Obtaining mathematical models for assessing efficiency of dust collectors using integrated system of analysis and data management STATISTICA Design of Experiments

A V Azarov, N S Zhukova, E Yu Kozlovteva, D R Dobrinsky

Volgograd State Technical University, 28, Lenina Ave., Volgograd, 400005, Russia
E-mail: artazarov@mail.ru

Annotation. The article considers obtaining mathematical models to assess the efficiency of the dust collectors using an integrated system of analysis and data management STATISTICA Design of Experiments. The procedure for obtaining mathematical models and data processing is considered by the example of laboratory studies on a mounted installation containing a dust collector in counter-swirling flows (CSF) using gypsum dust of various fractions. Planning of experimental studies has been carried out in order to reduce the number of experiments and reduce the cost of experimental research. A second-order non-position plan (Box-Bencken plan) was used, which reduced the number of trials from 81 to 27. The order of statistical data research of Box-Benken plan using standard tools of integrated system for analysis and data management STATISTICA Design of Experiments is considered. Results of statistical data processing with significance estimation of coefficients and adequacy of mathematical models are presented.

1. Introduction

The main purpose of the research conducted on installed laboratory facility is to determine influence of various parameters, aspiration system on the values of fractional efficiency of gypsum dust collection of various characteristics selected during gypsum binder production. For the research, various fractions of gypsum dust were selected during gypsum binder production: an aspirating system of a rotor crusher ($D_{50} = 30 \, \mu m$); an aspiration system for gypsum-welding boiler ($D_{50} = 50 \, \mu m$); a pneumatic conveying system of nealit line production ($D_{50} = 10 \, \mu m$).

Installed laboratory facility adjusting quality methods are: total air flow; ratio of air flow sent to the lower input of the counter swirling flows apparatus to the total flow rate; size of the gap between the alignment of the outlet branch pipe inside the separation chamber of the counter swirling flows apparatus and the baffle plate; a device parameter form for cleaning air ducts of aspiration and pneumatic transport systems (ratio air volume supplied through the tangential and coaxial branches).

As the factors of variation, the following was used:
1) $L_o$ - total flow in the system under study, m$^3$/h;
2) $L_o / L_o$ - ratio of air flow through the lower input of the counter swirling flows apparatus to the total flow rate;
3) $I_{i} / D_{a}$ - distance from the end of the axial branch pipe of the counter swirling flows apparatus to the bump washer, expressed by the ratio of the given distance to the diameter of the axial branch pipe;
4) $D_{50}$ - characteristic of gypsum dust weighed, expressed by the median diameter, $\mu m$. 


Factors and variation intervals are presented in Table 1.

| Factors                      | Variation intervals | Levels of factors |
|------------------------------|---------------------|-------------------|
| $L_o$ - total flow in the system under study, m$^3$/h | 200 | 300 | 500 | 700 |
| $L_o/L_o$ - ratio of air flow through the lower input of the counter swirling flows apparatus to the total flow rate | 0.10 | 0.15 | 0.25 | 0.35 |
| $I/D_i$ - distance from the end of the axial branch pipe of the counter swirling flows apparatus to the bump washer, expressed by the ratio of the given distance to the diameter of the axial branch pipe | 0.5 | 0.5 | 1.0 | 1.5 |
| $D_{so}$ - characteristic of gypsum dust weighed, expressed by the median diameter, µm | 20 | 10 | 30 | 50 |

2. Experimental planning and methods used during research

For the convenience of recording experimental conditions, processing experimental data, and obtaining regression models, the factor levels are reduced to a dimensionless form [1-4]:

\[
\begin{align*}
    x_1 &= \frac{T_o - \bar{T}_o}{\Delta T_o}; \\
    x_2 &= \frac{(L_o/L_o) - \bar{(L_o/L_o)}}{\Delta (L_o/L_o)}; \\
    x_3 &= \frac{(I/D_i) - \bar{(I/D_i)}}{\Delta (I/D_i)}; \\
    x_4 &= \frac{D_{so} - \bar{D}_{so}}{\Delta D_{so}},
\end{align*}
\]

where $T_o$, $(L_o/L_o)$, $(I/D_i)$, $D_{so}$ are values of the relevant factors in the center of the plan; $\Delta T_o$, $\Delta (L_o/L_o)$, $\Delta (I/D_i)$, $\Delta D_{so}$ are intervals for changing factor values.

