ISSN 0965-5441, Petroleum Chemistry, 2015, Vol. 55, No. 8, pp. 618–631. © Pleiades Publishing, Ltd., 2015.
Original Russian Text © M.B. Smirnov, N.A. Vanyukova, 2015, published in Neftekhimiya, 2015, Vol. 55, No. 6, pp. 473–486.

Distribution Patterns of Main Structural-Group Parameters of Crude Oils from the Volga–Urals Oil and Gas Basin According to NMR Data

M. B. Smirnov and N. A. Vanyukova

Topchiev Institute of Petrochemical Synthesis. Russian Academy of Sciences, Moscow, Russia

e-mail: m1952s@yandex.ru

Received September 18, 2014

Abstract—The character of the distribution of 16 main NMR-measurable parameters of the structural-group composition of Volga–Urals crude oils has been assessed. It has been shown that the distributions quite significantly differ from the normal distribution; they are in general asymmetrical and essentially different. Total aromaticity and the amount of *n*-alkyl structures in oils of the major stratigraphic plays are on average the same, although there are differences between the northern, central, and southern parts of the oil-and-gas basin. Regarding the concentration of unsaturated hydrocarbons, there are differences between both the plays and territories. Similarities and dissimilarities between the West Siberian and Volga–Urals basins in distributions have been revealed. Some features observed for Western Siberia oils have to be other in the Volga–Urals basin. Upper Carboniferous and Lower Permian carbonate reservoirs in the northeast of the basin have been found to contain oils with a uniquely high concentration of aromatics that evidently differ in composition from those normally present in crude oils. The advantages of interval plots of probability density over histograms for comparing the distributions have been demonstrated. Using the particular material, it has been shown that the Student *t*-test conventionally used to compare average values for oil groups leads to false conclusions. To determine systematic shifts between distributions, it is necessary to use the Mann–Whitney test and compare the medians of distributions.

Keywords: crude oil, statistical distribution, ^{1}H NMR, ^{13}C NMR, Volga–Urals oil-and-gas basin

DOI: 10.1134/S0965544115060122

A paradoxical situation has developed in petroleum chemistry in the last decades. A tremendous number of papers presenting data on the elemental, group, structural-group, and (first of all) individual compositions of a great variety of oils from almost all of the world oil-and-gas basins have been published. However, information on the frequency of occurrence of oils with a particular amount of the main elements is quantitatively given only by the data on elemental composition [1]. Regarding the other parameters, only ranges of possible or "most common" values for oils in general or their particular groups and, sometimes, estimated average values (e.g., see [1–4]) have been reported in the literature. At the same time, the knowledge of what is the probability that any quantity characterizing the composition falls in a certain range of values (i.e., distribution of the quantity) is of crucial importance at least for the correct use of methods of applied statistics in data processing [5–7]. This knowledge is also necessary for studying oil classification, modeling oil generation processes, and development of methods for analyzing the crude oil composition (e.g., see [8–11]).

In previous studies [12, 13], we considered the distribution patterns and the degree of interrelation of 14 main and most easily measured NMR characteristics of structural-group composition of Western Siberia crude oils, as well as equations describing the relation between pairs of most correlated quantities. The natural next step is to reveal which features are general in character and which reflect regional specifics, an objective that requires a similar study of oils from other basins. The first of the selected basins is the Volga–Urals oil-and-gas basin (OGB), which is adjacent to the Western Siberia basin geographically but radically differs in geological structure (see [14–18] and references therein).

In this work, we studied distributions of the same characteristics of the composition as in [12, 13], plus two other quantities derived from directly measured parameters.

EXPERIMENTAL

The majority of the test specimens were oil samples from the Topchiev Institute collection gathered in the 1960–1980s. By the storage conditions and the sample

preparation procedure adopted in connection with this [12], the total C_{8+} or C_{9+} fractions should be considered an object of study by structural-group analysis. The appropriateness of this restriction was substantiated in [12].

NMR spectra of oil solutions in $CDCl_3$ (~1 : 1 v/v) were recorded at 313 K on a Bruker DRX-400 spectrometer (400 MHz for ^1H and 100.6 MHz for ^{13}C) with a 5-mm detector without sample spinning. The recording conditions, the mode of reference for chemical shifts, and their processing were the same as previously [12].

RESULTS AND DISCUSSION

The objects of study were 167 samples of oil from 96 fields (Table 1). The samples represent almost all of the main productive plays; terrigenous Middle and Upper Devonian, carbonate Upper Devonian and Lower Carboniferous Tournasian, terrigenous Lower Carboniferous, carbonate Middle Carboniferous, and carbonate Upper Carboniferous and Lower Permian [18]. Oilfields and exploration areas located in the territory of Udmurtia; Bashkortostan; Tatarstan; and Perm, Samara, Saratov, Volgograd, and Orenburg oblasts are covered, as well as 18 tectonic structures [18]: Ablulinskii trough, Almet'evsk ridge, Babkinskaya saddle, Bashkir ridge, Belebey-Shkapovskaya ridge, Bel'skaya depression, Birsk saddle, Upper Kama trough, Zhiguli–Orenburg arch, Kos'vino-Chusovaya saddle, Nizhnekamsk linear dislocation system, Perm ridge, Sarailinskaya saddle, Solikamsk depression, Chermozskaya saddle, south-eastern part of Pachelma trough, south-eastern slope of the platform, and Yuryuzan–Sylvenskaya depression. The occurrence depth and reservoir temperature ranged from 370 to 4700 m and from 18 to 118°C, respectively.

The sample is not free of obvious defects. First, there are irrelevantly high representation of the Lower Carboniferous terrigenous play and little samples from Devonian reservoirs. Second, there are a higher proportion of oils from Perm oblast and a lower proportion from the south of the basin, especially, Volgograd, Saratov, and Orenburg oblasts [18]. Finally, the section of the productive stratum in Orenburg oblast is mainly represented by samples from one multibed oilfield. Accordingly, in the final analysis, it is necessary to assess in which way and to what extent this imperfectness of the sample will affect the results.

The following structural-group parameters (percentage of total hydrogen or total carbon of the sample, respectively, H_{db} ×100100) were measured for all of the samples:

H_γ – hydrogen of CH_3 groups separated by at least three C–C bonds from aromatic rings, carbonyl groups, or heteroatoms;

H_β – hydrogen of $-CH_2-$ and $-CH<$ groups, in β- and more distant positions with respect to the same structural units;

H_α – hydrogen of CH_3-, $-CH_2-$, and $-CH<$ groups in α-position to the same;

H_{ar} – hydrogen in the aromatic ring with division into two ranges:

$H_{ar,\,1}$ – hydrogen mainly in monoaromatic structures [12],

$H_{ar,\,2+}$ – hydrogen predominantly in bi- or polycyclic aromatic structures (i.e., $H_{ar} = H_{ar,\,1} + H_{ar,\,2+}$ +);

H_{db} – hydrogen at isolated double bonds;

C_{ar} – carbon in aromatic rings;

n(1)-C4 – carbon in terminal methyl groups of *n*-alkyl structures composed of four C atoms C*CCC (relevant C-atom hereinafter is marked with asterisk *);

n(3)-C6 – carbon C-3 in the structure of the CCC*CCC;

n(5+) – carbon in CH_2 groups of chains, beginning from C-5, separated from any substituent by at least three C–C bonds (–CCC–C*–CCC–).

In addition, the quantities $H_{ar,\,2+}/H_{ar,\,1}$, H_α/H_{ar}, Σ*n*-Alk (total carbon fraction in *n*-alkyl chain fragments; calculation procedure is described in [12]), Σ*n*-Alk/*n*(1)-C4, and *n*(3)-C6/*n*(1)-C4 were calculated. The last two quantities are added as compared to the set considered for Western Siberia oils. The first of them gives an idea on the average size of the alkyl chains, and the second shows how fast the amount of *n*-alkyl structures falls with an increase in their size.

