

В. В. БАВЫКИНА, канд. техн. наук, Ю. А. КОВАЛЬ, д-р техн. наук, А. А. ТКАЧУК,
канд. техн. наук, С. В. ТРАМБОВЕЦКИЙ

СТАТИСТИЧЕСКИЙ АНАЛИЗ ВЫБОРОК РЕЗУЛЬТАТОВ РАДИОМЕТЕОРНЫХ ИЗМЕРЕНИЙ СДВИГА ШКАЛ ЭТАЛОНОВ ВРЕМЕНИ

Радиометеорный метод синхронизации (РМС) в настоящее время применяется в Государственных службах единого времени и эталонных частот (ГСВЧ) Украины и России. Этот метод по точности синхронизации шкал времени не уступает перевозимым квантовым часам, спутниковым радионавигационным системам GPS и ГЛОНАСС, но обладает оперативностью и автономностью. РМС отличается независимостью от зарубежных систем, используемых для синхронизации шкал времени. Достоинства РМС определяют актуальность научных работ по его совершенствованию. Результаты измерений получены при штатных сравнениях шкал координированного времени UTC(UA) и UTC(SU) государственных эталонов Украины и России соответственно, проведенных в течение 2000 г. с использованием специальных радиометеорных комплексов аппаратуры «МЕТКА-6». Эти комплексы аппаратуры разработаны в Харьковском национальном университете радиоэлектроники (ХНУРЭ) и предназначены для автоматического высокоточного определения расхождения шкал времени разнесенных территориально хранителей. Аппаратура размещается в пунктах расположения хранителей времени, удаленных друг от друга примерно на 750 км. Было обработано 53 сеанса. Измерения проводились в среднем один раз в неделю. Начало измерений совпадало с началом суток. Средняя продолжительность сеансов была примерно один час. Основные статистические показатели радиометеорного канала синхронизации на трассе Харьков-Москва за 2000-й год приведены в табл. 1.

Таблица 1

Характеристика		Значение	Процент
Число сеансов с числом измерений	минимальным, равным единице	2	3,7
	малым (от 2 до 10)	17	32,1
	близким к среднему (от 10 до 100)	28	52,8
	равным или большим 100	6	11,3
Число сеансов с числом метеорных следов	меньшим 4	32	60,4
	близким к среднему (от 4 до 6 включительно)	18	33,9
	большим 6	3	5,7
Общее число обработанных сеансов		53	100
Характеристика		Значение	
Число метеорных следов в сеансе	Максимальное	17	
	Среднее	4,8	
	Минимальное	1	
Общее число «полезных» метеорных следов во всех обработанных сеансах		179	
Число результатов измерений в сеансе	Максимальное	1230	
	Среднее	66,9	
	Минимальное	1	
Число результатов измерений, полученных на метеорном следе	Максимальное	256	
	Среднее	48	
	Минимальное	1	
Общее число результатов измерений		3546	

В большинстве работ принималось (иногда по умолчанию), что распределение результатов сравнения подчиняется нормальному закону, а выборки таких результатов можно считать однородными и равноточными [1], в соответствии с этими предположениями создавались и алгоритмы обработки результатов, базирующиеся на традиционных классических процедурах. С нашей точки зрения такие подходы нуждаются в пересмотре, что и явилось целью настоящей работы.

Основные задачи работы заключаются в следующем: провести анализ данных и выполнить проверку статистических гипотез о нормальности распределения выборок, их однородности и равноточности.

Сеансы отличаются друг от друга продолжительностью, числом так называемых «полезных» метеоров, обеспечивающих отражение вперед (в пункт корреспондента), числом измерений, полученных от этих метеоров и общим числом измерений. Представление о распределении числа измерений за сеанс по дням 2000-го года дает рис. 1.

