Hindawi
Complexity
Volume 2022, Article ID 1862844, 9 pages
https://doi.org/10.1155/2022/1862844

*Research Article*

# Optimal Calculation Method of Mean Equivalent Diameter of Floc Particles Based on MCC

**Jun Liu**,[1] **Siqi Peng**,[1] **and Nan Zhou**[2]

[1]*School of Automation, Chengdu University of Information Technology, Chengdu 610225, China*
[2]*School of Electronic Information and Electrical Engineering, Chengdu University, Chengdu 610106, China*

Correspondence should be addressed to Jun Liu; liujun@cuit.edu.cn and Nan Zhou; nzhouuestc@126.com

Received 30 May 2022; Accepted 22 August 2022; Published 8 October 2022

Academic Editor: Jing Na



In the process of water treatment, coagulation is an important process to remove minerals and organic particles from raw water, which has typical time delay and nonlinearity. The effect of coagulation directly affects the turbidity of the effluent. According to the good correlation between the mean equivalent diameter of floc particles and the turbidity of effluent water, the image processing method is used to preprocess the floc image, the parameters of floc particles are quantitatively analyzed, and the mean equivalent diameter of floc particles is obtained, which is used as one of the bases for the control of coagulant dosage. However, no matter whether the coagulation process in actual situations interferes with the outside world or not, the equivalent diameters of floc particles after coagulation may have abnormal and invalid values, which may lead to a problem that there is a large difference between the mean equivalent diameter of floc particles calculated by a least square method and the mean valid equivalent diameter of floc particles. In response to this problem, this article proposes an optimal calculation method of the mean equivalent diameter of floc particles based on the maximum correntropy criterion (MCC) to reduce the negative impact of the abnormal and invalid equivalent diameters of floc particles on the mean equivalent diameter and provide an important reference data for the precise dosing control of coagulants. Finally, the feasibility of the theoretical results is verified by several numerical experiments.

## 1. Introduction

The traditional water treatment process mainly includes steps such as coagulation, sedimentation, filtration, and disinfection. Coagulation is an important process to remove minerals and organic particles in raw water, and its effect directly affects the turbidity of the effluent. When the coagulant is added to the raw water for a period of intense mixing, the coagulant is evenly and rapidly distributed in the water so that these colloidal particles in the water lose their stability. The interaction force between the destabilized colloidal particles will change from repulsion to attraction, so they collide continuously and slowly aggregate and finally form floc particles with certain strength, size, and density. This process is called coagulation, which has typical time-delay and nonlinear characteristics [1–3]. With the development of computer technology, the method of using image processing technology to detect the shape of floc particles and extract the characteristic parameters of floc particles has gradually become a current research focus [4] to control the coagulant dosage. The camera can capture a large number of floc images in the coagulation tank. After image preprocessing and feature extraction of the floc images, the equivalent diameters of the floc particles can be calculated by formula. The weighted average calculation of the equivalent diameters of floc particles can obtain the mean equivalent diameter of floc particles, which can be used as one of the important parameters to characterize the overall situation of coagulation. When the coagulant dosage is within an appropriate range, the floc particles slowly aggregate and become denser, which leads to fast sedimentation of floc particles, clear water, and low turbidity. At this time, the mean equivalent diameter of the floc particles is also within a reasonable range. The mean equivalent diameter of floc particles is adjusted by controlling the dosage of the coagulant so that the mean equivalent diameter of floc particles keeps approaching the set value of the mean equivalent diameter of floc particles and realizes the accurate control of the coagulant dosage and effluent turbidity [5]. Since there is a good correlation between the mean equivalent diameter of floc particles and the turbidity of effluent after flocculation and sedimentation, taking the mean equivalent diameter of floc particles as one of the control parameters for automatically controlling the dosage of coagulant can achieve better coagulation and sedimentation effect and obtain the expected effluent turbidity [6, 7].

Due to the important role of equivalent diameters of floc particles in modern water treatment automation control technology, scholars at home and abroad have carried out a lot of research on the extraction and application of the equivalent diameters of floc particles. To address the problem of hysteresis of obtaining water quality indicators in the water treatment process, Dai et al. [8] used micro-eddy current flocculation technology to detect floc particles. They combined the relationship between the equivalent diameter and the fractal dimension to determine the coagulant dosage and alleviate the time-delay problems in water treatment by analyzing and determining the coagulant dosage on the coagulation effect and the equivalent diameters of floc particles. Asensi et al. [9] developed a fully automatic activated sludge floc identification and morphological characterization toolbox based on digital image analysis and statistical processing. The toolbox could determine characteristic parameters such as the equivalent diameters of floc particles, which was mainly used to help study the characteristics of activated sludge flocs in urban sewage treatment plants. Khan et al. [10] used a microscope to collect floc images, used a state-of-the-art image segmentation algorithm to segment floc images with different equivalent diameters in different fluctuation ranges, and extracted morphological characteristic parameters for sludge volume index (SVI) and mixed liquor suspended solids (MLSS) modeling to explore the feasibility of applying the algorithm model to various plants in different regions. Chen et al. [11] used image acquisition system and data processing system to collect and process floc images in real time, took the mean equivalent diameter of floc particles calculated by the least square method as the target for controlling the coagulant dosage, and then automatically corrected the set system control parameters through flow and turbidity feedback so as to save the coagulant consumption and ensure the quality of the effluent from the sedimentation tank. Gao [12] used image processing technology to detect floc images, obtained various floc parameters that fluctuated within the correct range, and analyzed the effect of different flocculation time and coagulant dosage on the equivalent diameters of floc particles. It can provide a data reference for solving the problems of time delay and accuracy of dosing amount in the process of coagulation and dosing in water plants. Chen et al. [13] applied digital image processing technology to water treatment process control, improved the original water treatment experimental equipment, and used image processing algorithms to complete the extraction and analysis of the feature parameters of floc images in the coagulation process. The noncontact detection of floc properties in the flocculant addition control system provided an effective solution for water plants to improve the automation level of water treatment. Wang [14] used machine vision technology to explore the morphological change law of floc aggregation in the coagulation process and corresponded the characteristic parameters of floc particles such as porosity and equivalent diameters of floc particles with the factors that actually affected the flocculation process. Simulations were carried out to study the factors affecting the flocculation process and results, and the mean equivalent diameter of floc particles was determined as one of the key factors to regulate and control the flocculation process.

