National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Evaluation of equivalence of the MetOne BAM-1020 for the measurement of PM2.5 in ambient air

RIVM Letter report 2014-0078
T.L. Hafkenscheid | J. Vonk

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Evaluation of equivalence of the MetOne BAM-1020 for the measurement of $PM_{2.5}$ in ambient air

RIVM Letter report 2014-0078
T.L. Hafkenscheid | J. Vonk

Colophon

Theo Hafkenscheid, MIL-MLG
Jan Vonk, MIL-MLG

Contact:
Theo Hafkenscheid
MIL-MLG
theo.hafkenscheid@rivm.nl

This investigation has been performed by order and for the account of Ministry of Infrastructure and Environment, within the framework of within the framework of the monitoring of particulate matter in ambient air

This is a publication of:

National Institute for Public Health and the Environment
P.O. Box 1 | 3720 BA Bilthoven
The Netherlands
www.rivm.nl/en

Publiekssamenvatting

Evaluatie van de gelijkwaardigheid van de MetOne BAM-1020 voor het meten van PM2.5 in buitenlucht

Sinds een aantal jaren meet het Landelijk Meetnet Luchtkwaliteit (LML) van het RIVM op een groot aantal locaties de concentraties van PM2,5. Deze kleine fractie van fijn stof kan tot diep in de longen doordringen en is daardoor mogelijk schadelijker dan PM10. Sinds 2008 is PM2,5 vanwege Europese verplichtingen gemeten met behulp van een zogenoemde referentiemethode. Sinds 2010 wordt ook gemeten met een automatische monitor van het merk MetOne, type BAM-1020. Deze monitor genereert zelf continu resultaten, waardoor het minder arbeidsintensief is om de data uit te lezen.

Uit onderzoek van het RIVM blijkt dat de meetwaarden die met de BAM-1020 zijn verkregen gelijkwaardig zijn aan de referentiemeetwaarden. Daarom mag de BAM-1020 als gelijkwaardige methode worden gebruikt om PM2,5 te meten. Ook mogen de resultaten worden gebruikt voor zowel verplichte rapportages als om het publiek te informeren.

Om de meetresultaten van de monitoren voor deze doeleinden te mogen gebruiken, moest eerst worden aangetoond dat de resultaten overeenkomen met die van de referentiemethode. Hiertoe heeft het Landelijk Meetnet Luchtkwaliteit met de beide methoden op een aantal meetlocaties gemeten.

Trefwoorden:
Luchtkwaliteit, fijnstof, metingen, BAM-1020, equivalentie

Abstract

In the year 2009 the RIVM National Air Quality Monitoring Network (LML) has purchased 32 PM monitors of the type BAM-1020, manufactured by MetOne Instruments. These monitors are used to continuously measure PM2.5 at a number of monitoring sites of the LML. The aim of the monitoring is twofold:

- To provide information to the public about actual levels of PM2.5 in ambient air on an hourly basis.
- To replace (in part) the current PM2.5 monitoring network based on reference measurements. This network has been established primarily to measure concentrations of PM2.5 at 12 sites needed for the calculation of the Average Exposure Indicator (AEI) for the years 2009 to 2011. In total the PM2.5 network consists of 23 LML sites, and is supplemented by sites of DCMR Environmental Services Rijnmond and GGD Amsterdam.

The stand-alone operation of the BAM-1020 monitors requires the demonstration of equivalence between monitors and reference method. The requirements for the demonstration of equivalence are given in the Guide to the Demonstration of Equivalence of Ambient Air Monitoring Methods (GDE).
For the purpose of demonstrating equivalence parallel measurements have been performed with the BAM-1020 and reference equipment for a number of years at 14 locations.
The results of these measurements have been evaluated in accordance with the requirements of the GDE.

It is found that the measurement results of the BAM-1020 are equivalent to those of the reference method when a calibration is applied by using the following equation:

$$\text{BAM 1020 (cal)} = 0{,}91 \cdot \text{BAM 1020} - 1{,}6$$

The uncertainty of the measurement results of the BAM-1020 is 16%, which amply fulfils the requirement of the GDE (≤25%). Also the other requirements of the GDE for the demonstration of equivalence are met. This implies that the BAM-1020 may be used as an equivalent method in the LML for the measurement of PM2.5 in ambient air.

However, it is also found that:

- The relationship between the results is season-dependent. In the summer period the ratio of the results produced by the BAM-1020 and the reference results is relatively high compared to other seasons.
- The relationship between the results differs considerably from relationships obtained in other studies, where the results of the BAM-1020 and the reference method are in much closer agreement.

Both phenomena are the subject of ongoing research.

Keywords:
Air quality, particulate matter, measurements, BAM-1020, equivalence

Contents

Summary

In the year 2009 the RIVM National Air Quality Monitoring Network (LML) has purchased 32 PM monitors of the type BAM-1020, manufactured by MetOne Instruments. These monitors are used to continuously measure $PM_{2.5}$ at a number of monitoring sites of the LML.

In order to use the results produced by the BAM-1020 for reporting purposes and information of the public, the equivalence of the results with those obtained when using the reference method for measuring $PM_{2.5}$ (EN 14907:2005) has to be demonstrated. The requirements for the demonstration of equivalence are given in the Guide to the Demonstration of Equivalence of Ambient Air Monitoring Methods (GDE).

In order to demonstrate equivalence parallel measurements have been performed with the BAM-1020 and reference equipment for a number of years at 14 locations. Evaluation of the results of these measurements in accordance with the requirements of the GDE shows that the measurement results of the BAM-1020 are equivalent to those of the reference method when a calibration is applied by using the following equation:

$$BAM\ 1020\ (cal) = 0{,}91 \cdot BAM\ 1020 - 1{,}6$$

The uncertainty of the measurement results of the BAM-1020 is 16%, which amply fulfils the requirement of the GDE (≤25%). Also the other requirements of the GDE for the demonstration of equivalence are met. This implies that the BAM-1020 may be used as an equivalent method in the LML for the measurement of $PM_{2.5}$ in ambient air.

