ILVO

Test report:

Evaluation of five measuring devices for measuring PM_{10} and $PM_{2.5}$ concentrations in poultry houses

Vlaanderen
is landbouw & visserij

Authors:

Zhuang Shaojie

Laanen Loes

Instituut voor Landbouw, Visserij- en Voedingsonderzoek ILVO

Eenheid Technologie en Voeding, Onderzoeksdomein Agrotechniek

Burg. Van Gansberghelaan 115, bus 1

B-9820 Merelbeke

Contact:

- e-mail: loes.laanen@ilvo.vlaanderen.be
- Phone: +32 9 272 27 69

April 2023

Inhoud

INTRODUCTION 2

MATERIALS AND METHODS 3

- Dust measuring devices 3
  - Overview 3
  - Gravimetric method 3
  - Continuous measuring methods 4
- Test facility and test set-up 4
- Descriptive statistical estimates 5
- Standards EN 14793:2017 6

RESULTS AND DISCUSSION 7

- Overview of the measured dust concentrations 7
- Performance of impactor and cyclone samplers 12
- Determination of the acceptance criteria 15
- Demonstration of performance equivalence 15

CONCLUSIONS 19

REFERENCES 19

APPENDIX 1 – NOMENCLATURE 20

APPENDIX 2 – GRIMM VS. IMPACTOR SAMPLER (EXTRA) 20

APPENDIX 3 – EN 14793:2017 REPORT TABLE 22

INTRODUCTION

This test report is part of ILVO's reference tasks for the policy area Environment and more specifically within ***Task 7.3 Validation measurements for fine dust devices***.

According to the European standard *"EN 12341 – Ambient air – Standard gravimetric measurement method for the determinations of PM10 or PM2.5 mass concentration of suspended particulate matter"* (EN 12341:2014)., the reference measurement method for measuring particulate matter is the gravimetric method with impactor pre-separators. However, this standard was developed for measurements in ambient air. Due to possible overloading of the impactor plates in environments with high dust concentrations and therefore overestimation of the fine dust concentrations, this method is considered less suitable for use in livestock houses. Wageningen Livestock Research therefore recommends the use of cyclone pre-separators instead of impactor pre-separators (Ogink et al., 2011).

Despite being suitable for use in livestock housing, these cyclone pre-separators remain a gravimetric method that is very labour-intensive. Moreover, this method gives only 1 average concentration over the duration of the sampling, so fluctuations in time (e.g. day-night pattern for 24-hour measurements) are not visible.

There are also several measuring devices on the market today that measure particulate matter concentrations on a continuous basis. These devices not only provide more data and make fluctuations visible, but are also less labour-intensive. However, these measuring devices were developed for measurements in ambient air, which raises the question whether these devices could also be suitable for measurements in environments with high dust concentrations and high relative humidity as is the case in livestock houses. To investigate this, validation measurements were performed with different measuring devices for fine dust concentrations and the results of these devices were compared with the results of the gravimetric method.

This test report will summarise the validation measurements that have been carried out with different measuring devices for the determination of fine dust concentrations to verify their suitability for use in livestock houses.

Materials and Methods

Dust measuring devices

Overview

The tested measuring devices for particulate matter (PM10 and/or PM2.5) in this study included the gravimetric method with an impactor pre-separator (Imp, reference method), the gravimetric method with a cyclone pre-separator (Cyc), the DustTrak™ DRX Aerosol Monitor 8533EP (Dst), the Dust Decoder 11-D (Gri29, Gri62, Gri63), the EDM365-SVC (Gri365) and the Microdust Pro™ (Opt). These measuring devices are hereafter referred by their respect brand or trademark names. A summary of the analyser specifications is given in Table 1.

Table 1 Summary of the specifications of the measuring device for measuring fines dust concentrations (PM10 and/or PM2.5).

Device	Manufacturer	Device ID [a]	Measuring method	Calibration	Particle size
Gravimetric method with impactor pre-separator	Pump + impactor: Comde-Derenda GmbH, Stahnsdorf, Germany	Imp1 Imp2	Gravimetric	Jul 2020	PM2.5, PM10
Gravimetric method with cyclone pre-separator	Pump: Comde-Derenda GmbH, Stahnsdorf, Germany Cyclone PM10: 2000-30ENB, URG, Chapel Hill, North Carolina, USA Cyclone PM2.5: 2000-30EG, URG, Chapel Hill, North Caroline, USA	Cyc1 Cyc2	Gravimatric	Jul 2020	PM2.5, PM10
DustTrak™ DRX Aerosol Monitor 8533EP	TSI Incorporated, Minnesota, United States	Dst1 Dst2	Laser photometer	Jul 2020 Oct 2020	PM2.5, PM10
Dust Decoder 11-D	Grimm Aerosol Technik Ainring GmbH & Co. KG, Ainring, Germany	Gri29 Gri62 Gri63 [b]	Laser Spectro-meter	Jul 2020	PM2.5, PM10
EDM365-SVC	Grimm Aerosol Technik Ainring GmbH & Co. KG, Ainring, Germany	Gri365 [b]	Spectro-meter	Jul 2020	PM2.5, PM10
Microdust Pro™	Casella UK, Bedford, United Kindom	Opt1 Opt2	IR photometer	Aug 2020	PM10

[a] The ID of devices used for this study.
[b] The device was deployed in the field test but the measurements were excluded from the equivalence assessment. The measurement results are provided in Appendix 2.
[c] The method being used as the reference in this study.

Gravimetric method

For the gravimetric method, air is sampled by using a constant flow pump (MVS 6.1, Comde-Derenda Gmbh, Stahnsdorf, Germany). The volumetric flow rate is measured with an orifice plate between filter and vacuum pump and electronically controlled with an accuracy of ≤ 2 % deviation.

The climatic conditions are continuously monitored by temperature and humidity sensors. the intake air passes through an inlet, which contains a pre-separator (impactor or cyclone). This pre-separator ensures that only the desired fraction (PM10 or PM2.5) passes through to the filter and that the larger dust particles are captured. The air with the desired fraction is drawn through a filter (MN GF-3, ∅47mm, Macherey-Nagel GmbH & Co, Düren, Germany). The dust particles remain on this filter. The weight of the filter is determined both before and after sampling by weighing each filter (after acclimatisation in a climate chamber at T = 20°C ±1°C and RH = 50% ±5%) 4 times spread over 2 days on a precision balance with resolution of 10 µg (MAS225S-100-DI, Sartorius, Göttinger, Germany). The average of the 4 weights is used as the weight of the unloaded and loaded filter. The difference in weight between the loaded and unloaded filter gives the amount of fine dust that was in the extracted air. The concentration of fine dust in the air is determined by dividing the mass of fine dust captured on the filter by the sampled air volume.

The difference between the 2 gravimetric method is in the way the desired fraction is separated from the larger dust fractions. In an impactor pre-separator, this is achieved by using an impactor plate. On this greased plate, the larger dust particles will remain stuck and only the particles with the desired size pass through. The cyclone pre-separator uses the centrifugal principle to separate the larger particles from the smaller ones.

Continuous measuring methods

The continuous measuring methods (DustTrak™ DRX Aerosol Monitor 8533EP, Dust Decoder 11-D, EDM365-SVC and Microdust Pro™) all use light scattering to measure the concentration of PM10 and/or PM2.5. The used light (laser or IR) and the measuring method (photometer or spectrometer) can differ between the devices (Table 1).

Test facility and test set-up

The particulate matter concentration was determined simultaneously by all the different measuring devices. In total, 28 measurements with a duration between 2 and 48 hours, were carried out between 24/11/2020 and 20/05/2021. Measurements were carried out in poultry houses (both broilers and laying hens) and in ambient air. Table 2 summaries the measurements (location, animal category, duration and number of measurements).

Table 2: summary of the measurements in this study

Location	Animal category	Duration (h)	Number of measurements
ILVO test houses	Broilers	2	2
		12	2
		24	3
	Laying hens	24	5
ILVO Hangar	Ambient air	24	1
		48	1
Proefbedrijf Pluimveehouderij test houses	Broilers	2	1
		12	2
	Laying hens	2	6
		4	2
		12	3

During the tests, alle measuring devices were placed together, ensuring that the air inlets of all devices were positioned as close as possible to each other. Figure 1 shows the test set-up in the pressure chamber at Proefbedrijf Pluimveehouderij in Geel.

Figure 1: Set-up for simultaneous measurement with the different devices for measuring PM10 and PM2.5 concentrations in poultry houses.

Descriptive statistical estimates

The measured PM10 and PM2.5 concentrations were firstly analysed using basic statistical tools. The overall measurement range was described by the minimum, maximum, mean and median per device type. The agreement between the duplicate within each device type was described with four statistical estimates. The general agreement between the two devices was indicated by the coefficient of determination (i.e. R^2). The absence of measurement difference between the two devices of the same type was assessed by performing a Student's t-test (tested for $\alpha = 0.05$) on the pairwise difference. The mean difference being statistically zero would mean that there was no difference. The consistency of the pairwise difference was presented by the mean (Equation 1) and median (Equation 2) of the pairwise standard deviation.

