INFORMATION TECHNOLOGIES AND AUTOMATIC CONTROL IN FERROUS METALLURGY

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И АВТОМАТИЗАЦИЯ В ЧЕРНОЙ МЕТАЛЛУРГИИ

POSSIBLE CONSEQUENCES CAUSED BY REDUCING THE NUMBER OF SAMPLES DURING ACCEPTANCE TESTS OF A UNIT OF METAL PRODUCTS TO ASSESS ITS QUALITY

E. A. Sokolovskaya ✉, **E. V. Bosov, A. V. Kudrya,**

D. F. Kodirov, V. I. Alekseev

▌ **National University of Science and Technology "MISIS"** (4 Leninskii Ave., Moscow 119049, Russian Federation)

✉ sokolovskaya@misis.ru

Abstract. For 13G1S-U sheet steel and large forgings made of heat-treatable 38KhN3MFA-Sh steel, produced by existing technologies, the possible consequences for quality assessment of metal products were analyzed related to changes in the number of samples used in testing a single product unit (batch, forging). Based on the calculation of skewness and kurtosis coefficients, the authors estimated the change of distribution type of impact strength values accompanying the change in the number of samples. The sampling of impact strength range values obtained from testing two samples (three possible paired combinations) per unit of products were compared using the Student's and Smirnov's criteria, both among themselves and with the original sampling (three samples for evaluating one batch of sheet). The obtained results also showed that in conditions when the statistical nature of values distributions of the metal products quality parameters differs from the normal distribution type, it is necessary to use the criteria of nonparametric statistics. The risks of possible loss of information on the metal products' quality when reducing the number of samples tested within a single batch were assessed. In order to obtain adequate results of statistical analysis, it is necessary to identify and eliminate possible side effects that distort results of analysis: trends, seasonal fluctuations, and data recording errors. For metal products characterized by the developed heterogeneity of structures, obtaining objective information on the toughness reserve of steels can be obtained on the basis of micromechanical tests of samples whose dimensions are comparable to the scale of structural heterogeneity. The obtained results can be useful in the statistical analysis of production process and product control databases in metallurgy to obtain reasonable technological recommendations (within the framework of operation of the end-to-end quality management system) aimed at improving the uniformity of metal product.

Keywords: statistical analysis in metallurgy, quality assessment of metal products, informativeness of mechanical testing results, statistical nature of production control data, nonparametric and classical statistics, Big Data

О ВОЗМОЖНЫХ ПОСЛЕДСТВИЯХ, ВЫЗВАННЫХ УМЕНЬШЕНИЕМ КОЛИЧЕСТВА ОБРАЗЦОВ ПРИ ПРИЕМО-СДАТОЧНЫХ ИСПЫТАНИЯХ ЕДИНИЦЫ МЕТАЛЛОПРОДУКЦИИ ДЛЯ ОЦЕНКИ ЕЕ КАЧЕСТВА

Э. А. Соколовская ✉, **Е. В. Босов, А. В. Кудря,**

Д. Ф. Кодиров, В. И. Алексеев

Национальный исследовательский технологический университет «МИСИС» (Россия, 119049, Москва, Ленинский пр., 4)

✉ sokolovskaya@misis.ru

Аннотация. Для листовой стали 13Г1С-У и крупногабаритных поковок из улучшаемой стали 38ХН3МФА-Ш, изготовленных по существующим технологиям, были проанализированы возможные последствия для оценки качества металлопродукции, связанные с изменением количества образцов, используемых при тестировании единицы продукции (партии, поковки). На основе расчета коэффициентов асимметрии и эксцесса авторы оценили изменение вида распределения значений ударной вязкости, сопутствующее вариации количества образцов. Выборка значений диапазона ударной вязкости, полученных в результате тестирования двух образцов (трех возможных парных комбинаций) на единицу продукции, сравнивалась с использованием критериев Стьюдента и Смирнова как между собой, так и с исходной выборкой (три образца для оценки одной партии листа). Полученные результаты также показали, что в условиях, когда статистическая природа распределений значений параметров качества металлопродукции отличается от нормального вида распределения, необходимо использование критериев непараметрической статистики. Были оценены риски возможной потери информации о качестве металлопродукции при уменьшении количества образцов, испытываемых в рамках отдельной партии. Для получения адекватных результатов статистического анализа необходимо выявить и устранить возможные побочные явления, искажающие результаты анализа: тренды, сезонные колебания, ошибки в записи данных. Для металлопродукции, отличающейся развитой неоднородностью структур, объективная информация о запасе вязкости сталей может быть получена на основе микромеханических испытаний образцов, габариты которых сопоставимы с масштабом структурной неоднородности. Полученные результаты могут быть полезны при статистическом анализе баз данных производственного контроля процесса и продукта в металлургии для подготовки обоснованных технологических рекомендаций (в рамках функционирования сквозной системы управления качеством), направленных на повышение однородности качества металлопродукции.

Ключевые слова: статистический анализ в металлургии, оценка качества металлопродукции, информативность результатов механических испытаний, статистическая природа данных производственного контроля продукта, критерии непараметрической и классической статистики, Big Data

INTRODUCTION

Steel production is a complex, multi-stage process, with each stage well-equipped with tools for measurement and data collection. Depending on the intended application, the quality of metal products is typically assessed through mechanical testing, structural analysis, and fracture evaluation. Objective quality assessment is also essential for addressing the inverse problem – namely, establishing the relationship between structure and properties, identifying critical structural parameters responsible for variations in metal quality, and developing technological recommendations to improve product uniformity [1 – 2]. In this context, modern IT solutions – such as neural networks, Big Data algorithms, and machine learning – are increasingly used to process large datasets generated through production process and product quality control [3 – 6].

There is growing interest in enhancing methods for evaluating the quality of metal products, particularly given the new opportunities created by the digitalization of structural and fracture measurements and the automation of experimental data processing [7 – 8]. For example, despite decades of experience with mechanical testing, it remains unclear how variations in the number of test samples used in acceptance testing may affect the completeness of quality assessments – especially considering the wide range of heterogeneous structures that form under standard, well-established technologies and the significant resulting variability in properties, particularly impact strength.

Various approaches exist for determining the number of samples per unit of metal product during mechanical testing [9]. Such requirements may be specified in regulatory standards. For example, GOST 4543 – 2016 "Metal Products Made of Structural Alloy Steel" stipulates that one sample for tensile testing and one for impact testing (under each relevant test condition) must be taken from each bar, strip, or coil selected for quality control. In some cases, testing procedures are established by agreement between the manufacturer and the customer, based on the product's intended use. Typically, the number

of identical tests per control unit ranges from one to three, with impact testing usually performed at the upper limit due to the high variability of results. Sampling norms are the result of long-standing practical experience and have remained largely unchanged for decades. Their conservative nature is reflected in the update history of such standards – for example, GOST 4543 was revised in 1948, 1971, and 2016.

