ГОСУДАРСТВЕННЫЙ КОМИТЕТ СССР ПО СТАНДАРТАМ

ВСЕСОЮЗНЫЙ ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ ФИЗИКО-ТЕХНИЧЕСКИХ
И РАДИОТЕХНИЧЕСКИХ ИЗМЕРЕНИЙ

(ВНИИФТРИ)

Казанский филиал

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

ОЦЕНКА ДОСТОВЕРНОСТИ ЧИСЛЕННЫХ ДАННЫХ О СВОЙСТВАХ НЕФТИ И НЕФТЕПРОДУКТОВ

МИ 656—84

Москва
ИЗДАТЕЛЬСТВО СТАНДАРТОВ
1985

gost.e-ecolog.ru

РАЗРАБОТАНЫ Казанским филиалом Всесоюзного ордена Трудового Красного Знамени научно-исследовательского института физико-технических и радиотехнических измерений (ВНИИФТРИ); Всесоюзным научно-исследовательским центром по изучению свойств поверхности и вакуума (ВНИЦ ПВ)

ИСПОЛНИТЕЛИ:

В. Г. Гизатуллина, канд. техн. наук А. Д. Козлов (руководитель темы)

ВНЕСЕНЫ Казанским филиалом ВНИИФТРИ

Директор Н. М. Хусаинов

УТВЕРЖДЕНЫ И ВВЕДЕНЫ В ДЕЙСТВИЕ научно-техническим советом ВНИЦ ПВ от 11 декабря 1984 г. (протокол № 13)

УДК 665.6 : 53.081.7

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

ОЦЕНКА ДОСТОВЕРНОСТИ ЧИСЛЕННЫХ ДАННЫХ О СВОЙСТВАХ НЕФТИ И НЕФТЕПРОДУКТОВ

МИ 656—84

Настоящие методические указания устанавливают единые принципы оценки достоверности данных о физико-химических и эксплуатационных свойствах нефти и нефтепродуктов.

1. ОБЩИЕ ПОЛОЖЕНИЯ

1.1. Оценку данных о свойствах нефти и нефтепродуктов проводят с целью определения степени достоверности этих данных, установления числового интервала их погрешности и отбора наиболее надежных данных для обеспечения потребности народного хозяйства достоверной информацией о свойствах нефти и нефтепродуктов, необходимой при оценке их качества.

1.2. Под оценкой достоверности численных данных понимают:

выбор и обоснование методов отбора и анализа исходных данных;

нахождение оптимальных требований к методам и средствам измерений, применяемых при исследовании показателя или свойства;

выбор и обоснование методов анализа и обобщения данных для конкретной области их применения;

определение алгоритмов обработки данных.

1.3. При оценке достоверности численных данных о свойствах нефти и нефтепродуктов необходимо руководствоваться ГОСТ 8.344—79, ГОСТ 22013—76, ГОСТ 8.326—78, МУ 38.101003—81, РД 50-262—81, РД 50-326—82.

2. ОТБОР ИСХОДНЫХ ДАННЫХ

2.1. При отборе численных данных о свойствах нефти и нефтепродуктов необходимо учитывать следующие характеристики продуктов и их свойств:

марку продукта и его назначение;

технологию получения продукта;

условия транспортировки и хранения;

исходное сырье;

метод определения продукта и его свойств.

2.2. Исходной информацией при отборе численных данных о свойствах нефти и нефтепродуктов являются результаты испыта-

© Издательство стандартов, 1985

ний, полученные: стандартизованными методами; методами, включенными в комплексы квалификационной оценки; нестандартизованными методами, прошедшими метрологическую аттестацию в соответствии с МУ 38.101003—81; на специальных установках, прошедших метрологическую аттестацию в соответствии с ГОСТ 8.326—78.

2.3. При отборе данных предпочтительно использовать результаты, полученные по аттестованным методикам (РД 50-326—82).

3. ОБРАБОТКА ЧИСЛЕННЫХ ДАННЫХ

3.1. Настоящие методические указания рекомендуют оценивать данные двумя величинами: математическим ожиданием M и дисперсией σ^2 при заданном уровне значимости $\alpha=0{,}05$. В основу определения этих величин положено следующее предположение: результаты испытаний $x_1, x_2, \ldots, x_n$ являются независимыми в совокупности величинами и распределены по нормальному закону.