For all of the 16 parameters, the average characteristics of distributions: mean, median, and standard root-mean square deviation (RMSD) were calculated (Table 2). The values of criteria to distinguish a distribution from the normal distribution: ratio of the average unit value of the deviation from the mean to RMSD, eccentricity, and excess kurtosis [5, 7] were calculated. The distribution differs from the normal one with a statistical significance of >0.99 if at least one of the criteria goes outside the following boundaries: first, 0.758–0.837; second, (–0.486)–(+0.486); and third, –0.70–1.14. From Table 2 it follows that none of the parameters obeys the normal distribution law; the least difference is for H_α and *n*(3)-C6/*n*(1)-C4. Most distributions strongly asymmetrical, half of them have an excess kurtosis of ≥8; i.e., there are samples outlying by these parameters. Consequently, the RMSD in this case is not an unambiguous measure of scatter of data points relative to the mean values. Therefore, bounds within which 50%, 80%, and 90% values are retained were defined (Table 2). To visualize the spread relative to the median, the ratios and differences of the values at the boundaries of 50% and 80% ranges are also given in the table. Further analysis (see below) showed that among the test samples there are six unique samples (five, in the aromatics content and one in the amount of *n*-alkyl structures), which are

Table 1. Oilfields of the Volga–Ural oil-and-gas basin represented in the sample

No.	Field, area	Number of samples	Age (beds)	No.	Field, area	Number of samples	Age (beds)
1	2	3	4	1	2	3	4
Udmurtia (12, 9, 840–1757 m, 18–39°C)*							
1	Vyatskoe	1	C_{1jp}	6	Mishkinskoe	2	C_{1jp}, C_{2b}
2	Kyrykmasskoe’	2	C_{1jp}, C_{2pd}	7	Oblastanovskoe	1	D_{3psh}
3	Tarasovskaya area	1	C_{1jp}	8	Chutyrsko-Kiengopskoe	1	C_{2b}
4	Gremikhinskoe	2	C_{1tl}, C_{2b}	9	Arkhangel’skoe	1	D_{3psh}
5	Listvennoe	1	C_{1jp}				
Perm oblast (55, 29, 370–2260 m, 18–38°C)*							
10	Bardymskoe	1	D_{3fr}	25	Kozubaevskoe	1	C_{1tl}
11	Batyrbaiskoe	2	C_{1bb}, C_{2b}	26	Krasnokamskoe	2	D_{3kn}, C_{2b}
12	Kokuiskoe	1	C_{2b}	27	Osinskoe	1	C_{2b+nm}
13	Kuliginskoe	1	C_{1tl}	28	Shemetinskoe	3	C_{1t}, C_{1tl}, C_{2b}
14	Kamenskoe	1	C_{1tl}	29	Yarino-Kamenolozhskoe	11	C_{1t}, C_{1jp}, C_{2b}
15	Karamorskoe	1	D_{3psh}	30	Bel’skoe	1	C_{2b}
16	Krasnoyarsko-Kuedinskoe	4	D_{3psh}, C_{1tl}, C_{2b}, P_{1s}	31	Gezhskoe	3	D_{3fr}, C_{1t}, C_{1ok}
17	Kylasovskoe	2	C_{1bb}, C_{2vr}	32	Peschankovskoe	1	C_{2b}
18	Pavlovskoe	4	C_{1bb}, C_{1tl}, C_{2b}, C_{2vr}	33	Chashkinskoe	1	C_{1bb}
19	Shagirtskoe	2	C_{1tl}, C_{2vr}	34	Yaborovskoe	1	C_{2vr}
20	Gozhanskoe	1	C_{1bb}	35	Rusakovskoe	1	C_{1tl}
21	Nozhovskoe	1	C_{1jp}	36	Brusyanskoe	1	C_{1jp}
22	Tihovskaya area	3	C_{1jp}, P_1, P_{1ar}	37	Kopal’ninskoe	1	P_{1s+a}
23	Baklanovskaya	1	C_{1jp}	38	Luzhivskoe	1	C_{1jp}
24	Gorskoe	1	C_{1bb}				
Bashkiria (28, 19, 710–2506 m, 17–48°C)*							
39	Mustafinskoe	1	C_{1ksl}	49	Arslanovskoe	1	C_2
40	Buraevskoe	1	C_{2b}	50	Beketovskaya area	2	$D_{3psh+kn}$, C_{1t}
41	Teplyakovskoe	1	D_{3fm}	51	Iskanderovskoe	1	C_{1t}
42	Ust’-Aiskoe 2	2	D_{3dm}, C_{1bb-rd}	52	Severo-Urshakskoe 2	2	D_{2bs}, D_{3fm}
		3				3	4
43	Shav”yadinskoe	1	D_{3fm}	53	Sergeevskoe	1	$D_{3psh+kn}$
44	Shkapovskoe	1	D_{2vr}	54	Tavtimanovskoe	3	D_{2ml}, $D_{3psh+kn}$, C_1
45	Vvedenovskoe	1	P_{1ar-a}	55	Kyzylbaevskoe	4	D_{3dm}, C_{2b}, C_{2vr}, C_3
46	Kinzebulatovskoe	1	P_{1ar-a}	56	Muslyumovskoe	2	C_{2vr}, P_{1s}
47	Arlanskoe	1	C_{2m}	57	Yanybaevskoe	1	C_3
48	Mancharovskoe	1	C_{1bb}				
Tatarstan (24, 13, 810–1950 m, 20–39°C)*							
58	Aksubaevo-Mokshinskoe	2	C_{1bb}, C_1	65	Pionerskoe	2	C_{1bb}, C_{1tl}
59	Berket-Klyuchevskoe	1	C_{1bb}	66	Romashkino	7	D_{2vr}, D_{2ar}, D_{3psh}, C_{1ksl}, C_{1bb}
60	Enoruskinskoe	1	C_{1bb}	67	Uratminskoe	3	C_{1bb}, C_{1tl}, C_2
61	Ersubaikinskoe	1	C_1	68	Aktanyshskoe	1	C_{1bb}
62	Ivashkino-Malosul’chinskoe	1	C_{1bb}	69	Pervomayskoe	1	C_{1el}
63	Novo-Elkhovskoe	2	D	70	Bastrykskoe	1	C_{1t}
64	Novosheshminskoe	1	C_1				

Table 1. (Contd.)

No.	Field, area	Number of samples	Age (beds)	No.	Field, area	Number of samples	Age (beds)
1	2	3	4	1	2	3	4
Samara oblast (22, 14, 540–3270 m, 18–74°C)*							
71	Kanashskoe	1	C_{1ksl}	78	Kuleshovskaya	1	C_{2b}
72	Barinovskoe	2	D_{2zv} (bed D-3), C_{1ksl}	79	Novoklyuchevskoe	1	P_{1ku}
73	Bogatyrevskaya area	2	C1t (bed B-1, B-2)	80	Pokrovskoe	4	D_{3d-l}, C_{1up}, C_{1bb}, C_{1ok}
74	Verkhne-Vetlyanskoe	1	C_{1bb}	81	Syzranskoe	1	C_{2vr}
75	Dzerzhinskoe	1	C_{2b}	82	Utevskoe	3	C_{1ksl}, C_{1bb}, C_{1ok}
76	Zol'nenskoe	1	C_{1ksl}	83	Khilkovskoe	1	D_{3fm}
77	Il'menevskoe	2	C_{1bb}, C_{2b}	84	Yablonovyi Ovrag	1	D_3
Orenburg oblast (18, 5, 800–3900 m, 20–68°C)*							
85	Buguruslanskoe	1	C_1	88	Baytuganskoe	2	D_{3psh}, D_{3f2-3} (bed D_1, D_{kt1})
86	Baytuganskoe	1	C_{2b}	89	Rybkinskoe	1	D_{3f2-3}
87	Bobrovskoe	13	C_{1t}, C_{1bb}, C_{1ok}, C_{2b} (bed T_1, O_4, O_3, O_2, O_1, A_4)				
1	2	3	4	1	2	3	4
Saratov oblast (5, 4, ~800–~2200 m, 25–62°C)*							
90	Elshanskoe	1	C_{1vbb}	92	Sokolovogorskoe	2	D_{2vr}, D_{3psh}
91	Sovetskoe	1	D_{2ar}	93	Stepnovskoe	1	D_{2ar}
Volgograd oblast (3, 3, 2340–43700 m, 99–118°C)*							
94	Ol'khovskoe	1	D_{2vr}	96	Yuzhno-Kamyshlinskoe	1	D_{3sr}
95	Tersinskoe	1	D_{2ar}				

* /Figures in the parentheses successively present the number of samples examined for the region, the number of fields and prospecting areas, the occurrence depth interval, and the range of reservoir temperatures represented in the sample.

better to be considered separately, rather than combined with the others (hereinafter called the main part of oils). Accordingly, all the characteristics in Table 2 were calculated for the array of samples without the unique oils. Values that significantly differ from those calculated from the full sample are shown in Table 2 in parentheses.