However, it is also found that:

- The relationship between the results is season-dependent. In the summer period the ratio of the results produced by the BAM-1020 and the reference results is relatively high compared to other seasons.
- The relationship between the results differs considerably from relationships obtained in other studies, where the results of the BAM-1020 and the reference method are in much closer agreement.

Both phenomena are the subject of ongoing research.

1 Introduction

In the year 2009 the RIVM National Air Quality Monitoring Network (LML) has purchased 32 PM monitors of the type BAM-1020, manufactured by MetOne Instruments. These monitors are used to continuously measure $PM_{2.5}$ at a number of monitoring sites of the LML. The aim of the monitoring is twofold:

- To provide information to the public about actual levels of $PM_{2.5}$ in ambient air on an hourly basis.
- To replace (in part) the current $PM_{2.5}$ monitoring network based on reference measurements (filter sampling + differential weighing). This network has been established primarily to measure concentrations of $PM_{2.5}$ at 12 sites needed for the calculation of the Average Exposure Indicator (AEI) for the years 2009 to 2011 [1]. In total the $PM_{2.5}$ network consists of 23 LML sites, and is supplemented by sites of DCMR Environmental Services Rijnmond and GGD Amsterdam.

Because of the difference in principle between the reference measurements and the measurements by the BAM-1020 (see below), the stand-alone operation of the BAM-1020 monitors requires the demonstration of equivalence between monitors and reference method [1]. The requirements for the demonstration of equivalence are given in the Guide to the Demonstration of Equivalence of Ambient Air Monitoring Methods (GDE, [2]).

This report describes the performance and the evaluation of the tests needed to demonstrate this equivalence, and draws conclusions about the operation and calibration of the BAM-1020.

2 Description of the BAM-1020

2.1 Measurement principle

The measurement of mass concentrations of particulate matter by the BAM-1020 is based on the attenuation of beta-rays when passing a thin layer of material, according to the following equation:

$$c = \frac{A.10^6}{Q.\Delta t.\mu} \ln \frac{I_0}{I}$$

where:

c = mass concentration of particulate matter, in µg/m^3
A = sampling area for particles (filter spot), in cm^2
Q = air sampling flow rate, in L/min
Δt = sampling time, in min
µ = mass absorption coefficient, in cm^2/g
I_0 = beta-ray count rate at the beginning (blank filter spot)
I = beta-ray count at the end (sampled filter spot).

The mass determination is factory-calibrated and is checked hourly during operation at the zero point (clean filter spot) and at the reference point (built-in reference foil). With the help of the generated data measured values at zero and reference point can be easily affiliated. They can be compared with any stability requirements (drift) and with the nominal value for the reference foil (factory setting).

2.2 Functionality of the instrument

The sampled air passes a $PM_{2.5}$ pre-separator, consisting of an EPA PM_{10} sampling inlet and a $PM_{2.5}$
Sharp Cut Cyclone (SCC), at a flow rate of 1,0 m³/h at ambient conditions and arrives via a sampling tube at the measuring system. The measuring system is operated with a sampling tube heater BX-830 (“Smart Inlet Heater”). The heater always slightly heats (6% of its capacity, about 10 W) and switches to full power by control of the maximum relative humidity at the filter tape (factory setting: 45 %rh). As soon as the relative humidity is 1 % below the nominal value, the heater returns to its low-power mode.

The particles are collected on a glass-fiber filter tape for subsequent radiometric measurement. One measurement cycle (incl. automatic check of the radiometric measurement) consists of
the following steps.

1. The initial count I_0 of the clean filter tape is performed at the beginning of the cycle for
a period of 8 minutes.
2. The filter tape is advanced 4 windows and the sampling (vacuum pumping) begins
at the spot at which I_0 was measured. Air is drawn through this spot on the filter
tape for approximately 42 minutes.
3. At the same time a second count I_1 is being performed at a point on the tape 4 windows back for a period of 8 minutes.
4. Subsequently, a third count I_2 is performed with the reference membrane placed over the same spot on the tape. Its result combined

with that of step 3 provides a verification of the mass determination at the reference point.

5. Eight minutes before the end of sampling time, another count I_{1x} is performed at the same spot on the filter tape. Measurements of I_1 and I_{1x} serve to monitor the stability at the zero point (instrument drift caused by varying external parameters such as temperature and relative humidity).
6. After sampling, the filter tape is moved back 4 windows to measure the beta-ray absorption through the section that has collected particles (I_3).
7. The concentration calculation is performed to complete the cycle. The next cycle begins with step 1.

2.3 Practical operation

The configuration of the BAM-1020 as used in The Netherlands National Air Quality Monitoring Network consists of the elements given in Table 1.

Table 1. Configuration of the BAM-1020.

Sampling inlet	EPA PM_{10} sampling head + sharp cut $PM_{2.5}$ cyclone
Measurement time	60 minutes, of which 42 sampling and twice 8 for beta counts
Negative values	Allowed up to -5 $\mu g/m^3$
Flow rate	16,7 lpm (1 m^3/h)
Flow control and concentration reporting	Actual conditions
Heater control	6% idle, RH controlled to full power (setpoint 45%rh, 175 W)
Instrument firmware	V5.0.10

All monitors are placed inside temperature-controlled monitoring cabins. The temperature set point of each cabin is 22 °C. A temperature ranging between 17 and 27 °C is permitted before an error message is sent to the central operational control desk at RIVM.
The sampling inlet of the monitor is positioned at a height of 4 m above ground level. For this purpose, the sampling tube is inserted through a hole in the roof of the cabin and connected with the measurement system.
When placing a monitor, extreme care is taken to ensure that the monitor is horizontally placed on a support, and that the sampling tube is aligned absolutely straight. This prevents uneven pressure on the underlying nozzle and vane assembly, so it can close properly and risk of puncturing the filter tape is minimized. Finally the part of the sampling tube inside the monitoring cabin is thermally insulated, in order to avoid condensation.

Before putting a monitor into operation, a background measurement is performed by sampling zero air for a period of 72 hours and measurement of the background signal. This measurement is performed at the monitoring site to account for potential differences in background radiation. The background obtained is used in the calculation of the actual $PM_{2.5}$ concentration values.

