

10. Modeling of MnS precipitation during the crystallization of grain oriented silicon steel / D. Kalisz [et al.] // Metalurgija. – 2015. – № 54(1). – Pp. 139-142.

Рецензент: В.А. Маслов
д-р техн. наук, проф., ГВУЗ «ПГТУ»

Статья поступила 18.04.2016

УДК 669.184:519.22

© Бондарь В.И.*

ИССЛЕДОВАНИЕ ЗНАЧИМОСТИ РАЗЛИЧИЯ СРЕДНИХ ХИМИЧЕСКОГО СОСТАВА РАСПЛАВА И ЕГО ТЕМПЕРАТУРЫ В УСЛОВИЯХ LD-ПРОЦЕССА

На основе данных о химическом составе и температуре металлического расплава, полученного в условиях кислородного конвертирования с использованием пакета прикладных программ StatSoft Statistica 8.0, выполнен анализ значимости различия средних для конвертеров с верхней продувкой.

Ключевые слова: статистическая значимость, нормальность распределения, t-критерий, дисперсия, статистика Левена, критерий Манна-Уитни.

Бондар В.И. Дослідження значущості відмінності середніх хімічного складу розплаву і його температури в умовах LD-процесу. На основі даних про хімічний склад і температуру металевого розплаву, що одержаний в умовах кисневого конвертування з використанням пакету прикладних програм StatSoft Statistica 8.0, виконаний аналіз відмінності середніх для конвертерів з верхньою продувкою.

Ключові слова: статистична значимість, нормальність розподілу, t-критерій, дисперсія, статистика Левена, критерій Манна-Уітні.

V.I. Bondar. Study of the significance of differences of average chemical composition of the melt and its temperature in terms of the LD-process. Development of modern metallurgical production and engineer requires the scalene study of process of thermodynamics and kinetics of process of the oxygen converting. Creation of the determined model of the oxygen-converter melting is not possible, a statistical method was therefore used. The aim of the real research was determination of meaningfulness of distinction of mean values of chemical composition and temperature of metallic fusion of converter redistribution with the use of t - criterion for independent selections. For realization of statistical analysis the programmatic package of STATISTICA 8 was used. As a method of verification of statistical hypothesis about equality of middle two selections, the criteria of self-reactance and non-parametric statistics were used. Arrays of data, that analyses, were presented 88th melting from the first converter and 93th melting got on the second converter. It is set a calculation, that diagrams of scope for variables (T, °C), [C], [S], [P] answer distribution of variables under the law of normal distribution. At the same time distribution of variables [Mn] and a_{10} , appraised by the same method, it is fallen short of the law of normal distribution. It is set that for variables (T, °C) and [C] a hypothesis about equality of dispersions cannot be accepted. Calculation of comparison middle conducted on methodology of values of t - criterion with the separate estimations of dispersions. The results of calculation show that mean values of variable [C] meaningfully differ for fusions of both converters: 0,044 and 0,037% accordingly. At the same time mean values of variable (T, °C) meaningfully not distinguished. It is shown that hy-

* канд. техн. наук, доцент, ГВУЗ «Приазовский государственный технический университет», г. Мариуполь, bbvii.47@gmail.com

pothesis of equality of dispersions for variables [P] and [S] confirmed: the values of variables belong to the same general aggregates. Values middle for these variables: 0,0079 and 0,0085% for a variable [P], and for a variable [S] 0,0134 and 0,0128% accordingly for the first and second converters and does not differentiate meaningfully. The use of criterion Mann-Whitney for the case of the use of method of non-parametric statistics specifies on absence of meaningful distinctions in values middle for a variable a_{10} . Nevertheless, distinctions middle for a variable [Mn] appeared meaningful.

Keywords: *statistical significance, normal distribution, t-test, variance, stats Leuven, Mann-Whitney test.*

Постановка проблеми. Развитие современного металлургического производства и машиностроения требует разностороннего изучения процесса термодинамики и кинетики процесса кислородного конвертирования. Создание детерминированной модели кислородно-конвертерной плавки пока не представляется возможным ввиду термодинамической ее неустойчивости. Весьма эффективным способом исследования подобных неравновесных систем является статистический метод.

