты тропопаузы через нуль во все сезоны происходит значительно раньше (1,5—1,7 тыс. км).

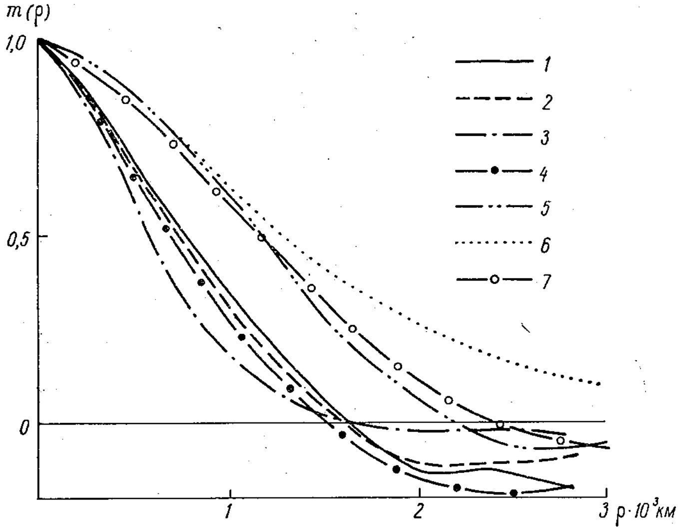

Рис. 3. Нормированные корреляционные функции высоты тропопаузы и геопотенциала изобарической поверхности 300 мбар.

1 — для января, 2 — для апреля, 3 — для июля, 4 — для октября, 5 — по данным [10] для зимы, 6 — по данным [10] для лета, 7 — по данным [11] для зимы.

В табл. 4 приведены и некоторые другие параметры корреляционной функции температуры и высоты тропопаузы. Мера ошибок наблюдений $\eta^2 = \frac{1 - m(0)}{m(0)}$, характеризующая случайные ошибки наблюдений, несколько меняется от месяца к месяцу, достигая максимума в июле и минимума в октябре, как для температуры, так и для высоты изобарических поверхностей. Наряду с величинами η^2 приведены также средние квадратические значения случайных ошибок δ, рассчитанные по известной формуле

$$\delta = \sigma \sqrt{1 - m(0)}\,.$$

В заключение выражаю глубокую признательность Л. В. Руховцу и К. М. Лугиной за ценные советы и внимание к работе.

СПИСОК ЛИТЕРАТУРЫ

1. Болтенков В. П. Исследование статистической макроструктуры температуры воздуха.— Труды ГГО, 1964, вып. 165, с. 16—26.
2. Болтенков В. П. Некоторые характеристики трехмерной макроструктуры температуры воздуха.— Труды ГГО, 1966, вып. 191, с. 47—57.
3. Гандин Л. С., Кузнецова Т. И. О пространственной статистической структуре поля геопотенциала.— Труды ГГО, 1965, вып. 168, с. 84—93.
4. Лугина К. М., Тараканова В. П. Пространственная структура поля средней месячной температуры.— В кн.: Исследования статистической структуры метеорологических полей, т. 1. М.: Гидрометеоиздат, 1975, с. 73—79.
5. Мартемьянов В. И. О статистической структуре поля геопотенциала и давления у земли.— Труды САНИГМИ, 1969, вып. 40 (55), с. 75—88.
6. Маховер З. М. Общие закономерности распределения тропопаузы над северным полушарием.— Труды ВНИИГМИ—МЦД, 1975, вып. 10, с. 3—12.
7. Маховер З. М. О распределении высот тропопаузы над земным шаром.— Труды ВНИИГМИ—МЦД, 1976, вып. 34, с. 3—15.
8. Маховер З. М. Временна́я изменчивость высоты и температуры тропопаузы.— В кн.: Применение статистических методов в метеорологии. Л.: Гидрометеоиздат, 1971, с. 153—163.
9. Олевская С. М. Пространственно-временна́я структура полей геопотенциала.— Труды Гидрометцентра СССР. 1969, вып. 39, с. 59—66.
10. Татарская М. С. Пространственная статистическая структура геопотенциала.— Изв. АН СССР. Физика атмосферы и океана, 1965, т. 1, № 7, с. 760—763.
11. Шуман Ф. Г. Многоуровенная модель по полным уравнениям.— В кн.: лекции по численным методам краткосрочного прогноза погоды. Л.: Гидрометеоиздат, 1969, с. 481—497.

