www.czasopisma.pan.pl PAN www.journals.pan.pl

ARCHIVES
of
FOUNDRY ENGINEERING

DOI: 10.1515/afe-2016-0061

Published quarterly as the organ of the Foundry Commission of the Polish Academy of Sciences

ISSN (2299-2944)
Volume 16
Issue 3/2016

115 – 120

The Use of Modern Statistical Methods to Optimize Production Systems in Foundries

J. Szymszal *, B. Gajdzik, G. Kaczmarczyk
Production Engineering Department, Silesian University of Technology, ul. Krasińskiego 8, 40-019 Katowice, Poland
* Corresponding author. E-mail address: jan.szymszal@polsl.pl

Received 31.03. 2016; accepted in revised form 23.04.2016

Abstract

Conducting reliable and credible evaluation and statistical interpretation of empirical results related to the operation of production systems in foundries is for most managers complicated and labour-intensive. Additionally, in many cases, statistical evaluation is either ignored and considered a necessary evil, or is completely useless because of improper selection of methods and subsequent misinterpretation of the results. In this article, after discussing the key elements necessary for the proper selection of statistical methods, a wide spectrum of these methods has been presented, including regression analysis, uni- and multivariate correlation, one-way analysis of variance for factorial designs, and selected forecasting methods. Each statistical method has been illustrated with numerous examples related to the foundry practice.

Keywords: Computer-aided foundry production, Quality management of castings, Statistical methods, Forecasting

1. Introduction

As demonstrated by practical experience, a reliable estimate, optimization and statistical interpretation of the obtained empirical results linked to the evaluation of functional production systems in foundries is for the majority of foundry staff managing these systems the process complicated and labour-intensive. At the same time, basing on the long experience of the authors, it can be concluded that in many cases statistical evaluation is ignored or completely useless as a result of improper selection of methods and subsequent misinterpretation of the results. This article presents a wide spectrum of modern statistical methods, starting with the selection and analysis of the level of measurement, discussing next the choice of appropriate statistical tests and use of relevant characteristics of descriptive statistics, and ending in the description with guidelines for the application of regression analysis, uni- and multivariate correlation, one-way analysis of variance for factorial designs, and several forecasting methods. The selected statistical methods have been illustrated with specific examples of casting practice and were presented during the 15th Scientific Conference held last year.

2. Elements of proper selection of the statistical methods

Conducting statistical analysis it is necessary to determine first the number of tests and/or the series of measurements, and choose the type of measurement scales, in which the analyzed data will be expressed. It is also necessary to determine the relationship between the measurement scales and the type of random variable, which may be discontinuous (discrete) or continuous [1-2]. The outcome of performed studies is one or several measurement series, where each of the values obtained is the realization of a random variable. Practical experience shows us that the diagnosis of the type of the analyzed random variable may create serious difficulties. Variables of the continuous type occur in the measurements of time, density, length, area, volume, pressure, concentration, etc., that is, the quantities which may take any value from a given range. Since measurements are made using a variety of devices, instruments or apparatus, there is a problem of accuracy with which it is possible to determine the value of the examined feature, i.e. a random variable. Of some significance is the order of magnitude related with the measurement. An example here is the measurement of distance expressed in kilometres or microns. When the cost is determined as a resultant feature, it is expressed in the units of a two-tier monetary system used in our country, and so there is one *zloty* and its equivalent equal to one hundred groshes. It is impossible, therefore, to determine the monetary value with an accuracy more precise than one grosh, and so the conclusion is that this measurable feature can not be classified as a continuous variable.

There are many variables which without any doubt belong to the family of continuous variables, but the available accuracy of measurements is of the order such that the obtained values resemble rather the variables of a discontinuous type. The examples of such variables are quite numerous, to mention only the measurement of the intensity of a signal, which assumes the values comprised in a given range, but its recording by conventional apparatus is done with a certain accuracy only. This has some consequences when it comes to the estimation of the characteristics of descriptive statistics (e.g. average or variance), since the discontinuous nature of the measurements determines the discontinuous nature of the estimators, and this has a significant impact on the quality of inference using parametric tests [1-3].

The discontinuous (discrete) variable adopts strictly determined values with probabilities that result from its probability distribution. Here it is easy to notice that the set of all the values of the discrete variable does not always belong to the group of finite (countable) sets. Among the examples of the discrete variables one can mention the popular Poisson distribution, which from a theoretical point of view is used to describe a discontinuous variable having an infinite number of values, and as such is frequently used to describe the number of occurrences of a specific event in a certain period of time [2]. As an example of the discrete variable one can mention the numbers drawn in a lottery, the number of pieces in a batch of castings which needs to be checked, the number of structural constituents which appear in the field of view during image analysis, etc.