Анализ последних исследований и публикаций. Современные научные исследования и производственная практика потребовали широкого применения математической статистики для анализа закономерностей массовых явлений во всех отраслях промышленности [1, 2].

Металлургическая промышленность несколько запоздала с внедрением методов математической статистики; в основном, их стали применять в последние десятилетия. Впервые эти методы использовали для анализа производственного процесса. Задача применения методов математической статистики обосновывается еще и тем, что современное развитие металлургии достигло такой стадии, при котором невозможно ограничиться только качественным анализом. Последний необходимо дополнить количественным описанием процесса, которое позволит в совершенстве изучить его, а, следовательно, и надежно управлять им. Эти обстоятельства являются одной из основных предпосылок для автоматизации металлургического производства.

Особого внимания в этой области заслуживают работы В.И. Кошелева, В.Н. Корнфельда, А.О. Войтова, Л.С. Штейнберга, в которых проводился анализ зависимостей между регулируемые и регулируемыми факторами в различные периоды плавки в мартеновских печах [3].

В настоящее время большое внимание уделяют анализу распределения случайных переменных, статистическим оценкам и проверкам гипотез, корреляционному и регрессионному методам анализа, а также некоторым другим методам статистической индукции [4-6].

Цель статьи – изучить значимость различия средних значений химического состава и температуры металлического расплава конвертерного передела с использованием t-критерия для независимых выборок.

Изложение основного материала. Создание детерминированной модели кислородно-конвертерной плавки не представляется возможным из-за термодинамической неустойчивости системы. Вследствие этого «судьбу» системы могут решать даже очень малые возмущения, часто выходящие за рамки экспериментального контроля [7]. В этой связи проявляется одно из наиболее глубоких следствий неравновесной термодинамики – дуалистичность необратимого процесса. В одном случае он – разрушитель порядка вблизи равновесия и создатель порядка вдали от равновесия. Для систем, далеких от равновесия, в частности системы «кислородно-конвертерный передел», общие экстремальные принципы равновесной термодинамики не выполняются. Сама система может развиваться непредсказуемо и при одном и том же наборе условий переходить к разным состояниям. Причиной тому могут быть флуктуации, малые неоднородности и другие случайные факторы.

Весьма эффективным способом исследования неравновесных систем является стохастический (статистический) метод. Несмотря на его недостатки, он весьма эффективен в случае анализа процессов в отдельно взятом агрегате, например, в кислородном конвертере, и имеет широкие возможности.

Цель настоящей работы – исследование значимости различия средних химического состава и температуры металлического расплава, произведенного раздельно в двух 350-т конвертерах ПАО МК «Азовсталь», работающих в режиме кислородного конвертирования с верхним дутьем. В качестве метода проверки статистической гипотезы о равенстве средних двух выбо-

рок використовувались критерії параметричної і непараметричної статистики.

Как известно, статистической гипотезой называется любое предположение о виде или параметрах некоторого закона распределения. Проверяемую гипотезу обычно называют нулевой и обозначают H_0 . Наряду с нулевой гипотезой рассматривают альтернативную гипотезу H_1 , являющуюся логическим отрицанием H_0 .

Вероятность совершить ошибку первого рода α , то есть отвергнуть гипотезу H_0 , когда она верна, называется уровнем значимости критерия. Уровень значимости p – это максимально приемлемая для исследователя вероятность отклонить нулевую гипотезу, когда на самом деле она верна. Величина уровня значимости обычно выбирается равной 0,05. Если p меньше либо равно 0,05, то результат статистически значим; если p меньше или равен 0,01, то результат считался статистически высоко значимым. Для проведения статистического анализа использовался программный пакет STATISTICA 8.

Массивы данных, что анализировались, были представлены 88-ю плавками от первого конвертера (Конв_1) и 93-я плавками, полученными на втором (Конв_2) конвертере. Массивы переменных формировались с учетом того, что содержание углерода не превышало 0,07% (масс.)¹, так как плавки с большим содержанием углерода подвергаются додувке. В качестве независимых переменных – предикторов – выбирались следующие: содержание углерода $[C]$, марганца $[Mn]$, фосфора $[P]$, серы $[S]$, активность кислорода $a_{[O]}$ и температура расплава (T , °C). Вводилась также группирующая переменная, с помощью которой кодировались два состояния системы – Конв_1 и Конв_2, соответственно.