Although the above studies were only scattered and preliminary discussions on the application of the equivalent diameters of floc particles in the automatic process of water treatment, they have already demonstrated the importance of the equivalent diameter of floc particles in the automatic control technology of coagulant dosage. Researchers at home and abroad not only extracted the internal relationship between geometric parameters such as the equivalent diameters of floc particles and the coagulant dosage but also explored the internal relationship between other parameters of floc particles and the coagulant dosage, which alleviated the lag of water quality indicators in water treatment to a certain extent. However, the problem of accurate calculation of relevant parameters such as the mean equivalent diameter of floc particles still needs to be solved urgently in the application of precise control of coagulant dose. In the actual water treatment process of the water plant, the effect of coagulation is greatly affected by the actual environment. The influence of microorganisms, abnormal coagulant dosage, equipment leakage, and air temperature will lead to abnormal settlement in the floc particles after coagulation [15, 16]. For example, when the weather is hot, the microbes in the sludge will decompose and produce gas, which will cause trace bubbles in the sludge. The gas leakage in the equipment will cause the gas content in the water to be too high, and the low temperature will affect the reaction speed of the coagulant. These objective factors will affect the effect of the aggregation and sedimentation of floc particles so that the captured floc images can not reflect the actual situation of the sedimentation tank, and the equivalent diameters of floc particles calculated based on these images will have a small number of abnormal and invalid values. In addition, no matter how the coagulation effect is, there may also be a small number of abnormal and invalid values in the calculated equivalent diameters of floc particles without interference from the objective environment. The abnormal and invalid values will lead to a large deviation between the mean equivalent diameter obtained by the least square method and the mean valid equivalent diameter of floc particles so that the current coagulation effect cannot be accurately evaluated. Therefore, a new data processing optimization method of the mean equivalent diameter of floc particles is examined, which is of great significance to eliminate the negative effect of the abnormal and invalid equivalent diameters of a small number of floc particles on the mean equivalent diameter of floc particles.

In view of the abnormal and invalid values of the equivalent diameters of floc particles during the coagulation process, maximum correntropy criterion (MCC) [17–19] will be introduced in this article, and an optimal method will be proposed to calculate the mean equivalent diameter of floc particles based on MCC. This method optimizes the process of solving the mean equivalent diameter of floc particles, reduces the calculation error caused by abnormal and invalid equivalent diameters, and achieves the purpose of accurately calculating the mean equivalent diameter of floc particles, which provides an important reference data for the precise dosing control of subsequent coagulants.

## 2. Preliminaries

The equivalent diameter means that when a particle has the same or similar physical properties to a spherical particle, we can replace the diameter of the particle with the diameter of the spherical particle. In water treatment, the sedimentation characteristics of floc particles are complicated. The floc particles are in a discrete state during the sedimentation process; their mass, size, and characteristics do not change; and the sedimentation velocity of the floc particles is not disturbed. The mathematical expression to characterize the particle settling motion often adopts the Stokes formula [11, 20], whose specific form is as follows:

$$v = \frac{(\rho - \rho_0) g}{18\mu} d_s^2, \tag{1}$$

where $v$ is the settling velocity of the floc particle, $\rho$ is the floc density, $\rho_0$ refers to the density of water, $g$ is the acceleration of gravity, $\mu$ is the viscosity coefficient of water, and $d_s$ is the diameter of the floc particle.

Further studies show that as the diameter of the floc particle changes, the density of the floc particle changes according to the following formula:

$$\rho - \rho_0 = d_s^{-k_p}, \tag{2}$$

where $k_p$ is a coefficient, whose value is generally $1.2 \sim 1.5$, depending on the coagulant filling rate and the quality of raw water. Combining the above formulas (1) and (2), it can be concluded that the relationship between the diameter and the sedimentation velocity of the floc particle can be obtained as

$$v = \frac{g d_s^{\left(2-k_p\right)}}{18\mu}. \tag{3}$$

The above analysis is based on the assumption that the floc particle is spherical, but we know that the actual floc particle is in an irregular state, and its sedimentation speed should indeed be slower than that of the spherical floc particle of the same volume. The size and shape of floc particles can be well reflected by the floc images collected by industrial cameras. Each floc region in the image reflects the state of floc particle movement during the coagulation process. The image of a floc particle in the 2D plane can be characterized by four parameters [11, 21]: the size-related area of the floc particle, the shape-related perimeter of the floc particle, the vacant area in the middle of the floc particle related to the degree of looseness, and the length to width ratio of the floc particle. These features represent the characteristics of the floc particle. The above four parameters can be converted into $\phi_i$ using the following formula:

$$\begin{aligned}\phi_i = {} & 2\sqrt{\frac{s_i}{\pi}}\left[1 - \left(1 - \frac{2\sqrt{s_i \pi}}{l_i}\right)k_1\right] \\ & \times \left[1 - \left(1 - \frac{1}{m_i}\right)k_2\right] \times \left(1 - \frac{s_{i0}}{s_i}k_3\right),\end{aligned} \tag{4}$$

where $\phi_i$ is the equivalent diameter of the $i$th floc particle; $s_i$ is the area of the $i$th floc particle; $l_i$ is the perimeter of the $i$th floc particle; $s_{i0}$ is the hollow area of the $i$th floc particle; $m_i$ is the length to width ratio of the $i$th floc particle; and $k_1$, $k_2$, and $k_3$ are the coefficients of the perimeter $l_i$, the length to width ratio $m_i$, and the hollow area $s_{i0}$, respectively. $k_1$, $k_2$, and $k_3$ are all decimals ranging from 0 to 1, which can be selected according to actual conditions.