Each monitor is checked and maintained following a dedicated standard operating procedure [3].

Measurement data and results of auxiliary measurements are acquired through Lumina processors using a user-developed protocol. Data go through an initial (technical) validation procedure, before being validated on the basis of plausibility. All steps of these operations are fully documented in a standard operating procedure [4].

The resulting valid one-hour average measurement results are aggregated into daily average results when a minimum of 18 hourly average results is available for each calendar day.

3 Reference measurements

3.1 General

Reference measurements of $PM_{2.5}$ are performed in conformity with EN 14907: 2005 [5].
Air is sampled through an EU $PM_{2.5}$ sampling inlet at a flow rate of 2,3 m^3/h at ambient conditions and is passed through a quartz-fiber filter. The filter has been conditioned and weighed prior to sampling.
$PM_{2.5}$ is collected on the filter for a nominal period of 24 hours (one calendar day). After sampling the filter is again conditioned and weighed. The mass concentration of $PM_{2.5}$ is calculated from the mass difference of the filter weighings and the total volume of air that has been passed through the filter.

3.2 Sampling

The sampling is performed using Derenda PNS 16T-6.1 sequential samplers. Each sampler has 14 positions for sampled filters and one position for a so-called field blank.
Flow setting and control is effectuated by using a critical orifice, in combination with continuous measurements of ambient temperature and air pressure for flow regulation.
The sampler is equipped with a filter storage compartment, the temperature of which is kept at ≤ 23 °C by means of a Peltier cooler.

Preferably, the sampler is placed inside a monitoring cabin, although depending on the cabin type and/or space therein, it can also be placed outside. In all cases, the height of the sampling inlet above ground level is between 3,5 and 4,5 m.

The sampler is designed to operate for an uninterrupted period of 14 days. After each 14-day sampling period the cassettes containing the sampled filters and the blank filter are replaced with cassettes containing freshly prepared and weighed filters. At the same time, the information of the 14-day sampling period that is stored onto interchangeable SD cards is collected. Data have subsequently been transmitted to protected storage media for further treatment. This consists of:

- Checking of operational parameters (temperatures, flows rates, volumes etc.) for anomalies;
- Combination of data for volumes with corresponding mass data for $PM_{2.5}$ obtained from differential filter weighings;
- Plausibility checks of resulting concentrations, e.g., by comparison with PM_{10} concentrations and/or with $PM_{2.5}$ concentrations from other nearby monitoring sites.

3.3 Filter conditioning and weighing

The filters used for the collection of the sampled particles are quartz-fiber filters (Whatman QM-A) of 47 mm diameter. Before use the filters are conditioned and weighed according to the following protocol.

- Filters are pre-conditioned at near 100 % relative humidity at (20±1) °C by placing them in opened boxes in a desiccator for a period of 3 weeks at minimum.
- Filters are then conditioned and weighed in accordance with the (regular) procedure prescribed in [5]:
    - conditioning at (50±5) % relative humidity and (20±1) °C for at least 48 hours, followed by a first weighing

- conditioning at (50±5) % relative humidity and (20±1) °C for at least 12 (unloaded filters) or 24 hours (loaded filters), followed by a second weighing.
- the average of the two resulting masses is taken as the filter mass before sampling.

The regular conditioning is performed in a dedicated weighing room of which the relative air humidity and temperature are kept within the above ranges.

After sampling the filters are subjected to the same regular conditioning and weighing procedure as described above. Again, the average result of the two weighing is used, this time as the filter mass after sampling.

In both cases, the difference between the two subsequent filter masses needs to fulfill the requirements of [5]:

- A maximum difference of 40 μg before sampling, and
- A maximum difference of 60 μg after sampling.

If these requirements are not met, filters are again conditioned for a minimum of 12/24 hours at (50±5) % relative humidity and (20±1) °C, followed by a third weighing. If the criteria are not met after a maximum of four weighings, the result is discarded.

The sample field blank is treated in the same way as the sampled filters. The difference of the average weighing results for this filter before and after placement in the sequential sampler should not exceed 60 μg. When this is the case, the cause of the exceedance must be investigated.

4 Measurement strategy

The GDE [2] prescribes that a full campaign for the demonstration of equivalence consists of a minimum of 4 comparisons at a minimum of 2 sites, preferably in different climatic seasons with particular emphasis on the following variables, if appropriate:

- Composition of the PM fraction, notably high and low fractions of semi-volatile particles, to cover the maximum impact of losses of semi-volatiles;
- Air humidity and temperature (high and low) to cover any conditioning losses of semi-volatiles during the sampling process;
- Wind speed (high and low) to cover any dependency of inlet performance due to deviations from ideal behaviour as dictated by mechanical design, or deviations from the designated sampling flow rate.

In each campaign, a minimum of 40 valid data pairs shall be obtained.

In The Netherlands parallel measurements between the BAM-1020 and the reference method for $PM_{2.5}$ have been performed from 2009 onwards within the frame of the validation of the performance of the BAM-1020. Since a substantial number of monitoring sites are available at which reference measurements were already performed for the determination of the AEI 2009-2011, parallel measurements have largely been performed at these sites. The sites are listed in Table 2.

Table 2. Sites for parallel measurements of $PM_{2.5}$.

Site code	*Site name*	*Site characteristics*
131	Vredepeel	Rural
230	Biest-Houtakker	Rural
240	Breda Tilburgseweg	Traffic
241	Breda Bastenakenstraat	Urban
247	Veldhoven	Urban
433	Vlaardingen	Traffic
444	De Zilk	Rural
538	Wieringerwerf	Rural
636	Utrecht Kardinaal de Jongweg	Traffic
643	Utrecht Griftpark	Urban
644	Cabauw	Rural
741	Nijmegen Graafseweg	Traffic
742	Nijmegen De Ruyterstraat	Urban
821	Enschede	Urban
934	Kollumerwaard	Rural

At these sites parallel measurements have been performed on a continuous basis. As can be seen from the information in Table 1, sites have been selected such that all site types are adequately represented.

Additional information has been added from shorter-term parallel measurements at the following sites:

- 547, Hilversum, urban;
- 641, Breukelen, traffic;
- 738, Wekerom, rural.