На первом этапе анализа была реализована проверка соответствия распределения случайных величин массивов закону нормального распределения. Нормальность распределения переменной – это слишком жесткое требование, поэтому в работе ограничились построением диаграмм размаха (рис. 1-6). Диаграммы размаха позволяют визуально судить о наличии или отсутствии симметричности в распределении каждой из переменных относительно ее медианы. Если симметричность распределения значений переменной относительно медианы подтверждается, то считается, что переменная распределена в соответствии с нормальным законом распределения.

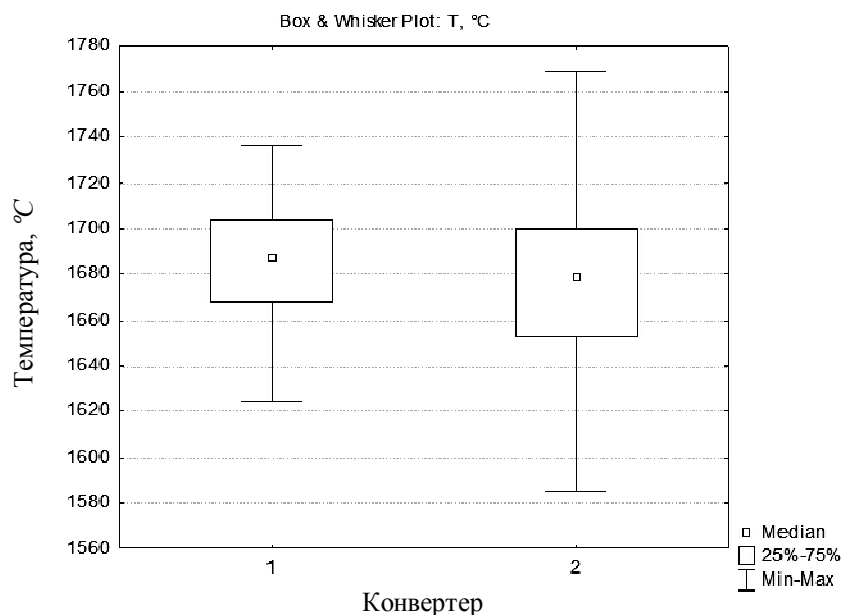


Рис. 1 – Диаграмма размаха для переменной (T , °C) Конв_1 и Конв_2

Относительно переменных (T , °C), $[C]$, $[S]$ и $[P]$ можно утверждать, что для значений группирующей переменной «1» и «2» распределение их случайных значений может быть признанным соответствующим нормальному.

¹ здесь и далее проценты массовые

Діаграми размаха переменных $[Mn]$ и $a_{[O]}$ существенно отклоняются от нормального распределения: как для значения группирующей переменной «1», так и для значения «2», «хвосты» распределения значений этих переменных несимметричны относительно их медиан. В этой связи для дальнейшего анализа использование параметрических критериев для оценки значимости различия средних не представляется возможным.