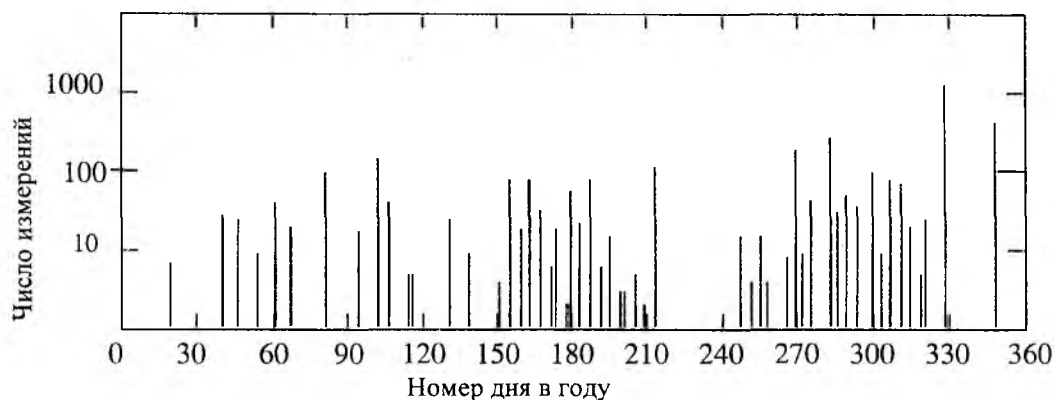


Рис. 1

Следует заметить, что большое число сеансов содержало малое число измерений. Рисунок 2 дает представление о кумулятивном распределении числа измерений по сеансам.

Результаты измерений за сеанс при статистической обработке рассматриваются как элементы выборок, а общий объем совокупной выборки определяется числом измерений за сеанс.

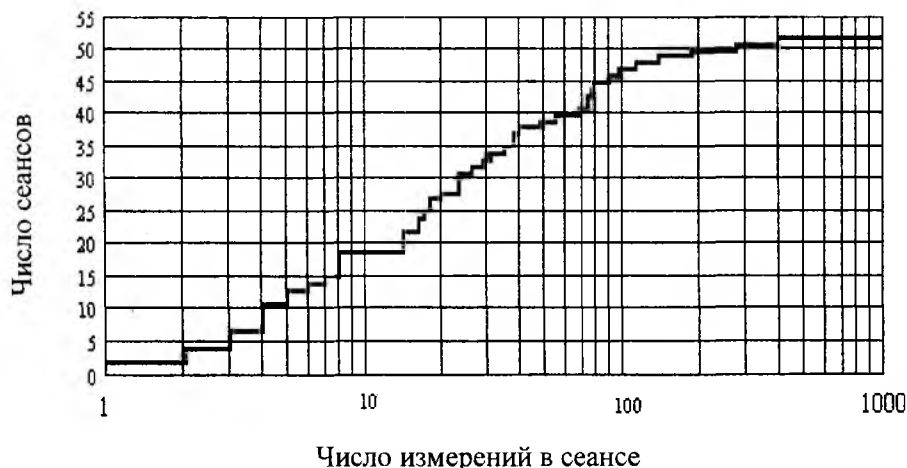


Рис. 2

Результаты измерений, полученные от «полезного» метеорного следа образуют группу, число таких измерений в группе составляют объем группы. На рис. 3 показано, как связаны число «полезных» метеорных следов в сеансе и число сеансов.

Рис. 4 дает представление о характере изменений результатов измерений во времени для метеорных следов, существующих длительное время. Типичными для таких реализаций яв-

ляются увеличение рассеяния к концу радиозо от метеорного следа и в некоторых случаях изменение среднего значения во времени. Вероятные причины такого поведения кроются в особенностях физики метеорных явлений: дифракции на образующемся метеорном следе, ветровом перемещении отражающей области, диффузном расширении метеорного следа, резонансе в метеорном следе, многолучевом распространении радиоволн и других аналогичных причинах [2 – 4]. Вид типичных реализаций результатов измерений, полученных на большом статистическом материале, позволяет сделать вывод, что имеет место существенная нестационарность процессов, связанных с метеорными явлениями, которая обусловлена эффектами формирования и разрушения метеорного следа, а также изменением притока метеорного вещества в течение суток и года.