The distributions of the values of 13 parameters for the main part of the oils also differ from the normal distribution with a statistical significance of >0.99, although there are no such large deviations of the parameters from the critical values. For three parameters (H_γ, Σn-Alk/n(1)-C4, and n(3)-C6/n(1)-C4, the statistical significance ranges between 0.95 and 0.99. Thus, the distributions remain non-Gaussian even when the sample does not include the unique oils. Consequently, the usual criteria given in all textbooks (such as Student's t-test used in [20]) cannot be applied when comparing different groups of oils (e.g., oils from different basins). It is necessary to use nonparametric statistical methods. Thus, instead of comparing the means, the medians of the distributions should be matched and the sign test should be used to assess the significance of differences [5, 7]. The grouping problem should also be addressed in another way [6].

From the data in Table 2 it follows that the array under study includes samples that greatly differ in composition of the main components. The total proportion of carbon in n-alkyl moieties varies from 6 to 47 abs. %, with the extreme values for polyatomic structures (parameter n(5+)) differing more than by a factor of 10. Aromaticity (C_{ar}) of the samples varies over a very wide range from 6 to 51 abs. %; the highest

Table 2. Numerical characteristics of the distribution of composition parameters of Volga–Ural crude oils[a]

		H_{ar}	H_{db}	H_{α}	H_{β}	H_{γ}	$H_{ar,\,2+}$	$H_{ar,\,1}$	$H_{ar,\,2+}/H_{ar,\,1}$	H_{α}/H_{ar}	C_{ar}, %	*n*(5+)	*n*(3)-C6	*n*(1)-C4	Σ*n*-Alk	Σ*n*-Alk/*n*(1)-C4	*n*(3)-C6/*n*(1)-C4
1		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Mean		4.23 (3.96)	3.1	9.8	59.4	26.5 (26.8)	2.21 (2.05)	2.01 (1.90)	1.10	2.40	16.8 (16.3)	8.1	1.99	2.58	17.0	6.59	0.76
Median		4.10	1.7	10.4	59.0	26.6	2.09	1.96	1.08	2.36	17.0	7.8	1.87	2.46	16.5	6.49	0.76
RMSD		1.9 (0.67)	3.6	2.7	3.6 (2.8)	2.4 (1.9)	1.12 (0.39)	0.81 (0.33)	0.18	0.43 (0.36)	5.0 (3.0)	2.8 (2.0)	0.65	0.71	5.1 (4.5)	0.75 (0.59)	0.06
Eccentricity		5.7 (−0.54)	1.9	−0.11	1.05 (0.63)	−0.85 (0.48)	5.8 (−0.19)	5.2 (−0.50)	1.09 (0.75)	−0.75 (0.04)	3.1 (−0.61)	4.2 (0.55)	1.5	1.14 (1.33)	1.9 (1.05)	3.1 (0.14)	−0.08 (−0.42)
Excess kurtosis		39.5 (0.66)	2.8 (2.15)	−0.31 (−0.67)	8.2 (0.04)	3.8 (0.18)	40.1 (1.27)	35.2 (−0.24)	2.8 (1.68)	1.35 (−0.65)	17.6 (0.22)	33.4 (0.96)	5.2 (6.0)	3.4	8.5 (3.0)	21.9 (−0.09)	1.67 (1.06)
Av. mod./RMSD[b]		0.42 (0.79)	0.72	0.84 (0.86)	0.70 (0.79)	0.73 (0.79)	0.43 (0.76)	0.48 (0.82)	0.75	0.79 (0.85)	0.60 (0.81)	0.63 (0.78)	0.74	0.76	0.71 (0.76)	0.67 (0.77)	0.76
50% bounds	From (a)	3.54	0.7	7.57	57.3	25.3	1.84	1.69	0.98	2.14	14.1	6.4	1.58	2.12	13.7	6.18	0.72
	To (b)	4.43	3.7	11.9	61.2	27.7	2.33	2.17	1.18	2.72	18.6	9.0	2.32	3.04	19.6	6.96	0.80
80% bounds	From (c)	3.16	0.3	6.41	56.0	23.9	1.56	1.44	0.92	1.97	12.1	5.8	1.30	1.79	12.1	5.91	0.70
	To (d)	4.76	7.6	12.80	63.8	29.3	2.54	2.31	1.32	2.95	19.7	10.9	2.64	3.39	22.0	7.34	0.83
90% bounds	From	2.76	0.0	5.25	55.5	23.3	1.47	1.28	0.86	1.72 (1.90)	10.9	5.5	1.19	1.63	11.3	5.57	0.67
	To	5.21 (4.83)	12.4	13.4	64.6	30.2	2.80 (2.61)	2.48 (2.35)	1.40	3.04	21.3	11.4	3.06	3.62	24.6	7.67	0.85
Least		1.42 (1.77)	<0.2	2.86 (3.40)	44.8 (53.5)	15.2 (17.7)	0.69	0.49 (0.98)	0.64	0.84 (1.52)	6.1 (7.0)	2.72	0.61	0.86 (1.18)	6.1	5.1	0.50
Greatest		20.37 (5.77)	15.3	17.2 (16.7)	80.5 (68.1)	33.0	11.61 (3.31)	8.76 (2.64)	1.92 (1.56)	3.26	51.0 (23.6)	32.3 (15.0)	5.52	6.07	47.5 (37.9)	12.5 (8.2)	1.00 (0.91)
Ratios	b/a	1.25	5.5	1.57	1.07	1.09	1.27	1.29	1.21	1.27	1.32	1.40	1.47	1.44	1.43	1.13	1.11
	d/c	1.51	25.4	2.00	1.14	1.23	1.63	1.61	1.44	1.50	1.62	1.89	2.04	1.90	1.82	1.24	1.20
Differences	b–a	0.89	3.1	4.34	3.8	2.3	0.49	0.48	0.20	0.58	4.52	2.55	0.74	0.93	5.9	0.78	0.08
	d–c	1.60	7.3	6.40	7.8	5.4	0.99	0.87	0.40	0.98	7.55	5.15	1.35	1.61	9.9	1.43	0.14

[a]/ The figures in parentheses are the values for the sample without six unique specimens (see text). All quantities expressed in percentage (mean, median, RMSD, etc.) refer to total hydrogen or total carbon of the sample, respectively; for H_{db}, the same ×100.

[b]/ Ratio of the average module of deviation from the mean to the standard deviation [5, 7].