Equation 1

$$RepS.avg == \frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{\sum_{j=1}^{n}(y_{ij}-\bar{y}_i)^2}{n}}, \text{ with } \bar{y}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij}$$

Equation 2

$$RepS.med = median\{\sigma_1, \sigma_2, \cdots, \sigma_i\}, \text{ with } \sigma_i = \sqrt{\frac{\sum_{j=1}^{n}(y_{ij}-\bar{y}_i)^2}{n}}$$

Standards EN 14793:2017

The assessment of the device equivalence followed the test procedure prescribed by the European standards "*EN 14793 – Stationary source emissions - Demonstration of equivalence of an alternative method with a reference method*" (EN 14793:2017). To demonstrate equivalence between a reference method and an alternative method, the norm mandates that four criteria have to be satisfied:

1. Correlation coefficient $r \geq 0.97$. This threshold is comparable to the $R^2 \geq 0.95$ criterion required by the European standards "*EN 12341:2014 Standard gravimetric measurement method for the determination of the PM10 or PM2.5 mass concentration of suspended particulate matter*".
2. Slope value C_1 is within an interval $[1 - \frac{s_R(\bar{\bar{z}})}{\bar{\bar{z}}}, 1 + \frac{s_R(\bar{\bar{z}})}{\bar{\bar{z}}}]$, with $\frac{s_R(\bar{\bar{z}})}{\bar{\bar{z}}}$ being a predetermined maximum acceptable relative uncertainty (reproducibility).
3. The offset C_0 is within an interval $[-s_R(\bar{\bar{z}}), s_R(\bar{\bar{z}})]$, with $s_r(\bar{\bar{z}})$ being a predetermined maximum acceptable uncertainty (reproducibility).
4. The repeatability standard deviation $s_r(\bar{\bar{x}}) \leq s_{r,limit}(\bar{\bar{z}})$, with $s_{r,limit}(\bar{\bar{z}})$ being a predetermined maximum allowable repeatability standard deviation of the reference method.

In these criteria $\bar{\bar{z}}$ standards for the grand mean of the measured concertation by the reference method. How $s_R(\bar{\bar{z}})$ and $s_{r,limit}(\bar{\bar{z}})$ should be determined are not explicitly defined by EN 14793. Instead, the norm states that "if $s_R(\bar{\bar{z}})$ is not specified in the reference method standard, the laboratory shall determine this value from the expanded uncertainty U_{RM} according to $s_R(\bar{\bar{z}}) = \frac{U_{RM}}{2}$", and "if $s_{r,limit}(\bar{\bar{z}})$ is not specified in the reference method standard, the laboratory shall determine its value from the standard deviation of the paired measurements of the reference method".

In this test, the $s_R(\bar{\bar{z}})$ is set to 10% of the grand mean of the measured concentration by the reference method on PM10 and PM2.5, respectively. The value of $s_{r,limit}(\bar{\bar{z}})$ is determined with a function of $\bar{\bar{z}}$, such that:

Equation 3

$$s_{r,limit}(\bar{\bar{z}}) = \exp\left(\beta_0 + \beta_1 \cdot \log(\bar{\bar{z}}) + k \cdot \delta\right)$$

where the $\exp(\beta_0 + \beta_1 \cdot \log(\bar{\bar{z}}))$ part of the equation predicts the average standard deviation of the reference method (i.e. impactor samplers) at a concentration $\bar{\bar{z}}$, and the $\exp(k \cdot \delta)$ part of the equation describes the expanded uncertainty of the prediction, with k being the coverage factor. The values of the parameters β_0 and β_1 were determined applying ordinary linear regression

Equation 4

$$\log(\sigma_z) = \beta_0 + \beta_1 \cdot \log(\bar{z})$$

where σ_z and $\bar{z}$ are the pairwise standard deviation and mean of the duplicated measurements of the reference method, respectively. The term δ in Equation 3 was the standard deviation of the residuals of Equation 4. Prior to the log-log transformation, the measurement uncertainty in the reference method increased with the concentration especially when the concentration exceeded 100 $\mu g/m^3$ (Figure 2), and this in turn led to a violation of the "equal variance" assumption required for the linear regression. Thus, the log-log transformation was adopted.

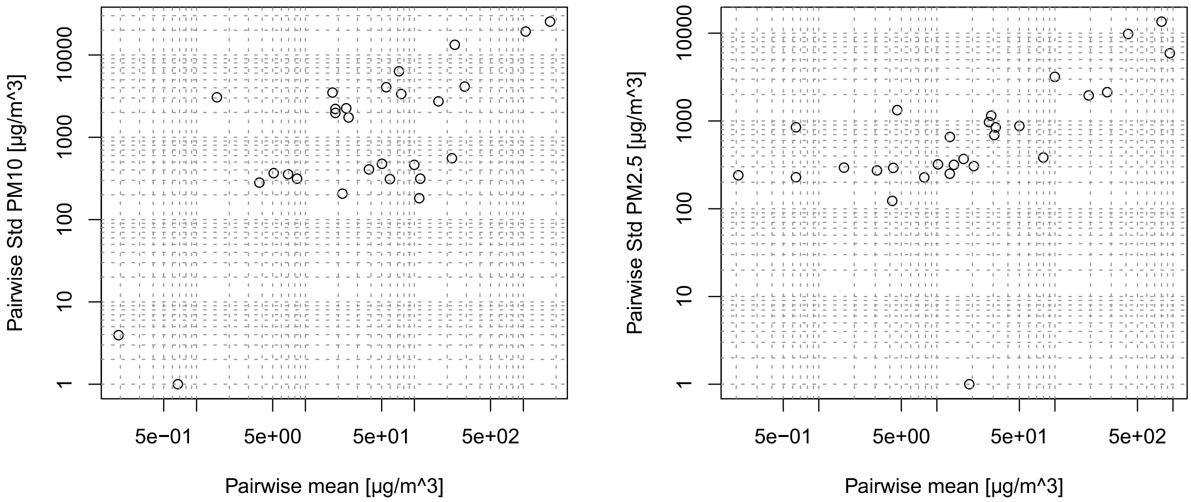

Figure 2 The pairwise mean and standard deviation of the impactor sampler measurements on PM10 (left) and PM2.5 (right). The pairwise mean on both plots were corrected for the offset to avoid negative values.

RESULTS AND DISCUSSION

Overview of the measured dust concentrations

The number of available measurements for PM10 and their observed concentrations is summarised in Table 33. For PM10 the observed dust concentrations were generally in the range of $0–3\times10^4$ $\mu g/m^3$ (Figure 3 and 4). The Optyl sensor measurements were noticeably lower than all the other device types especially for the high concentration range. Dust concentrations $>1\times10^4$ $\mu g/m^3$ were found in 3 out of the 28 trials (i.e. Trials 23, 25 and 27), and in the other 25 trials the measured concentrations were $<6\times10^3$ $\mu g/m^3$. In some of the trials (Trials 3, 15 and 16) both cyclone samplers reported negative PM10 concentrations, with a minimum of -453.8 $\mu g/m^3$. The impactor samplers also reported negative readings but only in one trial, and the lowest value (-3 $\mu g/m^3$) was close to zero. None of the automated sensor devices ever had a negative PM2.5 concentration reading. The Optyl sensors did not report concentration $>1\times10^3$ $\mu g/m^3$ in any of the trials. These observed concentrations were a lot wider than other evaluatory work. For example, the maximum values were 4×10^3 $\mu g/m^3$ PM10 and 168 $\mu g/m^3$ PM2.5 in Zhao et al. (2009) and 5×10^3 $\mu g/m^3$ PM10 in Winkel et al. (2015). The high concentration in this study could be a big challenge to all the measuring devices, including the impactors since the technique itself is not designed or verified for measuring high dust concentrations. The different measurement range from other studies and the unequal uncertainty across the observation range in the reference method as shown earlier (Figure 2) make it difficult to compare the performance of the measuring devices tested in this study directly against other studies.

Several device units were subjected to significant amount of data lose due to either system malfunctioning or suboptimal handling. DustTrak device 2 (Dst2) did not operate properly in 14

14 trials. Notice the boxplots of the two DustTrak units were seeming differently, and this difference was due to these lost trials. Measurements from Grimm 11D device 62 (Gri62) were invalid in 10 trials. Grimm EDM365 (Gri365) lost data from 11 trials. Both the Optyl devices missed 6 trials.