At the first stages of experiment planning, it is clear that there are four factors that can be varied at three levels. When performing a full factorial experiment $3^4$, taking into account combinations of all levels of factors, 81 experiments are required [1-4]. In addition, in order to take into account the peculiarities of a random process of changing the values of the integrated distribution functions of the dispersed dust composition at the input and at the output from the counter swirling flows apparatus, a series of at least 10 selected dust samples for further dispersion analysis will be held in each of the experiments. In accordance with this, it can be concluded that it is necessary to carry out an experiment of another kind with a view to its optimization [5].

The main task of the experiment planning was to establish the minimum required number of experiments. For the experiment, a non-composite plan of the second order was chosen [3, 4]. A similar plan of the experiment is a certain sample of lines from a full factorial experiment of type $3^4$.

Non-composition plan of the second order (Box-Benken plan) during the experiment with four factors makes it possible to reduce the number of experiments to 27. Unlike the central compositional plan, where the factors of variation are changed at five levels, non-composite variable varies only at the three levels, making experiment cheaper and easier. Also, a rotatable non-composite plan of the second order with four factors implies four experiments less than a rotatable central composition plan of the second order [5]. A matrix for planning a non-composite plan of the second order (Box-Benken plan) for current experimental studies in dimensionless form is presented in Table 2.

As the response functions (optimization parameters), the total dust concentration at the inlet of the dust collector, $C_{so}$, g/m$^3$; total dust concentration at the outlet of the dust collector, $C_{out}$, g/m$^3$; total purification efficiency, $\eta_o$, %; fractional cleaning efficiency (0 µm - 2.5 µm), $\eta_{2.5}$, %; fractional cleaning
efficiency (0 µm - 10 µm), η_{10}, %; fractional cleaning efficiency (0 µm - 40 µm), η_{40}, %; fractional cleaning efficiency (0 µm - 140 µm), η_{140}, % are selected.

The particle sizes in each of the response functions are selected in accordance with the existing classification of dust by dispersion: I - very coarse-grained dust, dimensions (more than 140 µm); II - coarsely dispersed dust (40 µm-140 µm); III - medium-dispersed dust (10 µm -40 µm); IV - fine dust (1 µm - 10 µm) [6]. Since fine dust with a particle size of up to 2.5 µm entering the atmosphere with venting emissions and easily adhering the respiratory organs represent a particular danger to humans, this fraction (0 µm - 2.5 µm) was also included into the list of functions under study response [7].