For the years 2009 and 2010, data available do not cover all seasons and station types with sufficient representativity. In 2009 measurements with the BAM-1020 have started sequentially at the various sites, in 2010 the number of data available for the reference measurements were unevenly spread over the seasons because of problems with the consistency of the Whatman QM-A filters. Consequently, the evaluation of equivalence has been performed with the results collected in the years 2011 to 2013.

By choosing this strategy, RIVM is preluding the requirement of TS 16450 [6] that for a yearly evaluation of equivalence/uncertainty of an automated monitoring system for particulate matter the results of parallel measurements over a 3-year period may be pooled.

5 Evaluation of equivalence

5.1 Formation of a consistent data set

From the data, first, missing values have been removed to form a consistent data set for the above-mentioned stations for the years 2011, 2012 and 2013. Subsequently, potential outlying results have been (tentatively) identified by using the following procedure.

- For all data pairs the ratios of the results of the BAM-1020 and the reference results have been calculated.
- For the resulting ratios the geometric mean (GM) and geometric standard deviation (GSD) have been calculated.
- From the GM and GSD the ±3σ-interval of the ratios has been calculated.
- Ratios outside of this interval have been marked.
- Subsequently, all ratios have been sorted according to the dates at which the measurements have been performed.
- For dates at which only a single ratio is outside the ±3σ-interval, the corresponding data pair has been removed.
- For dates with more than one outlying ratio, the magnitudes of and relationships between the results (e.g., from nearby stations) has been examined before deciding on the removal of the data. This is to exclude common effects of, e.g., the "PM climate" on the measurement results (the BAM-1020 is equipped with a different sampling inlet than the reference sampler).
- The remaining results have been plotted: BAM-1020 at the y-axis, reference at the x-axis. The resulting plot has been visually examined for further outlying values by eye-ball statistics.

After removal of outlying values, 14703 data pairs remain for further evaluation.

5.2 Evaluation of equivalence

The evaluation of equivalence has been performed using the Equivalence Tool v. 2.10, developed by Ruben Beijk [7]. This Excel workbook automatically performs the necessary calculations when a consistent data set has been entered. In addition, the tool permits the filtering of data, e.g., on year, site and site type when the information for these variables is available.

The "macro" characteristics of the data set are given in Table 3. The plot of results of the BAM-1020 versus those of the reference method is given in Figure 1.

Table 3. Macro characteristics of the data set for evaluation of equivalence.

Number of data pairs	14703
– *Winter season*	3661
– *Spring season*	3668
– *Summer season*	3829
– *Fall season*	3575
– *Rural sites*	5451
– *Urban sites*	5469
– *Traffic sites*	3783
Mean of reference results (μg/m^3)	13,8
Mean of BAM-1020 results (μg/m^3)	16,8
Number of reference results above the limit value (30 μg/m^3)*	1336
Number of reference results above 18 μg/m^3	3532

*The GDE uses a "pseudo"-daily limit value for $PM_{2.5}$ of 30 μg/m^3.

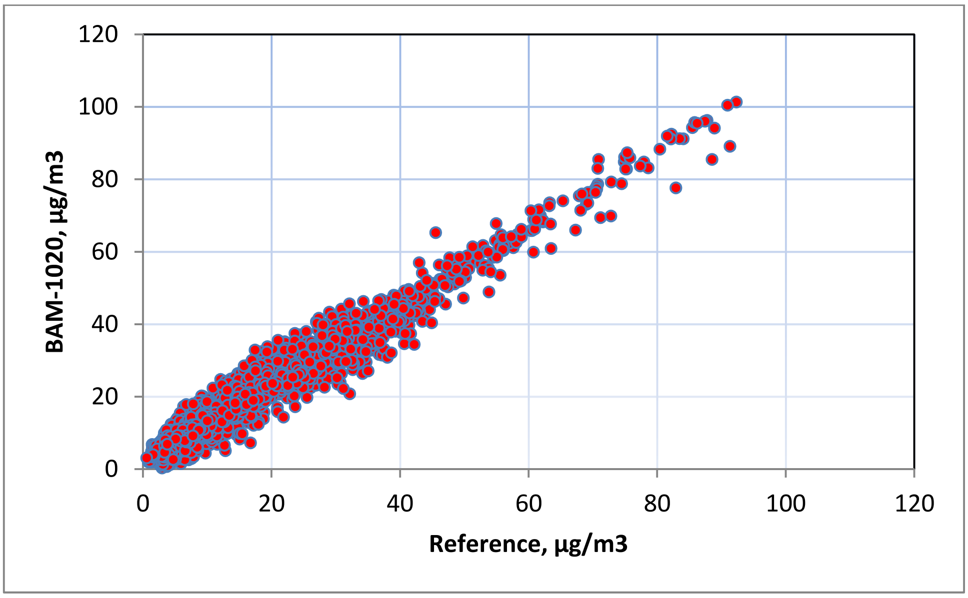

Figure 1. Plot of results of the BAM-1020 vs. reference results.

A number of observations may be made when studying the table and figure in more detail.

- The ratio between the mean results of the BAM-1020 and reference method is 1,22. This is much higher than expected on the basis of already known results from other studies (e.g., [8]). The ratio is also significantly higher than that found by other networks in The Netherlands.
- The relationship between the results of the BAM-1020 and the reference method appears to be slightly non-linear. At low concentration levels the slope is somewhat higher than at higher concentrations.
- The fraction of reference data with concentrations above 18 µg/m^3 is 0,24, which meets the requirement of the GDE (≥20%).

The results of the equivalence evaluation are given in Table 4.

Table 4. Results of the evaluation of equivalence.

