

Review of “Enhancing Low-Cost PM_{2.5} Sensor Reliability Through Multi-Model Calibration Against a Beta Attenuation Monitor” by Adira Prayoga et al.

Adira Prayoga et al., present an evaluation of a low-cost sensor against the BAM 1020 as a reference instrument. The authors investigate potential interference and biases in the low-cost sensor PM_{2.5} measurements and apply numerous calibration models to correct the low-cost sensor data relative to the BAM 1020. The authors conclude that machine learning models performed the best, and that there is a need to correct for environmental (namely RH) conditions to improve the low-cost sensor measurements.

While it is clear that the authors have put together a comprehensive study and done a lot of work, the novelty of the work is not clear. It is well-established that low-cost optical particle sensors need to be corrected for effects of relative humidity and that non-linear calibration methods perform better than a linear correction. The authors claim that a ‘comprehensive calibration framework is presented, however it is not clear to me what is the actual calibration framework is. For example, what of the presented models was found to perform the best and therefore be included in the framework? Random Forest? Support Vector Regression? While it may be useful to present and compare different calibration model approaches, the study needs to go a step further and articulate what approach is recommended and how it improves on previous work.

It is because of this lack of novelty that I am recommending that this work be rejected in its current format. The paper needs substantial work **to reframe the research and address some of the structural and analytical issues outlined below.**

Response to Reviewer 2 (General Comment)

We thank the reviewer for the careful reading of the manuscript and for the constructive feedback. We appreciate the reviewer’s recognition of the effort involved in conducting the experimental study and evaluating multiple calibration models. We also acknowledge the reviewer’s concern regarding the clarity of the novelty and the description of the proposed calibration approach.

We agree that the original manuscript did not sufficiently articulate the specific contribution of the study and that the use of the term “calibration framework” may have created ambiguity. In response to the reviewer’s comments, the manuscript has been **substantially revised and reframed.**

First, the terminology describing the study has been clarified. The manuscript no longer refers to the approach as a “comprehensive calibration framework.” Instead, it is now described as a **hybrid calibration approach combining linear bias correction and**

nonlinear machine-learning models, followed by an adaptive blending step. This change more accurately reflects the methodology implemented in the study.

Second, the revised manuscript now **clearly identifies the best-performing calibration approach and its practical implications**. The comparative evaluation of linear regression, Random Forest, Gradient Boosting, Support Vector Regression, and the hybrid adaptive-blend model has been retained, but the discussion now explicitly identifies the recommended model configuration based on performance metrics and robustness under high relative humidity conditions.

Third, the Results and Discussion sections have been reorganized to more clearly demonstrate how the proposed hybrid calibration strategy improves agreement between the low-cost sensor and the BAM reference measurements. The revised manuscript places greater emphasis on:

- bias characterization of the raw sensor measurements,
- environmental dependence of sensor errors,
- comparative performance of calibration models, and
- the practical implications of selecting nonlinear calibration approaches under high-humidity conditions.

Fourth, additional discussion has been added to place the results in the context of previous studies on low-cost PM_{2.5} sensor calibration and environmental compensation. This expanded discussion clarifies how the present study contributes empirical evidence from a tropical high-humidity environment and evaluates the relative performance of several calibration approaches under these conditions.

Finally, the conclusion has been revised to ensure that the claims made in the manuscript are fully supported by the presented analysis and that the recommended calibration strategy is clearly articulated.

We appreciate the reviewer's comments, which have helped us substantially improve the clarity, structure, and scientific positioning of the manuscript.

Further comments

Section 2.1: why are the GP2Y and SHS0 sensors described when only ZH03 sensor is seemingly tested section 3?

The BAM 1020 is utilized as a reference instrument for this study. While the BAM 1020 is a suitable choice, there are a couple limitations of the BAM 1020 that the authors have not mentioned or considered in their analysis.

Firstly, while the BAM 1020 does have FEM, this is only for 24-hr average measurements. Using it for 1-hr measurements comes with greater measurement uncertainty and the author need to consider this in their analysis.