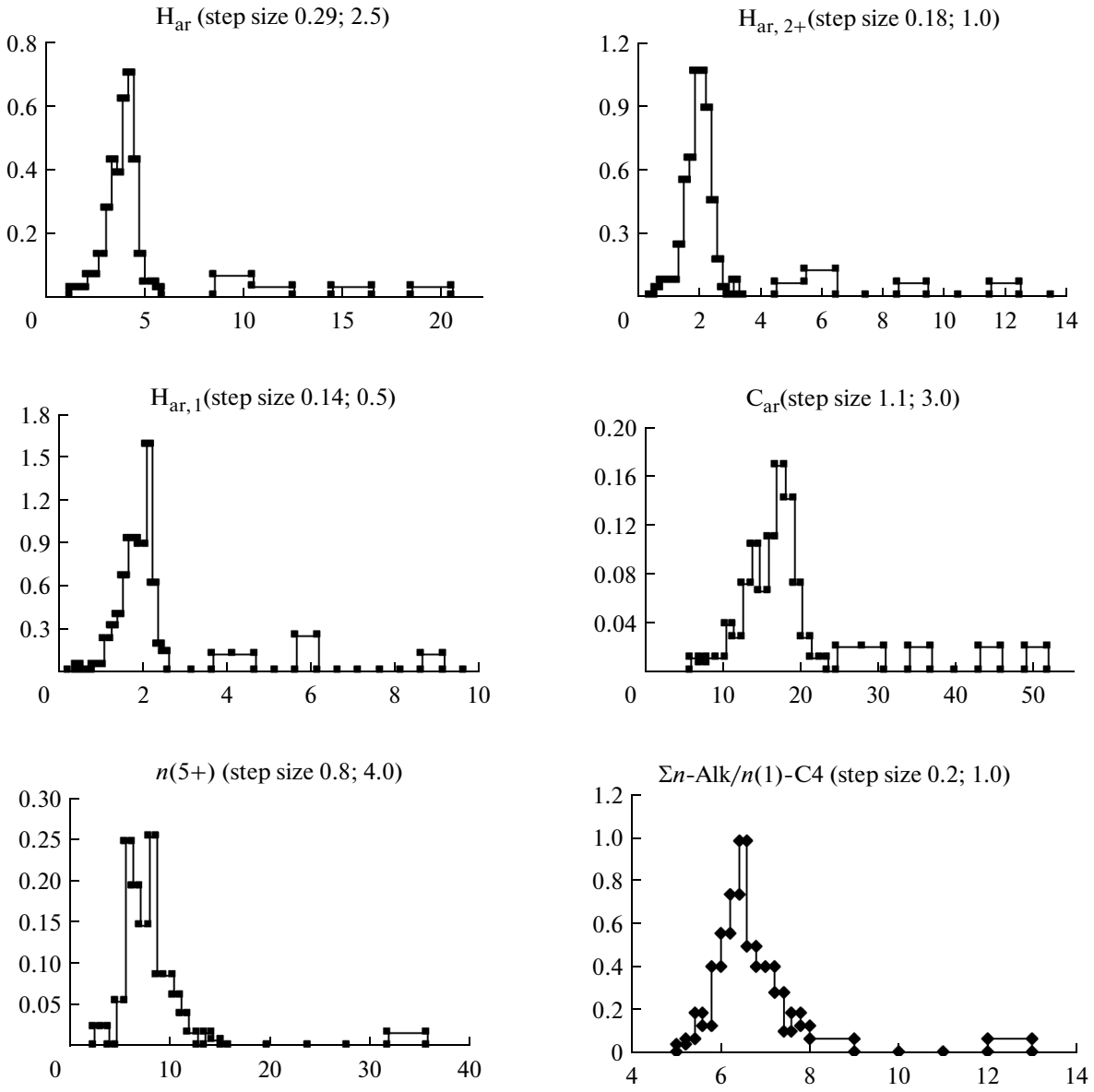

Fig. 1. Distribution density of composition parameters for the case when there are samples that dramatically stand out. The first in parentheses is the step size on the abscissa horizontal in the domain of values characteristic of the main part of the sample, and the second is the size in the domain of large values of the parameters (in %). For the latter, the probability density is multiplied by 10.

values are at least twice greater than those determined previously for oils and they are typical of polycycloaromatic fractions, resins, and asphaltenes [12, 19–21]. The ratio between the proportions of hydrogen in bi- + polycyclic aromatic and monoaromatic rings vary in extreme cases is about 3, and the average size of *n*-alkyl chains (Σ*n*-Alk/*n*(1)-C4) differs by a factor of 2.5. The proportions of hydrogen at isolated double bonds vary in the range of almost two orders of magnitude. The smallest differences between the oils are observed in parameter H_β, as in the case of Western Siberia[12], and somewhat larger differences are found in H_γ. Compared to Western Siberia, the ratios of the values on the boundaries of the intervals incorporating half of the oils are smaller for the parameters that characterize the aromatic components (H_{ar}, $H_{ar, 2+}$, etc. with the exception of H_α) and are approximately the same for the characteristics of *n*-alkyl structures.

None of the standard versions with either uniform or logarithmic scale proved suitable for detailed representation of the form of distributions. The failure is caused by the presence of samples with parameters that greatly differ in value from those of the majority of the samples. With the given number of samples, it is incorrectly to divide the entire range of values by more

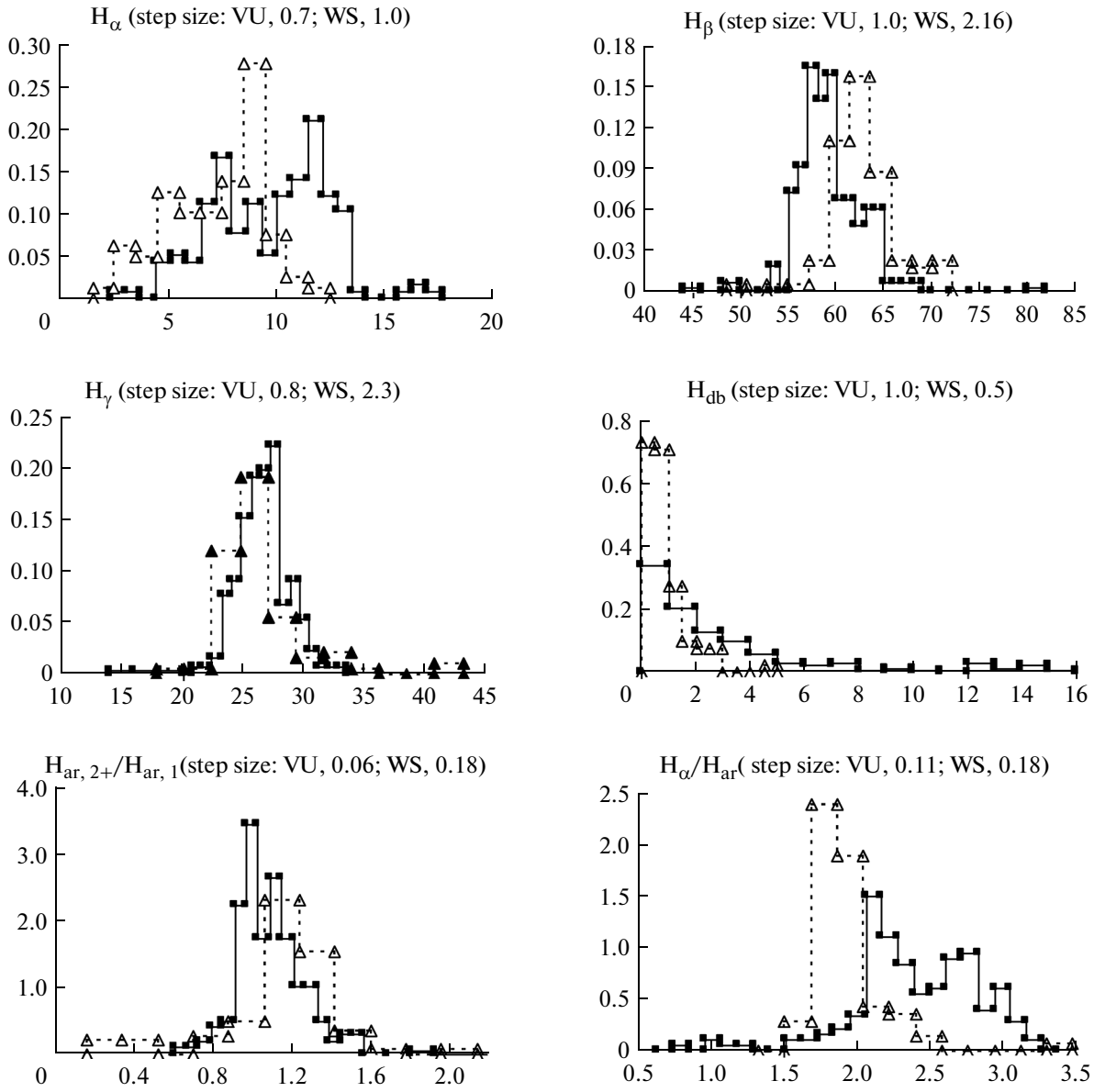

Fig. 2. Distribution density of composition parameters measured from ^{1}H NMR spectra (except those shown in Fig. 1). Step size: VU is for Volga–Urals, WS is fro Western Siberia. Filled squares and the solid line refer to Volga–Ural oils, and open triangles and the dashed line refer to West Siberian oils.

than 20 intervals [6, 7]. For example, such a division for H_{ar} (Fig. 1) gives a step of 1 abs. % on the uniform scale ("natural boundary" from 1 to 21 abs. %). Then the whole main domain of values (~ 1.5–5.5 abs. %) is divided into only five intervals, giving a very general picture of the distribution. Therefore, instead of the conventional histograms, we used plots showing the so-called probability density function p_{int} [6, 7, 22]. The value of p_{int} for each interval of values of the parameter is calculated by the formula

$$p_{int} = n_{int}/(d_{int} \times N), \quad (1)$$

where n_{int} is the number of samples with values of the parameter within the range under consideration, d_{int} is the width of the interval, and N is the total number of samples in the set. The plots are broken lines (Fig. 1) showing explicitly the width of each interval, which are not necessarily identical here, with the area under it being equal to 1. To obtain the usual quantities, the percentage of samples belonging to a given interval p_{int} is multiplied by the width of the interval and by 100. For example, H_{ar} (Fig. 1) in the interval with the highest probability density of $p_{int} \approx 0.7$, the step size is 0.29 and $0.7 \times 0.29 \times 100 \approx 20\%$ of the samples gets into this interval. Another advantage of this representation is that it allows direct comparison of distributions for groups with different numbers of samples when,

