Optimization methodology of product technical characteristics in the production of household appliances

M V Shanta, E G Semenova and M S Smirnova
Saint-Petersburg State University of Aerospace Instrumentation, Bolshaya Morskaia str. 67 A, Saint-Petersburg, 190000, Russian Federation

E-mail: marina.s_2004@mail.ru

Abstract: The paper presents optimization methodology of product technical characteristics, which was approbated by optimization of conductivity of water on washing machine production factory. This process ensures effectiveness of rinsing process during washing cycle and reduction of chemical ingredients after rinsing. Main factors, which influence conductivity function, were defined. The factors are analysed to define dependence of investigated values from each other. Developed linear multiple regression model, quality of developed model are checked. Due to usage of the developed methodology at washing machine factory quantity of chemical ingredients after rinsing process was reduced; as a result, norms of GOST standard 60456-2011 were reached.

1. Introduction
In the context of globalization and dynamic development of companies, it is necessary not only to produce high-quality products and meet the requirements of standards and norms, but also to constantly improve the products, develop new functions, and optimize existing technical properties of products.

The topic is urgent due to the fact that, despite a number of existing methods and tools for quality management (Ishikawa analysis, Kano model method, Kaizen and Cairio method, failure form and effects analysis (FMEA), six sigma methodology), there is no specific method for making technical changes in product characteristics, taking into account all factors influencing the desired result. Various manuals on mathematical modelling allow us to build models, but do not describe the specific steps required for the process of product optimization [1-2]. The international standards of the ISO series (ISO 9000, ISO 9001, ISO 9004, ISO 19011) do not contain recommendations as to when and in what sequence it is necessary to use one or another tool when optimizing parameters taking into account all the design and technological features of the product [3-6].

2. Optimization methodology of technical characteristics in production
The authors of the paper have developed a technique for optimizing technical characteristics for manufacturing enterprises, which allows identifying the main factors affecting the process under study, determining the optimal parameters of technical characteristics, and verifying the process. The methodology includes the steps, presented in table 1.
| Activities                                                                 | Method                          | Calculation formula                                                                 | Comments                                                                                                                                                                                                 |
|---------------------------------------------------------------------------|---------------------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Determination of factors influencing the process under study           | Fishbone Diagram                | -                                                                                     | Sequencing: 1. identifying potential influencing factors 2. grouping factors 3. ranking factors 4. elimination of insignificant factors 5. prioritization of factors by degree of significance |
| 2. Determining the minimum number of experiments                          | Full factor experiment          | \( N = p^k \), where \( N \) – number of experiments; \( p \) – number of factor levels; \( k \) – number of factors. | The level of the factor (p) represents the quantititative or qualitative states of the factor chosen for the experiment. In other words, these are the upper and lower levels. The minimum number of levels usually applied in the first stage of work equals to 2. |
|                                                                          | Fractional factor experiment     | \( n = N^{(k-p)} \), where \( n \) – minimal number of tests needed for realization of the experiment, \( N \) – number of factor levels, \( k \) – number of studied parameters, \( p \) – the number of linear effects equivalent to interaction effects. The \( k-p \) difference is called resolution. | Based on the theory of fractional factorial plans (FFP), it is possible to calculate the minimum number of tests necessary to draw a conclusion about the process. |
| 3. Statistics collection                                                  | Statistics observation          | -                                                                                     | -                                                                                                                                                                                                       |
| 4. Variable analysis                                                      | Time series plot                | -                                                                                     | Analysis of the dependence of the process on time                                                                                                                                                       |
|                                                                          | Probability plot                | -                                                                                     | Process analysis for normal distribution                                                                                                                                                                |
| 5. Model parameterization                                                 | Scatter plot/Probability plot   | -                                                                                     | The relationship between the variables x and y is represented by dots on the coordinate plane \((x, y)\) and is connected by a line. An assumption about the form of the dependence of the variable x on y: linear or nonlinear is made by the form of the empirical regression line. |
| 6. Model simulation                                                       | Paired regression               | \( y(x) = f^e(x) \), where \( y \) – dependent variable (effective | -                                                                                                                                                                                                       |
attribute); x – independent or explanatory variable (sign factor).