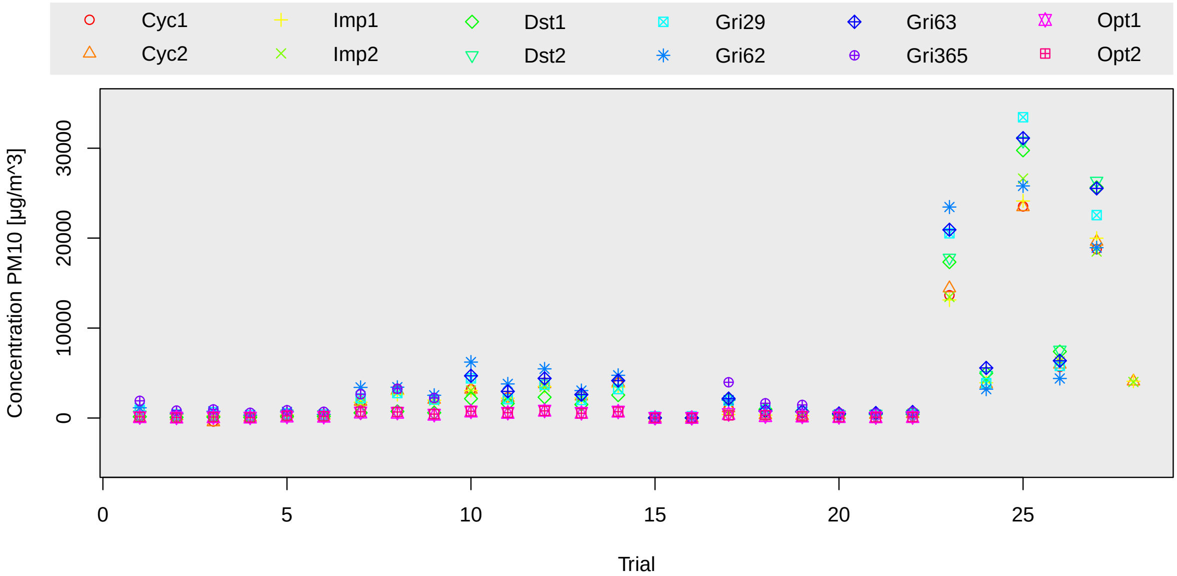

Figure 1 The measured PM10 concentrations in each trial by the dust measuring devices: Cyc – cyclone, Imp – impactor, Dst – DustTrak, Gri – Grimm, Opt – Optyl.

Table 3 Descriptive statistics of the measured dust concentrations (unit μg/m[3]) by each type of device.

Device	N_p	N_f	Grand Min.	Grand Median	Grand Max.	Grand Mean	Mean diff (p-value[e])	RepS.avg ± SE [a] (rel.% [b])	RepS.med (rel.% [c])	R^2 [d]
PM_{10}										
Cyclone	27	26	-453.8	1822.7	23522.1	3685.9	-36.8 (p>0.05)	170.3 ± 117.1 (4.6%)	18.3 (1.0%)	1.00
Impactor	28	28	-3	1214.8	26620.6	3600.1	-35.7 (p>0.05)	406.2 ± 292.1 (11.3%)	38.3 (3.2%)	0.99
DustTrak	27	14	25.8	576.9	30811.8	6350.1	-253.1 (p>0.05)	279.8 ± 205.1 (4.4%)	148.1 (25.7%)	1.00
Grimm	27	18	8.8	2027.4	33436.2	4785.1	-1.6 (p>0.05)	1195.8 ± 879.7 (25.0%)	341.5 (16.8%)	0.95
Optyl	22	22	8.1	193.3	818.7	312.9	-17.3 (p>0.05)	37.3 ± 20.4 (11.9%)	16.2 (8.4%)	0.96
$PM_{2.5}$										
Cyclone	27	27	-634.1	84.1	1092.1	178.9	-7.57 (p>0.05)	47.0 ± 28.5 (26.2%)	13.9 (16.6%)	0.96
Impactor	28	28	-253.4	302.6	13879.9	1510.4	-49.9 (p>0.05)	260.9 ± 168.2 (17.3%)	18.8 (6.2%)	0.99
DustTrak	27	14	24.5	238.5	9721.9	2048.2	-172.7 (p=0.01)	195.7 ± 161.8 (9.6%)	74.9 (31.4%)	1.00
Grimm	27	18	7.8	251.2	2666.7	425.7	14.9 (p>0.05)	50.1 ± 41.1 (11.8%)	10.3 (4.1%)	0.99

a Excluding trials containing no or only one device unit. SE – standard error.
b With respect to the grand mean. Interpret with caution when the device(s) have a non-zero offset.
c With respect to the grand median. Interpret with caution when the device(s) have a non-zero offset.
d Correlation of determinant between the device units of the same type.
e Student's t-test on the null-hypothesis: between-device difference is equal to zero.

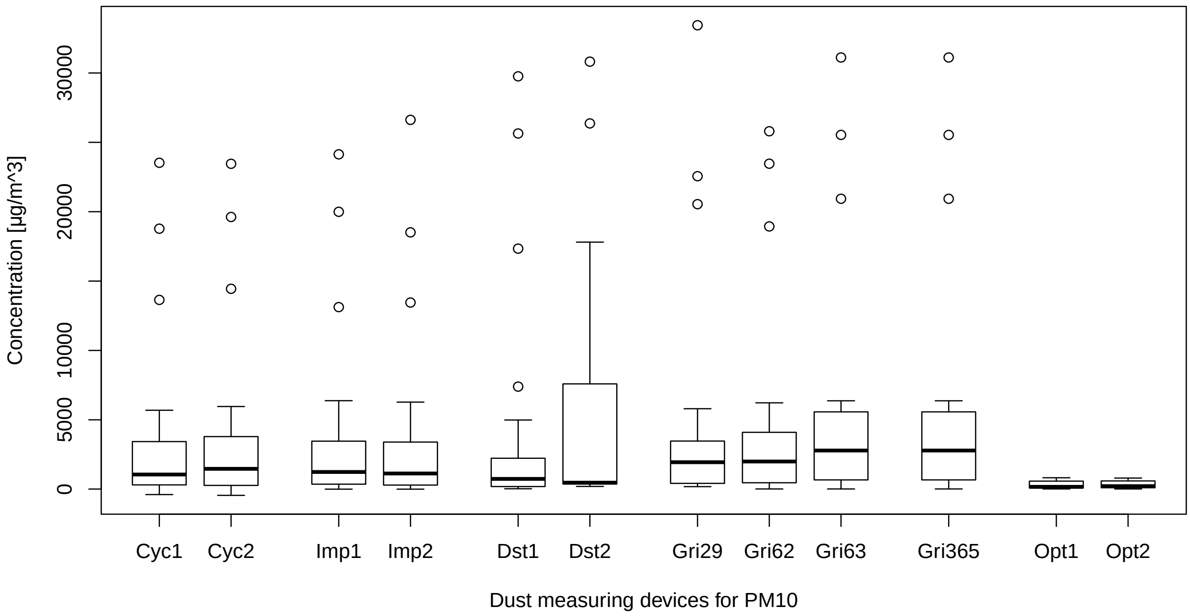

Figure 4 Summary of the measured PM10 concentrations by the dust measuring devices: Cyc – cyclone, Imp – impactor, Dst – DustTrak, Gri – Grimm, Opt – Optyl. The circles denotes outliers which are determined based on the 1.5×IQR rule.

The observed PM2.5 concentration ranges were generally consistent within each device type but with large differences between the device types (Figure 55 and Figure 6). Exceptionally high concentrations were measured in 3 out of the 28 trials (i.e. Trials 23, 25 and 27) by the impactor samplers and the DustTrak sensors. Negative PM2.5 concentrations were reported by both the cyclone samplers in 9 trials (minimum value = -634.1 $\mu g/m^3$), and by both the impactor samplers in 4 trials (minimum value = -253.4 $\mu g/m^3$). None of the sensor devices ever had negative PM2.5 concentration readings.

Reproducibility wise, the devices of the same type behaved similarly, indicated by the high R^2 values (≥ 0.96) (Table 3). The difference within most device pairs of the same type were not statistically different from zero. An exception was the two DustTrak sensors, where the measured values from one sensor (i.e. Dst1) tended to be lower than the other sensor (i.e. Dst2.) (p=0.01). This difference was likely linked to whether the sensor component is covered with a case; when both the DustTrak units were deployed without a case, the between-device difference was less pronounced (p=0.06, df=7) than when one unit was covered with a case (p<0.001, df=5). The Grimm sensors seemingly outperformed other devices types in terms of both the absolute and relative concentration measurements. The cyclone samplers had low absolute uncertainties, but the relative uncertainty was high. However, notice that the cyclone samplers were subjected to reporting negative values, and this caused the mean and median of the measured concentrations to become closer to zero. Consequently, the estimated relative uncertainty given in Table 3 is likely overestimated, and the offset has to be corrected for before the estimator can indicate the reproducibility of the cyclone samplers correctly. Similarly, the reproducibility impactor in terms of the relative uncertainty is also likely overestimated due to the non-zero offset (i.e. a positive offset). Due to the large differences in measured concentration ranges using the different device types, the reproducibility indicators should not be directly compared between the devices before recalibration (which should in principle bring the measured values to the same scale). For the same reason, R^2 between the devices of the same type should not be directly compared across different device types.