Slope *b* of BAM-1020 *vs*. reference	1,087
Uncertainty of *b*	0,002
Intercept *a* of BAM-1020 *vs*. reference (µg/m^3)	1,817
Uncertainty of *a* (µg/m^3)	0,036
Uncertainty of reference results (µg/m^3)	1,0
Random uncertainty (µg/m^3)	2,63
Bias at the limit value of 30 µg/m^3 (µg/m^3)	4,43
Combined uncertainty (µg/m^3)	5,16
Relative expanded uncertainty (%, 95% confidence)	34,4

The evaluation shows that the results measured by the BAM-1020 do not meet the uncertainty requirement of ≤ 25%.
In addition, the slope of the relationship of the results of the BAM-1020 and the reference results is significantly different from 1 and is outside the range of 0,98 to 1,02.
Lastly, the intercept of the relationship of the results of the BAM-1020 and the reference results is significantly different from 0 and is outside the range of -1,0 to 1,0.

Based on these observations, the measurement results of the BAM-1020 shall be corrected (calibrated) both for slope and intercept. The calibration function to be used is

$$BAM\ 1020\ (cal) = 0{,}92 \cdot BAM\ 1020 - 1{,}7$$

The additional uncertainty associated with the calibration is 0,070 µg/m^3.
After applying this calibration, the results of the equivalence test are as follows (Table 5).

Table 5. Results of the evaluation of equivalence after calibration.

Slope *b* of BAM-1020 *vs*. reference	0,998
Uncertainty of *b*	0,002
Intercept *a* of BAM-1020 *vs*. reference (µg/m^3)	0,001
Uncertainty of *a* (µg/m^3)	0,033
Uncertainty of reference results (µg/m^3)	1,0
Random uncertainty (µg/m^3)	2,39
Bias at the limit value of 30 µg/m^3 (µg/m^3)	-0,06
Combined uncertainty (µg/m^3)	2,39
Relative expanded uncertainty (%, 95% confidence)	15,9

After calibration, the results of the BAM-1020 amply meet the uncertainty requirement of ≤ 25%.

The next step in the demonstration of equivalence is to split the data set into several subsets and repeat the evaluation. The first is a subset with reference results ≥ 18 µg/m^3. The results are given in Table 6.

Table 6. Results of the evaluation of equivalence after calibration for the subset of data with reference results ≥ 18 µg/m^3.

Number of data pairs	3532
Slope *b* of BAM-1020 *vs*. reference	0,999
Uncertainty of *b*	0,005
Intercept *a* of BAM-1020 *vs*. reference (µg/m^3)	-0,472
Uncertainty of *a* (µg/m^3)	0,149
Uncertainty of reference results (µg/m^3)	1,0
Random uncertainty (µg/m^3)	3,20
Bias at the limit value of 30 µg/m^3 (µg/m^3)	-0,49
Combined uncertainty (µg/m^3)	3,24
Relative expanded uncertainty (%, 95% confidence)	21,6

From this evaluation it can be seen that the calibrated results of the BAM-1020 for reference concentrations ≥ 18 µg/m^3 meet the 25% uncertainty requirement.

The further investigation of the equivalence comprises the testing for subsets of individual monitoring sites for which parallel measurements have been performed during a prolonged period.

The results of the evaluation for individual sites, with equivalence testing after calibration, are given in Table 7. It is noted that for all sites the uncertainty meets the 25% criterion.

When plotting the regression lines (for the sake of clarity individual data points are not shown) the results in Figure 2 are obtained. The orange lines denoted CI+ and CI- represent the 95% confidence interval of the regression of the full data set. From the graph it is seen that the results for site 230 Biest-Houtakker

are outside the 95% CI for results approximately above 50 µg/m^3. All other lines fall (well) within the confidence interval.
An investigation into possible causes for this deviating behaviour at site 230 has revealed that from February 3, 2011 the reference sampler has not functioned properly.
For this reason, the evaluation of equivalence has been repeated, this time without the results for site 230.

The "macro" characteristics of the new data set are given in Table 8.
Again, the requirement for 20% of the reference results to be ≥ 18 µg/m^3 is met.
The results of the new equivalence evaluation are given in Table 9.

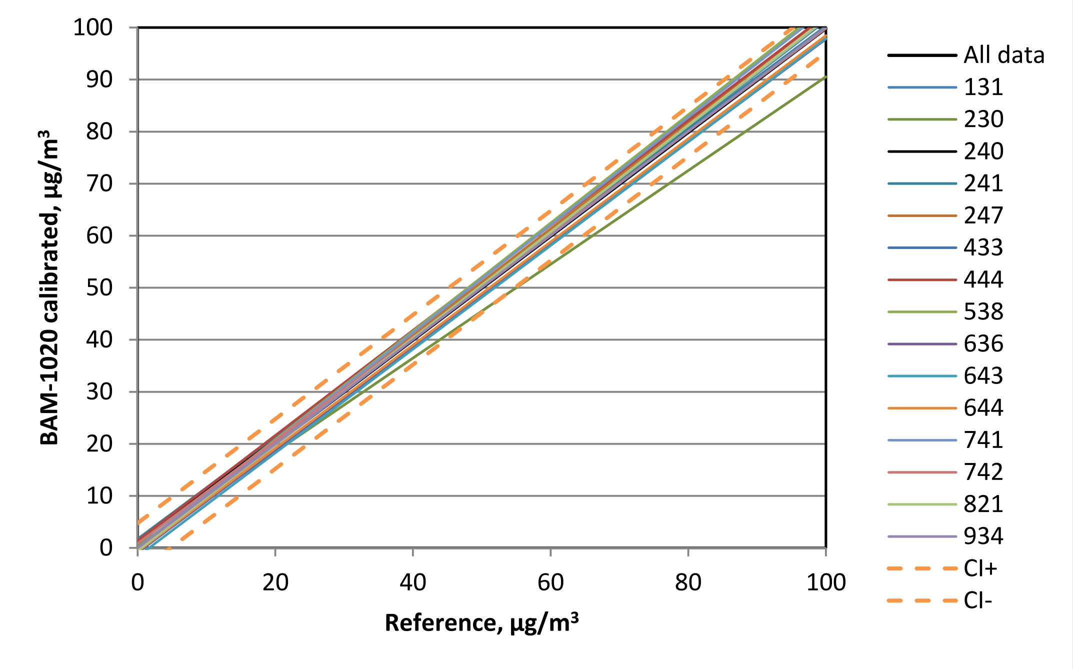

Figure 2. Regression lines for individual sites with 95% confidence interval.