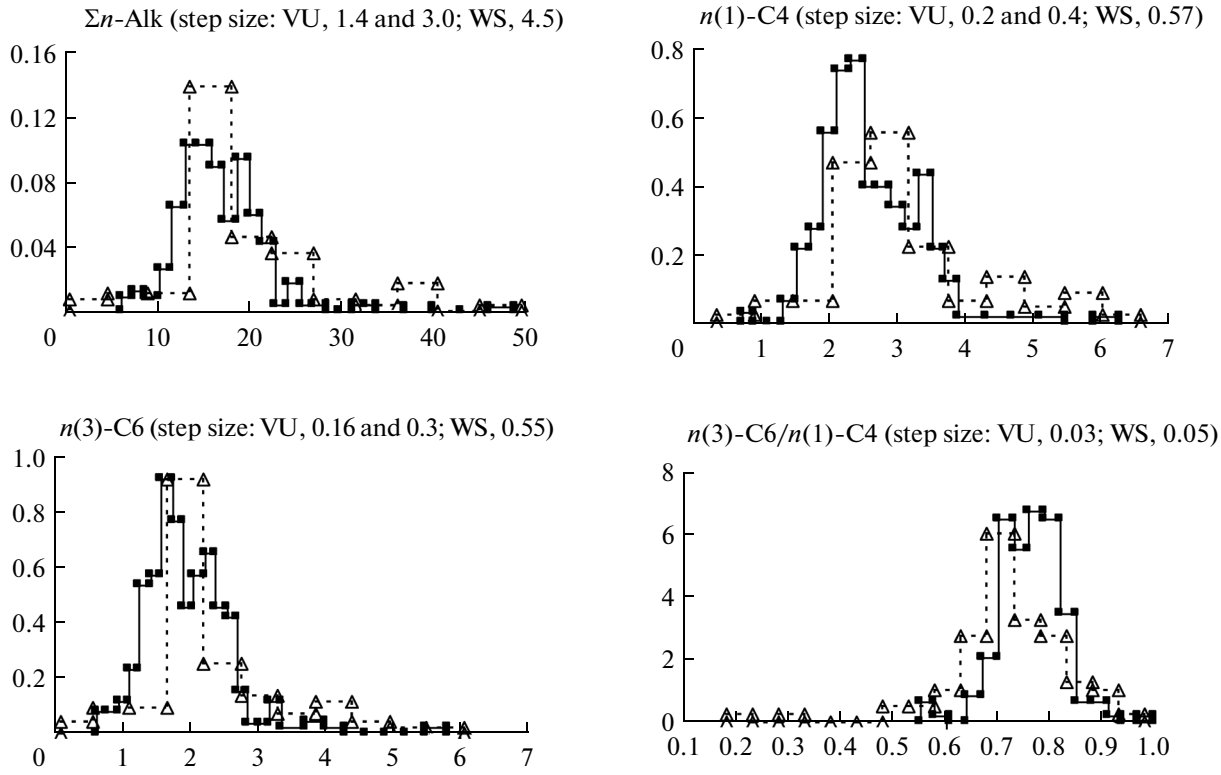

Fig. 3. Distribution density of parameters describing *n*-alkyl structures (except those shown in Fig. 1). The notation is the same as in Fig. The step size is 3.0 abs. % for Σ*n*-Alk in intervals of values greater than 33, 0.3 abs. % for *n*(3)-C6 on values of >3.3, and 0.4 abs. % for *n*(1)-C4 on values of >3.9.

accordingly, the entire range of values was divided into intervals in different ways (Figs. 2–4). The results are shown in Figs. 1–4.

The first thing that catches the eye is five oils with an unusually high aromatics content, which distinguishes them by parameters H_{ar}, C_{ar}, $H_{ar,2+}$, and $H_{ar,1}$ (Fig. 1). Compared with the most common oils, they have two to five times greater values of H_{ar} (average value at the distribution maximum is ~4.3 abs. %). All of them are from the Upper Carboniferous and Lower Permian carbonate play in the northeast of the basin. One sample (Ol'khovskoe field) outstands with an extremely high amount of *n*-alkyl structures, primarily in terms of parameters *n*(5+) and Σ*n*-Alk/*n*(1)-C4. In general, the majority of distributions have a complex pattern with two to three maximums and evident asymmetry. Frequently, there are rather large "wings" at the principal maximum (H_β; H_γ; H_{ar}, $H_{ap,2+}$, *n*(3)-C6, Σ*n*-Alk/*n*(1)-C4, *n*(3)-C6/*n*(1)-C4; and, at some stretch, $H_{ar,2+}/H_{ar,1}$) (Figs. 2–4). These features suggest the presence of a considerable number of factors that have comparable effects on the oil composition. The distribution of H_{db} is quite different: a decline with increasing values and an increase in the density in the region of 0.12–0.16 abs. % (0.1 abs. % H_{db} corresponds to ~3.4 abs. % unsaturated hydrocarbons [23, 24]). The simplest distribution pattern is for the *n*(3)-C6/*n*(1)-C4 ratio.

It was explored what features of these distributions may not be artifacts due to the above sample imperfections.

Corrections for the excessive number of samples from the section of Bobrovskoe oilfield lead to the following changes in Figs. 1–4. The C_{ar} distribution maximum at 12.6–14.8 abs.% remains as a shoulder, the maximum of H_α in the range of 6.55–7.95 abs. % decreases by half against the background, the same is true for Σ*n*-Alk (maximum at 18.6–20.0 abs. %), and the density at maximum at 8.1–8.9 abs. % for *n*(5+) drops from 0.25 to 0.20. The density is reduced over the entire range of 3.1 to 3.9 abs. % for *n*(1)-C4 (from 0.43 to 0.33 at maximum). For *n*(3)-C4, the maximum shifts to the neighboring interval (2.38–2.54 abs. %) and narrows into one interval. A slight decrease in density is observed for H_{ar} in shoulder at 3.02–3.89 abs. %, H_β at the maximum (57–59 abs. %), H_γ in the region of 28–32 abs. %, and H_α/H_{ar} at the principal maximum (2.06–2.28). The main peak of the $H_{ar,2+}$ distribution narrows because of the fall in density in the range of 1.53–1.71 abs. %.

The second type of sample imperfection is an uneven representation of samples from different productive plays and parts of the oil-and-gas basin. To

