



Gap-filling Fast Electrical Mobility Spectrometer Measurements of Particle Number Size Distributions for Eddy Covariance Application

Agnes Meyer-Kornblum*, Lars Gerling, Stephan Weber

Climatology and Environmental Meteorology, Institute of Geoecology, Technische Universität Braunschweig, 38106 Braunschweig, Germany

ABSTRACT

To estimate the spatial and temporal variation in urban particle number concentrations (PNCs), e.g., for exposure studies, a better knowledge of the exchange of particles between the urban surface and the atmosphere is important. Size-resolved fluxes of PNCs were quantified in Berlin, Germany, using the micrometeorological eddy covariance technique. The method requires measurements of particle number size distributions (PNSDs) by a fast particle spectrometer. The Engine Exhaust Particle Sizer (EEPS) Spectrometer 3090 (TSI Inc.) is designed for fast (10 Hz), high-concentration measurements of particles in the size range of 5.6–560 nm, e.g., in the exhaust plume of engines. In the urban background environment of Berlin, however, PNCs in some size channels can temporarily fall below the minimum threshold concentration of the analyser, resulting in missing concentrations that lead to gaps in the PNSD. In the present study, three gap-filling methods were applied to derive complete PNSDs: linear interpolation (LI), natural spline interpolation (NSI) and log-normal fitting (LNF). To evaluate the methods, different numbers of artificial gaps were inserted into 10^5 gapless PNSDs. Using three different data sets, the results demonstrate that LI and NSI (LI: $R^2 = 0.84$ – 0.94 ; NSI: $R^2 = 0.84$ – 0.95) outperform LNF ($R^2 = 0.78$ – 0.88). With regard to the Berlin data set, NSI is the recommended gap-filling method since it results in a lower average uncertainty of 10.5–21.8% vs. 13.3–22.9% for LI. It is advisable to reject the boundary areas of the PNSD from gap filling, i.e., $D_p < 10$ nm and $D_p > 200$ nm, as this considerably improves the gap-filling quality. This study provides recommendations on how to gap-fill incomplete PNSD data sets obtained with an EEPS.

Keywords: Micrometeorology; Engine Exhaust Particle Sizer Spectrometer; Berlin; Particle fluxes; Ultrafine.

INTRODUCTION

Epidemiological studies have indicated that particulate air pollution, especially ultrafine particles (UFPs) with a particle diameter $D_p < 100$ nm, might cause different health effects such as cardiovascular and respiratory diseases, lung cancer and increased mortality (Dockery *et al.*, 1993; Oberdörster *et al.*, 1995; Tie *et al.*, 2009; HEI Review Panel on Ultrafine Particles, 2013). Airborne concentrations of UFPs are strongly affected by the particle emission source strength, deposition, transformation processes, and atmospheric dilution and transport (e.g., Buzorius *et al.*, 2000; Birmili *et al.*, 2009; Choi *et al.*, 2012; Betha *et al.*, 2013; Meng *et al.*, 2015). For the development of measures to improve air quality and to minimize human exposure to fine and ultrafine particles, research on the temporal and spatial variation in the particle number concentration (PNC) and particle number size

distribution (PNSD) is of interest (Nel, 2005; Tsang *et al.*, 2008; Weber *et al.*, 2013; Ghassoun *et al.*, 2015; Kim *et al.*, 2015). In order to gain insight into the turbulent transport processes, which provide information on the strength of sources and sinks, measurements of particle number fluxes between surface and atmosphere using eddy covariance (EC) are a suitable tool (e.g., Buzorius *et al.*, 1998; Schmidt and Klemm, 2008; Ahlm *et al.*, 2009; Damay *et al.*, 2009; Ripamonti *et al.*, 2013; von der Heyden *et al.*, 2018). EC is a direct method to capture the turbulent exchange of a quantity, i.e., the turbulent flux density, between surface and atmosphere (Baldocchi, 2008).

A couple of studies on turbulent particle fluxes above different surfaces have been conducted in recent years. In natural ecosystems with tall vegetation such as forests, EC applications studying the vertical exchange and deposition velocities of particles are well established (e.g., Buzorius *et al.*, 1998; Buzorius *et al.*, 2000; Held and Klemm, 2006; Held *et al.*, 2006; Pryor *et al.*, 2008a, b; Damay *et al.*, 2009; Rannik *et al.*, 2009; Petroff *et al.*, 2018). The studies show that forest vegetation mainly acts as a sink of particles and that simultaneous emission and deposition fluxes of particles with different diameters may appear. Additionally, Held *et al.*

* Corresponding author.

Tel.: 0531/391-5910; Fax: 0531/391-5617

E-mail address: a.meyer-kornblum@tu-braunschweig.de

(2006) demonstrated that particle number fluxes of deposition events are dominated by UFPs.

The first urban study was conducted in Edinburgh (Dorsey *et al.*, 2002). Subsequently, particle number flux measurements followed in other European cities such as Stockholm (Mårtensson *et al.*, 2006; Vogt *et al.*, 2011a, b; Vogt *et al.*, 2013), Manchester (Longley *et al.*, 2004; Martin *et al.*, 2009), Gothenburg (Martin *et al.*, 2009), Helsinki (Järvi *et al.*, 2009; Ripamonti *et al.*, 2013), Münster (Schmidt and Klemm, 2008; Dahlkötter *et al.*, 2010; Deventer *et al.*, 2013; Deventer *et al.*, 2015), Lecce (Contini *et al.*, 2012), London (Martin *et al.*, 2009; Harrison *et al.*, 2012) and Innsbruck (Deventer *et al.*, 2018; von der Heyden *et al.*, 2018). Cities predominantly show net emission fluxes for the total particle number with a clear diurnal cycle that correlates with traffic activity and atmospheric stability (Deventer *et al.*, 2018). The investigation of size-resolved particle number fluxes indicate that UFPs contribute significantly to the total particle number fluxes (Schmidt and Klemm, 2008; Harrison *et al.*, 2012; Deventer *et al.*, 2018) and that simultaneous bi-directional fluxes, i.e., upward and downward fluxes, in different size bins are possible (Deventer *et al.*, 2013; Deventer *et al.*, 2015; Deventer *et al.*, 2018). However, the number of existing studies focusing on size-resolved and ultrafine particle number fluxes is still limited (Nemitz *et al.*, 2000; Deventer *et al.*, 2018).

The measurement of turbulent flux densities using EC requires fast data acquisition (≥ 10 Hz) of vertical wind fluctuations and a scalar quantity (Burba and Anderson, 2005), i.e., the PNSD in terms of turbulent size-resolved particle number fluxes. However, only a few spectrometers for fast PNSD measurements in the UFP size range are commercially available. The electrical mobility spectrometer used in the present study (Engine Exhaust Particle Sizer (EEPS) Spectrometer 3090; TSI Inc., Minnesota, USA), is one of the few which meets these requirements. With a measurement frequency of 10 Hz and a response time of about 0.5 s (Johnson *et al.*, 2003, 2004), most of the turbulent fluctuations contributing to the particle flux should be resolved by the instrument. However, the response time is limited in the sense that some contribution to the turbulent flux might be underestimated. This can be corrected during the flux calculations (e.g., Mårtensson *et al.*, 2006), which, however, is not within the scope of the present study. The device enables the simultaneous measurement of PNSD in the size range $5.6 < D_p < 560$ nm in 32 size bins. The instrument samples at a flow rate of 10 L min^{-1} , operates at ambient pressure to prevent evaporation of semi-volatile and volatile particles and was primarily developed for measurements of engine exhaust fumes (TSI Inc., 2016). Thus, the EEPS is optimized for particle concentrations higher than those usually observed in ambient conditions, e.g., in the urban background. As a result, in low-concentration or urban background conditions, the readings of specific size channels of a PNSD might temporarily fall below the analyser's minimum threshold. In this case, the specific channel concentration cannot be used for turbulent flux calculation. The missing concentration in this size channel is defined as a 'gap'. In particular, if the gap in the PNSD is recurring frequently within the time period used for block averaging in the EC method, i.e. 30 min (Aubinet

et al., 2012), it is not possible to calculate a flux for the size channel. To avoid rejecting the entire PNSD in case of gaps and to avoid large data gaps within the block-averaging interval, a gap-filling method is required to replace concentrations in missing size channels.

