A new model for the simplification of particle counting data

M. F. Fadal[1], J. Haarhoff[1], and S. Marais[2]

[1]Department of Civil Engineering Science, University of Johannesburg, South Africa
[2]Process Technology Department, Rand Water, South Africa

Correspondence to: J. Haarhoff (jhaarhoff@uj.ac.za)

Abstract. This paper proposes a three-parameter mathematical model to describe the particle size distribution in a water sample. The proposed model offers some conceptual advantages over two other models reported on previously, and also provides a better fit to the particle counting data obtained from 321 water samples taken over three years at a large South African drinking water supplier. Using the data from raw water samples taken from a moderately turbid, large surface impoundment, as well as samples from the same water after treatment, typical ranges of the model parameters are presented for both raw and treated water. Once calibrated, the model allows the calculation and comparison of total particle number and volumes over any randomly selected size interval of interest.

1 Introduction

The power of particle counters to provide a detailed description of the numbers and sizes of particles in a suspension is often not fully exploited, although its potential had been realised some decades ago in fields as diverse as phycology and water treatment (Sheldon, 1979; Lewis and Manz, 1991). The counters produce a count and the size limits for numerous channels, and the full meaning of the analysis is often obscured by a sheer weight of numbers. A method is required to compact the multitude of numbers from every count to as few as possible parameters to offer a reliable description of the particle size distribution. A second useful application of a generalised description is to allow the comparison of particle counts made by different particle counters with their own unique channel size settings. Such models have been proposed and used in the past. It is the objective of this paper to firstly propose the use of a new, improved model, aiming to overcome some of the weaknesses of earlier proposals. Secondly, the model will be applied to a large data set of particle counts collected before and after treatment at a large South African drinking water supplier.

2 Theoretical development

2.1 The power law

The commonly used power law is simply a straight line defining the normalised particle counts N (y-axis) in terms of the geometric mean size d of each counting channel (x-axis) on a log-log plane. The power law has some very real conceptual weaknesses – at small particle sizes, the particle number tends to infinity; at large sizes, the particle volume tends to infinity (Wilczak et al., 1992). The model and its calibration equations are, for n channels:

$$N = A \cdot d^{\beta}$$

$$\begin{bmatrix} \ln A \\ \beta \end{bmatrix} = \begin{bmatrix} n & \sum \ln d \\ \sum \ln d & \sum (\ln d)^2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum \ln N \\ \sum \{(\ln N)(\ln d)\} \end{bmatrix} \quad (1)$$

2.2 The variable-β model

To rectify the weaknesses of the power law, a variable-β model was proposed in conceptual form with no calibration data (Lawler, 1997). On a log-log plane, the variable-β model plots as an inverted parabola, centred about an axis at

$d = 1\,\mu\text{m}$. The variable-β model and its calibration equations are, for n channels:

$$N = A \cdot d^{\beta \ln d}$$

$$\begin{bmatrix} \ln A \\ \beta \end{bmatrix} = \begin{bmatrix} n & \sum (\ln d)^2 \\ \sum (\ln d)^2 & \sum (\ln d)^4 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum \ln N \\ \sum \{(\ln N)(\ln d)^2\} \end{bmatrix} \quad (2)$$

The variable-β model was calibrated and compared to the power law in an exhaustive study which used the particle counts from 1432 water samples, ranging from raw surface water to treated drinking water, including samples from the intermediate treatment steps (Ceronio and Haarhoff, 2005). It was conclusively demonstrated that the variable-β model provided a better fit than the power law.