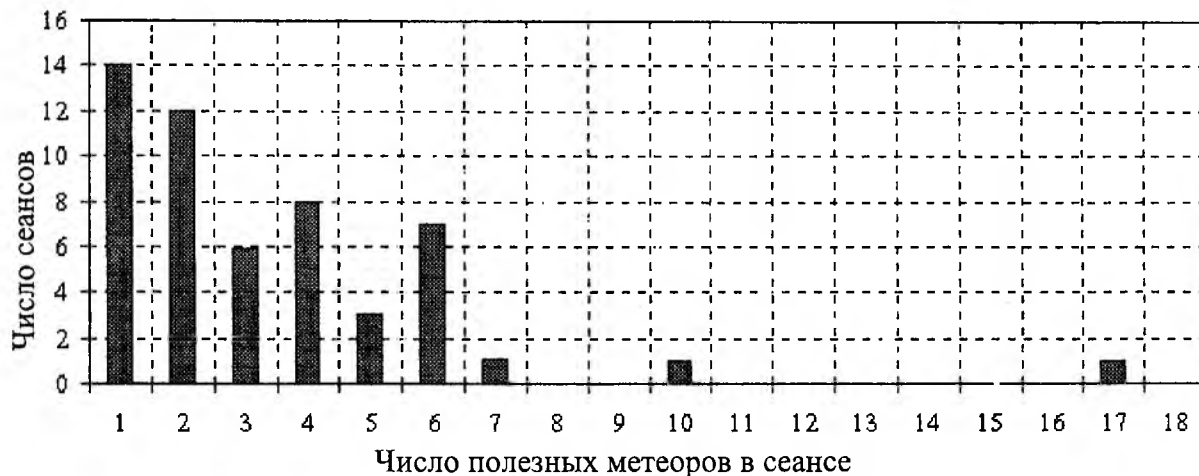


Рис. 3

Результаты измерений шкал времени, полученные от каждого «полезного» метеорного следа за сеанс, рассматривались как группа, подлежащая статистической обработке, и обозначались x_{ji} , где j – номер метеорного следа в сеансе, обеспечившего сличение, i – номер сличения в этой группе. Принималось, что число «полезных» метеоров в сеансе – L , общее число измерений в группе составляют объем группы – n_j , а общий объем совокупной выборки результатов измерений шкал времени за сеанс определяется формулой $n = \sum_{j=1}^L n_j$. Для

каждой группы: x_{j1}, \dots, x_{jn_j} , где $j = \overline{1, L}$ были вычислены выборочные средние

$$\bar{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ji} \text{ и выборочные дисперсии } s_j^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (x_{ji} - \bar{x}_j)^2. \text{ (При } n_j = 1 \text{ принималось}$$

$\bar{x}_j = x_{j1}$, а $s_j^2 = 0$). Затем с помощью аппарата проверки статистических гипотез устанавливалось, являются ли некоторые различия рассчитанных оценок существенными (значимыми) или случайными.

Предполагалось, что результаты измерений, полученные от отдельных метеорных следов, распределены по гауссовому (нормальному) закону. Следует заметить, что в соответствии с критерием согласия χ^2 эта гипотеза на уровне значимости 0.05 в 93% случаев не противоречит материалу наблюдения, очищенному от случайных выбросов.

В работе использован аппарат проверки статистических гипотез, хорошо разработанный для выборок с гауссовым законом распределения [5].

