Calculation models for the dispersed composition integral functions approximation of the dust in the air

N S Zhukova*, E A Kalyuzhina, A V Azarov, A B Strelyaeva
Volgograd State Technical University, 28, Lenin ave., 400005, Russia

E-mail: pns15@yandex.ru

Abstract. The danger of the small dust particles’ influence with sizes PM10 and PM2.5 is indicated. Computational models for approximating the integral functions of the dispersed dust composition are presented. The optimal planning problem is solved under the conditions of the Gauss-Markov theorem for determining the regression coefficients and a set of functions of a two-link and three-link spline.

1. The dust dispersed composition analysis
Numerous studies show that the dustiness of the working zone and atmospheric air, depending on the equipment condition, the nature of production operations, varies widely. The greatest danger is represented by the dust particles of small sizes PM10 and PM2.5, which can penetrate the human body [1, 2]. Thus, for an objective assessment of the danger to human health in the air of the working and the sanitary protection zone, it is necessary to know the percentage of particles with extremely small sizes. To do this, the research on the dispersed composition of the dust released from equipment in aspiration systems and the air in the working zone at various enterprises has been carried out [3].

To assess the dust particles’ size, and therefore, to select the most efficient dust collecting equipment, its adjustment and proper operation, a complex of dust measurements is carried out. In this complex, the analysis of the dust dispersed composition is most important [4].

For greater clarity and convenience of analysis, the mass distribution of particle sizes is represented graphically – Figure 1 [5, 6]. In this case, the laws of particle distribution are characterized by the integral distribution curves, but the dispersed composition of dust cannot be described in the probabilistic-logarithmic coordinates of one straight line, i.e. the log-normal law.

When analyzing the dust dispersed composition, a large scatter of data is often obtained. To assess the values of PM10 and PM2.5, it is important to approximate the results of the field studies.

For this, two options are proposed for approximating the experimental values of the integral function of the particle mass distribution over diameters, presented in the probability-logarithmic grid of the curve:
- using a two-link spline consisting of two functions - linear and nonlinear (hyperbolic) [7, 8];
- using a three-link spline, consisting of three functions: linear, parabola and hyperbola [9, 10].
2. **Two-spline approximation**

The use of a two-link spline approximation is due to the dust particles with a diameter of more than 10 microns.

For the graphical representation of the dust dispersed composition studies results the two-link spline is used, the range of values of all particle sizes $\delta$ is divided into two sections: the first section $x \leq x_{\text{nod}.1}$, the second section $x_{\text{nod}.1} \leq x < x_{\text{nod}.2}$ (Figure 2).

![Figure 2. The dispersed dust composition integral distribution function approximation by the two-link spline: 1 - linear function; 2 - nonlinear function; 3 - tangent.](image)

The function $y = \text{inverf}(x)$ is inverse to the probability integral function [11, 12]. So, the function $D(\delta)$ will be presented as (1):
The function by an increasing hyperbolic function having a vertical asymptote is approximated on 

\[ [x_{\text{nod}1}, x_{\text{nod}2}] \] at the location \( x_{\text{nod}2} = \log x_{\text{nod}2} \):

\[
y_2 = a_2 + \frac{a_3}{x_{\text{nod}2} - x}.
\]  

(3)

For the optimal curve matching \( y_1 \) and \( y_2 \) are set on the plots \( [0, x_{\text{nod}1}] \) and \( [x_{\text{nod}1}, x_{\text{nod}2}] \), we assume the equality \( Y_1(x_{\text{nod}1}) = Y_2(x_{\text{nod}2}) \), namely:

\[
\frac{d}{dx} Y_1(x) \bigg|_{x=x_{\text{nod}1}} = \frac{d}{dx} Y_2(x) \bigg|_{x=x_{\text{nod}1}}
\]  

(4)

The odds \( a_1, a_2, a_3, k \) selection can be carried out from the conditions:

1. The functions \( y_1 \) and \( y_2 \) at the point \( x_{\text{nod}1} \) equality:

\[
\begin{align*}
Y_1(x_{\text{nod}1}) &= a_1 + k x_{\text{nod}1}, \\
Y_2(x_{\text{nod}1}) &= a_2 + \frac{a_3}{x_{\text{nod}2} - x_{\text{nod}1}}.
\end{align*}
\]  

(5)

2. The derived functions \( y_1 \) and \( y_2 \) at the point \( x_{\text{nod}1} \) equality:

\[
\begin{align*}
Y'_1(x_{\text{nod}1}) &= k, \\
Y'_2(x_{\text{nod}1}) &= \frac{a_3}{(x_{\text{nod}2} - x_{\text{nod}1})^2}.
\end{align*}
\]  

(6)

The desired function (1) should be selected from the condition of minimum of the quadratic form

\[
\sum_{i=1}^{N} \epsilon_i^2 = \epsilon^T \cdot \epsilon
\]  

when changing three parameters \( x, y_0, k \). This minimum is defined as follows:

1. For each \( x \), it is necessary to find the value \( \min_{y_0, k} \sum_{i=1}^{N} \epsilon_i^2 \) and corresponding to this \( x \) value \( y_0, k \).