With the accumulation of representative production control databases, the growth of computational resources, and the development of specialized software tools, it has become possible to evaluate how changes in the number of test samples per product unit influence quality certification outcomes. While a reduction in the number of tests clearly leads to some loss of valuable information, the extent of that loss requires clarification. Understanding this is essential not only for evaluating the objectivity of quality assessments but also for making informed decisions on adjusting the technological process – particularly given the significant variation in acceptance testing results commonly observed in practice. Interest in this issue is also driven by the widespread use of mechanical testing and its contribution to the overall production cost of metal products.

A retrospective approach to this issue also allows for consideration of the statistical nature of the subject (i.e., large-scale production control datasets), making it possible to justify the selection of appropriate statistical tools that enhance the objectivity of the findings [10 – 11].

In this regard, the aim of this study is to assess the completeness of information obtained when varying the number of samples used in mechanical testing of individual metal product units. This analysis serves as a foundation for improving the effectiveness of metal quality forecasting through statistical analysis of production process and product control data in metallurgy.

RESEARCH OBJECTS AND METHODS

This study was based on production control databases for two types of metal products manufactured using established technologies over one to two years: large, variable cross-section forgings made of heat-treatable 38KhN3MFA-Sh steel, and sheet products made of 13G1S-U steel [1]. Each database was structured as a matrix $A_{m \times n}$, where the rows (m) corresponded to the number of heats (or batches/forgings), and the columns (n) included both technological parameters (n_t) and product quality characteristics (n_q). For 38KhN3MFA-Sh steel, the database contained $m = 342$ forgings derived from 40 individual heats, with the matrices linked to the chemical composition after electroslag remelting. For 13G1S-U sheet steel with thicknesses of 8, 10, and 12 mm, the number of batches was $m = 751$, 668, and 1281, respectively. the total number of columns n in the matrices was 91/20 and 33/16 for the two steels, corresponding to n_t/n_q.

values. the output parameters included, in particular, ultimate tensile strength (σ_u), yield strength ($\sigma_{0.2}$), elongation at break (δ), and impact strength (KCU/KCV), all measured at various test temperatures. Tangential mechanical test samples from 38KhN3MFA-Sh forgings were cut from end-face templates at maximum (D_l) and minimum (D_s) diameters. Two samples were tested per temperature condition: one for tensile testing at +20 °C , and two for impact strength at +20 (KCU_i and KCU_j) and –50 °C (KCU_i^{-50} and KCU_j^{-50}). From each batch of 13G1S-U sheet steel, one transverse sample was selected for tensile testing at room temperature, and three transverse samples were selected for impact strength testing at –40 °C and 0 °C, respectively (KCU_i^{-40}, KCU_j^{-40}, KCU_k^{-40} and KCV_i^0, KCV_j^0 and KCV_k^0) .

Statistical evaluation of sample groups (or batches) of acceptance test parameters was performed using Microsoft Excel. For each group, the maximum ($x_{i\,max}$), minimum ($x_{i\,min}$), mean $\overline{X}_i$ (with standard deviation s) and range ($\Delta = x_{i\,max} - x_{i\,min}$) were determined. the distribution type of each parameter was analyzed by constructing histograms using equal-width intervals. the number of intervals was set as the cube root of the number of measurements [12]. Skewness (A_s) and kurtosis (E_x) coefficients were then calculated, along with their respective standard errors [13 – 14].

Sample comparisons were performed using the Smirnov's and Student's tests (hereafter referred to as C_n and C_p, respectively), with the significance level of each hypothesis tested.

RESULTS AND DISCUSSION

One of the main barriers to the effective use of modern software tools in processing large-scale datasets is the insufficient attention paid to identifying factors that can significantly distort the results of statistical analysis. These distortions may affect both input variables and output parameters – such as the results of acceptance testing. For example, when chronological series of quality characteristics and technological parameters were constructed for 13G1S-U steel, correlated "seasonal" fluctuations were observed in both impact strength and niobium content (Fig. 1). This pattern indicates that the original dataset is, effectively, divided into two distinct subsets: one with low niobium content (Nb ≤ 0.03 wt. %) and one with high content (Nb > 0.03 wt. %). For sheet thicknesses of 8, 10, and 12 mm, the number of batches in the low-Nb subset was 269, 395, and 260, respectively; in the high-Nb subset, 489, 273, and 1021 batches. Conducting a combined statistical analysis of these subsets would result in averaging both input and output parameters, which would distort the actual shapes of the parameter distribution histograms and hinder the use of statistical tools such as regression analysis. Therefore, subsequent statistical processing was performed separately for
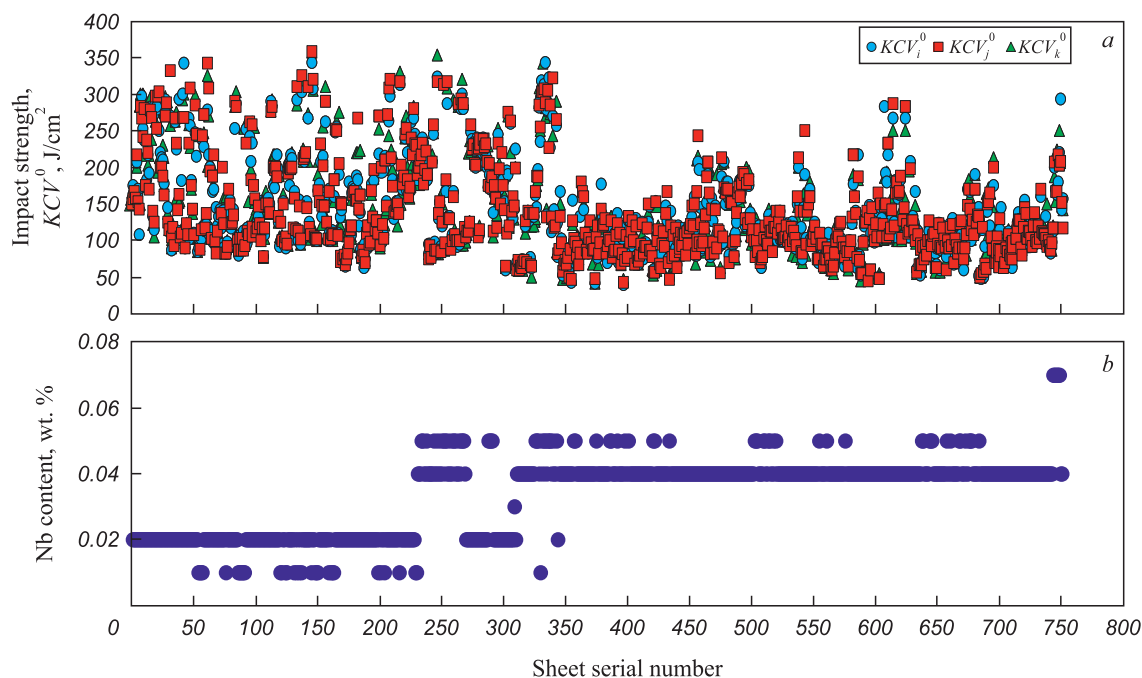
Fig. 1. Chronological series of distribution of impact strength KCV^0 (for three samples per batch) (*a*)
and niobium content (*b*) in a sheet 8 mm thick made of 13G1S-U steel

Рис. 1. Хронологические ряды распределения ударной вязкости KCV^0 (по трем образцам на партию) (*a*)
и содержания ниобия (*b*) в листе толщиной 8 мм из стали 13Г1С-У

the two subsets. the present paper reports the results for the subset with Nb > 0.03 wt. %, as it was the more representative in terms of sample size.