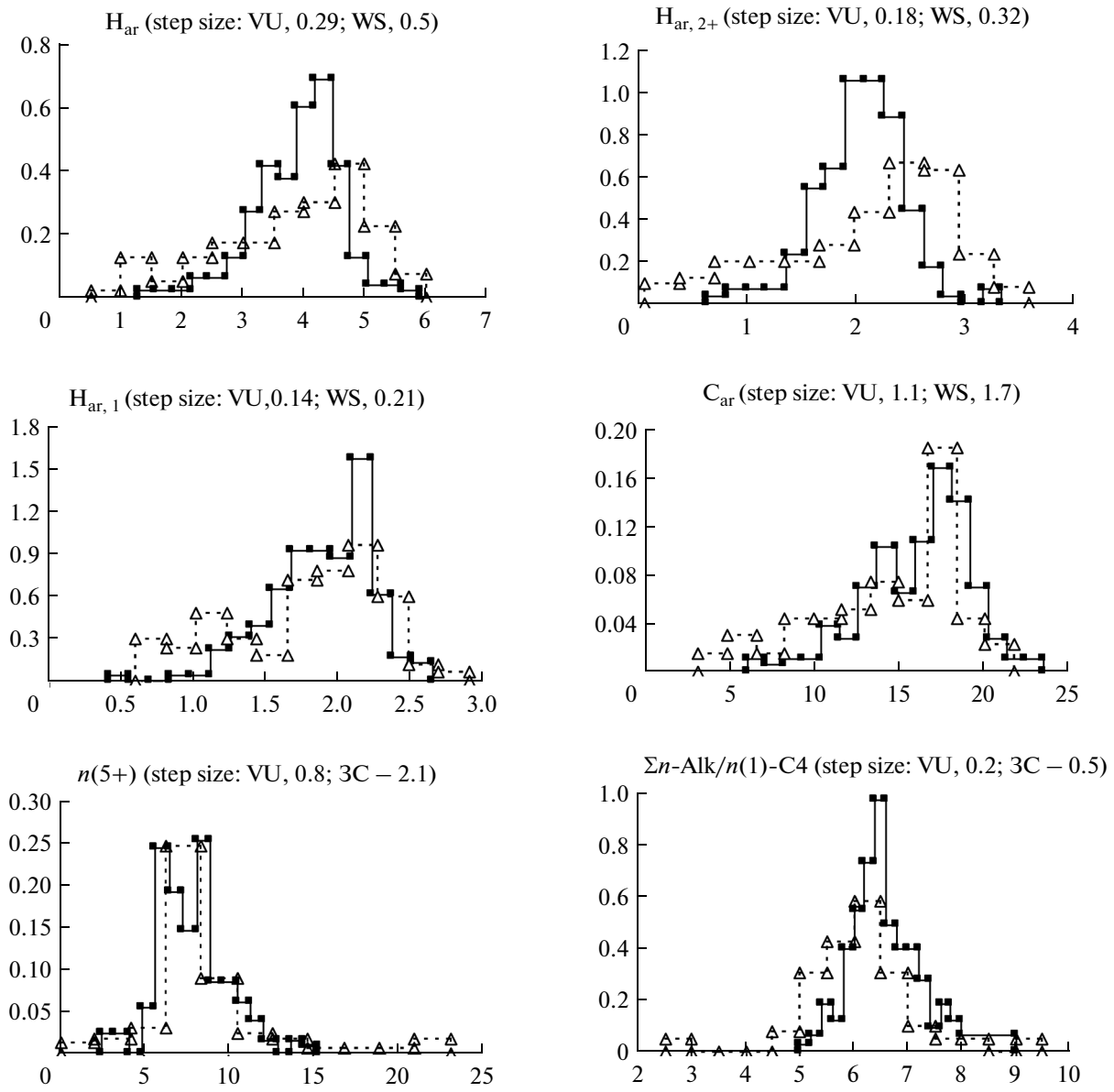

Fig. 4. Distribution density of parameters, shown in Fig. 1, for the main part of the samples. The notation is the same as in Fig. 2.

assess the possible consequences of this disproportion, we determined the existence, direction, and magnitude of a systematic shift of the distributions of each parameter between the corresponding oil groups. The Mann–Whitney U rank test [6, 7] was used, which makes it possible to determine probability P that the observed differences between two arbitrary distributions are accidental and which of the distributions is shifted towards lower values in comparison with the other. That is, we may assume the presence of a systematic shift with a probability $1 - P$ at low values of P (say, less than 0.05 or 0.01) and its absence at large values (for example, more than 0.3 or 0.5). The lower the value of P, the more reliable the conclusions about the presence of the shift and the greater the shift. Obviously, establishing the presence or absence of systematic shifts between oils from different producing plays and different parts of OGB also has a significant intrinsic value.

In the analysis of oils from different producing plays, the following pairs were considered: all terrigenous–all carbonate (T–C in Table 3); Lower Carboniferous terrigenous–Middle Carboniferous carbonate (TC_1–KC_2 ibid.); Lower Carboniferous terrigenous–Upper Devonian and Lower Carboniferous carbonate (TC_1–$K(C_1+D_3)$, as well as with its part, carbonates of the Lower Carboniferous only (TC_1–KC_1); Devonian terrigenous–Devonian carbonate (total each, TD–KD); Devonian terrigenous (total)–Lower Carboniferous terrigenous (TD–TC_1); Middle Carbonif-

Table 3. Comparison of the distributions of composition parameters of crude oils from the main producing plays of the Volga–Urals OGB (without six unique oils, see text) by the Mann–Whitney test

Parameter	Compared producing plays*						
	T–K	TC_1–KC_2	TC_1–$K(C_1+D_3)$	TC_1–KC_1	TD–KD	TD–TC_1	KC_2–$K(C_1+D_3)$
H_{ar}	2353	870	772	536	90	625	380
H_{db}	**1725**	823	**342**	**241**	**46**	554	**202**
H_{α}	2577	944	780	504	85	605	411
H_{β}	2517	953	741	523	75	660	391
H_{γ}	2540	926	746	509	91	576	392
$H_{ar,\ 2+}$	2334	862	745	499	97	637	411
$H_{ar,\ 1}$	2420	918	724	480	84	613	382
$H_{ar,\ 2+}/H_{ar,\ 1}$	2602	906	716	452	71	559	421
$H_{\alpha}//H_{ar}$	2623	886	735	538	68	617	349
C_{ar}	2504	978	755	527	89	649	416
n(5+)	2488	965	756	515	73	634	415
n(3)-C6	2450	989	728	528	73	654	376
n(1)-C4	2408	935	718	534	74	653	405
Σ*n*-Alk	2456	989	732	526	74	660	406
Σ*n*-Alk/*n*(1)-C4	2540	984	687	501	71	649	375
n(3)-C6/*n*(1)-C4	2568	929	744	522	77	657	380
Critical values of the Mann–Whitney test for different *P*'s							
M**	2624	990	780	540	99	660	429
P = 0.9	2592	974	767	529	96	648	421
P = 0.5	2453	906	708	483	84	596	385
P = 0.2	2299	830	643	432	69	537	345
P = 0.1	2206	784	604	401	61	502	320
P = 0.05	2125	744	570	373	53	471	299
P = 0.01	1965	664	501	319	31	410	256
P = 0.001	1775	569	420	253	15	336	204

* The figures in bold are the quantities for which the distribution is shifted toward lower values in the oils of the second member of the pair (at $p \le 0.1$).

** T—terrigenous, K—carbonate; C_1—Lower Carboniferous, C_2—Middle Carboniferous, and so on.

*** *M* is the mathematical expectation of *U* for the case when both samples are from the same general population.

erous carbonate–Upper Devonian and Lower Carboniferous carbonate (KC_2–$K(C_1+D_3)$); and Devonian terrigenous (total)–Middle Carboniferous carbonate. The six unique samples (see above) were excluded. The calculation results and critical values for different *P*'s (from 1 to 10^{-3}) are given in Table 3 (except the last pair, for which there are no differences at $P \le 0.2$).

From these data it follows that only the H_{db} distributions considerably vary in oils from different plays. The main differences are between the terrigenous play of the Lower Carboniferous and the Upper Devonian and Lower Carboniferous carbonate play, which has a substantially lesser amount of unsaturates. The medians are equal (in 10^{-2} abs. %): TC_1, 3.4; TD, 2.1; KC_2, 1.7; and $K(C_1 + D_3)$, 0.4. The statistical significance of the difference between TC_1 and TD is ~0.75 and that between TC_1 and KC_2 is ~0.8, which is insufficient for recognizing it reliable. Regarding the other parameters, the highest significance of the difference is found in the TD–KD pair for $H_{\alpha}//H_{ar}$ (0.80). That is, there are small differences, if at all. For two thirds of the values given in Table 3 (78 out of 112), there are no differences between the play pairs altogether. In terms of applied statistics, the probability that the differences are accidental is 50% or more (corresponding value is greater than or equal to the critical value at $P = 0.5$). In some cases, it makes more than 90% (for example, *n*(3)-C6 in the TC_1–KC_2 pair).

In the analysis of oils from different territories, the whole area of the basin was divided into three parts by administrative boundaries: northern (Udmurtia and Perm oblast), central (Bashkortostan and Tatarstan), and southern (Samara, Saratov, Volgograd, and Oren-

Table 4. Comparison of the distributions of the composition parameters of crude oils from the northern (N), central (C), and southern (S) regions of the Volga-Urals OGB (without six unique oils, see text) by the Mann–Whitney test

Parameter	OGB parts		
	N–C	N–S	C–S
H_{ar}	1485	**906**	**462**
H_{db}	**900**	**297**	**645**
H_{α}	***1247***	**914**	**224**
H_{β}	**1048**	1401	***478***
H_{γ}	***1224***	**637**	***577***
$H_{ar,\ 2+}$	1503	**913**	**628**
$H_{ar,1}$	1452	**1030**	**379**
$H_{ar,\ 2+}/H_{ar,1}$	**1251**	1276	1118
$H_{\alpha}//H_{ar}$	***987***	**1058**	**311**
C_{ar}	***1153***	**951**	**295**
n(5+)	**936**	1385	***416***
n(3)-C6	**961**	***1143***	***338***
n(1)-C4	**1060**	***931***	***346***
S*n*-Alk	**953**	***1219***	***319***
S*n*-Alk/*n*(1)-C4	**1145**	**951**	1106
n(3)-C6/*n*(1)-C4	**957**	**1166**	***918***
Critical values of the Mann-Whitney test for different *P*'s			
M**	1600	1504	1175
P = 0.9	1578	1483	1158
P = 0.5	1482	1391	1082
P = 0.2	1375	1289	997
P = 0.1	1311	1227	946
P = 0.05	1255	1173	901
P = 0.01	1143	1067	813
P = 0.001	1011	940	707

* The figures in bold or italic are the quantities for which the distribution is shifted toward lower or higher values, respectively, in the oils of the second member of the pair (at $P \le 0.1$).

burg oblasts). Table 4 shows Mann–Whitney test values for all of the three pairs: north–center, north–south, and center–south by all parameters, and Table 5 presents the medians of the corresponding distributions (all samples except the same six unique samples).

