

ГОСУДАРСТВЕННЫЙ КОМИТЕТ СССР
ПО ГИДРОМЕТЕОРОЛОГИИ
И КОНТРОЛЮ ПРИРОДНОЙ СРЕДЫ

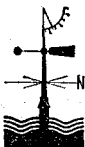
ТРУДЫ
ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ГЛАВНОЙ ГЕОФИЗИЧЕСКОЙ ОБСЕРВАТОРИИ
им. А. И. ВОЕЙКОВА

Выпуск

466

ПРИМЕНЕНИЕ
СТАТИСТИЧЕСКИХ МЕТОДОВ
В МЕТЕОРОЛОГИИ

Под редакцией
д-ра физ.-мат. наук / Р. Л. Кагана /



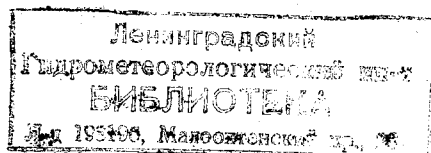
ЛЕНИНГРАД ГИДРОМЕТЕОИЗДАТ 1983

Основное содержание сборника составляют вопросы анализа метеорологических полей и применения данных об их статистической структуре к различным теоретическим и прикладным задачам. Выполнены оценки возможности эффективного определения климатических характеристик с учетом особенностей распределения сети станций. Рассмотрены также принципиальные вопросы регрессионных методов прогноза.

Рассчитан на научных работников, аспирантов и студентов гидрометеорологических и родственных специальностей.

The problems of the objective analysis of meteorological fields and the application of the data on their statistical structure to various theoretical and practical problems are the principal topics under consideration. The possibilities to determine correctly some climatic characteristics taking into account the peculiarities of actual station network distribution have been estimated. Some principal questions connected with the regression prediction methods have also been considered.

The book is designated for scientists, graduate and post—graduate students in meteorology, hydrology, oceanography and adjacent specialities.



Р. Л. Коган, Е. И. Хлебникова

О НАРУШЕНИИ ОДНОРОДНОСТИ МЕТЕОРОЛОГИЧЕСКИХ ПОЛЕЙ ПРИ ИХ ИНТЕРПОЛЯЦИИ

1. Известно, что при пространственной интерполяции метеорологических полей характеристики их статистической структуры изменяются. Степень искажения этих характеристик зависит от многих факторов, в том числе от статистической структуры исходного поля, применяемой схемы интерполяции и густоты сети станций, используемой при интерполяции. В отличие от статистической структуры исследуемого поля, схемы интерполяции с течением времени претерпевают значительные изменения. В большинстве случаев, особенно при рассмотрении данных за прошлые годы, вообще трудно говорить об определенных схемах, поскольку интерполяция производилась субъективно и каждый специалист выполнял ее по-своему, с учетом своего личного опыта.

В настоящее время, с внедрением методов численного анализа, используются четко заданные алгоритмы интерполяции. Однако в разных учреждениях эти алгоритмы разные, и различия их не всегда известны потребителям данных. Наконец, сеть станций, используемая при интерполяции одного и того же поля в различных географических районах, а также в различные периоды времени, имеет разную густоту. Это обстоятельство приводит к тому, что интерполированные значения метеорологических полей оказываются неоднородными как по пространству, так и по времени, что, разумеется, следует иметь в виду при пользовании данными в узлах регулярной сетки, получаемыми путем пространственной интерполяции данных наблюдений на сети метеорологических станций. Это тем более необходимо, что в последние годы банки данных в узлах регулярной сетки получают все более широкое применение, в том числе и для целей оценки характеристик изменчивости метеорологических полей [5, 8, 13].

В настоящей статье приведены некоторые количественные оценки возможного влияния густоты сети станций на характери-

стики изменчивости метеорологических полей. Оценки получены априорно по данным о статистической структуре этих полей для метеорологических элементов существенно разной структуры и для различных способов интерполяции.

Рассмотрим поле величины f , которое, вообще говоря, может быть неоднородным и неизотропным. Мы, однако, для простоты будем считать его однородным и изотропным, по крайней мере, в той области, в которой производится интерполяция. Будем полагать известными корреляционную функцию r , дисперсию σ^2 , а также меру ошибок данных наблюдений η^2 . Пусть по данным наблюдений f производится интерполяция в узлы регулярной сетки по формуле

$$\bar{f}_i = \sum_{k=1}^n p_k^i f_k, \quad (1)$$

где n — число станций, привлекаемых к интерполяции в i -й узел сетки; f_k — данные наблюдений на станциях; p — веса интерполяции; $\bar{f}_i$ — интерполированные значения.

Ковариация интерполированных значений в узлах i и j при этом имеет вид

$$\bar{R}_{ij} = \sigma^2 \left(\sum_{k=1}^n \sum_{l=1}^n p_k^i p_l^j r_{kl} + \eta^2 \sum_{k=1}^n p_k^i p_k^j \right), \quad (2)$$

где $r_{kl} = r(\rho_{kl})$, ρ_{kl} — расстояние между станциями k и l .

Для однородного и изотропного поля отсюда следуют простые формулы для оценки дисперсии $\tilde{\sigma}^2$ и корреляции $\tilde{r}$ интерполированных значений:

$$\tilde{\sigma}_i^2 = \chi_i^2 \sigma^2 = \sigma^2 \left(\sum_{k=1}^n \sum_{l=1}^n p_k^i p_l^i r_{kl} + \eta^2 \sum_{k=1}^n (p_k^i)^2 \right), \quad (3)$$

$$\tilde{r}_{ij} = \frac{1}{\chi_i \chi_j} \left(\sum_{k=1}^n \sum_{l=1}^n p_k^i p_l^j r_{kl} + \eta^2 \sum_{k=1}^n p_k^i p_k^j \right). \quad (4)$$

Величина χ_i , представляющая собой отношение средних квадратических отклонений (стандартов) интерполированных и фактических значений, будет называться нами мерой занижения изменчивости.

2. Для оценки характеристик интерполированных полей был выполнен ряд численных экспериментов. В первой серии экспериментов взаимное расположение узлов интерполяции и станций, в которых имеются данные наблюдений величины f , моделировалось на ЭВМ. Это позволяло производить оценки для большого набора гипотетических сетей станций заданной густоты, что давало возможность, в частности, оценить влияние неравномерности сети станций на характеристики структуры интерполированных величин. В качестве опорной задавалась квадратная сетка с шагом H . Расчеты велись применительно к двум способам ин-

терполяции: оптимальной интерполяции, для которой необходимы данные о статистической структуре исходного поля, и билинейной интерполяции, при которой поле в окрестности узла аппроксимируется полиномом первой степени по каждой из координат. Интерполяция в узлы сетки производилась по данным четырех ближайших станций. Поскольку поле интерполированных величин неоднородно и статистические характеристики его зависят от того, как узел расположен относительно станций, используемых при интерполяции, для выявления закономерных отличий поля интерполированных значений от исходного оценки производились для сетки густо расположенных узлов интерполяции с последующим осреднением. Расчеты велись для разных типов пространственной структуры исследуемого поля. Приведем некоторые результаты, полученные в том случае, когда она описывается экспоненциальной корреляционной функцией вида

$$r(\rho) = \exp(-\rho/\rho_0), \quad (5)$$

где ρ_0 — масштаб корреляции поля.

Зависимости такого вида широко используются для аппроксимации корреляционных функций многих метеорологических полей, например полей осадков, сумм радиации [10], общего содержания озона [4].

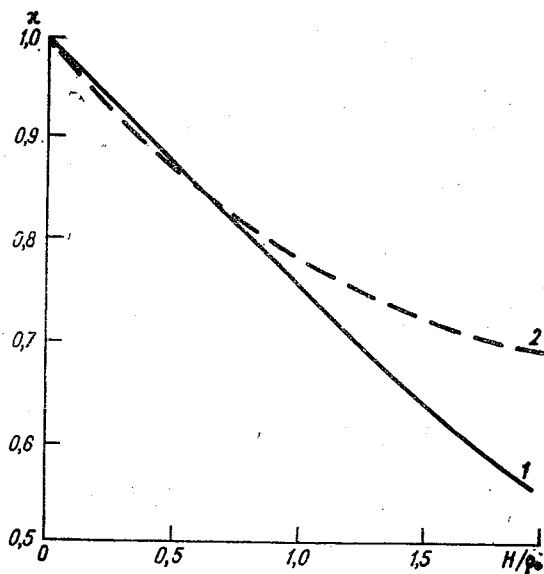


Рис. 1. Зависимость стандарта интерполированных величин от густоты сети ($r(\rho) = \exp(-\rho/\rho_0)$; $\eta^2=0$; $\alpha=0$):

1 — оптимальная интерполяция, 2 — билинейная интерполяция.

Рассмотрим прежде всего случай регулярного расположения станций. Очевидно, оценки характеристик статистической структуры поля интерполированных значений зависят не от расстояния между «станциями» H и масштаба корреляции ρ_0 в отдельности, а от их отношения H/ρ_0 . С увеличением этого отношения, т. е. при уменьшении густоты сети станций, дисперсия интерполированных величин уменьшается и при $H \approx \rho_0$ оказывается примерно вдвое меньше дисперсии исходного поля σ^2 ; соответственно мера занижения стандарта χ в этом случае будет составлять около 0,75.

Поскольку величина χ минимальна в центре квадрата, образуемого «станциями», и максимальна вблизи его углов, т. е. вблизи станций, для выявления закономерных отклонений дисперсии интерполированного поля от истинной удобно пользоваться осредненными по квадрату значениями величины χ^2 . Из рис. 1, на котором приведена зависимость χ от расстояния между станциями, видно, что до расстояний $H \approx \rho_0$ занижение изменчивости интерполированных величин оказывается примерно одинаковым для обоих способов интерполяции. При $H < 0,7\rho_0$ дисперсия оптимально проинтерполированных значений оказывается в среднем несколько выше дисперсии величин, полученных путем билинейной интерполяции. При более редкой сети станций занижение изменчивости оказывается большим при оптимальной интерполяции, причем различие между ними быстро возрастает с увеличением H .

Наличие ошибок наблюдений по-разному сказывается на характеристиках величин, интерполированных разными способами. При оптимальной интерполяции для фильтрации ошибок уменьшают интерполяционные веса, что приводит к дополнительному уменьшению дисперсии интерполированных величин. В результате до значений $H \approx 2\rho_0$ зависимость χ от H хорошо аппроксимируется выражением вида

$$\chi_{\text{опт}} = \frac{\sigma}{\sigma} \approx (1 - \eta^2/8) (1 - 0,23H/\rho). \quad (6)$$

Веса билинейной интерполяции не зависят от ошибок наблюдений. Учет ошибок приводит к увеличению дисперсии интерполированных величин, которая для достаточно близких к станциям точек может быть больше дисперсии исходного поля. В этом случае аналогичная (6) аппроксимационная формула имеет вид

$$\chi_{\text{бил}} \approx 1 + \frac{2}{9} \eta^2 - 0,23 \frac{H}{\rho_0}. \quad (7)$$

Эта формула применима при $H < \rho_0$.

В табл. 1 представлены оценки пространственной корреляционной функции интерполированных величин для некоторых значений H , полученные (с целью сглаживания мелкомасштабных флуктуаций) путем осреднения по полю соответствующих коэффициентов корреляции.

Из таблицы видно, что значения корреляционной функции интерполированного поля в целом больше соответствующих значений корреляционной функции исходного поля. При малых расстояниях ($\rho < H$), когда рассматриваемые точки находятся, как правило, в пределах одного квадрата, завышение корреляции в абсолютных цифрах сравнительно невелико, однако сглаживание поля f за счет интерполяции приводит к замене практически линейного убывания корреляции с расстоянием параболическим ходом. Естественно поэтому, что использование корреляционных функций $\tilde{r}$ вместо исходной корреляционной функции r в таких задачах, для

Таблица 1

Корреляция $\tilde{r}(\rho)$ интерполированных величин
при равномерной сети станций

η^2	Способ интерполяции	ρ (в долях от H)						
		0,25	0,5	1,0	1,5	2,0	2,5	3,0
$H=0,25 \rho_0$								
0	Оптимальная	0,990	0,962	0,881	0,781	0,704	0,619	0,563
	Билинейная	0,989	0,961	0,879	0,782	0,702	0,617	0,561
0,05	Оптимальная	0,986	0,960	0,890	0,791	0,717	0,628	0,574
	Билинейная	0,983	0,953	0,882	0,778	0,709	0,617	0,570
	$r(\rho)$	0,939	0,883	0,779	0,687	0,607	0,535	0,472
$H=\rho_0$								
0	Оптимальная	0,963	0,871	0,626	0,390	0,257	0,152	0,106
	Билинейная	0,962	0,866	0,617	0,380	0,252	0,147	0,105
0,05	Оптимальная	0,959	0,868	0,644	0,404	0,282	0,171	0,132
	Билинейная	0,955	0,858	0,630	0,388	0,275	0,163	0,181
	$r(\rho)$	0,779	0,607	0,368	0,223	0,135	0,082	0,050

которых требуется знание характеристик структуры для малых расстояний, например для оценок точности интерполяции, может привести к заведомо ложным выводам. Для больших расстояний ($\rho > 2-3H$) превышение $\tilde{r}$ над r является большим, однако характер убывания корреляционной функции на этом интервале значений ρ близок к истинному, особенно при небольших значениях H .

Из таблицы 1 видно также, что для представленного в ней случая равномерного размещения станций корреляция между интерполированными значениями практически не зависит от точности задания исходного поля (от величины η) и способа интерполяции.

Наряду с опорной сетью станций рассматривались также нерегулярные сетки, характеризующиеся той же средней плотностью станций на площади. Координаты станций нерегулярных сетей моделировались путем введения случайных смещений в координаты опорной сетки. В этом варианте интерполяция также производилась по данным четырех станций. В тех случаях, когда в одном из квадрантов поиск не обнаруживал станций, в качестве влияющей использовалась ближайшая станция из другого квадранта. Заметим, что в таких случаях билинейная интерполяция осуществляется с большой погрешностью и соответственно очень велико искажение характеристик структуры интерполированных значений. Поэтому для исключения этих случаев при осреднении характеристик интерполированных значений использовались оценки лишь для узлов, расположенных в центральной зоне моделируемого поля.

Выполненные оценки показывают, что при интерполяции по неравномерной сети станций дисперсия поля занижается, вообще говоря, в большей степени, чем в случае равномерной сети той же плотности. В связи с этим представляло интерес выяснить наличие связи величины κ с характеристиками неравномерности сети. В качестве такой характеристики был принят коэффициент неравномерности α , представляющий собой коэффициент вариации площадей, тяготеющих к станциям данной сети. Эти площади для каждой станции определялись, как это принято при осреднении методом полигонов, путем перебора всех точек вспомогательной густой регулярной сетки и отбора тех из них, расстояние от которых до станции меньше расстояния до всех других станций.

Результаты оценок для значений коэффициента неравномерности $\alpha = 0,1 \dots 0,6$ дают значения κ на несколько процентов меньше, чем для случая равномерной сети ($\alpha = 0$), однако четкой зависимости между κ и α не выявлено. Заметим, что такая зависимость хорошо выражена в случае, если оценки производятся не только по центральной, но и по периферийной зоне, для которой при неравномерной сети во многих узлах вместо интерполяции приходится проводить экстраполяцию, точность которой гораздо меньше; соответственно в этом случае степень сглаживания поля увеличивается.

Оценки корреляционных функций интерполированных значений $\tilde{r}$ для неравномерной сети показывают, что они мало отличаются от соответствующих корреляционных функций для равномерной сети той же средней плотности. В качестве примера в табл. 2 приводятся оценки искажений корреляционной функции, аналогичные представленным в табл. 1, для варианта размещения станций (рис. 2), характеризующегося коэффициентом неравномерности $\alpha = 0,5$.

3. Оценки характеристик изменчивости интерполированных величин, выполненные для случайных полей другой структуры, приводят к результатам, аналогичным представленным выше для случая экспоненциальной структуры. В частности, подтверждает-

ся вывод о том, что при редкой сети станций оптимальная интерполяция гораздо больше, чем другие способы интерполяции, сглаживает аномалии поля. В связи с этим возникает вопрос, в какой степени целесообразно использование этого способа интерполяции в различных целях.

Необходимо иметь в виду, что оптимальная интерполяция используется в оперативной практике Гидрометцентра СССР и некоторых других прогностических центрах, в том числе и зарубежных, и хорошо зарекомендовала себя в рамках объективного анализа для последующего численного прогноза погоды, когда осу-

Т а б л и ц а 2

Корреляция $\tilde{r}(\rho)$ интерполированных величин
при неравномерной сети станций ($\alpha = 0,5$)

η^2	Способ интерполяции	ρ (в долях H)						
		0,25	0,5	1,0	1,5	2,0	2,5	3,0
$H = 0,25 \rho_0$								
0	Оптимальная	0,988	0,960	0,878	0,790	0,720	0,640	0,557
	Билинейная	0,983	0,958	0,885	0,801	0,730	0,652	0,568
0,05	Оптимальная	0,986	0,959	0,886	0,800	0,731	0,655	0,573
	Билинейная	0,978	0,951	0,880	0,796	0,728	0,651	0,571
	$r(\rho)$	0,939	0,883	0,779	0,687	0,607	0,535	0,472
$H = \rho_0$								
0	Оптимальная	0,957	0,864	0,617	0,411	0,286	0,179	0,099
	Билинейная	0,933	0,851	0,642	0,438	0,309	0,199	0,114
0,05	Оптимальная	0,952	0,861	0,627	0,426	0,304	0,199	0,122
	Билинейная	0,928	0,845	0,643	0,444	0,321	0,215	0,133
	$r(\rho)$	0,779	0,607	0,368	0,223	0,135	0,082	0,050

ществляемое при анализе значительное сглаживание поля, по-видимому, является полезным. Может, однако, возникнуть сомнение в желательности использования оптимальной интерполяции для задач, в которых требуется минимальное искажение характеристик изменчивости, например, для целей мониторинга климата. Возможно, для этих целей оказалось бы более целесообразным использование других способов интерполяции, которые хотя и давали бы большую среднюю квадратическую погрешность интерполяции, но приводили бы к меньшему искажению характеристик изменчивости.

Очевидно, что для каждого способа интерполяции, для которого известна количественная зависимость меры занижения стандарта χ от исходных параметров, можно путем введения множителя $1/\chi$

к интерполированным значениям добиться равенства дисперсий исходного поля и поля интерполированных значений (специалисты по долгосрочному прогнозу называют такую коррекцию процедурой «инфляции» [11]). Оказывается, однако, что оптимальная интерполяция имеет в этом отношении большие преимущества. Во-первых, можно показать, что средняя квадратическая относительная погрешность оптимальной интерполяции ϵ_0 , которая автоматически определяется в процессе оптимальной интерполяции и которая меньше погрешностей любых других способов интерполяции, связана с мерой занижения стандарта κ простым соотношением

$$\epsilon_0^2 = 1 - \kappa^2. \quad (8)$$

Во-вторых, из всех возможных способов интерполяции именно оптимальная интерполяция после введения нормирующего множителя обеспечивает наибольшую точность интерполяции при отсутствии искажения дисперсии. В-третьих, для оптимальной интерполяции относительная средняя квадратическая погрешность интерполяции ϵ однозначно связана с погрешностью ϵ_0 интерполяции до коррекции:

$$\epsilon = \sqrt{2 \left| 1 - \sqrt{1 - \epsilon_0^2} \right|}. \quad (9)$$

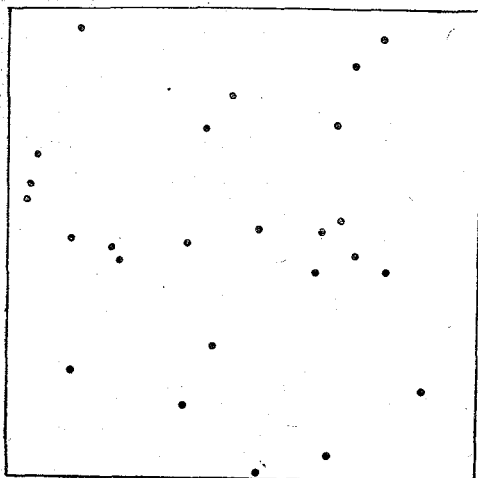


Рис. 2. Схема расположения станций.

Это позволяет, в частности, оценить географическое распределение меры занижения стандарта в различных районах земного

шара для многих метеорологических элементов, пользуясь выполненными ранее оценками точности оптимальной интерполяции [1].

При выяснении целесообразности проведения «равнодисперсной» интерполяции нужно иметь в виду, что, устраняя искажение дисперсии, она вовсе не исключает искажения характеристик корреляции и характеристик распределения. Кроме того, точность такой интерполяции меньше точности оптимальной интерполяции. Оценки по формуле (9) показывают, что до значений $\epsilon_0 \approx 0,5$ значения ϵ практически совпадают с ϵ_0 . В диапазоне $\epsilon \approx 0,5 \dots 0,8$ величина ϵ превышает ϵ_0 не более чем на 10 %. Однако при больших значениях ϵ_0 ($\epsilon_0 > 0,86$) относительные ошибки равнодисперсной интерполяции превышают единицу, т. е. ошибка интерполяции больше естественной изменчивости интерполируемой величины. Ясно, что в таких случаях использование подобной интерполяции нецелесо-

образно. Следует, однако, иметь в виду, что в этих условиях любые способы интерполяции не обеспечивают сколько-нибудь надежного описания пространственного распределения изучаемого поля. Вряд ли является существенным преимуществом менее точного способа интерполяции то обстоятельство, что дисперсия интерполированных величин для этого способа больше за счет более высокого уровня случайных ошибок интерполяции.

Указанные причины делают, на наш взгляд, использование оптимальной интерполяции предпочтительным и при формировании банков данных в узлах сетки для климатологических целей. Однако для правильной их интерпретации целесообразно хранить наряду с самими величинами и характеристики их точности в виде значений ϵ_0 .

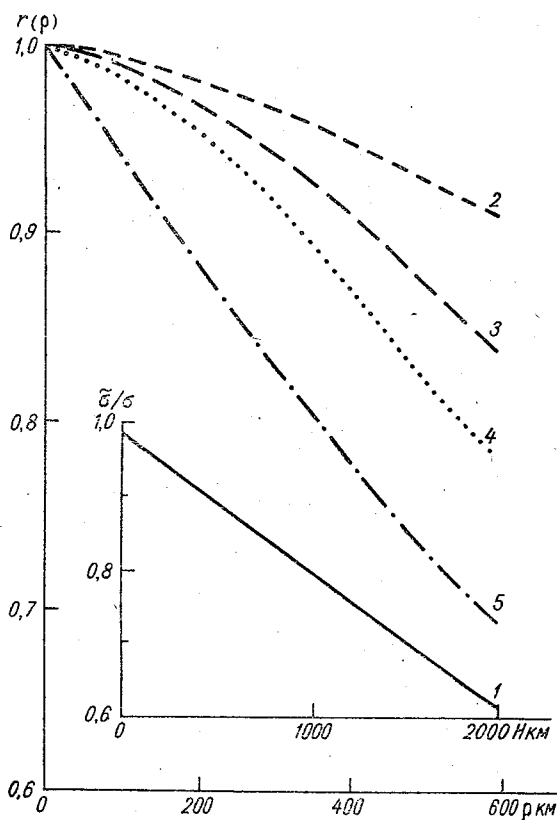


Рис. 3. Характеристики изменчивости интерполированных средних месячных значений общего содержания озона в атмосфере.

1 — $\kappa(H)$; 2 — $\tilde{r}(\rho)$ при $H=2000$ км; 3 — $\tilde{r}(\rho)$ при $H=800$ км; 4 — $\tilde{r}(\rho)$ при $H=400$ км; 5 — $\tilde{r}(\rho)$.

Не располагая такой информацией, для конкретного метеорологического элемента, можно, пользуясь приведенными выше оценками, получить общие выводы о характере искажения статистической структуры, без большой их детализации.

Так, на рис. 3 представлены оценки для пространственной корреляционной функции и стандарта интерполированных значений среднего месячного общего содержания озона в атмосфере. Как показывают выполненные в [7] оценки, пространственная корреляция этого элемента описывается экспоненциальной функцией (5) при масштабе корреляции $\rho_0 = 1600$ км и мере ошибок $\eta^2 = 0,1$. Это позволяет получить оценки искажения характеристик структуры путем непосредственного использования представленных в предыдущем разделе данных.

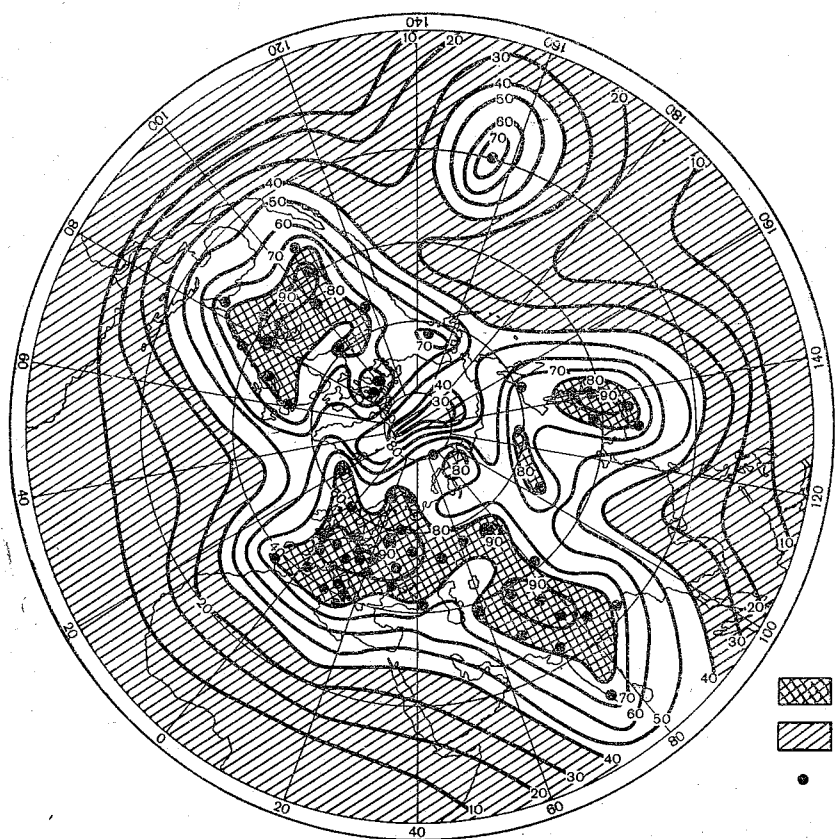


Рис. 4. Мера завышения стандарта κ для интерполированных средних месячных значений общего содержания озона в атмосфере.

1) $\kappa > 0,8$; 2) $\kappa < 0,4$; 3) станции озонометрической сети, учтывавшиеся при расчетах.

Из рисунка следует, что для $H=400$ км (характерное расстояние между озонотрическими станциями в некоторых районах Западной Европы) при равномерном расположении станций занижение стандарта находится в пределах 10 %. Для значений $H=800$ км, характерных для озонотрической сети СССР, оно составляет около 20 %. Для $H=2000$ км, что приблизительно соответствует густоте станций в Северной Атлантике, занижение σ составляет около 40 %.

Располагая данными о точности оптимальной интерполяции при конкретном расположении сети станций, можно детально описать географическое распределение меры занижения σ стандарта интерполированных величин.

В качестве примера на рис. 4 приведена карта распределения величины σ для среднего месячного содержания озона в атмосфере, построенная на основе выполненных К. М. Лугиной оценок точности оптимальной интерполяции по данным озонотрической сети станций на северном полушарии. При этом предполагалось использование данных 68 станций, оснащенных спектрофотометром Добсона или советскими озонотрирами М-83, регулярно публикуемых в [14].

Из рисунка видно, что наряду с большими значениями (малым занижением дисперсии) в Европе, Японии, Северной Америке имеются обширные зоны в полярных, экваториальных районах и на акваториях океанов, которые характеризуются значениями σ , составляющими 50 % и менее. Для таких районов, конечно, очень большим является и завышение пространственной корреляционной функции. Это обстоятельство следует иметь в виду, например, при интерпретации данных, полученных в [8] путем обработки значений в узлах сетки, содержащихся в атласе [12].

В случае отсутствия данных о точности интерполяции некоторое представление о степени искажения характеристик изменчивости может дать и использование приближенных формул типа (6) наряду с данными о реальной неравномерности сети в конкретных районах. Необходимо лишь иметь в виду, что формальное использование сведений о сети станций не всегда оправдано, так как в те или иные годы данные некоторых станций могут отсутствовать и фактическое число станций, на которых имеются данные к моменту обработки, всегда меньше списочного.

4. Из оценок, подобных представленным на рис. 4, следует, что в силу различной густоты сети станций в разных районах земного шара некритическое использование данных, интерполированных в узлы сетки, может привести к выводу о существенной пространственной неоднородности изучаемого поля даже в тех случаях, когда исходное поле является однородным.

В равной мере можно ожидать появления ложной временной нестационарности метеорологических рядов при изучении их по интерполированным данным, поскольку густота сети станций за период инструментальных наблюдений существенно возрасла, осо-

бенно в малообжитых периферийных зонах. Оценки возможных изменений характеристик структуры метеорологических полей за счет этого фактора производились нами в ходе второй серии экспериментов. В отличие от первой серии, в которой расположение «станций» моделировалось произвольным образом, в этой серии характеристики структуры интерполированных величин оценивались при задании фактического расположения станций, производивших измерения данного элемента в той или иной период.

Приведем некоторые результаты, относящиеся к характеристикам изменчивости термического режима в полярной зоне северного полушария. Известно [2, 3], что эта зона особенно чувствительна к возможным изменениям климата, поэтому слежение за фактическим вековым ходом ее температуры представляет важную практическую задачу. К сожалению, эта зона плохо освещена данными фактических наблюдений, что уменьшает возможности изучения хода температуры воздуха в ней.

Нами были выполнены оценки для трех периодов: 1900—1909, 1936—1945 и 1966—1975 гг. Если в течение первого периода севернее 70 с. ш. наблюдения за температурой воздуха ве-

Таблица 3
Параметры статистической структуры

Период осреднения	ρ_0 тыс. км	k	η^2
Январь	3,20	1,49	0,002
Год	1,95	1,62	0,021

ли лишь 3 станции, то в течение второго периода число их превысило 30, а в течение третьего оно увеличилось еще более чем вдвое. Хотя расположение станций является довольно неблагоприятным для освещения арктической зоны (они размещаются главным образом на ее периферии),

тем не менее увеличение их числа позволяет лучше оценивать нужные характеристики термического режима. При оценках, конечно, желательно было бы использовать данные о статистической структуре температуры воздуха именно арктической зоны. Таких данных в настоящее время очень мало по той же причине отсутствия длиннорядных температурных наблюдений.

Мы воспользовались при расчетах для характеристики структуры температуры в зимние месяцы данными [9], полученными для средней месячной температуры воздуха в январе в центре ЕТС. Если для зимних месяцев можно надеяться на не очень большие различия характеристик структуры арктической и умеренной зон, то для летних месяцев этого ожидать заведомо нельзя. Поэтому мы воспользовались полученными в ГГИ К. М. Лугиной и Н. А. Сперанской оценками пространственных корреляционных функций среднегодовых значений температуры воздуха для Арктики.

Корреляционные функции задавались в виде

$$r(\rho) = \exp(-(\rho/\rho_0)^k). \quad (10)$$

Параметры структуры, использованные при расчете, представлены в табл. 3.

Расчеты, как и ранее, производились применительно к оптимальной и билинейной интерполяции. Для периода 1900—1909 гг. билинейная интерполяция давала неудовлетворительные результаты, поскольку она требует наличия данных по крайней мере четырех станций, в то время как в этот период севернее 70 с. ш. действовали лишь три станции. Привлечение четвертой станции из более низких широт несущественно улучшало результаты из-

Таблица 4

Характеристики статистической структуры интерполированных значений средней температуры воздуха за январь и год в арктической зоне

Характеристика	Широта (разность широт) в град. мерид.	Январь				Год			
		Период			фактические	Период			фактические
		1900—1909	1936—1945	1966—1975		1500—1509	1936—1945	1966—1975	

Расчетные значения

$\frac{\sigma(\varphi)}{\sigma}$	75	0,79	0,98	0,99	1,00	0,67	0,96	0,96	1,00
	80	0,80	0,96	0,98	1,00	0,64	0,92	0,92	1,00
	85	0,79	0,94	0,97	1,00	0,61	0,87	0,87	1,00
	90	0,79	0,93	0,96	1,00	0,60	0,85	0,85	1,00
$r_{\lambda}(\Delta\varphi)$	5	0,99	0,97	0,97	0,95	0,98	0,94	0,94	0,91
	10	0,97	0,91	0,90	0,87	0,94	0,84	0,82	0,76
	15	0,93	0,82	0,81	0,77	0,87	0,68	0,67	0,59
$\sigma_{T_{\varphi}}/\sigma$	80	0,24	0,44	0,47	0,52	0,32	0,58	0,62	0,69
$\sigma_{T_{\lambda}}/\sigma$	80	0,22	0,38	0,43	0,52	0,27	0,49	0,56	0,69
σ_T/σ	80	0,33	0,58	0,64	0,73	0,42	0,76	0,84	0,98

По данным [5] для 80° с. ш.

$\sigma_V^{\circ\text{C}}$	2,6	3,4	3,4	—	2,0	2,4	2,6	—
$\sigma_T^{\circ\text{C}}/10 \text{ мер.}$	2,9	4,2	6,0	—	2,3	3,2	4,8	—

а ее неудачного расположения. Для двух более поздних периодов билинейная интерполяция давала вполне разумные результаты. Поскольку нас интересует изменение характеристик структуры за весь период 1900—1970 гг., мы ограничимся представлением данных для случая оптимальной интерполяции.

Расчеты производились для узлов широтно-долготной сетки арктической зоне с шагом по широте $\Delta\varphi=5^\circ$ и по долготе $\Delta\lambda=20^\circ$. Мы ограничимся представлением лишь некоторых дан-

ных, осредненных по всем 18 узлам каждой широтной зоны. Эти данные приведены в табл. 4. Здесь $\tilde{r}_\lambda(\Delta\varphi)$ представляет собой корреляционную функцию интерполированных значений в меридиональном направлении для расстояний, соответствующих 5, 10 и 15° меридиана; σ_{Γ_φ} и σ_{Γ_λ} — средние квадратические градиенты аномалий температуры в зональном и меридиональном направлениях в расчете на расстояние, соответствующее 10° меридиана; σ_Γ — средние квадратические значения модуля этого градиента ($\Gamma = \sqrt{\Gamma_\varphi^2 + \Gamma_\lambda^2}$).

Из таблицы видно, что с развитием сети станций в арктической зоне характеристики изменчивости приближаются к фактическим, что может привести к выводу о наличии их реального векового хода. Этот ложный ход прослеживается как на величинах стандартов и градиентов интерполированных значений, которые растут со временем, так и на значениях пространственной корреляционной функции, которая убывает со временем. Занижение дисперсии увеличивается с ростом широты, а занижение градиента температуры в меридиональном направлении больше, чем в зональном направлении. Сравнение оценок для января и для средних за год показывает, что в последнем случае сглаживание заметно больше. Это связано с более быстрым затуханием корреляционной функции для среднегодовых значений, отражающим меньшие пространственные масштабы синоптических процессов для летнего периода.

Следует отметить, что из рассмотренных характеристик более всего чувствительны к изменениям густоты сети оценки для градиентов температуры. Это легко понять, если учесть, что они определяются следующим образом:

$$\sigma_{\tilde{r}}(\rho) = \tilde{\sigma} \sqrt{2[1 - \tilde{r}(\rho)]}. \quad (11)$$

При увеличении густоты сети величина $\tilde{\sigma}$ возрастает, $\tilde{r}(\rho)$ убывает, оба эти фактора способствуют увеличению градиентов.

Если значения $\tilde{\sigma}$ за 70-летний период возрастают на 20—30 %, то значения градиентов за это время увеличиваются примерно вдвое.

В табл. 4 приводятся также для сравнения некоторые оценки характеристик изменчивости, полученные Г. В. Груза и Э. Я. Раньковой [5] путем фактической обработки данных о температуре воздуха в узлах широтно-долготной сетки. Здесь σ_ν представляет собой стандарт аномалий средней месячной температуры воздуха вдоль заданного круга широты. Она не совпадает с величиной $\tilde{\sigma}$ (дисперсия аномалий меньше полной дисперсии на дисперсию средней зональной температуры воздуха, которая в арктической зоне довольно велика), тем не менее ход σ_ν дает общее представление о возможном ходе $\tilde{\sigma}$.

Сравнение данных фактических оценок с расчетными показывает их хорошее согласование. Особенно это относится к оценкам градиентов, которые и в том и в другом случае увеличиваются

примерно вдвое за рассматриваемый 70-летний период. В связи с этим представляется, что наблюдающиеся тренды характеристик изменчивости в значительной мере могут объясняться изменением густоты сети станций. На возможность такого объяснения указывалось, впрочем, и в [6].

В заключение следует указать, что авторы вовсе не отрицают наличия пространственной неоднородности характеристик изменчивости метеорологических элементов или возможности их векового хода. Эти эффекты, безусловно, существуют и подлежат тщательному исследованию. Однако для этого необходимо уметь отфильтровывать те изменения, которые не связаны с реальными физическими причинами, а определяются лишь причинами методического характера, в частности использованием данных в узлах сетки. Более того, авторы не призывают отказываться от использования их, поскольку они имеют также определенные преимущества перед данными на станциях (пространственная равномерность сетки, отсутствие пропусков и т. д.), а в некоторых случаях являются единственно возможным источником информации. Необходимо лишь достаточно критически относиться к данным в узлах сетки и учитывать их особенности при практическом использовании.

СПИСОК ЛИТЕРАТУРЫ

1. Белоусов С. Л. и др. Обработка оперативной метеорологической информации с помощью электронных вычислительных машин. — Л.: Гидрометеоиздат, 1968. — 282 с.
2. Будыко М. И. Изменения климата. — Л.: Гидрометеоиздат, 1974. — 280 с.
3. Будыко М. И. Климат в прошлом и будущем. — Л.: Гидрометеоиздат, 1980. — 351 с.
4. Гриб Н. К., Поляк И. И. Некоторые статистические характеристики изменений атмосферного озона. — Метеорология и гидрология, 1973, № 4, с. 87—93.
5. Груза Г. В., Ранькова Э. Я. Данные о структуре и изменчивости климата. — Обнинск, ВНИИГМИ—МЦД, 1979. — 203 с.
6. Груза Г. В., Ранькова Э. В. Структура и изменчивость наблюдаемого климата. Температура воздуха северного полушария. — Л.: Гидрометеоиздат, 1980. — 72 с.
7. Каган Р. Л., Салль И. А., Щербакова Н. Н. О пространственной корреляции средних месячных значений общего содержания озона в атмосфере. — См. наст. сб.
8. Кайгородцев А. Е., Покровский О. М. О естественных составляющих поля общего содержания озона в Северном полушарии. — Метеорология и гидрология, 1980, № 8, с. 52—57.
9. 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.
10. Статистическая структура метеорологических полей. Будапешт, 1976. — 365 с.
11. Glahn H. R., Allen R. A. A note concerning the „Inflation” of regression forecasts. — J. Appl. Meteorol., 1966, vol. 5, N 1, p. 124—126.
12. London J., Bojkov R. D., Oltmans S. Atlas of the global distribution of total ozone. July 1957—June 1967. — NCAR Techn. Note, NCAR/TN/113+STR. — Boulder, Colorado, 1976. — 276 p.