It is important to emphasize that such effects are not unique to impact strength and may take various forms. For example, in a chronological series of yield strength values for pipe steel (strength category K65, wall thickness 27.7 mm), which correspond to the left and right peaks of a bimodal distribution histogram, a consistent alternation of values over time was observed [15].

For all datasets – regardless of how they were structured (e.g., whether trends were identified, data entry errors corrected, or other preprocessing steps applied) – the key statistical characteristics of the acceptance parameters were calculated. the analysis revealed a substantial spread in mechanical property values (Tables 1 and 2).

The greatest variability in property values was observed for impact strength, compared to strength parameters. For example, in 13G1S-U steel, the maximum toughness value exceeded the minimum by a factor of 5 – 6, while for 38KhN3MFA-Sh steel, the difference was twofold – similar trends were also observed in other types of metal products [10]. Such heterogeneity in toughness is attributed to differences in the scenarios of technological inheritance – the realization of diverse mechanisms of structural and defect evolution along the technological chain, ultimately leading to a wide variety of morphologies in nominally similar final structures [1; 16; 17]. This underscores the need for an objective assessment of the extent of variability in impact strength. Consequently, any variation in the number of tested samples (in this case, per individual forging

Table 1. **Scale of heterogeneity of 13G1S-U sheet steel quality**

Таблица 1. **Масштаб неоднородности качества листовой стали 13Г1С-У**

Sheet thickness, mm	Parmeter	KCU^{-40}, J/cm^2	KCV^0, J/cm^2	σ_u, MPa	$\sigma_{0.2}$, MPa	δ, %
8	$x_{i\,max} - x_{i\,min} = \Delta$	382 – 48 = 334	343 – 40 = 303	660 – 487 = 173	545 – 391 = 454	36 – 18 = 18
	$\overline{x}_i \pm s$	135 ± 51	121 ± 52	568 ± 29	450 ± 25	28 ± 3
10	$x_{i\,max} - x_{i\,min} = \Delta$	280 – 48 = 232	372 – 52 = 320	560 – 399 = 161	655 – 515 = 140	36 – 19 = 17
	$\overline{x}_i \pm s$	136 ± 40	118 ± 47	464 ± 26	577 ± 24	27 ± 3
12	$x_{i\,max} - x_{i\,min} = \Delta$	489 – 14 = 475	365 – 26 = 339	640 – 492 = 148	545 – 379 = 166	36 – 17 = 19
	$\overline{x}_i \pm s$	132 ± 50	117 ± 50	568 ± 20	455 ± 27	28 ± 2

Table 2. Scale of heterogeneity of quality of forgings from 38KhN3MFA-Sh steel according to the tests of samples cut in end templates of forgings with diameters D_l and D_s

Таблица 2. Масштаб неоднородности качества поковок из стали 38ХН3МФА-Ш по испытаниям образцов, вырезанных из торцевых темплетов поковок диаметром D_l и D_s

Tamplate diameter	Parameter	KCU, J/cm^2	KCU^{-50}, J/cm^2	σ_u, MPa	$\sigma_{0.2}$, MPa	δ, %
D_s	$x_{i\,max} - x_{i\,min} = \Delta$	$63 - 28 = 35$	$58 - 20 = 38$	$1580 - 1190 = 390$	$1490 - 1110 = 380$	$16.5 - 9.3 = 7.2$
	$\overline{x}_i \pm s$	47 ± 6	40 ± 6	1278 ± 40	1375 ± 35	$13.6 \pm 1,1$
D_l	$x_{i\,max} - x_{i\,min} = \Delta$	$56 - 31 = 25$	$51 - 20 = 31$	$1570 - 1340 = 230$	$1490 - 1230 = 260$	$17.5 - 8.8 = 8.7$
	$\overline{x}_i \pm s$	43 ± 4	33 ± 5	1483 ± 31	1377 ± 23	12.3 ± 1.1

or sheet batch) can become a significant factor affecting the reliability of the quality assessment.

However, comparing different datasets solely based on their mean values and ranges does not always provide a complete picture of the variability in the quality characteristics of metal products. However, comparing different datasets solely based on their mean values and ranges does not always provide a complete picture of the variability in the quality characteristics of metal products. For instance, a large range may result from the presence of isolated outliers, and mean values and their variances are meaningful only if the distribution of quality parameters approximates a normal pattern [10]. This consideration motivated the construction of distribution histograms. Separate histograms were constructed for the results of the first $\{KCU_i^{-40}\}$, second

$\{KCU_j^{-40}\}$, and third $\{KCU_k^{-40}\}$ tests for 13G1S-U steel, and for the first $\{KCU_i^{-50}\}$ and second $\{KCU_j^{-50}\}$ tests for 38KhN3MFA-Sh steel, with the test number corresponding to a specific column in the data matrix (Fig. 2).

