Supplementary Material for

Impact and Optimization of Calibration Conditions for Air Quality Sensors in the Long-term Field Monitoring

Han Mei[1], Peng Wei[2*], Meisam Ahmadi Ghadikolaei[1], Nirmal Kumar Gali[1], Ya Wang[1], Zhi Ning[1,*]

[1]Division of Environment and Sustainability, The Hong Kong University of Science and Technology, Hong Kong, China

[2]College of Geography and Environment, Shandong Normal University, Jinan, China

[*]Correspondence to Zhi Ning (zhining@ust.hk) , Peng Wei (pengwei@sdnu.edu.cn)

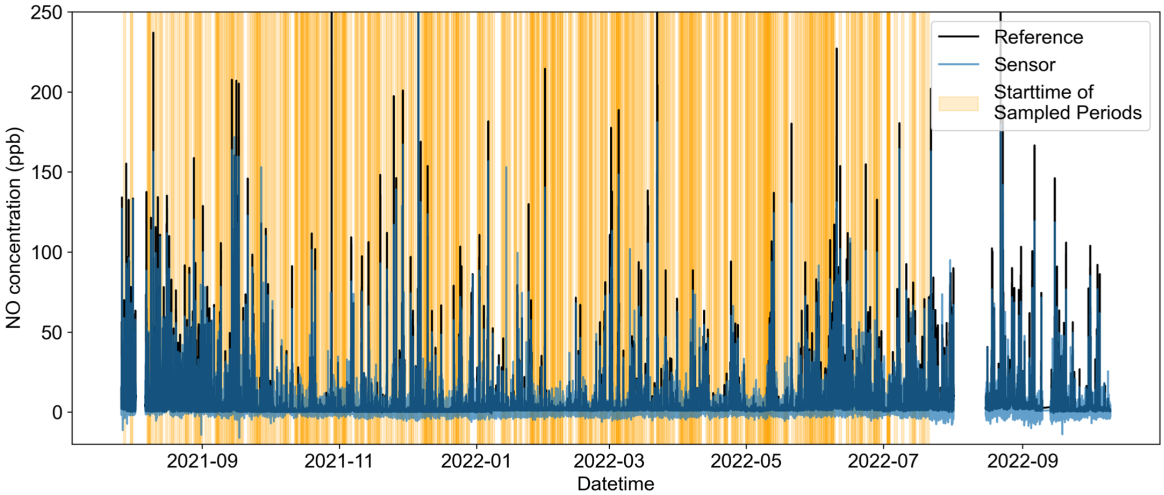

Figure S1. The start times of 500 randomly selected calibration periods for the MAS1 NO sensor.

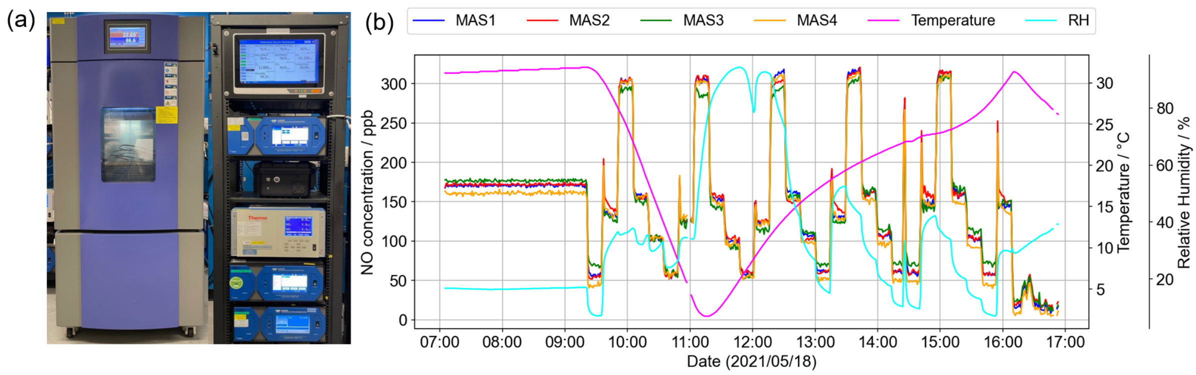

Figure S2. (a) Laboratory environmental chamber setup and (b) the response of 4 MASs' NO sensors under multiple point concentrations in laboratory temperature and humidity test.