There are significant differences in the majority of the parameters between all the territories; moreover, this significance in some parameters is very high even with temporal exclusion of unsaturated hydrocarbons (probability that the difference is accidental for all the parameters with $U < 330$ in the center–south pair is $<10^{-9}$). Only 10 out of 48 values in Table 4 correspond to the situation when the significance of the differences is less than 0.95.

By aromaticity (C_{ar}), the regions on average are arranged in the order: center > north > south. The same is with the parameter of substituents on aromatic rings (H_{α}) and the H_{α}/H_{ar} ratio. However, there is no difference between the oils of the north and center in H_{ar}, $H_{ar,\ 1}$, $H_{ar,\ 2+}$ (in the southern oils, their values are on average lower). But the $H_{ar,\ 2+}/H_{ar,\ 1}$ ratio is displaced toward higher values in the northern oils compared to oils from the central part of the basin; in the other two pairs (north–south, center–south), there is no reliably established difference. This implies some corollaries. For the Volga–Urals region as a whole, there will not be so unambiguous correlation between H_{ar} and C_{ar} as for Western Siberia oils [13], since there is no shift between the H_{ar} distributions of the northern and central parts of the basin, but it does exist for C_{ar}. Probably, in order to predict C_{ar} and other quantities using ^{1}H NMR data (like in [11, 19, 20]), it is better to separately consider different parts of the basin. The increased fraction of aromatic carbon in the Tatarstan and Bashkiria oils might be explained in terms of the well-known broad occurrence of deposits with high resin and/or asphaltene content, especially, in Carboniferous and Permian sediments in this region. However, if all of the differences between the aromatic components of the north and center were determined by this factor, the $H_{ar,2+}/H_{ar}$ distribution would also have to be shifted toward higher values in oils from the central part of the OGB, since typical values of this

Table 5. Medians of the composition parameters for oils from the northern, central, and the southern regions of the Volga–Ural OGB (without six unique oils, see text)

Parameter	North	Center	South	Parameter	North	Center	South
H_{ar}	4.27	4.18	3.63	$H_{\alpha}//H_{ar}$	2.36	2.71	2.23
H_{db}	3.7	1.9	0.7	C_{ar}	17.5	18.1	14.5
H_{α}	10.7	11.5	7.9	*n*(5+)	8.3	6.5	8.5
H_{β}	59.3	57.8	59.7	*n*(3)-C6	1.87	1.61	2.23
H_{γ}	25.7	26.7	28.4	*n*(1)-C4	2.46	2.19	2.85
$H_{ar,\ 2+}$	2.13	2.17	1.92	Σ*n*-Alk	16.7	13.9	18.8
$H_{ar,1}$	2.10	2.07	1.79	Σ*n*-Alk/*n*(1)-C4	6.7	6.4	6.4
$H_{ar,\ 2+}/H_{ar,\ 1}$	1.10	1.03	1.05	*n*(3)-C6/*n*(1)-C4	0.77	0.74	0.76

ratio for resins or asphaltenes and monocyclic aromatics are ~4 and ~0.1, respectively, according to our data. But the observed shift is in the opposite direction (Tables 4, 5). Therefore, the composition of the aromatic components of the northern and central oils differ in other characteristics as well, a fact that requires special consideration. In particular, since the shifts in C_{ar}, H_α, and H_α/H_{ar} are parallel, an increase in the proportion of aromatic carbon in this territorial consideration is accompanied by an increase in the degree of substitution on aromatic rings.

The oils from the central part of the basin are distinguished by a lower amount of *n*-alkyl structures (Tables 4, 5). The difference between the north and south in Σn-Alk is small, if at all (~ 1 : 10, i.e., the observed difference is accidental). The average chain length Σn-Alk/n(1)-C4) is the largest in oils of the northern regions and is about the same in the central and southern parts of the basin. It is noteworthy that the lower average length chain in Tatarstan and Bashkortostan oils is apparently due to the fact that their significant part is represented by relatively short substituents on cyclic compounds, whereas it value for the southern regions is associated with a lower average molecular weight of oil components.

Unsaturated hydrocarbons make the only group of the compounds considered here, whose average concentration in different parts of the basin differs by two or more times. If the entire range of H_{db} values is divided in half, the upper region ($H_{db} \geq 7 \times 10^{-2}$ abs. %) will include oils that mainly occur in Perm oblast (Yarino-Kamenolozhskoe, Yaborovskoe, Chashkinskoe, Rusakovskoe, Kuliginskoe, Krasnokamskoe, Karamorskoe, Luzhivskoe, Gezhskoe, and Brusyanskoe fields and Tikhovskaya area). Only two fields in Bashkiria (Kyzylbaevskoe and Arslanovskoe) and one in Tatarstan (Aktanysh) have deposits of such oil. On the other hand, there are more than half (30 out of 56) oils from the southern part of the basin among the samples belonging to the range of $H_{db} < 0.01$ abs. % (or <0.3–0.4 abs. % unsaturated hydrocarbons).

Thus, despite the well-known difference between the oils from carbonate and terrigenous sediments, it has no effect on their total aromaticity and content of *n*-alkyl structures. There are distinct differences found by averaging consideration of these generalized composition parameters between the oils of the northern, central, and southern regions of the basin. However, these differences are generally small (Table 5), so that serious distortions due to uneven representation of oils from different regions do not arise. The exception is H_{db}, for which the resulting distribution gives overestimated values of the probability density at high values (perhaps, by several times).

These distributions were compared with the published data on Western Siberia oils [12]. The results of the qualitative comparison are shown in Figs. 2–4. To establish the presence and degree of reliability of systematic shifts in the distributions, the Mann–Whitney test and the medians were used. The distributions obtained for all of the samples (variant I); with the rejected six unique Volga–Urals oils and two gas condensates that dropped out during the analysis of many relationships in Western Siberia [13] (variant II, which is of the greatest interest); and, when it was meaningful, "the main arrays of samples" (variant III) were compared.

By C_{ar}, the distribution for Western Siberia oils is shifted toward lower values in variants I and II (significances of 0.998 and 0.9985). In the region with the highest probability density (12–22 abs. %), the distributions are close: the positions of both maximums and the minimum in between coincide (Fig. 4). The shift is due to reduced aromaticity of the oils that are genetically related to Paleozoic sediments of Western Siberia [12]. There is no shift of the H_{ar} distribution in both of these variants and a slight shift only in version III for the region of 1.5–6.0 abs. % (significance is low, 0.93). However, a shift to the lower side for the Volga–Urals oils is clearly seen in the interval that includes the majority of values (2.5 to 5.5 abs. %). Hence, the equations relating H_{ar} to C_{ar} for the Western Siberia and Volga–Urals oils will be different and there cannot be a universal procedure for predicting C_{ar} values with good accuracy from the measured values of H_{ar} for oils; thus, the consideration is to be limited to regional relationships.

In the Western Siberia oils, the values of H_α on average are significantly lower (medians are 7.9 abs. % for Western Siberia and 10.1 abs. % for Volga–Urals in version II), as well as the values of H_α/H_{ar}, and the values of $H_{ar,\,2+}/H_{ar,\,1}$ are higher (significance is ~0.99999 for the main group of samples in which the ratio is 0.6 to 1.6 (Fig. 2)). The latter shift is due to smaller values of $H_{ar,\,2+}$. Consequently, aromatic rings in the Volga–Urals oils on average bear a greater number of substituents. A detailed study of the relations between different classes of aromatic compounds require a special investigation involving the isolation of fractions of mono- and bi- + polycyclic aromatics and use of mass spectrometric group analysis techniques [26]. The existing oil generation schemes [1, 2, 4] cannot give any explanation of these differences. So far we only note that, first, oils that form the main peak of the H_α distribution centered at ~12 abs. % are rare in Western Siberia. Second, the distribution of the H_α/H_{ap} ratio is unimodal in West Siberia and distinctly bimodal with comparable maximums in Volga–Urals. That is, there are at least two types distinguished in the Volga–Urals oils by this parameter.

