Journal of Aerosol Science 171 (2023) 106177

Contents lists available at ScienceDirect

Journal of Aerosol Science

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

Artificial neural network based coincidence correction for optical aerosol spectrometers

Lukas Oeser [a,b,*], Nakul Samala [a], Lars Hillemann [a], Jan Müller [a], Claudia Jahn-Wolf [a], Jens Lienig [b]

[a] Topas GmbH, Gasanstaltstraße 47, D-01237, Dresden, Germany
[b] Technische Universität Dresden, Institute of Electromechanical and Electronic Design, D-01062, Dresden, Germany

ARTICLE INFO

Handling Editor: Chris Hogan

ABSTRACT

Many applications such as filter testing, healthcare, quality monitoring, and environmental measurements require precise aerosol quantification by optical aerosol spectrometers. This type of measurement equipment is capable of in-situ measurements and provides easy access to the size distribution of the particles. Due to the coincidence error, optical aerosol spectrometers are limited to applications with relatively low concentrations. At high concentrations, the counting efficiency is reduced, while the size distribution is shifted towards larger particles. In 1984 Raasch and Umhauer proposed an analytical correction method for the size distribution. Although the approach is easy to implement, it has some disadvantages. In this work, an alternative correction method for the size distribution is presented, which is based on neural networks. The performance of both correction methods is evaluated on the cumulative distribution of raw detector voltages. The relative error of the median, as well as an error integral over the whole distribution is used as a measure. The neural network-based method gives a correction result that shows approximately half the relative median error, and a third of the error integral compared to the method of Raasch and Umhauer, for high concentrations.

1. Introduction

Optical aerosol spectrometers are used to measure particle sizes, ranging from 70 nm up to 25 µm (VDI 3867 Blatt 4:2011–16, 2011). The measuring principle is based on light scattering. An aerosol is passed through an illuminated measuring volume. If a particle is located in the measuring volume, the light is scattered in different spatial directions. For spherical particles, this effect can be described by the Mie theory (Mie, 1908). An optical system collects the scattered light and direct it to a photodetector, where the light is converted into an electrical signal. Every particle crossing the measuring volume generates a pulse on the output signal. The pulse amplitude contains the information about the particle size and depends on multiple variables such as geometrical properties, refractive index, and wavelength. All pulses can be counted and assigned into different amplitude classes. The resulting histogram can then be used to calculate the concentration and size distribution of the aerosol. However, the measuring principle is limited to relatively low number concentrations (max. 10^5 cm^{-3} (VDI 3867 Blatt 4:2011–16, 2011)). At high concentrations, the probability of more than one particle in the measuring volume at a time increases, which causes overlapping particle pulses on the detector signal that cannot be

* Corresponding author. Topas GmbH, Gasanstaltstraße 47, D-01237, Dresden, Germany.
E-mail address: loeser@topas-gmbh.de (L. Oeser).

https://doi.org/10.1016/j.jaerosci.2023.106177
Received 13 January 2023; Received in revised form 16 March 2023; Accepted 18 March 2023
Available online 28 March 2023

separated anymore (Oeser et al., 2022). As a result, the counting efficiency of the device is reduced and the measured size distribution is shifted towards larger particles [(Aerosol measurement, 2011), p. 62], (ISO 21501-1. ISO 21501-1:2009, 2150), (Raasch & Umhauer, 1984).

There are different techniques to minimize the coincidence error. Equation (1) can be used to calculate the counting efficiency η for a certain number concentration based on Poisson statistics [(Aerosol measurement, 2011), pp. 288–89], (ISO 21501-1. ISO 21501-1:2009, 2150), (VDI 3867 Blatt 4:2011–16, 2011). In Fig. 1 the counting efficiency function for one of the modified optical aerosol spectrometers used in this work is shown. Q_m is the sample flow rate through the measuring volume and τ is the recovery time (also known as dead time) of the electronics. Commonly, the recovery time is assumed to be as long as a single particle needs to pass through the measuring volume. So the product $Q_m \cdot \tau$ equals the measuring volume V_m. As can be seen, a small measuring volume minimizes the coincidence loss, because it decreases the probability of two or more particles in the measuring volume V_m. Practically, the measuring volume cannot be minimized infinitely due to physical and aerodynamic constraints. Modern signal processing techniques enable the deconvolution of overlapping particle pulses. This can reduce the detector dead time significantly. As a result, the devices can be used at higher number concentrations, without increasing the coincidence error. (Oeser et al., 2022).

In Equation (1) C_m represents the measured number concentration, which is usually known, and C_a the actual number concentration, which is usually the size of interest. Although the equation cannot be solved analytically, the exponential term can be approximated with a linear model, which allows to solve the equation for the actual number concentration. The model is precise as long as $C_a \cdot V_m < 0.1$. For higher number concentrations a numeric solution is recommended (Refer to Appendix 7.1 for further information).

$$\eta = \frac{C_m}{C_a} = e^{-C_a \cdot Q_m \cdot \tau} = e^{-C_a \cdot V_m} \approx 1 - C_a \cdot V_m \qquad \text{Equation 1}$$

While the correction of the measured number concentration is relatively simple, the correction of the measured size distribution is an ill-posed problem. To the best of our knowledge, the only existing method for this purpose was published by Raasch and Umhauer back in 1984 (Raasch & Umhauer, 1984). The authors developed an analytical correction formula for the size distribution. The proposed method is based on Poisson statistics and assumptions that can deviate from real-world applications. In this work, we present an alternative approach. The correction, i. e., the transformation from measured size distribution to corrected size distribution is done with an artificial neural network.

2. Materials and methods

In this section, the proposed correction method for a size distribution, measured with coincidence losses is described. As the method from Raasch and Umhauer is used as a reference, and their original work was published in the German language, a brief introduction to their analytical solution is provided. Before the methods are presented in detail, the following conventions are made, which apply to the rest of this publication.

The term concentration is always referred to particle number concentration i. e. the number of particles per volume.

In optical aerosol spectrometers, the detector output voltage is most commonly directly proportional to the light intensity scattered by the particles. According to Mie theory, the relation between the signal amplitudes and the respective particle sizes is highly nonlinear (Mie, 1908). The transfer function depends on the specific device design and is usually determined by calibration against particle standards (ISO 21501-1. ISO 21501-1:2009, 2150). In order to analyze the coincidence error in general, the density distribution of the

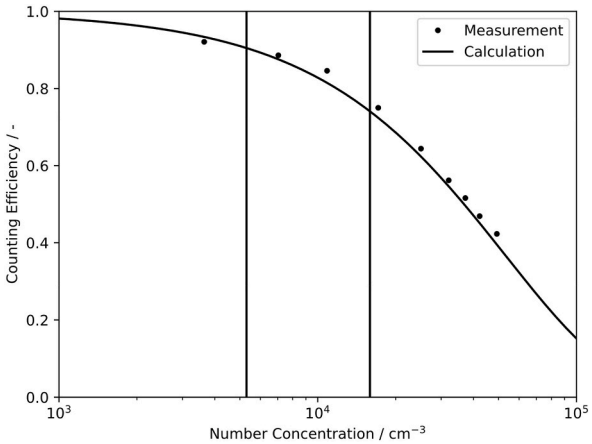

Fig. 1. Calculated and measured counting efficiency for one of the modified optical aerosol spectrometers used in this work. Measured against SMPS reference, without coincidence correction. The two vertical lines indicate the concentrations where $C_a \cdot V_m = 0.1$ and $C_a \cdot V_m = 0.3$. For concentrations smaller than $C_a \cdot V_m = 0.1$ no correction of the size distribution is required. Raasch and Umhauer suggest to use their method for number concentrations below $C_a \cdot V_m = 0.3$. At higher concentrations the correction result might differ significantly from the real density distribution.

pulse heights is used in this work, instead of the density distribution of the particle sizes. This allows general considerations, independent of the device-specific calibration function. Analogous to the conventional approach (ISO 9276–1:1998–06, 1998), the quantities in Table 1 are introduced (Cf. (Raasch & Umhauer, 1984)).