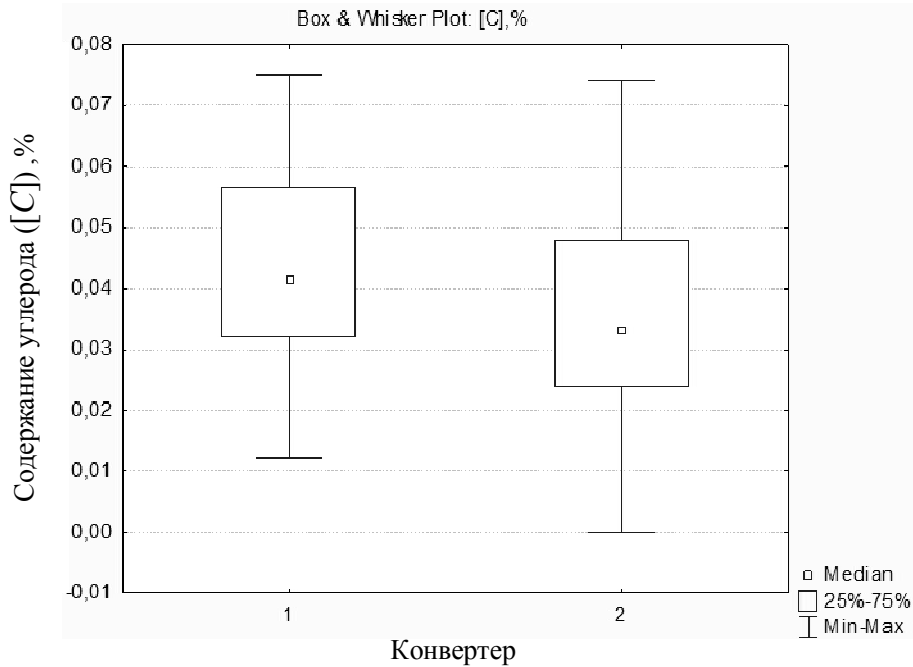


Рис. 2 – Диаграмма размаха для переменной $[C]$ Конв_1 и Конв_2

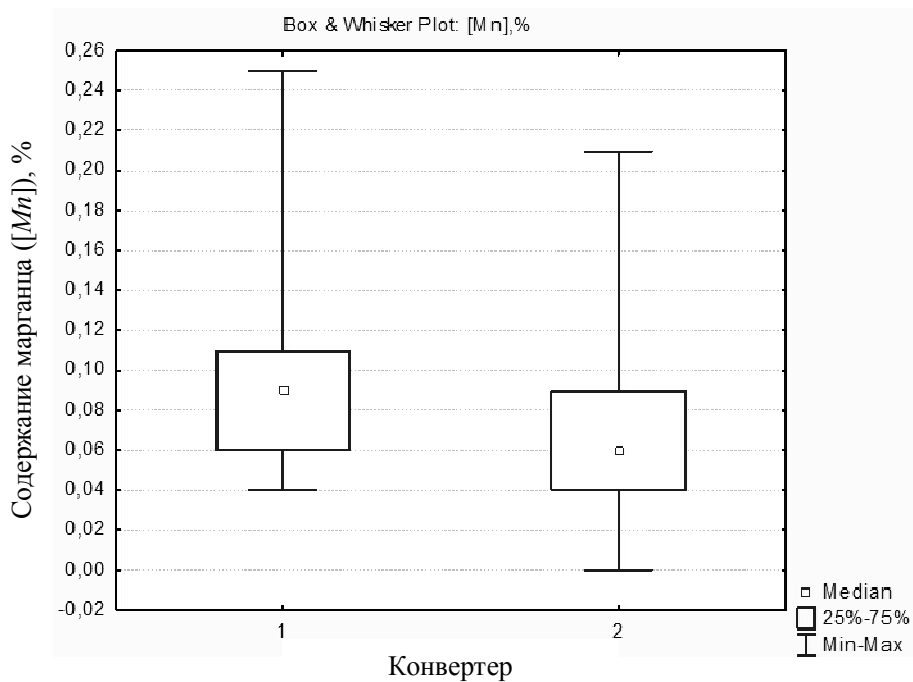


Рис. 3 – Диаграмма размаха для переменной $[Mn]$ Конв_1 Конв_2

Для случая соответствия распределения переменных нормальному распределению – это касается переменных $(T, ^\circ C)$, $[C]$, $[S]$ и $[P]$ – вначале следует установить значимо или незначимо

отличаются дисперсии каждой из переменных для значений «1» и «2» группирующей переменной. С этой целью использовалась статистика Левена. И, если устанавливается, что дисперсии переменных значимо не отличаются, делается вывод о различии средних переменной по величине уровня значимости p .

Статистика Левена для переменной (T , °C) дает значение p -уровня, равное 0,0046, что менее 0,05. Поэтому можно утверждать, что дисперсии выборок в состоянии «1» и «2» существенно отличаются друг от друга. Это подтверждается и сравнением величин стандартного отклонения переменной (T , °C): для состояния группирующей переменной «1» оно равно 25,42, а для состояния «2» – 35,65.