СОДЕРЖАНИЕ

Труды ГГО, вып. 431

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В МЕТЕОРОЛОГИИ

Редактор Л. И. Штанникова. Техн. редактор М. И. Брайнина. Корректор И. А. Крайнева

Сдано в набор 13.06.79. Подписано в печать 25.01.80. М-23224. Формат 60×90¹/₁₆. Бум. тип. № 1.
Лит. гарн. Печать высокая. Печ. л. 10. Уч.-изд. л. 9,89. Тираж 710 экз. Индекс МЛ-221.
Заказ № 473. Цена 70 коп. Заказное. Гидрометеоиздат. 199053. Ленинград, 2-я линия, д. 23.

Ленинградская типография № 8 ЛПО «Техническая книга» Союзполиграфпрома при
Государственном комитете СССР по делам издательств, полиграфии и книжной торговли.
190000, Ленинград, Прачечный пер., 6

ДК 551.509.1

О вертикальной статистической структуре относительного геопотенциала изо-арических поверхностей. Гандин Л. С., Тараканова В. П. Труды ГГО, 980, вып. 431, с. 3—11.

Разработана методика расчета характеристик вертикальной статистической труктуры относительного геопотенциала по данным о статистической структуре емпературы. Выполнены расчеты по разным выборкам для различных сезонов широтных зон. Проанализированы основные закономерности вертикальной труктуры относительного геопотенциала.

Предложен способ расчета вертикальной структуры абсолютного геопотен-иала с использованием данных о структуре относительного геопотенциала. Продемонстрировано преимущество такого способа по сравнению с обычным утем расчета структуры абсолютного геопотенциала.

Табл. 5. Библ. 8.

УДК 551.509.314

О вертикальном статистическом контроле данных косвенного температурного зондирования с метеорологических спутников. Гандин Л. С., Тараканова В. П. Труды ГГО, 1980, вып. 431, с. 12—22.

Предложен метод вертикального статистического контроля данных косвенного температурного зондирования, передаваемых в виде значений относительного геопотенциала главных изобарических поверхностей над поверхностью 1000 мбар. Процедура контроля состоит в расчете толщин слоев между каждой парой соседних поверхностей и оптимальной интерполяции этих толщин с последующим сопоставлением наблюденных и интерполированных значений. Вычислены веса и допустимые невязки для данного способа контроля, проанализирована их широтная и сезонная изменчивость, рассмотрена реакция вертикального контроля на различного рода грубые ошибки. Приведены примеры применения рассмотренного метода.

Табл. 3. Библ. 9.

УДК 551.509.314

Алгоритм комплексного статико-временно́го контроля аэрологической информации. Либерман Ю. М. Труды ГГО, 1980, вып. 431, с. 23—30.

Описывается усовершенствованный алгоритм автоматического контроля данных о геопотенциале и температуре изобарических поверхностей в процессе режимной обработки информации. Алгоритм объединяет статический и временно́й методы контроля со специальной процедурой принятия решения о качестве информации и ее возможном исправлении. Приводятся сведения о программной реализации алгоритма.

Табл. 2. Библ. 4.

УДК 551.509.1

Опыт статико-временно́го контроля аэрологической информации. Либерман Ю. М., Тараканова В. П. Труды ГГО, 1980, вып. 431, с. 31—43.

На материале 3294 телеграмм аэрологического зондирования отлажен алгоритм статико-временно́го контроля. Выявлено его преимущество по сравнению со статическим контролем. Привлечение временно́го контроля дает возможность не только убедиться в правильности однозначных исправлений, сделанных статическим контролем, но и выбрать один вариант исправления при неоднозначных исправлениях, а также отфильтровать ложное исправление. С помощью временно́го контроля удается обнаружить ошибки радиозондирования, чего нельзя сделать статическим контролем.

Табл. 8. Библ. 1.

