Estimating Exponential Concentration Decay Rates Using Time-Integrated Aerosol Sampling of Variable Duration

Simon Toby Parker[1]*, Zaheer Ahmad Nasir[2], Luiza Cintra Campos[2]

[1] *Dstl, Porton Down, Salisbury, Wiltshire SP4 0JQ, United Kingdom*
[2] *Healthy Infrastructure Research Centre, Department of Civil, Environmental and Geomatic Engineering, University College London, Gower Street, London, United Kingdom*

ABSTRACT

The effect of variable sampling duration on the determination of exponential decay rates using integrating samplers has been explored. The theoretical basis is developed and two practical methods are presented for processing sampled data, an iterative linear least squares approach and a non-linear least squares approach. These methods are compared to a simple, non-iterative, linear least squares approach that neglects the variation in sample duration, for idealised and experimental data. The degree of error introduced by the simple method, using three different reference times, is determined for a range of sampling parameters. Errors in the estimated decay rate were seen to depend on the product of the true decay rate and the initial sample duration, and also the ratio of sample duration to the previous sample. In addition, the error was affected by intervals in the sampling sequence when present. Errors as large as 80% were observed for the parameter range studied. Despite the introduction of large errors in the decay rate, neglecting variation in sample time did not lead to strong deviation from linearity in curve fitting and may not be apparent to the unwary.

Keywords: Integrated sampling; Exponential decay; Non-linear curve fitting; Aerosol sampling; Variable duration.

INTRODUCTION

The measurement of airborne aerosol concentrations can now be achieved with near real-time instrumentation for many applications (Fierz *et al.*, 2002; Agranovski *et al.*, 2003; Kimmel *et al.*, 2011; Jung and Lee, 2013). However, collection using single or staged impactors still plays a key role in aerosol science (Park *et al.*, 2009), whether for collecting a physical sample for subsequent analysis or to act as a reference method. These methods rely upon collecting a sample over a period of time and can be considered as integrating samplers. One application area where impaction is frequently used is the characterisation of bio-aerosol concentrations (Li *et al.*, 2009) and, in particular, the measurement of their rate of decay within a chamber.

In order to calculate the concentration decay rate in such circumstances it is necessary to know the change of concentration with time. When aerosol decay is governed by an exponential decay, the concentration can rapidly decay. In some circumstances there may be benefit from using longer sampling periods at later times to increase the number

of particles collected. However, the integrating nature of such samplers introduces additional considerations into the subsequent analysis of the data. This short paper explores the analysis of aerosol data for the measurement of exponential decay rates, sampled using integrating samplers and the implications of varying sampling periods within the same experiment. Following consideration of the underlying theory, calculations are presented for idealised data and for an example dataset. A practical method is presented for the calculation of decay rates from integrated data using variable sampling times. The authors are not aware of any previous work on this problem in the field of aerosol science.

THEORY

Fixed Sampling Durations

We assume that the airborne concentration being measured is subject to a first-order decay with a rate, λ [s^{-1}], which is unknown and is to be determined by the experiment. The equation describing its variation with time is therefore

$$C(t) = C_0 e^{-\lambda t} \tag{1}$$

where C is the concentration [particles/m^3] at time t and C_0 is the initial concentration.

A sampler collecting a sample of air at a fixed flow rate, Q [m^3/s] with a constant efficiency, ε, between times t_s and

---

* Corresponding author.
 Tel.: 44-1980-658954; Fax: 44-1980-658500
 E-mail address: stparker@dstl.gov.uk

$t_s + \Delta t$ will collect N particles, where N is given by

$$N = \int_{t_s}^{t_s + \Delta t} Q\varepsilon C(t)dt, \qquad (2)$$

t_s is the start of the sampling period, and Δt is the sampling duration.

Substituting (1) for $C(t)$, integrating and taking the natural logarithm gives

$$\ln(N) = \ln\left(\frac{Q\varepsilon C_0}{\lambda}\left(1 - e^{-\lambda\Delta t}\right)\right) - \lambda t_s. \qquad (3)$$

Considering first the case when the sampling duration, Δt, is constant, this expression can be compared with the equation for a straight line, $y = mx + c$. It can be seen that plotting $\ln(N)$ against t_s will result in a line with gradient, $m = -\lambda$. All of the parameters in the first term on the right hand side of the equation have been assumed to be constant and contribute to the intercept. In fact it can also be shown that plotting $\ln(N)$ against $t_s + \Delta t$ will also yield a line with the same gradient, although a different intercept. A simple linear least squares fitting of the data in this form will therefore lead to a value for λ.