The distribution of impact strength values showed noticeable deviations from the normal distribution. the extent of this deviation was assessed based on the corresponding skewness and kurtosis coefficients, which varied over a fairly wide range. For 38KhN3MFA-Sh steel, the skewness (A_s) and kurtosis (E_x) coefficients of impact strength values ranged from [–0.43; –0.31] and [0.70; 0.80], respectively for templates with diameter D_s and within [–0.31; 0.35] and [–0.14; 0.23] for templates with diameter D_l. For the distributions of KCV^0 and KCU^{-40} impact strength values in 13G1S-U steel – specifically for sheets with a thickness of 8 mm – the skewness
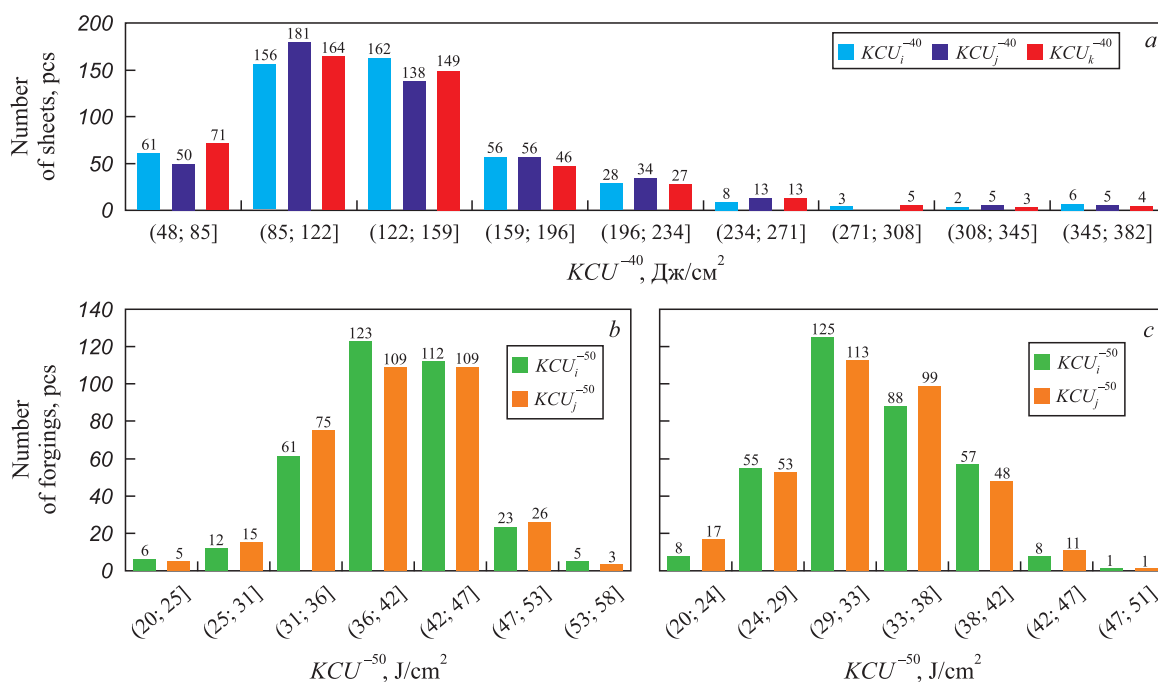

Fig. 2. Histograms of distribution of impact strength values (KCU^{-40}) of 13G1S-U steel (sheet thickness – 8 mm) (*a*) and 38KhN3MFA-Sh steel (KCU^{-50}) for end-temples with diameters D_s (*b*) and D_l (*c*)

Рис. 2. Гистограммы распределения значений ударной вязкости (KCU^{-40}) стали 13Г1С-У (толщина листа 8 мм) (*a*) и стали 38ХН3МФА-Ш (KCU^{-50}) для торцевых темплетов поковок диаметром D_s (*b*) и D_l (*c*)

and kurtosis coefficients ranged as follows: [1.64; 1.68]; [1.54; 1.81] and [3.12; 3.40]; [3.4; 5.29]. Differences in the statistical indicators derived from individual test results – $\{KCU_i^{-40}\}$, $\{KCU_j^{-40}\}$, $\{KCU_k^{-40}\}$ and $\{KCV_i^0\}$, $\{KCV_j^0\}$, $\{KCV_k^0\}$ (assuming that impact strength was assessed using only one sample per product unit – either a sheet batch or a forging) – reflect underlying differences in the statistical nature of these datasets. the absolute values of the skewness and kurtosis coefficients provide a quantitative measure of deviation from a normal distribution. This, in turn, underscores the need to account for the shape of the distribution when determining the appropriate number of samples required for metal product certification.

In statistical analysis, the reliability of results depends heavily on the sample size [1; 13]. For samples with the largest available volumes (V_i), the error in calculating the skewness and kurtosis coefficients was minimal – 0.23 and 0.77, respectively. As the sample size V_i decreased – starting from $V_i = 200 – 250$ for sheet steel (depending on thickness) and $V_i = 150 – 200$ for forgings – both the variability in skewness values and the associated estimation error increased noticeably. With further reductions in sample size – for example, in the case of KCU^{-40} impact strength values (based on three samples per batch) for 13G1S-U steel sheets 12 mm thick – the skewness coefficient (A_s), calculated across 20 subsamples (each containing 50 batches) sequentially extracted from an original

dataset of 1000 batches, ranged within [–0.2; 2.2], with an estimation error of 3.0. Clearly, at this level of variation of coefficient As and with such a high error margin, the resulting statistical estimates cannot be considered reliable. the same conclusion applies to impact strength samples for 38KhN3MFA-Sh steel forgings.

In this context, relying on mean values across samples and applying classical statistical tests for comparison may introduce uncertainty into the resulting assessments. For example, pairwise comparisons of impact strength values in 13G1S-U steel – based on the standard testing results from three samples per batch (Table 3) – revealed that the significance level of the hypothesis of sample equivalence differed considerably when evaluated using the Student's and Smirnov tests [18], with discrepancies reaching up to 30 %. In practical terms, this means that the three possible pairwise combinations of impact strength results from each batch (i.e., product unit) may follow different distribution patterns, suggesting differences in their underlying statistical nature. As a result, classical statistical criteria such as the Student's test may not always confirm the equivalence of samples – and in some cases, simultaneous use of the Student's and Smirnov's tests may even lead to contradictory conclusions.

It is clear that testing two samples per product unit improves the overall completeness of the quality assessment, compared to testing just one. However, this also

Table 3. Comparison of different impact strength samplings (obtained by their pairwise extraction from the results of standard toughness evaluation of three samples per batch) by Student's (C_p) and Smirnov's (C_n) criteria, 13G1S-U* steel

Таблица 3. Сопоставление различных выборок ударной вязкости (полученных при их попарном извлечении из результатов штатной оценки вязкости по трем образцам на партию) по критериям Стьюдента (C_p) и Смирнова (C_n), сталь 13Г1С-У*

Impact strength, J/cm²	Sheet thickness, mm	Experimental test statistics / significance levels for pairs of test results					
		$\{KCU(V^0)_i^{-40}\}$, $\{KCU(V^0)_j^{-40}\}$		$\{KCU(V^0)_i^{-40}\}$, $\{KCU(V^0)_k^{-40}\}$		$\{KCU(V^0)_j^{-40}\}$, $\{KCU(V^0)_k^{-40}\}$	
		C_n	C_p	C_n	C_p	C_n	C_p
KCU^{-40}	8	0.741 / 0.640	0.157 / 0.900	0.870 / 0.430	0.521 / 0.700	0.870 / 0.430	0.679 / 0.500
	10	0.513 / 0.950	0.091 / 0.900	0.427 / 0.990	0.115 / 0.900	0.470 / 0.980	0.025 / 0.900
	12	0.509 / 0.950	0.437 / 0.700	0.464 / 0.980	0.345 / 0.800	0.553 / 0.920	0.096 / 0.900
KCV^0	8	0.322 / 0.990	0.094 / 0.900	0.322 / 0.990	0.175 / 0.900	0.419 / 0.990	0.080 / 0.900
	10	0.557 / 0.910	0.217 / 0.900	0.514 / 0.950	0.215 / 0.900	0.729 / 0.660	0.429 / 0.700
	12	0.553 / 0.920	0.165 / 0.900	0.664 / 0.770	0.420 / 0.700	0.487 / 0.970	0.254 / 0.800

* Cells highlighted in color indicate discrepancies between the results of sample comparison by Student's (C_p) and Smirnov's (C_n) criteria.

raises the question of how to select a single value that adequately represents the product's quality. Calculating the mean $\overline{X}_i$ of two values may be inappropriate, as the same average can correspond to different ranges (Δ) between maximum and minimum values. Likewise, samples with identical ranges may differ in their absolute property levels, such as the median. One possible approach to statistical analysis is to select the minimum (i.e., worst-case) toughness value; however, it is generally preferable to base such decisions based on a comprehensive evaluation of all possible assessment options [19].