When the number of the observable values of the discrete random variable is relatively large, it is assumed that its distribution has certain characteristics typical of the distribution of a continuous random variable [2]. As an example can be given a universal in practical applications way to replace the binomial distribution (discontinuous variable) with normal distribution (continuous variable), provided that appropriate relationships are met [1-3].

A similar phenomenon of replacing with the normal distribution other distributions of discrete random variables occurs in the case of distributions of the sum of ranks, to mention as an example non-parametric Mann-Whitney and Wilcoxon tests [1]. In some cases, also distributions based on the sum of the squares of rank statistics in non-parametric Kruskal-Wallis and Friedman tests may be replaced by χ^2 distribution [2-3].

The basic condition for the applicability of parametric tests is compliance of distribution of the examined resultant feature with the normal distribution, and hence its continuity [2]. These limitations significantly reduce the possibility of verification of statistical hypotheses based on the results of parametric tests. However, as practical experience shows us, while the requirement for compliance of the distribution of the examined resultant feature with the normal distribution is obvious, the requirement for continuity of the examined random variable is often an excessive request. So, in the case of a sufficiently large resolution of the discontinuous random variable with distribution similar to a normal distribution, confirmed by relevant tests of compliance, the use of parametric tests should not raise major doubts. Thus, as recommended by some of the scientists dealing with statistical methods, a principle has been adopted that in taking a decision on the application of the chosen method, less attention will be devoted to studies of the continuity or discontinuity of resulting features, focusing attention rather on the type of measurement scales in which these features will be expressed. The basic condition for conducting proper statistical analysis is therefore the distinguishability and a good knowledge of the measurement scales: qualitative (nominal), ordinal, interval (differential) and the most powerful of all scales – the ratio scale [2-3].

Parametric tests are applicable when the values of the examined resultant feature are expressed in a differential or ratio scale with positive verification of the null hypothesis assuming normality of the distribution, that is, in the absence of any grounds to reject it. Otherwise, it is necessary to use appropriate non-parametric test, which should be a substitute for the parametric one. In the case when the values of resulting features are expressed in nominal or ordinal scale, only non-parametric tests are recommended for use.

Accurate diagnosis of the number of series and measurement scales in which the data is expressed is a second, next to the content of the problem solved, essential element deciding about the right choice of statistical method. When the case of one measurement series is considered, and we can be sure that there is only one test to deal with, the choice of a statistical method is determined, on the one hand, by the type of measurement scales in which the analyzed data is expressed, and on the other hand, by the content of the problem (task), which should indicate the type of hypotheses to be taken into consideration. Two series of measurements can represent equally well two tests or one test. Two series represent two tests only when their values refer to the same resultant feature, which means that they are expressed in the same measurement scale. Then, these are the two independent tests. Some indication may be in this case the unequal number of measurements in both runs (tests). If two series are composed of an equal number of measurements of the same resultant feature, the diagnosis if independent tests are dealt with must be based on a sound analysis of the content of the task. In the situation when there are two series of measurements and it is known that only one test has been made, it is clear that both series consist of an equal number of measurements. In this case, the choice of an appropriate statistical method will depend on the result of an investigation, if the series express the values of the same resultant feature or of two different features. In the former case it will probably be necessary to show the differences between these series, in the latter case - to establish some relationships or correlations. The most important factor that supports both the diagnosis and the correct choice of statistical method is analysis of the content of the problem, while details in this respect depend primarily on the measurement scales expressing the values of both data series.

Practical experience teaches us that analysis of more than two series of measurements requires substantially the same reasoning as the reasoning used in the case of two series. The first step is to investigate whether the examined series represent independent tests (with the same resultant feature), or one test. In the latter case, the test components can be described with several series of measurements of the same resultant feature or with several series of measurements of different resultant features. Proper diagnosis will depend on the content of the problem (task), and helpful in this respect can be equal or unequal number of all the measurement series. It is obvious that the statistical tools chosen for the two series of measurements will differ from the tools that are used for more than two series.

3. Statistical methods selected for the optimization of foundry production

The following chapters discuss the selection and use of different statistical tools to manage the production of castings.

3.1. Studies of the effect of W and Mo content on the properties of silumin 226

In [4], the authors describe the effect of the addition of W and Mo on the properties of silumin 226 (Table 1).