It is common practice to plot the natural logarithm of the mean concentration measured at each time against time. The mean concentration for a sample can be calculated as

$$\overline{C} = \frac{N}{\varepsilon Q \Delta t} \qquad (4)$$

Substituting this into (3) and rearranging gives

$$\ln(\overline{C}) = \ln(C_0) + \ln\left(\frac{1 - e^{-\lambda\Delta t}}{\lambda\Delta t}\right) - \lambda t_s. \qquad (5)$$

This equation shows that plotting $\ln(\overline{C})$ against t_s will also result in a line with gradient, $m = -\lambda$, when Δt is constant. It is interesting to note that the intercept is given by the first two terms on the right hand side and will not be equal to $\ln(C_0)$. However, the value of C_0 can be calculated once λ has been determined since Δt is known. Also, using the series definition of the exponential function and disregarding higher order terms, it can be shown that

$$\lim_{\lambda\Delta t \to 0} e^{-\lambda\Delta t} = 1 - \lambda\Delta t. \qquad (6)$$

Therefore, when $\lambda\Delta t$ is small Eq. (5) becomes

$$\lim_{\lambda\Delta t \to 0} \ln(\overline{C}) = \ln(C_0) - \lambda t_s \qquad (7)$$

and the intercept approaches $\ln(C_0)$. For example, when $\lambda\Delta t = 0.1$ the error in C_0 is equal to –4.8%.

Variable Sampling Durations

In the case of variable sample duration we revisit (3)

and separate the variable term from the constant term and make explicit the dependence of Δt on t_s as follows:

$$\ln(N) = \ln\left(\frac{Q\varepsilon C_0}{\lambda}\right) + \ln\left(1 - e^{-\lambda\Delta t(t_s)}\right) - \lambda t_s. \qquad (8)$$

In this expression the first term on the right hand side is constant, but now both the second and third terms vary with t_s, since Δt is variable. The values of Δt are known, but there is now a non-linear relationship between N and t_s.

One approach for calculating λ is to proceed iteratively. Rearranging Eq. (8) gives

$$\ln(N) - \ln\left(1 - e^{-\lambda\Delta t(t_s)}\right) = \ln\left(\frac{Q\varepsilon C_0}{\lambda}\right) - \lambda t_s. \qquad (9)$$

If an initial estimate is made for λ then the left hand side of (9) can be plotted against t_s as before and a least squares regression performed. The resulting value for λ can be used to update the estimate. The process can then be repeated until the change in the value of λ is below a given tolerance. Such a procedure can be easily performed in a spreadsheet environment.

Alternatively, a non-linear least squares method can be employed. For example, the Levenberg-Marquardt approach (Levenberg, 1944; Marquardt, 1963) is one algorithm that is widely implemented for non-linear curve fitting in numerical software (SciPy Community, 2012).

EXAMPLE DATA

Idealised Data

To demonstrate the influence of sampling duration we examine idealised data and the effect of integrated sampling. We compare the effect of variation in sampling duration using a simple non-iterative linear least squares approach and the correct non-linear approach.

We consider idealised decay data sampled with a sequence of five samples with increasing sample duration. The sampling sequence is defined by an initial sample duration, Δt_0, a factor, r, by which the sample duration increases, where $r = \Delta t_{i+1}/\Delta t_i$, and an interval, s, between the end of one sample and the start of the next. Real data may be affected by experimental noise. However, the nature of the noise will depend on the specific measurement method used and has not been considered in this study in order to keep the results as general as possible.