By the above analysis and calculation, the equivalent diameter of a floc particle can be extracted from the image of the floc particle. The equivalent diameter of the floc particle is an important characteristic parameter of the floc particle and has a good correlation with the turbidity of water. It not only reflects the quality of the coagulation effect but also relates to whether the subsequent effluent turbidity meets the water supply requirements. Taking it as one of the target values to control the coagulant dosage can achieve a good control effect. It can be seen from the above formula (3) that the larger the equivalent diameter of the floc particle, the faster the settling velocity of the floc particle, that is, the better the integrity of floc particles, the more sufficient the sedimentation and the smaller the turbidity of the sedimented water. The change of the equivalent diameters of floc particles can not only reflect the quality of the coagulation effect but also relate to whether the utilization of the coagulant can achieve the maximum benefit.

However, the parameter $\phi_i$ still cannot fully characterize the overall effect of coagulation. In practical application, the equivalent diameter of each floc particle is calculated according to formula (4), and then the mean equivalent diameter $\Phi$, which is regarded as a key parameter for the control of coagulant dosage, is calculated in real time according to the equivalent diameters of floc particles obtained within a certain time. The mean equivalent diameter $\Phi$ is expressed as follows:

$$\Phi = \frac{\sum_{i=1}^{N} (n_i \phi_i)}{\sum_{i=1}^{N} (n_i)}, \tag{5}$$

where $N$ represents the number of floc particles with different equivalent diameters, and $n_i$ represents the number of floc particles whose equivalent diameter is $\phi_i$.

According to the above analysis, we can display the floc image collected in real time on the computer and calculate $s_i$, $l_i$, $m_i$, and $s_{i0}$ of the $i$th floc particle in formula (4) by image preprocessing, image segmentation, and other image processing technologies, which lead to obtain $\phi_i$. Finally, we can obtain the mean equivalent diameter $\Phi$ of floc particles by substituting $\phi_i$ and $n_i$ into formula (5).

## 3. Main Results

In the actual water treatment process of water plant, the influence of microorganisms, too much or too little coagulant dose, equipment leakage, and air temperature will lead to abnormal aggregation and settlement of floc particles. This will make the captured floc images cannot reflect the real state of the sedimentation tank, and the calculated floc equivalent diameters will have some abnormal and invalid values. In addition, no matter whether the coagulation effect is disturbed by the environment or not, the equivalent diameters of floc particles calculated from floc images will also have some abnormal and invalid values, so that the mean equivalent diameter of floc particles calculated by the traditional calculation method (the least square method) cannot accurately characterize the coagulation effect. To weaken the influence of the invalid equivalent diameters of floc particles on the mean equivalent diameter of floc particles, MCC will be introduced to optimize the calculation process of the mean equivalent diameter of floc particles so as to provide important reference data for accurately describing the actual coagulation effect. Because of its good robustness [22], MCC is widely used in many fields, such as computer vision [23], feature extraction [24, 25], and signal processing [26–28]. It is mainly used to deal with non-Gaussian noise and outliers [27, 29, 30]. MCC is based on entropy [31], which in turn derives from information theory. Correlation entropy is used to measure the similarity between two variables, which is expressed as follows:

$$V_\sigma(A,B) = E\left[k_\sigma(A-B)\right], \tag{6}$$

where $E(*)$ is the expectation of $*$, $k_\sigma(\cdot)$ represents the Gaussian kernel function, and $\sigma$ represents the kernel width of $k_\sigma(\cdot)$. Usually, the joint probability distribution between variables $A$ and $B$ is unknown, and only a finite amount of data $A = (a_1, a_2, \ldots, a_N)$ and $B = (b_1, b_2, \ldots, b_N)$ can be available, which make the estimator of correlation entropy (6) can be represented as

$$\widehat{V}_\sigma(A,B) = \frac{1}{N}\sum_{i=1}^{N} k_\sigma(a_i - b_i), \tag{7}$$

where $k_\sigma(a_i - b_i) = e^{-(a_i-b_i)^2/2\sigma^2}$.

As we all know, the mean square error (MSE) is a measure that reflects the degree of difference between the estimator and the estimated value, and it is a global measure. Compared to the MSE, MCC is a local metric whose value mainly depends on the probability along the $A = B$ direction, and the local extent depends on the kernel size $\sigma$. The Gaussian kernel functions for various kernel widths are shown in Figure 1. As can be seen from Figure 1, the convergence rate of the function varies with the size of the kernel width. For large errors or outliers, the kernel function has better robustness. When the error between $A = B$ is large, the kernel function of $A = B$ gets a small value or even a value close to zero so that the calculation process can avoid the negative impact caused by outliers and has a good stable performance in the case of data anomalies caused by interference. Figures 2 and 3 illustrate the difference between the mean square error and the correlation entropy. In this article, an optimization model is constructed by MCC to eliminate the adverse effects of abnormal and invalid values (abnormal and invalid equivalent diameters of floc particles) on the mean equivalent diameter of floc particles.

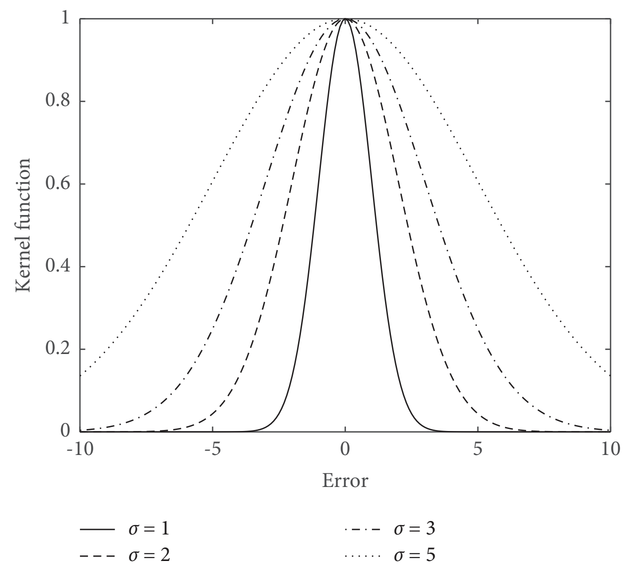


Figure 1: Gaussian kernel function with different kernel widths.