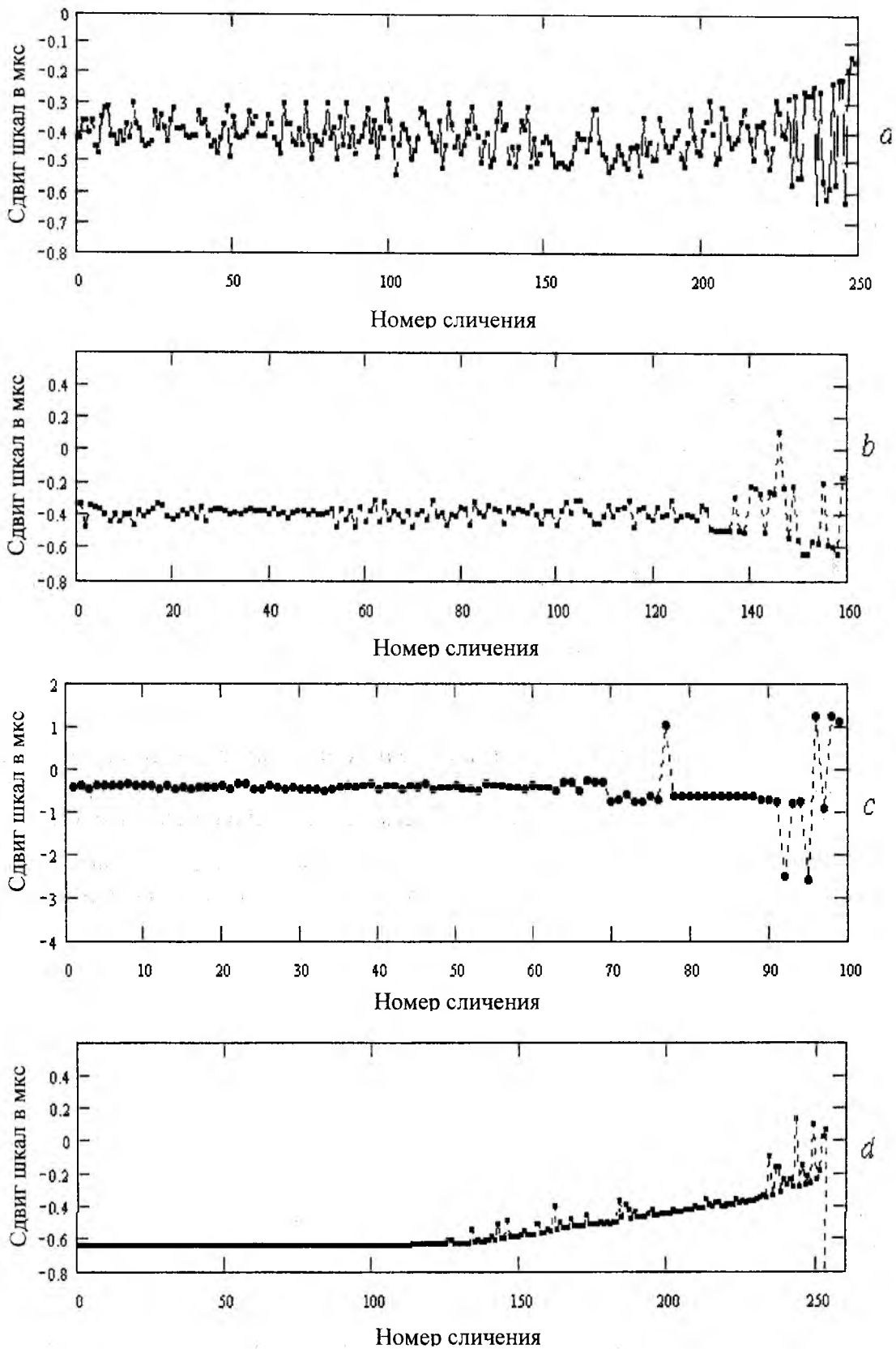


Рис. 4

Для проверки гипотезы H_0 относительно одного из параметров нормального распределения рассчитывалась некоторая статистика d , вычисляемая по различным формулам, приведенным в табл. 2, в зависимости от вида проверяемой гипотезы (о равенстве средних, о ра-

венстве дисперсий), от числа групп (для $L=1$, $L=2$ и $L \geq 3$) и от того, что известно о втором параметре. По заданной выборке вычислялось фактическое значение статистики d и сравнивалось с критическим значением d_α . Уровень значимости гипотезы α принимался равным 0.05. При $d > d_\alpha$ делался вывод, что гипотеза H_0 опровергнута опытом, а при $d < d_\alpha$ гипотеза признавалась не противоречащей имеющимся опытным данным.

Проверка однородности групп с нормальным распределением обычно выполняется путем сравнения средних значений и дисперсий выборок. Группы рассматриваются как однородные и равнозначные, если выполняются гипотезы равенства средних и выборочных дисперсий. Однородными не могут считаться группы, для которых гипотезы равенства средних отвергаются. Если же гипотеза равенства средних для всех групп, входящих в сеанс принимается, а гипотеза равенства выборочных дисперсий не выполняется, то такие сеансы можно считать однородными, но не равнозначными.