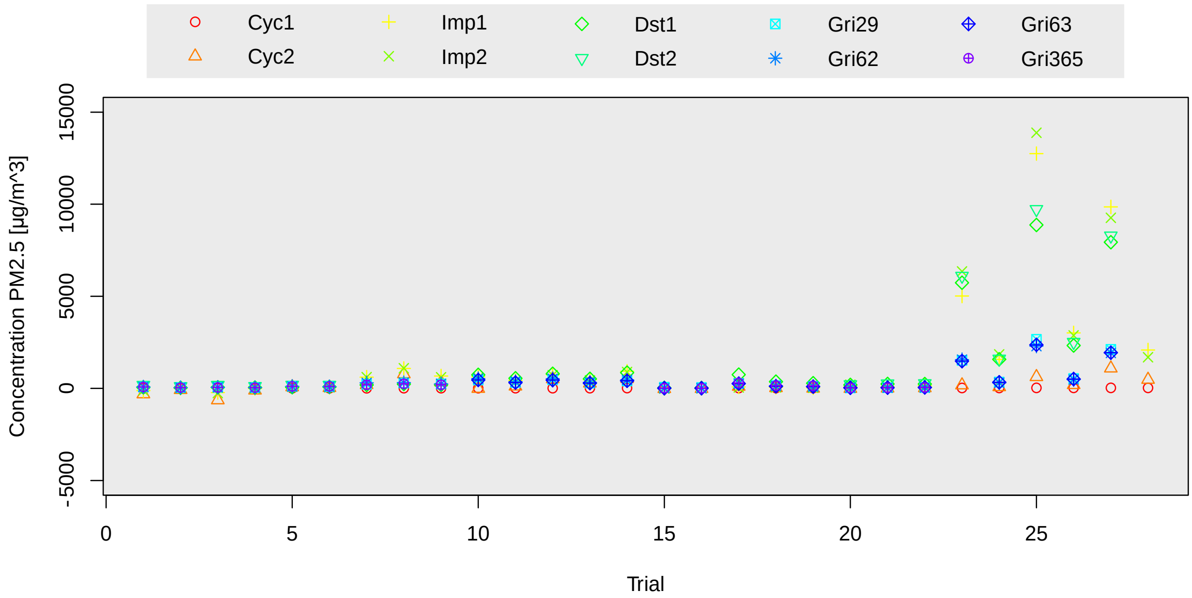

Figure 5 The measured PM2.5 concentrations in each trial by the dust measuring devices: Cyc – cyclone, Imp – impactor, Dst – DustTrak, Gri – Grimm, Opt – Optyl.

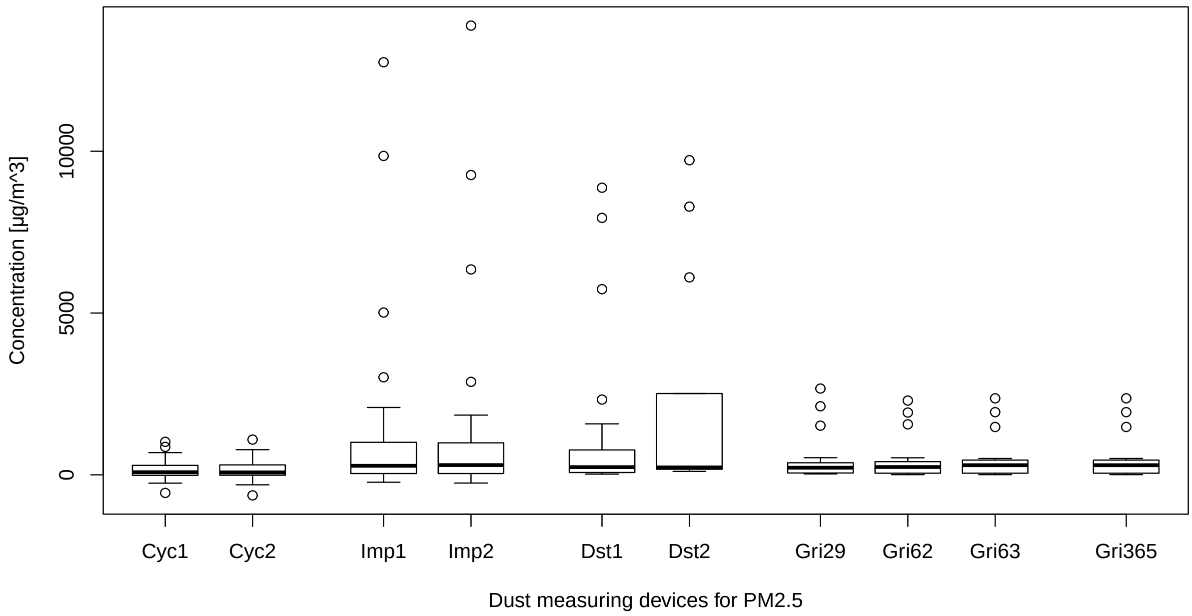

Figure 6 Summary of the measured PM2.5 concentrations by the dust measuring devices: Cyc – cyclone, Imp – impactor, Dst – DustTrak, Gri – Grimm. The circles denotes outliers which are determined based on the 1.5×IQR rule.

Performance of impactor and cyclone samplers

The measured PM10 and PM2.5 concentrations were graphically compared between the impactor samplers and cyclone samplers. More quantitative comparison will be given in the next section. The impactor sampler is commonly regarded as the reference method in office environment, and the cyclone samplers are currently considered a promising alternative method to the impactor samplers for dusty environments such as livestock barns (Winkel et al. 2015). For PM10, the measured concentrations by the impactor samplers and cyclone samplers exhibited a high correlation over the full measured concentration range, as shown in Table 33 and Figure 77. Still, there were some questionable measurements at concentrations <1000 μg/m^3 due to relatively large discrepancies either between or within the device types. These data points were regarded as outliers and were thereby excluded. In Zhao et al. (2009) a nonlinear response by the cyclone samplers on PM10 was observed, and the authors proposed to apply a two-stage linear calibration with a cut-off at 223 μg/m^3. This nonlinearity was not evident in our PM10 measurements (Figure 88). This could mean the nonlinearity found in Zhao et al. (2009) was not a fundamental characteristic of the cyclone samplers.

Both the impactor samplers and cyclone samplers were subjected to non-zero offsets on PM2.5 measurements and thereby yield negative readings. Moreover, large discrepancies in the measured PM2.5 concentrations were found between the two device types, and the measurement uncertainty seemed to increase at higher concentration levels (Figure). The linearity of the readings from the two device types was not ideal either. For <1500 μg/m^3 apparent PM2.5 concentration based on the impactor measurements, the cyclone sampler seeming had two calibration curves (Figure). The cut-off was not very obvious or certain since the impactor measurements were likely biased and the cut-off could change depending on whether or not to consider certain measurement points as outliers. In the example given in Figure 10, a cut-off of 50 μg/m^3 was used to illustrate the regression lines for “low” and “high” concentration ranges.

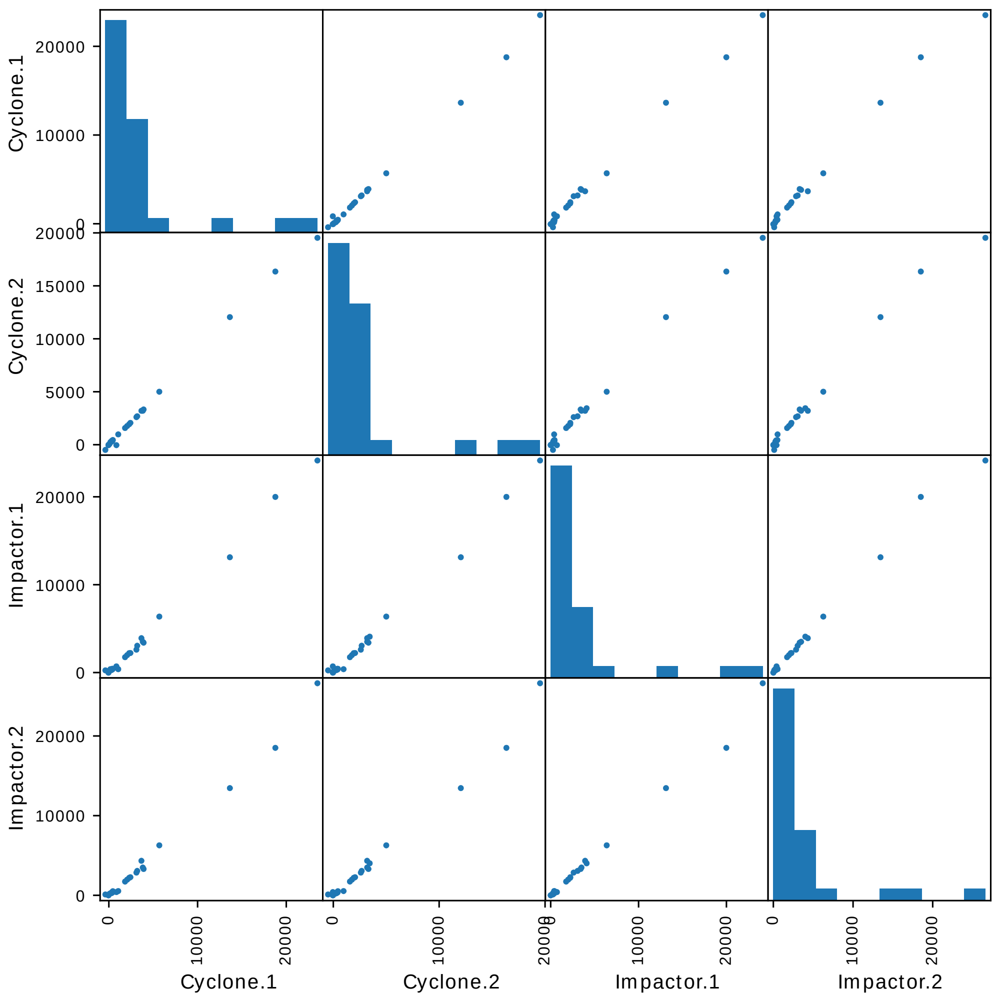

Figure 7 Scatter plots of the cyclone and impactor sampler measurement results on PM10 (unit = μg/m³). The number after the device name denotes the ID of the measuring unit. The presented dataset contains 27 measurement trials.