Table 1.
Studies of mechanical properties of the tested silumin including auxiliary variables [4]

LP	W_Mo	W_Mo_N	R_m	R_{p02}	A	HB	S_R_m	S_R_{p02}	S_A	S_HB	S_SUMA
1	0,0	0	246,2	127,1	3,5	117,0	-0,01	0,99	-0,27	1,85	2,56
2	0,0	0	255,9	112,4	3,9	115,0	0,40	-0,71	0,19	1,18	1,05
3	0,0	0	251,1	119,7	3,7	116,0	0,19	0,14	-0,04	1,51	1,80
4	0,1	1	228,5	123,1	2,6	112,0	-0,76	0,53	-1,47	0,17	-1,52
5	0,1	1	267,3	121,9	3,9	111,0	0,88	0,39	0,17	-0,17	1,28
6	0,1	1	274,2	134,7	4,1	111,5	1,17	1,88	0,50	0,00	3,55
7	0,2	2	264,0	125,3	4,2	108,0	0,74	0,79	0,60	-1,18	0,95
8	0,2	2	287,1	130,4	5,4	107,0	1,72	1,38	2,05	-1,51	3,64
9	0,2	2	251,5	110,1	4,4	107,5	0,21	-0,98	0,78	-1,34	-1,33
10	0,3	3	223,9	109,4	3,4	112,0	-0,96	-1,06	-0,46	0,17	-2,31
11	0,3	3	207,7	115,5	2,5	113,0	-1,64	-0,35	-1,52	0,50	-3,00
12	0,3	3	219,1	112,1	3,0	112,5	-1,16	-0,74	-0,87	0,34	-2,43
13	0,4	4	211,1	105,3	2,7	109,0	-1,49	-1,53	-1,24	-0,84	-5,11
14	0,4	4	260,1	120,2	4,7	111,0	0,58	0,19	1,13	-0,17	1,73
15	0,4	4	249,2	110,6	4,1	110,0	0,12	-0,92	0,45	-0,50	-0,86
		Mean	246,4	118,5	3,7	111,5					
		SD	23,62	8,63	0,80	2,98					

The statistical analysis was based on the independent variable: *total content of Mo and W in silumin [%] (variable W_Mo)*, which had five levels, namely 0.0; 0.1; 0.2; 0.3 and 0.4%, coded as 0, 1, 2, 3 and 4, respectively (*variable W_Mo_N*). The dependent variables were: *Rm* [MPa], $R_{p0.2}$ [MPa], *A* [%] and *HB*. The database was supplemented with standardized values of the dependent variables (*variables: S_Rm, S_*$R_{p0,2}$*, S_A and S_HB*) and with their sum (*variable S_SUMA*). The standardization procedure allows for the transformation of dimensional variable into the dimensionless one. As a result, the averages obtained from various sources and in different units can be compared with each other. The standardized value was calculated dividing the difference between the empirical value and average value by the value of standard deviation [1-3].

In the first stage, a univariate analysis of regression and linear correlation (according to Pearson) was carried out, capturing the effect of the independent variable *W_Mo* on each of the examined dependent features, i.e. R_m, $R_{p0.2}$, *A* and *HB,* separately. Only between *W* and *Mo* and *HB* there was a statistically significant negative linear correlation (Fig. 1).

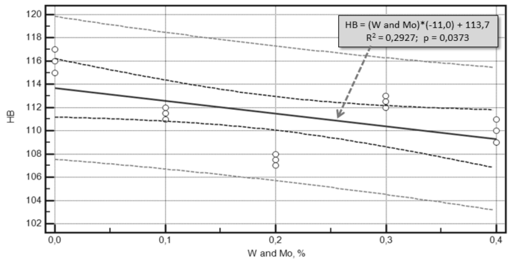

Fig. 1. The results of univariate analysis of regression and linear correlation capturing the effect of *W* and *Mo* content in the examined silumin on *HB*

To let the estimators of the linear regression function obtained by the method of least squares possess the required properties such as the efficiency and unbiasedness, a number of conditions (assumptions) should be met, the most important among them demanding that the random component (the residuals) had a normal distribution. The fulfilment of this assumption can be evaluated from the plot of normal residuals.

In the second stage, to evaluate the effect of the total content of *W* and *Mo* on R_m, $R_{p0.2}$, *A* and *HB*, the method of one-way analysis of variance (*ANOVA*) [3] was used. *ANOVA* is a technique that allows us to examine the results (experiments, observations), which depend on one factor (in this case it is the total content of *W* and *Mo* in silumin), coded with the respective numbers from 0 to 4, which denote the following five levels (variants): *0,0% =* ***0****; 0.1% =* ***1****; 0.2% =* ***2****; 0.3% =* ***3*** *and 0.4% =* ***4***. This is the, so called, grouping or classification factor.