ДК 551.501

Определение параметров комплексного контроля поля приземного давления. лягина Л. П. Труды ГГО, 1980, вып. 431, с. 44—50.

Создание банка данных дает возможность использования комплексного конт- оля. В работе на примере поля приземного давления предлагаются алгоритмы ризонтального контроля и временного контроля, которые основаны соответст- енно на оптимальной и линейной интерполяции.

Получен ряд относительных ошибок, который затем подвергался статистиче- кой обработке. В результате были найдены значения критерия отбраковки (K). ля горизонтального контроля $K=4$, для временного $K=2$. Если разность между асчетным значением и наблюденным будет превышать $KE_{\text{соп}}$, то информацию ожно считать сомнительной.

Библ. 7. Илл. 2.

УДК 551.524.1

Исследование возможности применения метода оптимальной интерполяции для контроля средней месячной температуры воздуха. Клягина Л. П., Скворцова Э. М. Труды ГГО, 1980, вып. 431, с. 51—55.

Приводятся результаты исследования возможности применения метода оптимальной интерполяции для контроля рядов средней месячной температуры воздуха по 158 станциям за период 1951—1970 гг. для января и июля.

Показано, что исследуемый метод недостаточно учитывает характер подстилающей поверхности, что приводит к отбраковке правильных данных при контроле.

Делается вывод о необходимости совершенствования метода подбора влияющих станций.

Илл. 1. Библ. 8.

ДК 551.509.1

К методике согласования наблюдений с данными численного прогноза.
андин Л. С., Тараканова В. П. Труды ГГО, 1980, вып. 431, с. 56—67.

Предлагается модификация разработанного Р. Л. Каганом метода согласо-
ния данных наблюдений и численного прогноза, базирующегося на идее
делирования свойств численного прогноза характеристиками статистического
огноза. Модификация использует инвариантную по отношению к повороту
ординат схему статистического прогноза, что позволяет не ограничиваться
ноточечным согласованием интерполированного и прогностического значений,
проводить многоточечное согласование данных наблюдений и прогноза.

Выполнены расчеты характеристик схемы статистического прогноза, влияю-
их на процедуру согласования и, в частности, корреляционных функций ощи-
к статистического прогноза. Сопоставление последних с данными об анало-
чных функциях для численных прогнозов позволило выбрать рациональные
ачения параметров схемы статистического прогноза. Показано, что вариация
их параметров в разумных пределах слабо влияет на величины, от которых
висит результат согласования.

Табл. 2. Илл. 2. Библ. 19.

ДК 551.501

**О расчете продолжительностей выбросов рядов средней суточной темпера-
ры воздуха.** Хлебникова Е. И. Труды ГГО, 1980, вып. 431, с. 68—73.

Рассматривается способ расчета продолжительностей выбросов случайных
следовательностей, основанный на предположении о марковости последова-
льности, являющейся импульсным преобразованием исходной. Применительно
гауссовским случайным процессам с экспоненциальной корреляционной функ-
ей, удовлетворительно описывающим вероятностную структуру рядов средней
точной температуры воздуха, изучается возможность оценки распределения
одолжительностей выбросов путем использования марковских моделей раз-
ичной связности. Для процессов рассматриваемого типа приведенные данные
зволяют легко оценивать распределения продолжительностей выбросов для
обых уровней и коэффициентов корреляции между последовательными наблю-
ениями.

Табл. 2. Илл. 1. Библ. 6.

ДК 551.501

К вопросу о выборе интервала осреднения периодически коррелированных
личин. Каган Р. Л., Сибир Е. Е. Труды ГГО, 1980, вып. 431, с. 74—89.

Рассматривается вопрос о влиянии выбора начала периода осреднения пе-
иодически коррелированных величин на статистические характеристики осред-
нных значений. Оценки выполняются для случайных процессов, нестационар-
сть которых определяется изменением дисперсии на периоде осреднения при
азличном задании корреляционных функций.

Значения средних квадратических отклонений осредненных величин и кор-
еляции их для смежных периодов зависят от выбора начала периода осредне-
ия. Для монотонно убывающих корреляционных функций минимум дисперсии
максимум корреляции осредненных величин имеют место при совпадении
ачала периода с максимумом дисперсии исходной величины.