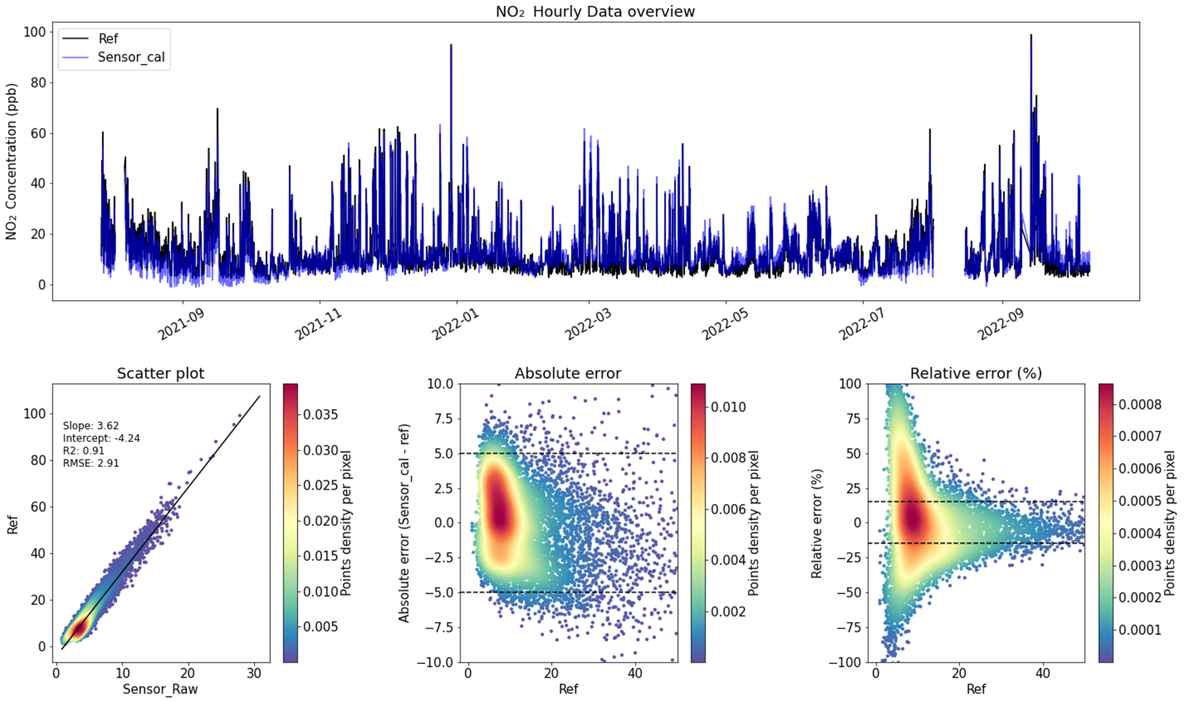

Figure S3. MAS1 NO_2 sensor calibrated data overview.