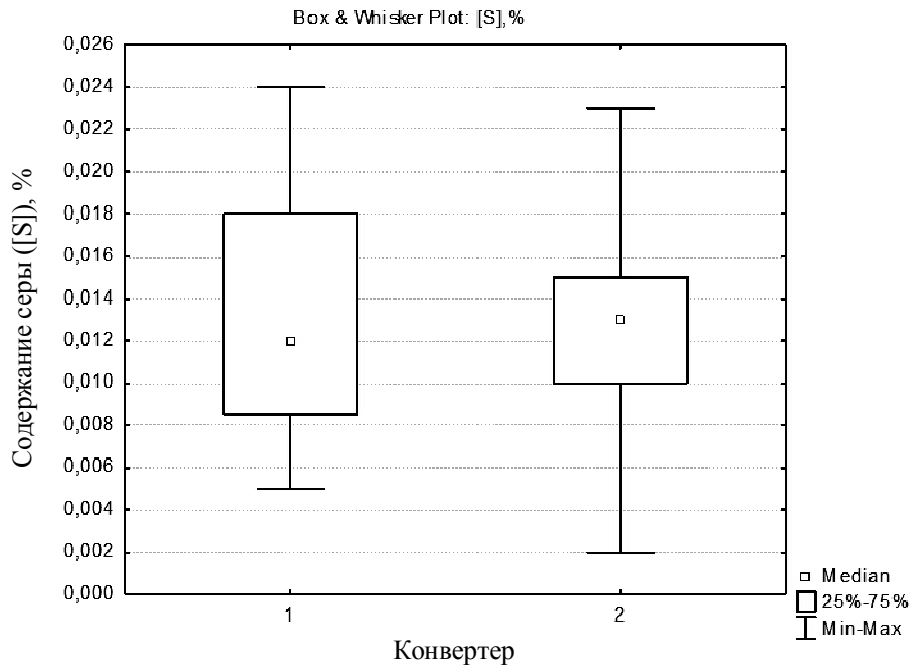


Рис. 4 – Диаграмма размаха для переменной [S] Конв_1 и Конв_2

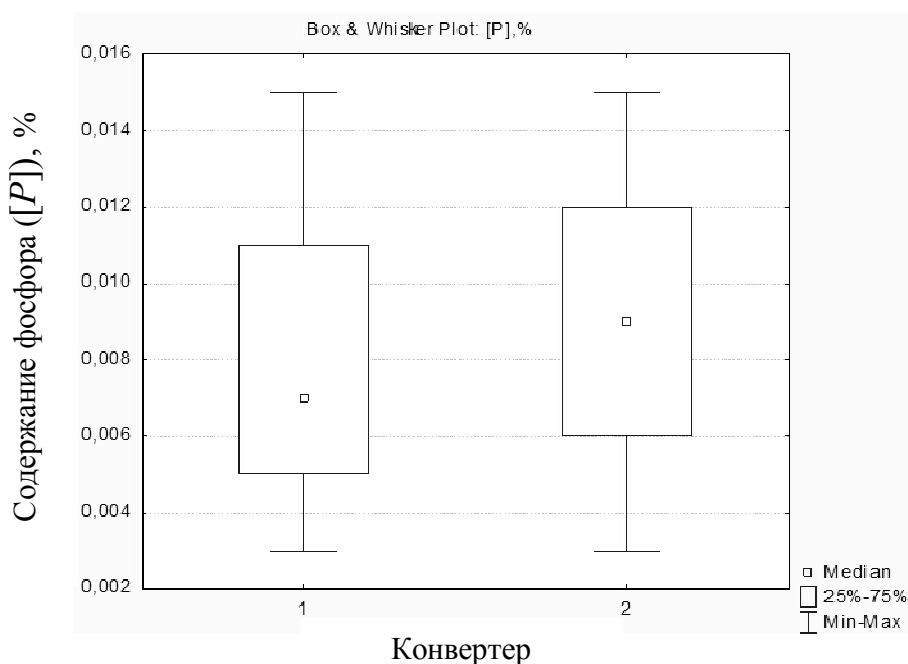


Рис. 5 – Диаграмма размаха для переменной [P] Конв_1 и Конв_2

Диаграмма размаха переменной (T , °C) – дополнительное тому подтверждение. Заслуга теста p -Левена в том, что отклонение дисперсий не является случайным (табл. 1). Статистика Левена для переменной $[C]$ определяет, что дисперсии выборок для состояния «1» и «2» также значимо отличаются: значение p -уровня составляет 0,029, что менее 0,05.