2.1. Coincidence correction by Raasch and Umhauer

In 1984 Raasch and Umhauer presented a correction method for measured size distributions which are affected by coincidence (Raasch & Umhauer, 1984). The approach is based on Poisson statistics. The probability that a certain number of particles is located in the measuring volume at a given time can be calculated by Equation (2).

$$P_\lambda(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \qquad \text{Equation 2}$$

The average number of particles λ in the measuring volume is determined from the aerosol concentration C_a and the size of the measuring volume V_m (Equation (3)).

$$\lambda = C_a \cdot V_m \qquad \text{Equation 3}$$

A constant aerosol velocity is assumed in the measuring volume. Thus, all particles need a constant time Δt to pass the volume. The probability that there are k particles in the measuring volume is therefore equal to the probability that a total of k particles enter the measuring volume during the time Δt. The probability that the difference of the entry times between two particles ($t_{i+1} - t_i$) is greater than the time Δt, corresponds to the probability that no particle is in the measuring volume. Using this relationship, the probabilities for different coincidence scenarios can be described by the equations shown in Table 2 (Cf. (Raasch & Umhauer, 1984)).

If two scattered light pulses with the amplitudes U_i and U_j overlap each other, the result is a pulse height $U_k = U_i + U_j$. Practically, this assumption is only valid for rectangular particle pulses. Most commercial devices use a Gaussian shaped laser beam which leads to Gaussian shaped particle pulses (Oeser et al., 2022). However, the particle pulses will be considered as rectangular shaped in the following. A measured particle pulse of a given amplitude must either be caused by a true single-particle pulse, or a combination of several particle pulses with different amplitudes, each smaller than the measured amplitude. The density distribution of the measured particle pulses can be divided into different amplitude classes of a given class width ΔU. The area under the curve corresponds to the relative frequency of this class (Fig. 2). For each class of the measured amplitude density distribution, this value is composed of several components. On the one hand from the probability of a true single-particle signal. On the other hand from the sum of all probabilities of all possible coincidence cases that would cause a signal with the corresponding amplitude.

With the considerations made above, the following equation (Equation (4) (Raasch & Umhauer, 1984),) can be derived to calculate a measured amplitude density distribution $q_0(U)$, i. e., with coincidence, from a known actual amplitude density distribution $q_0^*(U)$. The parameter α is the probability for a single particle count $e^{-\lambda}$. Raasch and Umhauer only considered coincidence scenarios caused by two particles. All other combinations are assigned to this case. Coincidences caused by more than two particles are much less likely as can be seen in Table 2.

$$q_0(U_k) \cdot \Delta U = \alpha \cdot q_0^*(U_k) \cdot \Delta U + (1-\alpha) \cdot \sum q_0^*(U_i) \cdot \Delta U \cdot q_0^*(U_j) \cdot \Delta U \qquad \text{Equation 4}$$

When calculating the sum, the condition $U_i + U_j = U_k$ must be met. From Equation (4) the more common integral form can be derived (Equation (5), (Raasch & Umhauer, 1984)).

$$q_0(U_k) = \alpha \cdot q_0^*(U_k) + (1-\alpha) \cdot \int_0^{U_k} q_0^*(U_j) \cdot q_0^*(U_k - U_j) \cdot \mathrm{d}U_j \qquad \text{Equation 5}$$

Equation (5) can be used to calculate a density distribution with a coincidence error $q_0(U)$, when the actual density distribution $q_0^*(U)$ is known. This can be used to estimate the error caused by coincidences. In practice, one is usually interested in the actual density distribution, and a density distribution with coincidence errors is given. For this purpose, Equation (5) can be rearranged as shown in

Table 1
Quantities used in this work.

Expression	Description	Unit
N	Sum of all detected particles	–
N_i	Number of particles, assigned to class i	–
U_{iL}	Lowest detector output voltage of class i	V
U_{iH}	Highest detector output voltage of class i	V
$U_i = \frac{U_{iH} - U_{iL}}{2}$	Mean detector output voltage of class i	V
$\Delta U_i = U_{iH} - U_{iL}$	Interval width of class i	V
$Q_0(U_{iH}) = \frac{\sum_{j=0}^{i} N_j}{N}$	Cumulative distribution of the pulse heights	–
$q_0(U_i) = \frac{1}{N} \cdot \frac{N_i}{\Delta U_i}$	Density distribution of the pulse heights	1/V

Table 2
Probabilities for different coincidence cases based on Poisson statistics.

Particle entry times	Probability	Description
$t_2 - t_1 > \Delta t$	$e^{-\lambda}$	No coincidence, single particle count
$t_2 - t_1 < \Delta t$; $t_3 - t_2 > \Delta t$	$(1 - e^{-\lambda}) \cdot e^{-\lambda}$	Single coincidence, two particles are counted as one.
$t_2 - t_1 < \Delta t$; $t_3 - t_2 < \Delta t; t_4 - t_3 > \Delta t$	$(1 - e^{-\lambda})^2 \cdot e^{-\lambda}$	Double coincidence, three particles are counted as one.
…	$(1 - e^{-\lambda})^n \cdot e^{-\lambda}$	n-fold coincidence, $n + 1$ particles are counted as one.

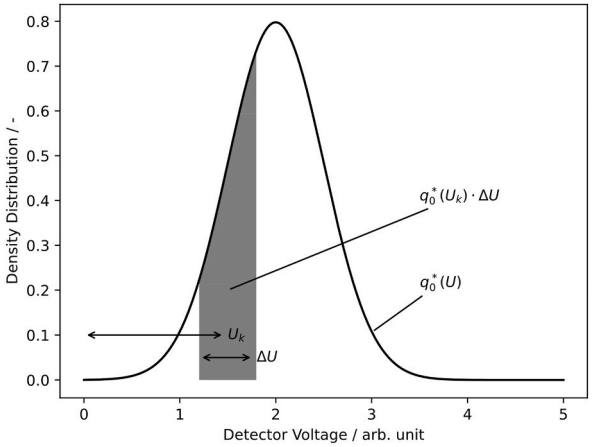

Fig. 2. Derivation of the correction method by Raasch und Umhauer. The actual density distribution is split into different amplitude classes. The probability (I. e. the area under the curve) for each amplitude class is the sum of the probability that a counted peak in this class was caused by a single particle, and the probabilities of all possible coincidence scenarios of multiple particles that would cause a peak amplitude of the respective class.

Equation (6).

$$q_0^*(U_k) = \frac{1}{\alpha} \cdot q_0(U_k) + \frac{(1-\alpha)}{\alpha} \cdot \int_0^{U_k} q_0^*(U_j) \cdot q_0^*(U_k - U_j) \cdot \mathrm{d}U_j \qquad \text{Equation 6}$$

The result is a nonlinear Volterra integral equation that has no analytical solution in general. However, the actual density distribution can be approximated iteratively. Therefore $q_0^*(U)$ is set to $q_0(U)$. In the next step, the result of Equation (6) is used as the actual amplitude density distribution. The procedure is repeated until a sufficient solution is available. In practice, the method converges after a few iterations (<10).

According to Raasch and Umhauer, the method is very accurate for concentrations in the range of $0.1 < C_a \cdot V_m < 0.3$. At lower concentrations, the correction is still possible, but not necessary, as the coincidence error has only a negligible effect on the particle size distribution. When the concentration is above $C_a \cdot V_m > 0.3$ the correction can still be applied, but larger deviations between the corrected and actual distribution are to be expected (Cf. Fig. 1).