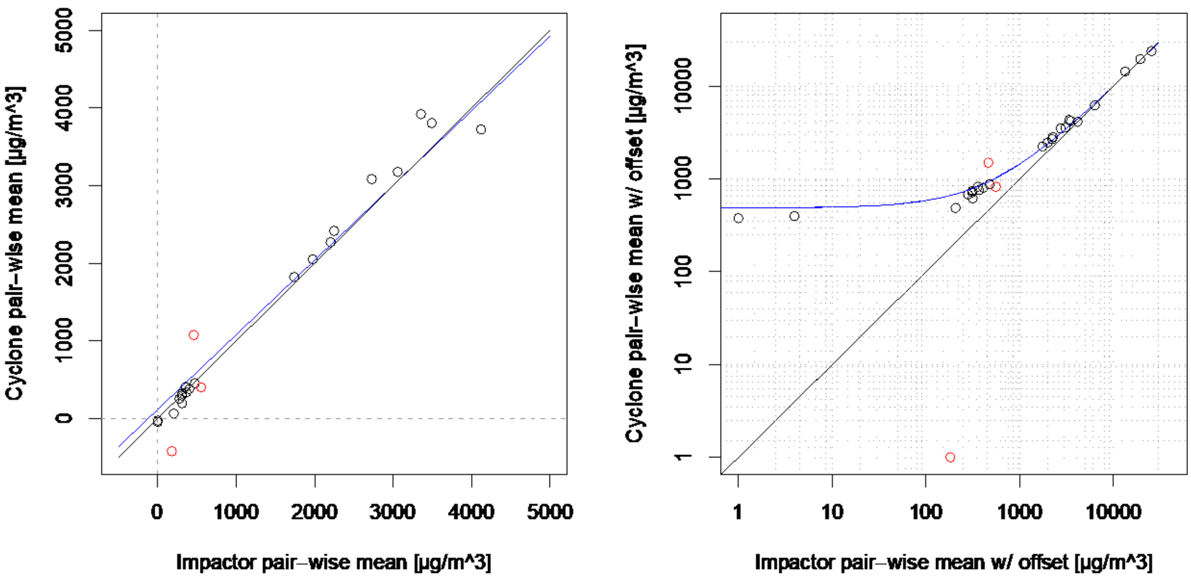

Figure 8 Comparison of the pair-wise mean between the impactor samplers and cyclone samplers on PM10 plotted in linear (left) and log-scale (right, with the offset shifted so that all values were ≥1). Blue solid line denotes the linear regression line (excluding outliers which are marked in red), and the black solid line denotes x = y.

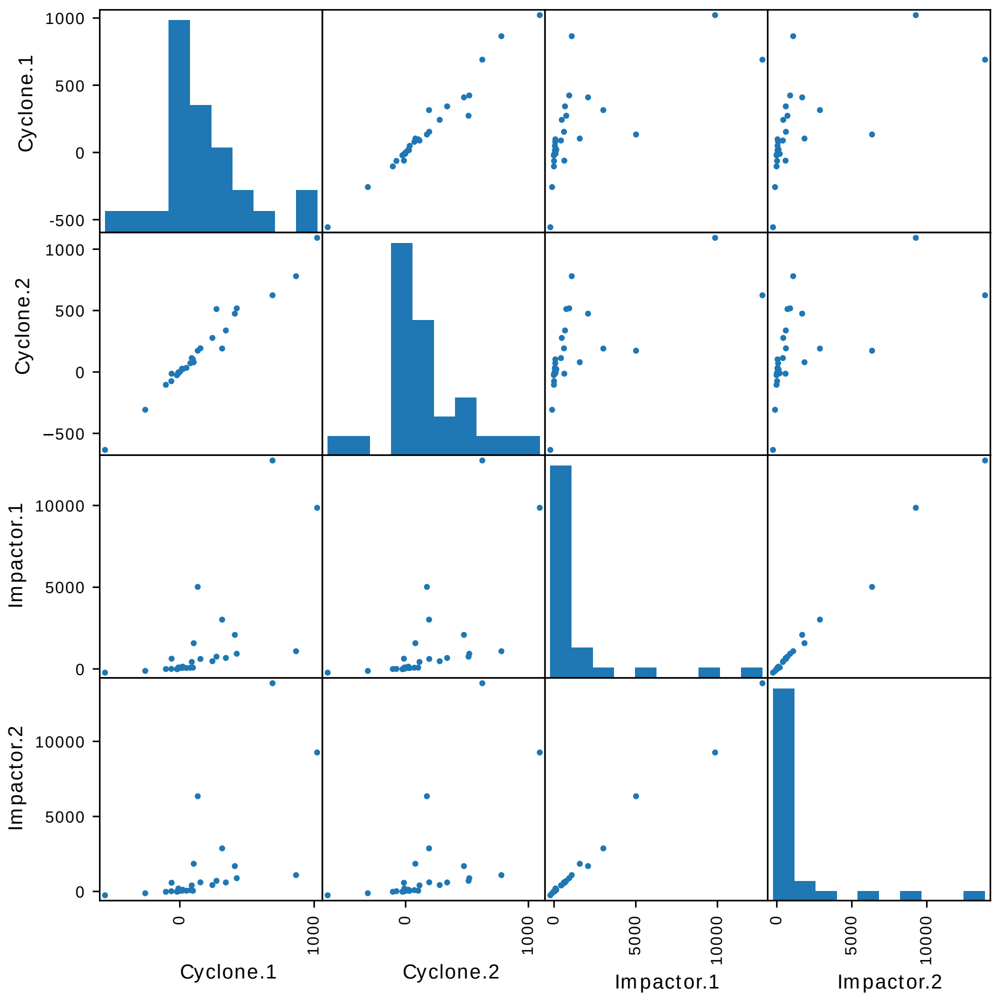

Figure 9 Scatter plots of the cyclone and impactor sampler measurement results on PM2.5 (unit = µg/m³). The number after the device name denotes the ID of the measuring unit. The presented dataset contains 27 measurement trials.

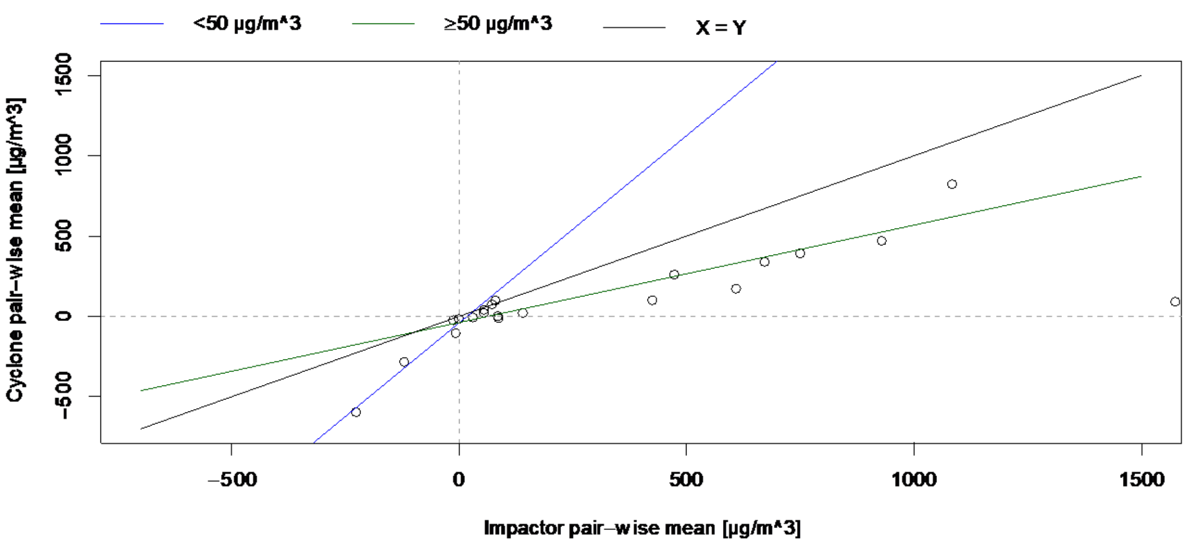

Figure 10 Comparison of the pair-wise mean between the impactor samplers and cyclone samplers on PM2.5. Blue (<50 µg/m³) and green solid (≥50 µg/m³) lines demonstrate the linear regression lines at two concentration ranges with respect to the apparent concentrations measured by the impactor samplers. The black solid line denotes x = y.