Р. Л. Казан, И. А. Салль, Н. Н. Щербакова

О ПРОСТРАНСТВЕННОЙ КОРРЕЛЯЦИИ СРЕДНИХ МЕСЯЧНЫХ ЗНАЧЕНИЙ ОБЩЕГО СОДЕРЖАНИЯ ОЗОНА В АТМОСФЕРЕ

1. Имеющиеся в настоящее время сведения о пространственной статистической структуре поля общего содержания озона (ОСО) в атмосфере относятся главным образом к их суточным значениям (см., например, [1, 9, 10, 12]). Данные о структуре средних месячных значений, которые представляют первостепенный интерес для многих приложений, практически отсутствуют. Это объясняется сравнительно малой длительностью наблюдений за озоном, в связи с чем коэффициенты корреляции между значениями ОСО в фиксированные месяцы на тех или иных парах станций оказываются гораздо менее надежными, чем соответствующие коэффициенты корреляции для их суточных значений. Наличие пропусков в данных наблюдений, неравномерность размещения сети станций, несопоставимость различных приборов еще более осложняет задачу прямого расчета характеристик статистической структуры месячных значений ОСО.

В работах [5, 15] сделаны попытки обойти эти трудности. В первой из них структура средних месячных значений ОСО определяется косвенным путем по данным о пространственной и временной структуре суточных средних. Однако делается это, по-видимому, не очень корректно. Достаточно указать, что использующиеся в [15] данные о пространственной и временной структуре суточного содержания озона не согласуются друг с другом. В частности, явно противоречат друг другу меры ошибок, принятые при аппроксимации пространственной и временной корреляционных функций ОСО. В работе [5] использованные характеристики структуры среднемесячных значений ОСО были получены путем обработки данных в узлах широтно-долготной сетки, опубликованных в [13]. Однако эти данные являются существенно сглаженными по сравнению с фактическим полем ОСО. Естественно поэтому, что рассчитанная по ним дисперсия оказывается заниженной, а корреляция завышенной, притом тем более, чем реже сеть станций, по которой производится интерполяция. Исправление соответствующих корреляционных функций является достаточно сложным делом. Проще выполнить расчет заново с использованием непосредственно наблюдаемых на станциях данных. Не-

которые результаты таких расчетов представлены в настоящей работе.

2. Для этой цели нами использовались данные, опубликованные в сводках [14]. Из числа станций северного полушария было отобрано 68 станций, которые в сводках за рассматриваемый период (1961—1977 гг.) представлены достаточно регулярно. Схема расположения этих станций представлена на рис. 1. К сожалению, они расположены далеко неравномерно, не позволяют получить полное представление о характере изменчивости поля озона на всем полушарии. Однако это те данные, на которые в настоящее время можно реально рассчитывать. Именно эта сеть является основой для построения оперативных карт ОСО, публикуемых Гидрометцентром СССР [8]. Заметим кстати, что при оперативной работе в действительности используется еще меньший объем данных, потому что часть материала неизбежно поступает с запозданием.

Из выбранных 68 станций 48 (в основном зарубежные) оснащены спектрофотометрами Добсона. На 20 станциях (19 совет-

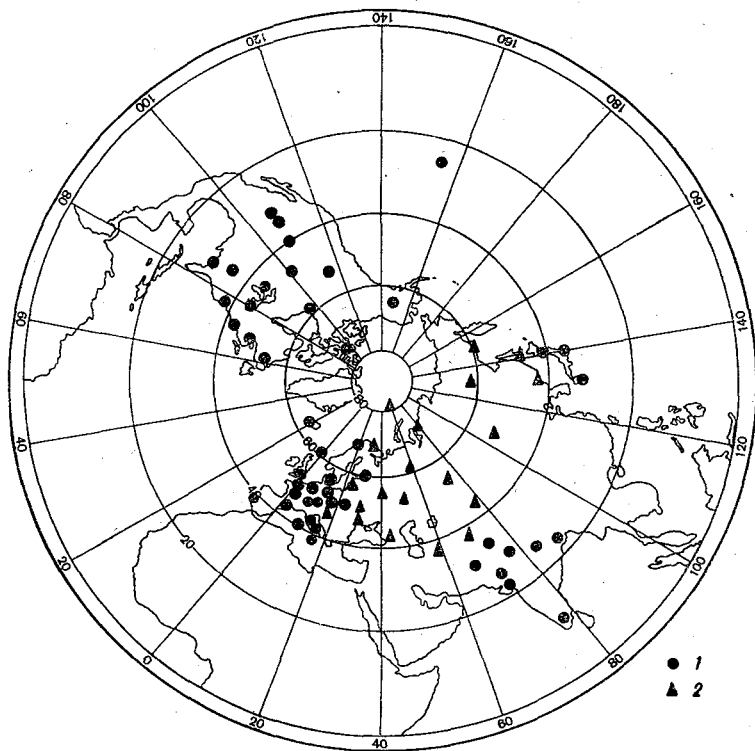


Рис. 1. Схема расположения озонметрических станций.
1 — наблюдения по спектрофотометру Добсона; 2 — наблюдения по фильтровому озонметру М-83.

ских станций и ст. Будапешт) измерения производятся по фильтровому озонметру М-83. На ст. Ленинград (Воейково) установлены и спектрофотометр Добсона и озонметр М-83, однако публикуются в [14] только данные измерений по спектрофотометру Добсона для обеспечения лучшей сравнимости данных наблюдений.

К числу используемых станций был отнесен также Париж. Впоследствии выяснилось, что наблюдения на этой станции ведутся по другому прибору, данные которого совсем несравнимы с данными упомянутых выше приборов. В результате станцию пришлось исключить из рассмотрения.

Расчеты проводились для центральных месяцев сезонов — января, апреля, июля и октября. Кроме того, для приближенной оценки структуры годовых значений ОСО рассматривались средние из значений ОСО для этих четырех месяцев.

Вычисления велись на ЭВМ ЕС-1022 по программе, являющейся вариантом базовой программы, описанной в [4] и реализованной в свое время на ЭВМ М-220 Е. Б. Журавлевой и И. И. Поляком [3]. Для каждой станции рассчитывались следующие величины:

- 1) среднее многолетнее значение ОСО $\bar{O}$;
- 2) среднее квадратическое отклонение (стандарт) ОСО σ_O .

Для каждой пары станций (i и k) определялись также:

- 1) коэффициент корреляции r_{ik} между значениями ОСО на этих станциях;
- 2) расстояние между станциями ρ_{ik} , рассчитываемое по дуге большого круга;
- 3) угол φ_{ik} между вектором расстояния и меридианом, проходящим через середину дуги большого круга, отсчитываемый по часовой стрелке от северного направления.

Выдавалось на печать также число лет n_{ik} , по которым имелись данные на обеих станциях. В случае большого числа пропусков, когда $n_{ik} < 5$, расчет коэффициентов корреляции не производился.

После расчета полной корреляционной матрицы по всем станциям сети рассчитывалась корреляционная функция в предположении однородности и изотропии этой характеристики. Программа позволяет вести расчеты с детализацией по широтным зонам и направлениям, однако на данном этапе попытка такой детализации корреляционной функции представляется преждевременной.

В результате расчета корреляционных функций для каждой градации расстояния выдаются на печать следующие характеристики:

- 1) $\bar{\rho}$ — среднее расстояние;
- 2) m — число пар станций, расстояние между которыми находится в пределах градации;
- 3) $\bar{n}$ — среднее число лет, по которым для этих пар получен коэффициент корреляции;

4) $\bar{r}(\rho)$ — средний эмпирический коэффициент корреляции для этой градации;

5) σ_r — разброс коэффициентов корреляции для градации.

3. Статистические параметры среднемесячных значений ОСО для отдельных станций получены по выборкам малого объема (лишь на немногих станциях имелись данные за все 17 лет наблюдений, на большинстве же станций за 12—15 лет, а на некоторых станциях, особенно на северных станциях зимой, всего за несколько лет), выборочные ошибки этих параметров являются довольно большими. Так, стандартная выборочная погрешность среднего квадратического отклонения $\sigma_{\bar{\sigma}}$ месячных значений ОСО составляет 18—25 % самой величины $\sigma_{\bar{\sigma}}$, стандартная ошибка среднего многолетнего значения для конкретных месяцев 0,25—0,35 $\sigma_{\bar{\sigma}}$.

В действительности в значениях параметров, помимо ошибок, связанных с ограниченностью выборки, содержатся еще значительные ошибки как систематического, так и случайного характера, связанные с особенностями прибора, методами наблюдений и калибровки, отличиями текущих условий погоды в пункте измерений от условий, в которых производилась калибровка. Различные аспекты этой сложной проблемы исследуются в ряде специальных работ, обзор которых содержится в [12]. Мы ее касаться не будем, но должны иметь в виду, что фактические погрешности значений $\bar{\Omega}$ и $\sigma_{\bar{\sigma}}$ могут быть значительно больше указанных выше выборочных погрешностей.

Поскольку данными наблюдений сколько-нибудь удовлетворительно освещены лишь некоторые районы северного полушария, построение соответствующих карт с целью сопоставления их с данными других авторов было затруднительно. Однако сравнение по ряду районов, производившееся с картами, построенными Л. А. Урановой [8] и Дж. Лондоном и др. [13], показывает разумное согласование значений $\bar{\Omega}$. Некоторые расхождения вполне могут объясняться несовпадением периода осреднения (у нас взят период 1961—1977 гг., карты в атласе [13] построены за 1957—1966 гг., карты Л. А. Урановой относятся к периоду 1964—1974 гг.), а также неодинаковым количеством станций, использованных в расчетах.

Для значений $\bar{\Omega}$ характерен четкий годовой ход с максимумом весной и минимумом осенью. Максимум $\sigma_{\bar{\sigma}}$ наблюдается зимой, минимум — летом.

Судить о закономерностях географического распределения $\sigma_{\bar{\sigma}}$ по указанным выше причинам затруднительно. Для января прослеживается довольно четкое увеличение значений $\sigma_{\bar{\sigma}}$ от умеренных широт к полярным. Возможно, впрочем, что это обусловлено не столько реальным увеличением изменчивости ОСО, сколько меньшей надежностью и регулярностью измерений ОСО в северных районах в зимних условиях.

Бросается в глаза резкое превышение значений σ_2 на станциях, оснащенных прибором М-83, над значениями σ_2 на станциях, на которых измерения производятся по спектрофотометру Добсона. Это превышение характерно для всех месяцев как при рассмотрении значений σ на близлежащих станциях (например, станции Саппоро, Татено, Владивосток, Большая Елань, Мурманск и Тромсе и т. д.), так и при рассмотрении значений σ_2 , осредненных в целом по группам станции и представленных в табл. 1.

В среднем значения σ на станциях, оснащенных М-83, оказываются для зимы и весны примерно вдвое, а для лета и осени втрое больше, чем на станциях, оснащенных прибором Добсона. Возможно, что эти различия отчасти связаны с особенностями изменчивости ОСО на континенте Евразии. Представляется, однако, что основной вклад в них вносят значительно большие слу-

Таблица 1

Осредненные по группам станций средние значения и дисперсии ОСО

Характеристика	Число станций	Прибор	Январь	Апрель	Июль	Октябрь	Год
$\bar{\sigma}$ Д. Е.	46	Добсона	336	370	320	292	327
	18	М-83	381	403	304	296	343
σ_2^2 (Д. Е.) ²	46	Добсона	484	357	133	200	100
	18	М-83	2295	1434	1657	1460	796

чайные ошибки измерений, характерные для прибора М-83 до замены их модернизированными приборами в 1973—1974 гг. [2, 11].

Значения статистических параметров для отдельных станций заслуживают более детального анализа. На данном этапе мы ограничились лишь предварительным просмотром этого материала, имея в виду главным образом выявление брака в исходных данных, подлежащего исключению. При такой отбраковке приходится проявлять большую осторожность, чтобы не получить искусственно чрезмерно сглаженное поле при таком субъективном подходе. Поэтому на данном этапе было сочтено целесообразным полностью исключить из рассмотрения лишь данные трех станций: Париж, Альбукерк и Душанбе. Для всех этих станций характерны большие систематические расхождения с близлежащими станциями, повышенное значение σ_2 , а также практическое отсутствие корреляции с соседними станциями. Для июля и года исключались дополнительно данные станций Бедфорд, Бракнелл, Орхус и Потсдам.

4. Значения коэффициентов корреляции для отдельных пар станций использовались нами лишь для отбраковки резко несогласующихся с соседними станциями рядов. Большие выборочные ошибки (каждый из них считался в среднем по ряду 10—12 лет) делали бессмысленным анализ каждого из коэффициентов корреляции в отдельности. Однако в совокупности они дают хорошее представление о характере пространственной связности поля ОСО.

Полученные корреляционные функции ОСО по данным измерений спектрофотометром Добсона представлены в табл. 2.

Таблица 2

Пространственная корреляционная функция ОСО
по данным 46 станций, полученным по спектрофотометру Добсона

$\bar{p}$ тыс. км	m	$\bar{n}$	$10^2 \bar{r}$	$10^2 \sigma_r$	m	$\bar{n}$	$10^2 \bar{r}$	$10^2 \sigma_r$	m	$\bar{n}$	$10^2 \bar{r}$	$10^2 \sigma_r$
Январь					Апрель				Июль			
0,1	2	7,0	93	3	2	7,5	90	7	1	10,0	82	—
0,3	4	7,0	60	19	6	9,0	71	17	4	10,2	82	18
0,5	11	10,4	63	14	12	10,1	58	24	8	12,2	56	20
0,7	19	10,0	54	31	21	10,5	51	24	15	11,6	21	25
0,9	25	9,8	58	19	27	10,1	61	22	19	11,2	28	32
1,2	38	9,4	40	37	46	10,1	43	32	36	10,7	28	34
1,6	46	10,2	48	19	50	10,6	39	31	38	11,6	26	31
2,0	28	10,3	23	28	33	10,2	26	33	28	10,7	21	29
2,4	21	9,3	29	40	26	9,6	11	34	21	10,4	17	27
2,8	19	10,6	7	35	22	10,5	24	30	21	11,3	26	29
3,3	17	9,8	16	37	24	9,3	28	32	23	9,8	38	24
Октябрь					Год				Октябрь (18 станций, М-83)			
0,1	2	8,5	95	8	1	7	97	—	—	—	—	—
0,3	6	10,0	59	21	4	8,2	70	26	—	—	—	—
0,5	12	11,6	45	32	7	9,6	49	24	1	14	72	—
0,7	21	11,7	41	33	14	9,6	34	31	1	14	78	—
0,9	28	10,9	53	33	17	9,0	50	22	5	11,8	55	27
1,2	47	11,0	27	37	31	8,8	31	42	6	9,7	53	31
1,6	50	11,2	26	37	35	9,0	46	35	7	13,1	52	33
2,0	36	10,2	16	31	20	8,8	26	37	13	11,5	47	21
2,4	25	10,0	12	33	16	7,6	21	49	12	12,9	49	27
2,8	22	11,1	16	35	16	9,3	25	55	8	10,8	40	34
3,3	23	9,4	—3	31	15	7,9	55	39	8	11,9	51	19

Эти корреляционные функции удовлетворительно описываются экспоненциальной зависимостью

$$\tilde{r}(\rho) = \tilde{r}(0) \exp(-\rho/\rho_0), \quad (1)$$

где ρ_0 — масштаб корреляции, который для месячных значений может быть принят равным 1600 км; $\tilde{r}(0)$ — значение, полученное путем экстраполяции эмпирической корреляционной функции до расстояния $\rho=0$.

При оценке величины $\tilde{r}(0)$ мы стремились к согласованию корреляционных функций для разных месяцев. Именно, полагая ошибки средних месячных значений ОСО случайными и характеризующимися средней квадратической погрешностью δ_{Σ} , имеем

$$\tilde{r}(0) = \frac{1}{1 + (\delta_{\Sigma}/\sigma_{\Sigma})^2}. \quad (2)$$

С учетом (1) и данных табл. 1 и 2 получаем оценку $\delta_{\Sigma} = 7$ Д. Е. Этому соответствуют значения $\tilde{r}(0)$ и меры случайных ошибок наблюдений $\eta^2 = (\delta_{\Sigma}/\sigma_{\Sigma})^2$, представленные в табл. 3.

Таблица 3

Значения $\tilde{r}(0)$ и η^2 для месячных и годовых значений ОСО
по данным измерений спектрофотометром Добсона

Характеристика	Январь	Апрель	Июль	Октябрь	Год
$\tilde{r}(0)$	0,91	0,88	0,73	0,81	0,91
η^2	0,10	0,14	0,37	0,24	0,10

Учитывая приведенные в табл. 1 оценки завышения дисперсии по данным М-83, можно было ожидать очень большого занижения пространственной корреляции, рассчитанной по этим данным, характеризующимся значениями $\eta^2 > 2...3$. В действительности корреляция данных ОСО на разных станциях оказывается довольно высокой. Это хорошо видно из табл. 2, где в качестве примера наряду с корреляционными функциями для станций, производящих измерения по спектрофотометру Добсона, приводится корреляционная функция месячных значений ОСО за октябрь для 18 станций, оснащенных М-83. Более того, последняя затухает с расстоянием даже медленнее, чем корреляция данных, полученных по прибору Добсона. Корреляция прослеживается на больших расстояниях, выходящих далеко за пределы представленных в табл. 2 расстояний.

Это обстоятельство свидетельствует, по-видимому, о том, что, по крайней мере до перехода на модернизированные приборы М-83, ошибки наблюдений на этих станциях являлись пространственно связными. Этот существенный факт заслуживает специального изучения. Не исключено, что учет связности ошибок на-

блюдений может оказаться полезным при ретроспективной обработке данных станций, оснащенных М-83. С другой стороны, если нас интересует возможность уточнения данных ОСО при его пространственном осреднении, следует иметь в виду, что в случае пространственной связности ошибок наблюдений точность осреднения существенно уменьшается по сравнению с точностью осреднения при независимых ошибках измерений в различных пунктах.

В заключение отметим, что полученные нами оценки характеристик структуры месячных значений ОСО могут рассматриваться лишь как предварительные. Из них следует, что масштаб r_0 пространственной корреляции месячных значений (1600 км) значительно больше принимаемого обычно для суточных значений (800—1000 км). Однако точность даже средних месячных значений в отдельных пунктах невелика, характеризуясь значением случайных ошибок σ_z порядка 7 Д. Е. (30—60 % естественной изменчивости, характеризующей величиной σ_z).

Для корректного описания географического распределения характеристик изменчивости ОСО необходимо более детальное их исследование. Для этой цели наряду с изучением данных судовых и спутниковых наблюдений представляется целесообразным первоочередное исследование статистической структуры поля ОСО на территории СССР по данным модернизированных приборов М-83, которые публикуются в справочниках [6].

СПИСОК ЛИТЕРАТУРЫ

1. Гриб Н. К., Поляк И. И. Некоторые статистические характеристики измерений атмосферного озона. — Метеорология и гидрология, 1973, № 4, с. 87—93.
2. Гушкин Г. П. Озонометрическая сеть СССР. — Метеорология и гидрология, 1979, № 3, с. 111—116.
3. Журавлева Е. Б., Каган Р. Л., Поляк И. И. Вычисление автокорреляционных и взаимных корреляционных функций по нескольким реализациям случайного процесса. — Труды ГГО, 1971, вып. 289, с. 20—28.
4. Каган Р. Л., Поляк И. И. О расчете характеристик пространственной корреляции метеорологических полей. — Труды ГГО, 1972, вып. 286, с. 18—25.
5. Кайгородцев А. Е., Покровский О. М. О естественных составляющих общего содержания озона в Северном полушарии. — Метеорология и гидрология, 1980, № 8, с. 52—57.
6. Общее содержание атмосферного озона и спектральная прозрачность атмосферы. Справочные данные за 1972—1973 гг., за 1974—1975 гг., за 1976—1977 гг./Под ред. Г. П. Гущина. — Л.: Гидрометеониздат, 1978, 157 с., 1979, 340 с., 1979, 316 с.
7. Перов С. П., Хргиан А. Х. Современные проблемы атмосферного озона. — Л.: Гидрометеониздат, 1980. — 287 с.
8. Уранова Л. А. Карты общего содержания озона за период с 1964 по 1974 г. Приложение к «Синоптическому бюллетеню Северного полушария». — М.: Изд. Гидрометцентра СССР, 1977. — 200 с.
9. Шафрин Ю. А. О статистических характеристиках озоносферы. — Изв. АН СССР. Физика атмосферы и океана, 1966, т. 2, № 6, с. 647—655.

10. Bojkov R. D. Some characteristics of total ozone deduced from Dobson-spectrophotometer and their application to a determination of the effectiveness of the ozone station network.—Ann. Geophys., 1966, vol. 25, N 1, p. 293—299.

11. Bojkov R. D. Differences in Dobson-spectrophotometer and filter ozonometer measurements of total ozone.—J. Appl. Meteorol., 1969, vol. 8, p. 362—368.

12. Fabian P. Eine Abschätzung der räumlicher Ausdehnung einheitlicher Lufipaketen in der unteren Stratosphäre aus Gesamt-ozonmessungen.—Arch. f. Meteorol. Geophys. u. Bioklim., 1967, A, B. 16, H. 4, S. 314—320.

13. London J., Bojkov R. D., Oltmans S. Atlas of the global distribution of total ozone July 1957—June 1967.—NCAR Techn. Note, NCAR/T+ N/113+STR, Boulder, Colorado, 1976.—276 p.

14. Ozone data for the world.—Toronto, Ontario, Canada: WMO, 1961—1977.

15. Wilcox R. W. Total ozone trend significance from space and time variability of daily Dobson data.—J. Appl. Meteorol., 1978, vol. 17, N 3, p. 405—409.

Е. И. Хлебникова

ОБ ИСКАЖЕНИИ РАСПРЕДЕЛЕНИЙ НЕГАУССОВСКИХ ВЕЛИЧИН ПРИ ИХ ПРОСТРАНСТВЕННОЙ ИНТЕРПОЛЯЦИИ

В исследованиях по теории климата в качестве исходных используются, как правило, не сами данные наблюдений, а характеристики, полученные в результате той или иной обработки этих данных: осредненные по времени или пространству значения, интерполированные величины и т. д. Поэтому весьма важной является задача изучения статистической структуры такого рода характеристик как для сравнения их особенностей со свойствами исходных рядов, так и для определения оценок точности этих величин, необходимых для выбора наилучших способов их расчета и получения обоснованных выводов, вытекающих из анализа значений таких параметров.

В случае, когда рассматриваются величины, полученные путем различных линейных преобразований исходных рядов наблюдений, изучение статистических особенностей таких «производных» характеристик применительно к гауссовским метеорологическим элементам сводится к анализу моментов до второго порядка включительно, поскольку характер распределения при таких преобразованиях не меняется. Наибольшие трудности возникают при изучении статистической структуры величин, являющихся производными от рядов наблюдений за метеорологическими элементами, распределение которых отклоняется от гауссовского, поскольку в этом случае характер распределения преобразованных величин отличен от исходного.

В работе [1] даны некоторые сведения о статистике ошибок осреднения негауссовских метеорологических элементов. Более детально вопрос о структуре осредненных величин изучается в работе Р. Л. Кагана и автора [2].

Предметом рассмотрения настоящей статьи, так же как и работы [3], являются величины, полученные путем пространственной интерполяции метеорологических элементов. Если в [3] производится сравнение с исходными таких статистических характеристик интерполированных величин, как дисперсия и корреляционная функция, то в настоящей работе рассматривается вопрос, возникающий при интерполяции негауссовских метеорологических элементов, а именно: вопрос об искажении в результате интерполяции самого характера распределения изучаемого элемента. В данной статье приводятся некоторые количественные оценки такого рода искажения на примере логарифмически-нормального распределения, часто используемого в статистике осадков. Аналитический расчет характеристик искажения представляет значительные математические трудности, поэтому нами для получения соответствующих оценок использовался метод статистического моделирования.

Пусть случайная величина Y имеет логарифмически-нормальное распределение, т. е. ее можно представить в виде

$$Y = b \exp(sX) + c, \quad (1)$$

где X — случайная величина, имеющая нормальное распределение с нулевым средним и единичной дисперсией.

Среднее $\bar{Y}$, дисперсия σ^2 , коэффициенты асимметрии A и эксцесса E величины Y определяются формулами

$$\bar{Y} = b \exp(s^2/2) + c, \quad (2)$$

$$\sigma^2 = b^2 \exp(s^2) (\exp(s^2) - 1), \quad (3)$$

$$A = (\exp(3s^2) - 3 \exp(s^2) + 2) / (\exp(s^2) - 1)^{3/2}, \quad (4)$$

$$E = (\exp(6s^2) - 4 \exp(3s^2) + 6 \exp(s^2) - 3) / (\exp(s^2) - 1)^2 - 3. \quad (5)$$

Поскольку логарифмически-нормальное распределение принадлежит к трехпараметрическому семейству кривых, для его задания можно воспользоваться первыми тремя моментами. Прежде всего, исходя из значения коэффициента асимметрии A , нужно найти параметр s [4]:

$$s = \sqrt{\ln(z^2 + 1)}, \quad (6)$$

где z — корень уравнения

$$z^3 + 3z = A, \quad (7)$$

после чего значения параметров b , c определить из формул (2) и (3) путем задания среднего и дисперсии рассматриваемой величины Y .

Можно показать, что корреляционная функция величины Y связана с корреляционной функцией величины X соотношением

$$r_Y(\rho) = (\exp(s^2 r_X(\rho)) - 1) / (\exp(s^2) - 1). \quad (8)$$

Плотность распределения величины Y , определяемой формулой (1), имеет вид

$$p(y) = \frac{1}{s(y-c)\sqrt{2\pi}} \exp\left\{-\frac{1}{2s^2} \ln^2\left(\frac{y-c}{b}\right)\right\}. \quad (9)$$

Пусть производится линейная интерполяция заданных на станциях значений метеорологического элемента Y , распределение которого может считаться логарифмически-нормальным, в узлы некоторой регулярной сетки точек по формулам вида

$$\bar{Y}_i = \sum_{k=1}^n p_k^i Y_k, \quad (10)$$

где n — число станций, Y_k — значения на станциях, $\bar{Y}_i$ — результат интерполяции в i -й узел сетки, p_k^i — веса интерполяции. Тогда при попытке изучения статистических особенностей поля Y по интерполированным значениям $\bar{Y}$ мы неизбежно получим искаженное представление о распределении значений Y , связанное не только с занижением дисперсии [3], но и с изменением в результате интерполяции самого характера распределения.

Для получения соответствующих количественных оценок такого рода искажений нами был выполнен ряд численных экспериментов, в ходе которых производилось статистическое моделирование значений метеорологического элемента на станциях и в узлах заданной сетки с последующей оценкой распределения интерполированных величин и его моментов и сопоставлением этих характеристик с характеристиками исходного метеорологического поля.

При этом статистическое моделирование логарифмически-нормальной величины Y с заданными тремя моментами распределения, имеющей экспоненциальную корреляционную функцию

$$r_Y(\rho) = \exp(-\rho/\rho_0), \quad (11)$$

где ρ_0 — масштаб корреляции поля, осуществлялось следующим образом. После определения по формулам (6), (7) параметра s моделировалось нормированное гауссовское случайное поле с корреляционной функцией

$$r_X(\rho) = \frac{1}{s^2} \ln(1 + (\exp(s^2) - 1) \exp(-\rho/\rho_0)). \quad (12)$$

Далее по формуле (1) производилось преобразование поля X в поле значений Y с учетом формул (2), (3).

В описанных опытах моделировалась сеть станций различной густоты, в качестве которой задавалась регулярная квадратная

сеть с варьирующимся шагом H . Поскольку характеристики интерполированных величин зависят от того, насколько узел интерполяции близок к опорной станции, для выявления закономерных отличий распределения интерполированных величин от исходного оценки выполнялись нами для группы плотно расположенных узлов интерполяции с последующим получением осредненных характеристик. В качестве алгоритма интерполяции использовался метод оптимальной интерполяции. Интерполяция в узлы сетки производилась по данным четырех станций, при отборе которых в каждом из квадрантов отыскивалась ближайшая станция.

Наряду со случаем отсутствия ошибок наблюдений производились оценки для полей, содержащих шумовую составляющую. Чтобы результирующее поле Y имело логарифмически-нормальное распределение, бессвязные ошибки заданного уровня вводились в гауссовское поле X , а именно, полагалось, что

$$X = \sqrt{1 - \eta^2} X_1 + \eta \dot{X}_2, \quad (13)$$

где X_1 — гауссовское поле с корреляционной функцией вида (12), X_2 — бессвязное гауссовское поле, η^2 — мера ошибок наблюдений величины X .

Таблица 1

Влияние ошибок наблюдений на искажение в результате интерполяции распределения логарифмически-нормальной величины Y ($\bar{Y}=0$, $\sigma^2=1$, $A=0,5$)

Интервал		Функция распределения		
от	до	истинная	после интерполяции при $H=\rho_0$	
			$\eta^2=0$	$\eta^2=0,05$
—3,0	—2,5	0,001	0,000	0,000
—2,5	—2,0	0,008	0,001	0,000
—2,0	—1,5	0,040	0,010	0,008
—1,5	—1,0	0,105	0,066	0,050
—1,0	—0,5	0,174	0,199	0,176
—0,5	0	0,204	0,268	0,282
0	0,5	0,182	0,246	0,250
0,5	1,0	0,130	0,133	0,149
1,0	1,5	0,079	0,055	0,063
1,5	2,0	0,042	0,017	0,018
2,0	2,5	0,020	0,004	0,003
2,5	3,0	0,009	0,001	0,000
3,0	3,5	0,004	0,000	0,000
3,5	4,0	0,001	0,000	0,000
4,0	∞	0,001	0,000	0,000

В этом случае корреляционная функция поля Y (с точностью до членов первого порядка относительно η^2) определяется формулой

$$r_Y(\rho) = \exp(-\rho/\rho_0) - \eta^2 g \ln g / (\exp(s^2) - 1), \quad (14)$$

где

$$g = 1 + (\exp(s^2) - 1) \exp(-\rho/\rho_0).$$

Некоторое представление о влиянии ошибок наблюдений на распределение интерполированных величин дает табл. 1, где для заданной густоты станций приведены распределения, относящиеся к случаям $\eta^2=0$ и $\eta^2=0,05$. Сопоставление этих распределе-

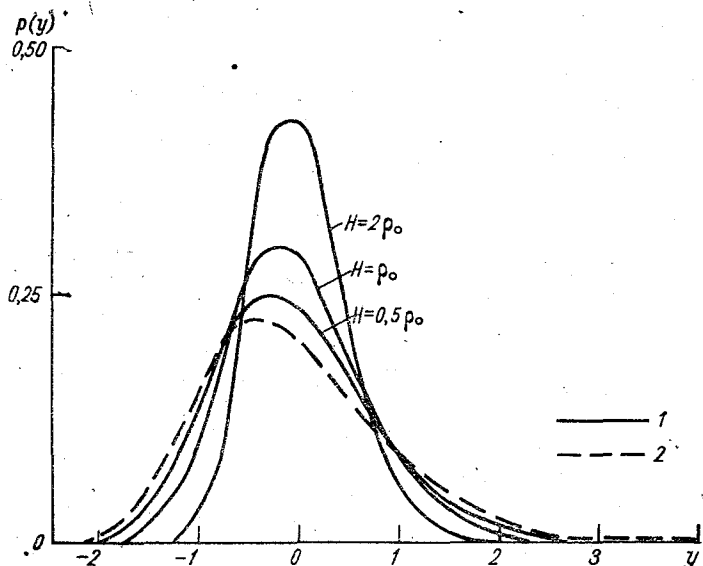


Рис. 1. Влияние густоты сети станций на искажение распределения логарифмически-нормальной величины.

1 — распределения, полученные по результатам интерполяции; 2 — истинное распределение.

ний с истинным показывает, что интерполяция при наличии ошибок наблюдений приводит к большему искажению распределения, чем интерполяция в условиях отсутствия ошибок наблюдений, однако различие данных распределений мало по сравнению с отклонением от истинного.

Наиболее же значительно распределения интерполированных величин зависят от густоты сети станций, используемой при интерполяции. На рис. 1 представлены некоторые результаты расчета распределений по интерполированным данным, относящихся к случаю логарифмически-нормального распределения с коэффи-

циентом асимметрии, равным 1, при мере ошибок наблюдений $\eta^2=0$ (среднее и дисперсия во всех опытах полагались равными соответственно нулю и единице). Эти результаты получены путем обработки не менее чем 1000 реализаций поля величины Y при числе станций $n=25$ (распределение интерполированных значений для $H=0,5 \rho_0$ получено в результате обработки 10 000 реализаций). Изображенные на рисунке кривые 1 представляют собой результат осреднения распределений, полученных для узлов интерполяции, различным образом расположенных по отношению к станциям.

Из рисунка хорошо видно, насколько сильно отличаются от истинного распределения, рассчитанные по интерполированным значениям, даже при сравнительно небольших значениях H/ρ_0 (0,5—1,0). Очень близкие к приведенным зависимости получены нами для распределений с коэффициентом асимметрии $A=0,5$.

Т а б л и ц а 2

Влияние густоты сети станций на искажение моментов распределения логарифмически-нормальной величины

Момент	Истинные значения	По результатам интерполяции		
		$H=0,5 \rho_0$	$H=\rho_0$	$H=2 \rho_0$
σ	1,00	0,87	0,73	0,47
A	1,00	0,83	0,77	0,65
E	1,82	1,26	1,17	0,72

Существенно обратить внимание на то, что различие истинного распределения и распределений, рассчитанных по интерполированным данным, проявляется не только в отличии дисперсий, но и в том, что с увеличением расстояния между станциями, по данным которых производится интерполяция, распределение интерполированных величин довольно быстро приближается к нормальному. О скорости этой сходимости можно судить по данным табл. 2, в которой приведены значения моментов распределений, представленных на рис. 1.

В заключение следует подчеркнуть, что хотя приведенные количественные оценки не дают полного представления о влиянии интерполяции на искажение распределений негауссовских величин, однако они безусловно свидетельствуют о значительности этого влияния и необходимости учета рассматриваемого фактора при проведении исследований, основанных на данных, полученных путем интерполяции (например, при использовании банков данных в узлах регулярной сетки).

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л., Лугина К. М. К вопросу об оценке точности осредненных по площади величин. — Труды ГГИ, вып. 247, 1977, с. 88—98.
2. Каган Р. Л., Хлебникова Е. И. К вопросу о сходимости распределения к нормальному для сумм связанных случайных величин. — Труды ГГО, вып. 431, 1980, с. 106—116.
3. Каган Р. Л., Хлебникова Е. И. О нарушении однородности метеорологических полей при их пространственной интерполяции. — См. наст. сб.
4. Кендалл М., Стьюарт А. Теория распределений. — М.: Наука, 1966.— 587 с.

Ю. М. Либерман

ЭФФЕКТИВНОСТЬ УЧЕТА ВЕТРА ПРИ АНАЛИЗЕ ГЕОПОТЕНЦИАЛА

В работе [2] описан алгоритм объективного анализа геопотенциала, позволяющий учесть дополнительную информацию об этом элементе, содержащуюся в данных о ветре. В настоящей статье представлены результаты испытания алгоритма на нескольких реальных случаях и приведены наиболее существенные сведения технологического характера.

Материалом для испытаний служат наблюдения аэрологических станций северного полушария, поступившие в Гидрометцентр СССР в сентябре 1977 г. в оперативном порядке. Эти данные за 8 последовательных сроков с интервалом 12 ч были записаны Е. А. Локтионовой на магнитную ленту в виде файла, любезно предоставленного в наше распоряжение. Сроки наблюдений и число использованных данных указаны в табл. 1. Все данные относятся к изобарической поверхности 500 гПа.

Таблица 1

Данные для испытаний

Номер срока	Час и дата	Всего данных	В том числе	
			геопотенциал	ветер
1	00 ч 12 IX	484	484	461
2	12 ч 12	365	364	341
3	00 ч 13	396	396	373
4	12 ч 13	406	406	382
5	00 ч 14	393	393	369
6	12 ч 14	409	406	381
7	00 ч 15	414	413	397
8	12 ч 15	404	404	386

Первичный файл многократно обрабатывался по нескольким программам с целью контроля информации. В результате тщательного статико-временного контроля данных о геопотенциале [1] обнаружены 42 грубо-ошибочных значения, из которых 37 удалось исправить, а остальные были исключены. Данные о ветре контролю не подвергались, поскольку алгоритмы контроля ветра находятся пока в стадии разработки. По-видимому, в исходных данных о ветре за каждый срок содержатся 5—10 грубых ошибок. Это обстоятельство необходимо учитывать при оценке наших результатов.

Объективный анализ геопотенциала выполнялся в узлы широтно-долготной сетки с шагом по широте и долготе 5° . Область анализа ограничена широтой 30° ; общее число узлов 864. Сравнивались два варианта анализа: одноэлементный — по данным только о геопотенциале (вариант 1), двухэлементный — с дополнительным привлечением данных о ветре (вариант 2). Оценивались различия двух вариантов анализа и теоретические средние квадратические ошибки оптимальной интерполяции E_1 , E_2 , а также теоретическое уточнение $\tilde{\gamma} = (E_1 - E_2)/E_1$ (индексы 1 и 2 здесь и далее означают номер варианта).

Наряду с анализом в узлы регулярной сетки рассчитывался анализ на станции. В этом случае, помимо теоретических ошибок, можно непосредственно найти эмпирические ошибки сопоставления $\delta_1 = H_1 - H$, $\delta_2 = H_2 - H$ результатов анализа H_1 и H_2 с наблюдением H , их средние квадратические значения $E_1 = \sqrt{\overline{\delta_1^2}}$, $E_2 = \sqrt{\overline{\delta_2^2}}$ и эмпирическое уточнение $\gamma = (E_1 - E_2)/E_1$. Анализ на станции позволяет, кроме того, рассчитать корреляцию фактического и проанализированного полей и оценить пространственное сглаживание.

В варианте 2 при анализе в каждый узел сетки ветер привлекался в дополнение к геопотенциалу, если теоретическая ошибка оптимальной интерполяции геопотенциала в данный узел ε превышала критическое значение $\varepsilon_{кр} = 0,20$; в противном случае выполнялся анализ только по геопотенциалу. Иначе строился анализ на станции: здесь был принят порог $\varepsilon_{кр} = 0$, т. е. привлечение ветра предусматривалось независимо от точности анализа геопотенциала. На основании предварительного эксперимента [2] ошибка наблюдений геопотенциала Δ была задана равной 2,0 гп. дам, а параметр λ , связывающий меры ошибок наблюдений геопотенциала и ветра, равная 3.

Рассмотрение результатов начнем с распределения узлов и станций по градациям ε . Это распределение, осредненное по 8 случаям, представлено в табл. 2. Уменьшение доли высоких значений за счет привлечения ветра как дополнительного источника информации является естественным: оно обеспечивается основным свойством оптимальной интерполяции. По той же причине снижается теоретическая абсолютная ошибка интерполяции

в узлы. Это снижение при осреднении по всем случаям характеризуется величинами $\bar{E}_1=4,2$ гп. дам, $\bar{E}_2=3,9$ гп. дам, $\gamma=8\%$.