3.2. Качественное заключение о нормальности распределения получают графическим путем, используя метод гистрограмм (рекомендуется для случая, когда $n>50$). Если объем имеющихся численных данных недостаточен ($n<50$), то гистограмму строят по совокупности малых выборок. Однако строить гистограмму в этом случае возможно, если выборки малого объема проверены попарно на однородность.

3.3. Однородность выборок проверяют по критерию Вилкоксона и его модификации — критерию Манна—Уитни. Пусть $x_1, \ldots, x_{n_1}$, $x'_2, \ldots, x'_{n_2}$ —данные двух выборок. Обозначим через $U_{x'<x}$ число пар $(x_i; x'_j)$, где $x'_j<x_i$. Приведенная статистика Манна-Уитни ($U_{x'<x}$) линейно связана со статистикой Вилкоксона (T_x; $T_{x'}$) соотношением

$$U_{x'<x}=n_1\cdot n_2+n_1(n_1+1)/2-T_{x'}=T_x-n_2(n_2+1)/2, \qquad (1)$$

где n_1, n_2 — число данных в каждой выборке.

Для нахождения статистики T_x и $T_{x'}$ из двух выборок составляется общий вариационный ряд $x_{(1)}<x_{(2)}<\ldots<x_{(n_1+n_2)}$, в котором для одного значения x_i определяют номер места* r_i — ранг. Аналогично определяют ранги a_j для одного значения x'_j.

Вычисляют значения T_x и $T_{x'}$

$$T_x=\sum_{i=1}^{n_1} r_i; \quad T_{x'}=\sum_{j=1}^{n_2} a_j \qquad (2)$$

* Если два или более значения совпадают, то им приписывают один и тот же ранг, равный среднему арифметическому рангов, несовпавших значений.

Однородность выборок не отвергается, если эмпирическое значение $U_{x'<x}$ превосходит табличное значение $U_{табл}$ (при уровне значимости $\alpha=0{,}05$).

Значения $U_{табл}$ представлены в табл. 1 приложения 1 при условии

$$P\{U_{x'<x} \leqslant U_{табл}\} \leqslant \alpha.$$

Критерий Вилкоксона применим для любого вида распределения и не требует большого числа данных в каждой выборке ($n \geqslant 5$).

Однородность выборок по критерию Вилкоксона и Манна—Уитни выявляется как по форме распределения, так и по средним.

3.4. Если гипотеза об однородности выборок не противоречит данным общей выборки, то выполняют проверку крайних значений по ГОСТ 11.002—73 при уровне значимости, равном $\alpha=0{,}05$, и строят гистограмму.

3.5. Проверку нормальности закона распределения выполняют с помощью коэффициентов асимметрии и эксцесса, согласно приложению I ГОСТ 22013—76.

3.6. Рассчитывают параметры распределения: среднее арифметическое $\bar{x}_i$ и среднее квадратическое отклонение S_i для каждой i-й выборки общей выборки.

4. ПРОВЕРКА СОГЛАСОВАННОСТИ

4.1. Проверку согласованности данных выполняют по сумме минимального отклонения оценок дисперсий каждой i-й выборки общей выборки ($n \geqslant 5$). Наилучшей оценкой в этом случае является среднее взвешенное $\bar{\bar{x}}$:

$$\bar{\bar{x}} = \frac{\sum\limits_{i=1}^{L} \bar{x}_i / S_i^2}{\sum\limits_{i=1}^{L} 1/S_i^2}, \qquad (3)$$

где S^2_i — дисперсия каждой i-й выборки; L — число всех выборок.

Согласованность данных не отвергается, если сохраняется условие

$$A^2_{min} \approx F, \qquad (4)$$

где F — число степеней свободы ($F=N-L$) при уровне значимости $\alpha=0{,}05$; N — число всех данных; A^2_{min} — вычисленное значение, равное

$$A^2_{min} = \sum_{i=1}^{L} \frac{(\bar{x}_i - \bar{\bar{x}})^2}{S_i^2}. \qquad (5)$$

4.2. Если вычисленное значение A^2_{min} превышает ожидаемое значение F, то рекомендуется выполнить анализ накопленных данных и выявить факторы, влияющие на их разброс.