Determination of the acceptance criteria

The equations for PM10 and PM2.5 were estimated based on the measurement of the impactor samplers according to Equation 3 with $k = 1$ (i.e. 1 × standard deviation), and the following equations were obtained:

Equation 5

$$s_{r,limit}(\bar{\bar{z}}) = \begin{cases} \exp(-0.6163 + 0.6174 \cdot \log(\bar{\bar{z}}_{PM10}) + 1 \cdot 1.4279) = 2.252 \cdot \bar{\bar{z}}_{PM10}^{0.6174} \\ \exp(-0.8906 + 0.7041 \cdot \log(\bar{\bar{z}}_{PM2.5}) + 1 \cdot 1.3531) = 1.587 \cdot \bar{\bar{z}}_{PM2.5}^{0.7041} \end{cases}.$$

Because most of the alternative device types did not participate in all field tests, the $\boldsymbol{s_{r,limit}(\bar{\bar{z}})}$ values were determined for each alternative device type separately based on the actual exposed dust concentrations (see Table Annex 1).

Demonstration of performance equivalence

The test procedure for demonstration of performance equivalence was carried out based on the European standards EN 14793 (Table 4). Because the number of available measurements of all alternative devices did not meet the minimum requirement, the statistical analysis results could only provide indicative performance of each device type and cannot be used as a proof of qualification of the alternative measuring method.

For PM10, cyclone samplers were the only device that could pass all four criteria, therefore the equivalence could be demonstrated. However, having comparative performance does not justify the correctness of the measurements. Recall that both the impactor and cyclone samplers were reporting negative values, which was obviously incorrect since mass cannot be negative. The DustTrak and Grimm sensor measurements correlated well with the impactor samplers, but the measured concentrations tended to be higher than the impactor samplers especially at the high concentration range. This might relate to the sensor calibration since the tendency was found on all devices of the same type. The Grimm sensors were in addition subjected to poor repeatability. The Optyl sensors underestimated the PM10 concentrations by almost 80%. The correlation with the impactor measurements was also insufficient. Nonetheless, four deviating data points (Trial 5, 7, 17 and 18) were found in the Optyl measurements, and these data points had the highest within-pair discrepancies. It was unclear if these measurements were associated with any exceptional operation conditions, but the r value could be improved to 0.988 if these data points were removed.

Table 4 Compliance of the device types in PM10 and PM2.5 measurement. The letter (P – pass, F – failed) behind the parameter estimates indicates the evaluation result based on the criteria of acceptance.

	Systematic deviation			
Device type	Correlation r	Slope C_1	Intercept C_0 [a]	Repeatability $s_r(\bar{\bar{x}})$ [a]
PM10				
Cyclone	0.998 | P	0.963 | P	136.3 | P	170.1 | P
DustTrak	0.997 | P	1.267 | F	-131.1 | P	287.3 | P
Grimm	0.985 | P	1.208 | F	146.4 | P	1423.7 | F
Optyl	0.947 | F	0.225 | F	57.9 | P	38.0 | P
Criterion	$r \geq 0.97$	$\|C_1 - 1\| \leq 0.1$	$\|C_0\| < s_R(\bar{\bar{z}})$	$s_r(\bar{\bar{x}}) \leq s_{r,limit}(\bar{\bar{z}})$
PM2.5				
Cyclone	0.626 | F	0.105 | F	9.1 | P	47.4 | P
DustTrak	0.987 | P	0.765 | F	239.7 | P	202.6 | P
Grimm	0.977 | P	0.193 | F	135.6 | P	65.2 | P
Criterion[1]	$r \geq 0.97$	$\|C_1 - 1\| \leq 0.1$	$\|C_0\| < s_R(\bar{\bar{z}})$	$s_r(\bar{\bar{x}}) \leq s_{r,limit}(\bar{\bar{z}})$

[a] The criterion are determined separately for each alternative device type. See Table Annex 1 for the actual threshold and calculation details.

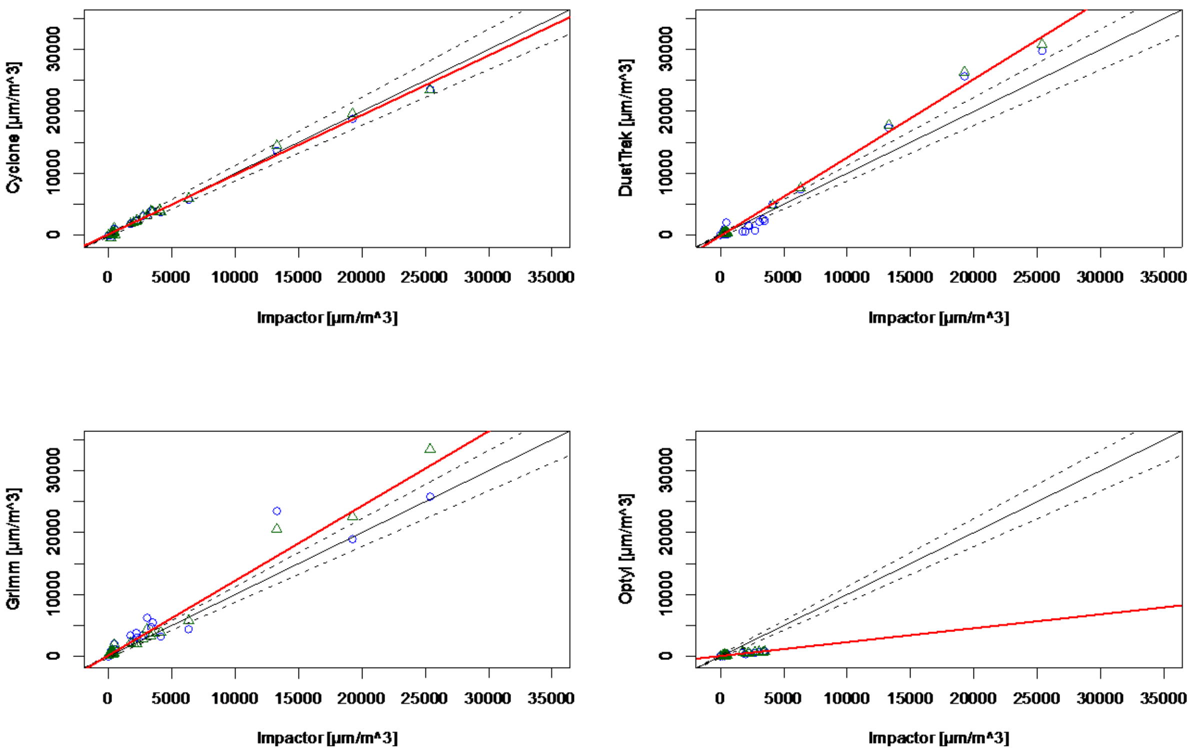

Figure 21 Comparisons between the impactor sampler measurements and four alternative methods on PM10. The blue circles and green triangles represent the measurement from one of the two units, respectively. The red solid line denote the orthogonal regression line. The black solid line denotes x=y regression line, and the dashed lines denote the interval of the desirable accuracy.

For PM2.5 none of the tested devices could show equivalence to the impactor samplers (Table 44). As shown earlier in Figure 10 there was a lack of linearity between the cyclone samplers and the impactor samplers, which eventually lead to a low *r* value (0.626). Moreover, the cyclone samplers significantly underestimated the PM2.5 concentrations by roughly 90%. Although the intercept was reasonably small, this would actually indicating a measurement bias since the impactor samplers themselves had a non-zero offset. The DuskTrak sensors correlated sufficiently with the impactor samplers, though according to Figure 2 there was a slight non-linearity for the >5000 μg/m^3 range. The repeatability standard deviation and offset was sufficiently small. Notice the intercept of the DustTrak was 239.7 and the lowest PM2.5 measurement was -253.4 μg/m^3, it was plausible that the true offset of the DustTrak sensors were actually close to zero. The measured PM2.5 concentrations were however lower than the impactor samplers by 25% according to the slope coefficient, and on this criterion the DustTrak failed to pass the equivalence. Similarly, the Grimm sensors had sufficiently high correlation coefficient against the impactors and low offset and repeatability standard deviation. However, the measured concentrations was roughly 80% lower than the impactor measurements.