By *n*-alkyl structures, a slight shift to lower values is observed in the Volga–Urals oils for Σn-Alk, n(3)-C6, and n(1)-C4 (significance in variant II is ~0.98, 0.97, or 0 99, respectively) and the absence of a significant difference for n(5+). The Σn-Alk/n(1)-C4 ratio in on average lower in Western Siberia oils, the n(3)-C6/n(1)-C4 distributions are almost identical except for the "wings" (see also Fig. 3). The main difference

between the distributions of these three parameters is in the width of the principal maximum, especially for *n*(3)-C6: the interval that contains 50% of its values in the West Siberia oils is half that in the Volga–Urals basin (Fig. 3). In view of the distribution of C_{ar}, it follows that the relationship between C_{ar} and Σ*n*-Alk observed in Western Siberia oils will be different in the Volga–Urals oils and, most likely, will not be as strong.

It should be pointed out that the comparative analysis as a whole in this paper is based on nonparametric methods applicable to all distributions, while comparison of the means and the *t*-test, whose domain of correct application is limited to values with Gaussian distribution [5–7], are conventionally used in this area. To illustrate the consequences of the conventional, but incorrect use of the mathematical apparatus, the results of the comparative analysis of the composition of the Volga–Urals and West Siberia oils in version I were compared with the picture obtained using the *t*-test. In two cases, this test showed the presence of differences with significances of 0.93 and 0.991 where these differences are absent (probability that the differences are accidental are >0.7 and >0.2, respectively, in the correct analysis). In two other cases, the opposite situation is observed: there is no difference according to the *t*-test ($P = 0.56$ or 0.21), but this probability is actually 0.002 or 0.001, respectively, i.e., the differences are significant. In two more cases, the values of *P* calculated by the Student's *t*-test differ from the correct values by an order of magnitude and in different directions again. Thus, the *t*-test gave a completely erroneous conclusion for a quarter of the parameters and an incorrect estimate of the significance for one eighth of them. This casts doubt on the results of many versatile works based on this comparative analysis.

In conclusion, a few words should be said about the six unique oils. Obviously, they are of particular interest and call for detailed investigation to determine their composition and generation environment. Some data on high-paraffin Ol'khovskoe oil are presented in [26]. In particular, it was found that this oil is strongly distinguished from other oils by a high concentration of T-branched hydrocarbons (with ethyl, *n*-propyl, etc. substituents on the alkyl chains). The available generalized characteristic of the composition of five oils with extremely high aromaticity only allows the suggestion that the aromatic rings in them on average are substituted to a substantially lesser extent. The absence of signal broadening in the spectra indicates a low asphaltene content. Regarding the average number of carbon atoms per molecule in aromatic cycles and the extremely low concentration of monoaromatic components to C_{10}, the aromatic compounds have to be substantially dominated by the bi- and polycyclic species. However, these oils are not distinguished from the others by the $H_{ar,2+}/H_{ar,1}$ ratio. This suggests that they contain a significant proportion of compounds in which benzene rings are linked either by an alkyl chain or through saturated cycles; in the latter case, these are mostly not fluorenes.

Thus, it has been shown that the distributions of 16 main NMR-measurable characteristics of the structural-group composition of Volga–Urals crude oils strongly differ from the normal distribution and, as a rule, are asymmetric and significantly dissimilar. The total aromaticity and the content of *n*-alkyl structures in the oils of the main stratigraphic plays on average are identical in the presence of differences between the northern, central, and southern parts of the oil-and-gas basin. In the amount of unsaturated hydrocarbons, there are differences both between the plays and territories. Similarities and dissimilarities in the distributions between the West Siberian and Volga–Urals basins have been revealed. A number of features found in the Western Siberia oils should be different in the Volga-Urals oils. This also applies to the relationship between H_{ar} and C_{ar}. In the Upper Carboniferous and Lower Permian carbonate reservoirs on the northeast of the OGB, oils with a uniquely high aromatics content were found, which strongly differ in the composition of aromatic compounds from common oils. The advantages of using interval probability density plots instead of the histograms for comparison of the distributions have been demonstrated. Using particular data, we have shown that the *t*-test conventionally used to compare average values for oil groups leads to false conclusions. In order to determine systematic shifts between distributions, it is necessary to use the Mann–Whitney test and match the medians of the distributions.

ACKNOWLEDGMENTS

This work was supported by Chembridge.

Special thanks to T.K. Bazhenova and S.B. Ostroukhov for technical assistance.

REFERENCES

1. B. Tissot and D. Welte, *Petroleum Formation and Occurrence* (Springer, Heidelberg, 1978).
2. A. A. Petrov, *Petroleum Hydrocarbons* (Springer, Berlin, 1987).
3. *Catalog of Parameter of Regional Characterization of Chemical and Individual composition of USSR Crude Oils*, Ed. by S. P. Maksimov and T. A. Botneva (Nedra, Moscow, 1981) [in Russian].
4. K. E. Peters, C. C. Walters, and J. M. Moldovan, *The Biomarker Guide: Biomarkers and Isotopes in Petroleum Systems and Earth History* (Cambridge University Press, Cambridge, 2005).
5. Bol'shev, L.N. and Smirnov, N.V., *Tables of Mathematical Statistics* (Nauka, Moscow, 1983).
6. A. I. Orlov, *Applied Statistics Textbook* (Ekzamen, Moscow, 2007) [in Russian].
7. M. B. Smirnov, *Lectures on Basics of Experimental Data Processing: Handout for Chemistry and Geochemistry*

Students (INKhS, Moscow, 2013), [in Russian]. www.ips.ac.ru

8. M. V. Giruts and G. N. Gordadze, Pet. Chem. **53**, 209 (2013).
9. D. A. Bushnev and N. S. Burdel'skaya, Pet. Chem. **53**, 145 (2013).
10. G. S. Pevneva, V. V. Savel'ev, and A. K. Golovko, Pet. Chem. **53**, 288 (2013).
11. B. R. Sadykov, V. P. Starikov, R. Kh. Sadykov, and G. A. Kalabin, Pet. Chem. **52**, 22 (2012).
12. M. B. Smirnov and N. A. Vanyukova, Pet. Chem. **54**, 16 (2014).
13. M. B. Smirnov and N. A. Vanyukova, Pet. Chem. **54**, 355 (2014).
14. S. P. Maksimov, V. A. Kirov, V. A. Klubov, et al., *Geology of Oil and Gas Fields in the Volga–Urals Province* (Nedra, Moscow, 1970) [in Russian].
15. S. P. Maksimov, G. Kh. Dikenshtein, and A. N. Zolotov, *Oil and Gas Geology of the East European Platform* (Nedra, Moscow, 1990) [in Russian].
16. A. E. Kontorovich, *Oil and Gas Basins and Regions of Siberia*, issue 2: *The Western Siberia Basin*, Ed. by A. E. Kontorovich, V. S. Surkov, A. A. Trofimuk, et al. (OIGGM, SNIIGGiMS, Novosibirsk, 1994) [in Russian].
17. I. V. Goncharov, *Geochemistry of Western Siberia Oils* (Nedra, Moscow, 1987) [in Russian].
18. *Handbook of USSR Oil and Gas Fields*, in two vols., Ed. by S. P. Maksimov (Nedra, Moscow, 1987) [in Russian].
19. G. A. Kalabin, L. V. Kanitskaya, and D. F. Kushnarev, *Quantitative NMR Spectroscopy of Natural Organic Feedstock and Its Processing Products* (Khimiya, Moscow, 2000) [in Russian].
20. D. Cookson and P. Lioyd, Fuel **65**, 1247 (1986).
21. L. S. Borisova and A. E. Kontorovich, in *Advances in Organic Chemistry: Proceedings of All-Russia Conference* (Novosibirsk, 2010), p. 52 [in Russian].
22. V. P. Lis'ev, *Theory of Probability and Mathematical Statistics: Handout for Students* (Moskovskii Gos. Univ. Ekonomiki, Statistiki i Informatiki, Moscow, 2006) [in Russian].
23. E. B. Frolov and M. B. Smirnov, Org. Geochem. **21**, 189 (1994).
24. E. B. Frolov, V. A. Melikhov, and M. B. Smirnov, Org. Geochem. **25**, 265 (1996).
25. E. S. Brodskii, A. A. Shelepchikov, G. A. Kalinkevich, et al., Pet. Chem. **54**, 28 (2014).
26. M. B. Smirnov and S. B. Ostroukhov, in *Prognosis and Development of Oil and Gas-Bering Structure of Lower Volga Region and Northern Caspian Sea: Collection of Articles* (Filial LUKOIL-Engineering Volgograd-NIPIneft', Volgograd, 2013), No. 72, p. 120 [in Russian].

Translated by S. Zatonsky