### Table 2. Levels and intervals of variation of the determining factors

| No. experience | Total flow in the system under study, L_{0}, m³/h | Cost ratio of the counter swirling flows apparatus, L_{0}/L_{0} | Distance from the axial nozzle to the bump washer, expressed by the ratio, I/D_{n} | Characteristics of used dust sample, expressed by the median diameter, D_{50}, µm |
|----------------|-------------------------------------------------|-------------------------------------------------|----------------------------------------|-------------------------------|
| 1              | -1                                              | -1                                              | 0                                     | 0                             |
| 2              | 1                                               | -1                                              | 0                                     | 0                             |
| 3              | -1                                              | 1                                               | 0                                     | 0                             |
| 4              | 1                                               | 1                                               | 0                                     | 0                             |
| 5              | 0                                               | 0                                               | -1                                    | -1                            |
| 6              | 0                                               | 0                                               | 1                                     | -1                            |
| 7              | 0                                               | 0                                               | -1                                    | 1                             |
| 8              | 0                                               | 0                                               | 1                                     | 1                             |
| 9              | 0                                               | 0                                               | 0                                     | 0                             |
| 10             | -1                                              | 0                                               | 0                                     | -1                            |
| 11             | 1                                               | 0                                               | 0                                     | -1                            |
| 12             | -1                                              | 0                                               | 0                                     | 1                             |
| 13             | 1                                               | 0                                               | 0                                     | 1                             |
| 14             | 0                                               | -1                                              | -1                                    | 0                             |
| 15             | 0                                               | 1                                               | -1                                    | 0                             |
| 16             | 0                                               | -1                                              | 1                                     | 0                             |
| 17             | 0                                               | 1                                               | 1                                     | 0                             |
| 18             | 0                                               | 0                                               | 0                                     | 0                             |
| 19             | -1                                              | 0                                               | -1                                    | 0                             |
| 20             | 1                                               | 0                                               | -1                                    | 0                             |
| 21             | -1                                              | 0                                               | 1                                     | 0                             |
| 22             | 1                                               | 0                                               | 1                                     | 0                             |
| 23             | 0                                               | -1                                              | 0                                     | -1                            |
| 24             | 0                                               | 1                                               | 0                                     | -1                            |
| 25             | 0                                               | -1                                              | 0                                     | 1                             |
| 26             | 0                                               | 1                                               | 0                                     | 1                             |
| 27             | 0                                               | 0                                               | 0                                     | 0                             |

One of the main results of the research is to obtain a mathematical model (regression equation) for each optimization parameter. Based on a priori information, it is known that the studied processes can be described by second-order polynomials. Nevertheless, the regression model was determined sequentially in order to take into account the effects of interaction of experimental factors. The results of experimental studies were processed in the integrated system of data analysis and management - STATISTICA Design of Experiments (Planning of experiments) [8, 9].

Instrumental measurements were carried out in accordance with standard methods.

The results of dispersion analysis were obtained according to the method of disperse composition measurement of dust using computers and digital means of processing the selected samples, computer programs "Dust 1".
Approximate representations of obtained results analysis of dispersed dust at the inlet and outlet of the dust collector are determined using a software package of computational models for approximate representation of the integrated functions of the particle mass distribution over dust diameters in computer algebra system, CAS Maple [10].

The random nature of the change in the values of the integral distribution functions of disperse composition of dust at the input and at the output from the counter swirling flows apparatus is taken into account by processing of the data for one point and one series of measurements in one data array of the proposed software complex computational models for approximating the submission of the subsequent application of the operator minimization of errors and cycle matching critical points of the unknown function.

3. Results and discussion

During experimental studies, the values of $C_{in}$, $C_{out}$, $\eta_0$, $\eta_{2.5}$, $\eta_{10}$, $\eta_{40}$, $\eta_{140}$ were obtained on mounted laboratory apparatus. The main results of the experimental studies are presented in Table 3.

| No. experience | Dust concentration at the inlet to dust collector, $C_{in}$, g / m$^3$ | Dust concentration at the outlet of dust collector, $C_{out}$, g / m$^3$ | Total purification efficiency of dust-gas cleaning installation, $\eta_0$, % | Fractional cleaning efficiency of dust-gas cleaning installation, $\eta_{2.5}$, % | Fractional cleaning efficiency of dust-gas cleaning installation, $\eta_{10}$, % | Fractional cleaning efficiency of dust-gas cleaning installation, $\eta_{40}$, % | Fractional cleaning efficiency of dust-gas cleaning installation, $\eta_{140}$, % |
|---------------|--------------------------------------------------|--------------------------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| 1             | 15.19                                            | 2.21                                             | 85.25                           | 10.87                           | 35.27                           | 39.25                           | 72.95                           |
| 2             | 6.09                                             | 1.03                                             | 82.82                           | 6.05                            | 26.99                           | 48.12                           | 74.47                           |
| 26            | 7.70                                             | 1.04                                             | 86.31                           | 14.86                           | 49                              | 60.96                           | 94.63                           |
| 27            | 9.12                                             | 0.49                                             | 94.48                           | 14.83                           | 48.92                           | 76.85                           | 89.33                           |