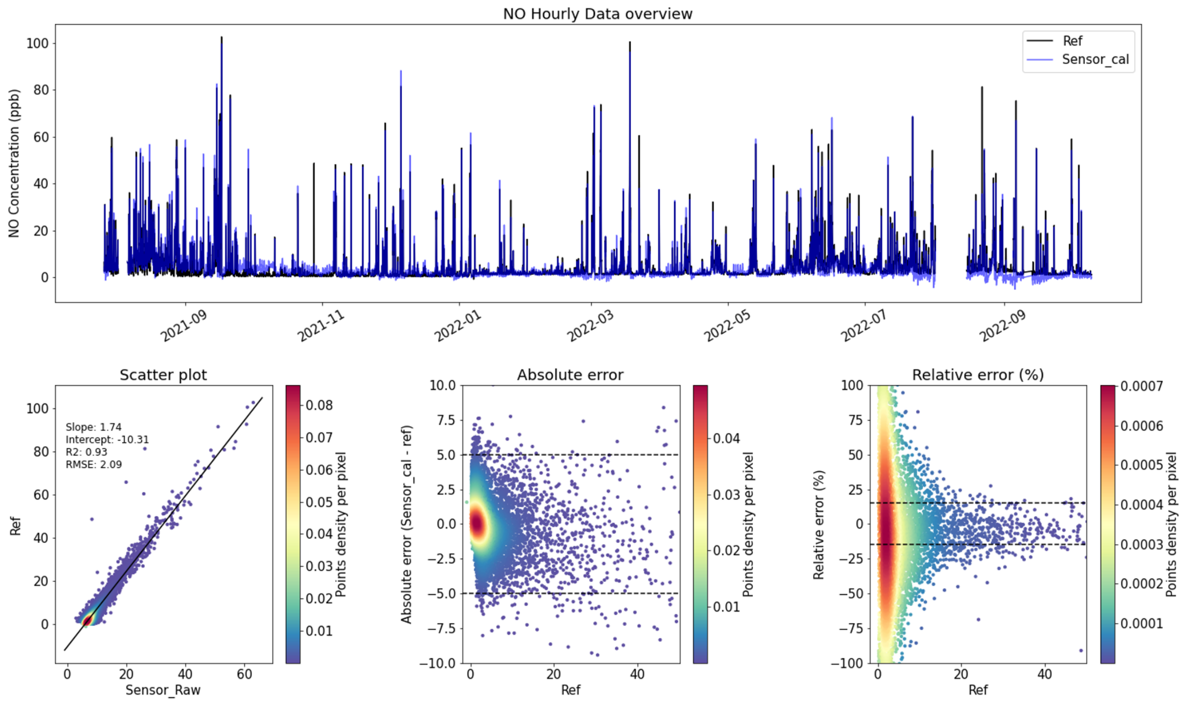

Figure S4. MAS1 NO sensor calibrated data overview.