В тех узлах, где анализ является двухэлементным, значения меняются в варианте 2 по сравнению с вариантом 1, как правило, на 1—2 гп. дам. В немногих узлах, находящихся в циклонических

Таблица 2

Распределение узлов и станций по градациям ϵ (%)

Градации ϵ	Узлы		Станции	
	Вариант 1	Вариант 2	Вариант 1	Вариант 2
0,0—0,1	5	5	2	3
0,1—0,2	44	51	61	71
0,2—0,3	19	22	25	18
0,3—0,4	13	9	7	4
0,4—0,5	8	5	2	2
0,5—0,6	4	2	1	1
0,6—1,0	7	6	2	1

областях и зонах струйных течений, изменения составляют 3—5 гп. дам, изредка превышая эти значения. Максимальные изменения достигают 13—15 гп. дам; они обнаружены в зонах особенно сильных ветров.

Особый интерес представляет сопоставление результатов интерполяции с данными наблюдений станций. Такое сопоставление

Таблица 3

Теоретические и эмпирические оценки двух вариантов анализа для станций

	Номер срока								Сред- нее
	1	2	3	4	5	6	7	8	
Теоретические оценки									
$\tilde{E}_1$ гп. дам	3,0	3,2	3,2	3,1	3,2	3,1	3,1	3,1	3,1
$\tilde{E}_2$ гп. дам	2,8	3,0	2,9	2,9	2,9	2,9	2,9	2,9	2,9
$\tilde{\gamma}$ %	7	6	9	6	9	6	6	6	7
Эмпирические оценки									
E_1 гп. дам	3,0	3,0	3,5	3,3	3,2	3,3	3,7	2,8	3,2
E_2 гп. дам	2,7	2,5	3,0	2,9	2,8	3,0	3,2	2,5	2,8
γ %	10	17	14	12	12	9	14	11	12

выполнено для всех фактических значений геопотенциала, общее число которых равно 3261. Полученные эмпирические оценки наряду с теоретическими приводятся в табл. 3 для каждого срока и в среднем по всей выборке. Они показывают, что привлечение ветра обеспечивает в среднем заметное уменьшение эмпирических ошибок. Этот отнюдь не тривиальный вывод подтверждается также распределением ошибок сопоставления по грациям (табл. 4).

Уточняющее воздействие ветра в процессе объективного анализа непосредственно связано с плотностью сети [2]. Оно слабо

Таблица 4

Распределение эмпирических ошибок

Градации δ гп. дам	Доля в процентах		Число случаев	
	Вариант 1	Вариант 2	Вариант 1	Вариант 2
0—2	64	67	2078	2169
2—4	22	24	715	767
4—6	9	6	293	218
6—8	3	2	94	67
8—10	1	—	33	14
>10	1	1	48	26

выражено как в районах густой сети, где ветер не содержит информации, дополнительной по отношению к геопотенциалу, так и в районах редкой сети, где данные практически отсутствуют. Но при средней плотности сети можно рассчитывать на существенное повышение точности. Среди всех станций полушария мы выделили три группы в зависимости от рассчитанного по геопотенциалу значения ϵ , именно станции, для которых $0 < \epsilon \leq 0,2$; $0,2 < \epsilon \leq 0,4$; $0,4 < \epsilon \leq 1$. Разумеется, в каждую группу вошли станции, расположенные на полушарии беспорядочным образом.

Таблица 5

Оценки анализа при разной плотности сети

Интервал ϵ	Объем данных		Теоретические оценки			Эмпирические оценки		
	число	%	$\tilde{E}_1$ гп. дам	$\tilde{E}_2$ гп. дам	$\tilde{\gamma}$ %	E_1 гп. дам	E_2 гп. дам	γ %
0,0—0,2	2043	63	2,8	2,7	4	2,7	2,6	5
0,2—0,4	1052	32	3,6	3,2	12	3,7	3,0	19
0,4—1,0	166	5	4,4	4,0	9	5,4	5,2	5
0,0—1,0	3261	100	3,1	2,9	7	3,2	2,8	12

Тем не менее эти группы характеризуют соответственно густую, среднюю и редкую аэрологическую сеть. Средние взвешенные оценки по группам, представленные в табл. 5, показывают, что реальное уточнение γ при средней плотности сети составляет 15—

20 %. Близость его к теоретическому уточнению γ подтверждает правильность выбора основных параметров нашей схемы анализа.

Эффективность привлечения информации о ветре к анализу геопотенциала можно проиллюстрировать следующим примером. Во всех сроках (с 1 по 8) ошибка сопоставления δ_1 превысила по абсолютной величине 8 гп. дам 81 раз. При анализе по варианту 2 учет ветра заметно (более чем на 1 гп. дам) изменил эту ошибку 51 раз. В 49 случаях из этих 51 ошибка сопоставления уменьшилась: $|\delta_1| < |\delta_2|$. При этом в 24 случаях произошло обострение реально существующего барического максимума и в 11 случаях углубилась реальная барическая ложбина.

Рассмотрим сглаживающий эффект оптимальной интерполяции подробнее. Этот эффект является следствием некоррелированности ошибки анализа в точку 0

$$\delta = H'_0 - H' \quad (1)$$

с результатом интерполяции H'_0 :

$$\overline{H'_0 \delta} = 0 \quad (2)$$

(штрих означает отклонение от нормы). Умножив (1) на δ и выполнив осреднение с учетом (2), легко убедиться в отрицательной корреляции ошибок анализа с истинным значением элемента:

$$\overline{H' \delta} = -\overline{\delta^2}. \quad (3)$$

Возведем (1) в квадрат:

$$\overline{H'^2_0} + \overline{H'^2} - 2\overline{H'_0 H'} = \overline{\delta^2}. \quad (4)$$

Подставив в (4) вытекающее из (1) и (2) соотношение

$$\overline{H'_0 H'} = \overline{H'^2_0}, \quad (5)$$

получим равенство

$$\overline{H'^2_0} = \overline{H'^2} - \overline{\delta^2}, \quad (6)$$

которое можно записать в виде

$$\sigma^2_0 = \sigma^2(1 - \varepsilon^2). \quad (7)$$

Последнее выражение показывает, что дисперсия проанализированного поля меньше фактической дисперсии. В этом и заключается сглаживание, свойственное оптимальной интерполяции. Разумеется, любой иной способ интерполяции также обладает свойством сглаживания. Но применительно к оптимальной интерполя-

ции удается весьма просто оценить этот эффект. Именно, как видно из (7), количественным показателем сглаживания служит непосредственно мера ошибки оптимальной интерполяции ε^2 .

Нас, однако, интересует не теоретическая, а эмпирическая оценка. Удобной эмпирической оценкой служит коэффициент корреляции фактических значений геопотенциала и ошибок анализа $r_{H,\delta}$. Разделив равенство (3) на $\sigma_0 \sqrt{\delta^2}$ получим соотношение

$$r_{H,\delta} = -\varepsilon, \quad (8)$$

показывающее, что этот коэффициент всегда отрицателен.

Рассчитанные для густой, средней и редкой сети средние взвешенные значения коэффициентов корреляции приведены в табл. 6.

Коэффициенты r_1 и r_2 характеризуют связь фактических значений геопотенциала на станциях H с ошибками анализа без учета ветра (δ_1) и с учетом его (δ_2). Коэффициенты r_3 описывают связь H с поправками к анализу, обусловленными учетом ветра, т. е. с разностями ошибок двух вариантов анализа $\delta H = \delta_2 - \delta_1$. Положительность коэффициентов r_3 означает, что в среднем положительным фактическим аномалиям H' отвечает положительная поправка δH . Эта статистическая связь подтверждает, что привлечение ветра действительно является компенсирующим фактором по отношению к сглаживающему эффекту интерполяции. О том же говорит и уменьшение абсолютных значений r_2 по сравнению с r_1 в районах густой и средней сети. Однако, в районах редкой сети, где сглаживание в процессе анализа особенно велико, воз-

Таблица 6

**Корреляция фактических значений
геопотенциала с ошибками анализа
при разной плотности сети**

Интервал ε	r_1	r_2	r_3
0,0—0,2	—0,24	—0,19	0,26
0,2—0,4	—0,21	—0,19	0,16
0,4—1,0	—0,49	—0,49	0,13
0,0—1,0	—0,22	—0,19	0,18

Таблица 7

Пример интерполяции (станция 296, срок 3)

Влияющие станции	p	H'	$\frac{\partial H'}{\partial x}$	$\frac{\partial H'}{\partial y}$	a	$\tilde{a}$	b	c
272	502	—3,7	23,9	—20,0	0,556	0,372	—0,154	0,032
295	626	—3,8			0,139	0,195		
297	718	1,6	—19,9	—10,8	0,251	0,259	0,072	0,062
273	815	7,4			0,224	0,154		
396	921	6,6			—0,149	0,036		
271	1002	7,3			0,037	0,001		

действие ветра выражено слабо, и рассчитывать на уточнение анализа не приходится.

Рассмотрим характерный пример интерполяции геопотенциала на контрольную станцию. Относящиеся к этому примеру данные представлены в табл. 7 и на рис. 1, где станции обозначены условными номерами. В табл. 7 используются следующие обозначения: ρ — расстояние между влияющей и контрольной станциями в километрах; H' — фактическая аномалия геопотенциала на влияю-

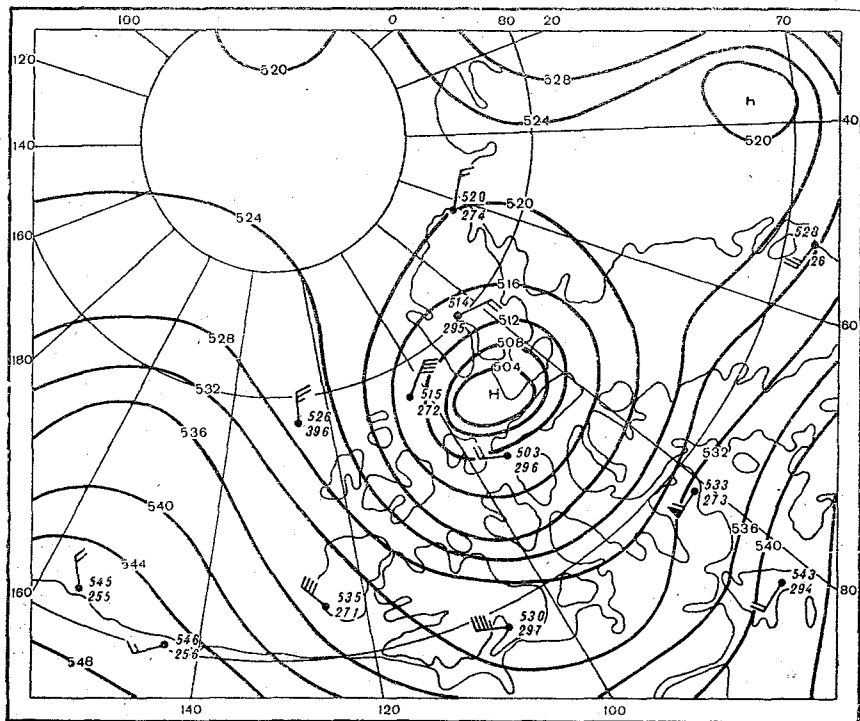


Рис. 1. К интерполяции на станцию 296. Срок 3.

щей станции в геопотенциальных декаметрах; $\frac{\partial H'}{\partial x}$, $\frac{\partial H'}{\partial y}$ — фактические аномалии градиентов геопотенциала на влияющей станции в безразмерных (отнесенных к масштабу корреляции L) координатах; a — вес геопотенциала в анализе по варианту 1; $\bar{a}$ — вес геопотенциала в анализе по варианту 2; b , c — веса градиентов $\frac{\partial H'}{\partial x}$, $\frac{\partial H'}{\partial y}$, в анализе по варианту 2.

В рассматриваемом примере интерполированная на станцию 296 аномалия H' , равная в одноэлементном варианте $-1,2$ гп. дам,

после привлечения ветра оказалась равной—6,8 гп. дам. Результат интерполяции H_0 изменился с 520 гп. дам на 515 гп. дам, а значение ε уменьшилось с 0,256 до 0,196. При этом ошибка сопоставления $\delta = H_0 - H$ снизилась с 17 гп. дам до 12 гп. дам. Этот пример иллюстрирует углубление циклона, происходящее при должном учете ветра. Здесь в обоих вариантах анализа использованы 6 влияющих станций. В одноэлементном варианте станция 396 оказалась экранированной станцией 272, расположенной между нею и контрольной станцией 296. В двухэлементном варианте учет ветра на станциях 272 и 297 снял экранирование и обеспечил уточнение анализа. Данные о ветре на этих станциях хорошо отражают циклонический характер циркуляции. Они вносят существенный вклад в проинтерполированное значение H_0 : как легко убедиться, он равен —6,7 гп. дам. Вероятно, точность анализа возросла бы в еще большей степени при дополнительном учете сильного ветра на станции 273, находящейся на южной периферии циклона. Однако включение этой информации повысило бы порядок системы до 12, что не допускается нашим алгоритмом.

Опишем некоторые особенности алгоритма, наиболее существенные для его реализации в оперативных условиях. Начнем с поиска влияющих станций, принципы которого изложены в работе [2].

Для двухэлементной оптимальной интерполяции в узел сетки требуется найти n станций с данными о геопотенциале и k станций с данными о ветре. Их поиск выполняется по квадратам так называемой машинной карты со стороной 750 км. Возможны три последовательных цикла обработки квадратов. Первый цикл представляет собой обработку одного центрального квадрата зоны поиска, т. е. того квадрата, в котором находится анализируемый узел. Во втором цикле обрабатывается расширенная область, состоящая из $3 \times 3 = 9$ квадратов, а в третьем — вся зона поиска, содержащая $5 \times 5 = 25$ квадратов. Переход к очередному циклу не происходит, если комплекс влияющих станций уже сформирован полностью в предшествующем цикле.

Каждый цикл включает следующие стадии обработки.

1. Выборка сведений о станциях, находящихся в квадратах рассматриваемой области. На этой стадии декартовы координаты станции x, y , значение геопотенциала H и компоненты ветра u, v поступают из общих информационных массивов в рабочие поля блока поиска.

2. Расчет расстояния R между узлом и станциями с учетом масштаба карты.

3. Упорядочение станций (т. е. принадлежащих станциям значений x, y, H, u, v) по возрастанию R .

4. Проба на наличие станции, для которой $R \leq 50$ км. Если такая станция найдена, то ее наблюдение H присваивается узлу без интерполяции.

5. Исключение из дальнейшего анализа станций, для которых $R > 2200$ км.

6. Отбор станций с данными о геопотенциале с учетом их симметричного расположения по отношению к узлу.

Если после этого число найденных станций оказывается менее 6, то осуществляется переход ко второму или третьему циклу, т. е. расширяется область поиска.

Таким образом отыскивается не более 6 влияющих станций с данными о геопотенциале. Разумеется, в районах редкой и средней сети число таких станций в действительности оказывается меньшим. Что касается станций с данными о ветре, то для каждого узла мы стремимся найти их не менее 2 и не более 5. Поиск таких станций предусмотрен на стадии 3: одновременно с упорядочением по расстоянию отбираются те станции, которые находятся в пределах «ветрового кольца» $500 \text{ км} < R < 1500 \text{ км}$. Требование симметричности к ним не применяется.

Как правило, в данных аэрологической станции содержатся сведения и о геопотенциале, и о ветре. Приняв для k и n наибольшие допустимые значения, соответственно $k_{\text{пр}}=5$ и $n_{\text{пр}}=6$, мы добились существенной экономии времени поиска. Действительно, поскольку $k_{\text{пр}} < n_{\text{пр}}$, а отбор каждой станции по ветру и геопотенциалу происходит одновременно, то время двухэлементного поиска лишь немногим больше одноэлементного. С иной, более трудной ситуацией пришлось бы иметь дело при $k_{\text{пр}} > n_{\text{пр}}$. В таком случае к моменту окончания поиска по геопотенциалу нужное число станций по ветру еще не будет найдено, так что потребуются расширение области. Так же обстоит дело, если среди исходных данных велика доля станций, наблюдающих лишь один из двух элементов: либо геопотенциал, либо ветер.

Для экономии времени счета важно ограничить число влияющих станций минимально необходимым. Выбор $n_{\text{пр}}=6$ обоснован многолетним опытом оперативного анализа геопотенциала. Для ветра же мы приняли условие $2 \leq k \leq 5$ с тем, чтобы порядок двухэлементной системы $n+2k$ не превосходил 11. Число станций — как по геопотенциалу, так и по ветру — ограничено только сверху. Естественным нижним пределом, при котором интерполяция не производится, является полное отсутствие станций ($n=k=0$); в любом другом случае интерполяция происходит (если только не обнаружена станция, отстоящая от узла не далее 50 км). В частности, возможно определение геопотенциала в узле по станционным данным только о ветре ($n=0, k \neq 0$).

Описанный алгоритм объективного анализа реализован на ЭВМ БЭСМ-6 на базе оптимизирующего алгольного транслятора. Он требует в среднем для обработки одного узла 0,10 с в одноэлементном варианте и 0,14 с в двухэлементном.

На рис. 2 заштрихованы расположенные севернее параллели 30° районы средней плотности сети, где целесообразен двухэлементный анализ. Это примыкающие к континентам части океанических акваторий, почти весь Центральный полярный бассейн и плохо освещенные области суши. Доля узлов, находящихся в указанных районах, составляет 32 % общего их числа 864

(см. табл. 2). Остановимся на возможностях расширения области анализа южнее 30° с. ш. Такое расширение весьма желательно. Действительно, в зоне $10-30^\circ$ с. ш. находятся обширные районы со средней плотностью аэрологической сети, для которых учет ветра является эффективным средством уточнения анализа. Расширение области анализа до параллели 10° увеличит долю узлов, приходящихся на районы редкой сети, с нынешних 19 % прибли-

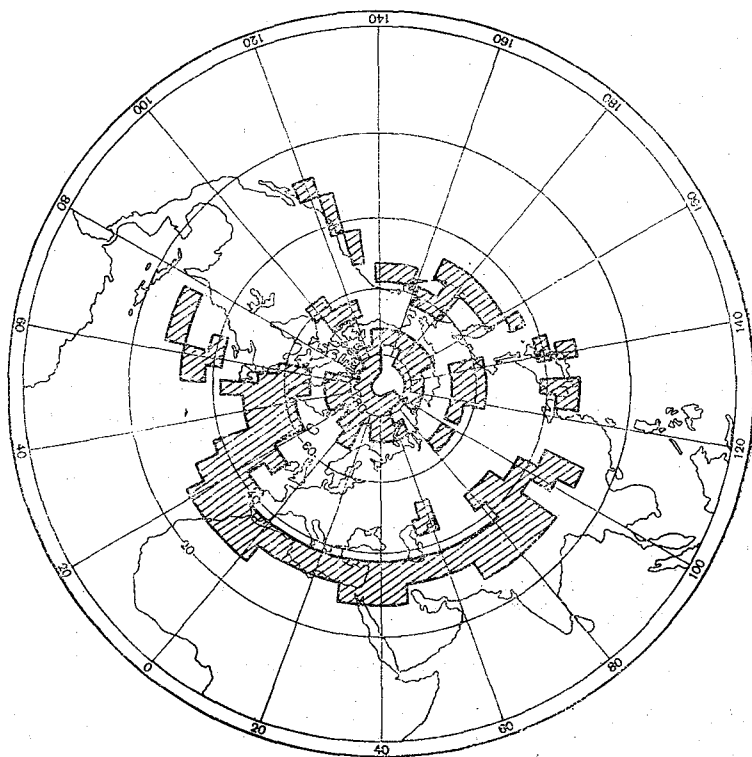


Рис. 2. Районы средней плотности сети.

зительно до 33 %, а доля узлов в районах густой сети составит в этом случае около 37 %. Привлечение ветра потребует, таким образом, лишь для анализа тех 30 % узлов, которые располагаются в районах средней сети. Естественно, все такие узлы должны быть заданы априори: это позволит избежать предварительной интерполяции для расчета величины ε . Привлечение ветра увеличит среднее время обработки одного узла не более чем в 1,5 раза и, следовательно, ожидаемое увеличение общего времени счета составит лишь около 15 % по сравнению с одноэлементным анализом. Вероятно, некоторая дополнительная экономия возможна

за счет рационализации начальных и конечных блоков алгоритма, не связанных непосредственно с оптимальной интерполяцией.

Испытания алгоритма двухэлементного анализа показали, что на его основе возможно повышение точности анализа на 15—20 % в значительных областях полушария со средней плотностью аэрологической сети. По-видимому, оперативная реализация такого анализа наиболее актуальна применительно к Северной Атлантике. Уточнение анализа геопотенциала путем использования ветровой информации заключается, в первую очередь, в снижении больших ошибок, число которых невелико, но искажающее влияние весьма значительно. Происходящее при этом обострение циклонов, барических ложбин и гребней заметно приближает результаты анализа к реальному барическому полю.

Отметим еще одно потенциальное преимущество двухэлементного анализа. Результатом такого анализа оказывается поле геопотенциала, построенное с учетом ветровой информации. На тех же принципах можно реализовать анализ ветра с привлечением фактических данных о геопотенциале. Проанализированные таким образом поля геопотенциала и ветра будут в значительной степени согласованы друг с другом, поскольку в процессе интерполяции существенно используются статистические связи обоих полей. Упомянутое согласование не является ни геострофическим, ни модельно зависимым. Можно надеяться, что оно сократит время инициализации исходных полей в цикле численного прогноза. Тем самым в некоторой степени компенсируется увеличение времени анализа.

Наконец, подчеркнем, что успешность двухэлементного анализа в значительной степени определяется качеством данных о ветре. Скорейшую разработку алгоритмов контроля ветра следует считать первоочередной задачей.

СПИСОК ЛИТЕРАТУРЫ

1. Либерман Ю. М. Алгоритм комплексного статико-временного контроля аэрологической информации. — Труды ГГО, 1980, вып. 431, с. 23—30.
2. Либерман Ю. М. Объективный анализ поля геопотенциала с использованием данных о ветре. — Труды Гидрометцентра СССР, 1980, вып. 242, с. 22—34.

Р. Л. Казан

К ВОПРОСУ ОБ АППРОКСИМАЦИИ ВРЕМЕННОЙ СТРУКТУРЫ БАРИЧЕСКОГО ПОЛЯ В ТРОПОСФЕРЕ

1. В «Справочнике по временным корреляционным связям метеорологических элементов в свободной атмосфере над СССР» [9] представлен обширный материал по временной корреляции

полей давления, температуры воздуха и скорости ветра в свободной атмосфере. Некоторые предварительные результаты его анализа приводятся в [6], однако для полного его освоения требуется еще значительное время. Одним из этапов такого освоения, на наш взгляд, должен явиться выбор вероятностных моделей, которые позволяли бы в аналитической форме описывать основные особенности статистической структуры исследуемых метеорологических полей.

Разумеется, нельзя рассчитывать в единой модели на удовлетворительное описание не только структуры всех полей, но даже структуры одного поля во всех возможных условиях. Например, как следует из [6], временная корреляция поля геопотенциала существенно различна в тропосфере и стратосфере. Заметно различается характер затухания корреляции со временем в течение первых суток-полтора и на четвертые-пятые сутки. Поэтому вопрос об аналитическом описании временной корреляционной функции этого элемента должен решаться отдельно для различных интервалов времени, а также в зависимости от задачи, для решения которой предполагается использование принятой модели.

Мы ограничимся далее подбором такой аналитической аппроксимации формулы для временной корреляционной функции поля геопотенциала (давления) в тропосфере, которая удовлетворительно описывала бы временную структуру этого поля для сдвигов времени в пределах полутора-двух суток. Наличие такой аппроксимации существенно облегчило бы решение ряда задач, связанных с четырехмерным объективным анализом [7], контролем аэрологических данных [10], их временным осреднением [2, 5] и т. д. Возможно также, что она могла бы в дальнейшем явиться основой для аналитического описания структуры других метеорологических полей.

2. В работе [7] было предложено описывать пространственно-временную корреляцию барического поля однородной и изотропной корреляционной функцией значений поля давления (геопотенциала) в различных точках пространства «горизонтальные координаты — время»

$$r(d) = (1 + d) \exp(-d). \quad (1)$$

Здесь

$$d = V(\rho/\rho_0)^2 + (\tau/T_0)^2, \quad (2)$$

$\rho = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ — расстояние между точками; $\tau = |t_1 - t_2|$ — временной сдвиг между ними; x_1, y_1, t_1 и x_2, y_2, t_2 — горизонтальные координаты и время для первой и второй точек соответственно; ρ_0 — масштаб пространственной корреляции, T_0 — масштаб временной корреляции.

В [3, 4] было показано, что модель (1) — (2) может быть уточнена путем учета наклона осей барических образований, анизотропии пространственной корреляции, а также смещения области максимальной корреляции давления в разных пунктах с увеличе-

нием сдвига во времени. Однако на данном этапе учет этих факторов представляется нам преждевременным.

В рамках модели (1)–(2) получаем формулу для аппроксимации временной корреляции барического поля

$$r(\tau) = (1 + \tau/T_0) \exp(-\tau/T_0). \quad (3)$$

Как следует из [4, 8], масштаб временной корреляции T_0 мало зависит от сезона и составляет около 30 ч. Нужно, однако, иметь в виду, что эти выводы, строго говоря, относятся лишь к полю приземного давления на Северо-Западе ЕТС и в Скандинавии. Применимость их к другим районам и к свободной атмосфере требует еще проверки.

Формула (3) предполагает расчет корреляции по точным значениям поля давления при отсутствии ошибок измерений. Фактически, при наличии случайных ошибок измерений эмпирическая корреляционная функция $\tilde{r}(\tau)$ при $\tau > 0$ отличается от $r(\tau)$ на постоянный множитель

$$\tilde{r}(0) = \frac{1}{1 + \eta^2}, \quad (4)$$

где η^2 — мера ошибок измерений (отношение дисперсии ошибок к дисперсии исследуемой величины), так что

$$\tilde{r}(\tau) = \tilde{r}(0) r(\tau). \quad (5)$$

Соответственно $\tilde{r}(0)$ представляет собой значение эмпирической корреляционной функции, полученное путем экстраполяции ее до значения $\tau = 0$.

В нашем случае

$$\tilde{r}(\tau) = \tilde{r}(0) \left(1 + \frac{\tau}{T_0}\right) \exp(-\tau/T_0). \quad (6)$$

Таким образом, в предположении справедливости нашей модели задача аппроксимации временной корреляции для заданного уровня на некоторой станции сводится к подбору двух параметров (T_0 и $\tilde{r}(0)$) по данным об эмпирической корреляционной функции $\tilde{r}(\tau)$ для нескольких сдвигов времени. Вероятно, наибольшей точности аппроксимации можно было бы добиться путем использования максимального числа сдвигов (в [9] данные о временной корреляции приводятся для сдвигов времени τ до 5 сут с дискретностью 6 ч), используя, например, метод наименьших квадратов с должным уменьшением весов при увеличении τ . Однако для этого необходимо специальное исследование. На данном этапе представляется возможным использование двух сдвигов во времени, τ_1 и $\tau_2 = 2\tau_1$.

Полагая известными $\tilde{r}_1 = \tilde{r}(\tau_1)$ и $\tilde{r}_2 = \tilde{r}(2\tau_1)$, имеем

$$\tilde{r}_2/\tilde{r}_1 = \tilde{\varphi}(\tau_1), \quad (7)$$

где

$$\begin{aligned}\tilde{\tau} &= \tau/T_0, \quad \tilde{\tau}_1 = \tau_1/T_0, \\ \varphi(\tilde{\tau}) &= \frac{1+2\tilde{\tau}}{1+\tilde{\tau}} \exp(-\tilde{\tau}).\end{aligned}\quad (8)$$

Зависимость φ и значений корреляционной функции $r(\tau)$ от отношения $\tilde{\tau}_1 = \tau_1/T_0$ представлена в табл. 1. Это позволяет установить однозначное соответствие как между φ и $\tilde{\tau}$, так и между

Таблица 1

Зависимость корреляционной функции r и отношения φ
(в тысячных долях единицы) от безразмерного сдвига времени $\tilde{\tau} = \tau/T_0$

$\tilde{\tau}$	$r(\tilde{\tau})$	$\varphi(\tilde{\tau})$	$\tilde{\tau}$	$r(\tilde{\tau})$	$\varphi(\tilde{\tau})$	$\tilde{\tau}$	$r(\tilde{\tau})$	$\varphi(\tilde{\tau})$
0	1000	1000	0,85	791	624	1,70	493	298
0,05	999	996	0,90	772	599	1,75	478	284
0,10	995	984	0,95	754	575	1,80	463	272
0,15	990	973	1,00	736	552	1,85	448	259
0,20	982	955	1,05	717	529	1,90	434	248
0,25	974	935	1,10	699	507	1,95	420	236
0,30	963	912	1,15	681	486	2,00	406	226
0,35	951	887	1,20	663	465	2,05	393	215
0,40	938	862	1,25	645	446	2,10	380	205
0,45	925	835	1,30	627	427	2,15	367	196
0,50	910	809	1,35	609	408	2,20	355	187
0,55	894	782	1,40	592	391	2,25	342	178
0,60	878	755	1,45	575	373	2,30	331	170
0,65	861	728	1,50	558	357	2,35	320	162
0,70	824	701	1,55	542	341	2,40	308	155
0,75	827	675	1,60	525	326	2,45	298	147
0,80	809	649	1,65	509	312	2,50	287	141

φ и r . В табл. 2 (см. также рис. 1) эта зависимость представлена в виде функции от φ . Этой таблицей можно пользоваться для оценки $r(\tau)$ по полученным отношениям $\varphi(\tau_1) = \tilde{r}(2\tau_1)/\tilde{r}(\tau_1)$.

Сравнивая эту оценку $r(\tau_1)$ с эмпирическими значениями можно получить

$$\tilde{r}(0) = \tilde{r}(\tau_1)/r(\tau_1). \quad (9)$$

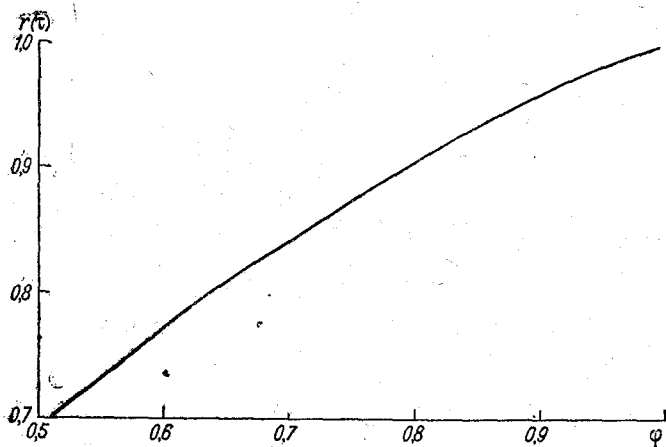


Рис. 1. Зависимость $r(\tau)$ от $\varphi = r(2\tau)/r(\tau)$ для корреляционной функции $r(\tau) = (1 + \tau/T_0) \exp(-\tau/T_0)$.

Таблица 2

Зависимость корреляционной функции r и безразмерного сдвига времени τ (в сотых долях единицы) от отношения

$$\varphi(\tilde{\tau}) = r(2\tilde{\tau})/r(\tilde{\tau})$$

φ	$r(\varphi)$	$\tilde{\tau}(\varphi)$	φ	$r(\varphi)$	$\tilde{\tau}(\varphi)$	φ	$r(\varphi)$	$\tilde{\tau}(\varphi)$
100	100	0	83	92	46	66	82	78
99	100	9	82	92	48	65	81	80
98	99	13	81	91	50	64	80	82
97	99	16	80	90	52	63	79	84
96	98	19	79	90	53	62	79	86
95	98	21	78	89	55	61	78	88
94	98	24	77	89	57	60	77	90
93	97	26	76	88	59	59	77	92
92	97	28	75	88	61	58	76	94
91	96	30	74	87	63	57	75	96
90	96	32	73	86	65	56	74	98
89	95	34	72	86	66	55	73	100
88	95	36	71	85	68	54	72	103
87	94	38	70	84	70	53	72	105
86	94	40	69	83	72	52	71	107
85	93	42	68	83	74	51	70	109
84	93	44	67	82	76	50	69	112

Оценка масштаба временной корреляции получается по формуле

$$T_0 = \tau_1 / \tilde{\tau}, \quad (10)$$

где $\tilde{\tau}$ — значение отношения τ/T_0 , соответствующее в табл. 1 эмпирическому значению $\varphi(\tau_1)$.

Таблица 3

Корреляционная функция $r(\tau)$ для различных масштабов корреляции

$T_0, ч$	$\tau (ч)$							
	6	12	18	24	30	36	42	48
20	0,96	0,88	0,77	0,66	0,56	0,46	0,38	0,31
21	0,97	0,89	0,79	0,68	0,58	0,49	0,41	0,33
22	0,97	0,90	0,80	0,70	0,61	0,51	0,43	0,36
23	0,97	0,90	0,82	0,72	0,63	0,53	0,45	0,38
24	0,97	0,91	0,83	0,74	0,64	0,56	0,48	0,41
25	0,98	0,92	0,84	0,75	0,66	0,58	0,50	0,43
26	0,98	0,92	0,85	0,77	0,68	0,60	0,52	0,45
27	0,98	0,93	0,85	0,78	0,70	0,62	0,54	0,47
28	0,98	0,93	0,86	0,79	0,71	0,63	0,56	0,49
29	0,98	0,94	0,87	0,80	0,72	0,65	0,57	0,51
30	0,98	0,94	0,88	0,81	0,74	0,66	0,59	0,52
31	0,98	0,94	0,88	0,82	0,75	0,68	0,61	0,54
32	0,98	0,94	0,89	0,83	0,76	0,69	0,62	0,56
33	0,99	0,95	0,89	0,83	0,77	0,70	0,64	0,57
34	0,99	0,95	0,90	0,84	0,78	0,71	0,65	0,59
35	0,99	0,95	0,91	0,85	0,79	0,72	0,66	0,60
36	0,99	0,96	0,91	0,85	0,80	0,74	0,67	0,62
37	0,99	0,96	0,91	0,86	0,81	0,75	0,68	0,63
38	0,99	0,96	0,92	0,87	0,81	0,75	0,70	0,64
39	0,99	0,96	0,92	0,87	0,82	0,77	0,71	0,65
40	0,99	0,96	0,92	0,88	0,83	0,77	0,72	0,66

В табл. 3 представлены значения корреляционной функции $r(\tau)$, соответствующие различным масштабам корреляции T_0 .

Приведенные в табл. 2 и 3 величины даны с точностью до одной сотой. Этой точности достаточно для анализа данных о временной корреляции, представленных в [9] также в сотых долях единицы. Следует иметь в виду, однако, что реальная точность этих эмпирических корреляционных функций в большинстве слу-

часов гораздо меньше. В самом деле, из табл. 2 Справочника [9], в которой представлены доверительные интервалы выборочных значений коэффициентов корреляции для различного числа независимых пар наблюдений, следует, что при фактически использованных объемах выборки (четыре наблюдения в сутки в течение 10 лет, что дает для каждого месяца при отсутствии пропусков около 1200 наблюдений) доверительный интервал при $r \leq 0,9$ составляет несколько сотых.

Поскольку ряды наблюдений являются связными, выборочные ошибки должны быть существенно больше. Мы не располагаем в настоящее время оценками выборочных погрешностей значений временных корреляционных функций специально для поля давления, однако можем примерно судить об их порядке по соответствующим оценкам, полученным в ГГО для скорости ветра и температуры.

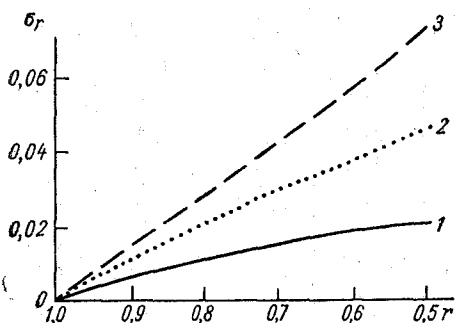


Рис. 2. Средняя квадратическая погрешность выборочного коэффициента корреляции при объеме выборки, равном 1200 членам.

1 — бесвязные ряды, 2 — четырехсрочные наблюдения за скоростью ветра, 3 — четырехсрочные наблюдения за температурой воздуха.

На рис. 2 представлена зависимость средней квадратической ошибки коэффициента корреляции от его значений, определенных по 1200 парам независимых измерений, а также в случае использования связных рядов того же объема, корреляция которых соответствовала бы характерной временной корреляции полей скорости ветра и температуры воздуха при шестичасовом интервале времени между наблюдениями.

Из рисунка видно, что погрешности коэффициентов корреляции для этих элементов примерно в два и три раза соответственно превышают погрешности для независимых рядов. До получения более точных оценок можно, по-видимому, считать, что погрешности выборочной корреляционной функции для поля давления близки к таковым для поля температуры. В этом случае средние квадратические погрешности коэффициентов корреляции для использованного объема выборки в диапазоне $0,5 \leq r \leq 1$ могут оцениваться по формуле

$$\sigma_r \approx 0,14(1 - r). \quad (11)$$

Соответственно и доверительные интервалы в табл. 2 Справочника [9] должны быть увеличены примерно в три раза.

Из (11) следует, что при $\tilde{r} < 0,9$ даже средние квадратические выборочные ошибки коэффициентов корреляции больше 0,01. Однако при значениях $r > 0,95$, которые вполне реальны для при-

земного давления при $\tau \leq 6$ ч, значение σ_r составляет несколько тысячных. Представление этих данных с двумя знаками после запятой заметно загроуляет их.

3. В качестве примера использования изложенной методики аппроксимации проанализируем приведенные в [9] данные о временной корреляции поля геопотенциала для Москвы. При этом ограничимся рассмотрением тропосферных уровней от Земли до 300 гПа включительно. Данные для стратосферных уровней требуют специального анализа с особым учетом их специфики.

Таблица 4

Оценка параметров T_0 и $\tilde{r}(0)$ временной корреляционной функции давления, Москва, январь

№ п/п	Характеристика	Уровень, гПа					
		Земля	850	700	500	400	300
1	$\tilde{r}(6)$	0,98	0,97	0,97	0,97	0,96	0,95
2	$\tilde{r}(12)$	0,93	0,93	0,93	0,92	0,91	0,90
3	$\varphi(6)$	0,95	0,96	0,96	0,95	0,95	0,95
4	$\tilde{\tau}(\varphi)$	0,21	0,19	0,19	0,21	0,21	0,21
5	T_0 ч	29	32	32	29	29	29
6	$\tilde{r}(\varphi)$	0,98	0,98	0,98	0,98	0,98	0,98
7	$\tilde{r}(0)$	1,00	0,99	0,99	0,99	0,98	0,97
8	η^2	0,005	0,01	0,01	0,01	0,02	0,03
9	σ_H гПа	11,7	9,1	10,8	14,4	16,7	18,6
10	σ_δ гПа	0,8	0,9	1,1	1,4	2,4	3,2

В табл. 4 представлена последовательность получения параметров T_0 и $\tilde{r}(0)$ временной корреляции поля давления на различных уровнях для января. При этом в качестве $\tilde{r}_1$ и $\tilde{r}_2$ использованы значения корреляционной функции для сдвигов времени $\tau=6$ ч и $\tau=12$ ч соответственно (строки 1 и 2). В строке 3 помещены значения отношения $\varphi=\tilde{r}_2/\tilde{r}_1$, в строке 4 — соответствующие им значения безразмерного сдвига $\tilde{\tau}$, взятые из табл. 2, а в строке 5 — оценки масштаба временной корреляции $T_0=6/\tilde{\tau}$ (см. формулу (10)).