2.3 A proposed refinement to the variable-β model

Despite the improved fit provided by the variable-β model, it was pointed out that the variable-β model has an important limitation (Ceronio and Haarhoff, 2005). Regardless of the values of A and β, the maximum N would always be found at a size of $d = 1\,\mu\text{m}$, regardless of the nature of the suspension. To remove this limitation, a further conceptual improvement was offered, without any further development or validation. The suggested three-parameter model is called the Ceronio model in this paper and plots as an inverted parabola on a log-log plane, without any constraints on the position of the vertical axis. The Ceronio model and its calibration equations are, for n channels:

$$N = A \cdot d^{\beta \ln d + C}$$

$$\begin{bmatrix} \ln A \\ \beta \\ \ln d \end{bmatrix} = \begin{bmatrix} n & \sum (\ln d)^2 & \sum \ln d \\ \sum (\ln d)^2 & \sum (\ln d)^4 & \sum (\ln d)^3 \\ \sum \ln d & \sum (\ln d)^3 & \sum (\ln d)^2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum \ln N \\ \sum \{(\ln N)(\ln d)^2\} \\ \sum \{(\ln N)(\ln d)\} \end{bmatrix} \quad (3)$$

It is noted in passing that the first matrix on the right-hand side of the calibration equations (for all the models above) is a function of the channel settings of the particle counter only, without being affected by the counts. The onerous inversion of the matrix has therefore only to be performed once for every instrument setting. Figure 1 illustrates the three models for a randomly selected particle count.

Once the models are calibrated, they can be used to rapidly obtain any desired property of the suspension. To obtain the total number of particles in any random size interval between d_1 and d_2, using the Ceronio model for illustration:

$$\#_{d_1, d_2} = A \int_{d_1}^{d_2} d^{\beta \cdot \ln d + C} \cdot dd \quad (4)$$

The corresponding total particle volume (assuming the particles to be spheres) in any random size interval between d_1 and d_2 is calculated with:

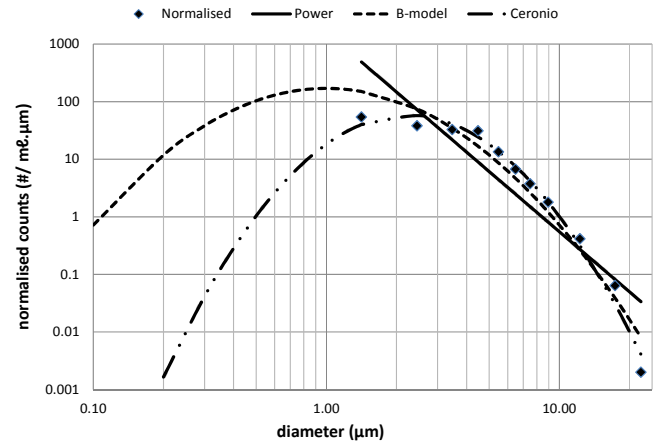$$V_{d_1, d_2} = \frac{\pi A}{6} \int_{d_1}^{d_2} d^{\beta \cdot \ln d + C + 3} \cdot dd \quad (5)$$

Figure 1. The power law, the variable-β model and the Ceronio model, fitted to the same particle count.

The power of the Ceronio model lies predominantly in its ability to model suspensions where the maximum normalised counts deviate from $d = 1\,\mu\text{m}$. The diameter where the normalised count reaches a maximum is provided by:

$$d_{\max} = e^{-C/2\beta} \quad (6)$$

3 Model application

3.1 Particle counting data collection

Rand Water is a drinking water supplier supplying about 3.7 million m^3/day to a population of roughly 11 million people in the Gauteng Province, as well as parts of the Mpumalanga, Free State and North West provinces of South Africa. Its primary water source is Vaal Dam, an impoundment of 2536 million m^3. Raw water is conveyed from the Vaal Dam by open channel and pipes to two treatment plants – Zuikerbosch (ZB) and Vereeniging (VG). At Zuikerbosch, the bulk of the raw water is first retained in a large balancing tank before it proceeds to treatment; at Vereeniging there is no balancing tank and the water is treated directly upon arrival.

Particle counting is performed on samples taken directly from the sampling taps on the incoming and outgoing pipelines. This paper utilises the counts for four sampling positions, namely the two raw water sampling points (ZB Raw and VG Raw) and two selected points on the treated water (pipelines B10 and A20, henceforth labelled ZB Final and VG Final), yielding a total of 321 samples. The data used was collected at roughly fortnightly intervals from August 2006 to September 2009, covering slightly more than three years. The samples were transported in glass containers to the laboratory and counted forthwith with the same particle counter.