The method published by Raasch and Umhauer is mathematically robust and easy to implement. Nevertheless, a few drawbacks have to be mentioned. For example, the result of Equation (6) can be negative in some cases. Further, the assumptions of rectangular-shaped particle pulses and the disregard of coincidences caused by more than two particles are critical. Most commercial devices do not include the coincidence correction by Raasch and Umhauer. The correction has to be done afterwards by the user, which requires detailed knowledge of the measuring cell. Unfortunately, most instrument manufacturers do not provide information about their measuring volume.

2.1.1. Coincidence correction with artificial neural networks

As shown above, the coincidence correction is a transfer function between measured and actual size distribution. The transfer function is too complex to describe all correlations exactly. Practically, a good approximation is sufficient in many cases. Therefore, the coincidence correction can be formulated as a multidimensional regression problem. Such problems can be solved by neural networks (Agatonovic-Kustrin & Beresford, 2000), (Lathuiliere et al., 2020). However, to the best of our knowledge, neural networks were not used for coincidence correction of a measured particle size distribution before. The major difference between the analytical approach and the neural network is that the neural network estimates the actual size distribution, based on the experience it learned in training, instead of a fixed equation. Thus, knowledge about the actual transfer function is not necessary. A significant advantage of neural networks is that they can, in theory, represent arbitrarily complex relationships between input and output data. Neural networks are

successfully used in numerous regression and pattern recognition tasks and often deliver superior results in comparison to conventional methods (Agatonovic-Kustrin & Beresford, 2000), (Lathuiliere et al., 2020), (LeCun et al., 1989).

Basically, neural networks consist of artificial neurons, which are connected by weights. Each neuron defines a relationship between a set of input values x_i and a corresponding output value y_i. A single artificial neuron can be described by Equation (7) (Agatonovic-Kustrin & Beresford, 2000), (Dasaradh, 2020), (Nwankpa et al., 2018). First, the weighted sum of all input values is calculated. The weights w_i are used to adjust the sensitivity of the neuron for the corresponding input value. The output of a neuron is calculated by the weighted sum using an activation function φ. In recent years, various activation functions such as sigmoid, ReLu, softsign, and hyperbolic tangent have been established (Nwankpa et al., 2018).

$$y = \varphi\left(\sum w_i \cdot x_i\right) \qquad \text{Equation 7}$$

Multiple neurons can be connected to each other. This means that the output of one neuron is used as the input of another neuron. In this way, arbitrary structures can be created. A convenient approach is to organize all neurons in several layers (Fig. 3). The output values of the previous layer are taken as input values for the following layer. Neurons of the same layer have no connection with each other. The number of layers, as well as the number of neurons per layer, can be chosen freely. This network topology is commonly referred to as a feedforward network.

Based on the previous considerations, the transfer function of the neural network depends on the structure of the neural network, the activation function of the individual neurons, and the weights of the individual neurons. Initially, the weights are set to random values. Known datapoints can be used to estimate the performance of the neural network, i. e., whether the weights are chosen properly or not. This requires a dataset that contains input values as well as the desired output values. For each input value of the dataset, the response of the network is calculated and compared with the desired output. The more efficient the network, the smaller the difference between the desired, and the actual output values. The deviation is measured with an error function, e. g. mean squared error. During training, the weights of the network are optimized iteratively, such that the error on the training dataset is minimal. For this purpose, different optimization algorithms can be used [(Goodfellow et al., 2016), pp. 271–325], (Kingma & Ba, 2014). Once the weights are converged, the network can predict the desired output values for new input values. After training, the neural network should be tested with some known datapoints, which were not used in training. This validation ensures that the network did not only learn the datapoints in the training dataset. In the case of overfitting, the network would give good results on the training set, but poor results on new data. Overfitting can be a challenge, especially when large networks are trained on small datasets.

Therefore, a large dataset is required for coincidence correction. It should contain a variety of size distributions from different aerosols at different concentrations. Ideally, all possible aerosol scenarios are present in the dataset. Each datapoint must always consist of a measurement with a coincidence error, and a measurement without a coincidence error for the same aerosol. The measurement with coincidence error is used as input data for the network, while the measurement without coincidence error is used as the desired output value. A neural network can be trained on such a dataset and used for coincidence correction.

In this work, the raw counts of the instrument are used for correction, since the coincidence error has a direct effect on them. However, the presented method can also be applied to the final measurement result in the form of e. g. a number or mass distribution. In this case, the correction depends on the calibration function between pulse height and particle size. If the correction should be applied to a particle system with different optical properties, the network must be trained again, or in the worst case, when the raw counts are not available, the dataset has to be recorded again.

2.2. Measurement setup

The training of the artificial neural network requires a dataset. In addition to a sufficient amount of datapoints, it is important that

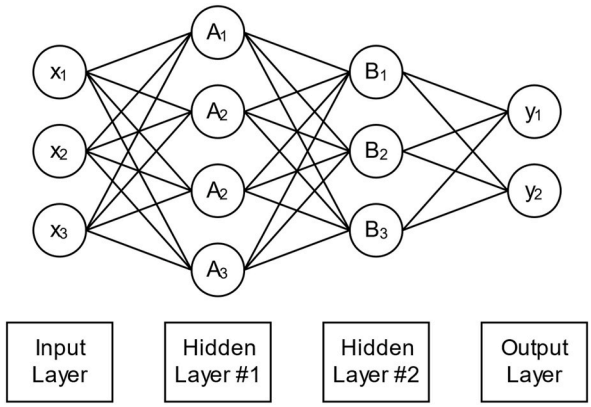

Fig. 3. Structure of a simple feedforward neural network. Each output node of a layer is connected to each input node of the following layer. For each set of input values a set of output values is calculated, based on the chosen structure of the neural network, the individual weights, and the activation functions of the neurons.

the dataset covers as many coincidence correction scenarios as possible. The dataset should contain samples of many differently shaped distributions at high and low concentrations. It is impossible to record a data set that completely covers all aerosol scenarios. Individual distributions may differ significantly in dispersity and modality. Due to the upper and lower detection limit with regard to the particle size, it is also possible that the measured distribution is truncated and shows only a part of the actual particle size distribution. All these cases should be present in an optimal dataset for coincidence correction. If the dataset should be recorded automatically, a fixed measurement setup is desired. In this work, we used the setup shown in Fig. 4.

The setup shown in Fig. 4 allows variation of concentration and size distribution over a wide range. A Collison type (May 1973) aerosol generator (ATM 222, Topas GmbH, Germany) is used to generate a Di-Ethyl-Hexyl-Sebacat (DEHS) aerosol. The aerosol concentration, as well as the output flow of the generator, can be adjusted with the nozzle pressure.

An electrostatic classifier (Model 3080, TSI Inc., USA) is placed after the generator to manipulate the size distribution. Here, the particles are charged by the integrated 170 MBq ^{85}Kr neutralizer (Model 3077A, TSI Inc., USA, 170 MBq, 2010) and passed through a differential mobility analyzer (DMA; Model 3081, TSI Inc., USA). Inside the DMA, an electric field is generated by a high-voltage source. This electric field is used to separate the particles, depending on their electrical mobility. Thus, the particle size of the aerosol can be adjusted by the voltage. Multiple charged particles can cause the aerosol to contain additional modes of larger particles. The quality of the classification depends on the ratio between aerosol and sheath air flow (Chen et al., 1999). While the aerosol flow is defined by the generator, the sheath air flow can be set by the electrostatic classifier, to vary the width of the size distribution of the aerosol.

The conditioned aerosol is diluted in a mixing chamber with clean air from a blower (RFU 564, Topas GmbH, Germany). Thus, the concentration can be changed without adjusting the actual distribution. This allows the measurement of an aerosol at different concentrations (with and without coincidences), without affecting the size distribution. The mixing chamber used in this work consists of a 500 mm long tube with an inner diameter of 20 mm. If the mixing chamber is selected too large, it may take a very long time for the desired aerosol to get to the spectrometer. A minimum flow rate of 3 l/min causes a dwell time of 3.2 s. A settling time of 30 s was awaited before each measurement.