To check the basic assumption about the uniformity (homogeneity) of variance in all groups (levels of the factor *W_MO_N*) Bartlett's test was used [1]. Should this assumption has not been met (or should each group comprise too few data - in the case under discussion only 3 measurements available each time), in the next step, to validate the hypothesis about the significance of the effect of the analyzed factor on the resultant feature, a non-parametric Kruskal-Wallis rank-sum test, which is an equivalent of the parametric one-way ANOVA, is applicable. The next validation covers the hypotheses assuming the same average values for the multiple comparisons in the examined five groups of the levels of factor W_MO_N using post-hoc Tukey test [3]. Sample results of analysis of the impact of factor *W_MO_N* in the examined silumin on *HB* are shown in Table 2.

The results of one-way ANOVA indicate that for the factor *W_MO_N* the average values of *HB* in all groups (and at all levels) differ significantly (p <0.0001). A similar result was obtained using Kruskal-Wallis rank-sum test (p = 0.0101). The results of intergroup comparisons (Tukey test) between the hardness values obtained in silumins containing *W and Mo* at a level of 0.1% and 0.3%, respectively (p = 0.5010) and *W and Mo* at a level of 0.1% and 0.4%, respectively (p=0.1722) have proved that there was no statistically significant difference in the average level of this property.

Table 2.
The results of statistical evaluation of the effect of *W and Mo* content in the examined silumin on *HB*

One-way analysis of variance					
Data: HB	**Sum of Squares**	**Degree of freedom**	**MeanSquare**	**F**	**p**
W_Mo_N	118,50	4	29,63	53,86	**0,0000**
Error	5,50	10	0,55		
Bartlett's test: $\chi 2$ = 1,949; p = 0,7450					
Post-hoc Tukey test					
Factor: W_Mo_N	**{0} 251,05**	**{1} 256,65**	**{2} 267,50**	**{3} 216,90**	**{4} 240,13**
{0}: W and Mo = 0,0%	p =	**0,0003**	**0,0002**	**0,0014**	**0,0002**
{1}: W and Mo = 0,1%	**0,0003**	p =	**0,0006**	0,5010	0,1722
{2}: W and Mo = 0,2%	**0,0002**	**0,0006**	p =	**0,0002**	**0,0138**
{3}: W and Mo = 0,3%	**0,0014**	0,5010	**0,0002**	p =	**0,0138**
{4}: W and Mo = 0,4%	**0,0002**	0,1722	**0,0138**	**0,0138**	p =
Kruskal-Wallis test: H (4, N= 15) = 13,26; p = 0,0101					

In studies of the effect of one independent factorial feature (*W_Mo_N* = 0; 1; 2; 3; 4) on several (more than one) dependent features (R_m, $R_{p0.2}$, *A*, *HB*), modern statistical tests, such as a multivariate analysis of variance (MANOVA), seem to be a good choice [3].

In the last stage of the optimization process of the combined effect of the explanatory variable, which is the addition of *W and Mo,* on the tested properties of silumin, a univariate analysis of regression and correlation was used [1-3]. Since the aim was to simultaneously achieve maximum values of all the investigated dependent variables, as an explained (dependent) variable, the sum of the standardized values of all the analyzed dependent variables *S_SUMA* (Table 3) has been adopted.

Table 3.
The results of linear regression capturing the effect of *W and Mo* content on *S_SUMA*

R= 0,5731; Coefficient of determination R^2= 0,3284						
F(1,13)=6,359 p<0,0255; Residual standard deviation 2,1973						
Parameter	**BETA**	**Std. error BETA**	**Coefficient**	**Std. error B**	**t(13)**	**p**
Intercept			2,0233	0,9827	2,0590	0,0601
W_Mo	-0,5731	0,2273	-10,1166	4,0117	-2,5218	**0,0255**

Finally, a statistically significant linear correlation (p=0.0255) between the total content of *W and Mo* and the sum of the standardized variables (the examined properties) was obtained (Fig. 3).

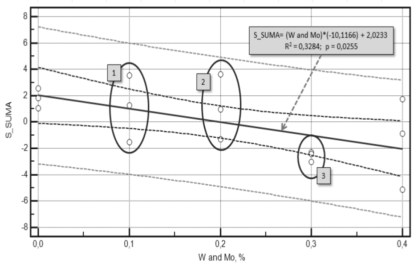

Fig. 2. Plotted linear correlation capturing the effect of *W and Mo* content in the examined silumin on *S_SUMA*

3.2. Evaluation of the effect of heat treatment conditions on the ADI impact resistance

Table 4 shows the results of impact resistance tests carried out on the unnotched specimens of ADI (Table 1).