Table 1. Percentage of samples that recorded zero counts in the channels indicated.

| Channel	d_{min}	d_{max}	ZB Raw	ZB Final	VG Raw	VG Final
9	10	15	0 %	0 %	0 %	0 %
10	15	20	0 %	0 %	0 %	0 %
11	20	25	6 %	14 %	16 %	9 %
12	25	30	37 %	44 %	46 %	38 %
13	30	40	52 %	46 %	64 %	46 %
14	40	50	87 %	71 %	86 %	78 %
15	50	100	92 %	88 %	93 %	86 %

3.2 Data screening

The particle counter used was a PAMAS 3116 FM with 16 channels, covering the range from 1 µm upwards. The channels used are separated at 1; 2; 3; 4; 5; 6; 7; 8; 10; 15; 20; 25; 30; 40; 50, and 100 µm. From these channel boundaries, the geometric mean of each channel was calculated to obtain the "d" required for the calibration matrices provided earlier. From the differential counts in each channel, the normalised counts were calculated by dividing them by the width of each channel, to obtain the "N" in the calibration matrices. The d- and N-values were used for further analysis.

For both the raw and treated water samples, there were very few counts in the higher size ranges. A necessary data screening step was to eliminate those larger channels which returned zero values, therefore not contributing to a meaningful fit of the data. The results are shown in Table 1. The channels in the lower half of Table 1 were eliminated from further consideration, based on the large percentage of samples having zero particle counts. There were therefore $15 - 4 = 11$ data points available for each calibration. After accounting for the three parameters in the Ceronio model, this left $11 - 3 = 8$ degrees of freedom, which is considered adequate for the purpose of reliable model calibration. The few zero counts in channel 11 were replaced by values of "1" to prevent the calibration procedure from trying to take the logarithm of zero.

3.3 Comparison of the Ceronio and variable-β models

Both the Ceronio and variable-β models were calibrated for each of the 321 samples. The goodness of fit for each sample was determined from the sum of squares SS, i.e. the sum of the squared differences between the logarithm of the actual count and logarithm of the modelled count. Figures 2 and 3 show the cumulative distributions for the SS for the Zuikerbosch and Vereeniging samples respectively. They clearly show the improvement in fit brought about by the Ceronio model. The sums of squares were reduced by 30 to 40 % in all cases, the improvement thus being about the same for both treatment plants, and for both raw and treated water.
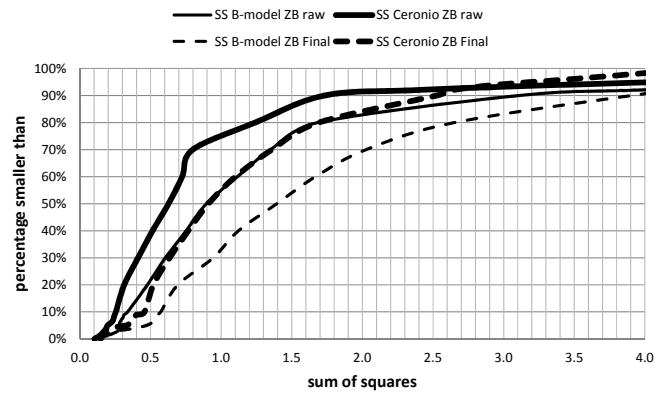
Figure 2. Cumulative sum of square for the ZB treatment plant.
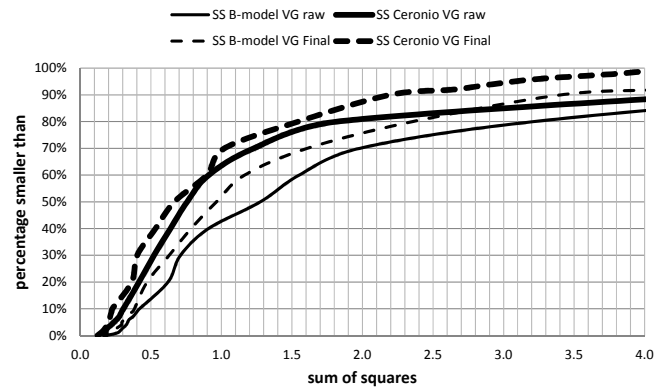
Figure 3. Cumulative sum of square for the VG treatment plant.