From a general standpoint, it is evident that testing three samples per product unit not only allows for estimating the standard error of the mean (s), but also provides a more objective assessment of quality heterogeneity – both in terms of standard deviation and range (Δ).

Distributions of impact strength ranges (ΔKCU^{-40}) were constructed based on the results of all three tests conducted for each batch of 13G1S-U steel sheets (12 mm thick), as well as on all possible pairwise combinations of these results (Δ_{i-j}, Δ_{i-k} and Δ_{j-k}, where indices

correspond to column numbers in the data matrix). A unified binning method was used to make these distributions comparable (Fig. 3).

All resulting histograms of range values exhibited right-skewed distributions, as confirmed by the calculated skewness (A_s) and kurtosis (E_x) coefficients (Table 4).

As expected, the absolute skewness and kurtosis values for the full three-sample distributions ($\Delta = KCU_{\max} - KCU_{\min}$) owere lower than those for the pairwise combinations. However, the absolute ranges based on three samples exceeded the pairwise ranges (Δ_{i-j}, Δ_{i-k} and Δ_{j-k}) in 56.4 – 67.7 % of cases across all sheet thicknesses and batches. In the pairwise data, range values tended to be smaller: the majority fell into the first bin, some were identical (zero range), and in subsequent bins, the number of values was 1.5 to 2.5 times lower than in the histogram based on three-sample data. All pairwise range datasets of range values (Δ_{i-j}, Δ_{i-k} and Δ_{j-k}) for KCU^{-40} and KCV^0 impact strength of 13G1S-U steel sheets 8, 10, and 12 mm thick) differed significantly from the original distribution: the experimental values
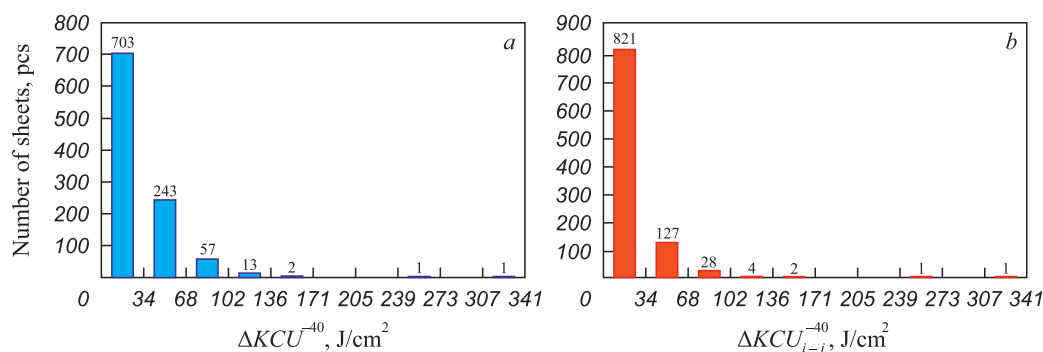

Fig. 3. Histograms of distribution of impact strength range values ΔKCU^{-40} for 13G1S-U steel (sheet thickness – 12 mm), calculated for three (ΔKCU^{-40}) – (*a*) and two (ΔKCU_{i-j}^{-40}) – (*b*) samples (one from each batch of controlled products)

Рис 3. Гистограммы распределения значений размахов ударной вязкости ΔKCU^{-40} стали 13Г1С-У (толщина листа 12 мм), рассчитанные по трем (ΔKCU^{-40}) (*a*) и двум (ΔKCU_{i-j}^{-40}) (*b*) образцам (от каждой партии – единицы контролируемой продукции)

Table 4. **Coefficients of kurtosis E_x and asymmetry A_s for samplings of impact strength ranges $\Delta = KCU(KCV)_{\max} - KCU(KCV)_{\min}$ of 13G1S-U sheet steel for three samples Δ and samplings of possible combinations of pairs (Δ_{i-j-k}) from the same samples**

Таблица 4. **Коэффициенты эксцесса E_x и асимметрии A_s выборок величин размахов ударной вязкости $\Delta = KCU(KCV)_{\max} - KCU(KCV)_{\min}$ листовой стали 13Г1С-У для трех образцов Δ и выборок возможных сочетаний пар (Δ_{i-j-k}) из этих же образцов**

Sheet thickness, mm	Skewness and kurtosis coefficients	KCU^{-40}				KCV^0			
		Δ	Δ_{i-j}	Δ_{i-k}	Δ_{j-k}	Δ	Δ_{i-j}	Δ_{i-k}	Δ_{j-k}
8	A_s	2.02	2.40	2.40	2.55	2.48	3.02	2.69	3.25
	E_x	5.19	8.17	7.53	8.57	7.37	11.28	9.37	13.89
10	A_s	1.20	1.67	1.72	1.45	1.75	1.94	1.76	2.26
	E_x	1.22	3.14	3.17	2.38	4.55	5.11	4.38	7.39
12	A_s	1.25	1.74	1.58	1.70	1.76	2.11	2.05	2.36
	E_x	1.34	3.62	2.56	3.48	3.52	5.72	5.12	7.32

of the Student's and Smirnov's test statistics were no less than 4.61 and 2.77, respectively (with a significance level of $p < 0.0001$).

However, when testing the significance of differences between the pairwise range distributions for KCU^{-40} and KCV^0, full consistency between the two statistical tests was no longer observed. In 9 out of the 18 possible pairwise comparisons (i.e., three combinations – Δ_{i-j}, Δ_{i-k} and Δ_{j-k} – for each of the three sheet thicknesses: 8, 10, and 12 mm, and for both types of impact strength: KCU^{-40} and KCV^0 for 13G1S-U steel), the risk levels for the hypothesis of sample difference varied between 0.22 and 0.50. Statistical equivalence was confirmed in these 9 comparisons; in the remaining 9, the samples were found to differ in 6 and 3 cases at risk levels not exceeding 0.20 and 0.30, respectively (according to at least one of the two tests), regardless of sheet thickness or type of impact strength.

Relative to the impact strength values initially obtained from two samples, the value from the third sample may occupy a different position along the corresponding impact strength axis – not only to the left or right of the minimum and maximum values, respectively, but also between them (if they are not equal), for example, to the left or right of the median value calculated from the initial pair of test results).