Secondly, the BAM 1020 can also suffer from interference from moisture. While this is typically mitigated for with the smart heater, its not infallible, especially in locations with high ambient RH this can result in a negative interference due to moisture on the tape. As there is no information on the sampling location, it is hard to judge if this may have been important.

Response to Reviewer 2 (Further Comments)

We thank the reviewer for the detailed and constructive comments regarding the experimental design and the use of the BAM 1020 reference instrument. These observations have helped us improve the clarity and methodological transparency of the manuscript. Our responses to the specific comments are provided below.

Comment:

Section 2.1: why are the GP2Y and SHS0 sensors described when only ZH03 sensor is seemingly tested in Section 3?

Response:

We appreciate the reviewer highlighting this source of confusion. In the original manuscript, the GP2Y and SHS0 sensors were briefly introduced as part of the broader hardware design of the low-cost monitoring unit. However, only the ZH03A sensor was used in the calibration analysis presented in Section 3. To avoid ambiguity, the manuscript has been revised to clearly state that the calibration study focuses exclusively on the ZH03A sensor. The descriptions of GP2Y and SHS0 have been shortened and moved to the supplementary material, while Section 2.1 now explicitly indicates that the ZH03A sensor is the device evaluated in the co-location experiment.

Comment:

The BAM 1020 is utilized as a reference instrument. While the BAM 1020 is suitable, there are limitations that are not mentioned.

Response:

We thank the reviewer for highlighting these important considerations regarding the BAM 1020 reference instrument. The manuscript has been revised to acknowledge these limitations and clarify how they may influence the interpretation of the results.

First, we now explicitly note that the Federal Equivalent Method (FEM) designation of the BAM 1020 applies to 24-hour averaged measurements, whereas the present study utilizes hourly data. Hourly BAM measurements may therefore exhibit greater measurement uncertainty compared to the regulatory averaging period. A discussion of this limitation has been added to the Methods section to clarify that the BAM measurements are used as a practical reference for relative comparison rather than as a perfectly uncertainty-free ground truth.

Second, we acknowledge that beta attenuation monitors can experience measurement interference from high humidity conditions due to moisture accumulation on the filter tape. Although the BAM 1020 includes an inlet heater to reduce this effect, the correction is not always perfect, particularly in environments with persistently high ambient relative humidity. A discussion of this potential negative humidity interference has now been included in the manuscript.

In addition, the revised Methods section now provides clearer information on the **sampling location, environmental conditions, and co-location setup**, which allows readers to better assess whether humidity-related interference could influence the BAM measurements in this study.

These revisions improve transparency regarding the limitations of the reference instrument and provide appropriate context for interpreting the calibration results.

There is too much technical detail in the methods section, especially 2.2 and 2.3

In section 2.3, there is too much irrelevant detail on the data collection and communication procedures. Key information that is missing include the actual sampling location, length of the sampling, dates (time of year).

Section 2.4. the removal of outliers based solely on statistics is in my opinion problematic. High concentration peaks can be real, and capturing these is important for a monitoring network from a public health perspective. Therefore, these should be left in unless there is a good operational reason (e.g., flow error, etc) as the correction model need to be able to handle these.

Section 3.1 and 3.2 could be significantly shortened and focused on the key findings. A short description of the meteorology would be useful as well during the study, to give context.

There is little reference to previous work in Section 3, notably in Section 3.3 where the authors may wish to review previous studies approaches to addressing hygroscopic growth using κ -Köhler theory.

Section 3.4. it is difficult to assess the performance of the models based on the presented data as it is not clear if training and validation datasets was used. that is, was the model performance tested on a separate (validation) dataset not used to develop the model? For example, was the (admittedly very short) time series in Fig 13 used to develop the calibration model or different dataset? That a machine learning model performs well on the dataset used to build is not surprising, and the real test is how it performs on different dataset.