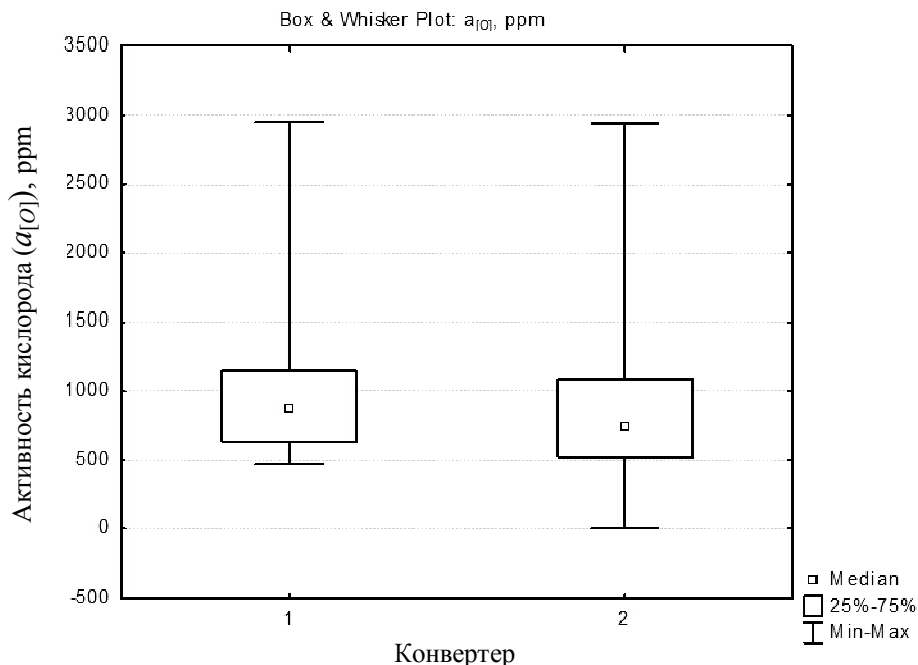


Рис. 6 – Диаграмма размаха для переменной a_{O_2} Конв_1 и Конв_2

Таблица 1

Значения критерия Левена

Переменная	Means_1	Means_2	p	Valid N_1	Valid N_2	Std.Dev._1	Std.Dev._2	p - Leven
(T , °C)	1684,3	1676,7	0,1061	88	93	25,41986	35,65004	0,00462
$[C]$	0,0441	0,0366	0,0029	88	93	0,01447	0,018586	0,02908
$[S]$	0,0134	0,0128	0,4273	88	93	0,005204	0,005018	0,15304
$[P]$	0,0078	0,0085	0,1837	88	93	0,003292	0,003377	0,42021

Значение p -уровня статистики Левена для переменных $[S]$ и $[P]$ для состояния группирующей переменной «1» и «2» значимо не отличаются. Для переменной $[S]$ расчетное значение p -уровня равно 0,153, что превышает 0,05. Для переменной $[P]$ его значение равно 0,4202, что значительно превышает 0,05. В двух последних случаях можно делать вывод о значимости или незначимости различия средних. Значение p -уровня для средних значений переменной $[S]$ составляет 0,4273, что больше, чем 0,05. Отсюда следует вывод о том, что средние значения обсуждаемой переменной значимо не отличаются: 0,0134 и 0,0128% для состояний группирующей переменной «1» и «2» при величине стандартного отклонения 0,052 и 0,0050. Значение p -уровня для переменной $[P]$ составляет, как и в предыдущем случае, величину, превышающую 0,05, а именно, p равно 0,1837. В этой связи корректен вывод о незначимости различия средних переменной для состояния «1» и «2»: 0,0078 и 0,0085% при значениях дисперсий 0,0032 и 0,0034. Таким образом, медианы переменных $[S]$ и $[P]$ признаются значимо не различающимися.