Table 8. Macro characteristics of the data set for evaluation of equivalence without site 230.

Number of data pairs	13782
– *Winter season*	3394
– *Spring season*	3430
– *Summer season*	3633
– *Fall season*	3325
– *Rural sites*	4530
– *Urban sites*	5469
– *Traffic sites*	3783
Mean of reference results (µg/m^3)	13,6
Mean of BAM-1020 results (µg/m^3)	16,6
Number of reference results above the limit value (30 µg/m^3)*	1204
Number of reference results above 18 µg/m^3	3218

*The GDE uses a "pseudo"-daily limit value for $PM_{2.5}$ of 30 µg/m^3.

Table 9. Results of the evaluation of equivalence without site 230.

Slope *b* of BAM-1020 *vs*. reference	1,099
Uncertainty of *b*	0,002
Intercept *a* of BAM-1020 *vs*. reference (µg/m^3)	1,752
Uncertainty of *a* (µg/m^3)	0,037
Uncertainty of reference results (µg/m^3)	1,0
Random uncertainty (µg/m^3)	2,61
Bias at the limit value of 30 µg/m^3 (µg/m^3)	4,71
Combined uncertainty (µg/m^3)	5,38
Relative expanded uncertainty (%, 95% confidence)	35,9

Again, the evaluation shows that the results measured by the BAM-1020 do not meet the uncertainty requirement of ≤ 25%. In addition, the slope of the relationship of the results of the BAM-1020 and the reference results is significantly different from 1 and is outside the range of 0,98 to 1,02. Lastly, the intercept of the relationship of the results of the BAM-1020 and the reference results is significantly different from 0 and is outside the range of -1,0 to 1,0.

Based on these observations, the measurement results of the BAM-1020 shall be corrected (calibrated) both for slope and intercept. The new calibration function to be used is

$$BAM\ 1020\ (cal) = 0{,}91 \cdot BAM\ 1020 - 1{,}6$$

This function is marginally different from the function obtained when results for site 230 are included.
The additional uncertainty associated with the calibration is 0,072 µg/m^3.

After applying this new calibration, the results of the equivalence test are as follows (Table 10).

Table 10. Results of the evaluation of equivalence without site 230 after calibration.

Slope *b* of BAM-1020 *vs*. reference	0,997
Uncertainty of *b*	0,002
Intercept *a* of BAM-1020 *vs*. reference (µg/m^3)	0,027
Uncertainty of *a* (µg/m^3)	0,033
Random uncertainty (µg/m^3)	2,33
Bias at the limit value of 30 µg/m^3 (µg/m^3)	-0,05
Combined uncertainty (µg/m^3)	2,33
Relative expanded uncertainty (%, 95% confidence)	15,5

After this new calibration, the results of the BAM-1020 again amply meet the uncertainty requirement of ≤ 25%. The uncertainty is slightly lower than that obtained for the data including site 230.

Evaluation of the data set with reference results ≥ 18 µg/m^3 leads to a similar result to that obtained for the original data set (see Table 11).

Table 11. Results of the evaluation of equivalence without site 230, after calibration, for the subset of data with reference results ≥ 18 μg/m³.

Number of data pairs	3218
Slope *b* of BAM-1020 *vs*. reference	1,000
Uncertainty of *b*	0,005
Intercept *a* of BAM-1020 *vs*. reference (μg/m³)	-0,505
Uncertainty of *a* (μg/m³)	0,151
Random uncertainty (μg/m³)	3,08
Bias at the limit value of 30 μg/m³ (μg/m³)	-0,51
Combined uncertainty (μg/m³)	3,12
Relative expanded uncertainty (%, 95% confidence)	20,8%

The results of the repeated evaluation for individual sites, with equivalence testing after calibration, are given in Table 12. It is noted that (again) for all sites the uncertainty meets the 25% criterion.

Table 7. Results of the evaluation of equivalence, after calibration, for individual sites with large data sets.

	131	*230*	*240*	*241*	*247*	*433*	*444*	*538*	*636*	*643*	*644*	*741*	*742*	*821*	*934*
Site type	R	R	T	U	U	T	R	R	T	U	R	T	U	U	R
Number of data pairs	980	921	982	1025	904	1048	956	956	1057	999	432	898	1051	871	969
Mean result BAM-1020 (µg/m^3)	17,5	19,1	19,3	19,7	16,5	16,7	15,3	14,3	17,3	15,4	15,4	18,6	18,3	15,5	13,9
Mean reference result (µg/m^3)	14,7	17,2	15,0	14,9	13,3	14,4	10,9	11,0	15,3	14,1	13,3	14,9	14,4	12,8	11,1
Mean result of BAM-1020 CM AC* (µg/m^3)	14,4	15,9	16,1	16,4	13,5	13,7	12,4	11,5	14,2	12,4	12,5	15,4	15,1	12,5	11,1
Ratio of means AC (BAM/reference)	0,981	0,926	1,068	1,102	1,019	0,951	1,141	1,047	0,928	0,880	0,940	1,035	1,050	0,979	1,003
Random uncertainty AC (µg/m^3)	2,40	2,15	2,25	2,34	2,76	2,26	2,25	1,72	1,88	2,38	2,04	2,31	2,03	1,94	1,57
Bias at limit value AC (µg/m^3)	0,01	-2,54	0,88	1,30	0,64	-0,29	1,71	1,27	-1,29	-1,76	-0,94	1,04	0,56	0,10	0,03
Combined uncertainty AC (µg/m^3)	2,40	3,32	2,41	2,67	2,83	2,28	2,83	2,14	2,28	2,96	2,24	2,53	2,10	1,94	1,57
Relative expanded uncertainty AC (%)	16,0	22,2	16,1	17,8	18,9	15,2	18,8	14,3	15,2	19,7	15,0	16,9	14,0	12,9	10,5

**AC means “after calibration” (*0,92 – 1,7).*

Table 12. Results of the evaluation of equivalence without site 230, after calibration, for individual sites with large data sets.