Взятые из табл. 2 «теоретические» значения коэффициента корреляции для $\tau=6$ ч представлены в строке 6, а в строке 7 — оценки $\tilde{r}(0)$, полученные по формуле (9). При этом следует иметь в виду, что оценка $\tilde{r}(0)=1,00$ для приземного давления получена

лишь вследствие принятой нами точности представления данных, в силу которой мы все значения $\tilde{r}(0) \geq 0,995$ округляем до 1,00.

В последних строках таблицы приводится оценка погрешностей случайных ошибок измерений, соответствующих полученным значениям $\tilde{r}(0)$. При этом мы исходили из следующих из (4) формул [2]

$$\eta^2 = \frac{1 - \tilde{r}(0)}{\tilde{r}(0)}, \quad (12)$$

$$\sigma_{\delta} = \sigma_H \eta = \tilde{\sigma}_H \sqrt{1 - \tilde{r}(0)}. \quad (13)$$

Эмпирические значения средних квадратических отклонений $\tilde{\sigma}_H$ геопотенциала для различных уровней (строка 9 табл. 4) взяты из Справочника [8].

Из табл. 4 видно, что стандарт случайных ошибок измерения геопотенциала растет от значений порядка 1 гПа в нижней тропосфере до 3 гПа в верхней тропосфере. Можно думать, что эти значения σ_{δ} действительно характеризуют ошибки радиозонди-

Таблица 5

Параметры временной структуры давления. Москва

Параметр	Месяц	Уровень, гПа					
		Земля	850	700	500	400	300
T_0 ч	I	29	32	32	29	29	29
	IV	21	23	29	25	25	25
	VII	25	28	28	28	28	28
	X	25	32	32	29	32	32
$\tilde{r}(0)$	I	1,00	0,99	0,99	0,98	0,97	0,97
	IV	0,99	0,99	0,98	0,98	0,97	0,97
	VII	1,00	0,98	0,98	0,97	0,96	0,95
	X	1,00	0,99	0,99	0,99	0,99	0,96
$\tilde{\sigma}_H$ гПа	I	11,7	9,1	10,8	14,4	16,7	18,6
	IV	8,4	7,1	8,3	11,5	13,6	15,3
	VII	6,0	5,7	7,3	8,7	10,1	12,1
	X	9,5	8,0	9,8	13,9	16,7	20,2
σ_{δ} гПа	I	0,8	0,9	1,1	1,4	2,4	3,2
	IV	0,8	0,7	1,2	1,6	2,3	2,6
	VII	0,4	0,8	1,0	1,5	2,0	2,7
	X	0,8	0,8	1,0	1,4	1,7	4,0
	Средн.	0,8	0,8	1,1	1,5	2,1	3,1

рования, тем более что для Москвы оценки σ_{δ} практически не зависят от сезона года. Это хорошо видно из табл. 5, в которой представлены значения оценок параметров временной структуры для центральных месяцев четырех сезонов, полученные аналогично оценкам, приведенным в табл. 4. Оценки σ_{δ} при переходе от сезона к сезону заметно устойчивее оценок $\tilde{r}(0)$ и $\tilde{\sigma}_H$. Это наводит на мысль, что для данного района можно использовать в качестве характерных осредненные по всем сезонам значения σ_{δ} . В этих условиях параметр $\tilde{r}(0)$ определяется по формуле

$$\tilde{r}(0) = 1 - (\sigma_{\delta}/\tilde{\sigma}_H)^2 \quad (14)$$

и по данным о временной корреляции определяется лишь масштаб T_0 .

Таблица 6

Расчетные и фактические значения временной корреляционной функции поля давления (в сотых долях единицы). Москва

τ, ч	Уровень, гПа								
	850			500			300		
	r	$\tilde{r}_p$	$\tilde{r}_\Phi$	r	$\tilde{r}_p$	$\tilde{r}_\Phi$	r	$\tilde{r}_p$	$\tilde{r}_\Phi$
Январь									
0	100	99	—	100	98	—	100	97	—
6	98	97	97	98	96	97	98	95	95
12	94	93	93	94	92	92	94	91	90
18	89	88	87	87	85	86	87	84	83
24	83	82	80	80	78	80	80	78	77
30	76	75	74	72	71	73	72	70	70
36	69	68	68	65	64	67	65	63	64
42	62	61	63	57	56	62	57	55	58
48	56	55	58	51	49	57	51	49	53
Июль									
0	100	98	—	100	97	—	100	95	—
6	98	96	96	98	95	95	98	93	93
12	93	91	91	93	90	90	93	88	88
18	86	84	85	86	83	85	86	82	83
24	79	77	79	79	77	80	79	75	77
30	71	70	71	71	69	74	71	67	70
36	63	62	64	64	61	67	63	60	63
42	56	55	58	56	54	62	56	53	58
48	49	48	53	49	48	57	49	47	53

Насколько хорошо аппроксимация (6) описывает фактические данные о временной корреляции барического поля в тропосфере, можно судить по табл. 6, в которой представлены наряду с фактическими значениями эмпирической временной корреляционной функции $\tilde{r}_f$ «теоретические» (r) и расчетные (r_p) значения, полученные при задании величин T_0 и $\tilde{r}(0)$ - из табл. 5.

Таблица 7

Параметры временной структуры поля давления. Владивосток

Параметр	Месяц	Уровень, гПа						
		Земля	1000	850	700	500	400	300
T_0 ч	I	19	23	29	32	32	45	45
	IV	17	38	23	29	21	21	25
	VII	20	—	25	29	29	25	29
	X	19	25	29	25	25	29	32
$\tilde{r}(0)$	I	0,98	0,78	0,86	0,85	0,83	0,81	0,82
	IV	0,99	0,73	0,78	0,79	0,80	0,80	0,80
	VII	1,00	—	0,79	0,79	0,81	0,83	0,83
	X	0,98	0,73	0,82	0,83	0,84	0,83	0,84
$\tilde{\sigma}_H$ гПа	I	6,7	4,5	5,2	7,0	10,8	12,8	13,9
	IV	6,9	4,2	5,6	7,6	11,4	14,1	17,1
	VII	4,6	—	3,9	4,5	6,0	7,5	10,2
	X	6,2	4,1	5,1	6,3	9,6	11,9	14,6
$\tilde{\sigma}_\theta$ гПа	I	0,9	2,1	1,9	2,7	4,5	5,6	5,9
	IV	0,7	2,2	2,6	3,5	5,1	6,3	9,3
	VII	0,3	—	1,8	2,0	2,6	3,1	4,2
	X	0,9	2,1	2,2	2,8	3,8	4,9	5,8
	Сред.	—	2,1	2,1	2,8	4,0	5,0	6,3

Из табл. 6 видно, что $\tilde{r}_p$ довольно близки к $\tilde{r}_f$. Наилучшим это совпадение является для $\tau=6$ ч и $\tau=12$ ч, что вполне естественно, поскольку данные именно для этих сдвигов использовались при аппроксимации. Существенно также, что расчетные значения практически совпадают с фактическими до сдвигов времени порядка полутора суток, после чего намечается систематическое занижение расчетных корреляционных функций.

Использование при расчете вместо табличных $\tilde{r}(0)$ значений $\tilde{r}(0)$, рассчитанных по формуле (14) при фиксированных средних значениях σ_θ , привело бы к несколько большим отличиям $\tilde{r}_p$ от $\tilde{r}_f$, особенно на верхних уровнях. Однако, учитывая сказанное выше о выборочных погрешностях эмпирических коэффициентов

корреляции, составляющих несколько сотых, вряд ли можно утверждать, что это отличие является существенным. С этой точки зрения, вероятно, можно не придавать значения также различиям масштаба корреляции для разных уровней и даже для разных сезонов и пользоваться для данного района единым его значением, например, $T_0=28$ ч. Заметим, что он оказался также близ-

Таблица 8

Расчетные и фактические значения временной корреляционной функции поля давления (в сотых долях единицы). Владивосток

τ_0 ч	Уровень (гПа)								
	850			500			300		
	r	$\tilde{r}_p$	$\tilde{r}_\Phi$	r	$\tilde{r}_p$	$\tilde{r}_\Phi$	r	$\tilde{r}_p$	$\tilde{r}_\Phi$

Январь

0	100	86	—	100	83	—	100	82	—
6	98	84	84	98	81	81	99	81	81
12	94	81	80	94	78	78	97	80	79
18	87	75	79	89	75	74	94	77	76
24	80	68	73	83	69	69	90	74	71
30	72	62	70	76	63	64	85	70	67
36	65	56	65	69	57	58	81	66	61
42	57	49	60	62	51	52	76	62	55
48	51	44	56	56	46	46	71	58	48

Июль

0	100	79	—	100	81	—	100	83	—
6	98	77	77	98	79	79	98	81	81
12	92	73	72	94	76	75	94	78	77
18	84	66	66	87	70	73	87	72	76
24	75	59	60	80	64	70	80	63	74
30	66	52	51	72	58	65	72	60	68
36	58	46	41	65	53	56	65	54	61
42	50	40	33	57	46	53	57	47	59
48	43	34	20	51	41	46	51	42	55

ким по порядку к оценкам T_0 (28—30 ч), полученным в [1, 7] для приземного давления в близлежащем районе.

Эти выводы, однако, не следует распространять на другие географические районы, для которых может иметь место существенный годовой ход характеристик временной структуры и значительные различия их на разных уровнях. В качестве примера

в табл. 7 представлены оценки этих характеристик, полученные для ст. Владивосток, где эти различия выражены довольно четко. Они заслуживают дополнительного анализа с целью выяснения того, какие из них являются реальными, а какие могут быть обусловлены причинами методического характера (например, возможным завышением дисперсии и корреляции за счет годового хода в переходные сезоны). Во всяком случае ясно, что пользование средними значениями параметров в таких условиях недопустимо. Может, однако, оказаться целесообразным определенное сглаживание параметров с целью получения более плавного хода их как по высоте, так и в течение года.

Как видно из табл. 8, в которой расчетные значения корреляционной функции сравниваются с фактическими, аппроксимация (6) для Владивостока оказывается заметно хуже, чем для Москвы. Представляется, однако, что до $\tau=24\ldots 36$ ч ею вполне можно пользоваться.

СПИСОК ЛИТЕРАТУРЫ

1. Антоненко В. С., Лугина К. М. Пространственно-временная структура поля давления для летнего сезона. — Труды ГГО, 1977, вып. 397, с. 70—77.
2. Каган Р. Л. Осреднение метеорологических полей. — Л.: Гидрометеиздат, 1979. — 213 с.
3. Каган Р. Л., Лугина К. М. К вопросу об использовании наземных данных при анализе барического поля на высотах. — Метеорология и гидрология, 1977, № 2, с. 32—38.
4. Каган Р. Л., Лугина К. М., Антоненко В. С. К вопросу об анизотропии пространственно-временной структуры барического поля. — Труды ГГО, 1980, вып. 412, с. 20—28.
5. Каган Р. Л., Сибир Е. Е. О точности осреднения периодически коррелированных величин. — Труды ГГО, 1980, вып. 431, с. 90—105.
6. Казначеева В. Д., Хвостова Р. Н. Временные корреляционные функции основных метеорологических элементов в свободной атмосфере над территорией СССР. — В кн.: Применение статистических методов в метеорологии: Труды II Всесоюзного симпозиума по применению статистических методов в метеорологии. Л.: Гидрометеиздат, 1977, с. 90—95.
7. Лугина К. М., Каган Р. Л. К вопросу о пространственно-временном анализе барического поля. — Труды ГГО, 1974, вып. 336, с. 75—94.
8. Справочник по вертикальным корреляционным связям метеорологических элементов в свободной атмосфере над СССР/Под ред. В. Д. Казначеевой. — М.: Гидрометеиздат, 1980. — 397 с.
9. Справочник по временным корреляционным связям метеорологических элементов в свободной атмосфере над СССР/Под ред. В. Д. Казначеевой. — М.: Гидрометеиздат, 1980. — 311 с.
10. Хвостова Р. И. Оценка эффективности различных алгоритмов контроля по материалам аэрологического архива на микрофильме. — Труды ВНИИГМИ—МЦД, вып. 35, 1977, с. 75—86.

К ВОПРОСУ ОБ ОЦЕНКЕ ЗНАЧИМОСТИ ЭКСТРЕМУМОВ В ВЕКОВОМ ХОДЕ МЕТЕОРОЛОГИЧЕСКИХ ЭЛЕМЕНТОВ

1. Проблема исследования экстремальных климатических характеристик имеет много аспектов. Ряд вопросов, относящихся к этой проблеме, был рассмотрен в работах [6, 9]. Были детально изучены возможности теоретического расчета как локальных характеристик экстремумов, так и обобщенных их характеристик — выбросов временных рядов. Широко использовался для этой цели также метод статистического моделирования. Исследовалась зависимость полученных результатов от типа статистической структуры изучаемого процесса и от точности задания исходных данных.

Оказывается, что результаты расчетов характеристик локальных экстремумов в большой степени зависят от точности задания исходной информации. Достаточно указать, что для получения некоторых характеристик необходимо задание производных корреляционных функций процесса до шестого порядка включительно. Вряд ли стоит доказывать, что такими сведениями метеорологи (да и не только метеорологи) не располагают в настоящее время и не будут располагать в течение обозримого будущего. Кроме того, имеющиеся в специальной литературе (см., например, [18]) многообещающие результаты относятся, как правило, к случаю отсутствия ошибок наблюдений. Для метеорологии этот случай представляет лишь сугубо теоретический интерес. Выполненные в [6, 9] оценки показали, что этот фактор существенным образом изменяет многие результаты по сравнению с результатами, получаемыми для идеально точных данных.

Наличие ошибок измерений, неточность необходимых для формулировки моделей данных о временной структуре, определяемая ограниченностью выборок, по которой они могут быть получены, приводят к тому, что результаты расчета характеристик, относящихся к локальным экстремумам, оказываются весьма ненадежными, и использование их для решения практических задач вряд ли может быть рекомендовано.

С этой точки зрения заслуживает предпочтения использование данных, относящихся к выбросам временных рядов, т. е. к периодам непрерывного пребывания изучаемых величин выше или ниже заданного уровня. Каждый выброс может содержать в себе несколько локальных экстремумов и в этом смысле воздействие выброса элемента представляет собой совокупное воздействие нескольких последовательных экстремумов. Естественно поэтому, что для расчета его характеристик требуется знание не столь детальных характеристик структуры, которые более надежно определяются по эмпирическим данным. В настоящей статье мы рассмотрим возможности использования для оценки значимости экстре-

мумов в вековом ходе метеорологических элементов характеристик продолжительности и мощности выбросов временных рядов.

2. Вопросы расчета характеристик выбросов временных рядов хорошо освещены в специальной литературе [12, 17, 18]. Применительно к специфическим задачам метеорологии ряд аспектов этой проблемы детально рассмотрен авторами [2, 5]. Однако в этих работах основное внимание уделялось определению среднего числа и в меньшей степени средней продолжительности выбросов. Распределению этих характеристик уделялось гораздо меньше внимания. Между тем такие распределения представляют существенный практический интерес, поскольку ясно, например, что режим работы ветроэнергетических установок или условия сельскохозяйственной деятельности должны организовываться по-разному в зависимости от того, определяется ли средняя продолжительность выброса одним длинным и многими короткими выбросами или каким-то количеством выбросов средней длины. Кроме того, нужно иметь в виду, что формальное определение выброса как пересечения некоего уровня недостаточно для суждения о возможной степени воздействия его на те или иные объекты. Ясно, например, что превышение уровня на несколько градусов гораздо существеннее, чем превышение на десятые доли градуса. С этой точки зрения представляло бы интерес изучение мощности выбросов. Эта характеристика, представляющая собой совокупное превышение значений элемента над уровнем во время выброса, широко используется в технике, где ее принято называть площадью выброса [18]. В метеорологических исследованиях эта характеристика, насколько нам известно, еще не использовалась, хотя она в известном смысле близка к используемым в прикладной метеорологии характеристикам типа сумм температур выше того или иного уровня.

Строгий аналитический расчет распределения продолжительности выбросов даже для гауссовских процессов представляет значительные математические трудности. Для процессов с дискретным временем расчет распределения продолжительности 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 продолжительности выбросов.

Один из возможных способов приближенного расчета распределения продолжительности выбросов случайной последовательности $f(t)$ через уровень C основан на предположении о том, что последовательность $y(t)$, принимающая значения y^+ , если $f(t) > C$, и y^- , если $f(t) < C$ (последовательность импульсов), является марковской цепью какого-нибудь порядка. Общий подход такого рода излагается в известной работе Б. Р. Левина и Я. А. Фомина [13]. Применительно к гидрологическим задачам этот подход излагается в обзоре [16]. В метеорологии такой подход применялся, например, в [19], где рассмотрение ограничивалось мар-

ковскими цепями первого порядка. Аналогично, следуя [19], Н. В. Кобышева и др. [11] оценивали продолжительность выбросов температуры воздуха. Поскольку, однако, события перехода через уровень даже для марковских процессов первого порядка не образуют марковской цепи, то, как показали выполненные авторами оценки, даже использование трехсвязной модели не всегда обеспечивает высокую точность расчетов.

Подобного рода аналитические расчеты могут производиться также и для случайных процессов, характеризующихся другими

Таблица 1

Вероятность появления выброса заданной длительности ($\Delta \leq l$) на реализации стационарной бессвязной гауссовской последовательности длиной $T=250$ членов

$c = \frac{f - \bar{f}}{\sigma}$	l							N
	1	2	3	4	5	7	10	
0	1,00	1,00	1,00	1,00	0,99	0,62	0,12	61897
0,5	1,00	1,00	1,00	0,76	0,38	0,05	0,00	52540
-0,5	1,00	1,00	0,99	0,78	0,39	0,03	0,00	52754
1,0	1,00	0,99	0,56	0,12	0,02	0,00		32917
-1,0	1,00	1,00	0,57	0,12	0,02	0,00		33030
1,5	1,00	0,63	0,06	0,00				15320
-1,5	1,00	0,64	0,07	0,00				15364
2,0	0,99	0,11	0,00					5370
-2,0	0,99	0,11	0,00					5580
2,5	0,78	0,00						1528
-2,5	0,78	0,01						1567
3,0	0,30	0,00						349
-3,0	0,28	0,00						336

свойствами, в частности для негауссовских процессов. Однако во многих случаях оказывается предпочтительным использование метода статистического моделирования, который уже применялся ранее авторами для оценок распределений числа и продолжительности выбросов и других детальных характеристик метеорологических процессов [3, 8]. Такое моделирование использовалось и при выполнении настоящей работы.

При обработке генерированных в соответствии с заданными моделями рядов использовались алгоритмы, общая схема которых подробно изложена в [7]. Особое внимание уделялось расчету характеристик продолжительности и мощности выбросов иссле-

дуемых процессов. В качестве примера, иллюстрирующего возможность такой обработки, в табл. 1 представлены данные, полученные для стационарной бессвязной гауссовской последовательности. Здесь f представляет собой значение моделируемой величины, $\bar{f}$ и σ — соответственно ее среднее значение (норму) и среднее квадратическое отклонение (стандарт), c — относительное отклонение f от нормы.

В таблице приводятся не сами распределения продолжительности, а вероятности появления выбросов заданной длительности

Таблица 2

Вероятность появления выброса заданной мощности ($s \geq k\sigma$) на реализации стационарной бессвязной гауссовской последовательности длиной $T=250$ членов

$c = \frac{f - \bar{f}}{\sigma}$	k							N
	0,0	0,5	1,0	2,0	3,0	4,0	5,0	
0	1,00	1,00	1,00	1,00	1,00	0,99	0,87	61897
0,5	1,00	1,00	1,00	0,99	0,73	0,29	0,08	52540
-0,5	1,00	1,00	1,00	1,00	0,72	0,30	0,11	52754
1,0	1,00	1,00	1,00	0,59	0,07	0,01	0,00	32917
-1,0	1,00	1,00	1,00	0,56	0,11	0,02	0,00	33030
1,5	1,00	0,99	0,81	0,08	0,00			15320
-1,5	1,00	1,00	0,82	0,10	0,01			15364
2,0	0,99	0,78	0,31	0,01	0,00			5370
-2,0	1,00	0,78	0,29	0,01	0,00			5580
2,5	0,78	0,30	0,05	0,00				1528
-2,5	0,78	0,28	0,05	0,00				1567
3,0	0,30	0,05	0,00					349
-3,0	0,28	0,05	0,01					336

на реализации фиксированной длины. Разумеется, эти вероятности связаны с распределением продолжительности, и их расчет по упомянутому распределению в данном случае можно произвести сравнительно просто. В общем случае связанных последовательностей такой расчет затруднителен. Между тем для практического использования эта характеристика представляется очень полезной.

В таблице для неотрицательных уровней представлены характеристики, относящиеся к выбросам вверх, а для отрицательных уровней — к выбросам вниз. Именно такие характеристики представляют, на наш взгляд, наибольший практический интерес. Разумеется, в стационарном варианте цифры, соответствующие вы-

бросам через уровни, одинаково удаленные от нормы, должны совпадать. Мы сочли целесообразным привести и те, и другие, поскольку расхождения между ними характеризуют возможную точность определения значений вероятности для выборок используемого объема (в данном случае оценки получены путем обработки 1000 реализаций длиной по 250 членов каждая).

Как видно из таблицы, эти расхождения могут оказаться довольно существенными лишь для достаточно удаленных от нормы уровней, для которых отмечается сравнительно небольшое число выбросов (число закончившихся выбросов, использованное при обработке полученной выборки, также представлено в таблице, где оно обозначено N).

Аналогичные данные, относящиеся к вероятности появления выброса заданной мощности, представлены в табл. 2.

Подобного рода оценки получались и для связанных последовательностей различной длины, притом не только для гауссовских, но и для логарифмически-нормально распределенных. Существенно также, что в этом случае не представляет труда получение оценок и для последовательностей с заданным трендом характеристик модели (например, нормы или стандарта ее), что при аналитических расчетах приводило бы к очень большим осложнениям. Некоторые примеры использования подобной методики расчетов приводятся нами ниже.

3. В последние годы все большее внимание уделяется проблеме изменений климата, которая становится одной из центральных проблем не только климатологии, но и всей современной науки. Этой проблеме посвящено множество работ, обзор которых можно найти, например, в [1, 20].

В этих работах убедительно показано, что климат Земли претерпевал очень большие изменения в прошлом. Однако, если говорить только о периоде инструментальных наблюдений, то сведения об изменениях климата за этот период не столь очевидны. Это связано, главным образом, с тем, что период инструментальных наблюдений невелик. По большинству метеорологических элементов он не превышает 100 лет, и даже по приземной температуре воздуха ряды длительностью более 100 лет имеются лишь по немногим станциям, не представляющим собой единой сети. Закономерная изменчивость метеорологических элементов в течение этого периода (если она и имела место) значительно меньше, чем та, которая наблюдалась при переходе от одной геологической эпохи к другой. К тому же она затушевывается сравнительно большой междугодовой изменчивостью, определяемой в значительной мере локальными метеорологическими процессами, не представляющими существенного интереса при изучении эволюции климатического режима.

Указанные обстоятельства делают крайне затруднительным статистический анализ данных наблюдений на предмет выявления возможных климатических трендов. Мы не будем останавливаться на истории этого вопроса, подробно освещенной, например,

в [14, 15]. Укажем лишь, что применение самых современных методов, будь то регрессионный или спектральный анализ, оказывается очень затруднительным ввиду того, что сам смысл задачи требует достаточно надежного выделения слабого полезного сигнала (искомого тренда) на фоне сильных шумов (мелкомасштабных локальных флуктуаций). При этом для каждого пункта мы располагаем выборкой небольшой длины. Неудивительно поэтому, что специалисты, занимающиеся этой проблемой, не находят достаточно убедительных доказательств существования или отсутствия таких трендов. В тех случаях, когда принимается гипотеза стационарности, т. е. отсутствия тренда, имеющиеся данные обычно не дают оснований отвергнуть эту гипотезу по данным одной исследуемой реализации.

Конечно, в этих условиях вряд ли можно рассчитывать на получение качественно иных результатов путем применения каких-либо новых критериев. Тем не менее представляется, что подобного рода критерии могут быть полезны хотя бы в том отношении, что они позволяют подходить к оценке этих рядов несколько в ином плане. Интересно с этой точки зрения было рассмотреть возможности оценки векового хода метеорологических элементов не в плане изучения его суммарного тренда, а в плане рассмотрения его отдельных особенностей.

При рассмотрении многолетних рядов метеорологических элементов наряду с пульсациями их от года к году наблюдаются и сравнительно длинные периоды относительно высоких и относительно низких значений. Естественным образом возникает вопрос, являются ли такие изменения следствием некоей закономерной эволюции, которая, будучи прослеженной, могла бы служить для целей сверхдолгосрочного прогноза, или они являются проявлением наличия случайных шумов на фоне стационарного климата. Нам представляется, что наряду со спектральными методами, которые обычно применяются для решения этого вопроса, заслуживает внимания и анализ выбросов рассматриваемого элемента. Очевидным достоинством такого подхода является его большая наглядность и возможность суждения о значимости отдельно взятого аномального периода.

Изложим особенности этого подхода на конкретном примере анализа данных о среднегодовой температуре воздуха в Центральной Англии, полученных в [21] путем приведения рядов на станциях, проводивших наблюдения в этом районе в течение более чем 300 лет. Образованный таким образом ряд детально анализировался и использовался в работе многих исследователей (см., например, [14, 22]).

Колебания температуры в течение периода 1659—1973 гг. представлены на рис. 1, из которого видно, что эти колебания весьма велики и на фоне их затруднительно судить о каких-либо закономерных изменениях. Для этой цели пришлось бы прибегнуть к сглаживанию ряда, что, как известно, может привести к появлению ложных циклов. Вместе с тем даже визуальный анализ дан

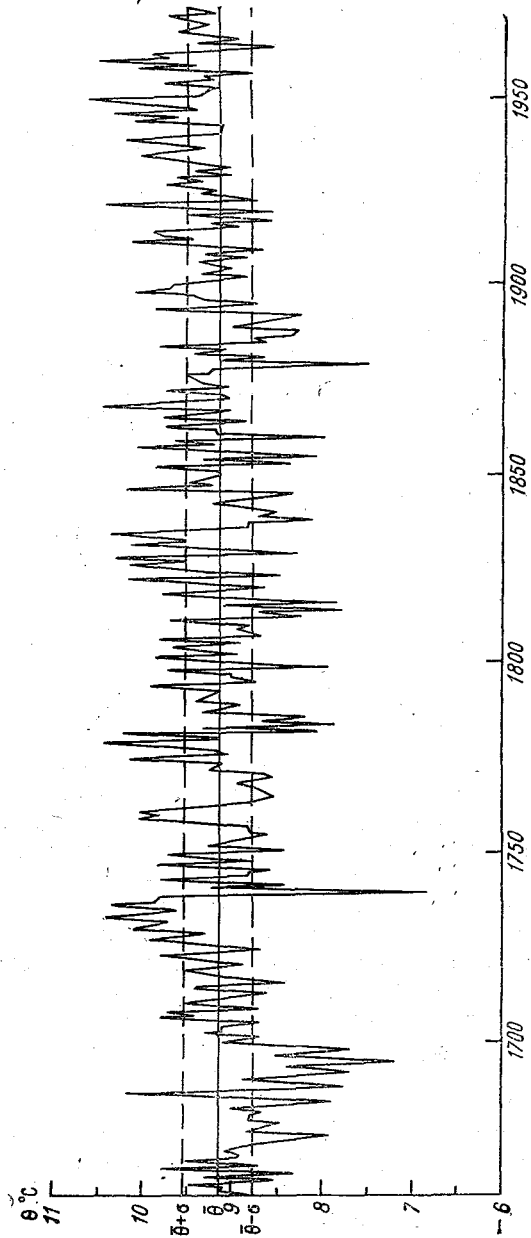


Рис. 1. Температура воздуха в Центральной Англии.

ных на рисунке позволяет обнаружить наличие довольно длительных периодов (пять и более лет) сравнительно высоких и низких температур, исследование которых представляет определенный интерес. Обработка данных указанного ряда приводит к значениям его статистических параметров, представленных в табл. 3, где приняты обычные обозначения.

Анализ выбросов температуры может вестись для разных уровней, при выборе которых приходится учитывать различные соображения. Во-первых, практический интерес представляют высокие и низкие уровни, значительно отклоняющиеся от нормы. С другой стороны, приходится иметь в виду, что такие выбросы

Таблица 3

Статистические параметры температуры воздуха (θ) по данным наблюдений в Центральной Англии за 1659—1973 гг.

$\bar{\theta}^{\circ}\text{C}$	$\sigma^{\circ}\text{C}$	A	E	Корреляционная функция со сдвигом τ лет						
				1	2	3	4	5	7	10
Фактические значения										
9,14	0,38	−0,37	0,49	0,26	0,28	0,12	0,11	−0,01	0,03	−0,07
Модельные значения										
9,14	0,38	0	0	0,30	0,23	0,17	0,12	0,09	0,05	0,02

очень редки и поэтому статистические оценки по таким выбросам получить довольно трудно. Кроме того, известно, что вероятность появления таких выбросов мало зависит от степени связности процесса [12] и от некоторых других гипотез о свойствах ряда. Само по себе это обстоятельство является благоприятным при расчетах соответствующих характеристик, поскольку позволяет получить их разумные оценки даже при не очень удачном выборе исходной модели. Однако именно это делает сомнительным использование высоких уровней в случае, если мы хотим воспользоваться данными анализа выбросов, чтобы отдать предпочтение той или иной модели. В силу сказанного нам представляется предпочтительным основываться при анализе на данных о выбросах за уровни, отстоящие от нормы на $0,5\sigma - \sigma$.

Используем приведенные в табл. 3 параметры в качестве входных для моделирования. При этом естественно выбрать гауссовскую модель, поскольку значения коэффициентов асимметрии (A) и эксцесса (E) невелики и находятся в пределах возможных ошибок для данного объема выборки. Примем также в качестве исходной гипотезу о стационарности ряда.

Для окончательной формулировки модели необходимо задание временной структуры ряда. Для этой цели мы полагаем, анало

гично [14], что ряд представляет собой сумму двух составляющих:

$$\theta(t) = \theta_1(t) + \theta_2(t), \quad (1)$$

где $\theta_1(t)$ — стационарная случайная функция с корреляционной функцией $r_1(\tau)$, $\theta_2(t)$ — некоррелированная с ней случайная величина (белый шум), дисперсию которой обозначим через σ_2^2 . Долю, которую эта дисперсия составляет от полной дисперсии температуры σ^2 , обозначим через ψ^2 . Она аналогична используемой в объективном анализе мере ошибок измерений η^2 , однако, по существу, определяется не ошибками измерений, а наличием мелко-масштабных локальных пульсаций метеорологических процессов. В рассматриваемом нами случае $\psi^2 = 0,59$. Соответственно дисперсия составляющей θ_1 есть

$$\sigma_1^2 = (1 - \psi^2) \sigma^2. \quad (2)$$

Корреляционная функция $r_1(\tau)$ аппроксимировалась зависимостью

$$r_1(\tau) = q^{\tau}, \quad (3)$$

где q оказалось равным 0,74.

Корреляционная функция, соответствующая принятой модели, также представлена в табл. 3.

Таблица 4

Распределение продолжительности и вероятности появления выбросов заданной длительности ($\Delta \gg l$) на реализации длиной $T=315$ членов

$c = \frac{f-\bar{f}}{\sigma}$	l									N
	1	2	3	4	5	6	7	8	9	
Расчетные распределения продолжительности (по 1000 реализациям модели)										
0,5	1,0	0,38	0,17	0,08	0,04	0,02	0,01	0,00	0,00	108502
1,0	1,0	0,25	0,08	0,03	0,01	0,00	0,00	0,00	0,00	69551
Расчетные значения вероятности (по 1000 реализациям)										
0,5	1,0	1,00	1,00	0,99	0,93	0,75	0,53	0,36	0,22	1000
1,0	1,0	1,00	0,95	0,70	0,38	0,18	0,08	0,03	0,01	1000
Фактические распределения продолжительности по данным о температуре в Центральной Англии										
0,5	1,0	0,32	0,19	0,08	0,04	0,02	0,02	0,02	0,02	55
-0,5	1,0	0,34	0,24	0,12	0,06	0,06	0,04	0,04	0,04	51
1,0	1,0	0,17	0,05	0,00	0,00	0,00	0,00	0,00	0,00	42
-1,0	1,0	0,18	0,03	0,03	0,03	0,03	0,03	0,03	0,00	33

В соответствии с описанной моделью по специальной программе, реализованной на БЭСМ-6, был сформирован массив данных, состоящий из 1000 реализаций длиной 315 членов каждая. Обработка этого массива позволила получить большой объем статистических данных. Ниже мы ограничимся приведением лишь некоторых из них, относящихся к уровням, отстоящим от нормы на $0,5\sigma$ и на σ .

В табл. 4 приводятся данные о распределении продолжительности выбросов температур и о вероятностях появления на реализации выбросов заданной длины. В соответствии с принятой гипотезой стационарности расчетные значения получены путем осреднения соответствующих характеристик выбросов вверх через уровни выше нормы и вниз через уровни ниже нормы. В этой же таблице представлена фактическая частота выбросов заданной длительности (отдельно для положительных и отрицательных уровней), полученная по данным имеющейся в нашем распоряжении одной реализации.

Данные этой таблицы могут быть использованы как для оценки значимости отдельных значительных выбросов, так и для выяснения общего соответствия принятой при построении модели гипотезы фактическим данным.

Как видно из таблицы, наиболее продолжительный выброс температуры с превышением нормы на $0,5\sigma (0,19^\circ\text{C})$ имел место в течение 9 лет, а наибольшая продолжительность непрерывного пребывания температуры ниже относительного уровня $c=-1$ составила 8 лет. Обращаясь к данным таблицы, видим, что вероятность появления таких выбросов на реализации длиной 315 членов составляет 0,22 для уровня $c=0,5$ и 0,03 для уровня $c=-1$.

Сравним также фактические и расчетные распределения продолжительности, используя критерий Колмогорова. Сопоставляя эти распределения, видим, что максимальное расхождение их составляет: 0,06—0,07 для $c=0,5$ и 0,07—0,08 для $c=\pm 1$. Учитывая количество выбросов N , по которому определялись фактические распределения, получаем оценку статистики Колмогорова $\lambda = D\sqrt{N}$. Она не превосходит 0,55 для $c=0,5$, а для $c=\pm 1$ $\lambda < 0,52$. Обращаясь к соответствующим таблицам, получаем, что значения вероятности случайного расхождения между распределениями больше 0,9 в обоих случаях. Таким образом, на основании критерия Колмогорова мы не можем отвергать гипотезу о стационарности рассматриваемого ряда.

Произведем аналогичные оценки применительно к распределениям мощности выбросов, данные о которых представлены в табл. 5.

Оценим вероятности наиболее мощных из наблюдавшихся фактически выбросов. Наибольшая мощность выброса вверх через уровень $c=0,5$ превысила 5σ . Такой же мощности достиг и выброс вниз через $c=-1$. Эмпирическая частота выбросов такой мощности составила 0,01 и 0,03 соответственно. Сопоставляя эти частоты с расчетными вероятностями появления выбросов такой

мощности, получаем вероятность выброса для уровней $c=\pm 0,5$ и $c=\pm 1$ соответственно $-0,72$ и $-0,13$.

Произведя сопоставление расчетных и фактических распределений мощности по критерию Колмогорова, имеем максимальное расхождение $D=0,07$ для уровней $c=\pm 0,5$ и $D=0,1$ для $c=\pm 1$. Соответственно для уровней $c=\pm 0,5$ $\lambda=0,52$, вероятность случайного расхождения распределений около 0,95, для $c=\pm 1$ $\lambda=0,57$, вероятность случайного расхождения больше 0,9. Следовательно, и рассмотрение распределений мощности выбросов не дает оснований отвергнуть гипотезу о стационарности ряда.

Таблица 5

Распределение мощности и вероятности появления выбросов заданной мощности ($s \geq k\sigma$) на реализации длиной $T=315$ членов

$c = \frac{f - \bar{f}}{\sigma}$	k							N
	0	0,5	1,0	2,0	3,0	4,0	5,0	
Расчетные распределения мощности (по 1000 реализациям модели)								
0,5	1,0	0,62	0,37	0,15	0,07	0,04	0,02	108502
1,0	1,0	0,50	0,23	0,06	0,02	0,00	0,00	69551
Расчетные значения вероятности (по 1000 реализациям)								
0,5	1,0	1,00	1,00	1,00	0,98	0,89	0,72	1000
1,0	1,0	1,00	1,00	0,89	0,57	0,29	0,13	1000
Фактические распределения мощности по данным о температуре в Центральной Англии								
0,5	1,0	0,69	0,39	0,13	0,06	0,02	0,02	55
-0,5	1,0	0,60	0,38	0,20	0,10	0,02	0,02	51
1,0	1,0	0,57	0,19	0,00	0,00	0,00	0,00	42
-1,0	1,0	0,57	0,33	0,06	0,03	0,03	0,03	33

Таким образом, из полученных нами четырех оценок вероятности появления единичных значительных выбросов лишь одна (вероятность наиболее продолжительного выброса через уровень $c=-1$) достаточно мала. Следует, однако, иметь в виду, что, как видно из рис. 1, рассмотренные крупные выбросы относятся к началу XVIII в. В данных для этого периода вполне могут содержаться ошибки методического характера, поэтому вряд ли стоит придавать чрезмерное значение этой единственной оценке. Что касается всех четырех оценок, полученных по критерию Колмогорова, то все они не позволяют отвергнуть гипотезу стационарности ряда.

Учитывая сказанное, нам представляется на данном этапе, до получения более убедительных результатов, возможным пользо-

ваться применительно к рассмотренному ряду стационарной моделью, т. е. полагать, что термический режим Центральной Англии за рассмотренный период существенно не менялся. Это согласуется с выводами, полученными путем анализа этих рядов другими способами [14].

4. Представляет интерес провести такой же анализ и применительно к рядам осадков. Для этой цели рассмотрим данные о суммах осадков в Бостоне за теплый сезон (апрель — октябрь) в течение 1818—1968 гг., любезно предоставленные нам А. С. Григорьевой. Обработка данных этого ряда приводит к значениям его статистических параметров, представленным в табл. 6.

Таблица 6

Статистические параметры сумм осадков (Q) по данным наблюдений в Бостоне за 1818—1968 гг.

$\bar{Q}$ мм	σ мм	A	E	Корреляционные функции со сдвигом τ лет						
				1	2	3	4	5	7	10
Фактические значения										
596	133	0,45	0,02	0,18	—0,00	0,22	0,21	0,24	0,26	0,23
Модельные значения										
596	133	0	0	0,29	0,28	0,27	0,26	0,25	0,23	0,20

Из таблицы видно, что при таком осреднении суммы осадков в Бостоне вполне могут считаться нормально распределенными. Заметим, что к аналогичному выводу приводят оценки для сумм осадков в других пунктах по данным, представленным в [20]. Временная структура ряда задавалась, как и для температуры, формулами (1)—(3) при значениях параметров $q=0,96$; $\psi^2=0,7$.

При реализации нашей схемы анализа на ЭВМ мы имели в виду также возможность применения к рядам осадков модели с заданием логарифмически нормального распределения. В данном случае необходимости в этом не возникло. Это связано, по-видимому, с тем, что нами рассматривались данные для станций, находящихся в зонах достаточного увлажнения. Можно думать, что для засушливых зон логарифмически-нормальное распределение оказалось бы более подходящим. К сожалению, надежными данными длиннорядных наблюдений по таким зонам мы не располагаем.