Table 4.
Fragment of a database with the results of impact resistance tests carried out on cast iron [5]

Lp	Podchla	Podchla_N	Tpi	Tpi_N	tau_pi	tau_pi_N	Kc
1	850	2	300	1	7,5	1	63,5
2	850	2	300	1	7,5	1	70,2
3	850	2	300	1	7,5	1	56,6
4	850	2	300	1	15	2	62,4
5	850	2	300	1	15	2	69,3
6	850	2	300	1	15	2	58,4
7	850	2	300	1	30	3	53,3
8	850	2	300	1	30	3	69,8
9	850	2	300	1	30	3	59,0
10	850	2	300	1	60	4	51,6
11	850	2	300	1	60	4	45,7
12	850	2	300	1	60	4	49,6
13	850	2	300	1	120	5	43,8
14	850	2	300	1	120	5	40,9
15	850	2	300	1	120	5	43,8
16	850	2	300	1	240	6	41,0
17	850	2	300	1	240	6	42,8
18	850	2	300	1	240	6	42,8
19	800	1	300	1	7,5	1	61,7
20	800	1	300	1	7,5	1	58,6
21	800	1	300	1	7,5	1	60,4
22	800	1	300	1	15	2	60,3

The statistical evaluation of the heat treatment effect on the impact resistance of cast iron (the measurable dependent variable *Kc* [J/cm^2]) included the following independent variables:

- cooling to intermediate temperature during austenitizing of cast iron expressed as a measurable variable with two levels of 800 and 850 [°C] (variable *Podchla*) and as a qualitative variable with levels 1 and 2, corresponding to the respective levels of temperature (*Podchla_N* variable);
- the temperature of austempering expressed as a measurable variable with two levels of 300 and 400 [° C] (*Tpi*) and as a qualitative variable with levels 1 and 2, corresponding to the respective levels of temperature (variable *Tpi_N*);
- the time of isothermal transformation expressed as a measurable variable with six levels of 7.5; 15; 30; 60; 120 and 240 [min] (variable *tau_pi*) and as a qualitative variable with levels from 1 to 6, corresponding to the respective levels of transformation time (variable *tau_pi_N*);

The adopted plan of experiment assumed the execution of full experiment, i.e. the execution of 24 (2·2·6 = 24) tests (called groups or subclasses). Each test consisted of three measurements of the impact resistance *Kc*. So, in total, 72 measurements of the impact resistance were made [5]. The effect of three independent variables treated as factor variables,i.e. *Podchla_N, Tpi_N and tau_pi_N,* on one measurable dependent variable *Kc* can be evaluated using the analysis of variance for factorial designs. The condition for use of this statistical tool is the homogeneity of variance of the examined feature in different subclasses, which can be checked with Levene's or Bartlett's test (Table 5). Based on the results of these two tests, it was found that there were no grounds to reject the hypothesis about the homogeneity (equality) of variance. The results of the analysis of variance for factorial designs (Table 5) have shown that the statistically significant effect on the impact resistance of the tested cast iron could be expected from all the single variables (factors), i.e. *Podchla_N, Tpi_N and tau_pi_N*, and also from some of their interactions, such as the interaction between *Podchla_N and Tpi_N,* and interaction between *Tpi_N and tau_pi_N.*

Table 5.
The results of the analysis of variance for factorial designs

Bartlett's test: $\chi 2$ = 27,6; p = 0,2304					
Levene's test: F = 1,627; p = 0,0773					
Effect	**Sum of Squares**	**Degree of freedom**	**Mean Square**	**F**	**p**
Intercept	350240,6	1	350240,6	10047,04	0,0000
Podchla_N	3040,4	1	3040,4	87,22	**0,0000**
Tpi_N	29010,3	1	29010,3	832,19	**0,0000**
tau_pi_N	6915,2	5	1383,0	39,67	**0,0000**
Podchla_N*Tpi_N	849,2	1	849,2	24,36	**0,0000**
Podchla_N*tau_pi_N	53,2	5	10,6	0,30	**0,9076**
Tpi_N*tau_pi_N	826,6	5	165,3	4,74	**0,0013**
Podchla_N*Tpi_N*tau_pi_N	152,3	5	30,5	0,87	**0,5056**
Error	1673,3	48	34,9		

It was also found that the strongest effect on changes in the cast iron impact resistance *Kc* had the factor (variable) *Tpi_N*, followed by *Podchla_N* (the effect nearly ten times weaker than that of *Tpi_N*) and *tau_pi_N* (the effect nearly twenty times weaker than that of *Tpi_N*). The information about the effect of individual factors (independent variables) on the impact resistance of cast iron can be derived from the values of descriptive statistics given in Table 6 and graphs based on these values, plotted in Figures 3-5.