При проверке групп на однородность использовалась статистическая гипотеза о равенстве средних. Для двух групп ($L=2$) при неизвестных и неравных дисперсиях, что соответствует нашему случаю, рассчитывалась величина, указанная в табл. 2 (строка 2). При проверке равенства средних для случая, когда число групп $L \geq 3$, сравнивались значения оценки межгрупповой дисперсии

$$s_{bg}^2 = \frac{1}{L-1} \sum_{j=1}^L n_j (\bar{x}_j - \bar{x})^2, \text{ где } \bar{x} = \frac{1}{n} \sum_{j=1}^L n_j \bar{x}_j, \text{ и среднее значение}$$

$$\text{внутригрупповых оценок дисперсий } s_{ig}^2 = \frac{1}{n-L} \sum_{j=1}^L (n_j - 1) s_j^2. \text{ Критерий основан на статисти-$$

стике s_{bg}^2 / s_{ig}^2 , которая имеет распределение Фишера с $L-1$, $n-1$ степенями свободы (строка 3. табл. 2).

При проверке групп на равнозначность использовалась статистическая гипотеза о равенстве дисперсий. Для проверки равенства дисперсий двух групп $H_0: \sigma_1^2 = \sigma_2^2$ используется

отношение оценок дисперсий s_1^2 / s_2^2 , которое имеет распределение Фишера с $n_1 - 1$, $n_2 - 1$ степенями свободы (строка 5 табл.2). При проверке равенства дисперсий $L \geq 3$ групп использовался приближенный критерий Бартлетта, который основан на сравнении оценок дисперсий групп s_j^2 с их средним значением s_{ig}^2 . Статистика Бартлетта приведена в табл. 2 в

строке 6, где $c = 1 + \left[\sum_{j=1}^L \frac{1}{n_j - 1} + \frac{1}{n - L} \right] / [3(L-1)]$; ее распределение аппроксимируется χ^2 - распределением с $f = L-1$ степенями свободы.

Использованные критерии сравнения при проверке статистических гипотез об однородности и равнозначности выборок приведены в табл. 2. В этой таблице приняты следующие обозначения: z_p - квантиль порядка P гауссовского распределения с нулевым математическим ожиданием и единичной дисперсией $N(0,1)$, определяемая условием $P\{|z| < z_p\} = P$;

t_p - квантиль порядка P распределения Стьюдента с указанным числом степеней свободы, т. е. $P\{|t| < t_p\} = P$; χ_α^2 и $\chi_{1-\alpha}^2$ - процентные точки χ^2 -распределения с $L-1$ числом степеней

свободы, которые удовлетворяют условиям: $P\{\chi^2 < \chi_\alpha^2\} = P\{\chi^2 > \chi_{1-\alpha}^2\} = \alpha$, $P = 1 - 2\alpha$;

F_α - процентные точки распределения Фишера с числом степеней свободы f_1 и f_2 и уровнем значимости гипотезы α . При этом вероятность отклонить верную гипотезу не более α . Указанные квантили и процентные точки определяются либо по статистическим таблицам

[6], либо непосредственно вычисляются в пакетах прикладных программ, используемых для статистической обработки.

Таблица 2

Номер гипотезы	Число метеороров L	Гипотеза H_0 : условие	Критическая область (правило отклонения гипотезы H_0)	Число степеней свободы	Число сеансов
Проверка на однородность (равенство средних значений)					
1	1	$a = a_0$: СКО не известно	$ \bar{x} - a_0 \sqrt{n} / s > z_p$	$f = n - 1$	14
2	2	$a_1 = a_2$: $\sigma_1 \neq \sigma_2$ - не известно	$ \bar{x}_1 - \bar{x}_2 / \sqrt{s_1^2 / n_1 + s_2^2 / n_2} > t_p$	$f = n_1 + n_2 - 1$	12
3	≥ 3	$a_j = a_2$: $\sigma_j = \sigma$ - не известно	$s_{bg}^2 / s_{ig}^2 \geq F_\alpha$	$f_1 = L - 1$; $f_2 = n - L$	27
Проверка на равнозначность (равенство дисперсий)					
4	1	$\sigma = \sigma_0$; $\sigma_0 = 20$ нс	$s^2 / \sigma_0^2 > \chi_\alpha^2$ или $< \chi_{1-\alpha}$	$f = n - 1$; $P = 1 - 2\alpha$;	14
5	2	$\sigma_1^2 = \sigma_2^2$	$s_1^2 / s_2^2 > F_\alpha$ или $< F_\alpha^{-1}$	$f_1 = n_1 - 1$; $f_2 = n_2 - 1$	12
6	≥ 3	$\sigma_j^2 = \sigma_1^2$	$\frac{2,303}{c} \sum_{j=1}^L (n_j - 1) \lg(s_{ig}^2 / s_j^2) > \chi_P^2$	$f = L - 1$	27

Необходимые расчеты были выполнены для всех 53 сеансов, из которых 14 состояло из одного «полезного» метеорного следа, 12 – из двух, а 27 сеансов содержало от трех до 17 метеорных следов. Во всех случаях рассчитанные фактические значения статистик, соответствующих той или иной проверяемой гипотезе, сравнивались с критическими при принятом уровне значимости гипотезы.