One out of three modified aerosol spectrometers (LAP 323, Topas GmbH, Germany) was used as a measuring device for each dataset shown in Table 4. The measuring volume of the modified devices was $V_m = 0.0188$ mm^3 (Cf. Fig. 1). Data acquisition was performed using a Field Programmable Gate Array (FPGA) prototyping platform (Eclypse Z7, Digilent, USA) in combination with an Analog-to-digital converter (ADC) module (Zmod Scope 1410–125, Digilent, USA). In this work, peak detection was done by a simple threshold-based algorithm. The peaks of the detector signal were assigned into 256 logarithmic equidistant classes from 5 mV to 5 V. Differences in the signal path between the three devices are only caused by tolerances of the individual components e. g. lasers, optics, and electronics, as they would typically occur in device manufacturing.

For each data point, a random operating point (generator nozzle pressure, DMA particle size, DMA sheath air, and blower flow) was selected. The parameter range for the device settings is shown in Table 3. For the chosen operating point two measurements where performed. One with coincidence error and one without coincidence error which is used as reference. It is not possible to completely eliminate the coincidence error in optical aerosol spectrometers. However, the effect of coincidences on the measured size distribution is negligible for concentrations below $C_a \cdot V_m = 0.1$. Thus the reference measurement requires a dilution of the aerosol. This was achieved by setting the controlled blower to the maximum flow rate which was 60 l/min. Note that the blower used in this work can provide flow rates of up to 100 l/min, but due to the pressure drops in our setup we had to select a smaller value. For the modified measuring cell, this coincidence limit corresponds to a concentration of 5590 cm^{-3} (Cf. Fig. 1). All datapoints with a reference concentration higher than this value must be sorted out. In this work, a significantly lower limit of 3000 cm^{-3} was used. To estimate the influence of device-specific tolerances on our method, a total of three datasets were recorded on each of the modified aerosol spectrometers (Device A, B, and C). Table 4 shows an overview of the datasets used in this work.

2.3. Training

As mentioned before, the measurement system used in this work has 256 raw-size channels. However, the resolution was reduced to 32 size classes, as most commercial devices provide a resolution in this range (VDI 3867 Blatt 4:2011–16, 2011).

It is not recommended to use the raw data from the aerosol spectrometer for the training of the artificial neural network. For better

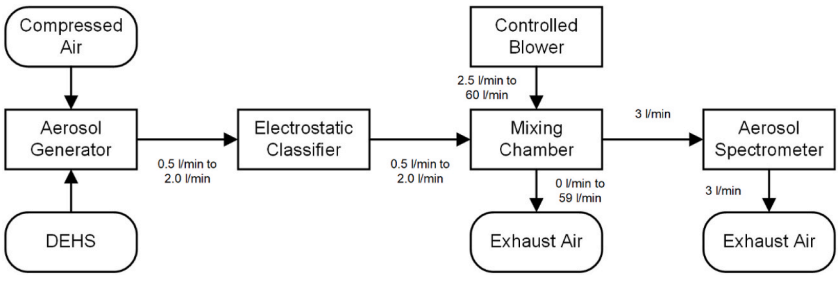

Fig. 4. Measurement setup to generate different aerosols. Concentration and size distribution of the aerosol can be varied by changing the operating parameters of the aerosol generator, electrostatic classifier, and the controlled blower.

Table 3
Parameter range used for aerosol generation. For each datapoint a random operating point was chosen.

Parameter	Device	Minimum	Maximum
Nozzle pressure	Aerosol generator	50 hPa	900 hPa
Sheath flow	Electrostatic classifier	4 l/min	20 l/min
Particle size	Electrostatic classifier	200 nm	2 µm
Dilution flow	Blower	3 l/min	30 l/min*

* For the measurment with coincidence error 30 l/min was used as the maximum dilution flow. All reference measurements where done with a dilution flow of 60 l/min.

Table 4
Overview of the datasets used in this work. Each dataset was obtained with a different optical aerosol spectrometer of the same model. Training data of device A was used for training of NN1 and NN2. Training data of device B was used for training of NN2. The dataset of device C was only used for evaluation.

Device	Data points (total)	Trained networks	Data points ($C_{ref} < 3000$ cm^{-3})			
			Total	Training	Validation	Test
A	1462	NN1, NN2	945	756	95	94
B	1500	NN2	1030	824	103	103
C	281	–	197	0	0	197

performance, the dataset has to be preprocessed first. Therefore, it is common to transform both, input and output data, into a fixed range of values [(Bishop, 1995), pp. 296–99]. In this work, the raw counts were normed by dividing each value by the maximum value of the vector. Unfortunately, the information about the particle concentration is lost in this normalization process. As the coincidence correction heavily depends on the number concentration of the aerosol, this information is feed as an additional value to the neural network. For the normalization of the measured concentration, a logarithmic scaling according to Equation (8) was used. Slope m and shift n are selected in such a way that all measured concentrations can be represented in a value range between 0 and 1.

$$x_0 = m \cdot \log(C_m) + n \qquad \text{Equation 8}$$

After preprocessing, the dataset has to be split into training, validation, and test data [(Ripley, 1996), p. 354]. This is important so that the neural network is not evaluated based on datapoints it was trained on, to prevent overfitting. Thus, the training data is used exclusively for training, and not for evaluation. The performance of the neural network depends on the chosen structure, which is defined by several hyperparameters (e. g. number of layers, number of neurons per layer, and activation function). One important step when training a neural network is to find a good set of hyperparameters. For this purpose, the validation data is used. When an adequate set of hyperparameters is found, the final model can be evaluated with the test dataset. In this work a common 80/10/10 (training/validation/test) split of the datasets from device A and B were used (refer to Table 4). The dataset of device C is significantly smaller and was only used for evaluation purposes.

In this work two neural networks, namely NN1 and NN2 were trained. As can be seen in Table 4, NN1 is trained with the data of device A only. In contrast, NN2 was trained with the data of device A and device B. The artificial neural networks used in this work were implemented in Python using Keras (Team, 2022). The application programming interface (API) contains many useful functions that simplify the training and modeling of the neural networks. For coincidence correction a simple feedforward neural network [(Ripley, 1996), pp. 143–80] is sufficient.

The network topology for coincidence correction used in this work is shown in Table 5. The number of layers, as well as the number of neurons per layer, and the activation functions were found by random grid search. For a regression task, the output layer does not necessarily need an activation function. However, we used the ReLu function (Nwankpa et al., 2018) to prevent negative concentration values.

Table 5
Neural network structure used for coincidence correction.

Layer	Neurons	Activation function
0	33 (input)	–
1	75†	Hyperbolic tangent†
2	69†	Softsign†
3	72†	Softsign†
4	32 (output)	ReLu

† Hyperparameter.

3. Results and discussion

First, a neural network was trained on the dataset of device A (NN1). NN1 was then used to correct the distributions of the test dataset. For a direct comparison, the method of Raasch and Umhauer was used on the same data. The diagrams in Fig. 5 show the results for several datapoints of the test dataset. Each of the four diagrams shows the density distribution and cumulative distribution of pulse heights for different aerosols generated by the setup from Fig. 4. As can be seen, the shape, size, and concentration of the aerosol were varied, as intended. The grey curve shows the result of the measurement with a coincidence error. The actual distribution, which was measured without coincidence error, is represented by the black line. The red and blue curves correspond to the corrected distribution by NN1 and the method of Raasch and Umhauer. For an ideal coincidence correction, the curve of the corrected distribution would match the curve of the actual distribution (black).