Fig. 1 shows the case where $\lambda = 0.1667$ min^{-1} (10 h^{-1}), $\Delta t_0 = 1$ min, $r = 2.0$ and $s = 1.0$ min. The graph shows how the instantaneous concentration and the average concentration, $\overline{C}$, for each sample, varies with time. It is clear that the average concentration underestimates the instantaneous concentration at the start time (t_s) and overestimates it at the end ($t_s + \Delta t$). Lines have been included that join the concentrations at the start (t_s), mean ($t_s + \Delta t/2$) and the end time of each sample ($t_s + \Delta t$). Whilst the instantaneous concentration follows a straight line, the connected sample points do not. The gradients of the line sections for the t_s
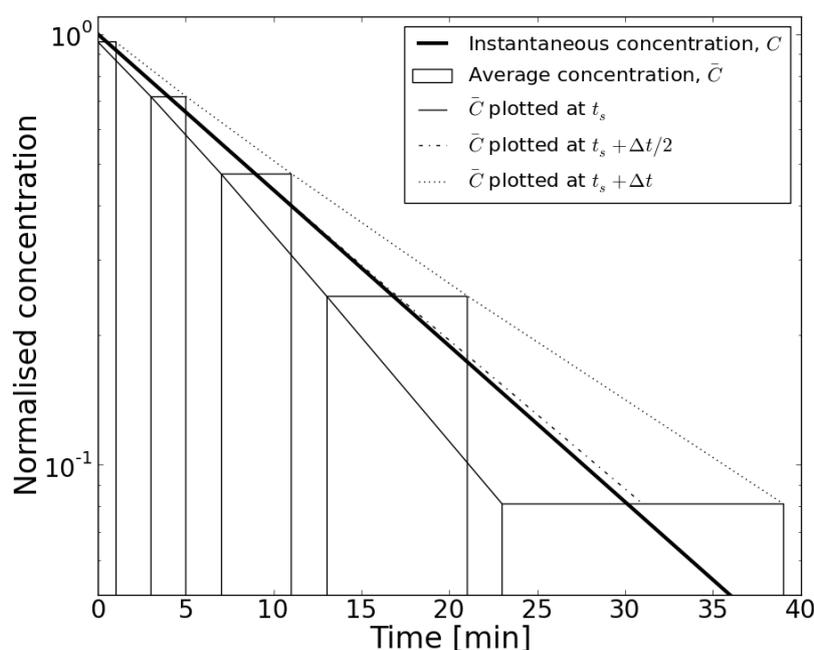
Fig. 1. Instantaneous concentration (thick line) and average concentration (bars) measured by an integrated sampler and the same data plotted using start, mean and end times. Concentration normalised by initial concentration and plotted on a logarithmic vertical scale.

and $t_s + \Delta t$ cases are considerably different to that of the instantaneous concentration line. This indicates that the use of uncorrected data in these cases would result in a considerable error in the decay rate estimate.

A linear least squares fit to the logarithm of the average concentration and each of the reference times based on Eq. (5) has been performed using the NumPy v1.6.2 (Oliphant, 2007) Python[TM] v2.7.3 (Beazley, 2009) library. A non-linear fit was also performed using the Levenberg-Marquardt method implemented in the SciPy v0.11.0 (SciPy Community, 2012) Python[TM] library. The logarithm of N was fitted to the right hand side of Eq. (3), with both λ and C_0 to be determined.

The resulting line fits and the decay rates are shown in Fig. 2. It can be seen that the non-linear method returns the true decay rate (0.167 min[-1]) whilst the linear fits either underestimate (–27%, –6.4%) or overestimate (31%) the rate depending on the reference time used. Although the linear least squares lines are not a perfect fit to the data, the deviations are not so large that they will necessarily indicate that the analysis approach is invalid. There is therefore a danger that the decay rate calculated by one of these methods may be used without recognition of the error introduced.

The case examined shows that, if not correctly accounted for, the use of variable sample duration can have a large effect on the derived decay rate. To examine a broad range of cases, the sampling parameters were varied over a wide range of values shown in Table 1. The length of the initial sample duration was related to the decay rate. A range of values for the dimensionless parameter $\lambda \Delta t_0$ was used to define these values. The number of samples was held constant at five. A total of 840 cases were examined and the decay rates calculated for each of the four methods.

The results showed that the non-linear method always

returned the correct decay rate for the idealised data as expected. The resulting decay rates determined from the linear least squares approach are shown in Table 2. For the case when $s = 0$ min the results were seen to depend only on $\lambda \Delta t_0$ and r. This can be shown to be consistent with analysis of the error in the last sample based on Eq. (5). Increasing r results in greater deviation from the true decay rate for $t_s + \Delta t/2$ and $t_s + \Delta t$. This is also true for t_s for values of $r \leq 1.5$, but the relationship becomes more complex for larger r and higher values of $\lambda \Delta t_0$. Larger values of $\lambda \Delta t_0$ for a given r decrease the calculated decay rate. This increases the error for $t_s + \Delta t/2$ and $t_s + \Delta t$, whilst it gets closer to the true value for t_s for the parameter range considered here. The example case considered above suggested that using the $t_s + \Delta t/2$ time yielded the least error. This wider analysis indicates that this is not always the case, particularly when $\lambda \Delta t_0$ is large.