Результаты проверки статистических гипотез о равенстве выборочных средних и равенстве выборочных дисперсий представлены в табл. 3.

Графическая интерпретация результатов проверки статистических гипотез об однородности и равнозначности выборок результатов измерений, полученных от метеорных следов в течение сеанса, дана на рис. 5 в виде круговой диаграммы, на которой отображены в процентах частоты появления сложившихся ситуаций 1...5. Позиции 1...5, принятые на рис. 5, соответствуют пяти типам возникающих ситуаций и расшифровываются в табл. 3.

Использование вероятностной логики принятия или отклонения гипотезы позволяет сделать следующие выводы: в 26 % случаев данные групп нельзя считать равнозначными; в 6% случаев группы не могут быть отнесены к одному и тому же закону распределения из-за того, что гипотеза о равенстве средних не подтверждается, а в 8 % случаев отклоняются обе

гипотезы о равенстве средних и о равенстве дисперсий.

Итак, только 60 % сеансов (32 сеанса из 53) составляют группы, которые могут быть отнесены к одному и тому же распределению, причем, в 26 % случаев сеанс состоял только из одного «полезного» метеорного следа, а в 40 % случаев результаты измерений шкал времени за сеанс нельзя считать однородной и равноточной выборкой.

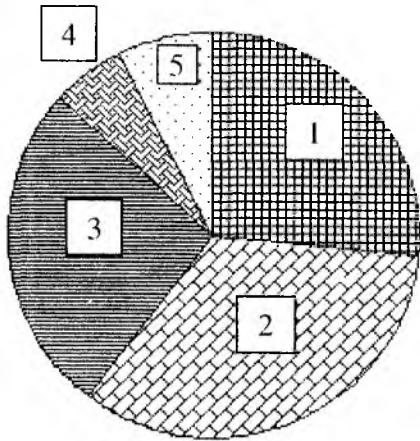


Рис. 5

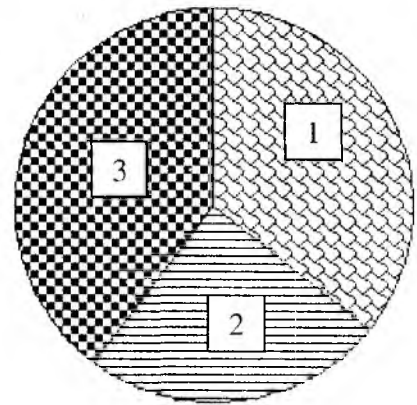


Рис. 6

Кроме того, в работе для каждого сеанса рассматривалось распределение результатов измерений в сеансе и проверялась гипотеза о нормальном законе распределения с параметрами (математическим ожиданием и дисперсией), оцененными по выборкам.

Таблица 3

Но- мер	Ситуация	Сеансы		Сеансы	
		Число	%	Число	%
1	Гипотеза о равенстве средних значений принимается	14	26	32	60
	Гипотеза о равенстве дисперсий принимается ($L = 1$)				
2	Гипотеза о равенстве средних значений принимается	18	34	21	40
	Гипотеза о равенстве дисперсий принимается ($L \geq 2$)				
3	Гипотеза о равенстве средних значений принимается	14	26	21	40
	Гипотеза о равенстве дисперсий отклоняется ($L \geq 2$)				
4	Гипотеза о равенстве средних значений отклоняется	3	6	21	40
	Гипотеза о равенстве дисперсий принимается ($L \geq 2$)				
5	Гипотеза о равенстве средних значений отклоняется	4	8	21	40
	Гипотеза о равенстве дисперсий отклоняется ($L \geq 2$)				
	Всего	53	100	53	100