The mean equivalent diameter of floc particles calculated by the least square method can be expressed as follows:

$$\Phi_c = \arg\min_{\Phi}\sum_{i=1}^{N} n_i(\phi_i - \Phi)^2, \tag{8}$$

which is an optimization problem whose solution of this optimization problem is

$$\Phi_c = \frac{\sum_{i=1}^{N}(n_i\phi_i)}{\sum_{i=1}^{N}(n_i)}. \tag{9}$$

According to formula (9), the mean equivalent diameter of floc particles calculated by the least square method is the mean of the sum of all equivalent diameters of floc particles, which cannot eliminate the abnormal and invalid equivalent diameter of floc particles. Considering the abnormal and invalid equivalent diameters of floc particles, we introduce maximum correntropy criterion [32] to weaken the influence of abnormal and invalid equivalent diameters of floc particles on the mean equivalent diameter of floc particles. Using MCC, the optimization problem is formulated as follows:

$$\Phi_{MCC} = \underset{\Phi}{\operatorname{argmax}}\sum_{i=1}^{N} n_i \exp\left(-\frac{(\phi_i - \Phi)^2}{2\sigma^2}\right). \tag{10}$$

Assuming $f(z) = z - z\ln(-z)$, it can be obtained as follows:

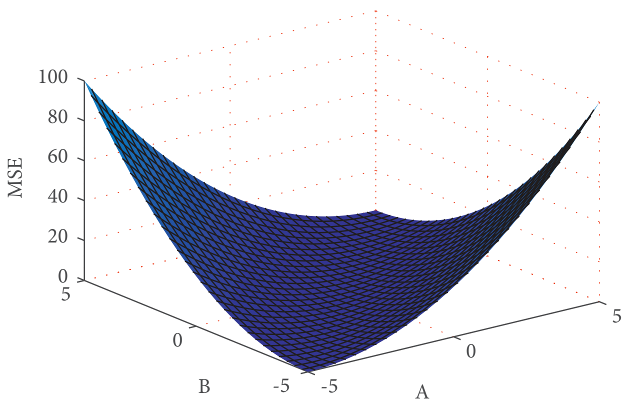


Figure 2: MSE in the joint space of variables $A$ and $B$.

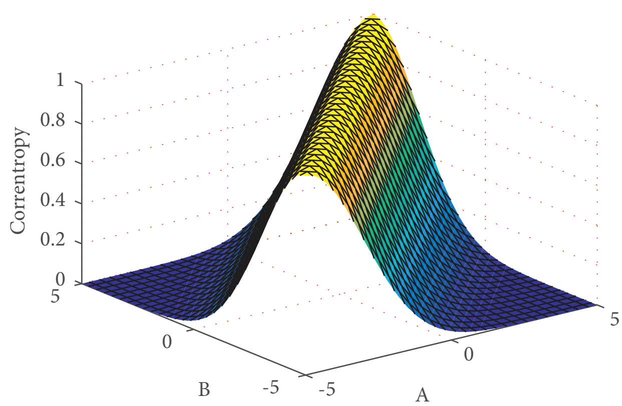


Figure 3: Correntropy in the joint space of variables $A$ and $B$.

$$\exp(-x) = \sup_z (zx - f(z)). \quad (11)$$

By $\partial/\partial_z (zx - f(z)) = x + \ln(-z) = 0$, we can find that the maximum value of $(zx - f(z))$ can be obtained when $z = -\exp(-x)$.

The above derivation process is called the half-quadratic (HQ) strategy [33, 34].

If let

$$z = \rho_i, \quad (12)$$

and

$$x = \frac{(\phi_i - \Phi)^2}{2\sigma^2}, \quad (13)$$

then it is easy to get

$$\Phi_{MCC} = \arg\max_{\Phi} \sum_{i=1}^{N} n_i\left(\rho_i \frac{(\phi_i - \Phi)^2}{2\sigma^2} - f(\rho_i)\right), \quad (14)$$

and

$$\rho_i = -\exp\left(-\frac{(\phi_i - \Phi)^2}{2\sigma^2}\right), \quad (15)$$

where $\rho_i$ is an auxiliary variable.

If we take the partial derivative of $\Phi$ of (14), then it is easy to obtain

$$\begin{aligned}
&\frac{\partial\left[\sum_{i=1}^{N} n_i\left(\rho_i\left((\phi_i - \Phi)^2/2\sigma^2\right) - f(\rho_i)\right)\right]}{\partial\Phi} \\
&= \frac{\partial\left[\sum_{i=1}^{N}\left(\rho_i n_i/2\sigma^2\right)(\phi_i - \Phi)^2 - \sum_{i=1}^{N} n_i f(\rho_i)\right]}{\partial\Phi} \\
&= \sum_{i=1}^{N}\left(\frac{\rho_i n_i}{2\sigma^2}\right)\frac{\partial\left[(\phi_i - \Phi)^2\right]}{\partial\Phi} \\
&= \sum_{i=1}^{N}\left(\frac{\rho_i n_i}{2\sigma^2}\right)\frac{\partial\left[\left(\phi_i^2 - 2\phi_i\Phi + \Phi^2\right)\right]}{\partial\Phi} \\
&= \sum_{i=1}^{N}\left(\frac{\rho_i n_i}{2\sigma^2}\right)(-2\phi_i + 2\Phi) \\
&= -\sum_{i=1}^{N}\frac{\rho_i n_i \phi_i}{\sigma^2} + \Phi\sum_{i=1}^{N}\frac{\rho_i n_i}{\sigma^2} = 0,
\end{aligned} \quad (16)$$

and it is easy to get

$$\begin{aligned}
\Phi_{MCC} &= \frac{\sum_{i=1}^{N}\rho_i n_i \phi_i/\sigma^2}{\sum_{i=1}^{N}\rho_i n_i/\sigma^2} \\
&= \frac{\sum_{i=1}^{N}\rho_i n_i \phi_i}{\sum_{i=1}^{N}\rho_i n_i}.
\end{aligned} \quad (17)$$

We can solve the problem (14) by alternative optimization.