Response to Reviewer 2 (Specific Comments)

We thank the reviewer for the detailed and constructive comments. The manuscript has been substantially revised to improve clarity, reduce unnecessary technical detail, and strengthen the methodological transparency of the analysis. Our responses to the specific comments are provided below.

Comment:

There is too much technical detail in the methods section, especially Sections 2.2 and 2.3.

Response:

We appreciate this observation and agree that the original manuscript contained excessive implementation-level detail that detracted from the scientific focus of the study. Sections 2.2 and 2.3 have been substantially revised and shortened. Descriptions related to hardware architecture, communication protocols, and system-level engineering details have been removed from the main text and either condensed or moved to the supplementary material. The revised Methods section now focuses primarily on the measurement setup, calibration methodology, and data analysis procedures relevant to the scientific objectives of the study.

Comment:

In Section 2.3, there is too much irrelevant detail on the data collection and communication procedures. Key information missing includes the actual sampling location, sampling duration, and dates.

Response:

We thank the reviewer for highlighting this important omission. The Methods section has been revised to include a clear description of the **sampling location, site classification, and the temporal coverage of the co-location campaign**. The revised text now specifies the monitoring location, sampling height, and the start and end dates of the measurement period, as well as a brief description of the environmental conditions during the campaign. At the same time, the detailed descriptions of data communication

procedures have been significantly reduced to maintain focus on the measurement and calibration aspects of the study.

Comment:

Section 2.4: removal of outliers based solely on statistics is problematic because high concentration peaks may be real and important.

Response:

We appreciate the reviewer's important point regarding the potential relevance of high concentration events. The manuscript has been revised to clarify the purpose and implementation of the outlier filtering procedure. In the revised version, the filtering approach is applied conservatively and designed to remove only clear statistical anomalies that are inconsistent with surrounding temporal behavior and likely attributable to measurement artefacts rather than genuine pollution events.

Furthermore, the manuscript now explicitly states that the outlier filtering was applied only to the **training dataset**, while the independent test dataset was left unfiltered. This ensures that model evaluation includes the full range of observed concentrations, including peak pollution events. A brief sensitivity discussion has also been added to explain that the calibration models were capable of capturing elevated PM_{2.5} concentrations.

Comment:

Sections 3.1 and 3.2 could be significantly shortened and focused on key findings. A short description of the meteorology during the study would be useful.

Response:

We agree with the reviewer and have substantially revised these sections. The descriptive reporting of raw measurement statistics has been reduced, and the sections now focus more directly on the key findings relevant to sensor bias and environmental effects. In addition, a short description of the **meteorological conditions during the monitoring period** has been added to provide context for interpreting the observed sensor behavior, particularly with respect to temperature and relative humidity variability.

Comment:

There is little reference to previous work in Section 3, particularly regarding hygroscopic growth and κ -Köhler theory.

Response:

We thank the reviewer for this suggestion. The Results and Discussion section has been expanded to include additional references to previous studies addressing humidity effects on optical particle sensors. In particular, the discussion now includes references to hygroscopic growth mechanisms and previous approaches for humidity correction, including studies based on κ -Köhler theory. These additions provide stronger scientific context for interpreting the environmental sensitivity observed in the low-cost sensor measurements.

Comment:

Section 3.4: it is difficult to assess the performance of the models because it is not clear whether a separate validation dataset was used.

Response:

We appreciate the reviewer highlighting the need for clearer explanation of the model evaluation procedure. The revised manuscript now provides a detailed description of the **training and validation strategy** used in the study. The dataset was divided chronologically, with the first 70% of observations used for model training and hyperparameter optimization, and the remaining 30% reserved as an independent test dataset. This sequential split prevents temporal leakage and ensures that model performance is evaluated using data that were not used during model development.

The manuscript now clearly states that the time series example shown in Figure 13 is used only for **visual illustration of model behavior**, and that the reported model performance metrics are based on the independent test dataset rather than the training data.

These revisions improve transparency and ensure that the calibration models are evaluated under conditions that better reflect real-world deployment scenarios.