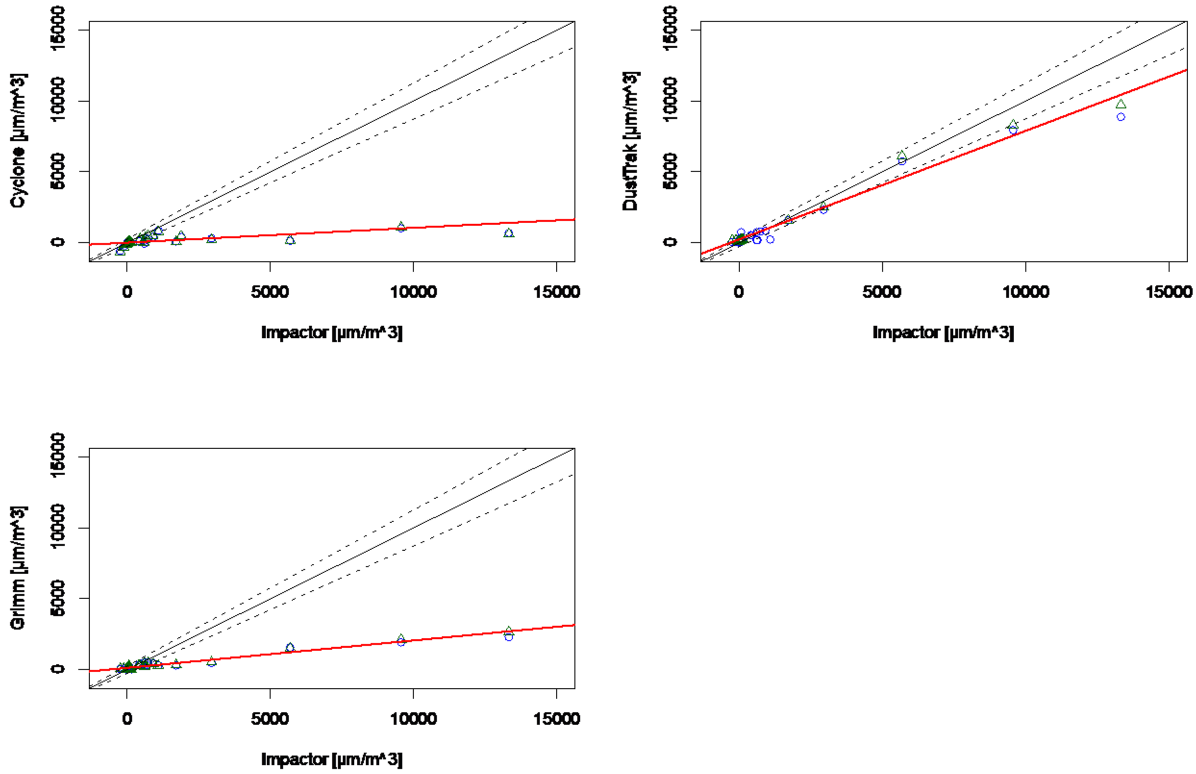

Figure 12 Comparisons between the impactor sampler measurements and four alternative methods on PM2.5. The blue circles and green triangles represent the measurement from one of the two units, respectively. The red solid line denote the orthogonal regression line. The black solid line denotes x=y regression line, and the dashed lines denote the interval of the desirable accuracy.

Although the impactor samplers were regarded as the reference, their reliability for livestock environment is questionable. For the very high PM2.5 concentrations there was a clear sign of overloading, meaning that the PM10 was also contributing to the PM2.5 measurement. Since the dust load in the sampler is theoretically time dependent, the duration of sampler period should also be accounted for. The following model was constructed to describe the numerical relationship between the cyclone measurements on PM2.5 and the PM10 measurements by the impactor samplers:

$$Cyc_{PM2.5} = \beta_0 + \beta_1 \cdot \log(IMP_{PM10}) + \beta_2 \cdot \log(Duration) + \beta_3 \cdot \log(IMP_{PM10}) \cdot \log(Duration),$$

where $Cyc_{PM2.5}$ was the mean PM2.5 concentration of the cyclone sampler duplicates, IMP_{PM10} was the mean PM10 concentrations of the impactor sampler duplicates, $Duration$ was the measurement duration in the field, and βs were the regression coefficients ($\beta_0 = -2096.3$, $\beta_1 = 299.7$, $\beta_2 = 580.9$, $\beta_3 = -78.1$). In order to allow taking the logarithm, the impactor sampler measurements were offset-shifted so that all values were ≥1. The absolute values of the slope coefficients were unlikely relevant to the actual PM2.5 concentrations, but the trend was reasonable: the slope coefficient for $\log(IMP_{PM10})$ decreased as the measurement time increased, meaning that more PM10 managed to bypass the PM10 filter and reached the PM2.5 collector. This model reached R^2 of 0.63 with respect to the cyclone measurements on PM2.5 (Figure 33). The actual PM2.5 concentration values measured by the impactor samplers was discarded from the model because it was unable to improve the model fit ($p = 0.59$).

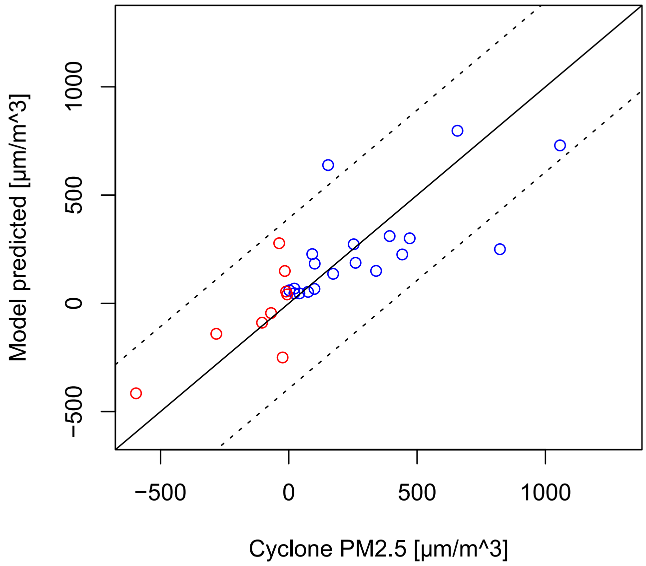

Figure 33 Comparison of PM2.5 measured by cyclone samplers and the prediction using PM10 measurements and duration of impactor samplers. The solid line denotes X=Y, and the dashed lines denote the 95%-CI of the model-fit residuals. Cyclone measurements of positive values were presented by blue dots, and negative values were by red dots.

Two Grimm sensor (Gri63 and Gri365) devices were not included in the test for demonstrating equivalence to the reference method, but based on a simple comparison the sensor performance could still be checked. The Gri63 device, similar to the units of the same type Gri29 and Gri62, tended to overestimate the PM10 concentrations by almost 30% (Figure 14), but on PM2.5 the measurements were underestimated by 80%. These values were comparable to the estimates given in Table 4 for the Grimm device, and the device would not be able to pass the equivalence test simply based on the slope coefficients. The performance of Gri365 was worse than the other Grimm sensors (Figure 15). The linearity and regression coefficients were poor for both PM10 and PM2.5 measurements. Since only one unit of this model was available, it is not possible to conclude whether the model per se has certain performance pitfalls or the sensor unit was not being handled properly.

Conclusions

In this study five devices, including duplicates, were tested for measuring PM10 and PM2.5 in poultry houses. Impactor samplers were regarded as the reference method. Cyclone samplers showed good agreement with the impactor samplers on PM10 measurements, but for PM2.5 the cyclone samplers were suffering from suboptimal linearity and non-zero offset. The negative dust concentrations in both gravimetric methods were presumably due to sub-optimal conditions during acclimatisation, indicating the importance of this step. Furthermore, it was shown by Zhao et al. (2009) that impactor pre-separators can be quickly overloaded when measuring in environments with high dust concentrations such as in poultry houses, leading to overestimated concentrations. This was especially the case for PM2.5. Cyclone pre-separators are more resistant to high dust concentrations.

The DustTrak sensors showed a high correlation to the impactor samplers and sufficient reproducibility. The sensor calibration should however be improved to avoid overestimation in PM10 and underestimation in PM2.5. Grimm sensors also showed a high correlation to the impactor samplers, and better calibration was needed to meet the accuracy requirement. The reproducibility of the Grimm devices on PM10 was also unsatisfying. The Optyl sensors showed high correlation to the impactor samplers and good reproducibility in PM10 measurements, but they also require better calibration. In this study, the reliability of the impactor samplers could be somewhat questioned due to negative dust concentration readings, and therefore the estimated performance indicators, as given in the equivalence test report, might not reflect the true performance of the other devices. Due to insufficient number of trials, the test results reported in this study could not be used to demonstrate the (non-)equivalence of the alternative devices to the impactor samplers.

References

EN 12341, 2014. *Ambient air – Standard gravimetric measurement method for the determination of PM10 or PM2.5 mass concentration of suspended particulate matter*, Brussels, Belgium: European Committee for Standardization.

Ogink, N.W.M., Hofschreuder, P., Aarnink, A.J.A., 2011. Protocol voor meting van finnstofemissie uit huisvestingssystemene in de veehoduerij 2010. https://edepot.wur.nl/179499

EN 14793, 2017. Stationary source emissions - Demonstration of equivalence of an alternative method with a reference method, Brussels, Belgium: European Committee for Standardization.

Ogink, N.W.M., Hofschreuder, P., Aarnink, A.J.A., 2011. Protocol voor meting van finnstofemissie uit huisvestingssystemene in de veehoduerij 2010. https://edepot.wur.nl/179499

Winkel, A. et al., 2015. Equivalence testing of filter-based, beta-attenuation, TEOM, and light-scattering devices for measurement of PM10 concentration in animal houses. Journal of Aerosol Science, 80, pp.11–26. Available at: http://dx.doi.org/10.1016/j.jaerosci.2014.10.009.

Zhao, Y. et al., 2009. Evaluation of an impaction and a cyclone pre-separator for sampling high PM10 and PM2.5 concentrations in livestock houses. Journal of Aerosol Science, 40(10), pp.868–878. Available at: http://dx.doi.org/10.1016/j.jaerosci.2009.06.001.