	131	*240*	*241*	*247*	*433*	*444*	*538*	*636*	*643*	*644*	*741*	*742*	*821*	*934*
Site type	R	T	U	U	T	R	R	T	U	R	T	U	U	R
Number of data pairs	980	982	1025	904	1048	956	956	1057	999	432	898	1051	871	969
Mean result BAM-1020 (µg/m^3)	17,5	19,3	19,7	16,5	16,7	15,3	14,3	17,3	15,4	15,4	18,6	18,3	15,5	13,9
Mean reference result (µg/m^3)	14,7	15,0	14,9	13,3	14,4	10,9	11,0	15,3	14,1	13,3	14,9	14,4	12,8	11,1
Mean result of BAM-1020 CM AC* (µg/m^3)	14,3	16,0	16,3	13,5	13,6	12,3	11,5	14,1	12,4	12,4	15,3	15,0	12,5	11,1
Ratio of means AC (BAM/reference)	0,976	1,062	1,095	1,014	0,946	1,136	1,043	0,924	0,876	0,936	1,029	1,044	0,975	0,999
Random uncertainty AC (µg/m^3)	2,37	2,22	2,31	2,72	2,24	2,22	1,70	1,85	2,35	2,01	2,28	2,00	1,91	1,54
Bias at limit value AC (µg/m^3)	-0,24	0,63	1,04	0,38	-0,53	1,44	1,01	-1,52	-1,99	-1,17	0,78	0,31	-0,15	-0,22
Combined uncertainty AC (µg/m^3)	2,38	2,30	2,53	2,75	2,30	2,65	1,98	2,40	3,08	2,33	2,41	2,02	1,92	1,56
Relative expanded uncertainty AC (%)	15,9	15,4	16,9	18,3	15,3	17,6	13,2	16,0	20,5	15,5	16,1	13,5	12,8	10,4

**AC means “after calibration” (*0,91 – 1,6).*

6 Further evaluation of performance of the BAM-1020

When performing the evaluation of equivalence, it has been found that, although the BAM-1020 after calibration meets all requirements for claiming equivalence, a remarkable phenomenon is observed when examining the results for the meteorological summer season (June, July, August). For these results the calibration using the function (*0,91 – 1,6) does not lead to an uncertainty ≤ 25% (see Table 13).

Table 13. Results of equivalence evaluation after calibration for all seasons.

	Winter	*Spring*	*Summer*	*Fall*
Number of data pairs	3394	3430	3633	3325
Mean result of BAM-1020 (μg/m^3)	18,8	19,9	12,0	16,2
Mean reference result (μg/m^3)	15,5	17,2	8,8	13,0
Slope *b* of BAM-1020 *vs*. reference	0,986	0,972	1,206	1,036
Uncertainty of *b*	0,003	0,004	0,008	0,004
Intercept *a* of BAM-1020 *vs*. reference (μg/m^3)	0,245	-0,211	-1,361	-0,325
Uncertainty of *a* (μg/m^3)	0,056	0,078	0,085	0,064
Random uncertainty (μg/m^3)	1,85	2,57	2,53	2,11
Bias at the limit value of 30 μg/m^3 (μg/m^3)	-0,18	-1,05	4,82	0,76
Combined uncertainty (μg/m^3)	1,85	2,78	5,44	2,25
Relative expanded uncertainty (%, 95% confidence)	12,4	18,5	36,3	15,0

The main cause is that the extrapolation to the level of the limit value leads to a strong bias of >4 μg/m^3.

When performing a regression analysis without calibration, first with data for the summer season only, and next without data for the summer season, the following results are obtained (Table 14).

These results confirm the difference in relationship of the results of the BAM-1020 and the reference results. In fact, the calibration function for the data leaving out the summer season is similar to that for the full data set (*0,91 – 1,6).

Table 14. Results of equivalence evaluation after calibration For all seasons excluding the summer and for the summer season.

	Summer excluded	*Summer*
Number of data pairs	10149	3633
Mean result of BAM-1020 (μg/m^3)	18,3	12,0
Mean reference result (μg/m^3)	15,2	8,8
Slope *b* of BAM-1020 *vs*. reference	1,090	1,337
Uncertainty of *b*	0,002	0,009
Intercept *a* of BAM-1020 *vs*. reference (μg/m^3)	1,698	0,158
Uncertainty of *a* (μg/m^3)	0,042	0,093
Random uncertainty (μg/m^3)	2,51	2,84
Bias at the limit value of 30 μg/m^3 (μg/m^3)	4,40	10,27
Combined uncertainty (μg/m^3)	5,06	10,66
Relative expanded uncertainty (%, 95% confidence)	33,8	71,0

When examining the two subsets with reference concentrations below and above 18 µg/m^3, a marked difference is observed in the ratios of the average results of the BAM-1020 and reference method. For results above 18 µg/m^3 the ratio is 1,14, for results below 18 µg/m^3 the ratio is 1,33.
In order to further investigate the possible cause of this phenomenon, both subsets have been split according to site type. This leads to the ratios for average results of the BAM-1020 and reference results shown in Table 15.

Table 15. Ratios of mean results of the BAM-1020 and the reference method, split into site type and concentration level.

Site type	*All data*	*Reference results < 18 µg/m^3*	*Reference results ≥ 18 µg/m^3*
Rural	1,263	1,377	1,146
Urban	1,237	1,330	1,146
Traffic	1,181	1,261	1,121

Apparently, the site type has an effect for low concentrations, but less so for higher concentrations.

When splitting the full data set according to meteorological season, a clearer picture emerges. The equivalence test results for each season are given in Table 16.

Table 16. Results of equivalence evaluation after calibration for low levels (excluding the summer season) and for rural sites in summer.

	Summer, rural sites	*Summer, other sites*
Number of data pairs	1151	2482
Mean result of BAM-1020 (µg/m^3)	11,2	12,3
Mean reference result (µg/m^3)	8,3	9,1
Slope *b* of BAM-1020 *vs*. reference	1,073	1,282
Uncertainty of *b*	0,012	0,011
Intercept *a* of BAM-1020 *vs*. reference (µg/m^3)	-0,274	-2,043
Uncertainty of *a* (µg/m^3)	0,119	0,113
Random uncertainty (µg/m^3)	2,15	2,66
Bias at the limit value of 30 µg/m^3 (µg/m^3)	1,91	6,41
Combined uncertainty (µg/m^3)	2,88	6,94
Relative expanded uncertainty (%, 95% confidence)	19,2	46,3

It is clear that in only the summer season problems exist in meeting the uncertainty criterion, caused by a “skewed” relationship between results of the calibrated BAM-1020 and reference results (slope 1,21, intercept -1,4). This coincides more or less with a majority of low measurement results. When splitting the summer season results in subsets < and ≥ 18 µg/m^3, no improvement is observed.