В табл. 7 представлена фактическая частота выбросов заданной длительности для уровней $s=\pm 0,5$ и $s=\pm 1$, полученная по данным наблюдений. Там же приведены для сравнения соответствующие значения, полученные путем обработки 1000 реализаций такой же длины, сформированных в соответствии со стационарной гауссовской моделью в предположении бессвязности ряда.

Таблица 7

аспределение продолжительности и вероятности появления выбросов заданной длительности ($\Delta \geq l$) на реализации длиной $T=151$ член

$c = \frac{f - \bar{f}}{\sigma}$	l						N
	1	2	3	4	5	6	
асчетные распределения продолжительности (по 1000 реализациям модели)							
0,5	1,0	0,36	0,16	0,02	0,05	0,03	52970
1,0	1,0	0,24	0,08	0,03	0,02	0,01	33800
Расчетные значения вероятности (по 1000 реализациям)							
0,5	1,0	1,00	0,93	0,77	0,60	0,44	1000
1,0	1,0	0,91	0,60	0,33	0,19	0,10	1000
Фактические распределения продолжительности по данным о суммах осадков в Бостоне							
0,5	1,0	0,37	0,12	0,12	0,08	0,00	26
-0,5	1,0	0,37	0,20	0,03	0,03	0,03	30
1,0	1,0	0,18	0,06	0,00	0,00	0,00	17
-1,0	1,0	0,20	0,05	0,00	0,00	0,00	20

Таблица 8

Распределение мощности и вероятности появления выбросов заданной мощности ($s \geq k$) на реализации длиной $T=151$ член

$c = \frac{f - \bar{f}}{\sigma}$	k							N
	0	0,5	1,0	2,0	3,0	4,0	5,0	
Расчетные распределения мощности (по 1000 реализациям модели)								
0,5	1,0	0,61	0,36	0,14	0,07	0,04	0,02	52970
1,0	1,0	0,49	0,24	0,06	0,02	0,01	0,01	33800
Расчетные значения вероятности (по 1000 реализациям)								
0,5	1,0	1,0	0,99	0,89	0,68	0,50	0,37	1000
1,0	1,0	0,99	0,90	0,51	0,25	0,13	0,07	1000
Фактические распределения мощности по данным о суммах в Бостоне								
0,5	1,0	0,63	0,40	0,24	0,08	0,04	0,04	26
-0,5	1,0	0,63	0,37	0,10	0,00	0,00	0,00	30
1,0	1,0	0,60	0,42	0,06	0,06	0,00	0,00	17
-1,0	1,0	0,45	0,10	0,00	0,00	0,00	0,00	20

Аналогичные данные, относящиеся к распределению мощности выбросов, представлены в табл. 8.

Мы не будем подробно останавливаться на анализе данных этих таблиц, поскольку он совершенно аналогичен проведенному выше анализу данных о температуре. Укажем лишь, что применение критерия Колмогорова дает значения вероятности случайного расхождения для мощности выбросов через уровни $c = \pm 0,5$ и $c = \pm 1$ более 0,9. Следовательно, принятая нами модель не противоречит фактическим данным. Рассмотрение же единичных наиболее мощных выбросов дает для уровня 0,5 вероятность 0,37, а для уровня $-0,25$ 0,25, что подтверждает удовлетворительность принятой модели.

Конечно, рассмотренные примеры не могут претендовать на сколько-нибудь серьезное заключение о действительном наличии или отсутствии существенных изменений климатического режима. Представляется, однако, что указанный подход позволяет взглянуть на исследуемые ряды с другой точки зрения. В то же время при анализе данных этим способом, как и другими способами, выявляются ограниченные возможности использования данных в отдельных пунктах. Представляется, что гораздо более надежные выводы могли бы быть получены путем анализа пространственно-осредненных рядов, в которых, можно надеяться на меньший уровень различных шумов. Однако это требует обеспечения надлежащей однородности пространственно сглаженных рядов, что, как следует из [4, 10], является отнюдь не простым делом.

СПИСОК ЛИТЕРАТУРЫ

1. Будыко М. И. Климат в прошлом и будущем. — Л.: Гидрометеониздат, 1980. — 351 с.
2. Каган Р. Л. О точности расчета характеристик выбросов средней суточной температуры. — Труды ГГО, 1977, вып. 397, с. 3—12.
3. 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.
4. Каган Р. Л., Лугина К. М. Некоторые оценки возможностей мониторинга общего содержания озона в атмосфере. — См. наст. сб.
5. Каган Р. Л., Федорченко Е. И. О расчете статистических характеристик выбросов случайной функции. — Труды ГГО, 1970, вып. 268, с. 146—172.
6. Каган Р. Л., Федорченко Е. И. О влиянии дискретности измерений на точность определения числа выбросов случайного процесса. — Труды ГГО, 1975, вып. 348, с. 78—98.
7. Каган Р. Л., Федорченко Е. И. Расчет характеристик выбросов и временной структуры случайных последовательностей. — Труды ГГО, 1975, вып. 364, с. 66—76.
8. Каган Р. Л., Федорченко Е. И. О применении статистического моделирования при решении метеорологических задач. — Применение статистических методов в метеорологии. — В кн.: Труды II Всесоюзного симпозиума по применению статистических методов в метеорологии. Л., Гидрометеониздат, 1977, с. 36—44.
9. Каган Р. Л., Хлебникова Е. И. О влиянии ошибок наблюдений на характеристики экстремумов метеорологических процессов. — Труды ГГО, 1980, вып. 412, с. 59—68.
10. Каган Р. Л., Хлебникова Е. И. О нарушении однородности метеорологических полей при их интерполяции. — См. наст. сб.

11. Кобышева Н. В., Кухтина Л. П., ЧмUTOва З. Е. Методы расчета непрерывной продолжительности экстремальных температур. — Труды ГГО, 1976, вып. 349, с. 50—56.
12. Крамер Г., Лидбеттер М. Стационарные случайные процессы. — М.: Мир, 1969. — 398 с.
13. Левин Б. Р., Фомин Я. А. Применение временной дискретизации случайного процесса для нахождения распределения длительности его выбросов. — Радиотехника, 1965, т. 20, № 10, с. 1—8.
14. Поляк И. И. Методы анализа случайных процессов и полей в климатологии. — Л.: Гидрометеиздат, 1979. — 255 с.
15. Рубинштейн Е. С., Полозова Л. Г. Современные изменения климата. — Л.: Гидрометеиздат, 1966. — 268 с.
16. Румянцев В. А., Бовыкин И. В. Основы теории случайных выбросов гидрологических рядов. Обзоры. Гидрология суши. — Обнинск, изд. ВНИИГМИ—МЦД, 1977. — 34 с.
17. Случайные процессы. Выборочные функции и пересечения. Математика. — В кн.: Новое в зарубежной науке. М., Мир, 1978. Вып. 10. — 280 с.
18. Тихонов В. И. Выбросы случайных процессов. — М.: Наука, 1970. — 392 с.
19. Gabriel K. R., Neumann J. A Markov chain model for daily rainfall occurrence at Tel Aviv. — Quart. J. Roy. Meteorol. Soc., 1962, vol. 88, N 375, p. 90—95.
20. Lamb H. H. Climate: present, part and future. Vol. 1, 2. — London: Methuen, 1973. — 612 p.; 1977. — 835 p.
21. Manley G. General England temperatures: monthly means 1659 to 1973. — Quart. J. Roy. Meteorol. Soc., 1974, vol. 100, N 425, p. 389—405.
22. Schönwiese C. D. Central England temperature and sunspot variability 1660—1975. — Arch. f. Meteorol., Geoph. Univ. Bioklim., 1978, Bd. 26, N 1, S. 1—16.

Р. Л. Казан, И. А. Салль

ОБ УТОЧНЕНИИ ОЦЕНОК СРЕДНИХ СУТОЧНЫХ ЗНАЧЕНИЙ МЕТЕОРОЛОГИЧЕСКИХ ПАРАМЕТРОВ

1. Для различных практических целей, в том числе для оперативного обслуживания народного хозяйства, используются средние суточные значения температуры воздуха, давления и других метеорологических параметров. Эти значения получаются путем осреднения данных наблюдений за отдельные сроки суток и точность их зависит от числа сроков наблюдений в течение суток, выбора моментов времени, на которые они приходятся, и принятого способа осреднения. Оценка точности, с которой могут быть получены средние суточные значения, легко может осуществляться априорно при использовании данных о временной структуре осредняемых величин [4]. В частности, для получения оценок дисперсии ошибки осреднения достаточно знать ход дисперсии осредняемой величины в течение суток и ее временную корреляционную функцию на этом интервале.

Некоторые аспекты оценки точности получения средних значений по данным измерений в дискретные моменты времени были

рассмотрены в [5]. При этом подробно исследовался случай, когда измерения на интервале осреднения производятся через равные интервалы времени, а осреднение производится методом среднего арифметического. Это соответствует практически используемым рекомендациям действующих официальных инструкций [7].

Полученные в [5] оценки показывают, что метод среднего арифметического далеко не всегда обеспечивает высокую точность осреднения. Эта точность зависит от размещения сроков наблюдения на интервале осреднения. Более того, оказывается, что предлагаемый в [7] выбор сроков наблюдений является наиболее неблагоприятным при принятом способе осреднения. Поскольку выбор сроков наблюдения определяется многими факторами и сама на их является сложным делом, естественно рассмотреть возможные способы уточнения осредненных величин путем выбора более совершенных способов осреднения. Такие уточненные величины могут быть получены хотя бы в процессе климатологической обработки, когда фактор времени не очень существен, как не существенно и некоторое усложнение расчетов при использовании ЭВМ.

Ниже рассматриваются некоторые возможности уточнения средних суточных значений метеорологических параметров. Приводимые примеры относятся к стационарному случаю. Выполнение соответствующих оценок для нестационарного случая не представляет принципиальных трудностей, однако для упрощения интерпретации данных этот случай не рассматривается, тем более что, как видно из [5], изменения в точности осреднения в нестационарном варианте по сравнению со стационарным относительно невелики.

2. Методика априорной оценки точности осреднения дискретных данных на основе использования характеристик статистической структуры в настоящее время достаточно хорошо разработана как для одномерных, так и для двумерных случайных процессов [4]. Относящиеся к одномерному процессу формулы детально описаны в [5] применительно к нестационарным периодически коррелированным случайным процессам. Мы ограничимся далее приведением соответствующих формул для стационарного случая.

Пусть в моменты времени $t_1, t_2, \dots, t_n$ производятся измерения стационарной случайной переменной $f(t)$, имеющей математическое ожидание $\bar{f}$ (черта сверху означает статистическое осреднение), дисперсию σ^2 и временную корреляционную функцию $r(|t_i - t_j|)$. Измерения производятся со случайными ошибками, мера которых $\eta^2 = \sigma_\epsilon^2 / \sigma^2$ считается известной. Необходимо оценить среднее значение $f(t)$ в течение периода T (например, в течение суток):

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

Обозначая через $f_i = f(t_i)$ значения f , через $\delta_i = \delta(t_i)$ — ошибки измерений в моменты t_i , а p_i — веса осреднения, можно получить оценку величины F_0 по данным измерений:

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

Ошибка осреднения составляет

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

Оценка систематической погрешности осреднения $\bar{\Delta}$ принципиальных трудностей не представляет. Будем считать, что эта погрешность устранена, так что $\bar{\Delta} = 0$, и будем характеризовать точность осреднения дисперсией величины Δ , $E^2 = \bar{\Delta}^2$. Практически в большинстве случаев удобнее пользоваться не абсолютными ее значениями, а мерой ошибки ε^2 , т. е. ее отношением к дисперсии осредняемого параметра. Для этой величины нетрудно получить формулу

$$\begin{aligned} \varepsilon^2 = E^2/\sigma^2 = & \sum_{i=1}^n \sum_{j=1}^n p_i p_j r_{ij} + \eta^2 \sum_{i=1}^n p_i^2 + \\ & + \frac{1}{T^2} \int_0^T \int_0^T r(|t_1 - t_2|) dt_1 dt_2 - \frac{2}{T} \sum_{i=1}^n p_i \int_0^T r(|t - t_i|) dt. \end{aligned} \quad (4)$$

Точность средней величины существенно зависит от способа осреднения, т. е. от выбора весов p_i . В [5] детально рассматривался наиболее простой случай арифметического осреднения, когда $p_1 = p_2 = \dots = p_n = 1/n$. Наряду с арифметическим возможно использование множества других способов осреднения. Значительный интерес, с нашей точки зрения, представляет оптимальное осреднение [4], которое, при прочих равных условиях, обеспечивает максимальную (в смысле метода наименьших квадратов) точность оценки средней величины.

Веса оптимального осреднения получаются путем решения линейной системы уравнений

$$p_i \eta^2 + \sum_{j=1}^n p_j r_{ij} = \omega_i \quad (i = 1, 2, \dots, n), \quad (5)$$

где $\omega_i = \frac{1}{T} \int_0^T r(|t - t_i|) dt$ — величина, пропорциональная коэф-

фициенту корреляции между значением f в момент времени t_i и средней F_0 .

Оптимальное осреднение может практически использоваться при наличии достаточно надежных данных о математическом ожидании $\bar{f}$. Поскольку при нем не налагается ограничений на сумму

весов p_i , оно является корректным в случае, когда $\bar{j}=0$, т. е. когда осреднению подвергаются отклонения от математического ожидания. В противном случае естественно требовать дополнительно выполнения условия нормировки весов

$$\sum_{i=1}^n p_i = 1. \quad (6)$$

Веса оптимального осреднения для варианта, обеспечивающего выполнение этого условия, могут быть получены путем решения, наряду с системой (5), вспомогательной системы

$$p_i r_i^2 + \sum_{j=1}^n p_j r_{ij} = 1 \quad (i=1, 2, \dots, n), \quad (7)$$

отличающейся от (5) лишь вектором правых частей. Обозначая p'_i и p''_i соответственно решения систем (5) и (7), получаем веса оптимального осреднения с нормировкой весов:

$$p_i = p'_i + \lambda p''_i; \quad \lambda = \frac{1 - \sum_{i=1}^n p'_i}{\sum_{i=1}^n p''_i}. \quad (8)$$

Оптимальное осреднение в обоих вариантах представляет интерес главным образом потому, что оно характеризует максимально возможную точность осреднения изучаемой величины и является тем самым своего рода эталоном, с которым можно сравнивать другие способы осреднения. Однако для практического его применения необходимо знание статистической структуры осредняемой величины, данными о которой мы не всегда располагаем. При отсутствии таких данных приходится использовать другие способы, формальные в том смысле, что они исходят не из данных об особенностях именно осредняемой величины, а, например, из соображений о возможности точного описания процесса полиномами той или иной степени. Из множества таких способов мы ограничимся рассмотрением одного из самых простых — метода трапеций, который, будучи достаточно простым в реализации, обеспечивает довольно большую точность осреднения. В этом случае веса осреднения определяются формулами

$$p_i = \begin{cases} \frac{t_1 + t_2}{2T} & \text{при } i=1; \\ \frac{t_{i+1} - t_{i-1}}{2T} & \text{при } i=2, \dots, n-1; \\ 1 - \frac{t_{n-1} + t_n}{2T} & \text{при } i=n. \end{cases} \quad (9)$$

При этом имеется в виду, что отсчет времени ведется от начала периода осреднения. При осреднении могут использоваться дан-

ные и за пределами интервала осреднения $(0, T)$, но лишь по одному наблюдению до его начала и после его конца.

Расчеты по приведенным выше формулам серьезных трудностей не представляют. Наиболее трудоемким является вычисление величин ω_i , а также двойного интеграла в (4), характеризующего дисперсию величины F_0 . При аналитическом задании корреляционной функции они в большинстве случаев могут быть также представлены аналитическими формулами. В противном случае, а также в случае табличного задания корреляционной функции интегрирование производится численно.

3. По приведенным формулам были выполнены оценки точности осреднения для случайных процессов, характеризующихся различной структурой, и при различном задании сроков измерений. Мы ограничимся приведением некоторых результатов, представляющих наибольший методический интерес.

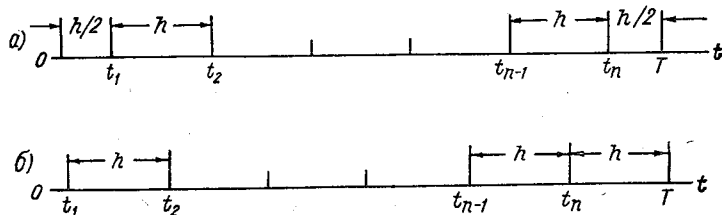


Рис. 1. Расположение моментов измерений на периоде осреднения.

а — оптимальное расположение, б — неблагоприятное расположение моментов измерений.

Как показано в [5], при арифметическом осреднении данных n наблюдений через равные интервалы времени $h = T/n$ наибольшая точность достигается при $t_1 = h/2$ (рис. 1 а). Этот вариант размещения сроков наблюдения будет далее называться оптимальным. Наименьшая точность осреднения достигается для варианта $t_1 = 0$ (рис. 1 б), который мы будем впредь называть неблагоприятным. Разумеется, эти названия чисто условные, как условно здесь и ограничение использованием данных измерения в пределах периода осреднения.

В табл. 1 приведены результаты оценок, выполненных для случая, когда временная структура осредняемой величины описывается корреляционной функцией вида

$$r(\tau) = (1 + |\tau|/T_0) \exp(-|\tau|/T_0), \quad (10)$$

где T_0 — масштаб корреляции.

Оценки даны для различных значений отношения T_0/T , характеризующего связность величины f на интервале осреднения. Погрешности ε_1 относятся к случаю оптимального, а ε_2 — к случаю неблагоприятного размещения сроков наблюдения.

Таблица 1

Средние квадратические (в сотых долях σ) ошибки определения средней по периоду по данным n наблюдений через интервал времени T/n . Корреляционная функция (10), $\eta^2=0,001$

n	T_0/T							
	0,5		1,0		2,0		5,0	
	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2
Арифметическое осреднение								
12	1	5	1	3	1	2	1	1
8	1	7	1	5	1	3	1	2
4	2	14	2	9	2	6	2	3
3	3	19	2	12	2	7	2	4
2	6	28	3	19	2	11	2	5
Оптимальное осреднение								
12	1	1	1	1	1	1	1	1
8	1	2	1	1	1	1	1	1
4	2	4	2	2	2	2	2	2
3	3	8	2	4	2	3	2	2
2	6	16	3	8	2	4	2	3
Оптимальное осреднение с нормировкой весов								
12	1	1	1	1	1	1	1	1
8	1	2	1	1	1	1	1	1
4	2	5	2	3	2	2	2	2
3	3	8	2	4	2	3	2	2
2	6	17	3	9	2	5	2	3
Метод трапеций								
12	1	1	1	1	1	1	1	1
8	1	2	1	1	1	1	1	1
4	2	5	2	3	2	2	2	2
3	3	8	2	5	2	3	2	2
2	6	18	3	10	2	6	2	3

Из таблицы видно, что при оптимальном размещении сроков наблюдений все методы осреднения дают практически одну и ту же точность. В этих условиях вполне удовлетворительные результаты дает арифметическое осреднение, осреднение методом трапеций в нашем варианте также сводится к арифметическому, оптимальное осреднение в обоих вариантах также очень близко к нему.

При неблагоприятном размещении сроков наблюдений переход от арифметического к оптимальному резко повышает точность осреднения. Средняя квадратическая погрешность определения средней величины при этом уменьшается в несколько раз. Особенно большим это увеличение точности является при сравнительно малых значениях отношения T_0/T .

Лишь незначительно уступает оптимальной интерполяции по точности интерполяция с нормировкой весов. Заметные различия между ними как при корреляционной функции (10), так и при других типах статистической структуры наблюдаются лишь при больших ошибках в данных измерений (больших значениях меры ошибки η^2). Учитывая это обстоятельство, а также то, что во всех рассмотренных нами случаях осреднение методом трапеций дает результаты, близкие к оптимальному осреднению с нормировкой весов, мы далее оценок для последнего приводить не будем.

Очень важной является уже упомянутая близость точности осреднения методом трапеций к точности оптимального осредне-

Таблица 2

Средние квадратические (в сотых долях σ) ошибки определения средней по периоду T по данным n наблюдений через интервал времени T/n . Корреляционная функция (11), $\eta^2 = 0,001$

n	T_0/T							
	0,5		1,0		2,0		5,0	
	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2

Арифметическое осреднение

12	5	7	4	6	3	5	2	3
8	7	11	5	9	4	7	3	5
4	14	23	10	17	7	13	5	9
3	19	29	14	23	10	18	6	12
2	29	43	21	35	15	26	9	18

Оптимальное осреднение

12	5	5	4	4	3	3	2	2
8	7	8	5	6	4	4	3	3
4	14	17	10	13	7	9	5	6
3	18	23	13	18	10	13	6	9
2	26	35	20	28	14	21	9	14

Метод трапеций

12	5	5	4	4	3	3	2	2
8	7	8	5	6	4	4	3	3
4	14	18	10	13	7	10	5	6
3	19	26	14	19	10	14	6	9
2	27	41	21	30	15	22	9	14

ния. Поскольку для метода трапеций нет необходимости пользоваться данными о структуре, это дает возможность получать сравнительно точные значения средних величин в самых различных условиях и без предварительного исследования временной корреляции.

Из табл. 2, в которой приводятся оценки для случая, когда временная корреляционная функция осредняемой величины описывается формулой

$$r(\tau) = \exp(-|\tau|/T_0), \quad (11)$$

видно, что сказанное выше о соотношении точности разных способов осреднения справедливо и для этого случая. Переход от

Таблица 3

Средние квадратические ошибки (гПа) определения среднего суточного давления воздуха по данным n наблюдений в сутки через интервал времени 24/ n . Северо-Запад ЕТС

E	Январь				Июль			
	n				n			
	3	4	8	12	3	4	8	12
Арифметическое осреднение								
E ₁	0,29	0,24	0,16	0,14	0,33	0,29	0,20	0,17
E ₂	1,54	1,15	0,58	0,39	0,86	0,66	0,36	0,26
Оптимальное осреднение								
E ₁	0,29	0,23	0,16	0,13	0,34	0,29	0,20	0,17
E ₂	0,47	0,32	0,18	0,15	0,43	0,35	0,22	0,17
Метод трапеций								
E ₁	0,29	0,23	0,16	0,13	0,34	0,29	0,20	0,17
E ₂	0,61	0,40	0,19	0,15	0,46	0,35	0,22	0,17

арифметического осреднения к оптимальному или к осреднению методом трапеций существенно повышает точность при неблагоприятном размещении сроков наблюдений, хотя это уточнение заметно меньше, чем для корреляционной функции (10).

4. Оценки, приведенные в табл. 1 и 2, относятся к довольно большому диапазону изменения входных параметров. Представляет интерес рассмотреть соответствующие оценки применительно к суточному осреднению конкретных метеорологических элементов.

В табл. 3 приводятся оценки средней квадратической погрешности определения среднего суточного давления воздуха на уровне моря для условий Северо-Запада ЕТС. По данным [1, 6], временная структура этого элемента при не слишком больших сдви-

гах времени удовлетворительно описывается формулой (10) при масштабе корреляции $T_0 \approx 30$ ч (следовательно, при суточном осреднении $T_0/T=1,25$). Среднее квадратическое отклонение давления для этого района составляет 14,6 гПа в январе и 7,7 гПа в июле. Мера ошибок измерения составляет 0,001 для января и 0,006 для июля, что соответствует средней квадратической случайной ошибке единичного измерения порядка 0,4 гПа для января и 0,6 гПа для июля. Эти случайные погрешности наряду с ошибками наблюдений по барометру включают в себя ошибки

Таблица 4

Средние квадратические (в сотых долях градуса) ошибки определения средней суточной температуры по данным n наблюдений за сутки через интервал времени 24/ n . Ленинград, ГМО

E	Январь				Июль			
	n				n			
	3	4	8	12	3	4	8	12
Арифметическое осреднение								
E_1	17	9	8	8	34	26	13	11
E_2	77	57	29	21	58	44	22	16
Оптимальное осреднение								
E_1	17	9	8	8	34	26	13	11
E_2	39	21	12	12	49	35	16	12
Метод трапеций								
E_1	17	9	8	8	34	26	13	11
E_2	40	23	12	12	52	36	16	12

приведения давления к уровню моря. Как и ранее, оценки представлены для оптимального (E_1) и неблагоприятного (E_2) размещений числа сроков на интервале осреднения.

Из табл. 3 видно, что при оптимальном выборе сроков измерений в течение суток арифметическое осреднение дает удовлетворительные результаты, средняя квадратическая погрешность средних суточных значений давления не превышает 0,4 гПа даже при четырехразовых наблюдениях и быстро уменьшается с увеличением числа сроков наблюдений. При неблагоприятном выборе сроков средние суточные значения давления получаются с гораздо большей погрешностью. При этом погрешность определения их по восьмисрочным наблюдениям оказывается вдвое больше погрешности определения их по четырем, но оптимально размещенным в течение суток срокам наблюдений. Это является очень важным обстоятельством, особенно если учесть, что действующим

«Наставлением» [7] рекомендуется именно неблагоприятный, с рассматриваемой точки зрения, выбор сроков, а суточные средние рекомендуется определять методом среднего арифметического. Разумеется, изменение сроков наблюдений на сети станций является сложным делом. Однако и для заданной системы сроков наблюдений точность среднего суточного давления может быть повышена в несколько раз уже при использовании метода трапеций.

В табл. 4 приведены результаты аналогичных оценок для температуры воздуха в январе и июле для ГМС Ленинград, ГМО. При этом были использованы данные о временной структуре, полученные в [8]. Данные эти задавались в табличной форме без предельной аналитической аппроксимации, а вычисление интегралов в формулах (4) и (5) производилось численно. При расчетах задавались значения $\sigma_1 = 6,4^\circ\text{C}$; $\sigma_{\text{вп}} = 3,3^\circ\text{C}$; $\eta_1^2 = 0,001$; $\eta_{\text{вп}}^2 = 0,004$.

Следует отметить, что для летнего сезона температура воздуха представляет собой существенно нестационарный процесс, так что выполненные нами для стационарного случая оценки являются не вполне корректными. Однако, как показано в [5], влияние нестационарности в данном случае сравнительно невелико, так что мы можем в первом приближении ориентироваться на оценки, полученные в стационарном варианте, и для июля.

Из табл. 4 видно, что и для температуры воздуха использование рекомендаций Наставления приводит к большой потере точности средних суточных значений, которые при восьмисрочных наблюдениях оказываются не более точными, чем единичные измерения (стандарт случайной ошибки их для температуры воздуха принят нами равным около $0,2^\circ\text{C}$). Конечно, для многих оперативных задач получающаяся при этом погрешность порядка $0,3^\circ\text{C}$ является вполне приемлемой. Следует, однако, иметь в виду, что приведенные нами оценки представляют собой средние квадратические погрешности, которые в отдельных случаях могут превышать в несколько раз. Кроме того, полученные средние суточные значения используются в дальнейшем для формирования средних декадных, средних месячных и т. д. значений, точность которых будет существенно зависеть от точности средних суточных значений. Поэтому уточнение средних суточных значений величин, хотя бы на этапе климатологической обработки, представляется весьма желательным. Как видно из табл. 4, использование метода трапеций, а тем более оптимального осреднения, позволяет добиться значительного уточнения, особенно для зимних условий.

5. До сих пор мы оценивали возможности получения средних суточных значений по заданному числу наблюдений в сутки, без привлечения наблюдений в смежные сутки. Очевидно, можно рассчитывать на получение более точных средних значений величин при привлечении данных из предыдущих или (и) последующих суток. В варианте расположения моментов отсчета, названном на-

ми неблагоприятным, когда первое наблюдения приходится на начало суток, можно привлечь приходящееся на конец данных суток первое наблюдение последующих суток (или — в случае, когда последнее наблюдение данных суток приходится на их конец — можно привлечь совпадающее с началом данных суток последнее наблюдение предыдущих суток). Тем самым можно увеличить информацию, используемую для расчета средних суточных значений, без какого-либо изменения сроков наблюдений. Некоторые оценки того, что может дать такой прием применительно к определению составляющих скорости ветра в свободной атмосфере, представлены в табл. 5. Эти оценки получены в предположении, что временная корреляция для составляющих скорости ветра на уровне 500 гПа описывается экспоненциальной формулой (11)

Таблица 5

Средняя квадратическая ошибка (в сотых долях σ) средних суточных значений составляющих скорости ветра на уровне 500 гПа по данным n наблюдений. Корреляционная функция (11). $T_0 = 48$ ч, $\eta^2 = 0,01$

Способ осреднения	n								
	2		3		4			8	
	ε_1	ε_2	ε_1	ε_2	ε_1	ε_2	ε_3	ε_1	ε_2
Арифметическое	16	27	11	19	9	14	10	5	8
Оптимальное	16	22	11	14	9	11	9	5	6
Метод трапеций	16	23	11	15	9	11	9	5	6

три масштабе корреляции $T_0 \approx 48$ ч (см., например, [2, 3]), а мера ошибок в исходных данных составляет около 0,01, что соответствует средней квадратической случайной ошибке для составляющих скорости ветра порядка 1 м/с.

Наряду с оценками, которые нами, как и ранее, обозначаются ε_1 и ε_2 , в табл. 5 для случая четырехсрочных измерений представлена величина ε_3 , соответствующая случаю привлечения данных межных суток, т. е. расчетам по данным наблюдений в 0, 6, 12, 8 и 24 ч (0 ч последующих суток). Привлечение пятого наблюдения заметно повышает точность уже при арифметическом осреднении. При использовании же метода трапеций или оптимального осреднения точность резко повышается и привлечение пятого наблюдения делает точность осреднения при неблагоприятном размещении сроков наблюдений близкой к точности осреднения при их оптимальном размещении.

Поскольку в настоящее время как наземные, так и аэрологические наблюдения производятся в единые синоптические сроки, о и метеорологические сутки принято задавать по единому, например гринвичскому или московскому декретному, времени. Это позволяет установить большее единообразие в соотношении между

сроками наблюдений и пределами метеорологических суток. В результате, однако, возникает опасность несогласования характеристик суточных значений, полученных для различных часовых поясов, в которых соотношение между сутками по солнечному времени и принятыми сутками по единому времени может быть различным. Поэтому при климатологической обработке может оказаться целесообразным определять средние суточные значения для суток, определяемых по местному солнечному времени, начало и конец которых при восьмисрочных наблюдениях могут быть на 1,5 ч, а при четырехсрочных наблюдениях даже на 3 ч смещены относительно сроков наблюдений. В этих условиях желательно использовать вместо арифметического точные способы осреднения, которые позволяют получить более точные и более сравнимые характеристики на различных станциях.

Таблица 6

Средняя квадратическая ошибка (в сотых долях σ) средних суточных значений составляющих скорости ветра на уровне 500 гПа по данным измерений через 6 ч. Корреляционная функция (11), $T_0=48$ ч, $\gamma^2=0,01$

Способ осреднения	Сроки наблюдений, ч											
	0 6 12 18 —	0 6 12 18 24	1 7 13 19 1	2 8 14 20 2	3 9 15 21 3	3 9 15 21 —	0 12 18 24 —	0 12 18 —	1 13 19 —	2 14 20 —	3 15 21 —	
Арифметическое	14	10	11	12	14	9	15	15	14	13	14	
Оптимальное	11	9	9	8	9	9	12	14	13	13	13	
Метод трапеций	11	9	9	9	9	9	13	14	13	13	13	

Насколько существенным это может оказаться, видно из табл. 6, в которой приведены оценки точности разных способов суточного осреднения по данным четырехразовых (с интервалом 6 ч) измерений. Рациональное использование пяти измерений вместо четырех позволяет при этом получить оценки средних суточных не менее точные, чем по четырем измерениям в оптимальные сроки, чего при арифметическом осреднении добиться не удастся.

Таким образом, можно сделать вывод, что принятое в настоящее время для получения средних суточных значений арифметическое осреднение во многих случаях приводит к нежелательному загроблению этих значений. При климатологической обработке, а возможно, и в оперативном порядке целесообразно производить осреднение более точно. Наибольшую точность, в принципе, обеспечивает оптимальное осреднение, для которого, однако, требуются детальные данные о временной структуре исследуемой величины. Как показывают выполненные нами оценки, весьма удовлетворительные результаты дает уже метод трапеций, который заслуживает первоочередного внедрения. Укажем в заключение, что

тот же метод трапеций оказывается очень удобным для получения оценок средних суточных значений при наличии пропусков в те или иные сроки наблюдений, что особенно актуально для аэрологических наблюдений на больших высотах. Это хорошо видно из табл. 6, в которой приведены также примеры оценок точности средних суточных значений составляющих скорости ветра при отсутствии наблюдений в те или иные сроки.

СПИСОК ЛИТЕРАТУРЫ

1. Антоненко В. С., Лугина К. М. Пространственно-временная структура поля давления для летнего сезона. — Труды ГГО, 1977, вып. 397, с. 70—77.
2. Гарифулин К. К. Изменчивость ветра в свободной атмосфере. — Л.: Гидрометеиздат, 1967. — 143 с.
3. Гутерман И. Г., Игнатьшина Е. И. Корреляционные функции характеристик ветра. В кн.: Применение статистических методов в метеорологии: Труды Всесоюзного симпозиума по применению статистических методов в метеорологии. Л., Гидрометеиздат, 1971, с. 67—89.
4. Каган Р. Л. Осреднение метеорологических полей. — Л.: Гидрометеиздат, 1979. — 213 с.
5. Каган Р. Л., Сибир Е. Е. О точности осреднения периодически коррелированных величин. — Труды ГГО, 1980, вып. 431, с. 90—105.
6. Лугина К. М., Каган Р. Л. К вопросу о пространственно-временном анализе барического поля. — Труды ГГО, 1974, вып. 336, с. 75—94.
7. Наставление гидрометеорологическим станциям и постам. Вып. 3, ч. 1. — Л.: Гидрометеиздат, 1969. — 308 с.
8. Федорченко Е. И. О временной структуре ежечасных наблюдений за температурой воздуха в Ленинграде. — Труды ГГО, 1975, вып. 348, с. 112—122.

Р. Л. Каган, К. М. Лугина

НЕКОТОРЫЕ ОЦЕНКИ ВОЗМОЖНОСТЕЙ МОНИТОРИНГА ОБЩЕГО СОДЕРЖАНИЯ ОЗОНА В АТМОСФЕРЕ

1. В последние годы появились сообщения о возможном антропогенном уменьшении озона в атмосфере [2, 8]. Такое уменьшение может иметь далеко идущие последствия как для условий существования биологических объектов на Земле, так и для климата планеты. Поэтому возникла необходимость тщательного изучения этой проблемы как путем теоретических расчетов, так и путем статистического анализа имеющихся данных наблюдений за озоном. Одним из аспектов такого анализа является рассмотрение имеющихся рядов озонометрических наблюдений с целью выявления в них возможных трендов.

В этом отношении заслуживают внимания работы Дж. Эндрюса и Дж. Коршвера [10, 11, 12], в которых делается попытка изучения хода общего содержания озона (ОСО) в атмосфере при-

мерно за 20-летний период существования глобальной озонометрической сети. Справедливо полагая, что данные наблюдений за озоном в отдельных пунктах не являются достаточно репрезентативными для исследования векового хода этого элемента, упомянутые авторы оперируют данными, осредненными по большим площадям.

Разумеется, для того чтобы судить об обоснованности результатов, получаемых в работах такого рода, необходимо представлять себе количественно возможную точность пространственного осреднения данных довольно грубых озонометрических наблюдений на сравнительно небольшом числе станций. В настоящей статье делается попытка выполнить такие оценки.

При этих оценках используется разработанный авторами [3, 4, 5] применительно к данным наблюдений за осадками и температурой воздуха статистический подход, сводящийся к априорной оценке точности пространственного осреднения по данным о пространственной структуре осредняемой величины. При наличии таких данных можно оценить точность средних, полученных по наблюдениям в любой мыслимой сети пунктов и при любой схеме осреднения. Эта методика позволяет применять наиболее эффективные способы осреднения данных существующей сети, а также оценивать различные варианты изменения сети на будущее.

Степень обоснованности получающихся при этом выводов, конечно, зависит от надежности данных о статистической структуре, используемых при расчетах, главным образом о пространственной корреляции осредняемого элемента.

В настоящее время имеется довольно много данных о пространственной структуре средних суточных значений ОСО (см., например, [1, 7, 9, 13]).

Однако применительно к средним месячным значениям ОСО, которые, на наш взгляд, представляют наибольший интерес для целей мониторинга климата, имеются лишь отрывочные данные. Так, например, в работе [16] приводятся коэффициенты корреляции средних значений ОСО лишь для нескольких пар станций. В работе Р. Уилкокса [19] приводятся некоторые оценки характеристик статистической структуры средних сезонных и даже средних годовых значений ОСО, полученные косвенным путем, по данным о пространственной и временной изменчивости суточных значений величин. Оценки эти, однако, получены по предложенной С. Лийсом [14] крайне упрощенной схеме и вряд ли заслуживают большого доверия. Заметим, кстати, что в связи с этим представляется неясной и возможность использования оценок значимости тренда, полученных Уилкоксом на основе этих данных о структуре.

Мы будем пользоваться для описания пространственной корреляционной функции (r) ОСО формулой вида

$$r(\rho) = \exp(-\rho/\rho_0), \quad (1)$$

где ρ — расстояние между точками, ρ_0 — масштаб корреляции.

В данных наблюдений содержатся шумы за счет ошибок измерения и мелкомасштабных флуктуаций, что приводит к занижению эмпирической корреляционной функции $\tilde{r}(\rho)$ по сравнению с истинной. Экстраполируя эту функцию на нуль, можно по значению $\tilde{r}(0)$ оценить меру ошибок наблюдений η^2 , представляющую собой отношение дисперсии случайных шумов к дисперсии σ^2 самой величины ОСО:

$$\eta^2 = \frac{1 - \tilde{r}(0)}{\tilde{r}(0)}. \quad (2)$$

Выполненные в ГГО проработки [6] показывают, что для месячных значений ОСО характерны существенно другие параметры структуры, чем для суточных. В частности, масштаб корреляции ρ_0 для месячных значений оказывается близким к 1600 км, что заметно больше, чем для суточных, для которых обычно получаются значения $\rho_0 \approx 1000$ км. Предварительные оценки дают для средних месячных значений ОСО $\eta^2 = 0,1 \dots 0,2$. Мы будем далее принимать η^2 равным 0,1, что соответствует точности средних месячных значений ОСО для января по данным измерений спектрофотометром Добсона.

Несомненно, что в дальнейшем параметры формулы (1) могут потребовать уточнения. Может быть введена детализация их по сезонам, различным географическим зонам, а также по разным направлениям, что позволит отказаться от неизбежного на данном этапе использования гипотезы об однородности и изотропии пространственных корреляционных функций ОСО. Имеющийся опыт показывает, однако, что хотя при такой детализации оценки точности несколько изменяются, порядок их остается прежним и, что самое главное, соотношение между точностью различных методов осреднения или между точностью осреднения по различным наборам станций остается практически неизменным. Можно надеяться поэтому получить разумные методические выводы уже на данном этапе.

