Powder Technology 325 (2018) 510–518

Contents lists available at ScienceDirect

Powder Technology

journal homepage: www.elsevier.com/locate/powtec

Non-parametric estimation of particle size distribution from spectral extinction data with PCA approach

Zhenzong He, Junkui Mao *, Xingsi Han

Jiangsu Province Key Laboratory of Aerospace Power System, College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, PR China

ARTICLE INFO

Article history:
Received 7 August 2017
Received in revised form 9 October 2017
Accepted 18 November 2017
Available online 22 November 2017

Keywords:
Tikhonov regularization method
Particle size distribution
Principle component analysis approach
Non-parametric estimation

ABSTRACT

Non-parametric estimation of particle size distribution (PSD) is carried out with the help of spectral extinction method and Tikhonov regularization method. Meanwhile, a selection principle of optimal measurement wavelengths on the basis of principle component analysis (PCA) approach is proposed to improve the retrieval accuracy. Results show that the optimally selected wavelengths can ensure retrieval accuracy, compared with randomly selected wavelengths. Then, four common PSDs, namely Log-Normal (L-N), Rosin-Rammer (R-R), Normal (N-N) and Gamma distributions, are estimated under different random measurement errors. Numerical tests show that the present methodology can be successfully used to reconstruct the PSDs with high stability in the presence of random errors and low susceptibility to the shape of distributions. Finally, the PSD of aerosol measured over Harbin in China and the PSD of microalgae cell are retrieved, respectively. The study shows that increasing the number of subintervals of the particle size range will improve the retrieval accuracy as well as increase the cost of experiment, so the number of subintervals should be selected reasonably. Moreover, to obtain more reliable results, the measurement sample size should be as large as possible. As a whole, the present methodology provides a reliable and effective technique to estimate the PSDs non-parametrically from the spectral extinction signals.

1. Introduction

Particle size distribution (PSD) usually plays a crucial role in influencing the production quality of industry, environmental pollution detection, and human health protection etc. [1,2]. Considerable research has been carried out on reconstructing the PSDs, and several professional organizations have been set up to study the PSDs of several special particles, such as the AERONET, MODIS, which are global ground-based aerosol observation networks established to study the properties of the atmospheric aerosols [3,4]. However, accurate prediction of the PSDs is still regarded as a very challenging problem and needs further research.

Generally, the PSD cannot be determined directly, but obtained by using non-contact optical measurement methods, e.g. spectral extinction method, angle light scattering method and dynamic light scattering method, combined with inverse problem solving techniques [5–10]. The spectral extinction method is regarded as one of the most promising techniques in studying the PSDs for only requiring a simple optical layout and can easily being realized by adapting of a commercial spectrophotometer. Moreover, the spectral extinction method can also provide a sufficient size-measuring resolution over a broad range of sizes [11–13].

According to whether there is prior information about the function type of PSDs, e.g. Log-Normal (L-N) distribution function, the reconstruction of PSDs can be divided into parametric estimation and non-parametric estimation approaches [14]. For the parametric estimation approach, the function type of the PSD is known beforehand, so there are usually only a few characteristic parameters in the distribution functions needed to be obtained. However, sometimes, there is not enough prior information about the function type of the PSD. If we still want to estimate the PSD parametrically, an assumed function type of the PSD should be determined first, and the deviation between the actual distribution and estimated results is inevitable [15]. To solve the problem, non-parametric estimation approach is proposed. Unlike the parametric estimation approach, the particle size range is divided into many subintervals and corresponding PSD is retrieved by measuring the multispectral transmittance signals. For being independent of any given priori information of the PSD, the non-parametric estimation approach can avoid retrieval error of the parametric estimation approach due to the deviation of assumed distribution function from the true distribution. The only obstacle for the non-parametric estimation approach is to solve the Fredholm integral equation of the first kind, a well-known ill-posed problem which leads to highly unstable solutions even under small noise components in the measured signals [14]. To offset the influences of ill-conditioning and solve the

* Corresponding author at: Jiangsu Province Key Laboratory of Aerospace Power System, College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, 29 Yudao Street, Qinhuai District, Nanjing 210016, PR China.
E-mail address: mjkpe@nuaa.edu.cn (J. Mao).

https://doi.org/10.1016/j.powtec.2017.11.050
0032-5910/

Nomenclature

D	particle diameter, μm
$f(D)$	unknown volume frequency distribution
F_{obj}	objective function
I	intensity of the laser, $W/(m^2 \cdot sr)$
I_0	total intensity of the laser, $W/(m^2 \cdot sr)$
L	geometric thickness of the particle system, m
m	complex refractive index of particle
N_D	total number concentration of the suspended particle system
N	number of subintervals
Q_{ext}	extinction efficiency
R_{ij}	correlation coefficients between the column vectors $\tau(\lambda_i)$ and $\tau(\lambda_j)$
S	number of the incident wavelengths
T	spectral transmittance signals matrix

Greeks symbols

λ	incident wavelength of the laser, μm
δ	relative deviation of the particle size distribution
τ	transmittance of the particle system
σ	narrowness indices of the distribution
α, β	characteristic parameters in Gamma distribution

Subscripts

est.	estimated value
L-N	Log-Normal distribution
N-N	Normal distribution
Gamma	Gamma distribution
max	maximum value
min	minimum value
R-R	Rosin-Rammer distribution
true	true value

problem, various methods were introduced, e.g. the generalized eikonal approximation (GEA) method [16], least squares QR decomposition (LSQR) method [14], Tikhonov regularization method [17] and the generalized cross-validation (GCV) method [18]. Among these methods, the Tikhonov regularization method is regarded as one of the most widely used strategies for solving the Fredholm integral equation of the first kind, and its performance in retrieving the PSDs is satisfactory [17].