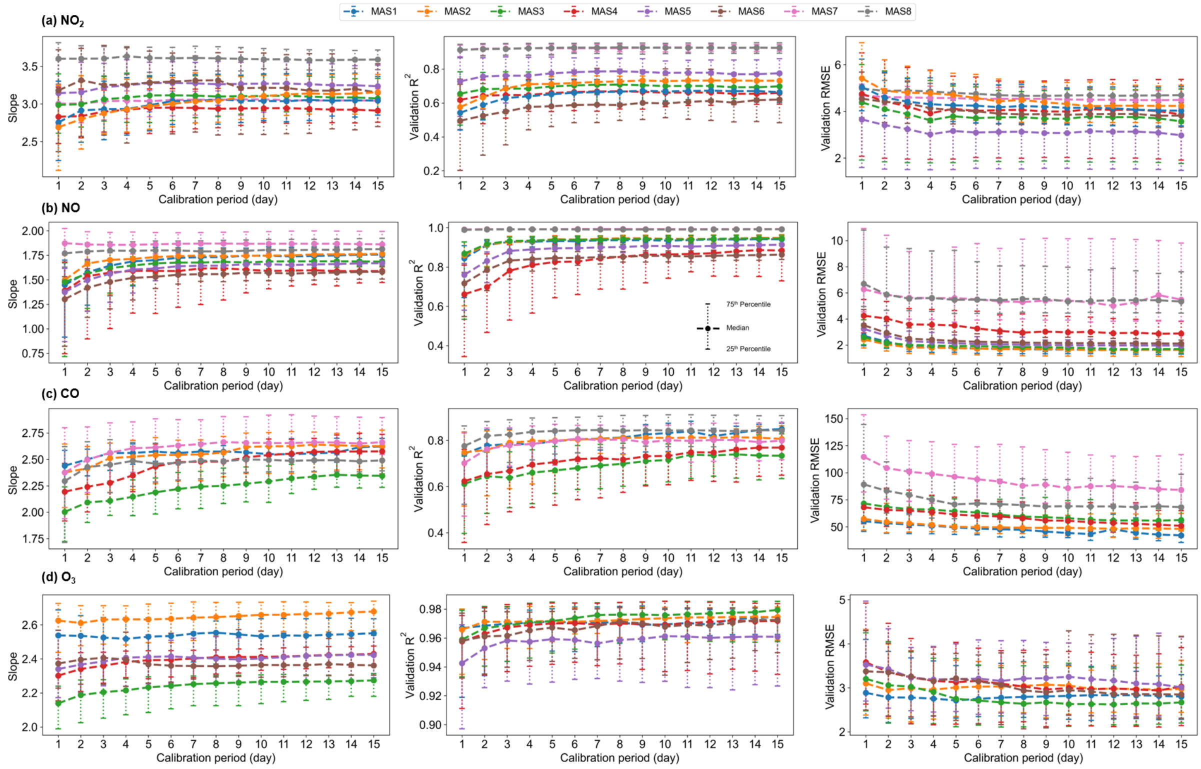

Figure S5. The potential range of the calibration slope, validation R^2 and RMSE for a given calibration period for each MAS (a) NO_2, (b) NO, (c) CO, and (d) O_3 sensors. Different colored lines represent the results of different MAS units. The vertical error bar is the 25%–75% distribution of the results under different calibration periods.

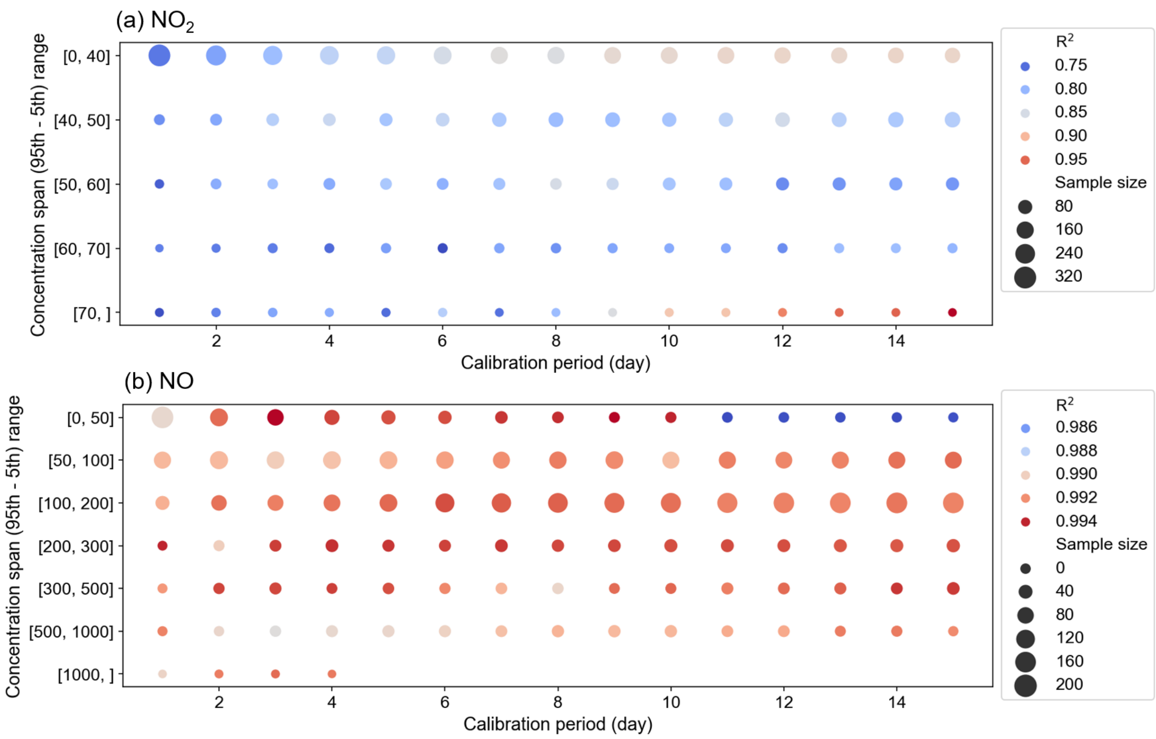

Figure S6. (a) NO_2 and (b) NO bubble plot of median R^2 of MAS units 7 and 8 (located in the high-concentration region (Shanghai)) and two factors: calibration period and concentration span. The size of the bubbles represents the number of samples. The color of these bubbles represents the median R^2 values in corresponding categories. Red represents higher R^2 values, while blue represents lower R^2 values.