The effect of varying the sampling interval is shown by the data in Table 3 for the case when $r = 2.0$. This data can be compared to the case when $s = 0$ min in Table 2. Increasing the interval between samples improves the estimate of the decay rate, especially for smaller values of $\lambda \Delta t_0$. However, no consideration has been given to the sensitivity of the sample analysis technique and an increased sample interval may result in lower concentrations adversely affecting the analysis.

Experimental Data

The importance of accounting for variable sample duration is illustrated with a small example experimental data set. Bio-aerosol concentrations were measured within a small chamber containing a portable air cleaning device following the dissemination of a test aerosol. The air within the chamber
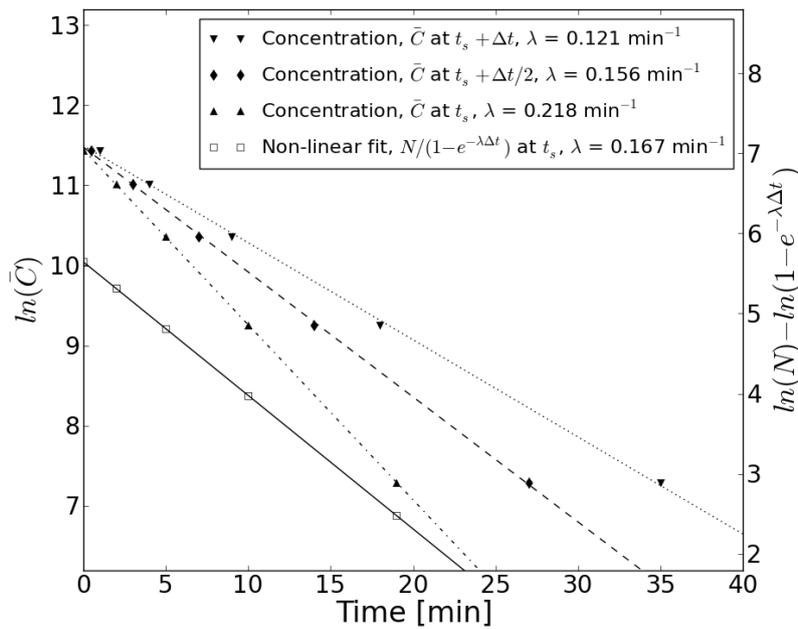
Fig. 2. Comparison of estimated decay rates using the linear least squares approach with start, mean and end times, and the non-linear Levenberg-Marquardt curve fitting. Initial sample duration 1 min, geometric increase in sample duration 2 and interval between samples 1 min.

Table 1. Sampling parameter ranges.

Decay rate	[h^{-1}]	1.0	2.0	5.0	10	20	
	[min^{-1}]	0.0167	0.0333	0.0833	0.167	0.333	
$\lambda\Delta t_0$	[-]	0.0333	0.100	0.333	1.00	3.33	
r	[-]	1.0	1.2	1.5	2.0	3.0	5.0
s	[min]	0.0	1.0	3.0	5.0		

Table 2. Predicted decay rate ratio (λ/λ_{true}) for six different sample duration ratios (r) and dependence on the product of decay rate and initial sample duration ($\lambda\Delta t_0$) using three reference times (t_s, $t_s + \Delta t/2$ and $t_s + \Delta t$) for s = 0 min.

$\lambda\Delta t_0$	r = 1.0			r = 1.2			r = 1.5		
	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$
0.033	1.000	1.000	1.000	1.099	0.999	0.916	1.246	0.997	0.831
0.100	1.000	1.000	1.000	1.097	0.998	0.915	1.237	0.990	0.825
0.333	1.000	1.000	1.000	1.091	0.992	0.910	1.209	0.967	0.806
1.000	1.000	1.000	1.000	1.075	0.978	0.896	1.143	0.914	0.762
3.333	1.000	1.000	1.000	1.039	0.944	0.865	1.057	0.845	0.704