Для анализа значимости различий средних в случае существенного различия дисперсий – переменные (T , °C) и $[C]$ – использовался прием оценки t -критерия с отдельными оценками дисперсий. Относительно переменной $[C]$ расчетное значение p -уровня составило величину 0,0029, что менее 0,05 (табл. 2). Следовательно, средние значения переменной $[C]$ значимо отличаются друг от друга: 0,0441 против 0,0366% при значениях дисперсий 0,0144 и 0,0186.

Таблица 2

Значения t-критерия с раздельными оценками дисперсий

Переменная	Mean_1	Mean_2	<i>p</i>	Valid N-1	Valid N-2
(<i>T</i> , °C)	1684,3	1676,8	0,1061	88	93
[<i>C</i>]	0,0441	0,036656	0,0029	88	93

Касательно переменной (*T*, °C), дисперсии которой для состояния «1» и «2» значимо отличаются друг от друга, расчетное значение *p*-уровня составило величину 0,1061 против 0,05. Следовательно, значение переменной (*T*, °C) для состояний «1» и «2» значимо не отличаются друг от друга, то есть значимого различия в средних 1684,3 и 1676,8°C не существует.

Как уже упоминалось ранее, диаграммы размаха для переменных [*Mn*], и особенно для $a_{[O]}$, существенно отличаются несимметричностью относительного среднего значения. Иначе говоря, распределение переменных [*Mn*] и $a_{[O]}$ не может быть признанным соответствующим нормальному распределению как для значения «1», так и «2». Для анализа значимости различия переменных в подобных случаях использовался аналог *t*-критерия, а именно непараметрический *t*-критерий – критерий Манна-Уитни (Mann-Whitney U-test). Для переменной $a_{[O]}$ уточненное расчетное значение *p*-уровня соответствует значению 0,082, что превышает 0,05 (табл. 3).

Таблица 3

Mann-Whitney Test

Переменная	Rank Sum-Group1	Rank Sum-Group2	U	Z	p-level	Z-adjusted	p-level	Valid N Group1	Valid N Group2
[<i>Mn</i>]	9167,50	7303,50	2932,5	3,2911	0,0010	3,3049	0,00095	88	93
$a_{[O]}$	8621,00	7850,00	3479,0	1,7399	0,0818	1,7400	0,0819	88	93

Таким образом, средние переменной $a_{[O]}$ для значений группирующих переменных «1» и «2» отличаются незначимо или значимое различие в значении переменной $a_{[O]}$ отсутствует. Значения средних составляет 920,48 и 842,74 ppm для значений группирующей переменной «1» и «2», соответственно. Расчетное значение *p*-уровня для критерия Mann-Whitney составляет 0,00098 для переменной [*Mn*], распределение которой не соответствует нормальному. Выполнение условия $0,00098 < 0,05$ указывает на значимость различий средних переменной [*Mn*]. Значения средних для переменной составляет величины 0,0891 и 0,0697 для значений группирующей переменной «1» и «2», соответственно.

В заключение следует отметить, что оба конвертера по ряду характеристик металлического расплава – по-разному работающие агрегаты, независимо от среднестатистического состава шихтовых материалов, объема и чистоты кислорода, израсходованного на конвертирование, и иные технологические характеристики плавки.

Углерод и марганец – окисляющиеся примеси металлического расплава и в заключительной стадии кислородного конвертирования окисление углерода сопровождается, как известно, окислением марганца [8]. Эта тенденция сохраняется для двух конвертеров, но эффективность обезуглероживания металлического расплава и окисление марганца, в нем содержащегося, для второго конвертера выше, чем для первого. Можно предположить, что снабжение расплава кислородом лучше во втором конвертере, чем в первом. Последнее предположение требует дополнительных исследований.

Выводы

1. Расчетом установлено, что диаграммы размаха для переменных ((*T*, °C), [*C*], [*S*], [*P*]) отвечают распределению переменных в соответствии с законом нормального распределения. Вместе с тем распределение переменных [*Mn*] и [*C*], оцененное тем же способом, закону нормального распределения не соответствуют.

2. Установлено, что для переменных (T , °C) и $[C]$ гипотеза о равенстве дисперсий не может быть принята. Расчет сравнения средних проводился по методике значений t -критерия с отдельными оценками дисперсий. Результаты расчета показывают, что средние значения переменной $[C]$ значительно отличаются для расплавов обоих конвертеров: 0,044 и 0,037% соответственно. В то же время средние значения переменной (T , °C) значительно не отличаются.