В качестве критерия точности осреднения мы будем пользоваться относительной средней квадратической ошибкой ε , представляющей собой отношение средней квадратической ошибки осреднения к среднему квадратическому отклонению (стандарту) точечных значений осредняемой величины σ , и величиной $\tilde{\varepsilon}$ — отношением стандарта ошибки осреднения к среднему квадратическому отклонению $\sigma_{\tilde{\sigma}}$ осредненной по площади величины.¹ Изменчивость последней меньше изменчивости точечных значений, так что $\tilde{\varepsilon} > \varepsilon$; для больших площадей осреднения эти различия могут быть очень заметными. Вывод формул, использующихся для расчета величин ε и $\tilde{\varepsilon}$, и соответствующие вычислительные алгорит-

¹ Эти ошибки следует отличать от величин, получающихся путем деления абсолютных значений ошибок не на стандарт, а на саму измеряемую величину, которую также нередко называют относительной ошибкой.

мы детально изложены в [3—5]. Поэтому мы далее ограничимся приведением лишь самых необходимых формул, без вывода.

2. Выше уже говорилось о недостаточной надежности локальных значений ОСО. Это связано с высоким уровнем ошибок измерений, в силу чего по наблюдениям на станции средние месячные значения на ней же получаются с относительной погрешностью ε более 30 %.¹ С невысокой точностью могут быть получены и значения ОСО, осредненные по площадям, на которых имеется не больше одной станции. Поскольку сеть озонметрических стан-

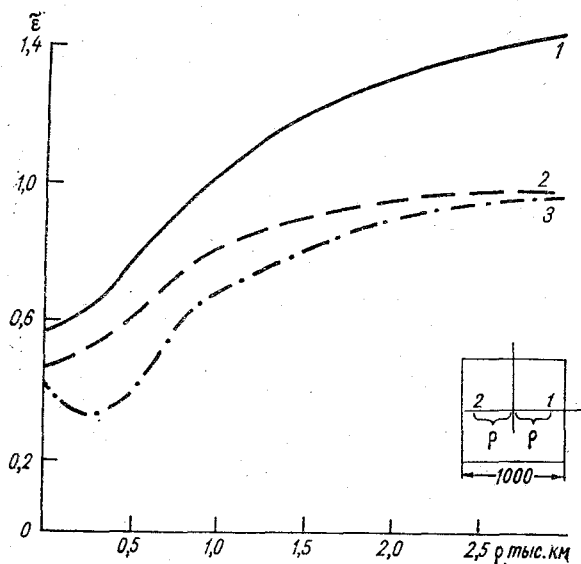


Рис. 1. Средняя квадратическая погрешность ε определения средних месячных значений ОСО на квадрате площадью 1 млн. км².

1 — замена средней по площади данными наблюдений на станции 1 (см. схему в правом нижнем углу); 2 — оптимальная оценка средней по площади по данным станции 1; 3 — оптимальная оценка средней по площади по данным станций 1 и 2.

ций относительно редка (одна станция может приходиться на несколько миллионов квадратных километров), такие площади могут быть довольно велики.

Так, для площадей порядка 1 млн. км², что близко к площади широтно-долготного «квадрата» $10 \times 10^\circ$ в субтропических широтах погрешность ε определения средней по данным одного измерения

¹ Здесь и далее речь идет лишь о случайной составляющей ошибок. В данных озонметрических наблюдений могут содержаться и значительные систематические ошибки. Мы этого вопроса касаться не будем, полагая, что систематические ошибки путем надлежащей обработки данных могут быть исключены из рассмотрения.

превышает 50 % даже в том случае, если станция находится в пределах площади осреднения. Это хорошо видно из рис. 1, на котором приведена зависимость относительных ошибок ε от расстояния между станцией и центром квадрата. Тем более велики ошибки в случае, если в интересующем нас квадрате нет ни одной станции и среднее значение по нему приходится оценивать путем экстраполяции или интерполяции данных станций, расположенных за его пределами, иногда на очень больших расстояниях. Такая ситуация является обычной для обширных районов земного шара, прежде всего для океанических акваторий.

Из рис. 1 видно, что относительная погрешность распространения на квадрат данных ближайшей станции (кривая 1) уже на расстоянии порядка 1000 км превышает единицу, т. е. она больше ошибки, которую мы допустили бы, заменив фактические данные наблюдений средними многолетними значениями, конечно, при условии, что последние известны точно. В последнем случае гораздо более эффективно не прямое распространение на окружающую территорию данных наблюдений на станции, а их приведение («редукция») к средней по интересующей нас площади с учетом корреляции осредненных и точечных значений. В этом случае (кривая 2 на рис. 1) значения ε (по крайней мере теоретические) не превышают единицу, и для площади рассматриваемого масштаба имеет смысл использование данных наблюдений на станциях, расположенных до расстояний порядка 2000 км от центра площади.

При отсутствии станций в интересующей нас области естественно пытаться привлекать данные наблюдений не с одной, а с нескольких близлежащих станций, т. е. заменять экстраполяцию данных наблюдений интерполяцией. Однако привлечение дополнительных станций оказывается и в этом случае эффективным лишь для достаточно густой сети. Это обстоятельство иллюстрируется кривой 3 рис. 1, которая характеризует ошибки получения средней по площади путем оптимального осреднения данных двух наблюдений, расположенных симметрично относительно центра квадрата на расстоянии ρ от него (т. е. на расстоянии 2ρ друг от друга). Привлечение второй станции резко повышает точность осреднения, пока станции находятся в квадрате или вблизи него. При расстоянии же между станциями более 2000 км добавление второй станции дает небольшой эффект. Малоэффективным в таких условиях является и привлечение для оценки средней по площади данных трех и более станций.

Таким образом, для получения надежных средних значений величин желательно осреднение производить по достаточно большим площадям, на которых располагалась бы не одна, а несколько станций. При этом наиболее выгодным с точки зрения анализа ОСО на территории вообще и, в частности, для получения возможно более точных средних значений величин на площади является равномерное размещение на ней пунктов наблюдений. Заметим,

что в этом случае особенно простым является осреднение, которое может производиться методом среднего арифметического. Естественно, что при этом сравнительно простыми оказываются и формулы для оценки точности осреднения, которые для случая экспоненциальной корреляционной функции (1) имеют вид

$$\varepsilon(n, s) \approx \sqrt{\frac{1}{n}(\eta^2 + 0,23\sqrt{s/\rho_0})}; \quad (3)$$

$$\tilde{\varepsilon}(n, s) \approx \exp(0,245\sqrt{ns/\rho_0}) \varepsilon(n, s). \quad (4)$$

Здесь n — число станций в области осреднения, s — площадь, приходящаяся на одну станцию.

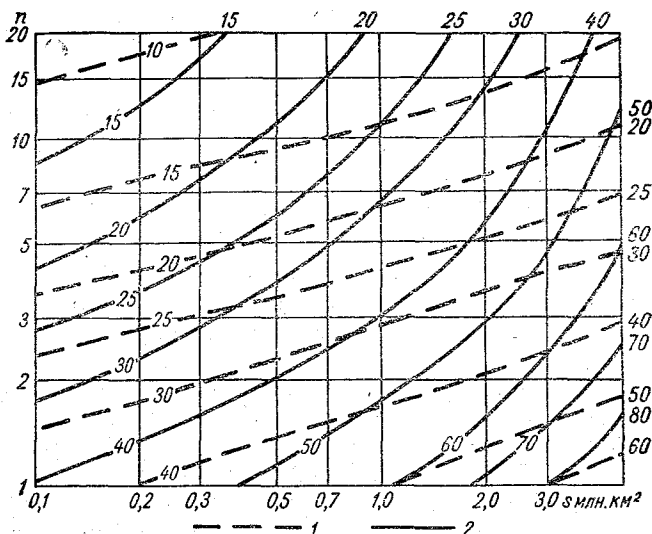


Рис. 2. Относительные средние квадратические ошибки (%) определения средних месячных значений ОСО на площади при равномерном размещении на ней n станций при плотности сети s млн. км² на станцию.

1 — $\varepsilon(n, s)$; 2 — $\tilde{\varepsilon}(n, s)$.

На рис. 2 представлена зависимость точности осреднения от n и s для ОСО. Из него видно, что погрешность ε (кривая 1), характеризующая точность осреднения по отношению к фиксированной величине σ , быстро уменьшается с увеличением n как при густой, так и при редкой сети. Однако для целей мониторинга климата больший интерес представляет погрешность $\tilde{\varepsilon}$, выраженная в долях изменчивости не точечных, а осредненных по площади величин (кривая 2). Эта изменчивость быстро уменьшается с ростом площади осреднения, т. е. с увеличением n при фиксированном s . Поэтому величина $\tilde{\varepsilon}$ всегда больше, чем ε , особенно

три больших s , для которых уменьшение ε с ростом n происходит очень медленно. В итоге оказывается, что для $s > 5$ млн. км² ε превышает 50 % при осреднении по областям площадью до 100 млн. км², т. е. для столь редкой сети информационная значимость пространственного осреднения повышается с увеличением площади очень мало.

Следует иметь в виду, что формула (3) и приведенные на рис. 2 оценки относятся к случаю равномерной сети станций, когда область осреднения может быть разбита на элементы равной площади s , в центре каждого из которых расположена станция. Реальная сеть станций, конечно, не удовлетворяет этим условиям, и при оценке возможной точности осреднения по данным этой сети приходится учитывать степень ее неравномерности. Для случая экспоненциальной структуры поля для этой цели может использоваться приближенная формула [4]

$$\varepsilon_1(n, s) = P\varepsilon(n, s), \quad (5)$$

где P — множитель, характеризующий увеличение средней квадратической погрешности осреднения при неравномерной сети по сравнению с погрешностью в случае равномерного размещения в этой области того же числа станций n ; s — средняя площадь, приходящаяся на одну станцию. Предполагается, что в случае неравномерной сети используется не средняя арифметическая, а средняя взвешенная, при нахождении которой данные станции входят с весами, пропорциональными площади тяготеющих к ним элементов области осреднения. Конкретно использовался метод полигонов [18]. При этом получено

$$P(\zeta, \alpha) = \sqrt{\frac{(1 + 1,6 \alpha^{1,4})^2 + 4,33(1 + \alpha^2)^2}{1 + 4,33 \zeta}}, \quad (6)$$

где $\zeta = \eta^2 \rho_0 / \sqrt{s}$ — безразмерный параметр, учитывающий плотность сети и точность наблюдений; α — коэффициент неравномерности сети, представляющий собой коэффициент вариации площадей элементов осреднения (или, что то же самое, весов осреднения методом полигонов).

В табл. 1 приведены полученные по формуле (5) оценки влияния неравномерности сети на точность осреднения. Из нее видно, что даже небольшая неравномерность сети ($\alpha < 0,5$) заметно уменьшает точность осреднения, особенно при редкой сети. Оказывается, например, что при $s = 5$ млн. км², $\alpha = 0,4$ та же точность осреднения может быть достигнута при использовании вдвое меньшего числа равномерно расположенных в области осреднения станций. Еще больше оказывается влияние неравномерности при больших значениях α .

3. Формулы (3) — (6) получены для случая двумерного осреднения на плоскости. При оценках точности осреднения по большим регионам или по широтным зонам необходимо учитывать сферичность Земли. В этом случае расчеты несколько усложняются.

ются. Они могут быть выполнены в соответствии с методикой, детально изложенной в [5, 6] в связи с оценкой точности осреднения температуры воздуха.

Применительно к ОСО оценки выполнялись нами для сети станций различной густоты и для разных способов осреднения. Наряду с погрешностью арифметического осреднения определялись погрешности осреднения методом полигонов и оптимального осреднения. Были проведены две серии расчетов. В первой из них рассматривались квазиравномерные¹ сети различной густоты на

Таблица 1

Влияние неравномерности сети станций ($P(\alpha, s)$) на точность осреднения
месячных значений ОСО ($\rho_0 = 1600$ км, $\eta^2 = 0,1$)

α	s млн. км ²			
	0,2	1,0	5,0	$(\eta_{\infty}^2 = 0)$
0	1,0	1,0	1,0	1,0
0,4	1,2	1,3	1,4	1,4
0,8	1,5	1,9	2,0	2,2
1,2	2,3	2,6	2,8	3,1
1,6	3,0	3,4	3,7	4,1
2,0	3,7	4,3	4,7	5,2

северном полушарии. Результаты этой серии оценок представлены в табл. 2, в которой приведены значения относительной средней квадратической ошибки ε для разных способов осреднения.

Будь сеть вполне равномерна, веса и соответственно точности для оптимального осреднения с нормировкой весов, арифметического и методом полигонов должны были бы совпадать. В действительности такого совпадения не наблюдается. Расхождения однако сравнительно невелики и не оправдывают перехода от арифметического к более сложным способам осреднения. Наряду с точными оценками погрешности осреднения, в табл. 2 приведены также оценки $\varepsilon(n, s)$, полученные по формуле (3). Они оказываются достаточно близкими к точным значениям. Несколько меньшие (на 10—15 % по сравнению с погрешностями осреднения методом полигонов) величины ε , полученные по формуле (3) можно отчасти приписать фактической неравномерности сети.

¹ В силу понятных причин размещение на сфере вполне равномерной сетки пунктов невозможно. Можно говорить лишь о большей или меньшей степени равномерности. В нашем случае выбор расположения станций по предложенному Л. Н. Магазенковым способу обеспечивал тем большую равномерность, чем больше число станций на полушарии.

При оценке информативности сети станций по данным табл. 2 не следует обольщаться малыми значениями ϵ , которые характеризуют отношение ошибки среднего по полушарию к σ точечных значений. Поскольку среднее квадратическое отклонение средней по полушарию составляет в нашем случае $0,222 \sigma$, то действительная относительная ошибка ϵ будет в 4,5 раза больше значений ϵ .

Выбор критерия значимости данных для целей мониторинга климата является сложным делом. Следует иметь в виду, что для этих целей нужно иметь надежные данные не только за отдельные годы, но и по изменениям от года к году, что повышает требования к исходным данным. Будем условно полагать, что данные за отдельные годы значимы в том случае, если средняя квадратическая погрешность их определения составляет не более половины их среднего квадратического отклонения, т. е. при $\epsilon < 0,11$.

Таблица 2

Точность осреднения средних месячных значений ОСО на северном полушарии по данным квазиравномерной сети станций

n	млн. км ²	Относительная средняя квадратическая погрешность (ϵ) осреднения				
		оптимального	оптимального с нормировкой	методом полигонов	арифметического	По формуле (3)
176	1,5	0,042	0,044	0,043	0,048	0,039
137	1,7	0,050	0,052	0,051	0,055	0,046
103	2,5	0,060	0,063	0,062	0,066	0,056
73	3,5	0,075	0,081	0,080	0,080	0,071
48	5,3	0,098	0,109	0,108	0,108	0,095
29	8,8	0,129	0,155	0,156	0,150	0,135
14	18,2	0,173	0,239	0,247	0,242	0,226

ратическая погрешность их определения составляет не более половины их среднего квадратического отклонения, т. е. при $\epsilon < 0,11$. В этом случае мы приходим к выводу, что для получения надежных значений ОСО для всего полушария необходимо иметь не менее 50 равномерно расположенных на нем станций.

В действительности, однако, реальная сеть станций далеко неравномерна. Поэтому во второй серии расчетов рассматривалась сеть, состоящая из действующих или действовавших ранее озонметрических станций, оснащенных спектрофотометром Добсона или озонметром М-83, данные которых публиковались в [17]. Максимальное число станций этой сети на северном полушарии составляло 173. Однако одновременно в каждый период времени на северном полушарии действует около 100 станций такого типа и лишь несколько десятков станций работают в течение длительного времени. Поэтому наряду с сетью максимальной густоты рассматривались различные выборки из нее. Ряд таких выборок был получен путем последовательного прореживания сети за счет

отсеивания наименее значимых станций. Значимость станции определялась по величине тяготеющей к ней площади, т. е. по весу при осреднении методом полигонов. Каждая последующая выборка получалась из предыдущей путем исключения станций с наименьшими весами.

Некоторые оценки точности определения средних месячных значений ОСО для северного полушария по всей сети и по выборкам из нее даны в табл. 3. Там же приведены оценки погрешностей полученных по приближенным формулам (3)—(6) для плоского

Таблица 3

**Точность осреднения средних месячных значений ОСО
на северном полушарии по данным сети озонметрических станций**

n	S млн. км ² на станцию	Средняя квадратическая погрешность (ε) осреднения				
		оптимального	оптимального с нормировкой	методом полигонов	арифмети- ческого	по формулам (3)—(6)

Сети, получающиеся путем последовательного прореживания исходной

α								
173	1,5	1,7	—	—	0,14	0,27	0,14	
150	1,7	1,5	—	—	0,14	0,23	0,14	
120	2,1	1,3	—	—	0,14	0,19	0,14	
90	2,8	1,0	—	—	0,14	0,16	0,14	
70	3,6	0,8	—	—	0,14	0,15	0,14	
50	5,1	0,6	—	—	0,14	0,15	0,15	
30	8,5	0,5	—	—	0,16	0,16	0,21	
20	12,8	0,3	—	—	0,20	0,20	0,22	

Сеть станций, данные которых регулярно публикуются в [17]

68	3,8	1,2	0,13	0,18	0,20	0,26	0,21	
----	-----	-----	------	------	------	------	------	--

Сеть станций, используемая в [10]

32	8,0	0,8	0,15	0,22	0,24	0,26	0,25	
----	-----	-----	------	------	------	------	------	--

варианта с учетом неравномерности. Согласование их с характеристиками точности осреднения методом полигонов представляется очень хорошим.

Из табл. 3 видно, что уменьшение числа станций на полушарии до 50 практически не уменьшает точности осреднения методом полигонов, а точность арифметического осреднения при этом даже увеличивается. Это связано с принятым нами алгоритмом прореживания, который исключает станции, расположенные в районах с густой сетью и оставляет станции, находящиеся в районах с редкой сетью, в результате чего сеть становится более равномерной. Соответственно при $n < 50$ ошибки осреднения методом

полигонов оказываются близкими к ошибкам арифметического осреднения. Отсюда следует, что для повышения точности осреднения по полушарию бесполезно привлекать дополнительные наблюдения в районах с относительно густой (по сравнению со средней густотой) сетью станций. Дополнительные наблюдения должны организовываться в районах, где в существующей сети станций имеются значительные пробелы, в первую очередь на акваториях океана. В качественном отношении этот вывод является, конечно, тривиальным, но сравнение данных табл. 3 и 2 позволя-

Таблица 4

Точность осреднения месячных значений ОСО по широтным зонам северного полушария по данным сети [10]

Широтная зона (°с. ш.)	$\frac{\sigma S}{\sigma}$	n	Характеристика	Относительная средняя квадратическая погрешность осреднения			
				оптимального	оптимального с нормировкой	методом полигонов	арифметического
По данным станций в зоне							
20—60	0,27	23	ε	0,18	0,25	0,26	0,27
			~				
			ε	0,66	0,93	0,99	1,01
60—80	0,45	7	ε	0,28	0,38	0,42	0,44
			~				
			ε	0,63	0,84	0,93	0,96
С привлечением станций из других зон							
20—60	0,27	35	ε	0,17	0,20	0,23	0,24
			~				
			ε	0,63	0,77	0,86	0,89
60—80	0,45	30	ε	0,27	0,28	0,34	0,39
			~				
			ε	0,59	0,62	0,75	0,87

ет количественно оценить выигрыш, получаемый за счет равномерного расположения озонметрических пунктов на полушарии.

Наряду с оценками, описанными нами выше, в табл. 3 приведены также значения погрешностей осреднения для двух других вариантов, которые также получены как выборки из сети 173 станций, но не путем формального прореживания на ЭВМ, а исходя из других соображений. В одной из них отобрано 68 станций, данные которых регулярно публиковались в сводках [17]. Они использовались в ГГО при оценке характеристик статистической структуры средних месячных значений ОСО. Вторая выборка состоит из 32 станций северного полушария, используемых в работе Энджела и Коршова [10] для исследования трендов общего

содержания озона в атмосфере. Сети эти характеризуются большой неравномерностью, в частности, они гораздо неравномернее полученных нами путем прореживания исходной сети сетей с примерно тем же количеством станций, представленных в верхней части табл. 3. Соответственно гораздо большими являются и ошибки осреднения. Не только для арифметического, но и для осреднения методом полигонов они имеют тот же порядок, что и естественная изменчивость среднего по полушарию ($\sigma_s = 0,22\sigma$). Это делает сомнительной возможность использования данных подобной сети для целей мониторинга ОСО на полушарии при использовании этих способов осреднения. Перспективным в этом отношении является использование оптимального осреднения, которое в условиях редкой сети позволяет получить заметно более точные средние значения величины. Однако для этого необходимо сначала иметь достаточно надежные значения норм общего содержания озона как в отдельных пунктах, так и по площадям. Насколько пригодны для этой цели имеющиеся в настоящее время статистические данные, представляется пока неясным.

Аналогичные выводы получаются и применительно к оценкам средних значений по широтным зонам по данным реально существующей сети. В качестве примера в табл. 4 представлены оценки погрешности осреднения ОСО по двум широтным зонам по данным сети [10]. Оказывается, при использовании традиционных методов осреднения эти погрешности имеют тот же порядок, что и изменчивость искомой величины (она характеризуется в табл. 4 отношением среднего квадратического отклонения σ_s осредненной по площади величины к стандарту точечных значений σ). Заметно повышает точность использование оптимального осреднения.

Определенного уточнения средней величины можно добиться, если не ограничиваться данными станций, расположенных непосредственно в области осреднения, а привлекать также станции из других зон, лежащие на не слишком большом расстоянии от области осреднения. В нашем случае привлечение станций из соседних зон (для зоны 20—60° с. ш. используются данные 3 станций из южного полушария) повышает точность осреднения арифметического и методом полигонов на 10—15 %. Однако и в этом случае сколько-нибудь приемлемую точность обеспечивает лишь оптимальное осреднение.

Таким образом, в настоящее время трудно рассчитывать на надежное определение глобальных характеристик общего содержания озона в атмосфере за отдельные годы и их межгодовой изменчивости по данным традиционных наблюдений озонметрической сети станций. Некоторого улучшения информативности этих характеристик можно добиться путем тщательной статистической обработки имеющегося материала, в частности, путем применения оптимального осреднения, использования связей ОСО с другими элементами, например с полем геопотенциала в стратосфере (см., например, [15]). Радикального улучшения ситуации можно ожи-

дать лишь после усовершенствования спутниковых озонметрических методов и разработки методики их эффективного использования в комплексе с наземными измерениями.

СПИСОК ЛИТЕРАТУРЫ

1. Гриб Н. К., Поляк И. И. Некоторые статистические характеристики измерений атмосферного озона.— Метеорология и гидрология, 1973, № 4, с. 87—93.
2. Заявление ВМО о воздействии на слой озона в результате деятельности человека и о некоторых вероятных геофизических последствиях.— Бюлл. ВМО, 1976, т. 24, № 1, с. 74—79.
3. Каган Р. Л. К оценке репрезентативности осадкомерных данных.— Труды ГГО, 1966, вып. 191, с. 122—132.
4. Каган Р. Л. Осреднение метеорологических полей.— Л.: Гидрометеониздат, 1979.— 213 с.
5. Каган Р. Л., Лугина К. М. О точности пространственного осреднения поля температуры воздуха.— Труды ГГИ, 1979, вып. 257, с. 50—66.
6. Каган Р. Л., Салль И. А., Щербакова Н. Н. О пространственной корреляции средних месячных значений общего содержания озона в атмосфере.— См. наст. сб.
7. Покровский О. М., Иwanyкин Е. Е., Кайгородцев А. Е. Численное моделирование систем наблюдения и анализа поля озона.— Метеорология и гидрология, 1979, № 1, с. 56—65.
8. Роулэнд Ф. С. Возможное влияние деятельности человека на стратосферный озон.— Бюлл. ВМО, 1976, т. 24, № 2, с. 97—106.
9. Шафрин Ю. М. О статистических характеристиках озоносферы.— Изв. АН СССР. Физика атмосферы и океана, 1966, т. 2, № 6, с. 647—655.
10. Angell J. K., Korshover J. Quasi-biennial and long-term fluctuations in total ozone.— Mon. Wea. Rev., 1973, vol. 101, N 5, p. 426—443.
11. Angell J. K., Korshover J. Global analysis of recent total ozone fluctuations.— Mon. Wea. Rev., 1976, vol. 104, N 1, p. 63—75.
12. Angell J. K., Korshover J. Global ozone variations: An update into 1976.— Mon. Wea. Rev., 1978, vol. 106, N 5, p. 725—737.
13. Bojkov R. D. Some characteristics of the total ozone deduced from Dobson-spectrophotometer and their application to a determination of the effectiveness of the ozone station network.— Ann. Geoph., 1966, vol. 25, N 1, p. 293—299.
14. Leith C. E. The standard error of time-average estimates of climatic means.— J. Appl. Meteorol., 1973, vol. 12, N 6, p. 1066—1069.
15. Miller A. J. e. a. Utilization of 100 mb midlatitude height fields as an indicator of sampling effects on total ozone variations.— Mon. Wea. Rev., 1979, vol. 107, N 6, p. 782—787.
16. Newell R. E. Wu M.-F. A pilot study of concomitant changes in total ozone and atmospheric general circulation.— Quart. J. Roy. Meteorol. Soc., 1978, vol. 104, N 442, p. 999—1003.
17. Ozone data for the world.— WMO, Toronto, Ontario, 1961—1977.
18. Thiessen A. H. Precipitation averages for large areas.— Mon. Wea. Rev., 1911, vol. 39, N 7, p. 1082—1084.
19. Wilcox R. W. Total ozone trend significance from space and time variability of daily Dobson data.— J. Appl. Meteorol., 1978, vol. 17, N 3, p. 405—409.

ОБ ОДНОМ АЛГОРИТМЕ СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ РЯДОВ МЕТЕОРОЛОГИЧЕСКИХ НАБЛЮДЕНИЙ

1. Статистическое моделирование данных метеорологических наблюдений с большой детальностью представляет значительные трудности, связанные с особенностями их корреляционной структуры для сравнительно малых (меньше суток) интервалов времени. Одна из этих трудностей связана с существенно различным ходом временной корреляционной функции $r(\tau)$ при различных временных сдвигах τ . Оказывается, что при больших τ , порядка 12 ч и больше, для ряда метеорологических элементов (например, для скорости ветра, температуры воздуха для зимнего сезона) имеет место практически экспоненциальное убывание $r(\tau)$, что, вообще говоря, характерно для недифференцируемых процессов. Однако в действительности изменение большинства параметров, измеряемых в стандартных условиях на метеорологических станциях, представляет собой довольно гладкий процесс, который должен иметь по крайней мере одну производную. Это практически выражается в том, что при малых τ корреляционная функция должна иметь вид квадратичной параболы. Действительно, тщательный анализ корреляционных функций, полученных, например, по данным достаточно точных температурных измерений [3], выявляет при малых τ ($\tau < 6$ ч) характерную «шапочку».

Для таких элементов любая разумная аппроксимация временной корреляционной функции должна давать параболический ход корреляции при малых τ , экспоненциальное убывание ее при больших τ и, конечно, непрерывность ее вместе хотя бы с первой производной. Подбор соответствующих формул осложняется еще требованиями положительной определенности спектра для соответствующей аппроксимации и достаточной экономичности алгоритма статистического моделирования рядов, соответствующих выбранной аппроксимации.

Оказывается, что этим требованиям удовлетворяет вероятностная модель процесса, получающегося путем скользящего осреднения простейшего недифференцируемого процесса. Ниже приводятся формулы, получающиеся при описании временной статистической структуры этой моделью, предлагается методика оценки параметров этих формул и рассматривается пример использования этой методики.

2. Пусть мы имеем процесс $f(t)$, который полагаем стационарным с дисперсией σ^2 и ковариационной функцией $R_f(t_1, t_2) =$

$=R_f(\tau)$, где $\tau=|t_1-t_2|$ — сдвиг во времени. Рассмотрим скользящую среднюю с периодом осреднения T

$$F(t) = \frac{1}{T} \int_t^{t+T} f(t) dt. \quad (1)$$

Ковариационная функция этой средней равна

$$R_F(\tau) = \frac{1}{T^2} \int_{t_1=0}^T \int_{t_2=\tau}^{\tau+T} R_f(|t_1-t_2|) dt_1 dt_2 \quad (2)$$

или

$$R_F(\tau) = \begin{cases} \frac{1}{T^2} \left\{ \int_0^{\tau} (T-\tau+t) R_f(t) dt + \int_0^{T-\tau} (T-\tau-t) R_f(t) dt + \right. \\ \left. + \int_{\tau}^{\tau+T} (T+\tau-t) R_f(t) dt \right\} & \text{при } \tau \leq T, \\ \frac{1}{T^2} \left\{ \int_{\tau-T}^{\tau} (T-\tau+t) R_f(t) dt + \int_{\tau}^{\tau+T} (T+\tau-t) R_f(t) dt \right\} & \text{при } \tau \geq T. \end{cases} \quad (2a)$$

В случае, когда $R_f(\tau)$ имеет вид экспоненциальной зависимости

$$R_f(\tau) = \exp\left(-\frac{\tau}{T_0}\right), \quad (3)$$

где T_0 — масштаб корреляции, формула (2a) приводится к виду

$$R_F(\tau) = \begin{cases} \left(\frac{T_0}{T}\right)^2 \left\{ 2 \frac{T-\tau}{T_0} - 2 \exp\left(-\frac{\tau}{T_0}\right) + \exp\left(-\frac{T}{T_0}\right) \times \right. \\ \left. \times \left[\exp\left(\frac{\tau}{T_0}\right) + \exp\left(-\frac{\tau}{T_0}\right) \right] \right\} & \text{при } \tau \leq T, \\ \left(\frac{T_0}{T}\right)^2 \left[\exp\left(\frac{T}{T_0}\right) + \exp\left(-\frac{T}{T_0}\right) - 2 \right] \exp\left(-\frac{\tau}{T_0}\right) & \text{при } \tau \geq T. \end{cases} \quad (4)$$

Из формулы (4) следует непрерывность R_F вместе с ее первой производной на всем интервале задания ее, в том числе и при $\tau=T$. При $\tau>T$ имеет место экспоненциальное убывание $R_F(\tau)$, а при малых τ она имеет «шапочку», первая производная ее при $\tau=0$ равна нулю, а вторая производная отрицательна, что обеспечивает наличие первой производной величины F . С другой стороны, $R_F(\tau)$ описывает физически легко реализуемый процесс и имеет неотрицательный спектр.

Таким образом, формула (4) удовлетворяет перечисленным выше требованиям. Для практического использования ее необходимо уметь определять по фактическим данным о временной корреляции исследуемых рядов входящие в (4) параметры — масштаб корреляции T_0 и период осреднения T .

При наличии достаточно детальных и надежных экспериментальных данных T могло бы определяться просто как значение τ , соответствующее точке перегиба корреляционной функции. Практически, однако, на это редко можно рассчитывать. Может оказаться даже, что T меньше минимального τ , для которого мы располагаем данными. В этих условиях надежную оценку параметров модели можно получить путем простой экстраполяции данных для достаточно больших сдвигов времени ($\tau > T$).

В самом деле, для таких сдвигов логарифм корреляционной функции линейно зависит от τ и масштаб корреляции T_0 может быть определен по значениям корреляционной функции для двух значений сдвига

$$T_0 = \frac{\tau_2 - \tau_1}{\ln(r_1) - \ln(r_2)}. \quad (5)$$

Практически удобно экстраполяцию производить графически на полулогарифмической бумаге. В этом случае входящие в формулу (5) и последующую формулу (8) значения r снимаются с прямой, аппроксимирующей корреляционную функцию при больших τ . Это позволяет уменьшить погрешности расчетов, связанные с выборочными ошибками индивидуальных значений r_1 и r_2 .

Для оценки параметра T учтем, что дисперсия величины, согласно (4), определяется формулой

$$\sigma_F^2 = R_F(0) = 2 \left(\frac{T_0}{T} \right)^2 \left[\left(\frac{T}{T_0} \right) + \exp\left(\frac{T}{T_0} \right) - 1 \right] \sigma^2, \quad (6)$$

а корреляционная функция имеет вид

$$r_F(\tau) = \begin{cases} \frac{2\beta - 2\frac{\tau}{T_0} - 2\exp\left(-\frac{\tau}{T_0}\right) + \exp(-\beta) \left[\exp\left(\frac{\tau}{T_0}\right) + \exp\left(-\frac{\tau}{T_0}\right) \right]}{2[\beta + \exp(-\beta) - 1]} & \text{при } \tau \leq T, \\ \frac{\exp(\beta) + \exp(-\beta) - 2}{2[\beta + \exp(-\beta) - 1]} \exp\left(-\frac{\tau}{T_0}\right) & \text{при } \tau \geq T, \end{cases} \quad (7)$$

где $\beta = T/T_0$ — относительный масштаб осреднения. Экстраполяция значений корреляционной функции с экспоненциального участка $\tau > T$ до $\tau = 0$ дает фиктивное значение r_0

$$\ln(r_0) = \frac{\tau_2}{\tau_2 - \tau_1} \ln(r_1) - \frac{\tau_1}{\tau_2 - \tau_1} \ln(r_2), \quad (8)$$

которое связано с β формулой

$$r_0 = \frac{\exp(\beta) + \exp(-\beta) - 2}{2[\beta + \exp(-\beta) - 1]}. \quad (9)$$

Зависимость (9) нетрудно затабулировать, что дает возможность по экстраполированным значениям r_0 находить β , а следовательно, и $T = \beta T_0$. Учитывая, что для исследуемых рядов $T \ll T_0$ и, следовательно, $\beta \ll 1$, удобно использовать разложение (9) в степенной ряд

$$r_0 = 1 + \frac{1}{3} \beta + \frac{1}{9} \beta^2 + \dots \quad (10)$$

Ограничиваясь квадратическими членами разложения и решая квадратное уравнение (10) относительно β , получаем

$$\beta \approx 3(r_0 - 1) - 3(r_0 - 1)^2 \approx 3 \ln r_0 (1 - \ln r_0), \quad (11)$$

$$T \approx 3 \ln r_0 (1 - \ln r_0). \quad (12)$$

Строго говоря, наша вероятностная модель должна была бы быть дополнена сведениями о функции распределения моделируемого элемента. Известно (см., например, [2, 3]), что нормальный закон лишь приближенно описывает распределение даже температуры воздуха, однако для многих задач использование его дает удовлетворительные результаты. На данном этапе мы им и ограничимся. В этом случае вероятностная модель полностью определяется моментами второго порядка, схема определения которых нами описана.

Статистическое моделирование на ЭВМ величин $F(t)$, соответствующих принятой модели (будем впредь ее называть «моделью F »), производится в такой последовательности:

а) моделируются бессвязные нормальные случайные величины φ_i с математическим ожиданием, равным нулю, и дисперсией, равной единице. Стандартные программы, обеспечивающие расчет таких псевдослучайных чисел, предусмотрены в системах математического обеспечения всех современных ЭВМ;

б) бессвязная последовательность φ_i преобразуется в связную последовательность f_i с математическим ожиданием, равным нулю, дисперсией, равной единице, и корреляционной функцией (3). Для этой цели используется известная рекуррентная формула

$$f_i = \begin{cases} \varphi_0 & \text{при } i = 0, \\ q f_{i-1} + \sqrt{1 - q^2} \varphi_i & \text{при } i > 0. \end{cases} \quad (13)$$

Здесь $q = \exp\left(-\frac{h}{T_0}\right)$ представляет собой коэффициент корреляции между смежными членами последовательности f_i , соответствующий интервалу времени h между ними;

в) последовательность f_i преобразуется в последовательность F_i с математическим ожиданием $\bar{F}$ и дисперсией σ^2 . При этом интегрирование производится численно по какой-либо из квадратурных формул. Как показано в [1], наибольшую точность интегрирования случайных функций такой структуры обеспечивают квад-

ратурные формулы типа трапеций. Использование такой формулы дает

$$F_i = \bar{F} + \frac{\tilde{\sigma}}{T} \sum_{j=0}^{m-1} \tilde{f}_{i+j} p_j, \quad (14)$$

где

$$\tilde{\sigma} = \frac{\sigma \beta}{\sqrt{2[\beta + \exp(-\beta) - 1]}},$$

$m = 1 + \text{entier}\left(\frac{T}{h} + \frac{1}{2}\right)$ — число ординат, используемых при осреднении; p_j — весовые множители, определяемые формулой

$$p_j = \begin{cases} \frac{h}{2} & \text{при } j=0; \\ h & \text{при } j=1, 2, \dots, m-2; \\ T - \left(m - \frac{3}{2}\right)h & \text{при } j=m-1; \end{cases} \quad (15)$$

г) к значениям F_i добавляются псевдослучайные гауссовские числа ξ_i с математическим ожиданием, равным нулю, и дисперсией, равной единице, имитирующие ошибки наблюдений. Тем самым можно моделировать не только «истинные» температурные ряды, но и ряды наблюдаемых величин

$$\tilde{F}_i = F_i + \eta \sigma \xi_i, \quad (16)$$

где η^2 — мера ошибок наблюдений, равная отношению их дисперсии к дисперсии истинных значений измеряемой величины.

3. Как уже указывалось, модель F обеспечивает наличие первой производной у моделируемого процесса, что вполне достаточно для ряда задач, например для расчета выбросов процесса. Однако в некоторых случаях желательно обеспечивать наличие второй производной. Так обстоит дело, например, при необходимости расчета характеристик экстремумов. В таких случаях желательно использовать повторное осреднение.