$\lambda\Delta t_0$	r = 2.0			r = 3.0			r = 5.0		
	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$
0.033	1.476	0.984	0.738	1.780	0.890	0.593	1.553	0.518	0.311
0.100	1.430	0.953	0.715	1.500	0.750	0.500	1.236	0.412	0.247
0.333	1.300	0.867	0.650	1.219	0.609	0.406	1.083	0.361	0.217
1.000	1.147	0.765	0.574	1.086	0.543	0.362	1.030	0.343	0.206
3.333	1.050	0.700	0.525	1.027	0.514	0.342	1.009	0.336	0.202

was well-mixed during the experiment. A microbial surrogate test aerosol was generated in a 1 m^3 chamber for 1 min and allowed to mix. After 5 min the air cleaning device was turned on and three sequential air samples taken; two of 1 min duration followed by a single sample of 2 min duration. There was an interval of 3 min between the end of one sample and the beginning of the next.

An Andersen six-stage viable sampler was used to collect samples from the chamber. The concentration data for the 0.65–1.10 μm diameter range was used for the subsequent analysis, expressed as colony forming units (CFU) per unit volume of air. The highest concentration

Table 3. Predicted decay rate ratio (λ/λ_{true}) for three different inter-sample intervals, *s*, for *r* = 2.0 and λ = 5.0 h^{-1}.

$\lambda\Delta t_0$	*s* = 1.0 min			*s* = 3.0 min			*s* = 5.0 min		
	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$	t_s	$t_s + \Delta t/2$	$t_s + \Delta t$
0.033	1.291	0.989	0.801	1.160	0.993	0.867	1.110	0.995	0.901
0.100	1.357	0.959	0.741	1.263	0.968	0.783	1.208	0.973	0.814
0.333	1.283	0.872	0.660	1.255	0.882	0.680	1.231	0.891	0.698
1.000	1.145	0.768	0.578	1.140	0.774	0.586	1.135	0.780	0.595
3.333	1.050	0.701	0.526	1.049	0.704	0.529	1.049	0.706	0.532

Table 4. Calculated decay rates (λ [min^{-1}]) and coefficients of determination (R^2) for example experimental measurements using non-linear and linear least squares approach with different reference times. Individual data (first three rows for each case), mean, standard deviation (S.d) and ratio to non-linear estimate shown.

	Non-linear		t_s		$t_s + \Delta t/2$		$t_s + \Delta t$	
	λ	R^2	λ	R^2	λ	R^2	λ	R^2
Case 1								
Data	0.411	0.990	0.435	0.996	0.408	0.991	0.384	0.985
	0.543	0.994	0.569	0.998	0.534	0.993	0.502	0.987
	0.501	0.986	0.527	0.993	0.494	0.986	0.464	0.978
Mean	0.485	0.990	0.510	0.995	0.479	0.990	0.450	0.983
S.d.	0.067	0.004	0.068	0.003	0.064	0.004	0.060	0.005
$\overline{\lambda} / \overline{\lambda_{nl}}$	1.000		1.052		0.987		0.928	
Case 2								
Data	0.682	0.985	0.714	0.992	0.669	0.984	0.628	0.976
	0.575	0.997	0.605	0.999	0.568	0.996	0.534	0.992
	0.668	0.971	0.701	0.964	0.663	0.976	0.628	0.984
Mean	0.641	0.984	0.673	0.985	0.633	0.985	0.597	0.984
S.d.	0.058	0.013	0.060	0.018	0.057	0.010	0.054	0.008
$\overline{\lambda} / \overline{\lambda_{nl}}$	1.000		1.049		0.987		0.930	

measured in this range was 6.5×10^4 CFU m^{-3} and the highest total concentration summed over all size ranges was 1.9×10^5 CFU m^{-3}. Two different air cleaning devices were examined (Cases 1 and 2). The details of the devices are not relevant to the interpretation of the data and are not provided for reasons of commercial confidentiality. The experiment was repeated three times for each air cleaning device.

The data were analysed by plotting the left hand side of Eq. (9) against t_s using an initial estimate of λ. The gradient of the line was used to refine the estimate of λ until there was a difference of less than 1×10^{-6} between the values. The values were also compared with values calculated using the non-linear Levenberg-Marquardt method and were found to be very similar (0.07%–0.8% error).

For comparison, values were also calculated by neglecting the variable sample duration and plotting the logarithm of the average concentration against each of t_s, $t_s + \Delta t/2$ and $t_s + \Delta t$. The resulting values are shown in Table 4.