Multiple regression

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m + \varepsilon, \]

where \( \beta_0, \beta_1, \beta_2, \ldots, \beta_m \) are regression coefficients (empirical regression coefficients); \( \varepsilon \) – deviation estimate. \( X_1, X_2, \ldots, X_m \) - vector of independent (explanatory) variables.

A model with a large number of factors. With the help of multiple regression, the influence of each of the factors separately on the result is determined, as well as their cumulative impact on the simulated indicator.

10. Correlation analysis

Paired correlation coefficients

\[ r_{xy} = \frac{\bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y}}{s(x) \cdot s(y)} \]

Partial correlation coefficients

\[ r_{yx_1/x_2} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1/x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1/x_2}^2)}} \]

The relationship between the two signs is distorted due to the fact that both signs are influenced by other factors. To obtain more accurate relationships between two variables, the influence of the third variable on them is excluded.

10. Quality check of the constructed regression equation

Determination coefficient

\[ R^2 = 1 - \frac{s_e^2}{\sum (y_i - \bar{y})^2} \]

\[ F = \frac{R^2}{1 - R^2} \frac{n - m - 1}{m} \]

The assessment of the significance of the multiple regression equation is carried out by testing the hypothesis of zero equality coefficient of determination calculated according to the general population.

If \( F < F_{kp} = F_\alpha ; n-m-1 \), there is no reason to reject the hypothesis \( H_0 \).

The term "technical characteristics" refers to the parameters of the equipment that ensure the correct functioning of the equipment throughout its entire service life. This group includes indicators of safety, reliability, functionality and harmlessness. In order to determine the possible factors of influence (x) on the dependent variable (y), it is advisable to use the Fishbone Diagram. Depending on the goals and conditions of the study, the following types of experiment are distinguished: a full factorial experiment and a fractional factorial plan ("semi-replica").

One of the methods for analysing variables is graphical analysis. To determine whether all factors are included in the analysis, it is necessary to exclude the fact that the process is time series plot and analyse the process for a normal probability plot. Parametrization means the choice of the type of model, including the composition and form of the relations between variables included in it. The relationship between the dependent variable (y) and a set of one or more variables of independent variables (x) may be linear and non-linear. One of the most commonly used statistical methods is a multiple linear regression. As a rule, the influence on the process is not limited to one factor, while this regression model can be used to predict Y, when only X is known. To estimate the significance of the multiple regression equation, use Fisher's F-test. If independent variables are different in their essence and have
different units of measurement, then the regression coefficients are incomparable. In this case, the closeness of the relationship can be measured using partial elasticity coefficients, $\beta$ – coefficients or partial correlation coefficients.

3. Improving the performance of technical characteristics in the production of washing machines

In the process of testing the methodology at an enterprise specializing in the production of washing machines, the functional characteristics of the products manufactured were analysed, such as washing efficiency, rinsing quality, spinning efficiency of laundry, wear of laundry after washing, and performance. It was found that the products have low rates associated with the quality of gargling. The amount of washing solution remaining in the laundry after rinsing exceeded the norm of 0.3 mg-eq/dm$^3$, which contradicts the requirements of GOST R IEC 60456-2011 [7]. The process of complete removal of powder and other chemical elements (rinsing) resulting from the use of cleaning agents is especially important for people suffering from reactions to household allergens (phosphates, flavours, bleaches, dyes). Powder allergy is particularly dangerous for children and pregnant women. In addition, if the laundry is badly rinsed, it shows traces of powder after washing. Properly selected software parameters for the rinsing stage will allow retaining the original properties and colour of the laundry, remove particles of detergent and dirt.

In the analysis of root causes using Fishbone Diagram, it was found that the technical characteristics of the electrical conductivity of water are not taken into account when testing instruments, but it is one of the priority factors influencing the quality of rinsing. The analysis of the factors revealed the following independent variables ($x_i$) that affect the dependent variable ($y$): washing time, washing temperature, rinsing time, amount of water consumed during rinsing. When solving the task, a full factor experiment $2^4$ was carried out, and therefore 16 washing machines were analysed. The test results are shown in table 2.