The EC method is characterized by the fact that parts of the data are repeatedly discarded due to quality checks or missing data, e.g., due to blocking (rain or fog) of open optical sensor paths (Burba and Anderson, 2005). As a result, the data availability in EC measurements can be as low as 40% (Aubinet *et al.*, 2012). In ecosystem studies focusing on the CO_2 -based carbon exchange between surface and atmosphere, continuous time series of CO_2 fluxes are necessary to calculate the annual carbon balance, i.e., the net ecosystem exchange (NEE; Falge *et al.*, 2001; Moffat *et al.*, 2007). Hence, gap-filling procedures are applied to calculate the missing fluxes, e.g., from linear regression approaches. This approach serves as a 'role model' in the present study. However, gap filling in this work is applied to the 10 Hz PNSD measurements in order to deduce complete PNSDs for the calculation of turbulent size-resolved particle number fluxes.

The motivation of this study is to test and assess three different gap-filling methods: a linear interpolation (LI), a natural spline interpolation (NSI) and a log-normal fitting (LNF). The latter method is based on an iterative procedure to fit three modes as log-normal distribution functions to different particle sizes of the PNSD. We hypothesise that the three gap-filling methods will perform differently in different size ranges of the PNSD and that the choice of the best gap-filling method may depend on the specific data set.

MATERIALS AND METHODS

Study Site and Instrumentation

The EC measurement site is located in the west of Berlin (Charlottenburg) near Ernst-Reuter-Platz on the rooftop of the main building of Technische Universität Berlin. The rooftop is at a height of 47 m above ground level. The building is located next to a busy 6-lane road (Straße des 17. Juni) with an average daily traffic intensity of between 3×10^4 and 4×10^4 vehicles day^{-1} (Umweltatlas Berlin, 2014). The site is part of a measurement network that was set up in the framework of the [UC]²-3DO project, which intends to collect an extensive data set for validation of the building-resolving urban climate LES model PALM-4U (Scherer *et al.*, 2019). The measurements in Berlin started in March 2017. In the present study, data of the one-year period from 15 March 2017 until 14 March 2018 is used.

The EC-system consists of a 3D ultrasonic anemometer (USA-1; Metek GmbH, Elmshorn, Germany) and a fast electrical mobility spectrometer (Engine Exhaust Particle Sizer Spectrometer 3090; TSI Inc., Minnesota, USA). The EEPS detects PNSDs in the diameter size range $5.6 \text{ nm} < D_p < 560 \text{ nm}$ in 32 size channels. Particles entering the EEPS sample inlet receive a positive electrical charge induced by a corona charger. The charge level is dependent on particle size. Subsequently, particles are transported through an electric field, where they are reflected outward according to

their electrical mobility. Particle charge is measured by 22 electrometers and converted into concentrations (TSI Inc., 2015).

Both the EEPS and the 3D ultrasonic anemometer sample data at a frequency of 10 Hz. For the measurement of PNSD, ambient air is drawn through a 9.1 m long stainless steel tube with an inner diameter of 10.7 mm. The sample line is located next to a 10 m tall mast resulting in an inlet height at 57 m above ground level. Inside a weatherproof housing installed at the bottom of the mast, the sample air is drawn through a Nafion dryer (0.9 m, MD-700; Perma Pure LLC) to keep the relative humidity of the sample < 40%. The total distance of the sample line from inlet to electrometers amounts to 10.75 m. With a flow rate of 10 L min⁻¹, this results in laminar flow conditions inside the sampling line (Reynolds number \approx 1300; Hinds, 1999). Although the use of turbulent flows within the sampling line has been discussed (Buzorius *et al.*, 1998), a larger number of (recent) studies have used laminar flow conditions for particle number flux measurements (Buzorius *et al.*, 1998; Deventer *et al.*, 2013, 2015, 2018). For a three-week period in July 2017, the ambient EEPS measurements were compared to a water-based condensation particle counter on site (WCPC 3787; TSI Inc., Minnesota, USA). Both concentration measurements were in very good agreement (data not shown here).

Gap-filling Strategy

If the particle concentration in a specific size channel fell below the minimum threshold concentration indicated by the EEPS firmware, the channel was rejected from the PNSD and needed to be filled by a gap-filling algorithm. To quantify the quality of the three gap-filling methods (LI, NSI and LNF), the complete PNSDs, i.e., without any gaps, were selected. A total of 2.7 million valid PNSDs were available, in which the concentration in every size channel was above

the minimum threshold concentration as defined by the EEPS firmware. We verified that the subset gives a representative sample of PNSDs in terms of daytime and season. Thereof, a randomly selected sample of 10⁵ PNSDs was chosen as the reference for method comparison, i.e., the ‘gold standard’ (Fig. 1).

For the calculation of the gap-filling uncertainty, the concentrations from gap-filled PNSDs were compared to the gold standard. For that purpose, artificial gaps were inserted into the gold standard to have a verifiable basis for gap filling (cf. Moffat *et al.*, 2007). Three different types (A, B and C) of artificial gaps were implemented to study the effects of gap occurrence. Different numbers of artificial gaps per data type and PNSD were inserted to calculate the uncertainty in the case of increasing gap numbers, i.e., 1–5 gaps, 6–10 gaps, 11–15 gaps and 16–20 gaps (Fig. 1).

As we defined a size channel with a missing concentration as a ‘gap’, we defined adjacent size channels with missing concentrations as ‘contiguous gaps’. To give an example, 4 missing concentrations in adjacent size channels were defined as 4 contiguous gaps (cf. Fig. 2). In Type A data sets, a maximum of 4 contiguous gaps were allowed. This was based on the manufacturers’ recommendation that interpolation of missing channels is possible if the number of contiguous gaps is < 5. However, in the Type B and C data sets, up to 5 contiguous gaps were allowed since contiguous gaps > 4 occurred in the Berlin data set. The artificial gaps were distributed randomly within the PNSDs of Data Types A and B. In Type C, the total number of gaps per PNSD and the likelihood of a certain size channel to be missing were based on the probabilities as detected in the data set measured in Berlin (Fig. 3). As evident from Fig. 3(a), the majority of gaps were located at the boundaries of the size spectrum, i.e., in the size range $D_p < 30$ nm and > 200 nm. In addition, Fig. 3(b) shows the specific number of gaps per PNSD.