The variation in sample duration is not as dramatic for these experimental data as some of the cases considered for the idealised data. However, the estimates of the decay rate show the same patterns. Specifically, the mean values ($\overline{\lambda}$) obtained using the logarithm of the average concentration and the start time (t_s) are higher than the mean non-linear value ($\overline{\lambda_{nl}}$) and are lower when the mean and end times are used. It is also interesting to note that,

although the calculated decay rates are different between the two cases, the ratios between the mean decay rates calculated by each method are very close between the cases (difference ≤ 0.003).

The examples explored here highlight the likely influence of variation in sample duration on the determined decay rate. However, this work is by necessity limited in scope. As discussed, the influence of experimental noise has not been systematically examined. The effect on the intercept has been explored theoretically, but not demonstrated in the example data as this is believed to be of less interest in the majority of studies. In addition, only one particular sampling sequence has been examined for the idealised data. It is possible to consider how a sampling sequence with variable duration might be optimised for a given experiment with a known sampling sensitivity. In some circumstances, it is also possible to overlap the sampling period of sequential samples. For the sake of brevity and clarity, neither of these more complicated aspects have been explored in this work.

CONCLUSIONS

The use of variable sampling duration when measuring an exponentially decaying concentration requires careful data analysis to accurately determine the decay rate. The theoretical basis for this analysis has been explored. An iterative approach using linear least squares curve fitting or

a non-linear curve fitting routine can be used to determine the decay rate.

The error introduced by neglecting the variable sampling duration has been explored using idealised data for a range of typical parameters. Plotting the logarithm of the average concentration using the start time tends to overestimate the decay rate, whilst the use of the mean and end times underestimates the rate.

When there is no interval between samples the error depends on the ratio of sample durations between subsequent samples and the product of the initial sample duration and the decay rate. For the parameter ranges considered, the error in decay rate can be as large as ± 80%. However, whilst neglecting the variable sample duration can introduce considerable error, this may not be apparent from the deviation from linearity in the curve fit. Examination of an experimental data set for concentration decay in a small chamber with variable sample duration shows the same pattern for the difference in derived decay rates.

While the application to the measurement of concentration decay has been explored, this approach may be relevant more broadly to other integrated measurements in the presence of a decay process.

ACKNOWLEDGMENTS

REFERENCES

Agranovski, V., Ristovski, Z., Hargreaves, M., Blackall, P. and Morawska, L. (2003). Real-time Measurement of Bacterial Aerosols with the UVAPS: Performance Evaluation. *J. Aerosol Sci.* 34: 301–317.

Beazley, D.M. (2009). *Python Essential Reference*, 4th Edition. Addison-Wesley, Upper Saddle River, NJ.

Fierz, M., Scherrer, L. and Burtscher, H. (2002). Real-time Measurement of Aerosol Size Distributions with an Electrical Diffusion Battery. *J. Aerosol Sci.* 33: 1049–1060.

Jung, J.H. and Lee, J.E. (2013). In Situ Real-time Measurement of Physical Characteristics of Airborne Bacterial Particles. *Atmos. Environ.* 81: 609–615.

Kimmel, J.R., Farmer, D.K., Cubison, M.J., Sueper, D., Tanner, C., Nemitz, E., Worsnop, D.R., Gonin, M. and Jimenez, J.L. (2011). Real-time Aerosol Mass Spectrometry with Millisecond Resolution. *Int. J. Mass Spectrom.* 303: 15–26.

Levenberg, K. (1944). A Method for the Solution of Certain Non-linear Problems in Least Squares. *Q. J. Appl. Math.* 2: 164–168.

Li, Y., Hao, L., Wang, S., Hou, L., Zhang, J. and Qi, J. (2009). An Experimental Study on Removal Efficiency of Bio-particles in an Airtight Decontamination Chamber. *Build. Environ.* 44: 2270–2275.

Marquardt, D. (1963). An Algorithm for Least-squares Estimation of Nonlinear Parameters. *J. Soc. Ind. Appl. Math.* 11: 431–441.

Oliphant, T.E. (2007). Python for Scientific Computing. *Comput. Sci. Eng.* 9: 10–20.

Park, J.M., Rock, J.C., Wang, L., Seo, Y.C., Bhatnagar, A. and Kim, S. (2009). Performance Evaluation of Six Different Aerosol Samplers in a Particulate Matter Generation Chamber. *Atmos. Environ.* 43: 280–289.

SciPy Community (2012). SciPy Reference Guide Release 0.11.0.