Как видно из рис. 2, большое число сеансов содержало малое число измерений. Так, 19 сеансов, что составляет 35,8 % от их общего числа, содержат число измерений меньше 10. Для таких сеансов проверять гипотезу о законе распределения по критериям согласия, требующим больших объемов выборок (χ^2 -квадрат критерий Пирсона, критерий Колмогорова), нецелесообразно. Остальные 34 сеанса (64,2 %) имеют число измерений $n \geq 14$, причем, 6 сеансов содержат более 100 измерений. Для этих 34 сеансов с числом измерений $n \geq 14$ проверялось по χ^2 -критерию, удовлетворяет ли рассматриваемая случайная величина X , представляющая результаты измерения сдвига шкал за сеанс и имеющая эмпирический закон распределения $F(x)$, заданному закону распределения $F_0(x)$. По выборке результатов изме-

рений за сеанс определялась контрольная статистика $\chi^2 = \sum_{i=1}^K \frac{(M_i - np_i)^2}{np_i}$, где K – число классов, на которые разбивается размах выборки, M_i – число элементов выборки в i -ом классе, p_i – теоретическая вероятность попадания в этот класс, n – объем выборки. Критическое значение χ_{α}^2 для заданного уровня значимости α и $m = K - 3$ степеней свободы находили, используя соответствующую функцию из использованного для статистических расчетов пакета прикладных программ. При $\chi^2 < \chi_{\alpha}^2$ принималось, что гипотеза о гипотетически нормальном законе распределения не противоречит экспериментальным данным.

В результате оказалось, что только для 13 сеансов из рассмотренных 34 сеансов с достаточным числом измерений гипотеза о нормальном законе распределения может быть принята, что составляет 24,5 % от общего числа сеансов. В табл. 4 приведены выполненные для всех сеансов результаты проверки статистических гипотез о законе распределения результатов измерений за сеанс,

Таблица 4

Номер	Ситуация	Сеансы		Сеансы	
		Число	%	Число	%
1	Гипотеза о нормальном законе распределения не может быть проверена	19	35,8	32	60,3
2	Гипотеза о нормальном законе распределения принимается	13	24,5		
3	Гипотеза о нормальном законе распределения отклоняется	21	39,7	21	39,7
	Всего	53	100	53	100

Графическая интерпретация результатов проверки статистических гипотез о законе распределения результатов измерений за сеанс в виде круговой диаграммы приведена на рис. 6.

Итак, в 24,5 % случаев гипотеза нормальном законе распределения не противоречит экспериментальным данным, в 35,8 % случаев из-за малого объема выборок проверка такой гипотезы не может быть выполнена, а в 39,7 % случаев наблюдаемые выборочные значения не согласуются с гипотетическим распределением.

Проверка статистических гипотез показала, что результаты измерений, представленные 53 сеансами за 2000 год, примерно в 40 % случаев не могут рассматриваться как нормально распределенные, однородные и равноточные выборки. Эти статистические свойства результатов измерений сдвига шкал эталонов времени по радиометеорному каналу связи обосновывают необходимость поиска помехоустойчивых процедур для их обработки.

Список литературы: 1. *Кащеев Б.Л., Коваль Ю.А., Горбач В.И. и др.* Метеоры сегодня. Киев: Техника, 1991. 196 с. 2. *Михайлов В. К., Сидоров В. В.* Фазовая нестабильность метеорного канала РРВ, вызванная ветровым перемещением отражающей области // Метеорное распространение радиоволн. Казань. 1976. Вып. 12. С. 3-14. 3. *Сидоров В. В., Михайлов В. К., Иванушкина Т. С.* Влияние резонанса в метеорном следе и диффузного расширения на фазовую нестабильность метеорного канала // Метеорное распространение радиоволн. Казань. 1980. Вып. 16. С. 53-54. 4. *Иванушкина Т. С., Михайлова Н. Ю., Сидоров В. В.* К вопросу о влиянии многолучевости на полосу пропускания метеорного канала // Метеорное распространение радиоволн. Казань. 1979. Вып. 14. С. 77-85. 5. *Уилкс С.* Математическая статистика/ Пер. с англ. М.: Наука, 1967. 632 с. 6. *Грановский В.А., Сирая Т.Н.* Методы обработки экспериментальных данных при измерениях. Л.: Энергоатомиздат, 1990. 288 с.

Харьковский национальный
университет радиоэлектроники

Поступила в редколлегию 08.01.2002