The positional statistics of the impact strength values of the third sample KCU_k^{-40} (relative to the corresponding values of the sample pairs KCU_i^{-40} and KCU_j^{-40}) exhibited a fairly typical pattern. For example, in the case of 13G1S-U steel sheets 12 mm thick, the proportion of KCU_k^{-40} values falling below and above the limits of the impact strength interval defined by individual sample pairs $\left[KCU_i^{-40}; \ KCU_j^{-40}\right]$ was approximately the same (Fig. 4, *a*). It is evident that even a slight deviation beyond the interval boundaries can significantly affect the median-based statistics when they are transformed into mean values – either upward or downward. When these deviations are more substantial, their influence becomes even more pronounced. the third sample's value fell within the pairwise impact strength interval $\left[KCU_i^{-40}; \ KCU_j^{-40}\right]$ in 373 to 402 cases (36.5 – 39.4 % of all batches); of these, 180 to 184 values were lower than the corresponding median values $\tilde{x}_i = \dfrac{KCU_i^{-40} + KCU_j^{-40}}{2}$, and 175 to 202 were higher (Fig. 4, *b*). Exact matches with the median values were observed in only 1.6 to 1.8 % of cases, effectively converting the median into the batch mean. All other values, to varying degrees, influenced the batch-level impact strength estimate. This highlights the potential risks when reversing the process – i.e., reducing the number of samples used to assess batch impact strength from three to two.

However, in steels with pronounced structural heterogeneity, even testing three samples may not guarantee

objective quality assessments – particularly with regard to impact strength. This is especially true for steels retaining cast structure (e.g., large forgings made of heat-treatable steel such as 38KhN3MFA-Sh or 15Kh2NMFA) and for high-strength rolled steels with ferrite–pearlite or ferrite–bainite banded microstructures [7; 20 – 22]. The presence of pronounced morphological heterogeneity – including non-metallic inclusions (NMIs) – both between samples and within individual samples, leads to a wide spread of impact strength values across the full range of test temperatures. This introduces uncertainty in the evaluation of toughness, including the determination of cold resistance.

In this context, the use of micro-samples appears promising: if their dimensions are comparable to the scale of structural heterogeneity, it becomes possible to evaluate the cold resistance of individual structural components and rank them by associated fracture risk, based on brittle fracture energy determined through acoustic emission measurements [7]. This is important for understanding the causes of toughness variability observed under standard testing protocols. It was precisely this approach that demonstrated that in large forgings of 38KhN3MFA-Sh heat-treatable steel, brittle fracture within the –130 to 100 °C range occurs only in the interdendritic regions, whereas below –130 °C, the dendrite axes themselves also undergo brittle fracture. Variations in dendritic structure patterns from one impact sample to another – including their downstream effects on microstructure and the morphology of non-metallic inclusions (NMIs) [1] – lead to increased scatter in toughness values at all test temperatures and reduce the reliability of cold resistance predictions.

For 15Kh2NMFA steel, such tests have made it possible to clarify the temperature range of the ductile-to-brittle transition and relate it to the fracture mechanism (transcrystalline, intergranular, or mixed). This is crucial for assessing potential cold resistance degradation during long-term operation when the number of witness samples is limited, as well as under conditions with a small number of samples available during acceptance testing.

The use of this approach proved valuable not only for evaluating the cold resistance of structural components but also for assessing the risk associated with structural anomalies. For instance, in high-strength pipe steels of strength grade K65, micromechanical testing made it possible to localize fracture within an extended interface region of the metal, accompanied by the formation of large facets approximately 500 μm in diameter each. This was identified as one of the possible causes of delamination (slate fracture appearance [20]), and the fracture energy was evaluated based on acoustic emission measurements [7].

Overall, the findings indicate that several factors – such as the varying statistical nature of property distributions in acceptance test results (particularly within
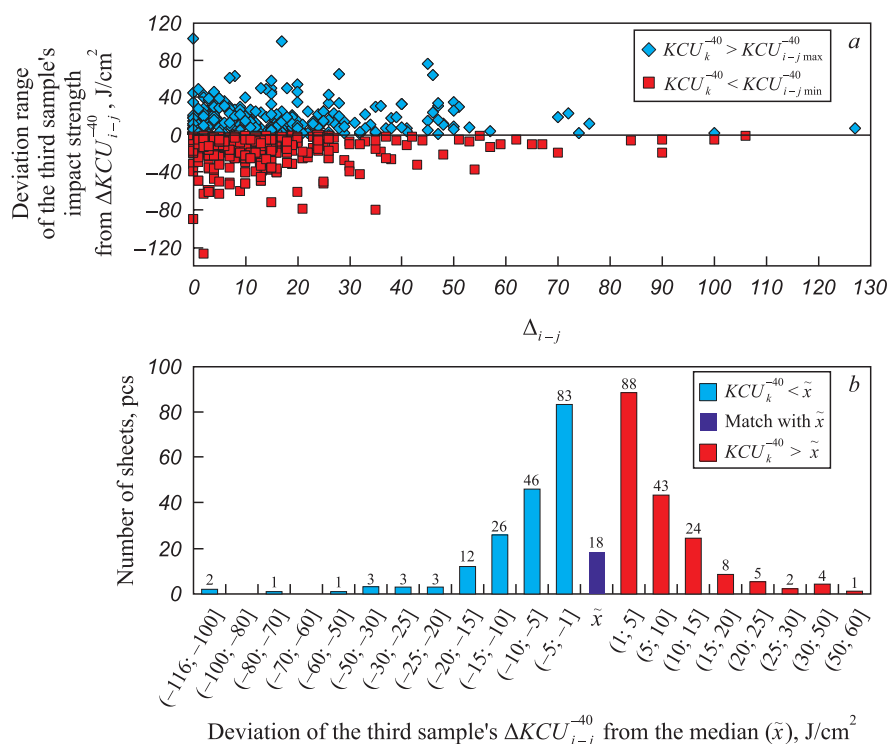
Fig. 4. Distribution of deviations of impact strength values of the third sample KCU_k^{-40} relative to the impact strength values obtained from testing two samples KCU_i^{-40} and KCU_j^{-40} (for samplings $\{KCU_i^{-40}\}$ and $\{KCU_j^{-40}\}$) – outside their value range (*a*)

and within this range – relative to the median value $\tilde{x}_i = \dfrac{KCU_i^{-40} + KCU_j^{-40}}{2}$ (*b*), 13G1S-U steel sheet thickness – 12 mm

Рис. 4. Распределение отклонений значений ударной вязкости третьего образца KCU_k^{-40} относительно значений ударной вязкости, полученных при испытании двух образцов KCU_i^{-40} и KCU_j^{-40} (для выборок $\{KCU_i^{-40}\}$ и $\{KCU_j^{-40}\}$) – вне интервала их значений (*a*)

и в пределах этого интервала – относительно медианного значения $\tilde{x}_i = \dfrac{KCU_i^{-40} + KCU_j^{-40}}{2}$ (*b*), сталь 13Г1С-У, толщина листа 12 мм

the examined chronological sequences), differences in the extent to which property variability is captured depending on the number of samples tested per product unit, the potential for conflicting outcomes in hypothesis testing depending on the statistical criteria applied, and the diversity of mechanisms of technological inheritance embedded in standard production processes (many of which remain insufficiently studied) [1; 7] – will inevitably and substantially constrain the effective application of modern software solutions in the development of end-to-end quality management systems for metal production.