Table 6.
Characteristics of descriptive statistics for individual subclasses (fragment)

Factors	Level Podchla_N	Level Tpi	Level tau_pi	N	Kc Mean	Kc Std. dev.	Kc Std. error	Kc: CI: 95,00%	Kc CI: +95,00%
Podchla_N*Tpi_N*tau_pi_N	1	1	1	3	60,23	1,56	0,90	56,37	64,10
Podchla_N*Tpi_N*tau_pi_N	1	1	2	3	56,13	4,86	2,80	44,07	68,20
Podchla_N*Tpi_N*tau_pi_N	1	1	3	3	48,10	2,25	1,30	42,50	53,70
Podchla_N*Tpi_N*tau_pi_N	1	1	4	3	39,50	1,47	0,85	35,84	43,16
Podchla_N*Tpi_N*tau_pi_N	1	1	5	3	42,03	3,65	2,11	32,97	51,09
Podchla_N*Tpi_N*tau_pi_N	1	1	6	2	36,55	2,19	1,55	16,86	56,24
Podchla_N*Tpi_N*tau_pi_N	1	2	1	3	80,87	4,32	2,49	70,15	91,59
Podchla_N*Tpi_N*tau_pi_N	1	2	2	3	104,20	5,41	3,12	90,77	117,63
Podchla_N*Tpi_N*tau_pi_N	1	2	3	3	86,17	5,06	2,92	73,61	98,73
Podchla_N*Tpi_N*tau_pi_N	1	2	4	3	76,03	2,47	1,43	69,90	82,17
Podchla_N*Tpi_N*tau_pi_N	1	2	5	3	68,27	0,71	0,41	66,50	70,03
Podchla_N*Tpi_N*tau_pi_N	1	2	6	2	68,75	4,74	3,35	26,18	111,32
Podchla_N*Tpi_N*tau_pi_N	2	1	1	3	63,43	6,80	3,93	46,54	80,33
Podchla_N*Tpi_N*tau_pi_N	2	1	2	3	63,37	5,51	3,18	49,67	77,06
Podchla_N*Tpi_N*tau_pi_N	2	1	3	3	60,70	8,38	4,84	39,88	81,52
Podchla_N*Tpi_N*tau_pi_N	2	1	4	3	48,97	3,00	1,73	41,51	56,42
Podchla_N*Tpi_N*tau_pi_N	2	1	5	3	42,83	1,67	0,97	38,67	46,99
Podchla_N*Tpi_N*tau_pi_N	2	1	6	4	40,40	3,70	1,85	34,51	46,29
Podchla_N*Tpi_N*tau_pi_N	2	2	1	3	105,83	7,22	4,17	87,89	123,77
Podchla_N*Tpi_N*tau_pi_N	2	2	2	3	122,17	8,86	5,12	100,15	144,19
Podchla_N*Tpi_N*tau_pi_N	2	2	3	3	102,83	9,57	5,52	79,07	126,60
Podchla_N*Tpi_N*tau_pi_N	2	2	4	3	97,03	7,44	4,30	78,54	115,53
Podchla_N*Tpi_N*tau_pi_N	2	2	5	3	87,77	8,50	4,91	66,65	108,88
Podchla_N*Tpi_N*tau_pi_N	2	2	6	4	89,08	10,56	5,28	72,27	105,88

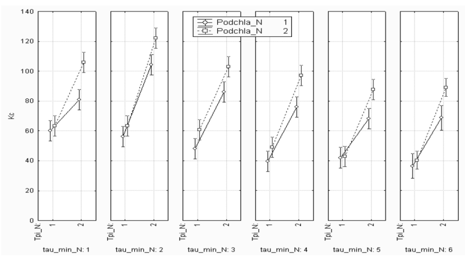

Fig. 3. The effect of the examined factors on the cast iron impact resistance *Kc*

From these graphs (Figs. 3-5) it can be concluded that increasing *Tpi* (from 300 to 400°C) increases the impact resistance (regardless of the value of *Podchla_N* and *tau_min_N*). Increasing *Podchla* (from 800 to 850°C) also increases the impact resistance (regardless of the value of *Tpi* and *tau_min_N*), while increasing *tau_min* (above 15 minutes) reduces the impact resistance (regardless of the value of *Tpi_N* and *Podchla_N*).