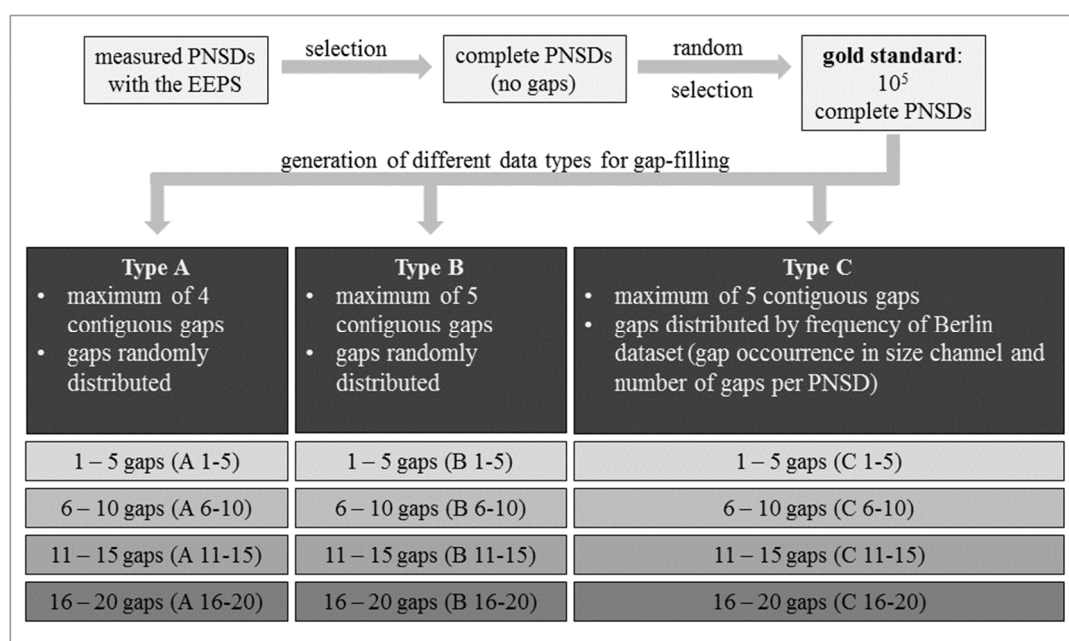


Fig. 1. Workflow for data handling in the gap-filling method comparison.

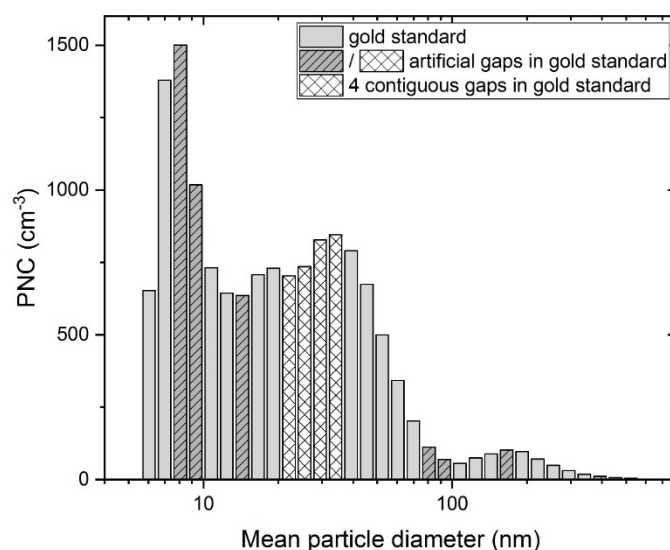


Fig. 2. Illustration of an example PNSD of the gold standard with artificial gaps. In the size range $20 \text{ nm} < D_p < 40 \text{ nm}$, 4 contiguous gaps were inserted.

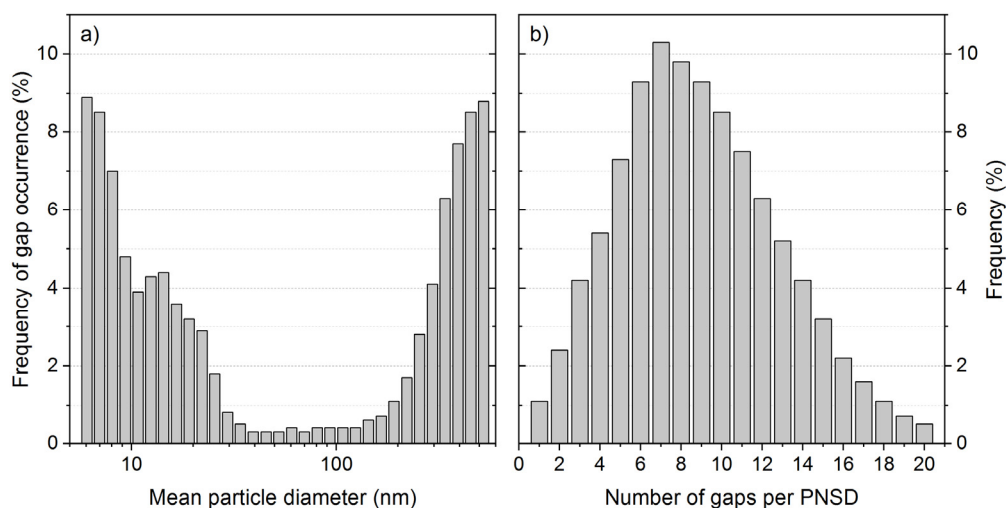


Fig. 3. Frequency distributions of (a) gap occurrence per size bin and (b) total number of gaps per PNSD of the data set measured in Berlin during the study period from 15 March 2017 to 14 March 2018. These frequencies were used to create the data sets of Type C.

Assuming that the gap-filling quality declines with a decreasing number of available data per PNSD, a maximum number of gaps to be allowed per PNSD was defined. This limit has been set to 20 gaps. Hence, concentration data for a minimum of 12 size channels per PNSD was always available. To evaluate whether the gold standard would be transferable to the remaining non-gold standard PNSDs, a brief analysis was conducted on a reduced data set with 2000 PNSDs. Hence, log-normal fittings were approximated to gap-afflicted PNSDs of both types (gold standard and non-gold standard). Log-normal approximations are often used in the modelling of aerosols and clouds, as well as in statistical or cluster analysis of PNSDs (Heintzenberg, 1994; Birmili *et al.*, 2001; Hussein *et al.*, 2005; von Salzen, 2006; Wegner *et al.*, 2012; Faghihi *et al.*, 2016; Ueda *et al.*, 2016). Assuming that gold standard and non-gold standard PNSDs behave physically

similarly or follow the same statistical distribution, we expect that the log-normal fitting will work equally well for both types of PNSDs. The analysis showed that despite lower concentrations of the non-gold standard PNSDs, the log-normal functions can be fitted with the same R^2 (0.96) and a similar RMSE (1323 cm^{-3} and 1181 cm^{-3} for gold standard and non-gold standard, respectively) compared to the gold standard PNSDs (data not shown here). Hence, we assume that the results of the gap-filling analysis are also applicable to the non-gold standard PNSDs.

Gap-filling Methods

To fill artificial gaps in the data sets, three methods were applied:

Linear interpolation (LI): The concentrations of adjacent upper and lower size channels were utilized to calculate the

concentration of the gap by fitting a linear function of the form:

$$PNC = a \times D_p + b \quad (1)$$

where $a, b \in \mathbf{R}$ are parameters: a is the gain and b , the offset. To set boundary conditions, the concentration at particle diameters of 0 nm and 600 nm were defined as 0 cm^{-3} .

Natural spline interpolation (NSI): A polynomial natural spline is defined as a function $s: [x_0, x_n] \rightarrow \mathbf{R}$ of odd degree $l = 2m - 1$, $m \geq 2$, which satisfies the following conditions (Abbasbandy and Babolian, 1998):

- $s \in C^{l-1}[x_0, x_n]$
- $s(x)$ is a polynomial of degree l for $x \in [x_i, x_{i+1}]$ $i = 0, 1, \dots, n-1$
- $s^{(v)}(x_0) = s^{(v)}(x_n) = 0$; $v = m, \dots, 2m-2$

with $n+1$ data points x_0, \dots, x_n . In the present study, natural cubic splines are used, thus $l = 3$ and $v = 2$. The 32 size channels of the PNSD are the data points ($n = 31$), and the calculated spline output is the PNC for the respective size channel. For size bins in which the spline reached values < 0 , the associated concentrations were set to 0 cm^{-3} .