Advancement in this field must be based on a deep understanding of structural and defect evolution throughout the entire technological chain, as well as the development of digital tools for quantitative analysis of microstructures and fracture surfaces, with their integration into industrial practice to enhance the completeness of product quality certification. It also requires the application of statistical procedures that take into account the statistical characteristics of the material under investigation, the identification of domains governed by dominant types of dependency (within the technological parameter space), and the evaluation of their combined effects [7; 15; 19].

CONCLUSIONS

Based on statistical analysis of representative production control datasets for the manufacturing processes of 13G1S-U steel sheets (8, 10, and 12 mm thick) and large forgings made from heat-treatable 38KhN3MFA-Sh steel, several factors have been identified that account for discrepancies in quality assessment outcomes during acceptance testing when different numbers of samples are used per product unit (batch or forging). These factors include: variation in the recorded property range (i.e., the spread of values); changes in the statistical characteristics of property value distributions (as reflected by variations skewness and kurtosis coefficients); and the size of the analyzed dataset.

It has been shown that, in the statistical analysis of production control databases, proper data preprocessing plays an important role in eliminating side effects that reduce the informativeness of acceptance testing results. Such preprocessing should aim to remove the influence of trends, seasonal fluctuations, outliers, and similar factors.

It was found that reducing the number of samples per batch from three to two in the acceptance testing of 13G1S-U steel sheets results in a 17 – 20 % increase

in the frequency of minimum impact strength range (Δ) values (0 to 34 J/cm^2) and a 2.0 – 3.5-fold decrease in the number of ranges within the 35 to 136 J/cm^2 interval. This leads to a distorted assessment of toughness heterogeneity in the steel. For 38KhN3MFA-Sh steel, even greater distortions can be expected due to the pronounced heterogeneity in structural morphology (including dendritic patterns, microstructural features, and nonmetallic inclusions).

When the number of tests per product unit varies – resulting in changes in the distribution pattern of quality indicators within the sample as a whole – the application of statistical hypothesis testing (e.g., the Student's and Smirnov's tests) may yield inconsistent results across different sample sets. This factor should be carefully considered when using modern software tools (such as big data analytics, machine learning, and related technologies) for retrospective analysis of production control databases in metallurgy.

■ References / Список литературы

1. Steel on the Threshold of Centuries. Karabasov Yu.S. ed. Moscow: MISiS; 2001:445–543. (In Russ.).

   Сталь на рубеже столетий / Под ред. Ю.С. Карабасова. Москва: МИСиС; 2001:445–543.

2. Pan G., Wang F., Shang C., Wu H., Wu G., Gao J., Wang S., Gao Z., Zhou X., Mao X., Advances in machine learning- and artificial intelligence-assisted material design of steels. *International Journal of Minerals, Metallurgy and Materials*. 2023;30(6):1003–1024. https://dx.doi.org/10.1007/s12613-022-2595-0

3. Wei J., Chu X., Sun X.Y., Xu K., Deng H.X., Chen J., Wei Z., Lei M. Machine learning in materials science. *InfoMat*. 2019;1(3):338–358. https://doi.org/10.1002/inf2.12028

4. Sandhya N., Sowmya V., Bandaru C.R., Raghu Babu G. Prediction of mechanical properties of steel using data science techniques. *International Journal of Recent Technology and Engineering*. 2019;8(3):235–241. https://doi.org/10.35940/ijrte.C3952.098319

5. Guo S., Yu J., Liu X., Wang C., Jiang Q. A predicting model for properties of steel using the industrial big data based on machine learning. *Computational Materials Science*. 2019;160:95–104. https://doi.org/10.1016/j.commatsci.2018.12.056

6. Sitek W., Trzaska J. Practical aspects of the design and use of the artificial neural networks in materials engineering. *Metals*. 2021;11(11):1832. https://doi.org/10.3390/met11111832

7. Kudrya A.V., Sokolovskaya E.A. Prediction of the destruction of materials with inhomogeneous structures. *Physics of Metals and Metallography*. 2022;123:1253–1264. https://doi.org/10.1134/S0031918X22601615

   Кудря А.В., Соколовская Э.А. Прогноз разрушения материалов с неоднородной структурой. *Физика металлов и металловедение*. 2022;123(12):1334–1346. https://doi.org/10.31857/S0015323022601088

8. Azimi S.M., Britz D., Engstler M., Fritz M., Mücklich F. Advanced steel microstructural classification by deep learning methods. *Scientific Reports*. 2018;8:2128. https://doi.org/10.1038/s41598-018-20037-5

9. Gerasimova L.P., Golubkov D.E., Guk Yu.P. Standard Methods for Quality Control of Metallic Materials, Welded and Soldered Joints. Moscow: Infra-Inzheneriya; 2024:668. (In Russ.).

   Герасимова Л.П., Голубков Д.Е., Гук Ю.П. Стандартные методы контроля качества металлических материалов, сварных и паяных соединений. Москва: Инфра-Инженерия; 2024:668.

10. Kudrya A.V., Sokolovskaya E.A., Kodirov D., Bosov E.V., Kotishevskiy G.V. On necessity of taking into account statistical nature of the objects using Big Data in metallurgy. *CIS Iron and Steel Review*. 2022;(1):105–112. https://doi.org/10.17580/cisisr.2022.01.19

11. Tripathi M.K., Kumarb R., Tripathib R. Big-data driven approaches in materials science: A survey. *Materials Today: Proceedings*. 2020;26(2):1245–1249. https://doi.org/10.1016/j.matpr.2020.02.249

12. Chentsov N.N. Statistical Decisive Rules and Optimal Conclusions. Moscow: Nauka; 1972:524. (In Russ.).

    Ченцов Н.Н. Статистические решающие правила и оптимальные выводы. Москва: Наука; 1972:524.

13. Gmurman V. E. Probability Theory and Mathematical Statistics. Moscow: Vysshaya shkola; 2003:479. (In Russ.).

    Гмурман В.Е. Теория вероятностей и математическая статистика. Москва: Высшая школа; 2003:479.

14. Shtremel' M.A. Engineer in the Laboratory. Moscow: Metallurgiya; 1983:128. (In Russ.).

    Штремель М.А. Инженер в лаборатории. Москва: Металлургия; 1983:128.