Рассмотрим наряду с $\tilde{f}(t)$ и $F(t)$ процесс

$$\Phi(t) = \frac{1}{T} \int_t^{t+T} F(t) dt. \quad (17)$$

В результате несложных, хотя и громоздких выкладок можно получить формулу для корреляционной функции этого процесса

$$r_{\Phi}(\tau) = \begin{cases} \frac{1}{A} \left\{ \frac{4}{3} \beta^3 + \tilde{\tau}^3 - 2\beta \tilde{\tau}^2 + 6\tilde{\tau} - 4\beta + 6 \exp(-\tilde{\tau}) - \right. \\ \quad \left. - 4 \exp(-\beta) [\exp(\tilde{\tau}) + \exp(-\tilde{\tau})] + \exp(-2\beta) \times \right. \\ \quad \left. \times [\exp(\tilde{\tau}) + \exp(-\tilde{\tau})] \right\} \quad \text{при } \tilde{\tau} \leq T, \\ \frac{1}{A} \left\{ \frac{8}{3} \beta^3 - \frac{1}{3} \tilde{\tau}^3 - 4\beta^2 \tilde{\tau} + 2\beta \tilde{\tau}^2 + 4\beta - 2\tilde{\tau} + \right. \\ \quad \left. + 6 \exp(-\tilde{\tau}) - 4 \exp(-\tilde{\tau}) [\exp(\beta) + \exp(-\beta)] + \right. \\ \quad \left. + \exp(-2\beta) [\exp(\tilde{\tau}) + \exp(-\tilde{\tau})] \right\} \quad \text{при } T \leq \tilde{\tau} \leq 2T, \\ r_0 \exp(-\tilde{\tau}) \quad \text{при } \tilde{\tau} \geq 2T. \end{cases} \quad (18)$$

Здесь обозначено

$$\tilde{\tau} = \frac{\tau}{T_0};$$

$$A = \frac{4}{3} \beta^3 - 4\beta + 6 - 8 \exp(-\beta) + 2 \exp(-2\beta);$$

$$r_0 = \frac{\left[\exp\left(\frac{\beta}{2}\right) - \exp\left(-\frac{\beta}{2}\right) \right]^4}{A}. \quad (19)$$

Величина $r_{\Phi}(\tau)$, в соответствии с формулами (18), непрерывна вместе с первыми тремя производными на всем интервале изменения τ . Она полностью определяется двумя параметрами (T_0 и T), которые, аналогично предыдущему случаю, могут быть найдены по данным для достаточно больших значений τ ($\tau > 2T_0$). В этом случае для определения масштаба корреляции можно воспользоваться той же формулой (5), а масштаб осреднения определяется по экстраполированному до $\tau=0$ по формуле (8) значению r_0 . Для этой цели удобно вместо точной формулы (19) воспользоваться разложением вида

$$r_0 = 1 + \frac{7}{15} \beta + \frac{49}{225} \beta^2 + \dots, \quad (20)$$

откуда следует приближенная формула

$$\beta \approx \frac{15}{7} [(r_0 - 1) - (r_0 - 1)^2] \approx \frac{15}{7} (\ln r_0 - \ln^2 r_0), \quad (21)$$

и соответственно

$$T \approx \frac{15}{7} T_0 (\ln r_0 - \ln^2 r_0). \quad (22)$$

Статистическое моделирование гауссовской последовательности Φ_i на ЭВМ по этой модели, которую будем называть «моделью Φ », может производиться аналогично схеме, изложенной выше для модели F . При этом формула (14) должна быть заменена формулой

$$F_i = \frac{1}{T} \sum_{j=0}^{m-1} f_{i+j} p_j, \quad (14a)$$

где весовые множители p_j по-прежнему определяются формулой (15), а величины Φ_i с математическим ожиданием $\bar{\Phi}$ и дисперсией σ^2 — по аналогичной (14) формуле

$$\Phi_i = \bar{\Phi} + \frac{\sigma}{T} \sum_{j=0}^{m-1} F_{i+j} p_j, \quad (23)$$

где

$$\sigma \approx \frac{\sigma \beta^2}{\sqrt{\frac{4}{3} \beta^3 - 4 \beta + 6 - 8 \exp(-\beta) + 2 \exp(-2\beta)}}.$$

Можно, однако, избежать повторного сглаживания и использовать и в этом случае сглаживание однократное, если учесть, что

$$\begin{aligned} \Phi(t) &= \bar{\Phi} + \frac{\sigma}{T^2} \int_{t_1=t}^{t+T} \int_{t_2=t_1}^{t_1+T} f(t_2) dt_1 dt_2 = \\ &= \bar{\Phi} + \frac{\sigma}{T^2} \left\{ \int_t^{t+T} (x-t) f(x) dx + \int_{t+T}^{t+2T} (t+2T-x) f(x) dx \right\}, \end{aligned} \quad (24)$$

или

$$\Phi(t) = \bar{\Phi} + \frac{\sigma}{T^2} \int_t^{t+2T} s(x) f(x) dx,$$

где

$$s(x) = \begin{cases} x-t & \text{при } t \leq x \leq t+T, \\ t+2T-x & \text{при } t+T \leq x \leq t+2T. \end{cases} \quad (25)$$

Тогда для моделирования двукратно дифференцируемого процесса Φ_i можно, минуя предварительный расчет F_i , производить сглаживание величин f_i по формуле

$$\Phi_i = \bar{\Phi} + \frac{\sigma}{T^2} \sum_{j=0}^{m-1} s_j f_{i+j} p_j, \quad (26)$$

где $m = 1 + \text{entier} \left(\frac{2T}{h} \right)$ (в отличие от (14), здесь при интегриро-

вании целесообразно использовать лишь моменты времени, попадающие внутрь интервала осреднения),

$$s_j = s|_{x=jh};$$

$$p_j = \begin{cases} \frac{h}{2} & \text{при } j=0; \\ h & \text{при } j=1, 2, \dots, m-2; \\ 2T - (m - \frac{3}{2})h & \text{при } j=m-1. \end{cases} \quad (27)$$

Аналогично (16), в ряды Φ_i могут быть также введены случайные числа, имитирующие ошибки наблюдений

$$\tilde{\Phi}_i = \Phi_i + \eta \sigma \xi_i. \quad (28)$$

4. Рассмотрим пример использования моделей F и Φ для аппроксимации данных о временной структуре температуры воздуха на ст. Ленинград, ГМО за январь, приведенных в [3]. В этой работе представлены данные о временной корреляции для сдвигов времени τ , равных 1, 2, 3, 6, 12 и 24 ч. Из рис. 1, на котором эти данные представлены в полулогарифмических координатах, видно, что экспоненциальное убывание корреляционной функции наблюдается практически уже при $\tau > 6$ ч. Точки, соответствующие

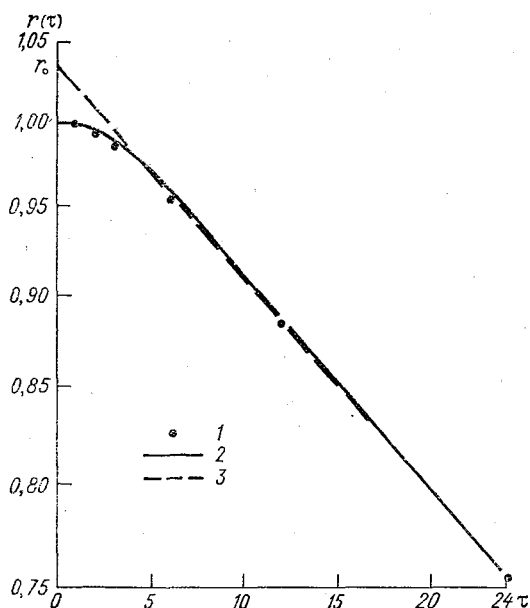


Рис. 1. Временная корреляционная функция температуры воздуха. Ленинград, ГМО, январь.
1 — по фактическим данным, 2 — расчетная по моделям F и Φ , 3 — аппроксимирующая прямая.

щие $\tau_1=12$ ч и $\tau_2=24$ ч, могут быть использованы для оценки параметров корреляционной функции. При такой оценке значение масштаба корреляции получаем равным $T_0=76$ ч. Экстраполяция до значений $\tau=0$ дает оценку $\ln r_0=0,034$. Учитывая занижение эмпирической корреляционной функции за счет случайных ошибок наблюдений, оценка меры которых дает значение $\eta^2=0,001$, получаем окончательно $\ln r_0=0,035$, или $r_0=1,035$. Поль-

Таблица 1

**Временные корреляционные функции, рассчитанные по фактическим и моделированным данным и по формулам (7) и (18).
Ленинград, ГМО, январь**

τ	Фактические данные	По формуле (7)	По формуле (18)	Моделированные			
				$\bar{r}_F$	σ_{r_F}	$\bar{r}_\Phi$	σ_{r_Φ}
1	0,997	0,9984	0,9985	0,995	0,001	0,996	0,000
2	0,991	0,9939	0,9940	0,989	0,001	0,992	0,001
3	0,983	0,9870	0,9869	0,981	0,002	0,984	0,001
4	—	0,9781	0,9778	0,971	0,003	0,975	0,002
5	—	0,9677	0,9672	0,959	0,004	0,965	0,003
6	0,953	0,9562	0,9558	0,947	0,005	0,953	0,004
7	—	0,9441	0,9438	0,934	0,006	0,941	0,005
8	—	0,9317	0,9316	0,922	0,007	0,929	0,006
9	—	0,9196	0,9195	0,910	0,008	0,917	0,007
10	—	0,9075	0,9075	0,898	0,009	0,905	0,008
12	0,883	0,8840	0,8839	0,874	0,011	0,881	0,010
14	—	0,8610	0,8610	0,851	0,013	0,858	0,012
16	—	0,8387	0,8386	0,829	0,015	0,836	0,014
18	—	0,8169	0,8168	0,808	0,017	0,814	0,016
20	—	0,7957	0,7956	0,787	0,019	0,793	0,017
22	—	0,7750	0,7750	0,766	0,021	0,772	0,019
24	0,755	0,7549	0,7548	0,746	0,022	0,752	0,020

зуюсь формулами (11) и (12) для случая однократного сглаживания, получаем оценку $\beta=0,102$; $T=7,75$ ч. Для случая двойного сглаживания подстановка в формулы (21) и (22) дает $\beta=0,0728$; $T=5,53$ ч. Подстановка в формулы (7) и (18) полученных параметров позволяет получить корреляционные функции, соответствующие простому и двойному сглаживанию. Эти корреляционные функции наряду с полученными в [3] фактическими корреляционными функциями представлены в табл. 1.

Из таблицы видно, что корреляционные функции для модели F и модели Φ различаются лишь в четвертом знаке после запятой. При сравнении их с фактическими данными следует иметь в виду необходимость введения поправочного множителя 0,999, соответствующего принятой мере ошибок $\eta^2=0,001$, после чего расчетные корреляционные функции хорошо согласуются с фактическими данными. Расхождения между расчетными и эмпирическими корреляционными функциями не превосходят возможных выборочных погрешностей за счет ограниченности ряда, использованного при расчетах (фактически в работе [3] использовались наблюдения за 22 года). Поскольку модель двойного сглаживания может использоваться в более широком круге приложений, применение ее, по-видимому, предпочтительнее. Реализация ее на ЭВМ осуществляется по формулам (13), (16)—(28) при задании параметров $\eta^2=0,001$; $q=\exp(-1/76)=0,9869$.

Насколько хорошо этот дискретный алгоритм позволяет имитировать непрерывное сглаживание, можно судить по представленным в табл. 1 результатам обработки полученных статистическим моделированием по этому алгоритму последовательностей, имитирующих ежеhourные ряды температуры воздуха за январь для Ленинграда. Всего обработке подвергались 50 выборок, каждая из которых состояла из 22 независимых реализаций (744 члена), что соответствует объему фактических данных, использованных в [3]. Расчет в этом случае также производился как для модели F , так и для модели Φ .

В табл. 1 приводятся осредненные по всем выборкам значения статистических характеристик $\bar{r}$ и средние квадратические отклонения σ_r , характеризующие погрешности их за счет ограниченности объема выборки. Из таблицы видно, что различия между временными корреляционными функциями, соответствующими дискретным алгоритмам F и Φ , несколько больше, чем они были бы в идеальном случае непрерывного сглаживания. Однако расхождения между ними, а также между корреляционными функциями для моделированных и фактических выборок находятся в пределах погрешности последней, которая вполне может характеризоваться величинами σ_{r_F} или σ_{r_Φ} .

Таким образом, предложенные алгоритмы позволяют моделировать на ЭВМ ряды, достаточно реалистично воспроизводящие основные особенности статистической структуры температуры воздуха для зимнего сезона. Эти алгоритмы уже используются в настоящее время для решения некоторых вопросов методики обработки этого элемента. Есть основания надеяться, что они окажутся пригодными и для моделирования других метеорологических элементов. Представляется целесообразным в дальнейшем рассмотреть возможность обобщения их на случай учета отклонений распределения от нормального, а также на случай моделирования нестационарных рядов, в первую очередь рядов периодически коррелированных.

1. Каган Р. Л. Осреднение метеорологических полей. — Л.: Гидрометеоиздат, 1979. — 213 с.
2. Тихомирова Л. В. Некоторые особенности временных рядов распределения температуры воздуха на территории СССР. — В кн.: Применение статистических методов в метеорологии: Труды Всесоюзного симпозиума по применению статистических методов в метеорологии. Л., Гидрометеоиздат, 1971, с. 290—304.
3. Федорченко Е. И. О временной структуре ежечасных наблюдений за температурой воздуха в Ленинграде. — Труды ГГО, 1975, вып. 348, с. 112—122.

Ю. В. Житорчук, Л. Е. Страшненко

ОПТИМИЗАЦИЯ СТРУКТУРЫ ПОСЕВНЫХ ПЛОЩАДЕЙ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР С УЧЕТОМ КЛИМАТОЛОГИЧЕСКОЙ ИНФОРМАЦИИ

В настоящее время все более возрастает необходимость учета метеоинформации при выработке оптимальных стратегий в сельском хозяйстве.

Рассмотрим задачу принятия экономически оптимального решения на основе климатологической информации на примере выбора оптимального соотношения посевных площадей двух сортов яровой пшеницы.

В большинстве исследований, посвященных вопросу оптимизации хозяйственных стратегий потребителя, используется принцип минимизации средних потерь [1]. Однако такой подход к решению хозяйственных задач встречает критическое отношение со стороны ряда экономистов и метеорологов [2, 3]. Как показано в этих работах, принцип минимизации средних потерь имеет определенную ограниченность и не учитывает возможность крупных потерь, которые могут привести к необратимым последствиям. Так, например, сильная засуха, приводящая к существенному неурожаю зерна, сказывается на уменьшении поголовья скота, для воспроизводства которого даже в случае благоприятных погодных условий требуется несколько лет. Поэтому более целесообразно с общеэкономической точки зрения использовать принцип минимизации повторяемости крупных потерь [2]. Этим принципом мы и будем руководствоваться при решении задачи нахождения оптимального соотношения посевных площадей двух сортов яровой пшеницы.

Пусть урожайности этих сортов характеризуются средними значениями $\bar{y}_1$ и $\bar{y}_2$, средние квадратические отклонения от которых соответственно равны σ_1 и σ_2 , а r — коэффициент корреляции между урожайностями первого и второго сортов рассматриваемой культуры. Если принять общую посевную площадь за 1 и считать,

что на долю первого сорта приходится α засеваемой площади, то на долю второго сорта будет приходиться $1-\alpha$ засеваемой площади и общая урожайность будет складываться из их суммы:

$$z = z_1 + z_2 = \alpha y_1 + (1 - \alpha)y_2.$$

Тогда задача выработки оптимальной хозяйственной стратегии сведется к нахождению такого соотношения засеваемых площадей, при котором общая урожайность z не была бы ниже величины c , т. е. такого значения α , при котором достигался бы минимум повторяемости крупных потерь $\min_{0 \leq \alpha \leq 1} P\{z < c\}$,

где c — уровень урожайности, ниже которого потери считаются крупными. Предполагая, что с достаточной степенью точности плотность распределения средних областных урожайностей сельскохозяйственных культур может быть описана при помощи нормального закона распределения, получим следующее выражение для расчета вероятности крупных потерь:

$$P\{z < c\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{c - \bar{z}}{\sigma_z}} e^{-\frac{u^2}{2}} du, \quad (1)$$

где

$$\frac{c - \bar{z}}{\sigma_z} = \frac{c - (\alpha \bar{y}_1 + (1 - \alpha)\bar{y}_2)}{\sqrt{\alpha^2 \sigma_1^2 + (1 - \alpha)^2 \sigma_2^2 + 2r \alpha(1 - \alpha)\sigma_1 \sigma_2}}. \quad (2)$$

Необходимым условием, при котором интеграл (1) принимает минимальное значение, является равенство нулю первой производной по α верхнего предела интегрирования:

$$\frac{\partial}{\partial \alpha} \left(\frac{c - (\alpha \bar{y}_1 + (1 - \alpha)\bar{y}_2)}{\sqrt{\alpha^2 \sigma_1^2 + (1 - \alpha)^2 \sigma_2^2 + 2r \alpha(1 - \alpha)\sigma_1 \sigma_2}} \right) = 0. \quad (3)$$

Производя очевидные преобразования, получаем для расчета оптимальной доли первого сорта в общей посевной площади следующее выражение:

$$\alpha_{\text{опт}} = \frac{r \sigma_1 \sigma_2 (c - \bar{y}_2) - \sigma_2^2 (c - \bar{y}_1)}{(\sigma_1^2 + r \sigma_1 \sigma_2) (\bar{y}_1 - \bar{y}_2) - (c - \bar{y}_1) (\sigma_1^2 + \sigma_2^2 - 2r \sigma_1 \sigma_2)}. \quad (4)$$

Удобнее перейти к безразмерным параметрам, вводя соответствующие обозначения:

$$l = \frac{\sigma_1}{\sigma_2}, \quad m = -\frac{c}{\bar{y}_2}, \quad n = \frac{\bar{y}_1}{\bar{y}_2}. \quad (5)$$

Тогда окончательное выражение для $\alpha_{\text{опт}}$ будет иметь вид

$$\alpha_{\text{опт}} = \frac{rl(m - 1) - (m - n)}{(1 - r)(n - 1) - (m - 1)(l^2 - 2rl + 1)}. \quad (6)$$

Урожайность сельскохозяйственных культур определяется как агротехническими, так и погодными факторами. Вследствие совершенствования земледелия она постепенно растет, в то время как резкое год от года изменение погодных условий приводит к сильному ее варьированию. Поэтому временной ряд урожайности можно разделить на низко- и высокочастотные составляющие. Низкочастотную составляющую урожайности, отражающую влияние на урожайность агротехники (так называемый тренд), в первом, очень грубом приближении можно аппроксимировать параболой второго порядка:

$$\hat{y} = a_1 t^2 + a_2 t + a_3,$$

где $\hat{y}$ — условное математическое ожидание урожайности как функция времени, a_1, a_2, a_3 — коэффициенты, которые находятся по методу наименьших квадратов. Тогда отклонения от условного математического ожидания урожайности в основном будут обусловлены влиянием погодных условий. Их статистические характеристики: математические ожидания, дисперсии и коэффициенты корреляции могут быть использованы при определении оптимального

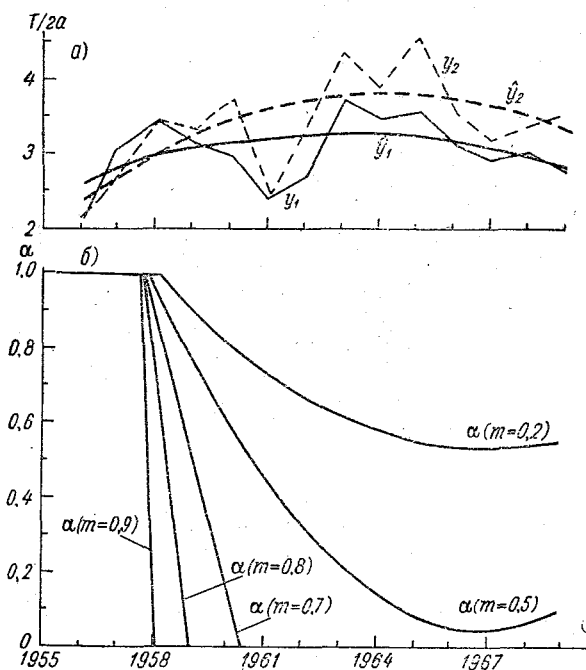


Рис. 1. Временная зависимость урожайностей y_1, y_2 , математических ожиданий урожайностей $\hat{y}_1, \hat{y}_2$ (а) и оптимального соотношения посевных площадей α сортов яровой пшеницы «Гражучай» и «Гарсас». (б) ($\sigma_1=3,81, \sigma_2=4,71, r=0,84$).

соотношения посевных площадей различных культур с учетом климатологической информации.

Были проведены статистические эксперименты по расчету оптимальных соотношений различных пар сортов яровой пшеницы с учетом климатологической информации на основе данных об урожайности за 16 лет по 13 участкам сортоиспытания некоторых областей с помощью предложенного алгоритма.

На рис. 1, 2 представлены некоторые примеры. Приведены временные зависимости урожайностей, их трендов и соответствующие оптимальные соотношения посевных площадей рассматриваемых сортов. Кривые, представленные на рис. 1 а, показывают изменение во времени урожайностей первого сорта яровой пшеницы «Гражучай» и второго сорта «Гарсаса». (Данные об урожайности этих сортов получены с Укмергского сортоиспытательного участка Литовской ССР.)

Как видно, до 1958 г. урожайность сорта «Гражучай» была в среднем выше, чем урожайность сорта «Гарсас». Однако с уве-

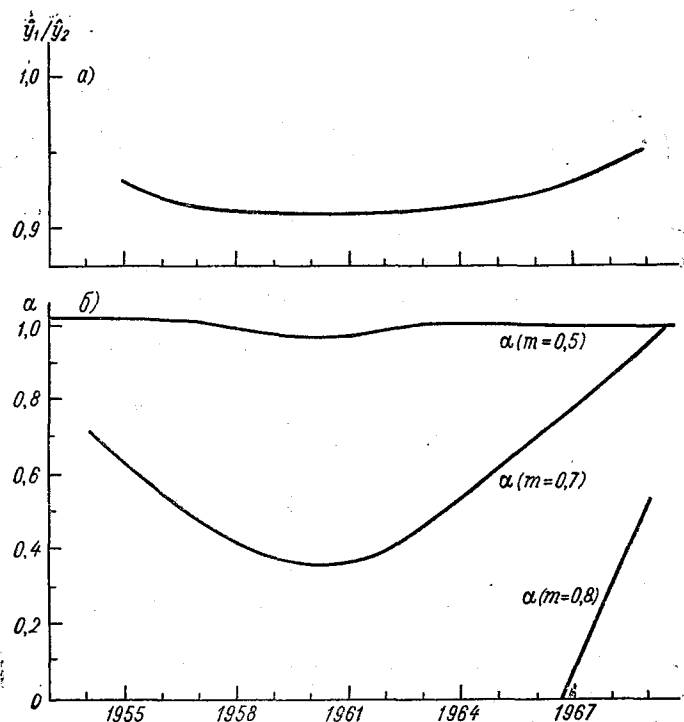


Рис. 2. Временная зависимость отношения средних урожайностей $\bar{y}_1/\bar{y}_2$ (а) и оптимального соотношения посевных площадей α сортов яровой пшеницы «Диамант» и «Минская» (б) ($\sigma_1=3,41$, $\sigma_2=4,78$, $r=0,88$).

личением уровня агротехники соотношение между урожайностями этих сортов изменилось на обратное. Соответственно изменилось и оптимальное соотношение посевных площадей этих сортов. Как следует из формулы (6), величина $\alpha_{\text{опт}}$ зависит не только от статистических характеристик рассматриваемых сельскохозяйственных культур, но и от экономических возможностей потреби-

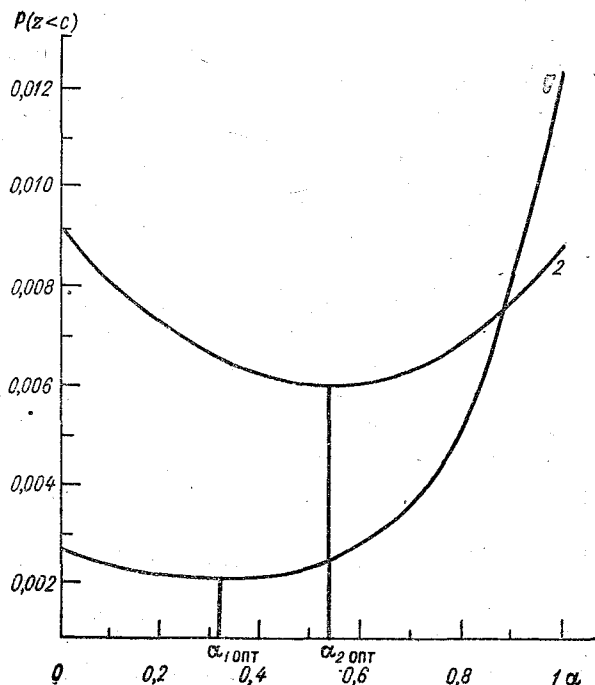


Рис. 3. Зависимость повторяемости крупных потерь от соотношения посевных площадей сортов.

1 — «Диамант» ($y_1=23,9$, $\hat{y}_1=26,97$, $\sigma_1=2,67$) и «Заря» ($y_2=25$, $\hat{y}_2=29,87$, $\sigma_2=3,06$), $r=0,53$, Ленинградская область, Гатчинский сортоиспытательный участок; 2 — «Лютесценс-758» ($y_1=28,3$, $\hat{y}_1=23,33$, $\sigma_1=4,6$) и «Харьковская-46» ($y_2=30,1$, $\hat{y}_2=24,38$, $\sigma_2=5,06$), $r=0,75$, Татарская АССР, сортоиспытательный участок.

теля, а именно от того, какую величину потерь потребитель считает крупными потерями.

На рис. 1 б дано несколько кривых, описывающих оптимальную стратегию потребителей, ориентирующихся на различные уровни крупных потерь.

Если потребитель всю площадь засеивает первым (вторым) сортом, то будем говорить, что он использует «чистую стратегию $\alpha=1$ ($\alpha=0$)», если сорта берутся в некоторых соотношениях, то «смешанную стратегию».

Для потребителей, ориентирующихся на относительно низкий уровень крупных потерь ($m=0,7 \dots 0,8$, т. е. потери считаются недопустимыми, если они превышают 20—30% всего урожая), чистая стратегия $\alpha=1$ на первом этапе довольно быстро переходит в чистую стратегию $\alpha=0$ на втором этапе. Для потребителей, ориентирующихся на уровень более крупных потерь ($m \leq 0,5$), чистая стратегия $\alpha=1$ на первом этапе переходит в смешанную на втором. Для потребителей, ориентирующихся на очень высокий уровень крупных потерь ($m=0,2$), оптимальной является такая смешанная стратегия, в которой доля сорта «Гражучай» становится преобладающей, несмотря на то, что урожайность этого сорта ниже. Это обусловлено тем, что урожайность «Гражучая» имеет меньшую дисперсию. Следовательно, при относительно низком уровне крупных потерь на оптимальную стратегию потребителя определяющее влияние оказывает соотношение урожайностей рассматриваемых культур. Чем выше этот уровень, тем больше оптимальная стратегия потребителя определяется соотношением дисперсий урожайностей.

Рассмотрим еще один пример. На рис. 2 а представлен график временной зависимости отношения урожайности двух сортов яровой пшеницы — «Диаманта» и «Минской» (Калининская область, Бежецкий сортоиспытательный участок). Средняя урожайность «Диаманта» ниже средней урожайности «Минской», дисперсия урожайности первого сорта меньше дисперсии урожайности второго сорта. В этом случае для потребителя, ориентирующегося на высокий уровень крупных потерь ($m=0,5$), оптимальной является либо частая стратегия $\alpha=1$, либо смешанная с явным преобладанием доли первого сорта «Диаманта». Но для потребителя, ориентирующегося на низкий уровень крупных потерь ($m=0,8$), кривая оптимального поведения имеет иной вид. Отношение урожайностей в этом случае является определяющим и предпочтение отдается «Минской», с более высокой средней урожайностью. Расчеты показали, что применение оптимальной стратегии, а не наилучшей чистой стратегии уменьшает вероятность крупных потерь на 20—30 % (рис. 3).

СПИСОК ЛИТЕРАТУРЫ

1. Жуковский Е. Е. Принятие оптимальных решений как метод повышения эффективности агрометеорологических прогнозов. — Метеорология и гидрология, 1973, № 1, с. 32—42.
2. Жуковский Е. Е., Чудновский А. Ф. Методы оптимального использования метеорологической информации при принятии решений. — Л.: Гидрометеиздат, 1978. — 52 с.
3. Можин В. Н. Оптимизация плановых решений в сельском хозяйстве. — М.: Экономика, 1974. — 150 с.
4. Юдин М. И. Долгосрочный прогноз погоды как средство управления хозяйственной деятельностью. — В кн.: Применение статистических методов в метеорологии: Труды II Всесоюзного симпозиума по применению статистических методов в метеорологии. Л., Гидрометеиздат, 1977, с. 5—10.

УЧЕТ АПРИОРНОЙ ИНФОРМАЦИИ В РЕГРЕССИОННЫХ МОДЕЛЯХ ПРОГНОЗА ПОГОДЫ

При построении статистических моделей прогноза, как правило, используются методы регрессионного анализа, распознавания образов, экспоненциального сглаживания, формальной экстраполяции и т. д. Достоинством этих методов является то, что они не требуют никаких априорных сведений о степени предсказуемости прогнозируемых процессов. Однако в ряде случаев такие сведения имеются или могут быть получены из имеющейся метеорологической информации. При этом перечисленные методы уже, как правило, не обладают свойством оптимальности. В этой ситуации для прогнозирования естественно воспользоваться методами теории несмещенных оценок [1], позволяющими учесть априорную информацию и получить оценки с минимальной дисперсией, у которых отсутствует систематическая ошибка. Статистическая теория несмещенных оценок, разработанная в работах А. Н. Колмогорова и С. Рао, нашла широкое применение в задачах контроля качества продукции. Однако при решении задач статистического прогноза случайных процессов эта теория применяется в гораздо меньшей степени.

Рассмотрим, как можно использовать априорную информацию о распределении информативности совокупности тест-предикторов для получения несмещенных оценок на примере простейшего уравнения регрессии, содержащего один предиктор X_t :

$$\hat{Y}_{t+\tau} = \alpha X_t, \quad (1)$$

где $\hat{Y}_{t+\tau}$ — прогностическое значение предиктанта $Y_{t+\tau}$, α — коэффициент регрессии.

Далее, без уменьшения общности выводов, будем полагать, что значения предиктора и предиктанта предварительно были центрированы и нормированы, т. е. $\bar{Y} = \bar{X} = 0$; $\sigma_X = \sigma_Y = 1$. В этом случае из теории наименьших квадратов известно, что коэффициент регрессии численно равен выборочному коэффициенту корреляции ($\alpha = r$).

Определим среднюю квадратическую ошибку уравнения регрессии (1), рассчитанного по выборке объема N , в случае, когда это уравнение неограниченное количество раз используется на независимой выборке для фиксированных выборочных (r) и генеральных (R) значений коэффициента корреляции, т. е. для одного конкретного уравнения регрессии

$$\epsilon_y^2 = M\{(Y - \hat{Y})^2\} = M\{Y^2\} - 2\alpha M\{XY\} + \alpha^2 M\{X^2\} = 1 + 2rR + r^2. \quad (2)$$

В конечном итоге нас интересует величина ошибки, полученная

с учетом плотности распределения значений r и R , т. е. величина ошибки, осредненная по большому числу уравнений регрессии

$$\epsilon^2 = M\{\epsilon_0^2\} = 1 - 2 \operatorname{cov}\{r, R\} - 2M\{r\}M\{R\} + \sigma^2\{r\} + M^2\{r\}. \quad (3)$$

Для того чтобы воспользоваться выражением (3), необходимо знать двумерное совместное распределение плотности вероятностей выборочных и генеральных значений коэффициентов корреляции. В общем случае это весьма сложная задача, однако, учитывая, что при долгосрочных прогнозах мы имеем дело со слабопредсказуемыми процессами, для которых $|R| \ll 1$, можно в качестве первого приближения использовать следующее приближенное равенство:

$$r \approx z = 0,5 \ln \frac{1+r}{1-r}. \quad (4)$$

Тогда плотность вероятности коэффициентов корреляции $g(r)$ приближенно может быть выражена через известную плотность вероятности $g(z)$ переменной z : $g(r) \approx g(z)$.

С учетом этого условия плотность вероятности выборочного коэффициента корреляции при фиксированном значении коэффициента в генеральной совокупности R^* имеет следующий вид:

$$g(r|R=R^*) \approx \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(r-R^*)^2}{2\sigma_0^2}}, \quad (5)$$

где

$$\sigma_0 = \sigma\{r|R=R^*\} = \frac{1}{\sqrt{N-3}}.$$

Смещенность множественного коэффициента корреляции зависит от вида распределения генеральных коэффициентов корреляции между предиктантом и выбранной совокупностью предикторов. Однако, до сих пор эта сторона обсуждаемой проблемы в литературе, по-видимому, никем не рассматривалась.

В этой связи нами исследовался вопрос о точности восстановления распределения выборочных коэффициентов корреляции (см. в качестве примера выражение (8)) при различных предположениях относительно вида функции, описывающей распределение генеральных коэффициентов корреляции. В результате было установлено, что лучше всего экспериментальным данным соответствует аппроксимация нормальным распределением

$$g(R) = \sqrt{\frac{\gamma}{\pi}} e^{-\gamma R^2}, \quad (6)$$

где γ — константа, характеризующая информативность выбранной совокупности тест-предикторов. Обозначив через $\beta = \gamma/\sigma_0$, запишем

двумерную плотность вероятности выборочных и генеральных коэффициентов корреляции в следующем виде:

$$g(R, r) = g(r|R) g(R) = \frac{V\beta}{\pi \sigma_0^2 \sqrt{2}} e^{-\left(\frac{1+2\beta}{2\sigma_0^2} R^2 - \frac{r}{\sigma_0^2} R + \frac{r^2}{2\sigma_0^2}\right)} \quad (7)$$

Здесь необходимо обратить внимание на то обстоятельство, что использование фишеровского приближения (4) возможно лишь при условии $\beta \gg 0$; $\gamma \gg 0$ и, следовательно, полученные выше формулы применимы для ограниченного числа задач, удовлетворяющих этому условию, в основном для задач долгосрочного и среднесрочного прогнозов погоды.

Определим условное математическое ожидание генерального коэффициента при фиксированном значении выбранного коэффициента корреляции

$$g(r) = \int_{-\infty}^{\infty} g(R, r) dR = \frac{V\beta}{V\pi \sigma_0 \sqrt{1+2\beta}} e^{-\frac{r^2 \beta}{\sigma_0^2(1+2\beta)}} \quad (8)$$

Из (8) следует, что дисперсия выборочных коэффициентов корреляции зависит от длины ряда N и показателя информативности совокупности тест-предикторов β :

$$\sigma^2\{r\} = \frac{\sigma_0^2(1+2\beta)}{2\beta} \quad (9)$$

Используя это свойство, можно сделать выборочную оценку величины β , вычислив статистику φ , представляющую собой отношение безусловной дисперсии выборочных коэффициентов к условной:

$$\varphi = \frac{\sigma^2\{r\}}{\sigma^2\{r|R=R^*\}} = \frac{N-3}{2n} \sum_{i=1}^n \left(\ln \frac{1+r_i}{1-r_i} \right)^2, \quad (10)$$

где n — число тест-предикторов в исследуемой совокупности. Из (10) следует

$$\beta = \frac{1}{2(\varphi-1)} \quad (11)$$

Статистические эксперименты, выполненные одним из авторов, показали, что тест-предикторы, используемые в долгосрочном физико-статистическом прогнозе [2], характеризуются при большой заблаговременности величиной $\varphi=1,2$. При нулевой заблаговременности для автокорреляционных связей индекс φ достигает величины $\varphi=4,0$. Заметим, что параметризация распределения плотности вероятности генеральных коэффициентов корреляции между предиктантом и предикторами (6) с учетом возможности

экспериментального определения величины параметра β открывает широкие перспективы для исследования случайных и систематических ошибок различных схем прогнозов погоды при помощи методов Монте-Карло.

Определим условное математическое ожидание генерального коэффициента корреляции при фиксированном значении выборочного коэффициента корреляции r^*

$$\bar{r} = M\{R|r=r^*\} = \frac{\int_{-\infty}^{\infty} Rg(R, r) dr}{g(r)} = \frac{r^*}{1+2\beta} = r^* \frac{\varphi-1}{\varphi}. \quad (12)$$

Из (12) видно, что оценки метода наименьших квадратов в приложении к рассматриваемой задаче являются смещенными. Однако этот вывод не противоречит известным доказательствам о несмещенности оценок методом наименьших квадратов, поскольку при их выводе в неявном виде используется предположение о том, что

$$g(R) = \text{const} \quad (-1 \leq R \leq 1). \quad (13)$$

Гипотеза (13) выполняется в том случае, если регрессионный анализ используется для решения большого числа разнообразных по своим свойствам задач, от чисто детерминированных до абсолютно случайных. Таким образом, метод наименьших квадратов является эффективным лишь в среднем. Если же мы имеем дело с определенным классом задач, при решении которого гипотеза (13) не соблюдается, то в этом случае в принципе может быть получено решение более эффективное, чем решение методом наименьших квадратов.

Оценим теперь на независимой выборке ожидаемую точность уравнений регрессии, рассчитанную по методу наименьших квадратов и методу несмещенных оценок. Для этого воспользуемся выражением (3). Запишем его в следующем виде:

$$g(r, R) = Ce^{-(a_{11}R^2 + a_{12}Rr - a_{22}r^2)}. \quad (14)$$

Тогда

$$\rho\{r, R\} = \frac{a_{12}}{\sqrt{a_{11}a_{22}}} = \frac{1}{\sqrt{1+2\beta}}, \quad (15)$$

$$\sigma^2\{R\} = \frac{a_{22}}{2(a_{11}a_{22} - a_{12}^2)} = \frac{\sigma_0^2}{2\beta}, \quad (16)$$

где $\rho\{r, R\}$ — коэффициент корреляции между выборочными и генеральными значениями коэффициентов корреляции. Точность уравнения регрессии в существенной мере зависит от используемого способа отбора предикторов.

Рассмотрим схему, при которой отбор предикторов осуществляется без учета информации о значении выборочного коэффициента корреляции. В этом случае

$$M\{R\} = M\{r\} = 0$$

и оценка точности прогностического уравнения (1) имеет вид

$$\sigma^2\{r\} = 1 + \frac{\sigma_0^2(2\beta - 1)}{2\beta} = 1 + \sigma^2(2 - \varphi). \quad (17)$$

Из (17) видно, что существует критическое значение индекса информативности предикторов $\varphi=2$, ниже которого точность уравнения регрессии, рассчитанного по методу наименьших квадратов, ниже точности климатологического прогноза. Это позволяет ввести количественное определение понятия слабопредсказуемых процессов.

Далее будем говорить, что процесс является слабопредсказуемым, если совокупность тест-предикторов, отобранных исходя из физических соображений, характеризуется значением $\varphi < 2$.

В этом случае, когда в качестве коэффициента регрессии будет использована несмещенная оценка (12),

$$\sigma^2\{\hat{r}\} = \frac{\sigma^2\{r\}}{(1 + 2\beta)^2} = \frac{\sigma_0^2}{2\beta(1 + 2\beta)}. \quad (18)$$

Отсюда следует

$$\varepsilon^2\{\hat{r}\} = 1 - \frac{\sigma_0^2}{2\beta(1 + 2\beta)} = 1 - \frac{\sigma^2(\varphi - 1)^2}{\varphi}. \quad (19)$$

Следовательно, коррекция уравнения регрессии (12) позволяет уменьшить относительную ошибку независимых прогнозов, значение которой в этом случае не может превосходить единицы. Однако такой вывод справедлив лишь в том случае, если значение φ известно абсолютно точно. В действительности же φ получается из выборочных данных и известно с некоторой погрешностью, в результате чего (19) несколько завышает ожидаемую точность уравнения регрессии. С практической точки зрения наибольший интерес представляют прогностические схемы, в которых на зависимой выборке отбираются «наиболее информативные» предикторы. В этом случае распределение каждого коэффициента корреляции, упорядоченного в порядке возрастания его выборочной информативности, существенно отличается от распределения всей совокупности.