Firstly, when $\Phi$ is fixed, the solution of $\rho_i$ is derived as follows:

$$\rho_i^{(k+1)} = -\exp\left(-\frac{\left(\phi_i - \Phi^{(k)}\right)^2}{2\left(\sigma^{(k+1)}\right)^2}\right), \quad (18)$$

where $k \geq 0$ is the number of iterations.

Secondly, when $\rho_i$ is fixed, the solution of $\Phi$ can be easily obtained as follows:

$$\Phi^{(k+1)} = \frac{1}{\sum_{i=1}^{N}\rho_i^{(k+1)} n_i}\sum_{i=1}^{N}\rho_i^{(k+1)} n_i \phi_i. \quad (19)$$

Furthermore, it is known from [23] that after each iteration $\sigma^2$ should be updated as follows:

$$\left(\sigma^{(k+1)}\right)^2 = \frac{1}{2\sum_{i=1}^{N} n_i}\sum_{i=1}^{N} n_i\left(\phi_i - \Phi^{(k)}\right)^2. \quad (20)$$

This update rule consists of the above three steps, which are repeated until the convergence condition is achieved. The procedure is summarized in Algorithm 1.

The proof is completed.

**Input:** The equivalent diameter $\phi_i$ $(1 \le i \le N)$ of the $i$th floc particle, and the number $n_i$ $(1 \le i \le N)$ of the $i$th floc particle whose equivalent diameter is $\phi_i$.
**Output:** The MCC-based samples' mean $\Phi_{MCC}$.
**Initialize:** $\Phi^{(0)} = \Phi_c$, $k = 0$.
**while** Not convergent **do**
  Update $(\sigma^{(k+1)})^2 \leftarrow$ (20).
  Update $\rho_i^{(k+1)} \leftarrow$ (18).
  Update $\Phi^{(k+1)} \leftarrow$ (19).
  $k \leftarrow k + 1$.
**end while**
$\Phi_{MCC} = \Phi^{(k)}$ as the MCC-based samples' mean.

ALGORITHM 1: The MCC-based samples' mean $\Phi_{MCC}$ = MCC-Mean $(\phi_i, n_i)$.

## 4. Simulation Results

There are four sets of simulation data generated for us to verify the effectiveness of Algorithm 1, and $n_i$ $(1 < = i < = N)$ is always equal to 1 for the convenience of simulation.

(1) The first set of simulation data is generated for us to verify the effectiveness of Algorithm 1. There are ninety points representing ninety valid equivalent diameters of floc particles randomly generated in 1D space by Gaussian distribution with mean $\Phi_v = 1.3$ mm and covariance $\sigma_v = 0.2$ and ten points representing ten abnormal and invalid equivalent diameters of floc particles by Gaussian distribution with mean $\Phi_a = 6$ mm and covariance $\sigma_a = 0.3$. Using these generated data, we can compute the conventional samples' mean $\Phi_c$ by (9), the MCC-based samples' mean $\Phi_{MCC}$ by Algorithm 1, and the valid samples' mean $\Phi_v$ by (9) with only considering the valid equivalent diameters. The positions of the three means are shown in Figure 4. It is obvious that the MCC-based samples' mean $\Phi_{MCC}$ and the valid samples' mean $\Phi_v$ are almost overlapped, whereas the conventional samples' mean $\Phi_c$ is seriously biased from the valid samples' mean $\Phi_v$ due to the existence of ten abnormal and valid equivalent diameters.

(2) The second set of simulation data is generated for us to verify the effectiveness of Algorithm 1. There are ten points representing ten abnormal and valid equivalent diameters of floc particles randomly generated in 1D space by Gaussian distribution with mean $\Phi_a = 1.3$ mm and covariance $\sigma_a = 0.2$ and ninety points representing ninety valid equivalent diameters of floc particles by Gaussian distribution with mean $\Phi_v = 6$ mm and covariance $\sigma_v = 0.3$. Using these generated data, we can compute the conventional samples' mean $\Phi_c$ by (9), the MCC-based samples' mean $\Phi_{MCC}$ by Algorithm 1, and the valid samples' mean $\Phi_v$ by (9) with only considering the valid equivalent diameters. The positions of the three means are shown in Figure 5. It is obvious that the MCC-based samples' mean $\Phi_{MCC}$ and the valid samples' mean $\Phi_v$ are almost overlapped, whereas the conventional samples' mean $\Phi_c$ is seriously biased from the valid samples' mean $\Phi_v$ due to the existence of ten abnormal and valid equivalent diameters.

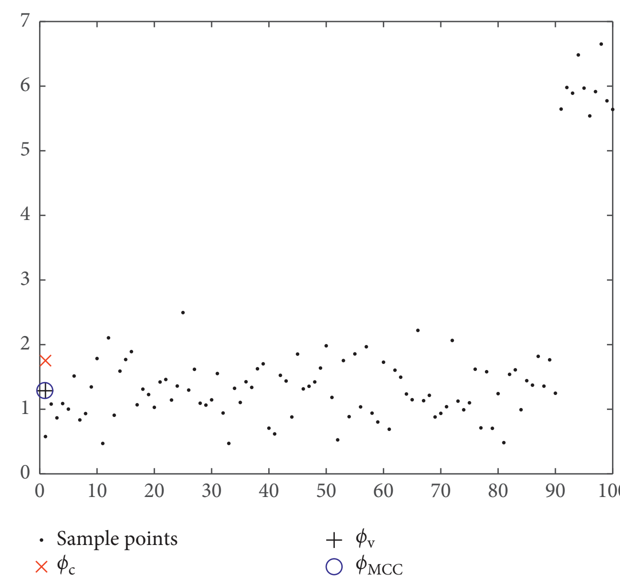


FIGURE 4: Toy problem with ninety inliers ranging from 0.3 mm to 2.6 mm and ten outliers ranging from 5.6 mm to 6.8 mm. The conventional samples' mean $\Phi_c$, the valid samples' mean $\Phi_v$, and MCC-based samples' mean $\Phi_{MCC}$.