4 Discussion of the Ceronio model

Some of the samples included in Figs. 2 and 3 were not modelled very accurately, as evidenced by their large sum of squares. For this section, where general guidelines for the Ceronio model constants are discussed, the data had to be further screened to include only those samples which could be modelled within a sum of squares of 2. This filtering step removed 56 (17 %) samples from the data set, about evenly spread amongst the four sampling positions, which left 265 samples (VG Raw: $n = 65$; VG Final: $n = 67$; ZB Raw: $n = 69$; ZB Final: $n = 64$). The parameter values of these samples were used to determine the cumulative distributions for the four sampling points discussed below.

Parameter A determines the height of the size distribution, as shown in Fig. 4. As expected for a surface water impoundment subject to sharp seasonal turbidity variations, this parameter covers a broad range. The value of A corresponds directly to the normalised count at $d = 1$ µm, similar to its interpretation for the variable-β model. From Fig. 5, the raw water samples had A-values about two orders of magnitude higher than the final treated water samples.

Parameter β, as shown in Fig. 6, determines the curvature of the size distribution. The interesting observation from
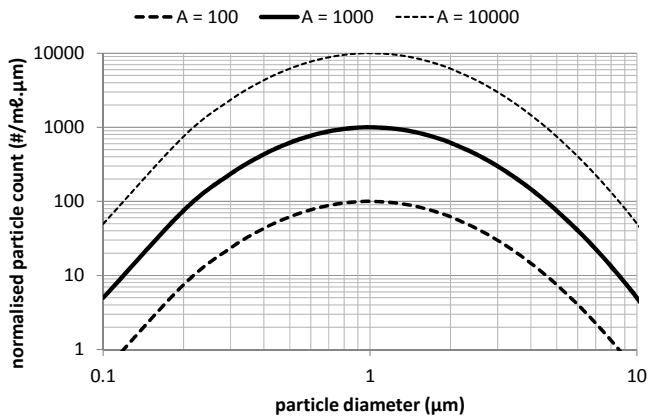
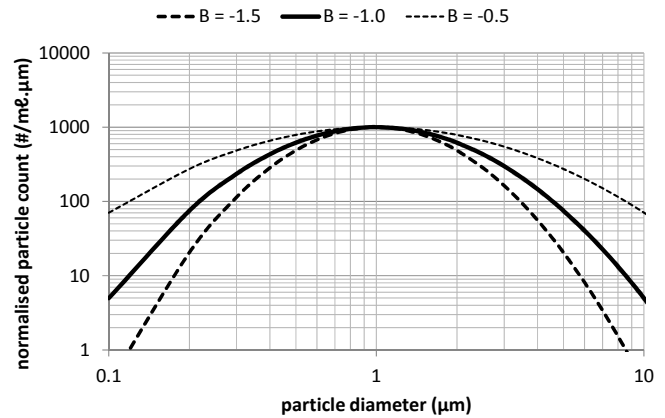**Figure 4.** The effect of model parameter A on the particle size distribution, with $\beta = -1$ and $C = 0$.

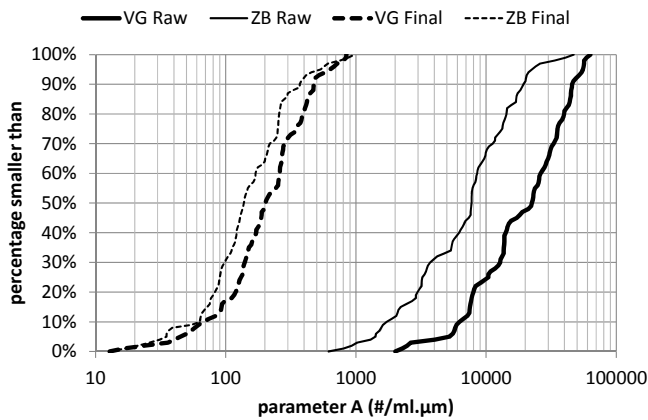**Figure 6.** The effect of model parameter β on the particle size distribution, with $A = 1000$ and $C = 0$.