Log-normal fitting (LNF): The LNF involves nonlinear fitting of the parameters of a sum of log-normal distribution functions (Hussein et al., 2005):

$$f(D_p, \bar{D}_{pg,i}, N_i, \sigma_{g,i}) = \sum_{i=1}^n \frac{N_i}{\sqrt{2\pi} \log(\sigma_{g,i})} \times \exp \left[-\frac{[\log(D_p) - \log(\bar{D}_{pg,i})]^2}{2 \log^2(\sigma_{g,i})} \right] \quad (2)$$

In Eq. (2), D_p is the diameter of an aerosol particle, i characterizes the number of the log-normal mode (in this

study $n = 3$), N_i is the mode number concentration, $\sigma_{g,i}^2$ is the geometric variance, and $\bar{D}_{pg,i}$ is the geometric mean diameter. As each mode is defined by three parameters, a total of nine parameters were fitted to each PNSD. The trust-region-reflective method was used to perform the least squares parameter fit (Coleman and Li, 1994, 1996). In addition, log transformation was performed in the least squares fit. In order to improve the flexibility of the model fit, overlapping of the modes was allowed. This means that $\bar{D}_{pg,1}$ ranged between 3 and 20 nm, $\bar{D}_{pg,2}$ between 15 and 80 nm and $\bar{D}_{pg,3}$ between 30 and 300 nm. Fig. 4 illustrates the concept of gap filling and the differences between the three methods for an example PNSD.

Performance of Gap-filling Methods

The gap-filling performance was investigated for six particle size ranges to gather information about different parts of the PNSD (Table 1). In every size range, the gap-filled concentrations were evaluated by comparison to the gold standard using the coefficient of determination (R^2) and the root mean square error (RMSE) as performance indicators. Since the absolute RMSE is not suitable for directly comparing the different size ranges due to varying concentrations, a relative metric for comparison was calculated, i.e., the quotient of RMSE and a normalised number concentration (NNC) of each size range (RMSE/NNC). This metric can be interpreted as a mean percentage error. The NNCs are defined as the annual mean concentrations in the respective size ranges normalised by the number of size channels in that size range (Table 1).

The focus of method evaluation is on Data Set C since it builds the basis for subsequent particle flux calculation with the Berlin data set in future research. Thus, the figures in the following section mainly focus on Type C data. The remaining figures for Data Sets A and B are shown in the supplementary material to this article.

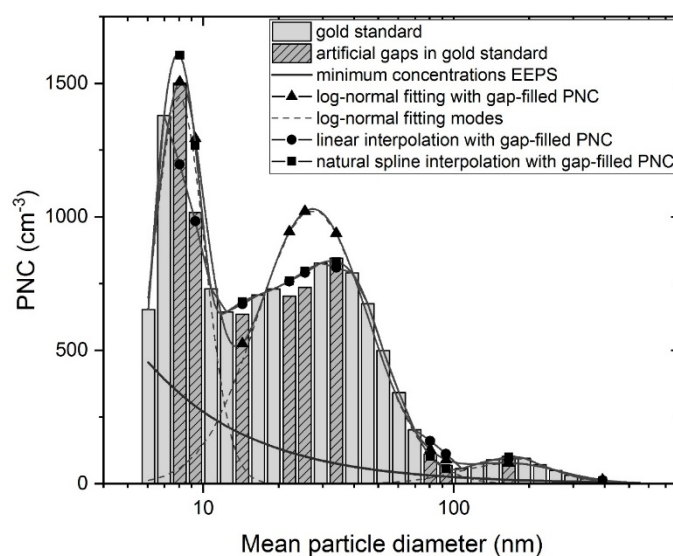


Fig. 4. Illustration of the method comparison for linear interpolation, natural spline interpolation and iterative log-normal fitting by means of an example PNSD. The various symbols illustrate the differences between the three gap-filling methods LI, NSI and LNF. Dashed bars indicate artificial gaps which were integrated into the gold standard.

Table 1. Normalised number concentrations (NNCs) based on annual mean concentrations in the different size ranges normalised by the number of size channels in this size range.

Range of particle size	NNC
5.6–10 nm	908.57 cm ⁻³
10–20 nm	675.04 cm ⁻³
20–50 nm	583.43 cm ⁻³
50–100 nm	498.11 cm ⁻³
100–200 nm	233.71 cm ⁻³
200–560 nm	29.97 cm ⁻³

RESULTS AND DISCUSSION

Observed Particle Number Concentrations

The measurement site is characterised by an average PNC ($10 \text{ nm} < D_p < 200 \text{ nm}$) of $8,300 \text{ cm}^{-3}$ (median = $7,300 \text{ cm}^{-3}$) for the one-year study period. The mean (median) PNCs of particles in the size ranges 10–30 nm, 30–100 nm and 100–200 nm are $3,900 \text{ cm}^{-3}$ ($3,100 \text{ cm}^{-3}$), $3,500 \text{ cm}^{-3}$ ($3,000 \text{ cm}^{-3}$) and 800 cm^{-3} (700 cm^{-3}), respectively. The half-hourly means are characterised by a maximum of $43,100 \text{ cm}^{-3}$ and a minimum of $3,300 \text{ cm}^{-3}$.

Gap-filling Method Comparison

The performance indicators show decreasing gap-filling

quality with an increasing number of gaps per PNSD for the different types of data (Fig. 5). This behaviour was expected, as less information is available for gap filling with an increasing number of gaps. Generally, the iterative LNF performs worse than the two methods LI and NSI. In the case of a small number of gaps, NSI is the most convincing method for each data set. However, with increasing gap number, LI works better. It is evident that quality indicators worsen from Data Set A to B and C (Figs. 5 and 6). Thus, the mean gap-filling quality declines with an increasing number of contiguous gaps and with more frequently occurring gaps in the boundary areas of the PNSD such as in Data Set C. The latter leads to a gap-filling quality decline with increasing gap number primarily in these regions (Fig. 7). In particular, R^2 shows relatively low values for small particle sizes ($< 10 \text{ nm}$). This might be due to high temporal variation in small particles and larger concentration differences between adjacent channels. The methods NSI and LI are characterized by higher R^2 values in the middle parts of the PNSD ($20 \text{ nm} < D_p < 200 \text{ nm}$) than at the edges. This might be due to missing information on adjacent channels next to the boundary size bins of the PNSD. By contrast, in the LNF, quality enhances with increasing particle size, presumably because an assumption of the log transformation in the least squares fit is that relative errors are associated with a constant standard deviation. Therefore, based on R^2 , the fit is more accurate for size bins with lower