(3) The third set of simulation data is generated for us to verify the effectiveness of Algorithm 1. Firstly, there are ninety points representing ninety valid equivalent diameters of floc particles randomly generated in 1D space by Gaussian distribution with mean $\Phi_v = 1.3$ mm and covariance $\sigma_v = 0.2$ and ten points representing ten abnormal and valid equivalent diameters of floc particles by Gaussian distribution with mean $\Phi_a = 6$ mm and covariance $\sigma_a = 0.3$. Secondly, the 100 points are randomly sorted. Finally, using these generated data, we can compute the conventional samples' mean $\Phi_c$ by (9), the MCC-based samples' mean $\Phi_{MCC}$ by Algorithm 1, and the valid samples' mean $\Phi_v$ by (9) with only considering

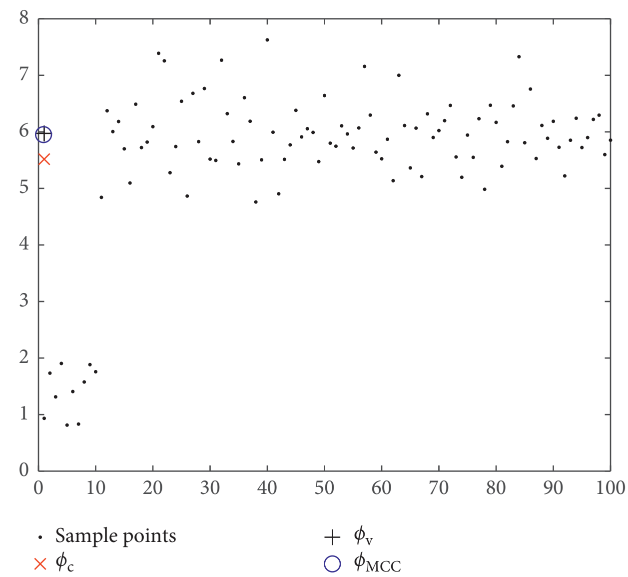


Figure 5: Toy problem with ninety inliers ranging from 4.6 mm to 7.7 mm and ten outliers ranging from 0.8 mm to 2 mm. The conventional samples' mean $\Phi_c$, the valid samples' mean $\Phi_v$, and MCC-based samples' mean $\Phi_{MCC}$.

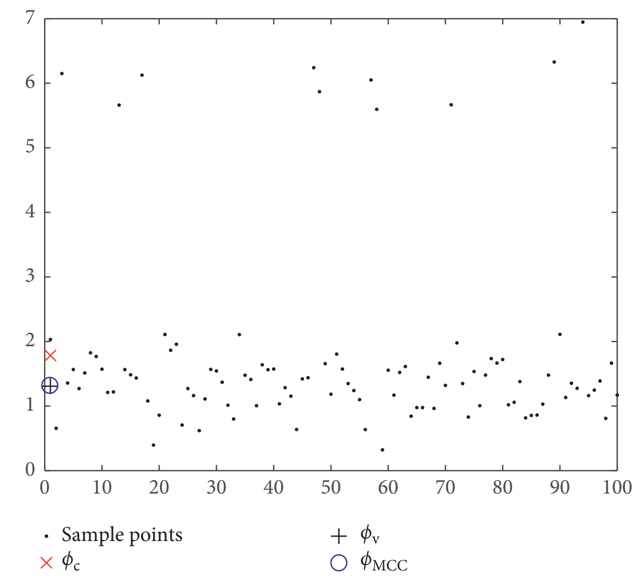


Figure 6: Toy problem with ninety inliers ranging from 0.3 mm to 2.3 mm and ten outliers ranging from 5.5 mm to 7 mm. The conventional samples' mean $\Phi_c$, the valid samples' mean $\Phi_v$, and MCC-based samples' mean $\Phi_{MCC}$.

the valid equivalent diameters. The positions of the three means are shown in Figure 6. It is obvious that the MCC-based samples' mean $\Phi_{MCC}$ and the valid samples' mean $\Phi_v$ are almost overlapped, whereas the conventional samples' mean $\Phi_c$ is seriously biased from the valid samples' mean $\Phi_v$ due to the existence of ten abnormal and valid equivalent diameters.

(4) The last set of simulation data is generated for us to verify the effectiveness of Algorithm 1. Firstly, there are ten points representing ten abnormal and valid equivalent diameters of floc particles randomly generated in 1D space by Gaussian distribution with mean $\Phi_a = 1.3$ mm and covariance $\sigma_a = 0.2$ and ninety points representing ninety valid equivalent diameters of floc particles by Gaussian distribution with mean $\Phi_v = 6$ mm and covariance $\sigma_v = 0.3$. Secondly, the 100 points are randomly sorted. Finally, using these generated data, we can compute the conventional samples' mean $\Phi_c$ by (9), the MCC-based samples' mean $\Phi_{MCC}$ by Algorithm 1, and the valid samples' mean $\Phi_v$ by (9) with only considering the valid equivalent diameters. The positions of the three means are shown in Figure 7. It is obvious that the MCC-based samples' mean $\Phi_{MCC}$ and the valid samples' mean $\Phi_v$ are almost overlapped, whereas the conventional samples' mean $\Phi_c$ is seriously biased from the valid samples' mean $\Phi_v$ due to the existence of ten abnormal and valid equivalent diameters.