**Table 2.** The results of testing the function of electrical conductivity for the process of rinsing clothes.

| N  | Electrical conductivity (y) | Washing time ($x_1$) | Washing temperature ($x_2$) | Rinsing time ($x_3$) | Water consumption ($x_4$) |
|----|----------------------------|----------------------|-----------------------------|---------------------|--------------------------|
| 1  | 629                        | 123                  | 37                          | 5                   | 45                       |
| 2  | 620                        | 175                  | 37                          | 5                   | 45                       |
| 3  | 623                        | 123                  | 45                          | 5                   | 45                       |
| 4  | 616                        | 175                  | 45                          | 5                   | 45                       |
| 5  | 618                        | 123                  | 37                          | 8,5                 | 45                       |
| 6  | 621                        | 175                  | 37                          | 8,5                 | 45                       |
| 7  | 615                        | 123                  | 45                          | 8,5                 | 45                       |
| 8  | 614                        | 175                  | 45                          | 8,5                 | 45                       |
| 9  | 618                        | 123                  | 37                          | 5                   | 55                       |
| 10 | 593                        | 175                  | 37                          | 5                   | 55                       |
| 11 | 625                        | 123                  | 45                          | 5                   | 55                       |
| 12 | 596                        | 175                  | 45                          | 5                   | 55                       |
| 13 | 607                        | 123                  | 37                          | 8,5                 | 55                       |
| 14 | 593                        | 175                  | 37                          | 8,5                 | 55                       |
| 15 | 612                        | 123                  | 45                          | 8,5                 | 55                       |
| 16 | 598                        | 175                  | 45                          | 8,5                 | 55                       |

The previously selected sample of measurements of the electrical conductivity was made with a data logging interval of 1 day, the measurement results are presented in figure 1.
Figure 1 shows that the time series of values is not peculiar to cyclo-stationarity, the data structure on weekdays is somewhat different from the data structure on weekends. An absolutely random distribution of electrical conductivity values was found. There is no general or partial tendency to decline or increase in values, which indicates the absence of special reasons, such as, for example, weather conditions, humidity in the room, dependence on the work shift.

During the graphical analysis of data for probability plot, it was revealed that the graph of the series shows a uniform random hit of points in the range on both sides of the conditional centre line, which indicates the normal probability plot, and also indicates that the regression type is linear (figure 2).

Figure 2. Graph of water electrical conductivity probability plot.

The proximity of values to a straight red line also indicates that the relationship between the variables is linear. Using multiple linear regression, the paper developed a model of the electrical conductivity function for the rinsing process, determined the influence of each of the factors separately on the result, as well as their cumulative effect on the simulated indicator. The vectors of regression coefficient estimates are calculated according to the least squares method using the formula:
\[ Y(X) = (X'X)^{-1}X'Y, \]  

where \( Y \) is the dependent variable, \( x \) is the independent variable, \((X'X)^{-1}\) is the inverse matrix of the product of the transposed matrix of the independent variable and the matrix of the independent variable, \( X'Y \) is the product of the transposed matrix of the independent variable and the dependent one.

The regression equation can be presented as follows:

\[ Y = 728.1346 -0.2389X_1 +1.5002X_2 +1.425X_3, \]

where \( y \) – water conductivity, \( x_1 \) – washing time, \( x_2 \) – rinsing time, \( x_3 \) – water consumption during rinsing.

To estimate the closeness of the linear correlation connection between the dependent random variable and the independent random variable, it is necessary to calculate the pairwise correlation coefficients. The coefficient of pair correlation (Pearson) is a number from \( -1 \) to \( 1 \), characterizing the closeness of the linear correlation connection. On the Cheddock scale, the interpretation of the possible values of the correlation coefficient is carried out according to the following principle: if \(|r| > 0.3\) - the connection is almost absent; \( 0.3 \leq |r| \leq 0.7 \) - connection is average; \( 0.7 \leq |r| \leq 0.9 \) - strong bond; \(|r| > 0.9\) - the connection is very strong. In the same way it can be judged by the coefficient of pair correlation on the existence of multicollinearity in the model. If there is an interfactor correlation coefficient in the matrix, \( r_{x_ix_i} > 0.7 \), then there is multicollinearity in this multiple regression model.