4.3. Для приведения численных данных к согласию в ряде случаев (для фундаментальных констант) рекомендуется использовать расчетный коэффициент Берджа. Согласование численных данных по коэффициенту Берджа выполняют в соответствии с рекомендованным приложением 2.

5. ДОВЕРИТЕЛЬНЫЕ ГРАНИЦЫ СРЕДНЕЙ АРИФМЕТИЧЕСКОЙ И СРЕДНЕГО КВАДРАТИЧЕСКОГО ОТКЛОНЕНИЯ

5.1. Анализ точности данных чаще основан на подсчете точечных оценок. Однако ошибки, допускаемые в процессе испытаний, приводят к большому разбросу данных, который не позволяет объективно оценить исследуемое свойство. Поэтому наряду с вычислением точечных оценок рекомендуется использовать интервальные. Интервальное оценивание в данном случае основано на построении доверительных интервалов для выборочной средней $\bar{x}_i$ и для выборочной дисперсии S^2_i при доверительной вероятности $P=0{,}95$.

В табл. 1 приложения 4 дан пример вычисления точечных и интервальных оценок.

5.1.1. При построении доверительного интервала для выборочной средней $\bar{x}_i$ используют распределение Стьюдента. Из таблиц распределения Стьюдента с n_i-1 степенями свободы находят величину t_{n_i-1}, для которой справедливо равенство

$$P\left\{-t_{\alpha/2} < \frac{(\bar{x}_i - M)}{S_i} \cdot \sqrt{n_i} < t_{\alpha/2}\right\} \geqslant 1-\alpha. \qquad (6)$$

Преобразуя выражение (6), строят двусторонний доверительный интервал для каждой i-й выборки:

$$\bar{x}_i - \frac{s_i \cdot t_{\alpha/2}}{\sqrt{n_i}} < M < \bar{x}_i + \frac{S_i \cdot t_{\alpha/2}}{\sqrt{n_i}}, \qquad (7)$$

где $t_{\alpha/2}$ — квантили уровня $1-\alpha/2$, $\alpha/2$; n_i — число данных в i-й выборке.

5.1.2. Доверительный интервал для выборочной дисперсии S^2_i строят на основании того, что случайная величина $(n_i \cdot S_i^2)/\sigma^2$ имеет распределение χ^2 с n_i-1 степенями свободы, т. е. $n_i \cdot S_i^2/\sigma^2 \sim \chi^2_{n_i-1}$. где $\sigma^2 > n_i \cdot S_i^2/\chi^2_{n_i-1}$.

Доверительный интервал для выборочной дисперсии строят из условия

$$P\left\{\chi^2_{n_i-1,\ \alpha/2} \leqslant \frac{n_i \cdot S_i^2}{\sigma^2} \leqslant \chi^2_{n_i-1,\ \alpha/2}\right\} \geqslant 1-\alpha. \qquad (8)$$

Преобразуя выражение (8) в $P\left\{\frac{1}{\underline{\chi}^2_{n_i-1,\ \alpha/2}} < \sigma^2 < \frac{1}{\overline{\chi}^2_{n_i-1,\ \alpha/2}}\right\} \geqslant 1-\alpha$,

строят двусторонний доверительный интервал

$$P\left\{\frac{n_i \cdot S_i^2}{\underline{\chi}^2_{n_i-1,\ \alpha/2}} < \sigma^2 < \frac{n_i \cdot S_i^2}{\overline{\chi}^2_{n_i-1,\ \alpha/2}}\right\} \geqslant 1-\alpha, \qquad (9)$$

где $\underline{\chi}^2_{n_i-1,\ \alpha/2}$; $\overline{\chi}^2_{n_i-1,\ \alpha/2}$ — квантили, которые определяются по χ^2 с n_i-1 степенями свободы при уровне значимости $\alpha=0{,}05$.

Значения χ^2 представлены в табл. 1 приложения 3.