In the top/left diagram, the curves match each other perfectly. At low concentrations, the reference is reproduced well by both correction methods. As the concentration is well below the coincidence limit of the used measurement setup, no correction is needed in this case. At higher concentrations (two diagrams on the right) the coincidence error causes a significant difference between measured and actual distribution. It can be seen, that the correction by NN1 gives a good approximation of the actual distribution. The analytical correction of Raasch and Umhauer seems to reproduce the actual distribution slightly worse, than the neural network does. The result has even negative values in the density distribution of the peak heights. This causes local extrema on the cumulative distribution. Practically, this could be prevented by setting the values of the corresponding channels to zero. In the top/right diagram, the cumulative distribution has values above 1 for the analytical result.

The individual evaluation of hundreds of test datapoints would be time-consuming and confusing. For this reason, the performance of the two methods is compared by two parameters. Namely, the relative shift of the median voltage U_{50}, and the error integral between the correction result, and the actual cumulative distribution of the pulse heights ERR_{Q_0}. The median value is an important parameter for the description of a particle size distribution. Due to coincidences, this value is shifted towards larger particles (higher signal amplitudes). Therefore, the relative error of the median voltage is used as the first performance criterion. The parameter has two major drawbacks. If the cumulative distribution has a flat slope at the median value, a small error of the cumulative distribution might result in a large relative error, even if the correction result does match the actual distribution quite well. In addition, the corrected density distribution of the pulse heights from Raasch and Umhauer can have negative values, which might cause ambiguous median values. For such cases, the smallest median value was considered in the following evaluation. However, the second performance criterion used for the evaluation of the correction methods is the error integral of the cumulative distribution (Equation (9)).

$$ERR_{Q_0} = \frac{1}{U_{max} - U_{min}} \cdot \int_{U=U_{min}}^{U_{max}} \left| Q_{0,est}(U) - Q_{0,ref}(U) \right| \cdot dU \qquad \text{Equation 9}$$

The error integral is visualized in Fig. 6. For an optimal correction, the estimated cumulative distribution $Q_{0,est}(U)$ does match the actual cumulative distribution $Q_{0,ref}(U)$. If the curves deviate from each other, an area between the two curves is created. This area is inverse proportional to the similarity of the two distributions. A small error integral corresponds to a small difference. Note that a

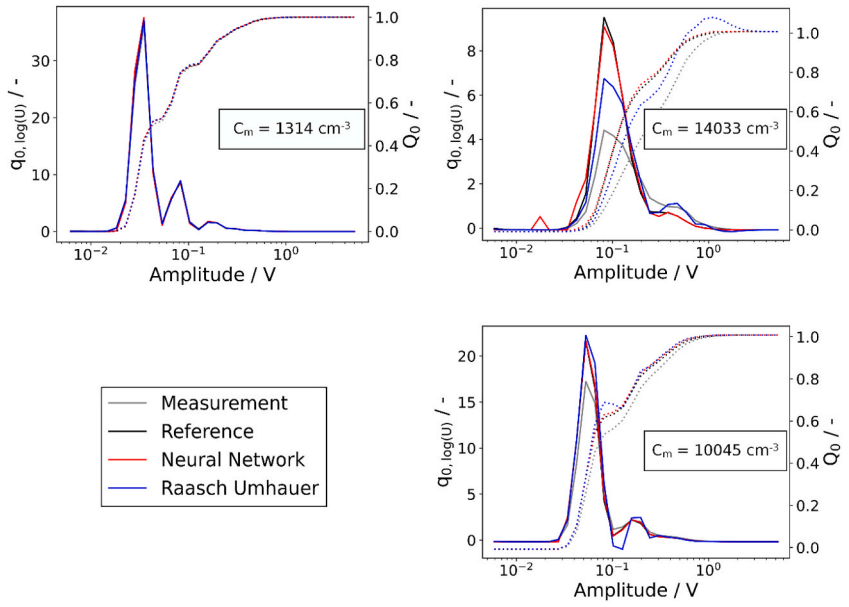

Fig. 5. Example datapoints of test dataset from device A. Coincidence correction with neural network and the method of Raasch and Umhauer. At low concentrations (top/left) all curves match each other. At higher concentrations (right side) the neural network gives the better correction result.

perfect correction, which deviates extremely at one point, can reach the same value as another correction, which deviates over the complete range, but only very slightly.

The two performance parameters were calculated for each datapoint of the test dataset. The results for the coincidence correction with NN1, as well as the analytical method of Raasch and Umhauer are shown in Figs. 7 and 8. Fig. 7 shows the relative shift of the median voltage in dependence on the aerosol concentration, while Fig. 8 shows the error integral. The black points represent the measured datapoints without coincidence correction. For better visualization, all datapoints were assigned to different logarithmic equidistant concentration classes. The grey vertical lines represent the limits of each class. The exact values of the boundaries can be found in Table 6 and Table 7. For each concentration class, the mean value and standard deviation were calculated and plotted as an error bar. The mean values are connected with lines. Note that the test dataset of device A did not contain datapoints with concentrations below 300 cm^{-3}. The third concentration class did only contain one datapoint. Thus, the shown standard deviation remains 0.

For concentrations below the coincidence limit, the shift of the median values in Fig. 7 is small, even without any correction. The differences become noticeable at concentrations above 4000 cm^{-3}. As expected, the median value shifts towards higher voltages at high concentrations. Both correction methods clearly reduce this deviation. A relative median error of 20% appears to be very high at first glance. Considering the correlation between particle size and associated signal amplitude, this corresponds to a relative error of only 2.5% for Rayleigh scattering, or 9.5% for larger particles (VDI 3867 Blatt 4:2011–16, 2011). It seems that the method of Raasch and Umhauer, in contrast to the correction by the neural network, tends to overcorrect the distribution. This can be explained by the assumption of rectangular particle pulses, which is wrong for most optical aerosol spectrometers. For example, if two Gaussian shaped particle pulses overlap each other, the resulting amplitude is significantly lower than the sum of both amplitudes, as soon as there is a small time lag between the pulses. Accordingly, the coincidence error is smaller than assumed in the model of Raasch and Umhauer and does not have to be corrected that much in reality. However, for the test dataset, both, the mean and the standard deviation of the relative median error are smaller for the neural network in comparison to the analytical approach. The correction of Raasch and Umhauer generates some significant outliers, which are caused by the effect of ambiguous, or sensitive, median values, as described above.

Fig. 8 shows the same scattering plot, but for the error integral instead of a median error. By definition, the error integral cannot be negative. For small concentrations, the error integral stays small. Below the coincidence limit, no coincidence correction is required. At higher concentrations, the error integral of the measured cumulative distribution increases significantly due to the coincidences. The analytical approach of Raasch and Umhauer shows only little advantage, compared to the raw measurement data without correction. Especially at concentrations above 20000 cm^{-3} the method seems to give unstable results. The neural network on the other hand still keeps the error integral close to zero. Thus, the presented method is well suited for the correction of measured particle distributions with coincidence error.

One major challenge of the proposed method is that the acquisition of the training dataset takes a lot of time. Even if this process is automated, the training of the neural network requires a certain level of expertise. The question arises whether a neural network, which was trained with a dataset from a specific device can be used on another device. Tolerances of the device, the measuring cell, optical components, and the signal processing chain might influence the data basis. Therefore, the same neural network was evaluated on a dataset recorded by a second device. The evaluation results for the NN1 on the dataset of device B are shown in Tables 6 and 7. The corresponding diagrams, similar to Figs. 7 and 8, can be found in Appendix 7.2. As can be seen, the correction by the neural network shows poor performance on this dataset. Especially at low concentrations, the relative median error, as well as the error integral, have much higher values than the correction by Raasch and Umhauer, or even the measurement data without any correction. Thus, the proposed correction with the neural network must not be used on a device, when the neural network was only trained on data from a single other device.