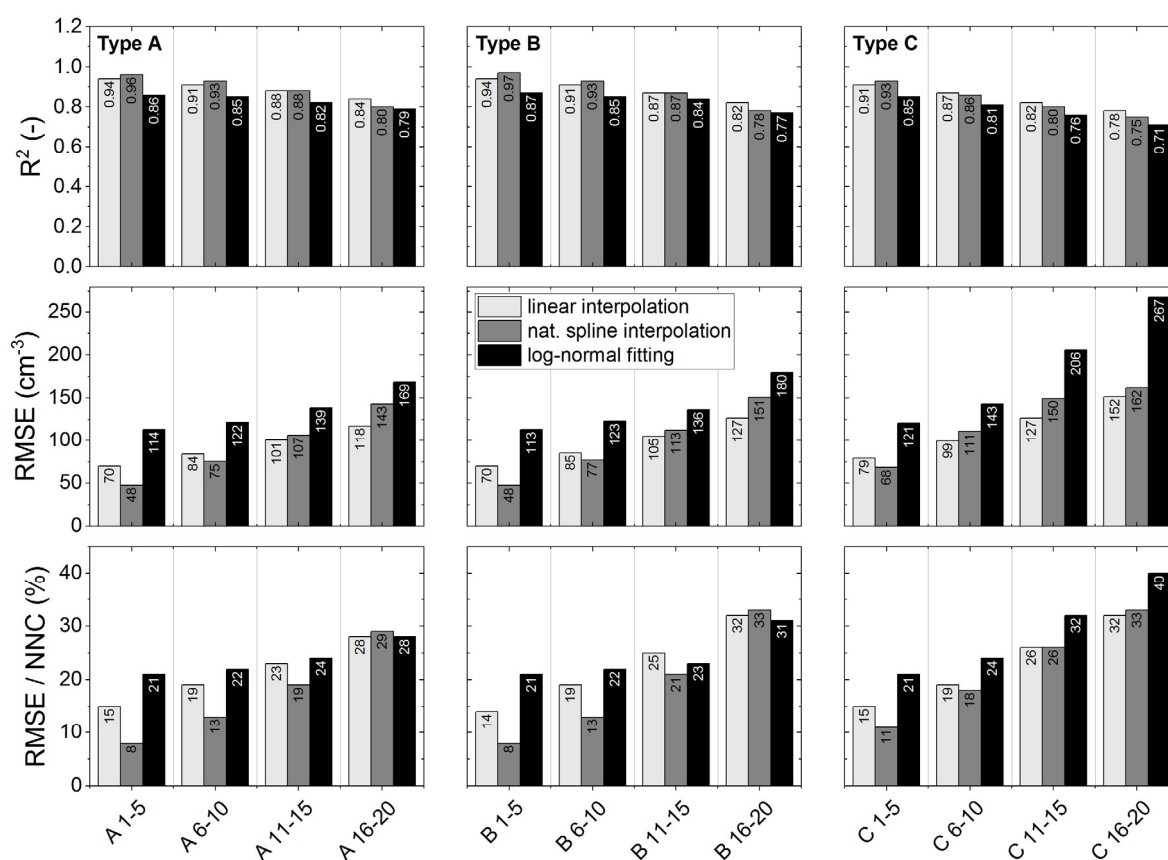


Fig. 5. Coefficient of determination (R^2), root mean square error (RMSE) and mean percentage error (RMSE/NNC) for all three interpolation methods, subdivided according to the data sets of Types A, B and C. The values are averages of the six size ranges of particle diameter.

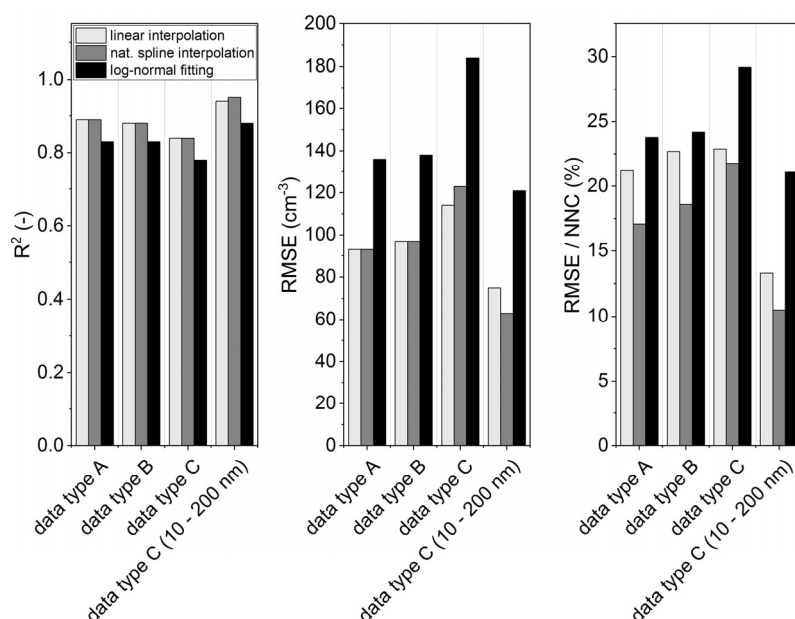


Fig. 6. Quality indicators (R^2 , RMSE and RMSE/NNC) for Type A, B and C data sets. Additionally, the analysis of data Type C (10–200 nm) is shown.

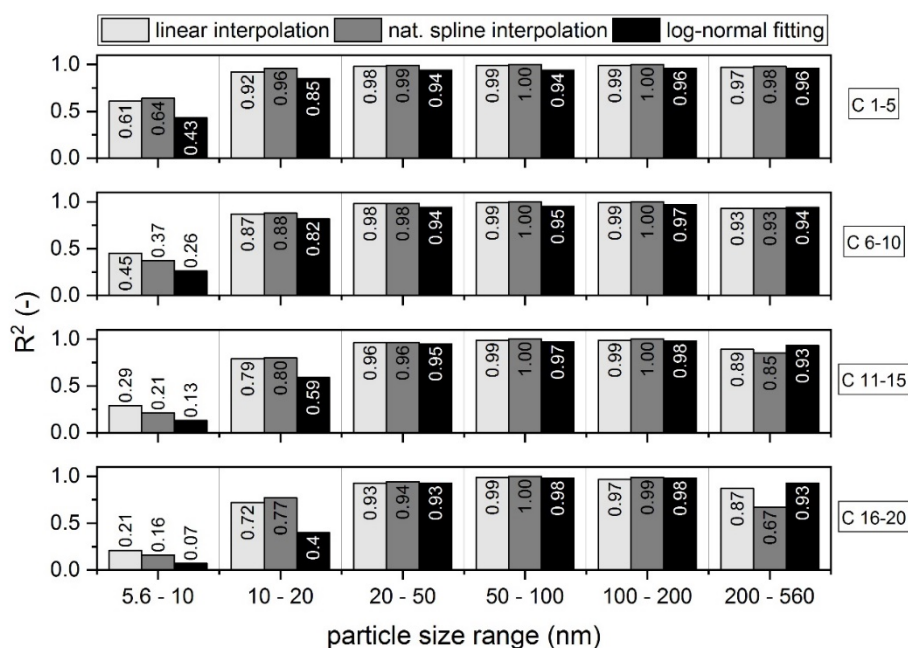


Fig. 7. Coefficient of determination (R^2) for the three interpolation methods and the data sets of Type C, subdivided according to the six ranges of particle diameter.

PNCs, as is the case for the largest size range (Table 1), than for size bins with higher PNCs. NSI reaches its highest R^2 for $10 \text{ nm} < D_p < 200 \text{ nm}$, while LI is best suited for the smallest and LNF for the largest size ranges of the PNSD (Fig. 7).