During the analysis of the obtained statistical data, a matrix of independent variables \( X \) was obtained:

\[
\begin{vmatrix}
\sum n & \sum x & \sum x_1 & \sum x_2 & \sum x_3 \\
\sum y & \sum y^2 & \sum xy & \sum x_1y & \sum x_2y \\
\sum x_1 & \sum x_1y & \sum x_1^2 & \sum x_1x_2 & \sum x_1x_3 \\
\sum x_2 & \sum x_2y & \sum x_2x_1 & \sum x_2^2 & \sum x_2x_3 \\
\sum x_3 & \sum x_3y & \sum x_3x_1 & \sum x_3x_2 & \sum x_3^2
\end{vmatrix}
\]

(3)

The calculation of paired correlation coefficients is made according to the formula:

\[ r_{xy} = \frac{\sum xy - \sum x \sum y}{s(x)s(y)}, \]

(4)

The matrix of paired correlation coefficients according to the calculations made is:

\[
\begin{vmatrix}
y & x_1 & x_2 & x_3 \\
x_1 & 1 & -0.5338 & -0.2335 & -0.6338 \\
x_2 & -0.5338 & 1 & 0 & 0 \\
x_3 & -0.2335 & 0 & 1 & 0 \\
x_3 & -0.6338 & 0 & 0 & 1
\end{vmatrix}
\]

(5)

From the magnitude of paired correlation coefficients (all paired correlation coefficients \(|r| < 0.7\)), we can conclude that the influence of factors on each other is absent, which indicates the absence of multicollinearity of factors. According to the calculated t-statistics values, the correlation coefficients for \( r_{x_1x_1} \) and \( r_{x_1x_3} \) are statistically significant, the correlation coefficient for \( r_{x_2x_2} \) is not significant. Due to the fact that some of the coefficients are significant, and some are not significant, there are no grounds for a violation of adequacy. This means that part of the regressors can adequately describe the object. Since the factor signs are different in their essence and have different units of measurement, it is necessary to measure the closeness of the relationship between the factors and the result. To assess the closeness of the relationship of the factors with the result, as well as in order to screen out the factors carrying no useful information, we will calculate the partial correlation coefficients. The partial correlation coefficients measure the effect on the result of factor \( x_i \) with a constant level of other factors. If the value of the coefficient is small or insignificant, then the relationship between this factor and the resultant variable is either very weak or not at all, so the factor can be excluded from the model.

The coefficient of private correlation is calculated according to the formula:
\[ r_{yx_1/x_2} = \frac{r_{yx_1} - r_{yx_2} r_{x_1 x_2}}{\sqrt{(1-r^2_{yx_2})(1-r^2_{x_1 x_2})}} \]  

To calculate the significance of the correlation coefficient \( r_{yx_1/x_2} \), it is necessary to calculate the observed values of t-statistics by the formula:

\[ t_{nabt} = \frac{r_{yx_1/x_2}}{\sqrt{\frac{n-k-2}{1-r^2_{yx_1/x_2}}}} \]

where \( k = 1 \) - the number of fixed factors.

If \( t_{\text{observ}} > t_{\text{crit}} \), the hypothesis of equality 0 correlation coefficient is rejected and the correlation coefficient is statistically significant. If \( t_{\text{observ}} < t_{\text{crit}} \), we accept the hypothesis of equality 0 of the correlation coefficient. In other words, the correlation coefficient is not statistically significant. Table 3 shows the results of calculations of partial correlation coefficients.

Table 3. Partial correlation coefficients of the electrical conductivity function.