In present study, based on the spectral extinction method, the Tikhonov regularization method is used to estimate the PSDs non-parametrically. Besides, to improve the accuracy of retrieval results, a selection principle of optimal measurement wavelengths on the basis of principle component analysis (PCA) approach is proposed. The remainder of this research is organized as follows. First, the principle of spectral extinction method for non-parametric estimation of PSDs and the selection principle of optimal measurement wavelengths are introduced. Then, four common PSDs, i.e. the Log-Normal (L-N), Rosin-Rammer (R-R), Normal (N-N), and Gamma distributions are retrieved under different measurement errors added. Meanwhile, the PSD of aerosol measured over Harbin, China and the PSD of microalgae cell available in Ref. [19] are reconstructed, respectively. The main conclusions and prospects for further research are provided finally.

2. Methodology

2.1. The spectral extinction method

The fundamental principle of spectral extinction method is based on the Mie scattering theory and the Lambert-Beer Law. When a collimated light beam of intensity I_0 impinges on a suspension particle system with a refractive index which is different from that of the dispersant medium, the transmitted light is scattered and absorbed by the particles, which results in the attenuation of the light. If the optical thickness is thin and the independent scattering dominates, the transmitted light intensity I at different incident wavelengths λ_i can be calculated, and the spectral transmittance signals can be derived as follows [20,21]:

$$\begin{cases} \ln[\tau(\lambda_1)] = -\dfrac{3}{2}\times L\times N_D\times\displaystyle\int_{D_{min}}^{D_{max}}\dfrac{Q_{ext}(\lambda_1,m_1,D)}{D}f(D)dD \\ \cdots\ \cdots \\ \ln[\tau(\lambda_i)] = -\dfrac{3}{2}\times L\times N_D\times\displaystyle\int_{D_{min}}^{D_{max}}\dfrac{Q_{ext}(\lambda_i,m_i,D)}{D}f(D)dD \\ \cdots\ \cdots \\ \ln[\tau(\lambda_S)] = -\dfrac{3}{2}\times L\times N_D\times\displaystyle\int_{D_{min}}^{D_{max}}\dfrac{Q_{ext}(\lambda_S,m_S,D)}{D}f(D)dD \end{cases} \tag{1}$$

where $\tau(\lambda_i) = I(\lambda_i)/I_0(\lambda_i)$ is the spectral transmittance at wavelength λ_i, which can be measured by the spectrograph or actinometer; D_{max} and D_{min} denote the upper and lower integration limits, respectively; $f(D)$ is the unknown volume frequency distribution needs to be determined; L is the geometric thickness of the particle system; N_D is the total number concentration of the suspended particle system; S denotes the number of measurement wavelengths; $m = n + ik$ is the spectral complex refractive index of particles, where n is the real part and k is the imaginary part; $Q_{ext}(\lambda_i,m_i,D)$ is the extinction efficiency factor of a single particle, which can be derived in terms of the Mie scattering theory [22,23].

To solve Eq. (1) numerically, the formula should be discretized as follows:

$$\ln[\tau(\lambda_i)] = -\frac{3}{2}\times L\times N_D \times\sum_{j=1}^{N}\frac{Q_{ext}(\lambda_i,m_i,D_j)}{D_j}f(D_j)dD_j \quad (i = 1,2,\ldots,S) \tag{2}$$

where N denotes the number of subintervals which the particle size range $[D_{min}, D_{max}]$ is divided into. Therefore, Eq. (1) can be described as a discrete version by matrix equation as follows:

$$\mathbf{B} = \mathbf{A}\cdot\mathbf{f} \tag{3}$$

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & & \cdots & & a_{1,N} \\ a_{2,1} & \cdots & a_{i-1,j} & \cdots & \\ \vdots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \vdots \\ & \vdots & a_{i+1,j} & \vdots & \\ a_{S,1} & & \cdots & & a_{S,N} \end{bmatrix} \tag{4}$$

$$\mathbf{B} = [b_1, \ldots, b_k, \ldots, b_S]^T \tag{5}$$

$$\mathbf{f} = [f(D_1), \ldots, f(D_j), \ldots, f(D_N)]^T \tag{6}$$

where $\mathbf{A}$ denotes the coefficient matrix, $a_{i,j} = -3/2\times L\times N_D Q_{ext}(\lambda_i,m_i,D_j)/D_j$; $\mathbf{B}$ denotes the spectral transmittance signals matrix, $b_i = \ln[\tau(\lambda_i)]$.

In present work, four common PSDs, i.e. the R-R, N-N, L-N and Gamma distributions are estimated. The mathematical representations of their volume frequency distributions are as follows [21,24]:

$$f_{R-R}(D) = \frac{\sigma}{\overline{D}}\times\left(\frac{D}{\overline{D}}\right)^{\sigma-1}\times\exp\left[-\left(\frac{D}{\overline{D}}\right)^{\sigma}\right] \tag{7}$$

$$f_{N-N}(D) = \frac{1}{\sqrt{2\pi}\sigma}\times\exp\left[-\frac{(D-\overline{D})^2}{2\sigma^2}\right] \tag{8}$$