Пусть на зависимой выборке были рассчитаны n коэффициентов корреляции между предиктантом и n предикторами. Расположим их в порядке возрастания:

$$r_{(1)} \leq r_{(2)} \leq \dots \leq r_{(i)} \leq \dots \leq r_{(n)}. \quad (20)$$

В таком упорядоченном ряду i -й по порядку член называется i -й порядковой статистикой. Распределение порядковых статистик

имеет весьма сложный вид, поэтому при решении многих практических задач, как правило, используются упрощенные асимптотические формулы [3], позволяющие оценить математическое ожидание и дисперсию порядковых статистик:

$$M\{r_{(i)}\} = \lambda_{(i)}^*, \quad (21)$$

$$\sigma^2\{r_{(i)}\} = \frac{i - 0,375}{n + 0,25} \left(1 - \frac{i - 0,375}{n + 0,25}\right) \frac{1}{(n + 1,5)(g\{\lambda_{(i)}^*\})^2}, \quad (22)$$

где $\lambda_{(i)}^*$ — квантиль исходного распределения

$$F(\lambda_{(i)}^*) = \int_{-\infty}^{\lambda_{(i)}^*} g(r) dr = \frac{i - 0,375}{n + 0,25}. \quad (23)$$

Обычно удобнее рассматривать нормированные величины, поэтому далее через $\lambda_{(i)}$ обозначим квантиль нормированного нормального распределения $N(0, 1)$, тогда

$$M\{r_{(i)}\} = \lambda_{(i)} \sigma\{r\}. \quad (24)$$

В отличие от схемы случайного отбора предикторов, в рассматриваемом варианте математическое ожидание выборочных порядковых статистик и соответствующее условное математическое ожидание коэффициента корреляции генеральной совокупности не равны нулю, поэтому относительная ошибка уравнения регрессии с учетом (3) имеет вид

$$\varepsilon^2\{r_{(i)}\} = 1 + \sigma_0^2 (2 - \varphi)(\lambda_{(i)}^2 + k_{(i)}^2), \quad (25)$$

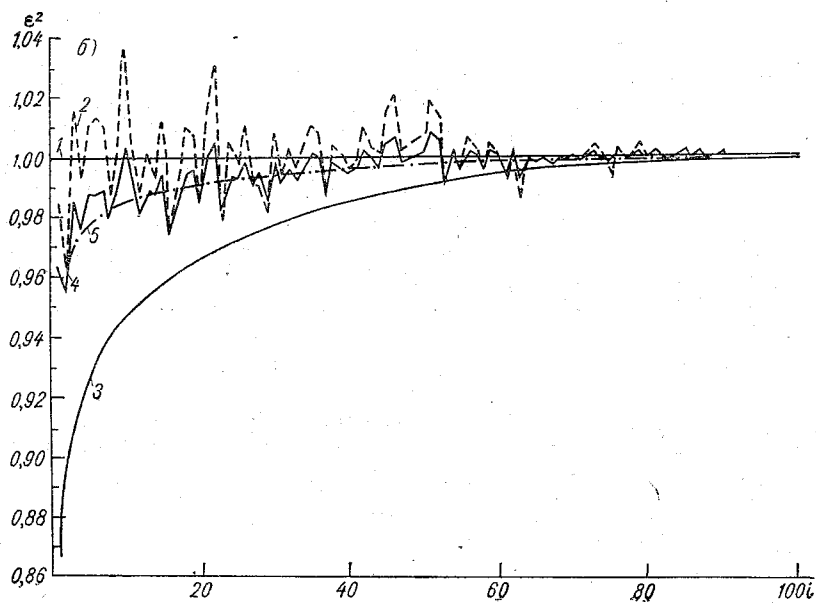
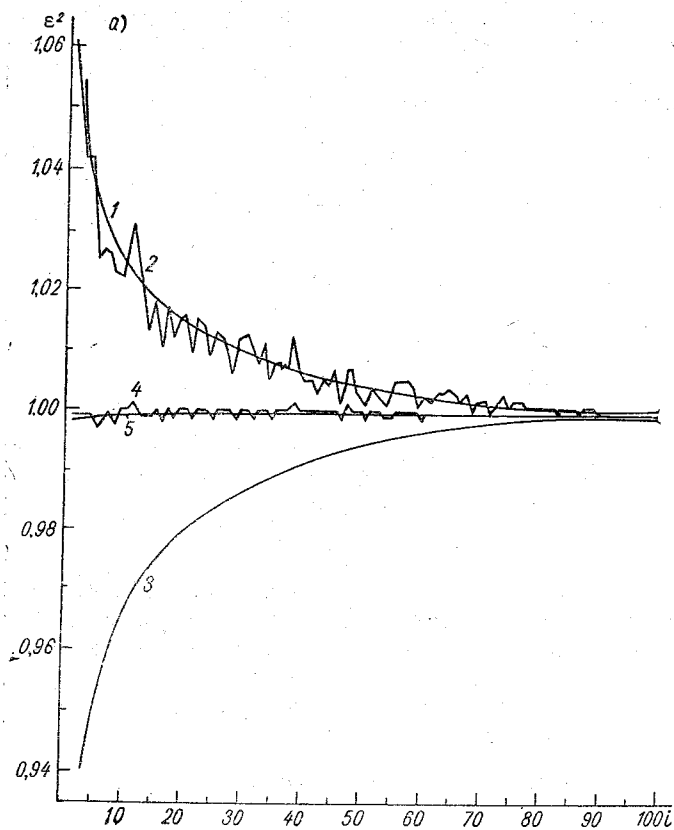
где

$$k_{(i)} = \frac{\sigma\{r_{(i)}\}}{\sigma_0} < 1.$$

Результаты расчета по формуле (25) для φ , равных 1,2; 2,0 и 4,0, представленные на рис. 1, обнаруживают интересную особенность — для слабопредсказуемых процессов ($\varphi < 2$) отбор наибольших по абсолютной величине выборочных коэффициентов корреляции приводит не к уменьшению, как это обычно предполагается, а к увеличению ошибки независимых прогнозов. Это означает, что использование разнообразных процедур просеивания предикторов в этом случае может лишь ухудшить результаты независимых прогнозов. Совсем по-другому дело обстоит, когда $\varphi > 2$. В этом случае отбор информативных предикторов увеличивает точность не только зависимых, но и независимых прогнозов.

В случае, когда в качестве коэффициента регрессии в схеме с отбором информативных предикторов используется несмещенная оценка (12), то с учетом (25) получим

$$\varepsilon^2\{r_{(i)}\} = 1 - \sigma_0^2 \frac{(\varphi - 1)^2}{\varphi} (\lambda_{(i)}^2 + k_{(i)}^2). \quad (26)$$



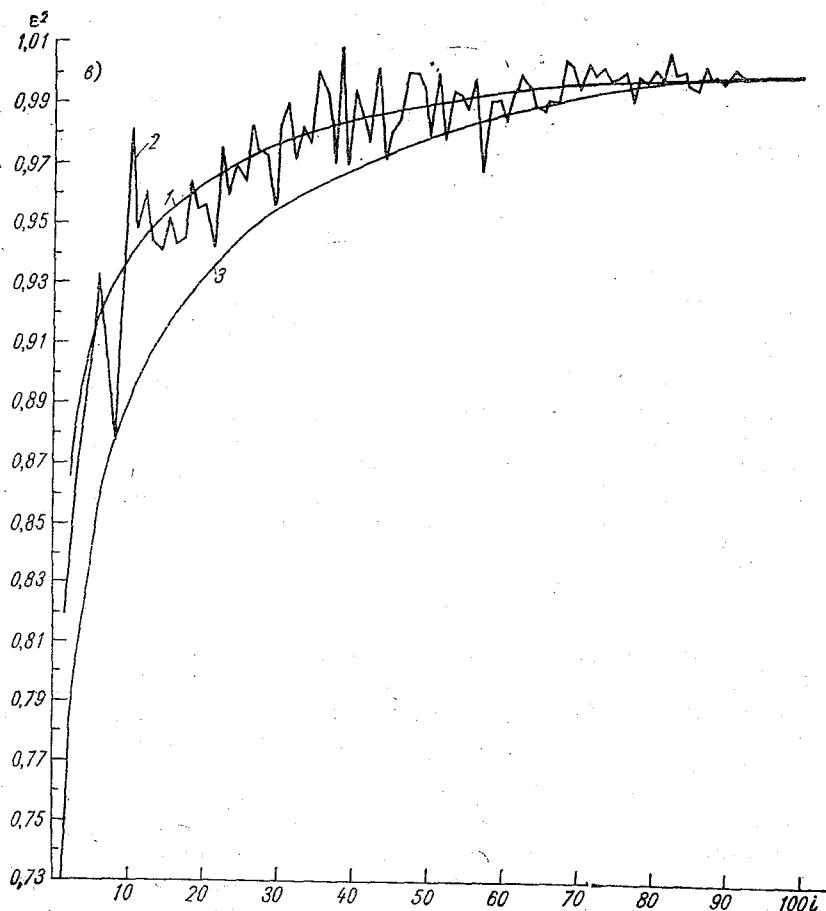


Рис. 1. Зависимость математического ожидания относительной ошибки прогноза от порядкового номера i , ряда ранжированных по абсолютной величине коэффициентов корреляции между предиктантом и предикторами.

а) $\varphi=1,2$; б) $\varphi=2,0$; в) $\varphi=4,0$; 1 — теоретическая; 2 — экспериментальная оценка относительной ошибки $E^2\{r(i)\}$ независимых прогнозов, выполненных по методу наименьших квадратов; 3 — экспериментальная оценка относительной ошибки независимых прогнозов, выполненных по методу наименьших квадратов; 4 — экспериментальная; 5 — теоретическая оценка относительной ошибки $E^2\{\hat{r}(i)\}$ независимых прогнозов, выполненных по методу несмещенного оценивания (12).

Как видно из выражения (26), введение коррекции позволяет эффективно использовать процедуру просеивания предикторов при любом значении φ . Поскольку при выводе полученных формул использовался ряд допущений прежде всего приближенного равенства (4), то для проверки их точности был использован метод статистических испытаний.

Для моделирования предиктанта и совокупности n тест-предикторов использовался стандартный датчик случайных чисел, обладающий нормальным распределением с нулевым математическим ожиданием и единичной дисперсией.

На первом этапе формировался ряд «предиктанта» Y_i , состоящий из N случайных чисел. На втором этапе формировалась совокупность предикторов, обладающих по отношению к моделируемому ряду предиктанта заданным распределением информативности вида (6)

$$X_{t,i} = R_i + \xi_{t,i} \sqrt{1 - R_i^2}, \quad (27)$$

где ξ — нормально распределенные случайные числа с нулевым математическим ожиданием и единичной дисперсией.

Значения генеральных коэффициентов корреляции R_i по условию задачи должны обладать нормальным распределением с нулевым математическим ожиданием и дисперсией, определяемой выражением (16). В соответствии с этим в качестве генерального коэффициента корреляции между предиктантом и i -м предиктором использовались следующие величины:

$$R_i = \xi \frac{\sigma_0}{\sqrt{2\beta}} = \xi \frac{\sqrt{\varphi-1}}{\sqrt{N-3}}. \quad (28)$$

Статистические испытания проводились следующим образом.

1. Генерировалась зависимая выборка объема N , на которой рассчитывались выборочные коэффициенты корреляции r_i , выборочные коэффициенты регрессии и выборочное значение индекса информативности φ .

2. Все предикторы переупорядочивались в порядке убывания абсолютной величины выборочных коэффициентов корреляции. В соответствии с этим вводилась новая нумерация предикторов и для каждого из них рассчитывались зависимые прогнозы $\hat{Y}_{t,i}$.

3. В соответствии с новой системой упорядочивания генерировалась независимая выборка объемом $N_{\text{нез}}$ и на ней для каждого из предикторов рассчитывались $N_{\text{нез}}$ прогнозов $\hat{Y}_{t,i}$. Качество прогнозов на зависимой и независимой выборках оценивалось при помощи относительной ошибки прогноза

$$\varepsilon_i^2 = \frac{\sum_{t=1}^N (Y_t - \hat{Y}_{t,i})^2}{\sum_{t=1}^N (Y_t - \bar{Y})^2}. \quad (29)$$

Поскольку полученные таким образом оценки не обладают свойством эргодичности, процедура 1—3 проводилась K раз. В результате была получена осредненная по K реализациям оценка зависимости относительной ошибки прогноза от номера порядковой статистики i , в качестве которой рассматривались абсолю-

ные величины рассчитанных на зависимой выборке коэффициентов корреляции между предиктантом и n предикторами. Номер порядковой статистики $i=1$ соответствует максимальному по абсолютной величине коэффициенту корреляции, т. е. «наилучшему» с этой точки зрения предиктору, соответственно при $i=n$ мы имеем дело с «наихудшим» предиктором.

В ходе статистических испытаний оценивалась точность зависимых и независимых прогнозов, составленных по методу наименьших квадратов и методу несмещенных оценок (12).

Эксперименты проводились при следующих значениях параметров: длина зависимой и независимой выборок $N=N_{\text{нез}}=100$, количество предикторов в совокупности $n=100$; величины относительной ошибки прогнозов осреднялись по 200 реализациям.

Результаты расчетов по методу Монте-Карло для характерных величин индекса информативности предикторов $\varphi=1,2; 2,0$ и $4,0$, приведенные на рис. 1, хорошо отражают обсуждавшиеся выше особенности расчетов, выполненных по формулам (25) и (26).

Из этого следует, что принятое в настоящей работе в качестве первого приближения равенство плотностей вероятности коэффициентов корреляции и их фишеровских преобразований не оказывает существенного влияния на точность расчетов по крайней мере до величины $\beta>0,5$.

Заметим, что при коррекции коэффициента регрессии в методе Монте-Карло, в отличие от теоретических оценок (26), использовались выборочные, а не генеральные значения индекса φ . Тем не менее данные экспериментов свидетельствуют о возможности получения относительной ошибки независимых прогнозов меньше единицы даже при $\varphi=1,2$. Этот эффект может быть увеличен при использовании теории несмещенного оценивания применительно к уравнениям регрессии, содержащим не один, а несколько предикторов.

СПИСОК ЛИТЕРАТУРЫ

1. Розенблит П. Я. Статистическое оценивание характеристик надежности и эффективности сложных систем. — Рига: Зинатне, 1979. — 272 с.
2. Юдин М. И., Мещерская А. В., Блажевич В. Г. Характеристика гидрометеорологической информации, использованной при долгосрочном физико-статистическом прогнозе температуры и осадков для районов недостаточного увлажнения. — Труды ГГО, 1970, вып. 236, с. 45—63.
3. Боярский Э. А. Порядковые статистики. — М.: Статистика, 1972. — 119 с.

О РАСПРЕДЕЛЕНИИ КЛИМАТИЧЕСКИХ ВЕТРОЭНЕРГОРЕСУРСОВ ПО ТЕРРИТОРИИ СССР

В последние годы в связи с постепенным истощением невозполнимых источников энергии резко возрос интерес к использованию атмосферных источников энергии, в частности энергии ветра. Для ветроэнергетики метеорологические факторы имеют решающее значение. Они определяют целесообразность установки ветродвигателя в том или ином месте, выбор параметров ветродвигателя и планирование режима его работы. Климатологическая информация необходима при определении потенциальных и утилизируемых¹ ветродвигателем ветроэнергоресурсов. Вопросы климатических ветроэнергоресурсов рассмотрены в работе [2]. В этой работе предложена методика оценки потенциальных и утилизируемых ветродвигателями определенного типа климатических ресурсов энергии ветра по данным о статистическом распределении повторяемости различных скоростей ветра, приведенных в Справочнике по климату СССР.

Если ветродвигатель полностью использует энергию ветра, то сила давления ветра на лопасти ветрового колеса пропорциональна ветровому напору $\rho v^2/2$ (v — скорость ветра, ρ — плотность воздуха), а скорость движения лопастей пропорциональна скорости ветра. Поэтому мощность ветродвигателя пропорциональна $\rho v^3/2$.

Зная дифференциальную функцию распределения скорости ветра, легко оценить потенциальные ветроэнергоресурсы по формуле

$$\overline{N}_n = \bar{\rho} \int_0^{\infty} v^3 f(v) dv. \quad (1)$$

В качестве ρ можно использовать среднее многолетнее среднемесячное значение плотности.

Данная методика была использована для расчетов ветроэнергоресурсов на ряде станций территории Советского Союза как для отдельных месяцев, так и для года в целом. Однако полученные на близлежащих станциях значения заметно отличаются друг от друга и с трудом поддаются картированию. Представляется целесообразным использование для расчетов не самих таблиц повторяемостей, а полученных на их основе параметров распреде-

¹ Потенциальная — это ветроэнергетическая мощность, которую вырабатывал бы ветродвигатель, если бы он одинаковым образом реагировал на любую скорость ветра. Утилизируемая — это ветроэнергетическая мощность, которую вырабатывает ветродвигатель с данными параметрами.

ления, поля которых являются более гладкими, что должно заметно облегчить их картирование.

Как известно, распределение скорости ветра для открыто расположенных станций удовлетворительно описывается двухпараметрическим законом Вейбулла:

$$f(v) = \exp\left\{-\left(\frac{v}{\beta}\right)^\gamma\right\}, \quad (2)$$

где $f(v)$ — функция распределения скорости ветра; β — масштабный множитель, близкий к средней скорости ветра; γ — параметр, характеризующий форму распределения.

Такая аппроксимация применительно к скорости ветра широко использовалась для статистической экстраполяции максимальных скоростей ветра (см., например, [1, 7]). Наиболее массовые расчеты были проведены С. Д. Кошинским [3, 5].

Однако, насколько нам известно, в литературе не имеется достаточно полных сведений о распределении этих параметров по территории Советского Союза. Они опубликованы лишь для некоторых районов С. Д. Кошинским.

В этой связи полученные ранее значения параметров β и γ не могут быть использованы для оценки ветроэнергоресурсов и для их определения потребовались специальные расчеты.

Для определения параметров β и γ существует несколько способов. Наиболее широкое применение получил описанный в [1] графический метод, заключающийся в использовании специальной вероятностной клетчатки.

С. Д. Кошинский рекомендовал и использовал метод наименьших квадратов, позволяющий по заданным значениям $f(v)$ и v объективно оценить значения β и γ . Параметры β и γ можно получить также, используя для аппроксимации метод максимального правдоподобия [6].

Сопоставление возможной точности различных способов аппроксимации распределения скорости ветра показывает, что применение двухпараметрического закона Вейбулла для аппроксимации распределения скорости ветра дает вполне удовлетворительные результаты. Наиболее надежные значения параметров β и γ могут быть получены методом максимального правдоподобия. При этом следует иметь в виду, что оценка точности аппроксимации проводилась применительно к расчетам ветроэнергоресурсов, для которых решающее значение имеет правильное описание распределения скорости ветра в диапазоне больших ее значений.

Для получения значений β и γ были использованы данные 200 станций на территории Советского Союза о повторяемости различных скоростей ветра, помещенные в Справочнике по климату СССР. Полученные при аппроксимации распределения скорости ветра методом максимального правдоподобия значения β и γ были использованы для картирования. Следует, однако, иметь в виду, что их применение значительно шире.

Зная параметры β и γ , можно получить моменты распределения скорости ветра. Для начального момента порядка k имеет место соотношение

$$M_k(\beta, \gamma) = \beta^k \Gamma\left(1 + \frac{k}{\gamma}\right). \quad (3)$$

Отсюда следует, что средняя скорость $\bar{v}$, среднее квадратическое отклонение σ_v и коэффициент вариации c_v могут быть определены по формулам

$$\bar{v} = \beta \Gamma\left(1 + \frac{1}{\gamma}\right), \quad (4)$$

$$\sigma_v = \beta \sqrt{\Gamma\left(1 + \frac{2}{\gamma}\right) - \Gamma^2\left(1 + \frac{1}{\gamma}\right)}, \quad (5)$$

$$c_v = \frac{\sigma_v}{\bar{v}} = \frac{\sqrt{\Gamma\left(1 + \frac{2}{\gamma}\right) - \Gamma^2\left(1 + \frac{1}{\gamma}\right)}}{\Gamma\left(1 + \frac{1}{\gamma}\right)}. \quad (6)$$

Ветроэнергоресурсы пропорциональны величине $\bar{v}^3$, которая определяется как третий начальный момент:

$$\bar{v}^3 = \beta^3 \Gamma\left(1 + \frac{3}{\gamma}\right). \quad (7)$$

Коэффициент вариации ветроэнергоресурсов может быть получен по формуле

$$c_{\bar{v}^3} = \frac{\sqrt{\Gamma\left(1 + \frac{6}{\gamma}\right) - \Gamma^2\left(1 + \frac{3}{\gamma}\right)}}{\Gamma\left(1 + \frac{3}{\gamma}\right)}. \quad (8)$$

На рис. 1 приведено распределение средних годовых значений параметра β по территории Советского Союза. Параметр β имеет размерность скорости ветра, близок к $\bar{v}$ по величине и его распределение соответствует распределению средней скорости ветра. Разумеется, рисунок носит осредненный характер и показывает фоновое распределение параметра β . На отдельных станциях значения β могут существенно различаться.

Отношение параметра β к $\bar{v}$

$$\frac{\beta}{\bar{v}} = \frac{1}{\Gamma\left(1 + \frac{1}{\gamma}\right)} \quad (9)$$

зависит от параметра γ . Из рис. 2 видно, что оно составляет 1—1,13 и несколько растет с увеличением γ .

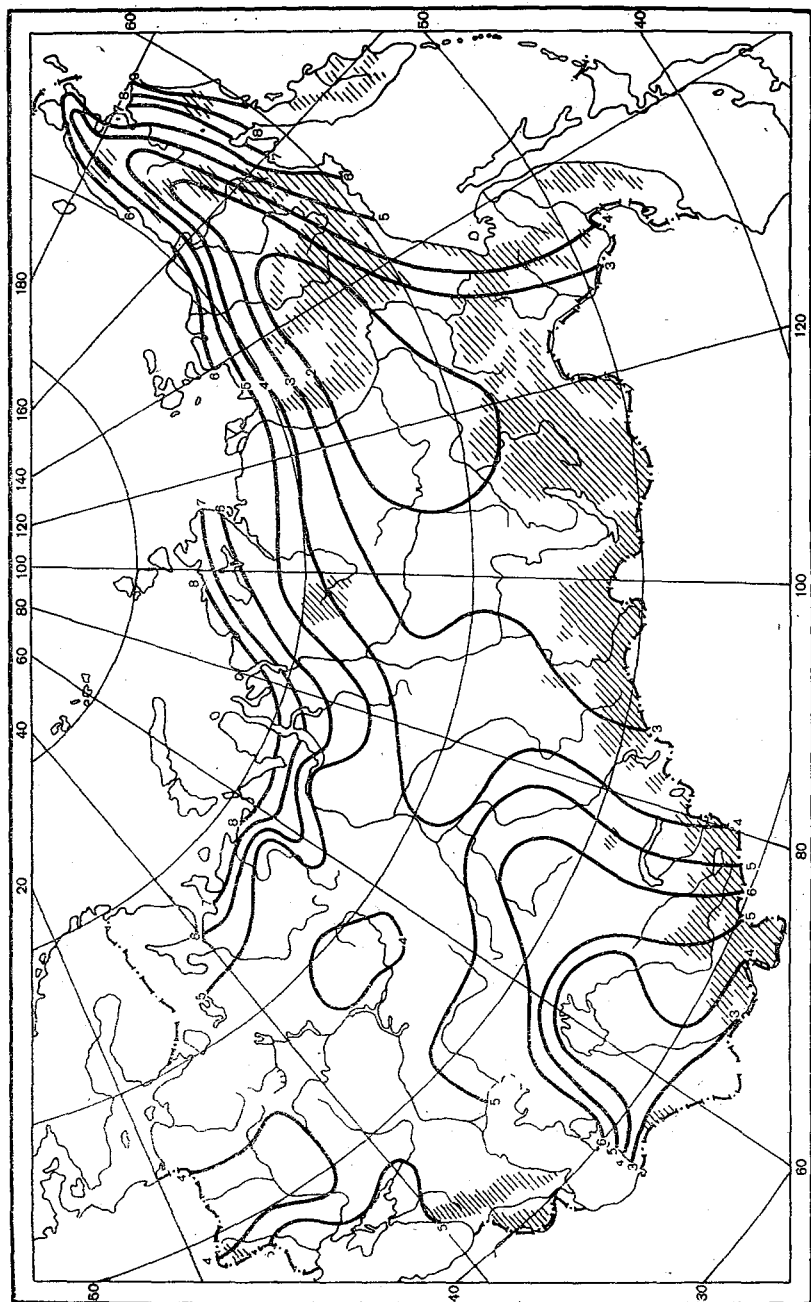


Рис. 1. Распределение значений параметра β по территории Советского Союза.

Как видно из формул (6) и (8), любые безразмерные характеристики скорости ветра являются функцией параметра γ . На рис. 3 приведена зависимость коэффициента вариации c_v от параметра γ . Эта зависимость обратно пропорциональна γ и приблизительно можно считать, что

$$c_v = \frac{1}{\gamma}.$$

Распределение параметра γ по территории Советского Союза (рис. 4) показывает, что этот параметр изменяется от 1,1 до 1,6. На отдельных станциях γ достигает значений 1,8—1,9. На территории Восточной Сибири γ меняется незначительно, в среднем составляя 1,1—1,2, повышаясь до 1,4 на побережье Охотского моря и Северного Ледовитого океана.

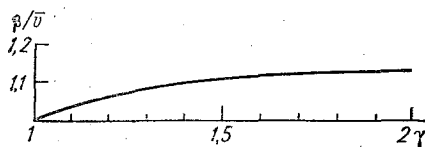


Рис. 2. Зависимость отношения $\beta/\bar{v}$ от параметра γ .

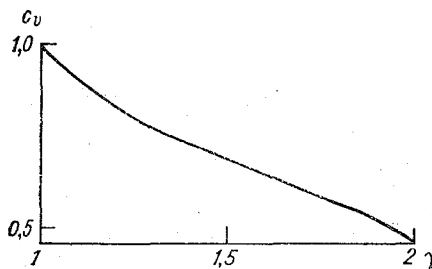


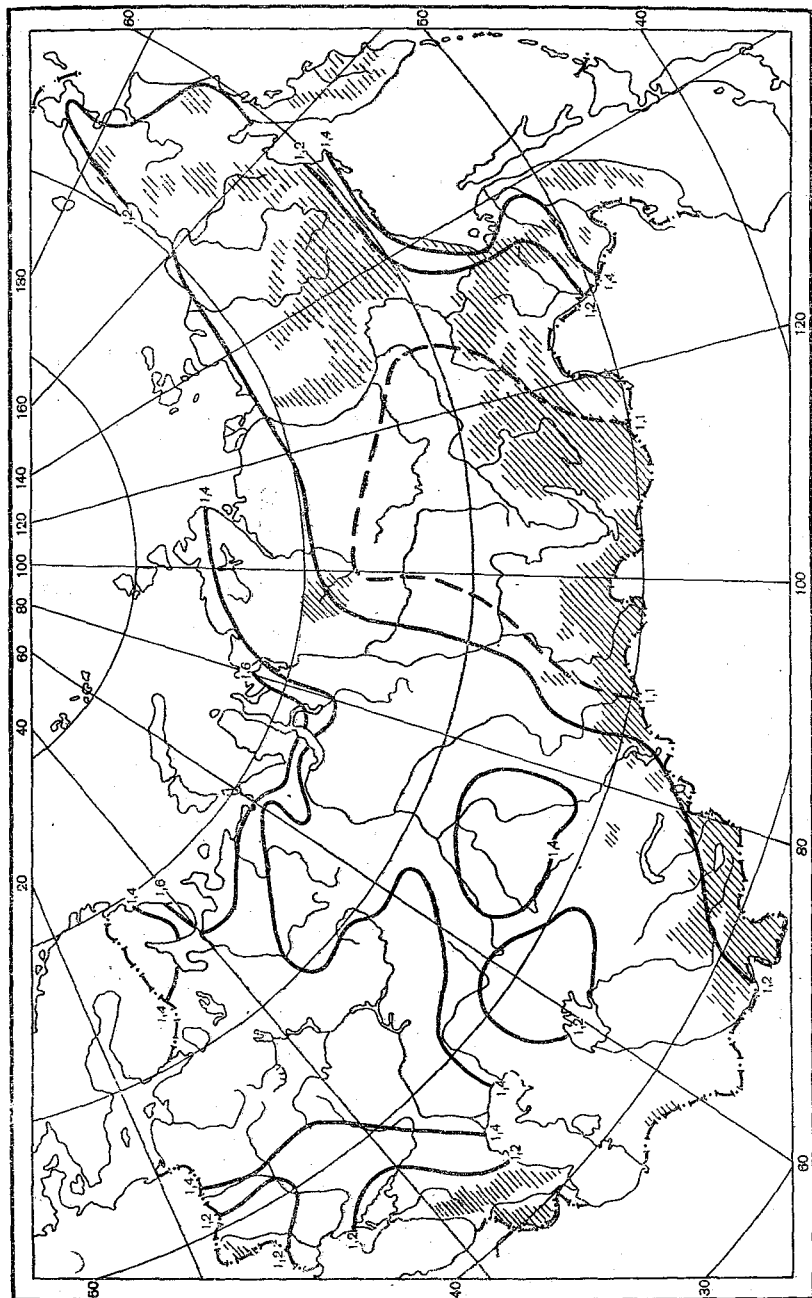
Рис. 3. Зависимость коэффициента вариации c_v от параметра γ .

В центральной части ЕТС γ в среднем составляет 1,4—1,6, уменьшаясь к югу до 1,2.

Рисунки 1 и 4, где представлено распределение параметров β и γ по территории Советского Союза, были использованы для получения этих параметров в узлах регулярной сетки на расстоянии 400 км. По полученным значениям β и γ рассчитаны $\bar{v}^3$ в узлах регулярной сетки, нанесены на карту (рис. 5) и проанализированы. Наибольшие запасы $\bar{v}^3$ (до 1200 м³/с³) отмечаются на побережье Северного Ледовитого океана и Охотского моря. Существенными запасами энергии ветра (400—600 м³/с³) обладают север Казахстана и северо-восток ЕТС.

Следует иметь в виду, что на рис. 5 представлена не ветроэнергетическая мощность, определяемая формулой (1), а $\bar{v}^3$, т. е. величина, пропорциональная $\bar{N}_n$. В большинстве случаев для определения ветроэнергетической мощности можно воспользоваться стандартным значением ρ .

Учитывать отличие ρ от стандартного его значения 1,275 кг/м³ важно лишь для пунктов, расположенных достаточно высоко над уровнем моря. Например, на высоте 1500 м плотность меньше стан-



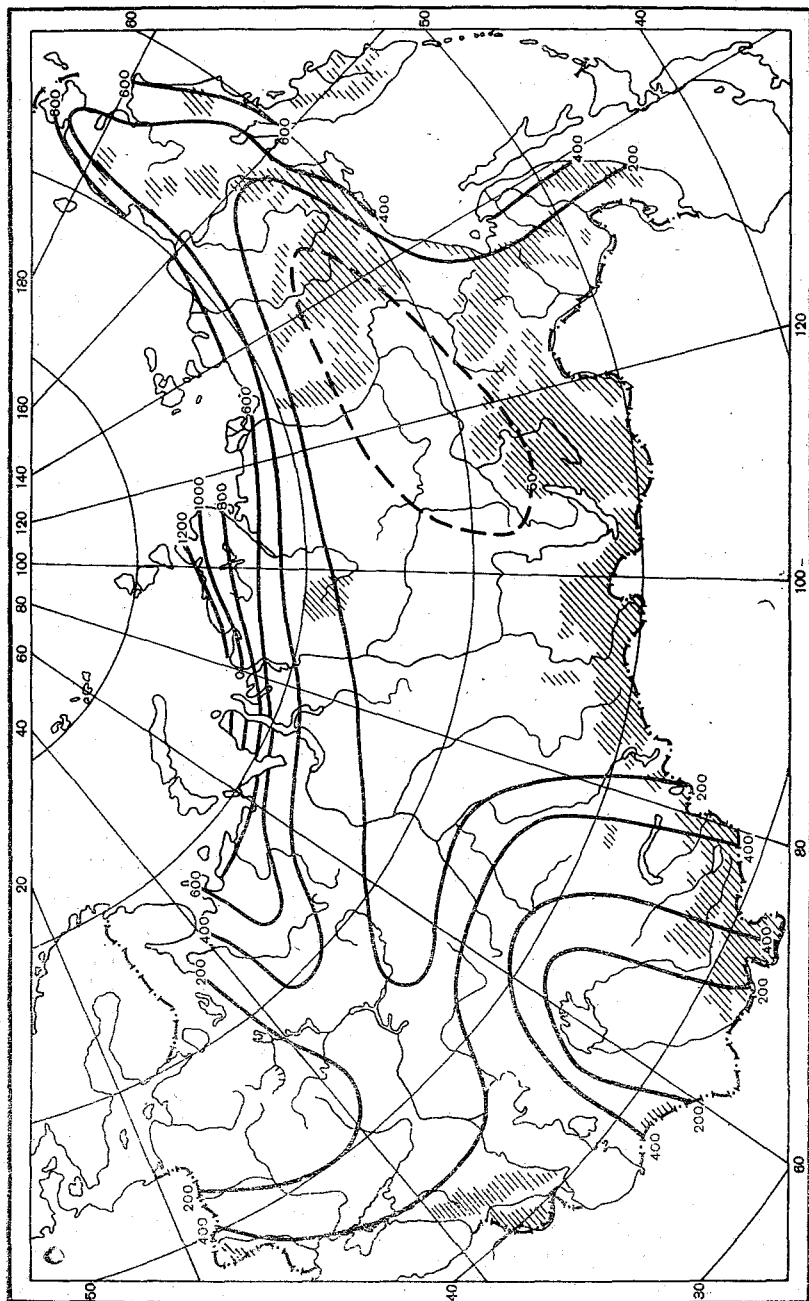


Рис. 5. Распределение изотерм по территории Советского Союза.

дартного значения примерно на 10 %. Соответственно меньшими получаются и ветроэнергетические ресурсы.

Годовой ход ρ обусловлен годовым ходом температуры, его амплитуда составляет несколько процентов. Тем не менее, рассматривая месячные значения ветроэнергоресурсов, годовой ход плотности воздуха следует принимать во внимание. Неучет его приведет к систематическим ошибкам — занижению ветроэнергоресурсов зимой и завышению их летом.

Представленное на рис. 5 распределение ветроэнергоресурсов по территории Советского Союза может быть использовано при крупномасштабном планировании ветроэнергетической системы. Для принятия решения о целесообразности установки ветродвигателей в том или ином месте и определения оптимальных параметров ветродвигателей этого не достаточно. В частности, необходимо оценить утилизируемые ветродвигателем энергетические ресурсы, которые зависят как от характеристик ветрового режима, так и от

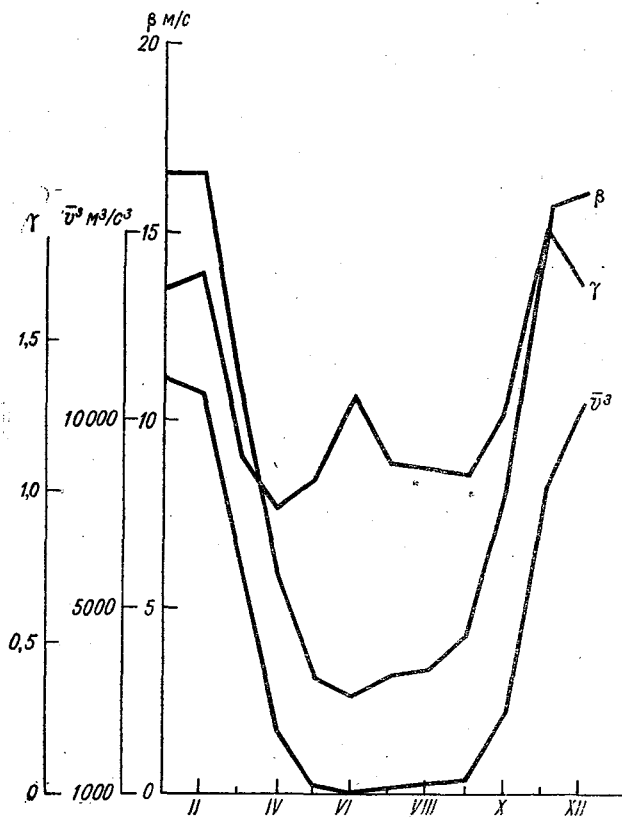


Рис. 6. Годовой ход $\bar{v}_z$, β , γ . Пестря Дресва

параметров ветродвигателя. Некоторые результаты опубликованы в [4].

На рис. 6 в качестве примера приведен годовой ход $\bar{v}^3$, а также параметров β и γ на ст. Пестрая Дресва. Годовой ход параметра β соответствует годовому ходу средней скорости ветра с максимумом зимой и минимумом летом. Что касается параметра γ , то он, как правило, не имеет выраженного годового хода. На ст. Пестрая Дресва вследствие особенностей ее расположения средняя месячная скорость ветра с ноября по февраль превышает 14 м/с. Потенциальные запасы энергии ветра здесь огромны, в зимние месяцы они достигают 11 000 м³/с³. Следует заметить, что современные ветродвигатели могут использовать только незначительную часть этих возможностей.

СПИСОК ЛИТЕРАТУРЫ

1. Анапольская Л. Е. Режим скоростей ветра на территории СССР. — Л.: Гидрометеиздат, 1961. — 199 с.
2. Анапольская Л. Е. Ветроэнергетические ресурсы и методы их оценки. — Метеорология и гидрология, 1978, № 7, с. 11—17.
3. Кошинский С. Д. Режимные скорости ветра на морях Советского Союза. Ч. 1. — Каспийское море, Л., Гидрометеиздат, 1975, 412 с.
4. Кошинский С. Д. Критические замечания о методах расчета максимальных скоростей ветра различной обеспеченности. Труды НРГМЦ СССР, 1969, вып. 2, с. 3—32.
5. Кошинский С. Д. Из опыта расчета некоторых статистических характеристик ветра и параметров функции распределения вида $f(x) = \exp \left\{ - \left(\frac{x}{\beta} \right)^\gamma \right\}$ на электронно-вычислительных машинах. — Труды НРГМЦ СССР, 1969, вып. 2, с. 44—53.
6. Крамер Г. Математические методы статистики. — М.: Мир, 1975. — 648 с.
7. Справочник по климату СССР. — Л.: Гидрометеиздат, 1965—1969. Ч. III Вып. 1—34.

СОДЕРЖАНИЕ

Р. Л. Каган, Е. И. Хлебникова. О нарушении однородности метеорологических полей при их интерполяции	3
Р. Л. Каган, И. А. Салль, Н. Н. Щербакова. О пространственной корреляции средних месячных значений общего содержания озона в атмосфере	18
Е. И. Хлебникова. Об искажении распределений негауссовских величин при их пространственной интерполяции	26
Ю. М. Либерман. Эффективность учета ветра при анализе геопотенциала	32
Р. Л. Каган. К вопросу об аппроксимации временной структуры барического поля в тропосфере	42
Р. Л. Каган, Е. И. Хлебникова. К вопросу об оценке значимости экстремумов в вековом ходе метеорологических элементов	55
Р. Л. Каган, И. А. Салль. Об уточнении оценок средних суточных значений метеорологических параметров	69
Р. Л. Каган, К. М. Лугина. Некоторые оценки возможностей мониторинга общего содержания озона в атмосфере	81
Р. Л. Каган, Е. Е. Сибир. Об одном алгоритме статистического моделирования рядов метеорологических наблюдений	94
Ю. В. Житорчук, Л. Е. Страшненко. Оптимизация структуры посевных площадей сельскохозяйственных культур с учетом климатологической информации	104
А. Т. Житорчук, Ю. В. Житорчук. Учет априорной информации в регрессионных моделях прогноза погоды	110
Л. Л. Брагинская. О распределении климатических ветроэнергоресурсов по территории СССР	120

Труды ГГО, вып. 466

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ
В МЕТЕОРОЛОГИИ

Редактор Г. И. Слабкович. Технический редактор М. С. Костакова. Корректор Т. И. Черненко

Н/К Сдано в набор 27.09.82. Подписано в печать 18.08.83. М-29156. Формат 60×90¹/₁₆, бумага тип. № 2.
Гарнитура литературная. Печать высокая. Печ. л. 8,5. Кр.-отт. 8,75. Уч.-изд. л. 9,55. Тираж 630 экз.
Индекс МЛ-193. Заказ № 172. Цена 70 коп. Заказное.

Гидрометеиздат. 199053, Ленинград, 2-я линия, д. 23.

Сортавальская книжная типография Государственного комитета Карельской АССР по делам издательств, полиграфии и книжной торговли. Сортавала, Карельская, 42.