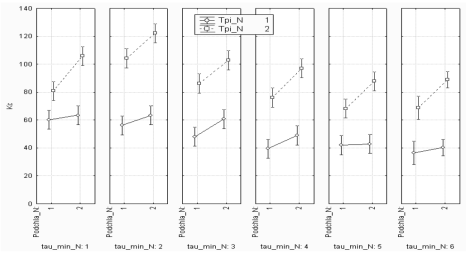

Fig. 4. The effect of the examined factors on the cast iron impact resistance *Kc*

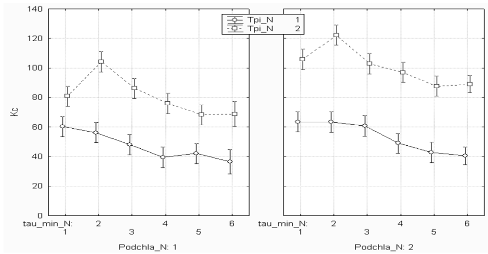

Fig. 5. The effect of the examined factors on the cast iron impact resistance *Kc*

The next step may involve interclass comparisons made with post-hoc Tukey test (Table 7).

Table 7.
The results of interclass comparisons made with post-hoc Tukey test (fragment)

Number of the subclass	Factors			Number of the subclass											
	Podchla_N	Tpi_N	tau_pi_N	1	2	3	4	5	6	7	8	9	10	11	12
1	800°C	300°C	8	p =	1,0000	0,6590	**0,0146**	0,0615	**0,0111**	**0,0154**	**0,0002**	**0,0006**	0,1966	0,9893	0,9944
2	800°C	300°C	15	1,0000	p =	0,9893	0,1349	0,3801	0,0877	**0,0012**	**0,0002**	**0,0002**	**0,0238**	0,6590	0,7736
3	800°C	300°C	30	0,6590	0,9893	p =	0,9771	0,9998	0,8773	**0,0002**	**0,0002**	**0,0002**	**0,0003**	**0,0204**	0,0532
4	800°C	300°C	60	**0,0146**	0,1349	0,9771	p =	1,0000	1,0000	**0,0002**	**0,0002**	**0,0002**	**0,0002**	**0,0002**	**0,0006**
5	800°C	300°C	120	0,0615	0,3801	0,9998	1,0000	p =	1,0000	**0,0002**	**0,0002**	**0,0002**	**0,0002**	**0,0005**	**0,0021**
6	800°C	300°C	240	**0,0111**	0,0877	0,8773	1,0000	1,0000	p =	**0,0002**	**0,0002**	**0,0002**	**0,0002**	**0,0003**	**0,0005**
7	800°C	400°C	8	**0,0154**	**0,0012**	**0,0002**	**0,0002**	**0,0002**	**0,0002**	p =	**0,0029**	1,0000	1,0000	0,5914	0,8263
8	800°C	400°C	15	**0,0002**	**0,0002**	**0,0002**	**0,0002**	**0,0002**	**0,0002**	**0,0029**	p =	0,0671	**0,0003**	**0,0002**	**0,0002**
9	800°C	400°C	30	**0,0006**	**0,0002**	**0,0002**	**0,0002**	**0,0002**	**0,0002**	1,0000	0,0671	p =	0,8945	0,0720	0,2161

In the next stage, to evaluate the effect of three independent variables treated as numerical variables, i.e. *Podchla, Tpi* and *tau_pi*, on one measurable dependent variable *Kc,* the method of multiple (multivariate, multi-dimensional) regression can be used.

Table 8.
The results of multivariate (linear) regression showing the effect of *Podchla, Tpi* and *tau-pi* on the cast iron impact resistance *Kc*

R= 0,9298; R^2= 0,8646; F(3,68)=144,75 p<0,0001 Standard error: 9,3046						
Independent variables	**BETA**	**Std. error BETA**	**B**	**Std. error B**	**t(68)**	**p**
Constant			-290,0007	37,0232	-7,8330	0,0000
Podchla	0,2768	0,0446	0,2721	0,0439	6,2025	**0,0000**
Tpi	0,8346	0,0446	0,4102	0,0219	18,7037	**0,0000**
tau_pi	-0,3025	0,0446	-0,0915	0,0135	-6,7797	**0,0000**

Analyzing the linear factors only (no interaction), it can be concluded that statistically significant effect on the impact resistance improvement has an increase in *Tpi* and *Podchla* (the effect of the latter one being almost three times weaker than the effect of *Tpi*) and decrease in *tau_pi*. As in other ANOVA tests, the condition for the correct statistical evaluation by this method is the distribution of residuals close to normal, which can be assessed from the normal probability plot of residuals. A very good fit of the function:

$$Kc = Podchla*0,2721+Tpi*0,4102 - tau\_pi*0,0915 - 290,0$$

is confirmed by the value of the coefficient of determination close to unity R^2 = 0.8646 and a scatterplot comparing the values predicted and observed (Fig. 6)