By contrast, the size-dependent behaviour is not typical for Data Sets A and B (Figs. S1 and S2). Instead, a transition region is apparent, in which the boundary between NSI and LI as the preferred method shifts towards larger D_p with an increasing number of gaps. However, for all data sets, LNF

can outperform the other methods only in the size range $200 \text{ nm} < D_p < 560 \text{ nm}$. In the middle size range of the PNSD ($20 \text{ nm} < D_p < 200 \text{ nm}$), all methods perform almost equally well. Considering NSI, the R^2 for this size range is > 0.93 , whereas the mean percentage error is between 2.0% (C 1–5, 50–100 nm) and 17.2% (C 16–20, 20–50 nm; Fig. S8). The size range 10–20 nm represents the transition between the boundary area $< 10 \text{ nm}$ and the middle part of the PNSD. In this size range (10–20 nm), gaps occur more often than in the middle part of the PNSD. Hence, the

maximum R^2 is $0.77 < R^2 < 0.96$ (NSI), whereas the mean error is between 13.2% and 31.6%.

These size ranges with low R^2 are characterized by high RMSE and an increased mean percentage error (Figs. S5 and S8). For instance, in the case of 16–20 gaps per PNSD in the Type C data set, the mean percentage error in the smallest and largest size channels is $> 60\%$. Hence, it may be advisable to discard the boundary parts of the PNSD from gap filling to reduce the maximum mean error of PNC to $< 32\%$ (NSI, 10–20 nm). The mean percentage error for the whole PNSD then decreases from 21.8% to 10.5% (NSI; Fig. 6). If the size range is limited to $20 \text{ nm} < D_p < 200 \text{ nm}$, the maximum mean error can be reduced to $< 17\%$ (NSI, 20–50 nm; Fig. S8).

Generally, the two methods LI and NSI perform well in the gap-filling procedure, while LNF shows larger deviations. LNF is a widely used tool for various applications regarding PNSDs (Heintzenberg, 1994; Birmili et al., 2001; Hussein et al., 2005; von Salzen, 2006; Wegner et al., 2012; Faghghi et al., 2016; Ueda et al., 2016). Thus, we expected that LNF would be one of the preferred gap-filling methods. On the contrary, LNF is only convincing in the size range $> 200 \text{ nm}$ compared to the two other methods. One important point is that LNF is used for gapless data sets generally (Heintzenberg, 1994; Birmili et al., 2001; Hussein et al., 2005; Faghghi et al., 2016). Due to gaps within a PNSD, the relation between the number of data points and the number of parameters used in the fitting weakens. Hence, the log-normal modes often do not correspond to the true course of the distribution, resulting in over- or underestimation of PNCs. Furthermore, the LNF algorithm always fits three modes to the PNSDs, whereas distributions can have only one or two modes. This may impede the determinability of the function parameters and even impede the fit in PNSDs with gaps due to overfitting. The log-normal fitting algorithm of Hussein et al. (2005)

was evaluated with data from different environments (boreal forest, polar region and urban area) and conditions (low and high pollution). However, in the urban data set, distinct differences between modal PNCs and measured concentrations were visible due to the higher temporal variation from urban emissions, e.g., traffic (Hussein et al., 2005). The PNSDs measured in Berlin are influenced by traffic activity in the area of Ernst-Reuter-Platz. Consequently, this could also be a reason for the poorer performance of the LNF compared to the two other methods. Additionally, computing time for LNF is much longer in comparison to LI and NSI due to the iterative procedure. This is of interest for large data sets, such as the 10 Hz data from Berlin.

Fig. 8 summarises the best methods for all data sets and size ranges. The RMSE is not shown since the pattern is similar to RMSE/NNC. In the case of R^2 , two methods are defined as equally good if R^2 is the same up to the second decimal place. For the percentage error, we define both methods as equally good when the difference is smaller than one percentage point. It is evident that in the size range $50 \text{ nm} < D_p < 200 \text{ nm}$, only NSI is advisable for all data sets. For smaller size channels ($< 50 \text{ nm}$), NSI and LI perform nearly equally well. LNF is applicable for particles $> 20 \text{ nm}$ but outperforms the other methods only in the size range $> 200 \text{ nm}$. With regard to the calculation of size-resolved particle fluxes for the Berlin data set in a future study, NSI is recommended for gap filling when the smallest ($< 10 \text{ nm}$) and largest ($> 200 \text{ nm}$) size ranges of the PNSD are excluded.

SUMMARY AND CONCLUSION

In this study, three gap-filling methods were investigated and compared using PNSDs measured with a fast electrical mobility spectrometer at an urban rooftop site in Berlin.

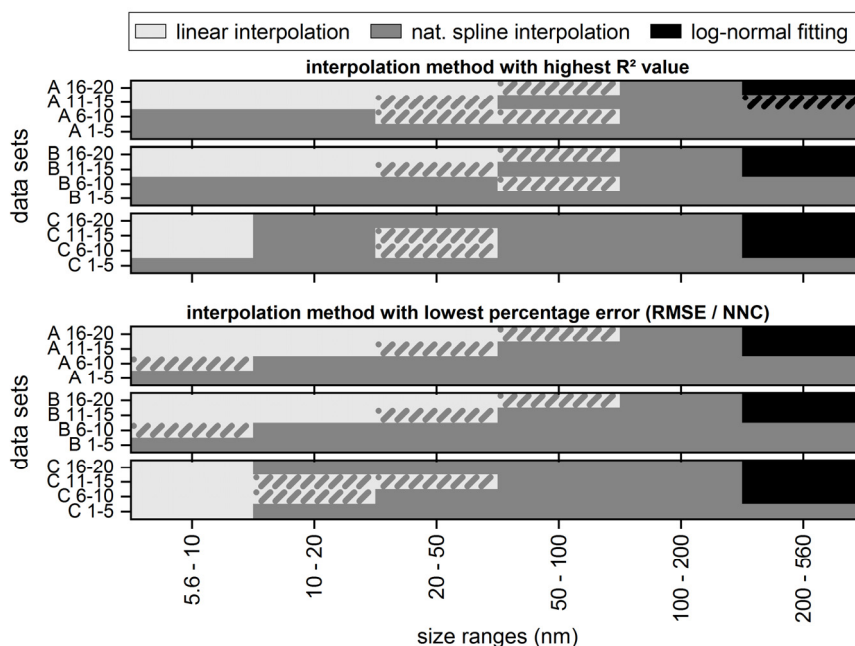


Fig. 8. Best gap-filling methods with regard to R^2 and RMSE/NNC for all data sets in the six defined particle size ranges of the PNSD. The striped pattern signalizes that two methods perform equally well.

Three types of data sets, each with a different number of gaps per PNSD, were generated from a gold standard. The gap-filled data sets were compared using different quality indicators. Based on our results, gap filling across a wide size range of the PNSD is best achieved by either NSI or LI. NSI performs particularly well in the middle range of the size distribution ($50 \text{ nm} < D_p < 200 \text{ nm}$) and with few gaps per PNSD, whereas LI improves gap filling for the smaller size ranges and as the number of gaps increases. For the current EEPS data from Berlin, in which the gaps are non-randomly distributed and frequently occur in the boundary regions of the PNSD, we recommend using NSI and discarding the smallest and largest size channels. Additionally, PNSDs with > 5 contiguous gaps or > 20 total gaps should be discarded.

Generally, the number of gaps should be lower when the EEPS is used in environments with higher particle concentrations. However, the present study offers recommendations on achieving complete PNSDs with moderate uncertainty when gaps in data sets do occur.

ACKNOWLEDGEMENTS

This study was supported by the German Federal Ministry of Education and Research (BMBF) under Grant FKZ 01 LP 1602 D (Urban Climate under Change, Module 3DO). The authors furthermore thank the project partners of the TU Berlin, Institute for Ecology, Chair of Climatology, for being able to attach the measuring systems to their existing measuring site.

SUPPLEMENTARY MATERIAL

Supplementary data associated with this article can be found in the online version at <http://www.aaqr.org>.