Appendix 1 – Nomenclature

N_p	Number of trials containing measurements taken by at least one unit of the given device type
N_f	Number of trials containing measurements taken by at least two units for the given device type
n	The number of units rested of the given device type
y	The measured dust concentration
i	The index of the N trials
j	The index of the n device units
$\bar{y}_i$	The mean of the ith trial; $\bar{y}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij}$
P	The total number of measurements per device type; including duplications
$\bar{\bar{z}}, \bar{\bar{x}}$	Grand mean of the measured concentrations by the reference (z) and alternative (x) method, respectively.
$\bar{z}, \bar{x}$	The pairwise mean of the measurements
$\bar{\sigma}_z$	The pairwise standard deviation of the reference measurements

Appendix 2 – Grimm vs. impactor sampler (extra)

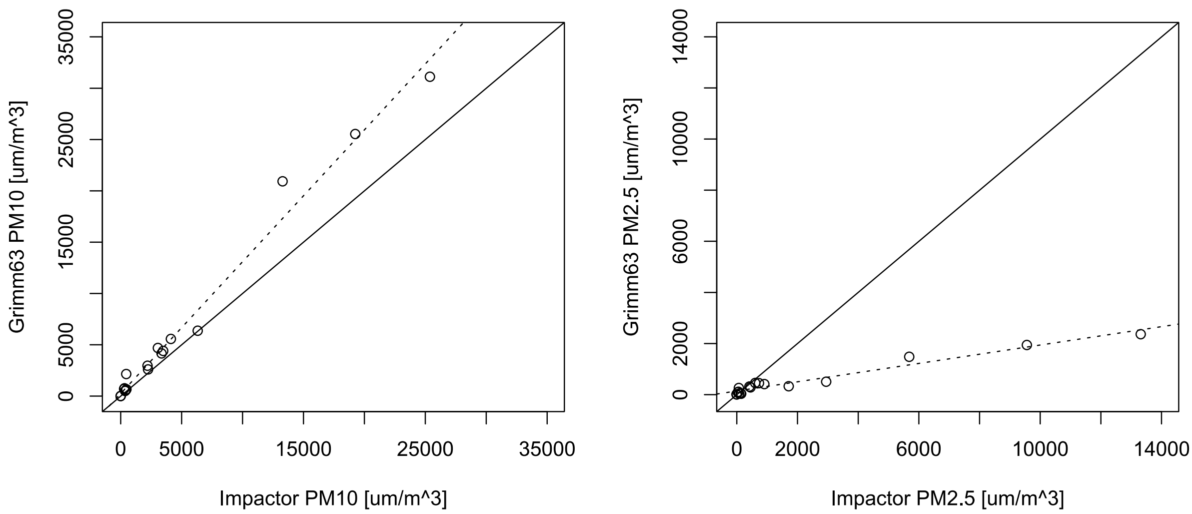

Figure 14 Comparisons between Grimm 63 and impactor samplers on PM10 (left) and PM2.5 (right) measurements. The solid lines denote x=y. The dashed lines denote the regression lines (unit μm/m³): $\boldsymbol{Gri63_{pm10} = 251.7 + 1.29 \times Impactor}$ and $\boldsymbol{Gri63_{pm2.5} = 140.4 + 0.18 \times Impactor}$.

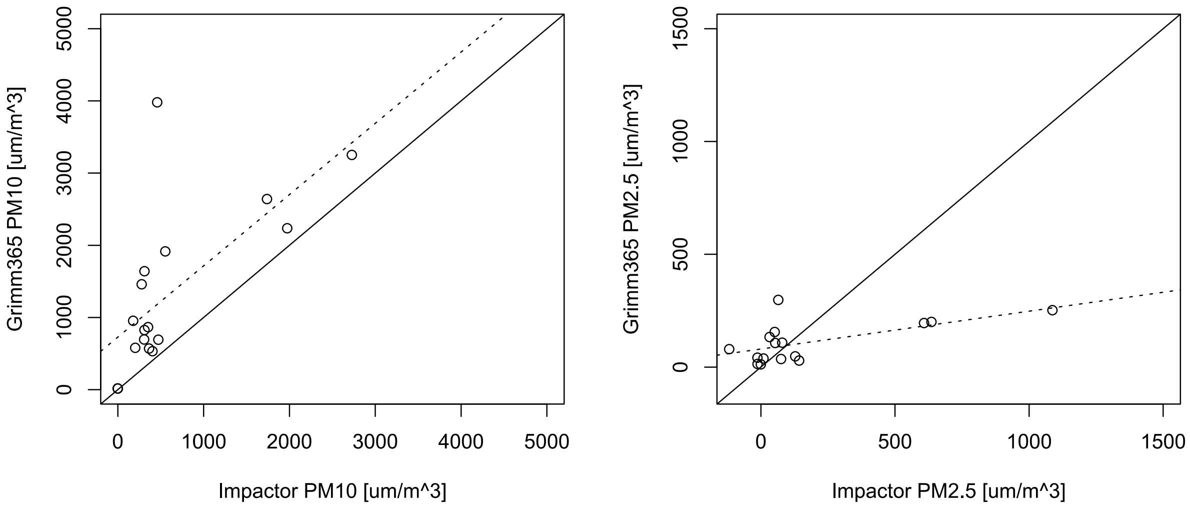

Figure 15 Comparisons between Grimm 365 and impactor samplers on PM10 (left) and PM2.5 (right) measurements. The solid lines denote x=y. The dashed lines denote the regression lines (unit μm/m³): $\boldsymbol{Gri365_{pm10} = 728.1 + 0.99 \times Impactor}$ and $\boldsymbol{Gri365_{pm2.5} = 80.3 + 0.17 \times Impactor}$.

Appendix 3 – EN 14793:2017 report table

Table Annex 1 Statistical results of the tests for equivalence based on EN 14793:2017. The impactor sampler (Impactor) is regarded as the reference method (denoted by "z") for all the alternative methods (denoted with "x"), and the cyclone sampler (Cyclone) is used as the reference to determine the statistical estimates for the impactor sampler.

		PM10					PM2.5			
		Impactor	Cyclone	DustTrak	Grimm	Optyl	Impactor	Cyclone	DustTrak	Grimm
Systematic deviation										
Grand mean	$\bar{\bar{x}}$ $\bar{\bar{z}}$	3704 3701.7	3701.7 3704	6821.3 5486	4835.6 3881.4	323.2 1180.9	1545.2 170.7	170.7 1545.2	2200.1 2562.9	447.8 1621.1
Repeatability										
Repeatability standard deviation	$s_r(\bar{\bar{x}})$ $s_r(\bar{\bar{z}})$	419.7 170.1	170.1 419.7	287.3 582.3	1423.7 430.7	38 74.1	265.8 47.4	47.4 265.8	202.6 367.4	65.2 270.9
Variation of repeatability	$s_r^2(\bar{\bar{x}})$ $s_r^2(\bar{\bar{z}})$	176162.6 28921.5	28921.5 176162.6	82548.7 339015.9	2026991 185531.9	1440.9 5484.1	70657.9 2250.4	2250.4 70657.9	41042.2 134961.8	4245 73382.2
Total number of measurements	P	50	50	26	48	42	52	52	26	48
Total number of trials	N_f	25	25	13	24	21	26	26	13	24
Correlation *r*	$cov(\bar{x},\bar{z})/(\sigma_{\bar{x}} \cdot \sigma_{\bar{z}})$	0.998	0.998	0.997	0.985	0.947	0.626	0.626	0.987	0.977
	Equ. $s_{r,limit}(z)$[1]		$2.25 \cdot z^{0.6174}$					$1.59 \cdot z^{0.7041}$		
	$s_{r,limit}(\bar{\bar{z}})$		359.3	457.9	369.8	117.4		279.7	399.5	289.4
	Equ. $s_R(z)$		$0.1z$					$0.1z$		
	$s_R(\bar{\bar{z}})$		370.4	548.6	388.1	118.1		154.5	256.3	162.1
	$s_R(\bar{\bar{z}})/\bar{\bar{z}}$		0.1	0.1	0.1	0.1		0.1	0.1	0.1
	$s(\bar{\bar{x}})$ $s(\bar{\bar{z}})$	6335.3 6098.4	6098.4 6335.3	10715.1 8455.2	7743.8 6409.8	271.9 1210.5	3211.4 336	336 3211.4	3328.9 4351.9	639.5 3320.8
	$C_1 = s(\bar{\bar{x}})/s(\bar{\bar{z}})$	1.039	0.963	1.267	1.208	0.225	9.557	0.105	0.765	0.193
	$C_0 = \bar{\bar{x}} - \bar{\bar{z}} \cdot s(\bar{\bar{x}})/s(\bar{\bar{z}})$	-141.6	136.3	-131.1	146.4	57.9	-86.5	9.1	239.7	135.6

[1] A rewritten form of the equation $s_{r,limit}(\bar{\bar{z}}) = \exp(\beta_0 + \beta_1 \cdot \log(\bar{\bar{z}}) + k \cdot \delta)$ after substitute the estimated parameter values.