To overcome this, the training data of multiple devices of the same type can be combined. So the neural network has the possibility

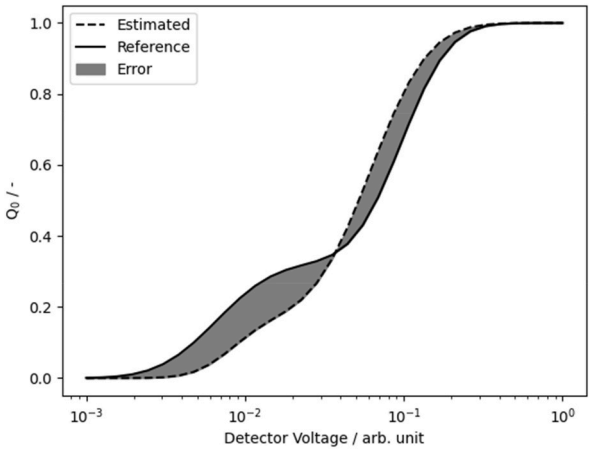

Fig. 6. Error integral. The area between corrected and reference cumulative distribution is used as a performance parameter to evaluate the quality of the correction.

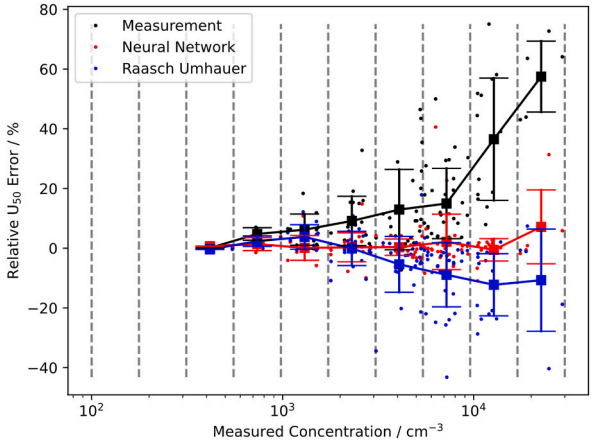

Fig. 7. Relative shift of the median voltage. NN1, test dataset from device A (Table 6, Table 7). While the method of Raasch and Umhauer tends to overcorrect the size distribution, the neural network is able to keep the shift of the median voltage low, even at high concentrations.

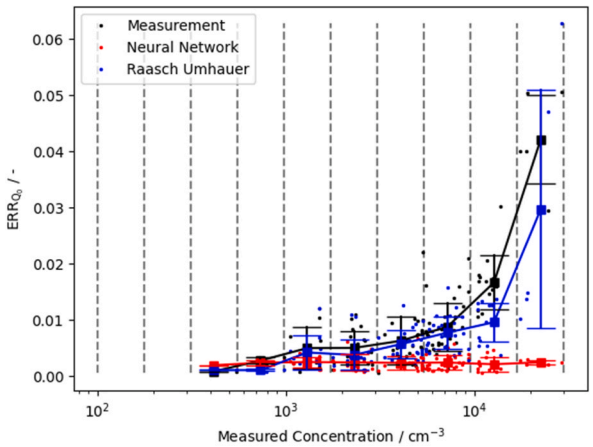

Fig. 8. Error integral of the cumulative distributions of the pulse heights. NN1, test dataset from device A (Table 6, Table 7). At high concentrations, the error integral between measured and reference distribution increases. Compared to the method of Raasch and Umhauer the neural network is able to keep the error low at high concentrations.

to see different device tolerances, which helps generalizing the coincidence correction. For this purpose, we recorded another dataset (device C), and trained a second neural network on the datasets from device A and device B (NN2). The corresponding diagrams of the results from NN2 on the dataset from device C can be found in Appendix 7.3.

As can be seen in Tables 6 and 7, the neural network correction gives better results on data from an unknown device (device C), when datasets of multiple devices (device A and device B) are used. The error integral still shows a deviation, especially at low aerosol concentrations. But for higher concentrations, the error of the median, as well as the error integral, are much smaller than the correction from Raasch and Umhauer for the second neural network. Especially for high concentrations ($c > 16959$, Class 10), the relative error of the median can be minimized from −9% for the analytical approach, down to 5% with the neural network. The error integral is reduced to almost a third, for the same concentration. Unfortunately, only three similar devices were available for this work. It can be assumed, that the coincidence correction on new devices is more precise, if datasets from more devices are used for training. In this way, only a single neural network has to be trained once. This neural network can then be used on devices of the same design, regardless of the specific device tolerances.

The proposed method for coincidence correction works well for different distributions that are similar to those included in the dataset. These may differ in particle size, concentration, mode number, and distribution width. Nevertheless, all aerosols are quite similar, since they were all generated with the same measurement setup. Due to the electrostatic classifier, a clear major peak can be found in all distributions. Multiple charged particles cause more or less pronounced secondary peaks (Cf. Fig. 5). Optical aerosol spectrometers are often used in research, and thus are also used to measure completely different distribution types. If the distribution of the aerosol differs strongly from the distributions included in the dataset, the neural network may output incorrect results. Fig. 9 shows such a case. The distribution was also recorded with the setup from Fig. 4, but the electrostatic classifier was bypassed. Thus, the broadly distributed aerosol from the generator is measured directly. As a part of the particles is smaller than the detection limit of our

Table 6
Relative median error. Values in percent. Methods: Measured (No correction, M), neural network (NN), Raasch Umhauer (RU).

N. Network	Test Data	Method	Class									
			1	2	3	4	5	6	7	8	9	10
			C_{min}/cm^{-3}									
			100	177	313	554	979	1732	3064	5420	9587	16959
			C_{max}/cm^{-3}									
			177	313	554	979	1732	3064	5420	9587	16959	30000
NN1	**A**	**M**	–	–	0 ± 0	5 ± 2	6 ± 5	9 ± 8	13 ± 13	15 ± 12	37 ± 20	57 ± 12
		NN	–	–	1 ± 0	1 ± 2	0 ± 4	0 ± 5	0 ± 3	2 ± 9	−1±4	7 ± 12
		RU	–	–	0 ± 0	2 ± 2	4 ± 4	0 ± 6	−5±9	−9±11	−12 ± 10	−11 ± 17
NN1	**B**	**M**	–	1 ± 0	0 ± 3	5 ± 5	7 ± 10	10 ± 11	11 ± 11	15 ± 13	33 ± 17	58 ± 14
		NN	–	−58 ± 2	−20 ± 18	−20 ± 24	−20 ± 22	−9±15	−10 ± 14	−8±11	−7±13	−6±25
		RU	–	0 ± 0	0 ± 4	2 ± 5	1 ± 12	0 ± 11	−4±9	−7±8	−11 ± 12	−9±14
NN2	**C**	**M**	–	–	1 ± 3	3 ± 5	5 ± 8	9 ± 12	8 ± 8	11 ± 8	26 ± 10	54 ± 16
		NN	–	–	9 ± 10	0 ± 5	0 ± 7	1 ± 9	0 ± 6	−1±7	0 ± 7	5 ± 7
		RU	–	–	−1±4	1 ± 3	2 ± 8	2 ± 11	−3±8	−7±9	−11 ± 12	−9±12

Table 7
Error integral. The shown values are multiplied by 1000 for better visualization. Methods: Measured (No correction, M), Neural network (NN), Raasch Umhauer (RU).