REFERENCES

- Abbasbandy, S. and Babolian, E. (1998). Interpolation of fuzzy data by natural splines. *Korean J. Comput. Appl. Math.* 5: 457–463.
- Ahlm, L., Nilsson, E.D., Krejci, R., Mårtensson, E.M., Vogt, M. and Artaxo, P. (2009). Aerosol number fluxes over the Amazon rain forest during the wet season. *Atmos. Chem. Phys.* 9: 9381–9400.
- Aubinet, M., Vesala, T. and Papale, D. (2012). *Eddy covariance: A practical guide to measurement and data analysis*. Springer Netherlands, Dordrecht.
- Baldocchi, D. (2008). 'Breathing' of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems. *Aust. J. Bot.* 56: 1–26.
- Birmili, W., Wiedensohler, A., Heintzenberg, J. and Lehmann, K. (2001). Atmospheric particle number size distribution in central Europe: Statistical relations to air masses and meteorology. *J. Geophys. Res.* 106: 32005–32018.
- Burba, G. and Anderson, D. (2005). *A brief practical Guide to eddy covariance flux measurements: Principles and workflow examples for scientific and industrial applications*. LI-COR, Inc.
- Buzorius, G., Rannik, Ü., Mäkelä, J. M., Vesala, T. and Kulmala, M. (1998). Vertical aerosol particle fluxes measured by eddy covariance technique using condensational particle counter. *J. Aerosol Sci.* 29: 157–171.
- Buzorius, G., Rannik, Ü., Mäkelä, J.M., Keronen, P., Vesala, T. and Kulmala, M. (2000). Vertical aerosol fluxes measured by the eddy covariance method and deposition of nucleation mode particles above a Scots pine forest in southern Finland. *J. Geophys. Res.* 105: 19905–19916.
- Coleman, T.F. and Li, Y. (1994). On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. *Math. Program.* 67: 189–224.
- Coleman, T.F. and Li, Y. (1996). An interior trust region approach for nonlinear minimization subject to bounds. *SIAM J. Optim.* 6: 418–445.
- Contini, D., Donato, A., Elefante, C. and Grasso, F.M. (2012). Analysis of particles and carbon dioxide concentrations and fluxes in an urban area. Correlation with traffic rate and local micrometeorology. *Atmos. Environ.* 46: 25–35.
- Dahlkötter, F., Griessbaum, F., Schmidt, A. and Klemm, O. (2010). Direct measurement of CO₂ and particle emissions from an urban area. *Meteorol. Z.* 19: 565–575.
- Damay, P.E., Maro, D., Coppalle, A., Lamaud, E., Connan, O., Hébert, D., Talbaut, M. and Irvine, M. (2009). Size-resolved eddy covariance measurements of fine particle vertical fluxes. *J. Aerosol Sci.* 40: 1050–1058.
- Deventer, M.J., Griessbaum, F. and Klemm, O. (2013). Size-resolved flux measurement of sub-micrometer particles over an urban area. *Meteorol. Z.* 22: 729–737.
- Deventer, M.J., El-Madany, T., Griessbaum, F. and Klemm, O. (2015). One-year measurement of size-resolved particle fluxes in an urban area. *Tellus B* 67: 25531.
- Deventer, M.J., Heyden, L. von der, Lamprecht, C., Graus, M., Karl, T. and Held, A. (2018). Aerosol particles during the Innsbruck Air Quality Study (INNAQS). Fluxes of nucleation to accumulation mode particles in relation to selective urban tracers. *Atmos. Environ.* 190: 376–388.
- Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G. and Speizer, F.E. (1993). An association between air pollution and mortality in six U.S. cities. *N. Engl. J. Med.* 329: 1753–1759.
- Dorsey, J.R., Nemitz, E., Gallagher, M.W., Fowler, D., Williams, P.I., Bower, K.N. and Beswick, K.M. (2002). Direct measurements and parameterisation of aerosol flux, concentration and emission velocity above a city. *Atmos. Environ.* 36: 791–800.
- Faghihi, E.M., Martin, D., Salimi, F., Edwards, G., He, C. and Morawska, L. (2016). What are the driving factors influencing the size distribution of airborne synthetic clay particles emitted from a jet milling process? *Aerosol Air Qual. Res.* 16: 25–35.
- Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.O., Katul, G., Keronen, P., Kowalski, A., Lai, C.T., Law, B.E., Meyers, T., Moncrieff, J., Moors, E., Munger, J.W., Pilegaard, K., Rannik, Ü., Rebmann, C.,

- Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K. and Wofsy, S. (2001). Gap filling strategies for defensible annual sums of net ecosystem exchange. *Agric. For. Meteorol.* 107: 43–69.
- Ghassoun, Y., Ruths, M., Löwner, M.O. and Weber, S. (2015). Intra-urban variation of ultrafine particles as evaluated by process related land use and pollutant driven regression modelling. *Sci. Total Environ.* 536: 150–160.
- Harrison, R.M., Dall'Osto, M., Beddows, D.C.S., Thorpe, A.J., Bloss, W.J., Allan, J.D., Coe, H., Dorsey, J.R., Gallagher, M., Martin, C., Whitehead, J., Williams, P.I., Jones, R.L., Langridge, J.M., Benton, A.K., Ball, S.M., Langford, B., Hewitt, C.N., Davison, B., Martin, D., Petersson, K.F., Henshaw, S.J., White, I.R., Shallcross, D.E., Barlow, J.F., Dunbar, T., Davies, F., Nemitz, E., Phillips, G.J., Helfter, C., Di Marco, C.F. and Smith, S. (2012). Atmospheric chemistry and physics in the atmosphere of a developed megacity (London). An overview of the REPARTEE experiment and its conclusions. *Atmos. Chem. Phys.* 12: 3065–3114.
- HEI Review Panel on Ultrafine Particles (2013). Understanding the Health Effects of Ambient Ultrafine Particles. HEI Perspectives 3. Health Effects Institute, Boston, MA.
- Heintzenberg, J. (1994). Properties of the log-normal particle size distribution. *Aerosol Sci. Technol.* 21: 46–48.
- Held, A. and Klemm, O. (2006). Direct measurement of turbulent particle exchange with a twin CPC eddy covariance system. *Atmos. Environ.* 40: 92–102.
- Held, A., Nowak, A., Wiedensohler, A. and Klemm, O. (2006). Field measurements and size-resolved model simulations of turbulent particle transport to a forest canopy. *J. Aerosol Sci.* 37: 786–798.
- Hinds, W.C. (1999). *Aerosol technology: Properties, behavior, and measurement of airborne particles*. 2. ed., Wiley, New York.
- Hussein, T., Dal Maso, M., Petäjä, T., Koponen, I.K., Paatero, P., Aalto, P.P., Hameri, K. and Kulmala, M. (2005). Evaluation of an automatic algorithm for fitting the particle number size distributions. *Boreal Environ. Res.* 10: 337–355.
- Järvi, L., Rannik, Ü., Mammarella, I., Sogachev, A., Aalto, P.P., Keronen, P., Siivola, E., Kulmala, M. and Vesala, T. (2009). Annual particle flux observations over a heterogeneous urban area. *Atmos. Chem. Phys.* 9: 7847–7856.
- Johnson, T., Caldow, R., Pöcher, A., Mirme, A. and Kittelson, D.B. (2003). An engine exhaust particle sizer spectrometer for transient emission particle measurements. 9th Diesel Engine Emissions Reduction (DEER) Workshop 2003, 24–28 August 2003, Newport, RI (US).
- Johnson, T., Caldow, R., Pöcher, A., Mirme, A. and Kittelson, D. (2004). A new electrical mobility particle sizer spectrometer for engine exhaust particle measurements. SAE Technical Paper Series 2004-01-1341 (SAE-SP, 1801).
- Kim, K.H., Woo, D., Lee, S.B. and Bae, G.N. (2015). On-road measurements of ultrafine particles and associated air pollutants in a densely populated area of Seoul, Korea. *Aerosol Air Qual. Res.* 15: 142–153.
- Longley, I.D., Gallagher, M.W., Dorsey, J.R. and Flynn, M. (2004). A case-study of fine particle concentrations and fluxes measured in a busy street canyon in Manchester, UK. *Atmos. Environ.* 38: 3595–3603.
- Mårtensson, E.M., Nilsson, E.D., Buzorius, G. and Johansson, C. (2006). Eddy covariance measurements and parameterisation of traffic related particle emissions in an urban environment. *Atmos. Chem. Phys.* 6: 769–785.
- Martin, C.L., Longley, I.D., Dorsey, J.R., Thomas, R.M., Gallagher, M.W. and Nemitz, E. (2009). Ultrafine particle fluxes above four major European cities. *Atmos. Environ.* 43: 4714–4721.
- Moffat, A.M., Papale, D., Reichstein, M., Hollinger, D.Y., Richardson, A.D., Barr, A.G., Beckstein, C., Braswell, B.H., Churkina, G., Desai, A.R., Falge, E., Gove, J.H., Heimann, M., Hui, D., Jarvis, A.J., Kattge, J., Noormets, A. and Stauch, V.J. (2007). Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agric. For. Meteorol.* 147: 209–232.
- Nel, A. (2005). Atmosphere. Air pollution-related illness. Effects of particles. *Science* 308: 804–806.
- Nemitz, E., Fowler, D., Dorsey, J.R., Theobald, M.R., McDonald, A.D., Bower, K.N., Beswick, K.M., Williams, P.I. and Gallagher, M.W. (2000). Direct measurements of size-segregated particle fluxes above a city. *J. Aerosol Sci.* 31: 116–117.
- Oberdörster, G., Gelein, R. M., Ferin, J. and Weiss, B. (1995). Association of particulate air pollution and acute mortality. Involvement of ultrafine particles? *Inhalation Toxicol.* 7 111–124.
- Petroff, A., Murphy, J.G., Thomas, S.C. and Geddes, J.A. (2018). Size-resolved aerosol fluxes above a temperate broadleaf forest. *Atmos. Environ.* 190: 359–375.
- Pryor, S.C., Barthelmie, R.J., Sørensen, L.L., Larsen, S.E., Sempreviva, A.M., Grönholm, T., Rannik, Ü., Kulmala, M. and Vesala, T. (2008a). Upward fluxes of particles over forests. When, where, why? *Tellus B* 60: 372–380.
- Pryor, S.C., Gallagher, M., Sievering, H., Larsen, S.E., Barthelmie, R.J., Birsan, F., Nemitz, E., Rinne, J., Kulmala, M., Grönholm, T., Taipale, R. and Vesala, T. (2008b). A review of measurement and modelling results of particle atmosphere–surface exchange. *Tellus B* 60: 42–75.
- Rannik, Ü., Mammarella, I., Aalto, P., Keronen, P., Vesala, T. and Kulmala, M. (2009). Long-term aerosol particle flux observations part I. Uncertainties and time-average statistics. *Atmos. Environ.* 43: 3431–3439.
- Ripamonti, G., Järvi, L., Mølgaard, B., Hussein, T., Nordbo, A. and Hämeri, K. (2013). The effect of local sources on aerosol particle number size distribution, concentrations and fluxes in Helsinki, Finland. *Tellus B* 65: 19786.
- Scherer, D., Ament, F., Emeis, S., Fehrenbach, U., Leidl, B., Scherber, K., Schneider, C. and Vogt, U. (2019). Three-dimensional observation of atmospheric processes in cities. *Meteorol. Z.* 28: 121–138.
- Schmidt, A. and Klemm, O. (2008). Direct determination of highly size-resolved turbulent particle fluxes with the disjunct eddy covariance method and a 12 – stage electrical low pressure impactor. *Atmos. Chem. Phys.* 8: 7405–7417.

- Tie, X., Wu, D. and Brasseur, G. (2009). Lung cancer mortality and exposure to atmospheric aerosol particles in Guangzhou, China. *Atmos. Environ.* 43: 2375–2377.
- Tsang, H., Kwok, R. and Miguel, A.H. (2008). Pedestrian exposure to ultrafine particles in Hong Kong under heavy traffic conditions. *Aerosol Air Qual. Res.* 8: 19–27.
- TSI Inc. (2015). Engine Exhaust Particle Sizer Spectrometer Model 3090/3090AK. Operation and Service Manual.
- TSI Inc. (2016). Engine Exhaust Particle Sizer Spectrometer Model 3090. Model 3090 EEPS Spectrometer Spec Sheet A4. www.tsi.com/getmedia/a01ec52e-f39f-4312-8f9a-fded71631a69/3090_Engine_Exhaust_Particle_Sizer_A4_2980351_WEB?ext=.pdf.
- Ueda, S., Miura, K., Kawata, R., Furutani, H., Uematsu, M., Omori, Y. and Tanimoto, H. (2016). Number-size distribution of aerosol particles and new particle formation events in tropical and subtropical Pacific Oceans. *Atmos. Environ.* 142: 324–339.
- Umweltatlas Berlin (2014). Verkehrsmengen 2014. https://fbinter.stadt-berlin.de/fb/index.jsp?loginkey=zoomStart&mapId=wmsk_07_01verkmeng2014@senstadt&bbox=385330,5818907,387245,5819832, Last Access: 14 March 2019.
- Vogt, M., Nilsson, E.D., Ahlm, L., Mårtensson, E.M., Johansson, C. (2011a). Seasonal and diurnal cycles of 0.25–2.5 μm aerosol fluxes over urban Stockholm, Sweden. *Tellus B* 63: 935–951.
- Vogt, M., Nilsson, E.D., Ahlm, L., Mårtensson, E.M. and Johansson, C. (2011b). The relationship between 0.25–2.5 μm aerosol and CO_2 emissions over a city. *Atmos. Chem. Phys.* 11: 4851–4859.
- Vogt, M., Johansson, C., Mårtensson, M., Struthers, H., Ahlm, L. and Nilsson, D. (2013). Heated submicron particle fluxes using an optical particle counter in urban environment. *Atmos. Chem. Phys.* 13: 3087–3096.
- von der Heyden, L., Deventer, M. J., Graus, M., Karl, T., Lamprecht, C. and Held, A. (2018). Aerosol particles during the Innsbruck Air Quality Study (INNAQS). The impact of transient fluxes on total aerosol number exchange. *Atmos. Environ.* 190: 389–400.
- von Salzen, K. (2006). Piecewise log-normal approximation of size distributions for aerosol modelling. *Atmos. Chem. Phys.* 6: 1351–1372.
- Weber, S., Kordowski, K. and Kuttler, W. (2013). Variability of particle number concentration and particle size dynamics in an urban street canyon under different meteorological conditions. *Sci. Total Environ.* 449: 102–114.
- Wegner, T., Hussein, T., Hämeri, K., Vesala, T., Kulmala, M. and Weber, S. (2012). Properties of aerosol signature size distributions in the urban environment as derived by cluster analysis. *Atmos. Environ.* 61: 350–360.

Received for review, June 7, 2019

Revised, September 27, 2019

Accepted, October 13, 2019