15. Kudrya A.V., Shabalov I.P., Velikodnev V.Ya., Sokolovskaya E.A., Akhmedova T.Sh., Vasil'ev S.G. Possibilities of statistical analysis of acceptance test results for determining the scale of pipe steel quality inhomogeneity. *Metallurgist*. 2018;62: 1167–1172. https://doi.org/10.1007/s11015-019-00769-z

    Кудря А.В., Шабалов И.П., Великоднев В.Я., Соколовская Э.А., Ахмедова Т.Ш., Васильев С.Г. Возможности статистического анализа результатов приемо-сдаточных испытаний для определения масштабов неоднородности качества трубных сталей. *Металлург*. 2018;(11):64–68.

16. Chang Y., Haase C., Szeliga D., Madej L., Hangen U., Pietrzyk M., Bleck W. Compositional heterogeneity in multiphase steels: Characterization and influence on local properties. *Materials Science and Engineering: A*. 2021;827:142078. https://doi.org/10.1016/j.msea.2021.142078

17. Klein D.V., Faleskog J. Influence of heterogeneity due to toughness variations on weakest-link modeling for brittle failure. *Engineering Fracture Mechanics*. 2023;292:109643. https://doi.org/10.1016/j.engfracmech.2023.109643

18. Bolshev L.N., Smirnov N.V. Tables of Mathematical Statistics. Moscow: Nauka; 1965:464. (In Russ.).

    Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. Москва: Наука; 1965:464.

19. Bosov E.V., Kodirov D.F., Sokolovskaya E.A., Kudrya A.V. Evaluation of cold resistance of large forgings made of 38KhN3MFA-Sh improved steel based on «data mining» of industrial control of process and product. *Deformatsiya i razrushenie materialov*. 2025;(4):29–39. (In Russ.). https://doi.org/10.31044/1814-4632-2025-4-29-39

Босов Е.В., Кодиров Д.Ф., Соколовская Э.А., Кудря А.В. Оценка хладостойкости крупных поковок из улучшаемой стали 38ХН3МФА-Ш на основе «раскопок данных» производственного контроля процесса и продукта. *Деформация и разрушение материалов*. 2025;(4):29–39. https://doi.org/10.31044/1814-4632-2025-4-29-39

20. Arabei A.B., Pyshmintsev I.Yu., Shtremel' M.A., Glebov A.G., Struin A.O., Gervas'ev A.M. On the structural causes of shingled brittle fractures in plate steel. *Izvestiya. Ferrous Metallurgy*. 2009;52(9):9–15. (In Russ.).

Арабей А.Б., Пышминцев И.Ю., Штремель М.А., Глебов А.Г., Струин А.О., Гервасьев А.М. О структурных причинах шиферности вязких изломов толстолистовой стали. *Известия вузов. Черная металлургия*. 2009;52(9):9–15.

21. Efron L.I. Metal Science in "Big" Metallurgy. Pipe Steels. Moscow: Metallurgizdat; 2012:696. (In Russ.).

Эфрон Л.И. Металловедение в «большой» металлургии. Трубные стали. Москва: Металлургиздат; 2012:696.

22. Gurovich B.A., Kuleshova E.A. Reactor pressure vessel steels: Structure, properties, radiation embrittlement. *Materialovedenie*. 1999;(11):33–45. (In Russ.).

Гурович Б.А., Кулешова Е.А. Стали корпусов ядерных реакторов: структура, свойства, радиационное охрупчивание. *Материаловедение*. 1999;(11):33–45.

Information about the Authors / Сведения об авторах

Elina A. Sokolovskaya, Cand. Sci. (Eng.), Scientific Secretary of the UMO Council for Education in Metallurgy, Assist. Prof. of the Chair "Metallography and Physics of Strength", National University of Science and Technology "MISIS"
ORCID: 0000-0001-9381-9223
E-mail: sokolovskaya@misis.ru

Egor V. Bosov, Analyst of the Project Management Office, National University of Science and Technology "MISIS"
ORCID: 0009-0001-5225-6070
E-mail: bosov.e@misis.ru

Aleksandr V. Kudrya, Dr. Sci. (Eng.), Prof. of the Chair "Metallography and Physics of Strength", National University of Science and Technology "MISIS"
ORCID: 0000-0002-0339-2391
E-mail: AVKudrya@misis.ru

Davronjon F.U. Kodirov, Cand. Sci. (Eng.), Assistant of the Chair "Metallography and Physics of Strength", National University of Science and Technology "MISIS"
ORCID: 0009-0003-5380-5558
E-mail: kodirov.df@misis.ru

Valerii I. Alekseev, Postgraduate, Engineer of the Chair "Metallography and Physics of Strength", National University of Science and Technology "MISIS"
E-mail: kiilobit@edu.misis.ru

Элина Александровна Соколовская, к.т.н., ученый секретарь Совета УМО по образованию в области металлургии, доцент кафедры металловедения и физики прочности, Национальный исследовательский технологический университет «МИСИС»
ORCID: 0000-0001-9381-9223
E-mail: sokolovskaya@misis.ru

Егор Владимирович Босов, аналитик подразделения Офиса управления проектами, Национальный исследовательский технологический университет «МИСИС»
ORCID: 0009-0001-5225-6070
E-mail: bosov.e@misis.ru

Александр Викторович Кудря, д.т.н., профессор кафедры металловедения и физики прочности, Национальный исследовательский технологический университет «МИСИС»
ORCID: 0000-0002-0339-2391
E-mail: AVKudrya@misis.ru

Давронжон Фарходжон Угли Кодиров, к.т.н., ассистент кафедры металловедения и физики прочности, Национальный исследовательский технологический университет «МИСИС»
ORCID: 0009-0003-5380-5558
E-mail: kodirov.df@misis.ru

Валерий Игоревич Алексеев, аспирант, инженер научного проекта кафедры металловедения и физики прочности, Национальный исследовательский технологический университет «МИСИС»
E-mail: kiilobit@edu.misis.ru

Contribution of the Authors / Вклад авторов

E. A. Sokolovskaya – preparation of the final text, discussion of results.

E. V. Bosov – conducting statistical analysis of production control databases, constructing graphical dependencies, participation in the article preparation.

A. V. Kudrya – conceptualization, justification for conducting the research, participation in the article preparation.

D. F. Kodirov – analysis of the state of the issue based on literary sources, statistical analysis of results.

V. I. Alekseev – assistance in structuring production control databases, conducting tests to assess cold resistance on samples with dimensions comparable to the scale of structural inhomogeneity of structural steel.

Э. А. Соколовская – подготовка окончательного текста статьи, обсуждение полученных результатов.

Е. В. Босов – проведение статистического анализа баз данных производственного контроля, построение графических зависимостей, участие в подготовке статьи.

А. В. Кудря – концепция статьи, обоснование проведения исследования, участие в подготовке текста статьи.

Д. Ф. Кодиров – анализ состояния вопроса по литературным источникам, статистический анализ полученных результатов.

В. И. Алексеев – помощь в структурировании баз данных производственного контроля, проведение испытаний по оценке хладостойкости на образцах, габариты которых сопоставимы с масштабом структурной неоднородности конструкционной стали.