2. We find the minimum form \( x \), and for each value \( x \) where this minimum is reached, we find the values \( y_0, k \) and, therefore, we obtain the desired spline function (1).

The first stage is reduced to the optimal planning problem [13]. We introduce the following notation:

\( Y = (y_1...y_N)^T \) – is the integral function values’ vector (T – is the transpose symbol); \( \theta = (y_0, k)^T \) – is the vector of the unknown parameters; \( \epsilon = (\epsilon_1...\epsilon_N)^T \) – is the deviation vector; \( F(x) \) – is the matrix size \((N \times 2)\), dependent on \( x \) (\( x_i \leq x \) at \( i \leq n \), \( x_i \geq x \) at \( n < i \leq N \)).
The problem of finding a vector $\theta(y_0, k)$ for each fixed $x$ reduces to the optimal planning problem [13], which in matrix form has the form:

$$
inverf Y = F(x_0) \bar{\theta} + \bar{\varepsilon}$$

(8)

According to the Gauss-Markov theorem, provided that $\det F^T \cdot F \neq 0$, the least squares methods assessments $\bar{\theta}$ are uniquely determined, are the best linear unbiased estimates and have the form:

$$
\bar{\theta}(x_0) = (F^T F)^{-1} F^T Y
$$

(9)

Therefore, for each $x$ the least by the parameters $y_{nod1}$, $k_1$, $k_2$ value $\bar{\varepsilon}^T \varepsilon$ can be defined as follows:

$$
\min_{y_0, k_1, k_2} \varepsilon^T \varepsilon(x_0) = (Y - F(F^T F)^{-1} F^T Y)^T (Y - F(F^T F)^{-1} F^T Y)
$$

(10)

At the second stage, the quantities $\bar{\varepsilon}^T \varepsilon$ by $x$ are determined. Equating the derivative to zero then the determines $x$, at which a minimum of shape is achieved $\bar{\varepsilon}^T \varepsilon(x)$, and, therefore, in which the spline function (1) is the best approximation of the experimental values:

$$
x = \frac{k_1 \sum_{i=1}^{n} (y_i - y_0 - k_1 x_i) + \sum_{i=1}^{N} (y_i - y_0 - k_2 x_i)}{k_1^2 n + k_2^2 (N-n)}
$$

(11)

where $y_0$, $k_1$, $k_2$ are defined as the vector components $\bar{\theta}(x_0)$ out of equality (9).

3. Three-spline approximation

The use of a three-link spline approximation is a typical characteristic of the fine dust with a diameter of less than 10 microns. For a graphical representation of the dust dispersed composition study results, using a three-link spline, the obtained measurement results are divided into three sections as follows: $\delta_1...\delta_{N1}$ - located on the first segment $[0; \exp x_{nod1}]$; $\delta_{N1+1}...\delta_{N2}$ is on the second segment $[\exp x_{nod1}; \exp x_{nod2}]$; $\delta_{N2}$ is on the third segment $[\exp x_{nod2}; \exp x_{nod3}]$ (fig. 3).

In this case, the function $D(\delta_i)$, will be represented by the following set of functions:
Figure 3. The dispersed dust composition integral distribution function approximation by the three-link spline: 1 - linear; 2 - parabola; 3 - hyperbole

By the analogy with the presentation of a two-link spline, the first stage is reduced to finding a minimum $\sum_{i=1}^{N} \varepsilon_i^2$. The matrix $F(x_{nod.1}, x_{nod.2}, x_{nod.3}, S)$, the size $(N \times 6)$ will have the form:

$$F(x_{nod.1}, x_{nod.2}, x_{nod.3}, S) = \begin{pmatrix}
    x_1 & 1 & 0 & 0 & 0 & 0 \\
    x_2 & 1 & 0 & 0 & 0 & 0 \\
    \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
    x_{N-1} & 1 & 0 & 0 & 0 & 0 \\
    0 & 0 & 1 & x_{N+1} & x^2_{(N+1)} & 0 \\
    0 & 0 & 1 & x_{N+2} & x^2_{(N+2)} & 0 \\
    0 & 0 & 1 & x_{N+3} & x^2_{(N+3)} & 0 \\
    0 & 0 & 0 & 0 & 0 & 1 \\
    0 & 0 & 0 & 0 & 0 & 1 \\
    0 & 0 & 0 & 0 & 0 & 1 \\
\end{pmatrix}$$

$$\text{inverf}\left(D\left(\delta_i\right)\right) = \begin{cases}
(k \log \delta_i + p) + \varepsilon_i, & \text{if } \delta_i < \exp x_{nod.1} \\
(c + b \log \delta_i + a \log^2 \delta_i) + \varepsilon_i, & \text{if } \exp x_{nod.1} \leq \delta_i < \exp x_{nod.2} \\
\frac{S}{x_{nod.3} - \log \delta_i}, & \text{if } \exp x_{nod.2} \leq \delta_i < \exp x_{nod.3}
\end{cases}$$