When examining site types the phenomenon appears to be most prominent for urban and traffic sites. For rural sites, when applying the calibration function (*0,91 – 1,6), the uncertainty passes the criterion of ≤ 25% (see Table 16).

In conclusion, it is found that problems exist particularly when measuring at urban and traffic sites in the summer season.

7 Discussion

The results for the measurement of $PM_{2.5}$ using the BAM-1020 in the Netherlands National Air Quality Monitoring Network (LML) may be claimed to be equivalent to those obtained when using the reference measurement method (filter sampling + differential weighing) provided that the raw results of the BAM-1020 are calibrated using the function

$$BAM\ 1020\ (cal) = 0{,}91 \cdot BAM\ 1020 - 1{,}6$$

However, results for the meteorological summer for urban and traffic sites do not meet the 25% uncertainty criterion when using this calibration function.

The latter phenomenon calls for further investigations into the nature of the relationship between the results provided by the BAM-1020 and the reference method under “summer” conditions.

In 2007-2008 RIVM, GGD Amsterdam and DCMR Environmental Services Rijnmond have performed a joint study into the comparability of measurement results of (e.g.) the BAM-1020 and reference method for $PM_{2.5}$ [9]. This study then led to the conclusion that both methods gave equivalent results even without the need for calibration of the results of the BAM-1020. The ratio of the mean results of the BAM-1020 and the reference method was 1,039. This is considerably lower than the ratios found in the present equivalence study. Unfortunately, only one data pair was obtained during the summer season. This previous study can therefore not be used to gain further insight into possible “summer season” effects.

Another relatively extensive data set is available from the GGD Amsterdam. This network measures $PM_{2.5}$ in the same way RIVM does, with instruments of the same type. When examining the data produced for parallel measurements over the same period (2011 to 2013) the same “summer season” phenomenon is observed. Results of the evaluation of equivalence are presented in Table 17 for their full data set and data sets for the summer period and all other periods separated.

Table 17. Results of equivalence evaluation of data from GGD Amsterdam.

	Full data set	*Summer*	*Summer excluded*
Number of data pairs	2794	746	2048
Mean result of BAM 1020 (µg/m^3)	17,2	13,0	18,7
Mean reference result (µg/m^3)	15,4	10,1	17,4
Slope *b* of BAM 1020 *vs*. reference	1,029	1,182	1,033
Uncertainty of *b*	0,004	0,016	0,005
Intercept *a* of BAM 1020 *vs*. reference (µg/m^3)	1,357	1,073	0,826
Uncertainty of *a* (µg/m^3)	0,087	0,192	0,101
Random uncertainty (µg/m^3)	2,77	2,67	2,68
Bias at the limit value of 30 µg/m^3 (µg/m^3)	2,22	6,54	1,81
Combined uncertainty (µg/m^3)	3,55	7,06	3,23
Relative expanded uncertainty (%, 95% confidence)	23,6	47,1	21,5

From the results of the evaluation it is also apparent that the results of the BAM 1020 and the reference method are in closer agreement than those for the RIVM network. For the results of the GGD Amsterdam the resulting calibration function would be:

$$BAM\ 1020\ (cal) = 0{,}97 \cdot BAM\ 1020 - 1{,}3$$

Since both networks use the instruments of the same type, this feature is remarkable, and is the subject of ongoing research.
Findings from other studies (e.g., [8]) also indicate that the results of the BAM 1020 and the reference method should be in much closer agreement than found in this evaluation.

8 Conclusions

The results for the measurement of $PM_{2.5}$ using the BAM-1020 in the Netherlands National Air Quality Monitoring Network may be claimed to be equivalent to those obtained when using the reference measurement method (filter sampling + differential weighing) provided that the raw results of the BAM-1020 are calibrated using the function

$$BAM\ 1020\ (cal) = 0{,}91 \cdot BAM\ 1020 - 1{,}6$$

However, it is found that:

- The relationship between the results is season-dependent. In the summer period the BAM-1020 produces relatively higher results compared to other seasons.
- The relationship between the results differs considerably from relationships obtained in other studies, where the results of the BAM-1020 and the reference method are in much closer agreement.

Both phenomena are the subject of ongoing research.

9 Acknowledgements

The authors wish to thank Guus Stefess and Hans Berkhout for their helpful comments, and Dave de Jonge (GGD Amsterdam) for making available their results of parallel measurements of PM2.5.

10 References

[1] EC Directive 2008/50/EC on ambient air quality and cleaner air for Europe.

[2] Guide to the Demonstration of Equivalence of Ambient Air Monitoring Methods.

[3] LM-O-203. Maintenance and checks of the BAM 1020 $PM_{2.5}$ monitor (in Dutch).

[4] IL-W-015. Validation of PM2.5 measured using the BAM-1020 (in Dutch).

[5] EN 14907:2005. Ambient air quality - Standard gravimetric measurement method for the determination of the PM2,5 mass fraction of suspended particulate matter.

[6] CEN TS 16450:2012. Ambient air - Automated measuring systems for the measurement of the concentration of particulate matter (PM10; PM2,5).

[7] Version 2.9 of the tool is available at: http://ec.europa.eu/environment/air/quality/legislation/assessment.htm

[8] TUEV Report 936/21209919/A:2010. Report on the suitability test of the ambient air quality measuring system BAM 1020 with $PM_{2.5}$ pre-separator from the company MetOne, Inc. for the measurement of $PM_{2.5}$.

[9] RIVM Report 680709001:2008. $PM_{2.5}$ performance test. Comparison of automated $PM_{2.5}$ monitoring equipment.

RIVM
Committed to *health and sustainability*