5.2. Чем больше взят объем выборки, тем выше надежность оценки, т. е. тем более узкий доверительный интервал может быть принят при оценке данных.

ПРИЛОЖЕНИЕ 1
Обязательное

Значение наибольших $U_{табл}$ при $(U_{x'<x} \leqslant U_{табл}) \leqslant \alpha = 0{,}05$

n_2	n_1 = 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	—																			
2	—	—																		
3	—	—	—																	
4	—	—	—	—																
5	—	—	—	—	0															
6	—	—	—	0	1	2														
7	—	—	—	0	1	3	4													
8	—	—	—	1	2	4	6	7												
9	—	—	0	1	3	5	7	9	11											
10	—	—	0	2	4	6	9	11	13	16										
11	—	—	0	2	5	7	10	13	16	18	21									
12	—	—	1	3	6	9	12	15	18	21	24	27								
13	—	—	1	3	7	10	13	17	20	24	27	31	34							
14	—	—	1	4	7	11	15	18	22	26	30	34	38	42						
15	—	—	2	5	8	12	16	20	24	29	33	37	42	46	51					
16	—	—	2	5	9	13	18	22	27	31	36	41	45	50	55	60				
17	—	—	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70			
18	—	—	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81		
19	—	0	3	7	12	17	22	28	33	39	45	51	57	63	69	74	81	87	93	
20	—	0	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

СОГЛАСОВАНИЕ ДЫННЫХ

1. Согласование данных нарушается, если вычисленное значение A^2_{min} превышает ожидаемое значение F согласно п. 4.1.

Для устранения несогласованности данных рекомендуется использовать расчетный коэффициент Берджа

$$R_Б = \sqrt{A^2_{min}/F}. \quad (1)$$

В этих случаях вычисленное по п. 3.6 значение среднего квадратического отклонения S_i каждой i-й выборки общей выборки изменяют на $R_Б$ число, равное коэффициенту Берджа, и находят измененные значения среднего квадратического отклонения S'_i каждой i-й выборки:

$$S'_i = S_i \cdot R_Б \quad , \quad (2)$$

с помощью которых, аналогично величине A^2_{min}, вновь вычисляют величину A^2_{1min}, согласно п. 4.1.

Если вычисленное значение A^2_{1min} не превышает ожидаемое F (при $\alpha=0{,}05$), то согласованность данных принимается.

2. В случае, если необходимо вычислить коэффициент Берджа R_i для каждой i-й выборки, используют следующую формулу:

$$R_i = \frac{\sqrt{\sum_{i=1}^{n_i} (x_i - \bar{x}_i)^2/S_i^2}}{n_i - 1} \quad . \quad (3)$$

Затем находят измененные оценки дисперсий $(S''_i)^2$ каждой i-й выборки

$$(S''_i)^2 = S^2_i R_i \quad (4)$$

и вычисляют величину A^2_{2min}

$$A^2_{2min} = \sum_{i=1}^{L} \frac{(\bar{x}_i - \bar{\bar{x}})^2}{(S''_i)^2} \quad , \quad (5)$$

где $\bar{\bar{x}}$ — среднее взвешенное, вычисленное с помощью измененных оценок $(S''_i)^2$, согласно п. 4.1.

Согласованность данных принимается, если сумма квадратов относительных изменений оценок дисперсий $(S''_i)^2$ минимальна:

$$\sum_{i=1}^{L} [1 - (S''_i/S^2_i)] = \min$$

Значение χ^2 в зависимости от вероятности P и числа степеней свободы с (n_1-1) $\chi^2_{n_1-1}$ — распределение χ^2