After receiving the results and compiling the final table, statistical processing was performed in the integrated data analysis and management system - STATISTICA Design of Experiments [8, 9]. The purpose of the processing is to obtain and determine the significance of the regression equations coefficients of the following form:

$$ y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_k + b_4 x_{1x2} + b_{k-1x3} x_k + b_{kx1} x_{k-1} + b_{kk} x_k^2 $$

According to the results of statistical processing of the obtained data, the coefficients of regression equations are determined; their evaluation, as well as an assessment of the main effects and effects of interaction, is carried out. The main results of the statistical processing of experimental research data are presented in Table 4.
Table 4. Results of statistical data processing

| Designation of the regression model coefficient | Coefficient value | Estimation of regression coefficient (*Std. Err.) | Estimation of regression coefficient (t-test) | Effect Value | Effect evaluation (*p-value) | Estimation of regression coefficient (t-test) | Estimation of regression coefficient (*p-value) |
|-----------------------------------------------|------------------|-----------------------------------------------|-----------------------------------------------|--------------|-------------------------------|-----------------------------------------------|-----------------------------------------------|
| Average free member \((b_0)\)                | 95.523           | 1.574                                         | 60.694                                        | 0.000        | 83.748                        | 0.909                                         | 92.165                                        |
| \((1) L_{\text{total}} (L)\)                 | -1.005           | 0.787                                         | -1.277                                        | 0.218        | -2.010                        | 1.574                                         | -1.277                                        | 0.218                                        |
| \((2) L_{\text{total}} (Q)\)                 | -6.138           | 1.180                                         | -5.200                                        | 0.000        | 6.138                         | 1.180                                         | 5.200                                         | 0.000                                        |
| \((3) L_{\text{total}} (L)\)                 | -0.330           | 0.787                                         | -0.419                                        | 0.680        | -0.660                        | 1.574                                         | -0.419                                        | 0.680                                        |
| \((4) L_{\text{total}} (Q)\)                 | -7.263           | 1.180                                         | -6.153                                        | 0.000        | 7.263                         | 1.180                                         | 6.153                                         | 0.000                                        |
| \((5) L_{\text{total}} (L)\)                 | -1.180           | 0.787                                         | -1.288                                        | 0.218        | -2.010                        | 1.574                                         | -1.277                                        | 0.218                                        |
| \((6) L_{\text{total}} (Q)\)                 | -3.065           | 1.180                                         | -2.596                                        | 0.018        | 3.065                         | 1.180                                         | 2.596                                         | 0.018                                        |
| \((7) L_{\text{total}} (L)\)                 | 0.728            | 0.787                                         | 0.924                                         | 0.367        | 1.455                         | 1.574                                         | 0.924                                         | 0.367                                        |
| \((8) L_{\text{total}} (Q)\)                 | -1.197           | 1.180                                         | -1.014                                        | 0.324        | 1.197                         | 1.180                                         | 1.014                                         | 0.324                                        |

Fractional cleaning efficiency of dust-gas cleaning installation, \(\eta_{2.5\%}\) %