If the mean equivalent diameter of floc particles corresponding to the turbidity of the effluent within the normal range is about 1 mm-2 mm, then it can be seen from Figures 4 and 6 that the corresponding coagulation effect is good, which indicates that the turbidity of the effluent is within the normal range, and then it can be seen from Figures 5 and 7 that the corresponding coagulation effect is not good, which indicates that the turbidity of the effluent is outside the normal range.

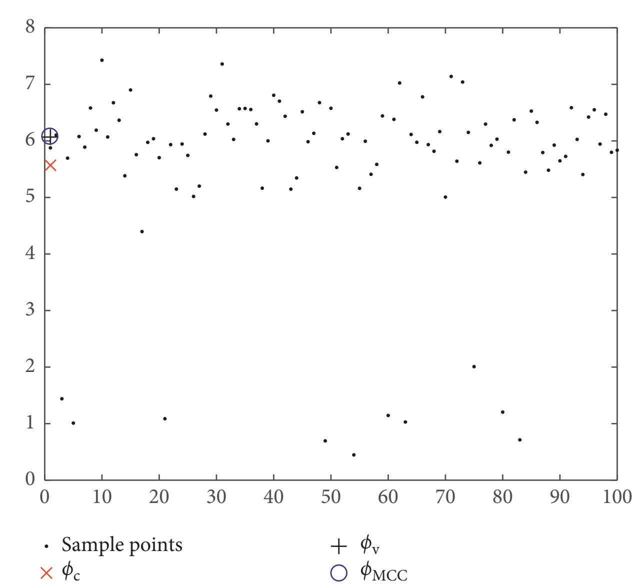


Figure 7: Toy problem with ninety inliers ranging from 4.3 mm to 7.5 mm and ten outliers ranging from 0.3 mm to 2.2 mm. The conventional samples' mean $\Phi_c$, the valid samples' mean $\Phi_v$, and MCC-based samples' mean $\Phi_{MCC}$.

Several sets of mean equivalent diameters of floc particles calculated by the conventional method and the MCC-based optimization method are presented in Table 1, and the corresponding error rates are given. As can be seen from

Table 1: Calculation of the mean equivalent diameters of floc particles.

| The number of images | The number of floc particles | $\Phi_v$ (mm) | $\Phi_c$ (mm) | Error rate 1 (%) | $\Phi_{MCC}$ (mm) | Error rate 2 (%) |
|---|---|---|---|---|---|---|
| 1 | 30 | 0.8700 | 0.9317 | 7.09 | 0.8939 | 2.75 |
| 2 | 36 | 0.8400 | 0.8858 | 5.45 | 0.8492 | 1.10 |
| 3 | 29 | 1.0900 | 1.1540 | 5.87 | 1.1022 | 1.12 |
| 4 | 45 | 1.2100 | 1.2442 | 2.83 | 1.2205 | 0.87 |
| 5 | 52 | 0.7900 | 0.8289 | 4.92 | 0.7989 | 1.13 |

Table 1, the error rates of the mean equivalent diameters of the floc particles are reduced by means of the MCC algorithm.

## 5. Conclusion

According to the characteristics that the mean equivalent diameter of floc particles is closely related to the turbidity of the effluent, we use image processing technology to process and extract the characteristics of the collected floc images to obtain the relevant important parameters such as the mean equivalent diameter of floc particles and feed them back to the control system of coagulant dose. This can not only effectively improve the utilization rate of coagulants but also liberate human resources and reduce production costs. The mean equivalent diameter of floc particles is an important parameter used to describe the characteristics of floc precipitation after coagulation in the water treatment process. In the actual operation of the water plant, no matter whether the coagulation effect is good or bad, there will be abnormal and invalid values in the equivalent diameters of floc particles. To avoid the deviation of the mean equivalent diameter and the mean valid equivalent diameter caused by the abnormal and invalid equivalent diameters of floc particles, the MCC algorithm is introduced in this article. The MCC algorithm optimizes the solution of the mean equivalent diameter of floc particles to eliminate or reduce the influence of the abnormal and invalid values of equivalent diameters on the overall sedimentation in the actual situation and provides a reference method for water treatment fields such as tap water treatment and sewage treatment. Finally, the validity of the theoretical results is verified by numerical experiments.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported in part by the Natural Science Foundation of Sichuan of China under Grant 2022NSFSC0462, the Major Industrial Projects of Sichuan Science and Technology Department of China under Grant 21ZDYF2965, and the National Natural Science Foundation of China under Grant 61802036.

## References

[1] Y. S. Vadasarukkai and G. A. Gagnon, "Characterization, fate and transport of floc aggregates in full-scale flocculation tanks," *Environmental Sciences: Water Research & Technology*, vol. 2, no. 1, pp. 223–232, 2016.

[2] E. N. Banadda, I. Y. Smets, R. Jenne, and J. F. Van Impe, "Predicting the onset of filamentous bulking in biological wastewater treatment systems by exploiting image analysis information," *Bioprocess and Biosystems Engineering*, vol. 27, no. 5, pp. 339–348, 2005.

[3] R. K. Chakraborti and J. F. Atkinson, "Settling velocity analysis of natural suspended particles using fractal approach," *Journal of Environmental Engineering*, vol. 146, no. 12, Article ID 04020138, 2020.

[4] M. Lapointe and B. Barbeau, "Characterization of ballasted flocs in water treatment using microscopy," *Water Research*, vol. 90, pp. 119–127, 2016.

[5] Y. G. Perez, S. G. F. Leite, and M. A. Z. Coelho, "Activated sludge morphology characterization through an image analysis procedure," *Brazilian Journal of Chemical Engineering*, vol. 23, no. 3, pp. 319–330, 2006.