3. Показано, что гипотеза равенства дисперсий для переменных $[P]$ и $[S]$ подтверждается \square значения переменных принадлежат к одним и тем же генеральным совокупностям. Значения средних для этих переменных: 0,0079 и 0,0085% для переменной $[P]$, а для переменной $[S]$ 0,0134 и 0,0128% соответственно для первого и второго конвертеров и значительно не различаются.

4. Использование критерия Манна-Уитни для случая использования метода непараметрической статистики указывает на отсутствие значимых различий в значениях средних для переменной $a_{[O]}$. Тем не менее, различия средних для переменной $[Mn]$ оказались значимыми.

Список использованных источников:

1. Кнотек М. Анализ металлургических процессов методами математической статистики / М. Кнотек, Р. Войта, И. Шефиц. – М. : Металлургия, 1968. – 212 с.
2. Пустыльник Е.И. Статистические методы анализа и обработки наблюдений / Е.И. Пустыльник. – М. : Наука, 1968. – 288 с.
3. Ковшов В.Н. Постановка инженерного эксперимента / В.Н. Ковшов. – Киев-Донецк : Вища школа, 1982. – 120 с.
4. Корнеева А.А. Непараметрическое моделирование конвертерной плавки / А.А. Корнеева, М.Е. Корнет // Известия ВУЗов. Черная металлургия. – 2013. – № 10. – С. 24-28.
5. StatSoft. Inc. (2007). Электронный учебник по статистике : в 3-х т. [Электронный ресурс]. – М. : StatSoft, 2007. – 2595 с. – (<http://www.StatSoft/home/textbook/default.htm>).
6. Халафян А.А. STATISTICA 6. Статистический анализ данных : учебник / А.А. Халафян. – М. : ООО «Бином-Пресс», 2007. – 512 с.
7. Боровиков В.А. STATISTICA. Искусство анализа данных на компьютере : учебное пособие / В.А. Боровиков. – СПб. : Питер, 2003. – 688 с.
8. Меджибожский М.Я. Основы термодинамики и кинетики сталеплавильных процессов : учебное пособие для ВУЗов / М.Я. Меджибожский. – Киев-Донецк : Вища школа, 1979. – 280 с.

Bibliography:

1. Knotec M. Analysis of metallurgical processes the methods of mathematical statistics / M. Knotec, R. Voyta, I. Shefits. – M. : Metallurgiya, 1968. – 212 p. (Rus.)
2. Pustyl'nik E.I. Statistical methods of analysis and treatment of supervisions / E.I. Pustyl'nik. – M. : Nauka, 1968. – 288 p. (Rus.)
3. Kovshov V.N. Raising of engineering experiment / V.N. Kovshov. – Kiev-Donetsk : Vischa shkola, 1982. – 120 p. (Rus.)
4. Korneeva A.A. Non-parametric design of the converter melting / A.A. Korneeva, M.E. Kornet // Izvestiya VUZov. Chernaya metallurgiya. – 2013. – № 10. – P. 24-28. (Rus.)
5. StatSoft. Inc. (2007). Electronic textbook on statistics : in 3 vols. [Electronic resource]. – M. : StatSoft, 2007. – 2595 p. – (<http://www.StatSoft/home/textbook/default.htm>).
6. Halafyan A.A. STATISTICA 6. Statistical analysis of data : textbook / A.A. Halafyan. – M. : ООО «Binom-Press», 2007. – 512 p. (Rus.)
7. Borovikov V.A. STATISTICA. Art of analysis of data on a computer : textbook / V.A. Borovikov. – SPb. : Piter, 2003. – 688 p. (Rus.)
8. Medzhibozhskiy M.Y. Bases of thermodynamics and kinetics of steel-smelting processes : manual for institutions of higher learning / M.Y. Medzhibozhskiy. – Kiev-Donetsk : Vischa shkola, 1979. – 280 p. (Rus.)

Рецензент: В.А. Маслов
д-р техн. наук, проф., ГВУЗ «ЛГТУ»

Статья поступила 05.05.2016