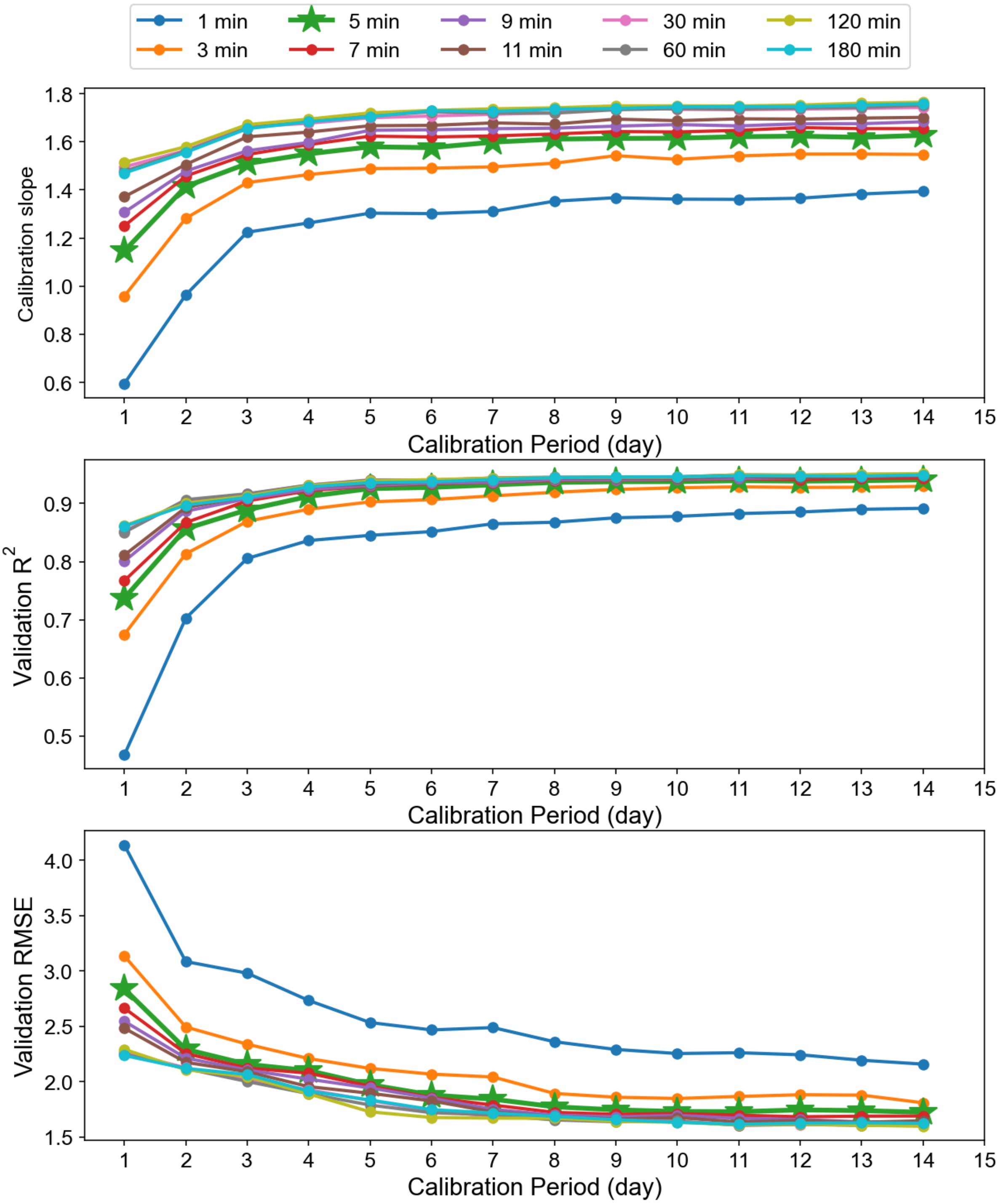

Figure S7. Calibration slope median, the R^2 median of the validation set, and the RMSE median of the validation set under different time averaging for a given calibration period for the MAS1 NO sensor.