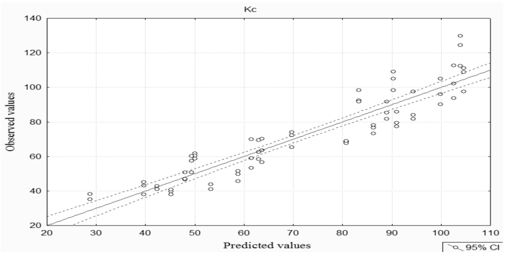

Fig. 6. Scatterplot comparing the values predicted and observed

3.3. The forecast share of ductile iron in the total production of castings in Poland

To better highlight the possibilities offered by modern methods of forecasting in relation to the time series associated directly with the production of castings, data on the share of ductile iron in the total production of castings in the years 1989 – 2013 has been used [6] (Fig. 7).

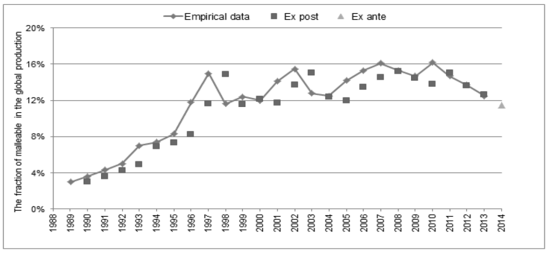

Fig. 7. The share of ductile iron in the total production of castings in Poland in the years 1989-2013

In this series it is easy to notice large random fluctuations with upward trend in the starting period and a clear downward trend in recent years. With this course of changes it is difficult to build a forecast of high accuracy and reliability. After a detailed analysis, for the construction of forecast, the following methods were selected [7]:

- Multinomial econometric model of second degree;
- Autoregressive model AR1;
- Model of naive forecasting for a series focusing on the development trend;
- Model of simple moving average for a series focusing on the development trend (k = 2 and k = 3);
- Simple model of exponential smoothing;
- Exponential-autoregressive model;
- Model of weighted moving average for a series focusing on the development trend (k = 3);
- Model of single exponential smoothing;
- Holt's linear model with additive trend (two types of start-up);
- Holt's linear model with multiplicative trend (two types of start-up);
- Holt's linear model allowing for the trend suppression effect (additive - two types of start-up);
- Holt's linear model allowing for the trend suppression effect (multiplicative - two types of start-up);
- Holt's square model (two types of start-up).

For all of the above presented forecasting methods, the point forecast (*ex-ante*) for the year 2014, the average approximation forecast error (*ex-post),* and the root mean square error (*ex-post*) were estimated.

Figure 7 shows an example of Holt's multiplicative linear forecast model allowing for the trend suppression effect.

4. Summary and conclusions

It is quite obvious that the statistical methods described in this article do not exhaust the very vast subject. What the authors wanted to achieve was only an easy-to-understand presentation of the possibilities offered by modern statistical tools, based on the selected research problems related to foundry practice with correct interpretation of the obtained results. An advantage of the presented methods and tools, promoting their practical application, is easy implementation, customization and processing of results in the popular Excel spreadsheet.

Acknowledgements

The authors would like to thank the Staff Members at the Universities of Czestochowa, Lodz and Bydgoszcz for sharing the test results and kind assistance in their implementation.

References

[1] Szymszal, J., Maliński, M. (1999). *Contemporary mathematical statistics in spreadsheets.* Katowice: Ed. SlAM. (in Polish).

[2] Aczel, A.D. (2000). *Statistics in the management.* Warszawa: PWE. (in Polish).

[3] Stanisz, A. (2006). *Accessible course in statistics*. Vol. 1-3. Kraków. StatSoft Polska Sp. z o.o. (in Polish).

[4] Szymczak, T., Gumienny, G., Walas, K. & Pacyniak, T. (2015). Effect of Tungsten and Molybdenum on the Crystallization, Microstructure and Properties of Silumin 226. *Archives of Foundry Engineering*. 15(3), 61-66.

[5] Giętka, T., Ciechacki, K., Szykowy. T. (2015). Influence of the temperature of breaking on the ADI cast irons impact resistance. *Archives of Foundry Engineering*. 15(2), 87-92. (in Polish).

[6] Soiński, M.S., Skurka, K., Jakubus, A. & Kordas, P. (2015) . Structure of Foundry Production in the World and in Poland over the 1974-2013 Period. *Archives of Foundry Engineering*. 15(2), 69-76.

[7] Krawiec, S. (2014). *Adaptive models of exponential smoothing as instruments of short-term forecasting quantitative phenomena*. Katowice: Ed. Silesian University of Technology. (in Polish).