| Correlation coefficient | Value | Correlation | T_{\text{observ}} | T_{\text{crit}} | Correlation coefficient value | Comments |
|-------------------------|-------|-------------|------------------|----------------|-------------------------------|----------|
| \( r_{yx_1/x_2} \)     | -0.549| Moderate    | 2.37             | 2.16           | \( t_{\text{observ}} > t_{\text{crit}} \) - significant | The correlation of \( y \) and \( x_1 \) provided that \( x_2 \) will be included into the model, became stronger |
| \( r_{yx_1/x_3} \)     | -0.69 | Moderate    | 3.44             | 2.16           | \( t_{\text{observ}} > t_{\text{crit}} \) - significant | The correlation of \( y \) and \( x_1 \) provided that \( x_2 \) will be included into the model, became stronger |
| \( r_{yx_2/x_1} \)     | -0.276| Low         | 1.04             | 2.16           | \( t_{\text{observ}} < t_{\text{crit}} \), not significant | The correlation of \( y \) and \( x_2 \) provided that \( x_1 \) will be included into the model, became stronger |
| \( r_{yx_2/x_3} \)     | -0.302| Not strong  | 1.14             | 2.16           | \( t_{\text{observ}} < t_{\text{crit}} \), not significant | The correlation of \( y \) and \( x_2 \) provided that \( x_1 \) will be included into the model, became stronger |
| \( r_{yx_3/x_1} \)     | -0.75 | Strong      | 4.08             | 2.16           | \( t_{\text{observ}} > t_{\text{crit}} \) - significant | The correlation of \( y \) and \( x_3 \) provided that \( x_1 \) will be included into the model, became stronger |
| \( r_{yx_3/x_2} \)     | -0.652| Moderate    | 3.1              | 2.16           | \( t_{\text{observ}} > t_{\text{crit}} \) - significant | The correlation of \( y \) and \( x_3 \) provided that \( x_2 \) will be included into the model, became stronger |

Based on the results obtained, factor \( x_3 \) (water consumption during rinsing) has the greatest influence on the process; factor \( x_1 \) (rinsing time) has the least influence. Table 4 presents the results of the analysis of the inter-factorial relationship of the partial correlation coefficients.

Table 4. Inter-factor coupling of partial correlation coefficients.

| Correlation coefficient | Value | Correlation | T_{\text{observ}} | T_{\text{crit}} | Correlation coefficient value | Comments |
|-------------------------|-------|-------------|------------------|----------------|-------------------------------|----------|
| \( r_{x_1 x_2 y} \)    | -0.152| Weak        | 0.55             | 2.16           | \( t_{\text{observ}} < t_{\text{crit}} \), not significant | The correlation of \( y \) and \( x_2 \) provided that \( y \) is included into the model, has lowered. The input |
It follows from the analysis that the input of the factor $x_3$ into the regression equation remains inappropriate. However, after the exclusion of the $x_3$ factor, the coefficient of determination decreased significantly, and therefore the factor was not excluded from the equation.

To check the overall quality of the regression equation, the coefficient of determination $R^2$ is used, which is generally calculated by the formula:

$$ R^2 = 1 - \frac{s_e^2}{\sum(y_i - \bar{y})^2} $$

Thus,

$$ R^2 = 1 - \frac{547.43}{2021.75} = 0.7292 $$

Analysis of the statistical significance of the coefficient of determination is carried out on the basis of testing zero hypothesis $H_0: R^2 = 0$ versus alternative hypothesis $H_1: R^2 > 0$. To test this hypothesis, the following $F$ – statistics is used:

If $F < F_{kp} = F_{\alpha}$; $n-m-1$, there is no reason to reject the hypothesis $H_0$.

$$ F = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m} = \frac{0.7292}{1-0.7292} \cdot \frac{16-3-1}{3} = 10.77 $$

The tabular value with degrees of freedom $k_1 = 3$ and $k_2 = n-m-1 = 16 - 3 - 1 = 12$, $F_{kp}(3;12) = 3.49$. Since the actual value is $F > F_{kp}$, the coefficient of determination is statistically significant and the regression equation is statistically reliable (that is, the coefficients $b_i$ are jointly significant).
As a result of the calculations, a multiple regression equation was obtained: 
\[ Y = 728.1346 - 0.2389X_1 - 1.5002X_2 - 1.425X_3. \]

The following interpretation of the model parameters is possible: an increase in \( X_1 \) (washing time) by 1 unit leads to a decrease in \( Y \) by an average of 0.239 units; increase in \( X_2 \) (rinse time) by 1 unit leads to a decrease in \( Y \) by an average of 1.5 units; increase in \( X_3 \) (water consumption during rinsing) by 1 unit leads to a decrease in \( Y \) by an average of 1.425 units. The statistical significance of the equation is verified using the coefficient of determination and the Fisher criterion. It is established that in the situation under study, 72.92% of the total variability of \( Y \) is explained by a change in \( X_j \) factors.