K	Вероятность P												
	0,005	0,010	0,025	0,05	0,10	0,20	0,80	0,90	0,95	0,975	0,990	0,995	0,999
1	0,0393	0,0157	0,0982	0,0393	0,0158	0,0642	1,64	2,71	3,84	5,02	6,63	7,88	10,8
2	0,0100	0,0201	0,0506	0,103	0,211	0,446	3,22	4,61	5,99	7,38	9,21	10,6	13,8
3	0,0717	0,115	0,216	0,352	0,584	1,00	4,64	6,25	7,81	9,35	11,3	12,8	16,3
4	0,207	0,297	0,484	0,711	1,06	1,65	5,99	7,78	9,49	11,1	13,3	14,9	18,5
5	0,412	0,554	0,831	1,15	1,61	2,34	7,29	9,24	11,1	12,8	15,1	16,7	20,5
6	0,676	0,872	1,24	1,64	2,20	3,07	8,56	10,6	12,6	14,4	16,8	18,5	22,5
7	0,989	1,24	1,69	2,17	2,83	3,82	9,80	12,0	14,1	16,0	18,5	20,3	24,3
8	1,34	1,65	2,18	2,73	3,49	4,59	11,0	13,4	15,5	17,5	20,1	22,0	26,1
9	1,73	2,09	2,70	3,33	4,17	5,38	12,2	14,7	16,9	19,0	21,7	23,6	27,9
10	2,16	2,56	3,25	3,94	4,87	6,18	13,4	16,0	18,3	20,5	23,2	25,2	29,6
11	2,60	3,05	3,82	4,57	5,58	6,99	14,6	17,3	19,7	21,9	24,7	26,8	31,3
12	3,07	3,57	4,40	5,23	6,30	7,81	15,8	18,5	21,0	23,3	26,2	28,3	32,9
13	3,57	4,11	5,01	5,89	7,04	8,63	17,0	19,8	22,4	24,7	27,7	29,8	34,5
14	4,07	4,66	5,63	6,57	7,79	9,47	18,2	21,1	23,7	26,1	29,1	31,3	36,1
15	4,60	5,23	6,26	7,26	8,55	10,3	19,3	22,3	25,0	27,5	30,6	32,8	37,7
16	5,14	5,81	6,91	7,96	9,31	11,2	20,5	23,5	26,3	28,8	32,0	34,3	39,3
17	5,70	6,41	7,56	8,67	10,1	12,0	21,6	24,8	27,6	30,2	33,4	35,7	40,8
18	6,26	7,01	8,23	9,39	10,9	12,9	22,8	26,0	28,9	31,5	34,8	37,2	42,3
19	6,84	7,63	8,91	10,1	11,7	13,7	23,9	27,2	30,1	32,9	36,2	38,6	43,8
20	7,43	8,26	9,59	10,9	12,4	14,6	25,0	28,4	31,4	34,2	37,6	40,0	45,3

ПРИЛОЖЕНИЕ 4
Справочное

Обработка данных, полученных при определении плотности топлива ТС-1 по ГОСТ 3900—47

Номер выборки	Результаты определения плотности при 20°С ρ, кг/м³	Среднее арифметическое результатов испытаний $\overline{x_i}$	Дисперсия S_i^2	Среднее взвешенное $\overline{\overline{x}}$	Вычисленное значение A^2_{min}	Коэффициент Берджа $R_Б$	Пределы значения плотности при 20°С ρ, кг/м³	Требования по ГОСТ 3900—47
1.	778,0; 778,0; 778,0; 778,0, 779,0; 780,0; 779,7; 778,0; 779,0, 779,0, 779,0, 780,0	778,18	0,64	779,79	2,46	0,33	778,4—779,9	Не менее 775,0
2.	779,5; 779,8; 779,6; 781,0, 781,0, 781,0, 781,0, 781,0, 780,0, 780,0, 781,0, 780,0	780,41	0,40				780,1—780,7	

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

Оценка достоверности численных данных о свойствах нефти и нефтепродуктов

МИ 656—84

Редактор *Т. Ф. Писарева*
Технический редактор *Н. В. Белякова*
Корректор *С. И. Ковалева*

Н/К

Сдано в наб. 22.04.85 Подп. в печ. 19.11.85 Т—19714 Формат 60×90/16 Бумага типографская литературная Печать высокая 0,75 усл. п. л. 0,75 усл. кр.-отт. 0,48 уч.-изд. л. Тир. 2.000 Зак. 2119 Цена 3 коп. Изд. № 8659/4

Ордена «Знак Почета» Издательство стандартов, 123840, Москва, ГСП, Новопресненский пер., д. 3.
Вильнюсская типография Издательства стандартов, ул. Миндаугo, 12/14.

gost.e-ecolog.ru