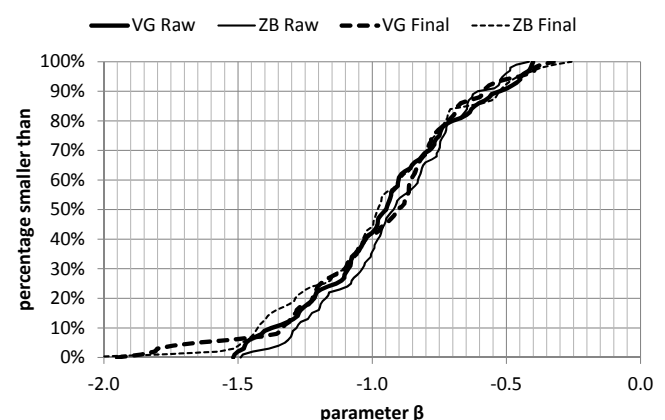**Figure 5.** The cumulative distribution of model parameter A.

Figure 7. The cumulative distribution of model parameter β.

Fig. 7 is that the cumulative distributions for the raw and treated waters are not different. Using the 10th and 90th percentiles as guidelines, the range of β was from -1.4 to -0.5.

Parameter C moves the size distribution from left to right, as shown in Fig. 8. Figure 9 indicates that the range of C is between -1.5 to $+1.5$. The C-values of the raw water samples lie consistently to the left of the C-value of the treated water samples, indicating that the raw waters had relatively more small particles than the treated water samples. The median value of d_{max} for the raw water samples is at $0.8\,\mu m$ (10th percentile $0.4\,\mu m$, 90th percentile $1.6\,\mu m$). Although the median value of d_{max} for the treated water samples is at $1.0\,\mu m$ (10th percentile $0.3\,\mu m$; 90th percentile $1.7\,\mu m$), its variability emphasises the weakness of the variable-β model, which forces d_{max} to be $1\,\mu m$ in all cases.

It is pointed out that the d_{max} values discussed in the previous paragraph are values predicted from the Ceronio model, and are not directly supported by the particle counting data. The smallest particles that could be counted, due to the technological limitations of the particle counter, are in the interval between 1 and $2\,\mu m$, which are characterised by a mean

particle diameter of $1.4\,\mu m$. This means that the values for d_{max} are mostly just smaller than the smallest particles that could actually be counted, with no data points to validate the shape and position of the apex of the predicted particle count. This is an unfortunate limitation, which of course applies equally to the validation of the variable-β and Ceronio models. This implies that, until this validation can be done by including smaller particles, not too much weight should be lent to the exact value of d_{max}. For purposes involving particle sizes above $2\,\mu m$, the general conclusions regarding the Ceronio constants are firmly supported by the data sets used.

5 Summary and conclusions

– Every sample processed by a modern electronic particle counter, produces an extensive list of numbers specific to the particular instrument setting. Such numbers are not easy to interpret, and also do not allow comparison
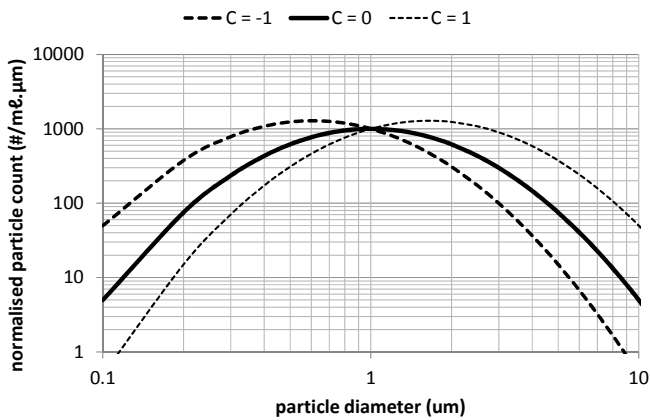
Figure 8. The effect of model parameter C on the particle size distribution, with $A = 1000$ and $\beta = -1$.