4. Conclusion

The developed methodology can be applied at various stages of the product life cycle, starting from marketing research (identifying a set of factors affecting demand), developing a design (determining optimal software parameters), optimizing technical characteristics and processes, analysing the causes of defective products, and ending with reclamations analysis on the market (identifying the main causes of a service call). The methodology can be used as a tool for a decision support system in the production of innovative products [8-10]. The methodology can act as an addition to the risk mitigation technique in the field of high-tech products [11–13]. Also, elements of the methodology can be part of a decision-making system in the field of information technology [14].

The company specializing in the production of household appliances, namely, an automatic washing machines, developed the technique used and tested on the example of optimization of technical characteristics - the electrical conductivity of water. When developing new software, the conductivity parameter was improved from 629 S/cm to 599 S/cm by increasing the amount of water during rinsing to 56 litres and increasing the washing time to 175 minutes. The duration of the laundry rinse process was reduced from 8 to 5 minutes, since it was found that the influence of the factor on the process is insignificant. This model allows for maximum removal of chemical components contained in the powder and other cleaning products, removal of contaminants and other foreign particles. As a result of the introduction of the developed technique in the production of household appliances, starting from April 2018, the indicator responsible for rinsing quality (the amount of washing solution remaining in the laundry after the full wash cycle) was improved to 0.2 mg– eq/dm³, as a result, the requirements of the GOST standard P IEC 60456-2011 were achieved. The number of losses from the 1st and 2nd kind errors also decreased by 10% compared with the previous year, which led to an improvement in the internal quality indicator (level of weighted errors) by 4% and an improvement in the external indicator (number of repairs in the market) by 6%.

References

[1] Dyakonov V P 2003 MATLAB 6/6.1/6.5 + SIMULINK 4/5 in mathematics and modelling. Complete user guide (Moscow: SOLON-Press)
[2] Ryazantsev A I 2016 Product quality assessment by mathematical methods Information technologies and problems of mathematical modeling of complex systems 57-69
[3] GOST ISO 9000 2015 Quality management systems. Fundamentals and vocabulary (Moscow: Standartinform)
[4] GOST R ISO 9001 2015 Quality management systems. Requirements (Moscow: Standartinform)
[5] GOST R ISO 9004 2010 2011 Management to achieve sustainable success of the organization. Approach based on quality management (Moscow: Standartinform)
[6] GOST R ISO 19011-2012 2013 Management system audit guidelines (Moscow: Standartinform)
[7] GOST R MEK 60456-2011 2013 Methods for measuring functional characteristics (Moscow: Standartinform)
[8] Semenova E G and Smirnova M S 2008 Decision support system in multi-criteria tasks of production management of innovative products Scientific and technical bulletin of the St. Petersburg State Polytechnic University 56 57-9
[9] Milskaya E, Mednikov M and Loginova N 2016 Development of the innovative technologies
transfer conceptual model Journal of Applied Endgineering Science 14(3) 383-90

[10] Ryapukhina V N, Suprun E V, Doroshennko Y A, Bukhonova S M and Somina I V 2015 Strategy of effective management for small business at different stages of innovation activity Journal of Applied Endgineering Science 13(2) 117-25

[11] Batkovskiy A M, Semenova E G, Fomina A V, Khrustalev E I and Khrustalev O E 2016 The methodology and mathematical tools to assess and mitigate the risk of creating high-tech products Indian Journal of Science and Technology 9(28) 97659

[12] Batkovskiy A M, Batkovskiy M A, Klochkov V V, Semenova E G and Fomina A V 2017 Analysis of the efficiency of specialization centers formation in high-tech industry Journal of Applied Economic Sciences 12(3) 671-86

[13] Napolskikh D 2015 Clustering of high-tech industrial production: factors and trends Journal of Applied Endgineering Science 16(2) 166-72

[14] Semenova E G, Smirnova M S and Tushavin V A 2014 Decision making support system in multi-objective issues of quality management in the field of information technology Research Journal of Applied Sciences 9(12) 1078-81