| Average free member \((b_0)\)                | 14.313           | 1.301                                         | 11.000                                        | 0.000        | 12.412                        | 0.673                                         | 18.439                                        |
| \((1) L_{\text{total}} (L)\)                 | -1.961           | 0.651                                         | -3.014                                        | 0.011        | -3.588                        | 1.166                                         | -3.078                                        | 0.010                                        |
| \((2) L_{\text{total}} (Q)\)                 | -4.492           | 0.976                                         | -4.603                                        | 0.001        | -4.450                        | 0.874                                         | 5.089                                         | 0.000                                        |
| \((3) L_{\text{total}} (L)\)                 | 1.240            | 0.651                                         | 1.906                                         | 0.081        | 5.313                         | 1.166                                         | 4.557                                         | 0.001                                        |
| \((4) L_{\text{total}} (Q)\)                 | -1.841           | 0.976                                         | -1.886                                        | 0.084        | 0.424                         | 0.874                                         | 0.645                                         | 0.636                                        |
| \((5) L_{\text{total}} (L)\)                 | -2.521           | 0.651                                         | -2.387                                        | 0.002        | -2.508                        | 1.166                                         | -4.467                                        | 0.001                                        |
| \((6) L_{\text{total}} (Q)\)                 | -0.922           | 0.976                                         | -0.945                                        | 0.363        | 1.005                         | 0.874                                         | 1.150                                         | 0.273                                        |
| \((7) L_{\text{total}} (L)\)                 | -3.497           | 0.651                                         | -3.575                                        | 0.000        | -6.993                        | 1.166                                         | -5.998                                        | 0.000                                        |
| \((8) L_{\text{total}} (Q)\)                 | 2.987            | 0.976                                         | 3.060                                         | 0.010        | -3.028                        | 0.874                                         | -3.463                                        | 0.005                                        |
| \((9) D_{50} (L)\)                           | -0.205           | 1.127                                         | -0.182                                        | 0.859        | -0.410                        | 2.019                                         | -0.203                                        | 0.843                                        |
| \((10) D_{50} (Q)\)                          | -1.765           | 1.127                                         | -1.566                                        | 0.143        | -4.530                        | 2.019                                         | -2.243                                        | 0.045                                        |
| \((11) D_{50} (L)\)                          | 2.713            | 1.127                                         | 2.407                                         | 0.033        | 5.425                         | 2.019                                         | 2.686                                         | 0.020                                        |
| \((12) D_{50} (Q)\)                          | -0.585           | 1.127                                         | -0.519                                        | 0.613        | -1.670                        | 2.019                                         | -0.827                                        | 0.424                                        |
| \((13) D_{50} (L)\)                          | 0.195            | 1.127                                         | 0.173                                         | 0.865        | -0.610                        | 2.019                                         | -0.302                                        | 0.768                                        |
| \((14) D_{50} (Q)\)                          | 0.208            | 1.127                                         | 0.184                                         | 0.857        | -0.585                        | 2.019                                         | -0.290                                        | 0.777                                        |

* Std.Err. - Value of the difference between average value of the sample and the mean of General population;
* p-value - Value of the actually achieved significance level;
* Significant coefficients and effects of interaction in the table (bold).

The results of statistical data processing and a visual representation of the most significant effects on response functions were demonstrated using Pareto diagrams. On the data charts, the authors presented the assessment of effects; effects of the columns were located in the absolute magnitude of
the values from largest to smallest. The value of each effect was represented by a column, the bars were intersected by a line indicating what value of the effect should be in order to be statistically significant [8].

Estimates of the main linear effects (L) can be interpreted as the difference between the average response at high and low settings of the relevant factors. The estimation of the principal quadratic (nonlinear) effect (Q) can be interpreted as the difference between the average response at the center (middle point) and the response combination for the high and low settings of the corresponding factor.

Estimates of one-dimensional linear interaction effects can be interpreted as half the difference between the linear principal effects of one factor at the high and low levels of another factor. Similarly, interactions of quadratic effects can be interpreted as half the difference between the quadratic effect of one factor on the corresponding settings of the other (quadratic-linear interaction) or a combination of response at medium, high and low settings (quadratic-quadratic interaction).

The adequacy of the obtained mathematical models was tested with the help of Fisher's calculated F-test, which was compared with the table value [1-4].

If the found value of the Fisher criterion is less than the table value at the accepted level of significance (in technique most often the significance level is \( q = 5\% \)) and the corresponding degrees of freedom, then the hypothesis of the adequacy of the obtained model is accepted. Below are the main adequate mathematical models for optimization parameters (response functions) with the exception of insignificant coefficients:

1. Total purification efficiency of dust-gas cleaning installation, \( \eta_0 \), %:
   \[
   \eta_0 = 95.52339 - 6.13833x_1^2 - 7.26333x_2^2 - 1.80083x_3 - 3.06458x_4^2
   \]
2. Fractional cleaning efficiency of dust-gas cleaning installation, \( \eta_{2.5} \), %
   \[
   \eta_{2.5} = 14.31333 - 1.79417x_1 - 4.45042x_1^2 + 2.65667x_2 - 2.60417x_3 - 3.49667x_4 + 4.45042x_4^2 - 2.265x_1x_3 + 2.71250x_1x_4
   \]
   where \( x_1 \) - total flow in the system under study, \( L_0 \), \( m^3/h \); \( x_2 \) - cost ratio of the counter swirling flows apparatus, \( L_0/L_0 \); \( x_3 \) - distance from the axial nozzle to the bump washer, expressed by ratio, \( L/D_n \); \( x_4 \) - characteristic of used dust sample, expressed by median diameter, \( D_{50} \), \( \mu m \); factors of variation in dimensionless quantities.

4. Conclusion
When planning technical experiments to obtain mathematical models, a non-composite plan of the second order was used (Box-Bencken plan). Application of this plan has made it possible to reduce the costs of conducting experimental studies by reducing the required number of trials from 81 to 27.

Integrated system for analysis and data management - STATISTICA Design of Experiments has standard tools for statistical processing of experimental data on Box-Bencken scheduling matrix.

According to the results of statistical processing of the obtained data, the coefficients of regression equations were determined by the program, their evaluation and estimation of the main and interaction effects. As a result of statistical data processing, it is possible to visualize most significant interaction effects on the response function (Pareto diagrams). The use of these diagrams makes it possible to clearly evaluate the significance of interaction effects of variation factors.

On the basis of obtained mathematical models with the help of profiling function of integrated system for analysis and data management - STATISTICA Design of Experiments, diagrams of response surfaces (3D diagrams) of optimization parameters from the pair wise combination of factors were constructed. This graph reflects the forecast equation, which gives the corresponding response surface. In these diagrams, the independent variables (variation factors) are presented in pairs; each region of the resulting response surface is a different combination of levels of two variation factors. On axes of 3D diagram, values of the factor levels in dimensionless quantities are plotted, and desired value of the response functions is 1.

Thus, obtaining mathematical models for efficiency evaluating of the dust collectors using an integrated system for analysis and data management - STATISTICA Design of Experiments - made it possible not only to obtain the desired mathematical models but also to carry out a comprehensive
analysis of the results obtained and their visual representation.

References

[1] Voznesenskiy V A 1981 *Statistical methods of experiment planning in technical and economic studies* (Moscow: Finance and Statistics)

[2] Zazhigaev L S, Kish'yan A A and Romanikov Yu I 1978 *Methods for planning and processing the results of a physical experiment* (Moscow: Atomizdat)

[3] Dmitrieva L S, Kuz'mina L V and Moshkarev L M 1984 *Planning an experiment in ventilation and air conditioning* (Irkutsk : Irkut. un-t)

[4] Baptizmanskiy V I, Yakovlev Yu N and Paniotov Yu S 1992 *Organization of the experiment* (Kiev: UMK VO)

[5] Kel'ner R 2004 *Analytical chemistry. Problems and approaches* (Moscow: Mir)

[6] Vetoshkin A G 2005 *Processes and apparatus for dust removal* (Penza: PSU)

[7] Azarov V N, Esina E Yu 2009 The application of the "dissection" method in the analysis of the dispersion of dust in the air in the working zone of enterprises of the construction industry and machine building *Mechanical engineering and technosphere of the XXI century (XVI International Scientific and Technical Conference)* (volume 1) (Donetsk: DonNTU) pp 30-33

[8] Howard J Seltman 2015 Experimental Design and Analysis

[9] Bureeva N N 2007 *Multidimensional statistical analysis using the PPP STATISTICA* (Nizhny Novgorod)

[10] Azarov A V, Zhukova N S and Sidorenko V F 2016 Improving the Computational Model for Approximation of Particle Functions over Diameter of Dust in the Work Area and at the Border of the Sanitary Protection Zone *2ND International conference of industrial engineering (ICIE 2016)* (Vol. 150) ed A A Radionov (Procedia Engineering) pp 2073-2079