Figure 9. The cumulative distribution of model parameter C.

amongst particle counting results done with different instrument settings.

– Mathematical models present an opportunity to compress the particle counting data into two or three model parameters, which are much more amenable to analysis and interpretation. Moreover, the models are calibrated using all the channels with non-zero counts, thus providing a more robust description than could be obtained from using single counts from individual channels.

– The power law, commonly used, has two serious shortcomings which prohibit its use at both ends of the size spectrum. As the size gets smaller, the number of particles approaches infinity; as the size gets larger, the volume of particles approaches infinity.

– The variable-β model removes the shortcomings of the power law, but has its own limitation. Regardless of the sample, the maximum normalised particle count is always found at a size of 1 μm.

– This paper developed an earlier proposal to remove the limitation of the variable-β model, here called the Ceronio model. By introducing a third model parameter, the Ceronio model can describe a distribution with the maximum normalised count at any size.

– Using 321 samples of both raw and treated water from a large South African drinking water supplier, the Ceronio model consistently provided a better fit than the variable-β model, reducing the sum-of-squares by between 30 % and 40 %.

– A systematic analysis of the three parameters describing the Ceronio model provided useful insights. Parameter A is a parameter closely related to the number of particles in a sample; parameter B provides a measure of the ratio of smaller to larger particles; and parameter C
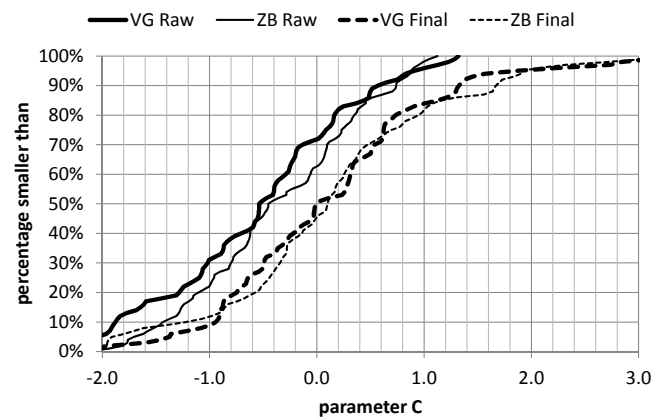
predominantly influences at which size the maximum normalised particle count is found.

– Potentially, the most important predictions of the Ceronio model deals with those small particles between 0.1 μm and 2 μm. This is the size range where the maximum normalised particles counts are found, and also where the fundamental transport mechanisms of particles in water are at their weakest. Unfortunately, the particle counter used could not measure sufficient data points in this range to provide solid experimental verification of the Ceronio model, indicating an important research need.

– For larger particles, the Ceronio model provides a robust tool for the computation and comparison of particle numbers and volumes at any desired particle size interval.

References

Ceronio, A. D. and Haarhoff, J.: An improvement on the power law for the description of particle size distribution in potable water treatment, Water Res., 39, 305–313, 2005.

Lawler, D. F.: Particle size distributions in treatment processes: theory and practice. Proceedings of the Fourth International Conference: The role of particle characteristics in separation processes, IAWQ-IWSA Joint Specialist Group on Particle Separation, Jerusalem, 29–30 October 1996, 1997.

Lewis, C. M. and Manz, D. H.: Light-scatter particle counting: Improving filtered-water quality, J. Environ. Eng., 117, 15 pp., 1991.

Sheldon, R. W.: Measurement of phytoplankton growth by particle counting, Limnol. Oceanogr., 24, 760–767, 1979.

Wilczak, A., Howe, E. W., Aleta, E. M., and Lee, R. G.: How pre-oxidation affects particle removal during clarification and filtration, Journal of the American Water Works Association, 84, 85–94, 1992.