[6] K. Strom and A. Keyvani, "An explicit full-range settling velocity equation for mud flocs," *Journal of Sedimentary Research*, vol. 81, no. 12, pp. 921–934, 2011.

[7] A. Vahedi and B. Gorczyca, "Predicting the settling velocity of flocs formed in water treatment using multiple fractal dimensions," *Water Research*, vol. 46, no. 13, pp. 4188–4194, 2012.

[8] H. L. Dai, Z. M. Qiu, F. P. Hu, C. Gao, Y. F. Chen, and Z. Zhou, "Floc performance parameters during water treatment in a micro-vortex flocculation process determined by machine vision," *Environmental Technology*, vol. 40, no. 23, pp. 3062–3071, 2019.

[9] E. Asensi, D. Zambrano, E. Alemany, and D. Aguado, "Effect of the addition of precipitated ferric chloride on the morphology and settling characteristics of activated sludge flocs," *Separation and Purification Technology*, vol. 227, Article ID 115711, 2019.

[10] M. B. Khan, H. Nisar, and C. A. Ng, "Generalized modeling of the sludge volume index and suspended solids to monitor activated sludge bioflocculation using image analysis," *Environment Protection Engineering*, vol. 46, no. 3, pp. 17–37, 2020.

[11] F. Chen, H. Y. Ruan, and Q. Yang, "Technology of automatic control for coagulant dosing based on equivalent diameter of flocs," *Industrial Water and Wastewater*, vol. 36, no. 6, pp. 50–52, 2005.

[12] C. Gao, *The Study of Floc Measurement and Performance in the Flocculation Process of Water Treatment Based on Machine Vision*, East China Jiaotong University, China, 2016.

[13] K. G. Chen, X. J. Hang, and Y. Zhang, "Coagulation control method based on image processing," *Water Technology*, vol. 6, no. 1, pp. 40–42, 2012.
[14] Y. Wang, "Study on Treatment of Microplastics in Drinking Water by Coagulation-Sedimentation Process," Harbin Institute of Technology, China, 2020.
[15] Y. Rang, R. P. Liu, D. W. Zhang, and H. Xu, "Investigation of the coagulation process in the NingDong water treatment plant during abnormal water quality in spring," *Water Technology*, vol. 11, no. 5, pp. 11–16, 2017.
[16] Y. Q. Xu, H. J. Zhao, and Y. G. Kong, "Analysis and Solution of Abnormal Flocs in Precipitation Process," *Journal of Shijiazhuang Institute of Railway Technology*, vol. 11, no. 4, pp. 60–63, 2012.
[17] X. Q. Chen, Y. H. Hu, and Y. R. Huang, "Image thresholding segmentation based on two-dimensional mcc," *Journal of Infrared and Millimeter Waves*, vol. 24, no. 5, pp. 397–400, 2005.
[18] B. B. Wang and T. Hu, "Online gradient descent for kernel-based maximum correntropy criterion," *Entropy*, vol. 21, no. 7, p. 644, 2019.
[19] J. C. Yen, F. J. Chang, and S. Chang, "A new criterion for automatic multilevel thresholding," *IEEE Transactions on Image Processing*, vol. 4, no. 3, pp. 370–378, 1995.
[20] N. H. G. Rahmani, T. Dabros, and J. H. Masliyah, "Fractal structure of asphaltene aggregates," *Journal of Colloid and Interface Science*, vol. 285, no. 2, pp. 599–608, 2005.
[21] R. Y. Song, "Selection of automatic process mode for coagulant filling," *City and Town Water Supply*, no. 3, pp. 9–13, 1996.
[22] J. C. Principe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*, Springer, New York, NY, USA, 2010.
[23] R. He, W. S. Zheng, and B. G. Hu, "Maximum correntropy criterion for robust face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1561–1576, 2011.
[24] J. J. Y. Wang, X. L. Wang, and X. Gao, "Non-negative matrix factorization by maximizing correntropy for cancer clustering," *BMC Bioinformatics*, vol. 14, Article ID 107, 2013.
[25] L. Du, X. Li, and Y. D. Shen, "Robust nonnegative matrix factorization via half-quadratic minimization," in *Proceedings of the IEEE 12th International Conference on Data Mining*, Brussels, Belgium, December 2012.
[26] B. Chen, J. Wang, H. Zhao, N. Zheng, and J. C. Principe, "Convergence of a fixed-point algorithm under maximum correntropy criterion," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1723–1727, 2015.
[27] B. Chen, L. Xing, H. Zhao, N. Zheng, and J. C. Principe, "Generalized correntropy for robust adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3376–3387, 2016.
[28] B. D. Chen, L. Xing, J. L. Liang, N. N. Zheng, and J. C. Principe, "Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion," *IEEE Signal Processing Letters*, vol. 21, no. 7, pp. 880–884, 2014.
[29] N. Zhou, B. D. Chen, Y. H. Du, T. Jiang, J. Liu, and Y. Y. Xu, "Maximum correntropy criterion-based robust semi-supervised concept factorization for image representation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 10, pp. 3877–3891, 2020.
[30] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: properties and applications in non-Gaussian signal processing," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5286–5298, 2007.
[31] A. Renyi, *On Measures of Entropy and Information*, University of California Press, Berkeley, California, 1961.
[32] N. Zhou, H. Cheng, J. Qin, Y. H. Du, and B. D. Chen, "Robust high-order manifold constrained sparse principal component analysis for image representation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 7, pp. 1946–1961, 2019.
[33] J. C. Jiang, H. Tao, R. D. Fan, W. Z. Zhuge, and C. P. Hou, "Incomplete multi-view learning via half-quadratic minimization," *Neurocomputing*, vol. 443, pp. 106–116, 2021.
[34] P. P. Zhao and Y. M. Huang, "Conjugate gradient method preconditioned with modified block ssor iteration for multiplicative half-quadratic image restoration," *Calcolo*, vol. 57, no. 3, Article ID 31, 2020.