Выполненные для конкретных метеорологических элементов количественные
ценки показывают, что этот эффект практически мало сказывается на их ста-
истических характеристиках как при суточном, так и при годовом осреднении.
Табл. 3. Илл. 2. Библ. 6.

УДК 551.501

О точности осреднения периодически коррелированных величин. Ка-
ан Р. Л., Сибир Е. Е. Труды ГГО, 1980, вып. 431, с. 90—105.

Рассматривается влияние выбора числа и размещения сроков наблюдения
а периоде и нестационарности процесса на точность определения его осреднен-
ых значений. Оценки выполняются для случая арифметического осреднения
анных наблюдений в равноотстоящие моменты времени периодически коррели-
ованных процессов, нестационарность которых определяется изменением дис-
ерсии на периоде осреднения. Они определяются теоретически по данным о ста-
истической структуре процесса при различном задании его корреляционной
ункции.

Наибольшая точность осреднения в этом случае обеспечивается при симмет-
ичном расположении моментов наблюдения на периоде, при котором первый
последний моменты отстоят от его концов на половину интервала. При таком
асположении моментов измерения нестационарность процесса практически не
казывается на точности осреднения. При несимметричном их расположении
естационарность приводит к увеличению ошибки осреднения в случае, если
аксимум дисперсии приходится на концы периода, и к уменьшению ошибки
случае, если он приходится на середину периода осреднения.

В качестве примера представлены оценки точности определения средней
уточной температуры воздуха на станции. Ленинград ГМО по данным, полу-
енным при различном выборе сроков наблюдений.
Табл. 7. Илл. 1. Библ. 12.

К 551.501

**К вопросу о сходимости распределения к нормальному для сумм связных
учайных величин.** Каган Р. Л., Хлебникова Е. И. Труды ГГО, 1980,
п. 431, с. 106—116.

На примере процесса $y(t)=x^2(t)$, где $x(t)$ — стационарный гауссовский про-
сс, рассматривается вопрос о быстроте убывания асимметрии процесса при его
еменном осреднении. Показано, что распределение ошибок представления сред-
й по фиксированному интервалу по данным дискретных отсчетов на нем
одится к нормальному с увеличением числа отсчетов гораздо быстрее, чем
ми средние арифметические. Приводятся некоторые оценки точности опреде-
ния характеристик погрешности осреднения методом прореживания эталонной
следовательности.

Табл. 2. Илл. 2. Библ. 6.

ДК 551.501

**Об амплитуде суточного хода статистических параметров температуры воз-
ха.** Сибир Е. Е. Труды ГГО, 1980, вып. 431, с. 117—122.

Приводятся данные об относительных амплитудах хода нормы и дисперсии
мпературы воздуха для 42 станций Советского Союза.

Обсуждаются закономерности годового хода амплитуд суточного хода дис-
рсии.

Табл. 2. Илл. 2. Библ. 6.

ДК 551.510.53 : 629.13

**Статистическая оценка чувствительности одной модели атмосферного пере-
оса к возмущениям ее параметров.** Андронова Н. Г. Труды ГГО, 1980,
ып. 431, с. 123—135.

На основе одномерной численной среднезональной модели переноса и фото-
иссоциации N_2O в стратосфере опробована методика оценки статистической
ачимости реакции модели на вводимые малые возмущения. Формулируются
екомендации по использованию указанной методики при оценке внутренней
зменчивости сложных моделей атмосферных процессов.

Табл. 3. Илл. 6. Библ. 5.

УДК 551.501

О пространственной статистической структуре температуры и высоты тропопаузы. Курбанов Б. Т. Труды ГГО, 1980, вып. 431, с. 136—144.

Исследуется пространственная статистическая структура высоты и температуры тропопаузы. Приводятся графики нормированных корреляционных функций для срединных месяцев каждого сезона, значения радиусов корреляции, значения средних квадратических отклонений, а также некоторые другие параметры корреляционных функций высоты и температуры тропопаузы.

Табл. 4. Илл. 3. Библ. 11.