N. Network	Test Data	Method	Class									
			1	2	3	4	5	6	7	8	9	10
			C_{min}/cm^{-3}									
			100	177	313	554	979	1732	3064	5420	9587	16959
			C_{max}/cm^{-3}									
			177	313	554	979	1732	3064	5420	9587	16959	30000
NN1	**A**	**M**	–	–	8 ± 0	3 ± 1	5 ± 4	5 ± 3	6 ± 4	9 ± 4	2 ± 5	42 ± 8
		NN	–	–	2 ± 0	2 ± 0	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 1	2 ± 0
		RU	–	–	1 ± 0	1 ± 0	4 ± 3	4 ± 3	6 ± 3	7 ± 3	10 ± 3	30 ± 21
NN1	**B**	**M**	–	1 ± 0	2 ± 2	3 ± 3	4 ± 3	5 ± 4	6 ± 4	9 ± 4	20 ± 6	38 ± 12
		NN	–	18 ± 1	25 ± 15	14 ± 7	19 ± 12	18 ± 12	17 ± 12	18 ± 12	18 ± 13	22 ± 30
		RU	–	1 ± 2	3 ± 2	3 ± 3	4 ± 3	4 ± 3	5 ± 2	8 ± 3	10 ± 5	25 ± 14
NN2	**C**	**M**	–	–	2 ± 2	2 ± 1	5 ± 5	6 ± 9	7 ± 5	9 ± 4	18 ± 6	40 ± 14
		NN	–	–	−9±6	8 ± 4	9 ± 4	8 ± 4	8 ± 4	8 ± 4	8 ± 3	9 ± 5
		RU	–	–	3 ± 1	2 ± 2	5 ± 4	5 ± 8	6 ± 3	8 ± 4	10 ± 4	27 ± 15

measurement setup, a part of the distribution is cut off. Thus, the neural network has to handle a completely unknown type of distribution, which causes a very poor correction.

In conclusion, the proposed correction method must only be applied to distribution types, which were included in the training dataset. The misbehavior of the artificial neural network in Fig. 9 is caused by an insufficiently generalized dataset. Practically, it will not be possible, or only with extreme effort, to record a dataset that covers all aerosol scenarios. In other neural network applications it is common to extend the dataset synthetically [(Nikolenko, 2021), pp. 11–12]. This can be done in different ways. On the one hand, recorded data can be manipulated. For the proposed coincidence correction, this would mean adding noise to the raw counts of the datapoints. Further, it would be conceivable to shift the recorded distributions towards smaller or larger particles. Since the size classes of the aerosol spectrometer are logarithmic equidistant, and thus the relative distances of the channels to each other are equal for every class in the spectrum, this is easily possible. However, this technique would not change the fact that the dataset contains only the recorded distribution types. No significant improvement was achieved with an appropriately prepared dataset. Another way to extend the dataset, is to generate datapoints completely synthetically. The advantage of this approach is that the distribution shapes can be generated by combining e. g. different normal-, and log-normal distributions. The drawback is that the distribution with coincidence error has to be computed from the actual distribution for each datapoint. For this purpose, a coincidence model, such as Equation (4) or Equation (5) from Raasch and Umhauer has to be made. The dataset would include all assumptions, simplifications, and so the resulting errors. An artificial neural network trained on such a dataset would not be superior to the method of Raasch and Umhauer.

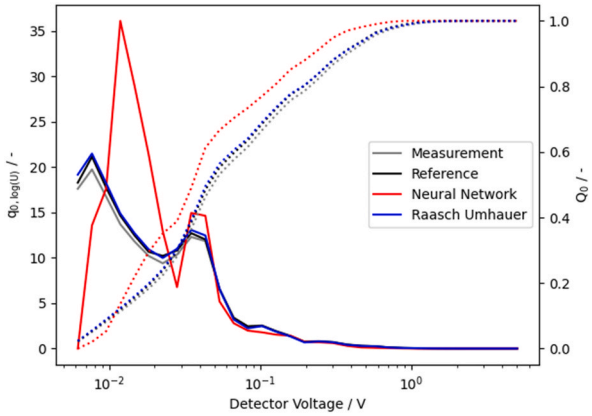

Fig. 9. Coincidence correction for an unknown distribution. If the neural network is used on a distribution type, which was not included in the training data, the correction becomes unstable. The analytical correction of Raasch and Umhauer still gives a good result.

As a result, the neural network can only be used for coincidence correction if the distribution type of the aerosol is known and was included in the training dataset. In this case the neural network gives better correction results than the analytical approach of Raasch and Umhauer, which is the only known method for this purpose. So the method proposed in this work is especially interesting for industrial applications. In process monitoring or quality management tasks, the type of the test aerosol remains constant over several years.

The correction of the measured size distribution is done independently to the correction of the measured concentration. Since the integral of a particle density distribution is 1.0 per definition (ISO 9276–1:1998–06, 1998), the density distribution does not contain information about the concentration of the aerosol. Thus, the coincidence correction method presented in this work does not improve the counting efficiency. The actual concentration can be calculated according to Equation (1) or the numerical solution provided in Appendix 7.1. However, the counting efficiency can be improved by reducing the measuring volume, or utilizing advanced peak detection algorithms, as shown in our recent paper (Oeser et al., 2022).

4. Conclusion

The use of optical aerosol spectrometers is limited by the coincidence error. At higher concentrations, the counting efficiency is reduced and the measured size distribution is shifted toward larger particles. While the counting efficiency is easy to correct, a neural network-based method for the correction of the size distribution was proposed.

The training of the artificial neural network requires a sufficiently large and well-generalized dataset that contains datapoints from aerosols with different distributions. Each datapoint must consist of a distribution measured with coincidence error, and a distribution of the same aerosol without coincidences. A measurement setup was shown that can be used to automatically record such a dataset. After training, the neural network was evaluated and compared with the analytical correction method of Raasch and Umhauer, which is the only known method for this purpose so far. The performance of both methods was evaluated by two parameters, i. e. relative median error, and the error integral of the cumulative distribution of the peak amplitudes. The neural network showed better correction results on the test dataset, especially at high concentrations. A drawback of the proposed method is that the recording of the dataset is very time-consuming. Due to device tolerances a trained neural network can only be used on another device of the same type to a limited extent. If the neural network is used to correct a particle size distribution that differs significantly from the distributions contained in the training dataset, the correction result is erroneous. Nevertheless, the method remains interesting, especially for the monitoring of industrial applications, where the type of aerosol remains constant over years. Here, the coincidence correction by an artificial neural network is an alternative to the method of Raasch and Umhauer.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Lukas Oeser, Nakul Samala, Lars Hillemann, Jan Müller, Claudia Jahn-Wolf, and Jens Lienig have no conflict of interest. Lukas Oeser, Nakul Samala, Lars Hillemann, Jan Müller and Claudia Jahn-Wolf are employees of the Topas GmbH, a company that develops, produces and markets aerosol technologies for research and industrial applications.

Data availability

The data that has been used is confidential.

Appendix

7.1 Numerical calculation of the actual aerosol concentration

The coincidence error results in a reduced counting efficiency η. The ratio between measured C_m and actual C_a concentration can be described by the following equation.

$$\eta = \frac{C_m}{C_a} = e^{-C_a \cdot V_m}$$

This equation cannot be solved for the actual concentration, but rearranged as follows.

$$0 = f(C_a) = C_a \cdot e^{-C_a \cdot V_m} - C_m$$

The actual concentration must be a zero of this function. The function is negative for small and large C_a. It has a global maximum at $C_a = 1/V_m$, which has a value of $\frac{1}{V_m \cdot e} - C_m$. Accordingly, there is only a solution if $C_m \cdot V_m < 1/e \approx 0.368$. If the measured concentration is above this value, the calculation of the actual concentration is not possible. The function has two zeros. The actual concentration must lie in the range between $C_m < C_a < 1/V_m$. The other zero of the function gives an invalid solution. The calculation of the valid zero can be done iteratively with the given python code.

```
def calc_conc(Cm, Vm, precision):
    lower = Cm
    upper = 1 / Vm

    while(upper / lower > 1 + precision):
        half = (upper + lower) / 2
        print(half)
        value = half * np.exp(-half*Vm) - Cm
        if value >= 0:
            upper = half
        else:
            lower = half

    return half
```

Fig. 7.1. Python function for the calculation of the actual aerosol concentration.