(12)
The problem of finding a vector \( \vec{b} (k, p, c, b, a, S) \) for each fixed set \( x_{\text{nod}1}, x_{\text{nod}2}, x_{\text{nod}3} \) is reduced to the optimal planning problem, which in matrix form has the form (9). At the second stage, the quantities \( x_{\text{nod}1}, x_{\text{nod}2}, x_{\text{nod}3} \) from the conditions are:

1. The functions \( y_1 \) and \( y_2 \) at the point \( x_{\text{nod}1} \) equality:
   \[
   \begin{cases}
   y_1(x_{\text{nod}1}) = k x + p, \\
   y_2(x_{\text{nod}1}) = ax^2 + bx + c.
   \end{cases}
   \]
   \[ (14) \]

2. The derived functions \( y_1' \) and \( y_2' \) at the point \( x_{\text{nod}1} \) equality:
   \[
   \begin{cases}
   y_1'(x_{\text{nod}1}) = k, \\
   y_2'(x_{\text{nod}1}) = 2ax + b.
   \end{cases}
   \]
   \[ (15) \]

3. The functions \( y_2 \) and \( y_3 \) at the point \( x_{\text{nod}2} \) equality:
   \[
   \begin{cases}
   y_2(x_{\text{nod}2}) = ax^2 + bx + c, \\
   y_3(x_{\text{nod}2}) = \frac{S}{x_{\text{nod}3} - x_{\text{nod}2}}.
   \end{cases}
   \]
   \[ (16) \]

Based on the obtained values of the coefficients, depending on the fact which values of the sections \([0; exp \ x_{\text{nod}1}], [exp \ x_{\text{nod}1}; exp \ x_{\text{nod}2}], [exp \ x_{\text{nod}2}; exp \ x_{\text{nod}3}]\) are found \( \delta = 2.5 \ \mu m \) and \( \delta = 10 \ \mu m \) the corresponding calculation formula is selected from (12) to find the finely divided fractions fraction \( D(\delta = 2.5) \) and \( D(\delta = 10) \). Then the values PM10 and PM2.5 are found.

**Summary**

The optimal planning problem under the conditions of the Gauss-Markov theorem made it possible to uniquely determine the regression coefficients and the functions set of the two-link and three-link spline approximating the dust particle mass distribution function over the diameters. For the different types of dust, it is necessary to vary the type approximating the distribution function of the dust particles’ mass along the curve diameters. This allows us to describe the fine dust as accurately as possible, which is especially relevant in assessing the proportion of PM10 and PM2.5, as well as for solving a number of environmental safety and labor protection problems.

**References**

[1] Azarov V N, Tertishnikov I V, Kalyuzhina E A and Marinin N A 2011 On the assessment of the concentration of fine dust (PM10 and PM2.5) in air Bulletin of VolgGASU. Ser. Construction and architecture 25 (44) 402-06.

[2] Azarov V N, Tertishnikov I V and Marinin N A 2012 Rationing of PM10 and PM2.5 as social standard of quality of life in areas where construction industry enterprises are located Housing construction 3 20-23.

[3] Azarov V N 2013 On the study of the dispersed composition of dust in the air V International Scientific Conference "System Synthesis and Applied Synergetics" FSAEI HPE "North-Caucasian Federal University" 1 6-14.

[4] Marinin N A, Azarov V N, Bogomolov A N, Sergina N M and Chizhov N I 2011 On the dispersed composition of dust in aspiration and pneumatic cleaning systems Bulletin of VolgGASU. Ser. Construction and architecture 25 (44) 218-24.

[5] Allen T 1997 Particle Size Measurement V.1 Powder sampling and particle size measurement (CHAPMAN & HALL)

[6] Kouzov P A 1987 Fundamentals of analysis of the dispersed composition of industrial dusts and crushed materials (Chemistry, Leningrad).
[7] Ponomareva N S, Gladkov E V, Shibakov V A and Gvozdkov I A 2013 About the calculation model of the dispersed composition of dust in the air of the working area and environmental engineering systems Indoor air and environmental quality: Indoor air and environmental quality: Mater. XI Int. scientific conf. Hanoi. VolgGASU. Volgograd 29-35.

[8] Bogomolov A N, Zhukova N.S., Azarov D.V. and Khegay D.S. 2016 Evaluation of the dust factor in the air of the working area at chalk production enterprises Bulletin of VolgGASU. Ser. Construction and architecture 43 175-88.

[9] Kalyuzhina E A, Kireeva A I, Marinin N A, Gladkov E V and Mikhailovskaya Yu S 2013 On the approximation of the integral functions of the dispersed composition of dust in air Alternative energy and ecology 14 52-56.

[10] Azarov A V, Zhukova N S, 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 Procedia Engineering (Elsevier publishing) 2073-2079.

[11] Philip J R 1960 The function inverf Austral. J. Phys. 13 13 – 20.

[12] Strecok A J 1968 On the calculation of the inverse of the error function Math. Comput. 22 144-58.

[13] Azarov V N 2004 A comprehensive assessment of the dust environment and the development of measures to reduce the dustiness of the air of industrial enterprises: dis. ... Dr. tech. sciences. Rostov-on-Don.