First, an interval is defined, which must contain the solution. The actual concentration will always be above the measured concentration. So the measured concentration is used as the lower limit. The upper limit is defined by the position of the maximum of the function. Then the function value is calculated for the center of the interval. For a negative value, the zero must be between the center of the interval and the upper limit. For a positive value, the zero must be between the lower limit and the center of the interval. With this information, the interval can be narrowed and the next iteration step can be performed. This process is repeated, until the solution is precise enough. If lower and upper limit are close together, i. e. the interval is smaller than the required precision, the solution (center of the interval) is returned by the function.

7.2 Results NN1, test dataset from device B

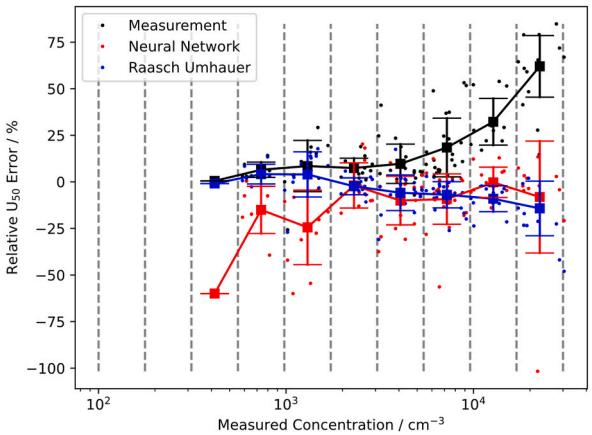

Fig. 7.2. Relative shift of the median voltage. Results for NN1 (trained on dataset from device A) on the dataset from device B..

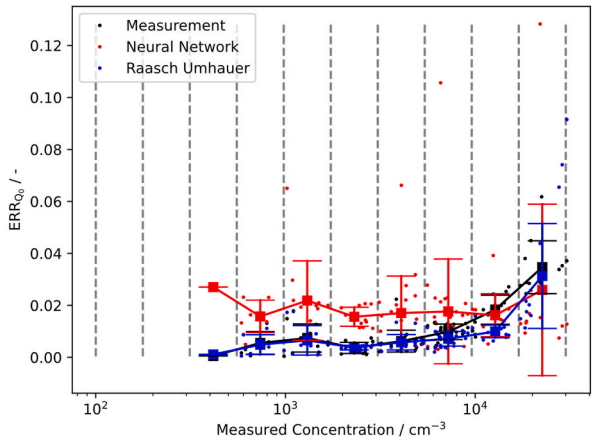

Fig. 7.3. Error integral of the cumulative distributions of the pulse heights. Results for NN1 (trained on dataset from device A) on the dataset from device B..

7.3 Results of NN2, test dataset from device C

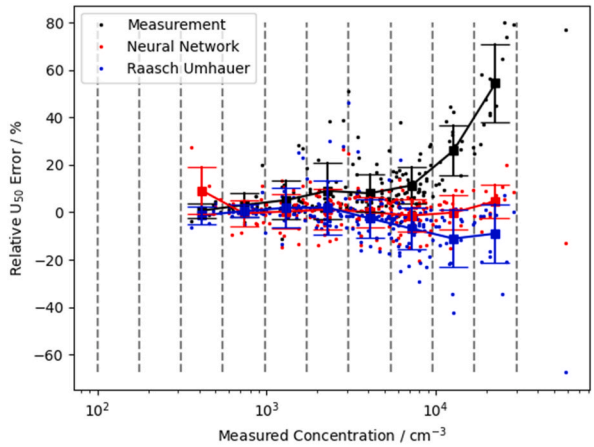

Fig. 7.4. Relative shift of the median voltage. Results for NN2 (trained on dataset from device A and device B) on the dataset from device C..

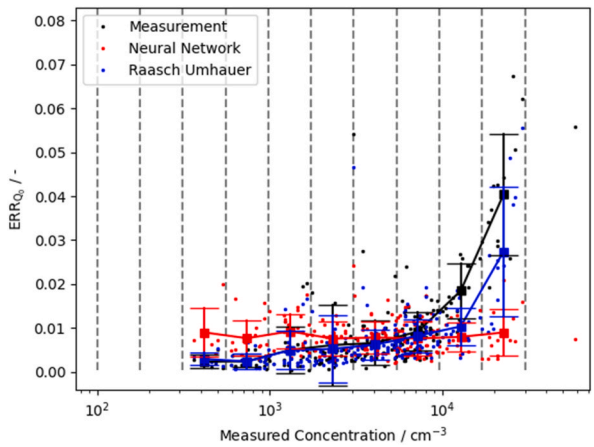

Fig. 7.5. Error integral of the cumulative distributions of the pulse heights. Results for NN2 (trained on dataset from device A and device B) on the dataset from device C..

References

Aerosol measurement: Principles, techniques, and applications. In . *Engineering professional collection* (3rd ed.), (2011)). Hoboken, N.J: Wiley https://onlinelibrary.wiley.com/doi/book/10.1002/9781118001684.

Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis, 22*(5), 717–727.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford: Oxford University Press; Clarendon Press. http://www.loc.gov/catdir/enhancements/fy0605/95040465-d.html.

Chen, D.-R., Pui, D. Y., Mulholland, G. W., & Fernandez, M. (1999). Design and testing of an aerosol/sheath inlet for high resolution measurements with a DMA. *Journal of Aerosol Science, 30*(8), 983–999.

Dasaradh, S. (2020). *An introduction to mathematics behind neural networks | towards data science*. Towards Data Science. https://towardsdatascience.com/introduction-to-math-behind-neural-networks-e8b60dbbdeba. (Accessed 29 August 2022).

Goodfellow, I., Courville, A., & Bengio, Y. (2016). *Deep learning. Adaptive computation and machine learning*. Cambridge, Massachusetts: The MIT Press. https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=2565107.

ISO 21501-1. ISO 21501-1:2009. https://www.iso.org/standard/42728.html. (Accessed 18 October 2022).

ISO 9276-1:1998-06 Representation of results of particle size analysis - Part 1: Graphical representation.

Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization*. http://arxiv.org/pdf/1412.6980v9.

Lathuiliere, S., Mesejo, P., Alameda-Pineda, X., & Horaud, R. (2020). A comprehensive analysis of deep regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 42*(9), 2065–2081. https://arxiv.org/pdf/1803.08450.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation, 1*(4), 541–551.

May, K. R. (1973). The collison nebulizer: Description, performance and application. *Journal of Aerosol Science, 4*(3), 235–243. https://www.sciencedirect.com/science/article/pii/0021850273900062.

Mie, G. (1908). Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Annals of Physics, 330*(3), 377–445.

Nikolenko, S. I. (2021). *Synthetic data for deep learning* (1st ed.). Cham: Springer International Publishing; Imprint Springer, 2021. Springer eBook Collection 174.

Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). *Activation functions: Comparison of trends in practice and research for deep learning*. https://arxiv.org/pdf/1811.03378.

Oeser, L., Samala, N., Hillemann, L., Rudolph, A., & Lienig, J. (2022). Minimizing the coincidence error in particle size spectrometers with digital signal processing techniques. *Journal of Aerosol Science, 165*, Article 106039.

Raasch, J., & Umhauer, H. (1984). *Der Koinzidenzfehler bei der Streulicht-Partikelgrössen-Zählanalyse: Fortschrittberichte der VDI-Zeitschriften ; Reihe 3*. Düsseldorf: VDI-Verlag. *Verfahrenstechnik ; 95* http://slubdd.de/katalog?TN_libero_mab214254889 ER.

Ripley, B. D. (1996). *Pattern recognition and neural networks*. Cambridge: Cambridge University Press.

Team, K. (2022). *Keras: The Python deep learning API*. https://keras.io/. (Accessed 29 August 2022).

VDI 3867 Blatt 4:2011-16. Measurement of particles in ambient air - determination of the particle number concentration and particle size distribution of aerosols - optical aerosol spectrometer.