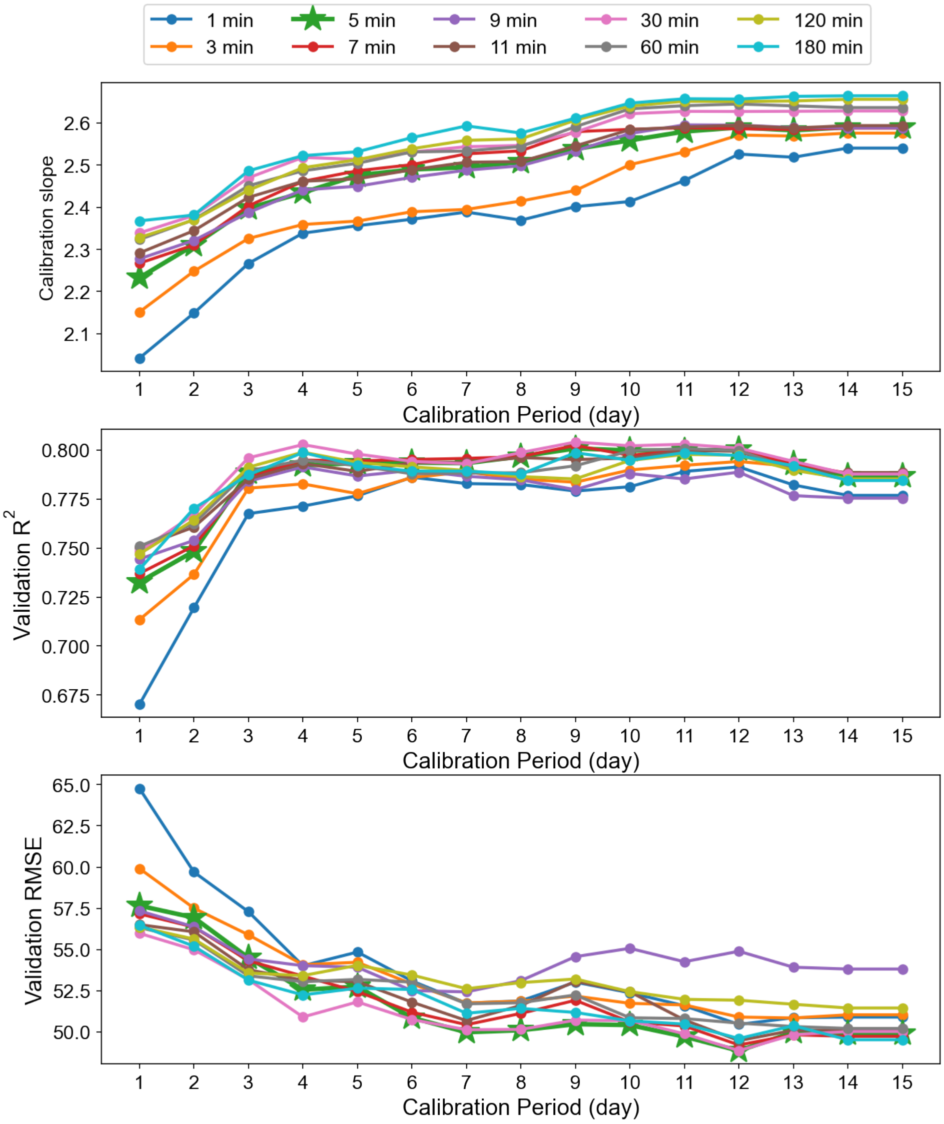

Figure S8. Calibration slope median, the R^2 median of the validation set, and the RMSE median of the validation set under different time averaging for a given calibration period for the MAS1 CO sensor.

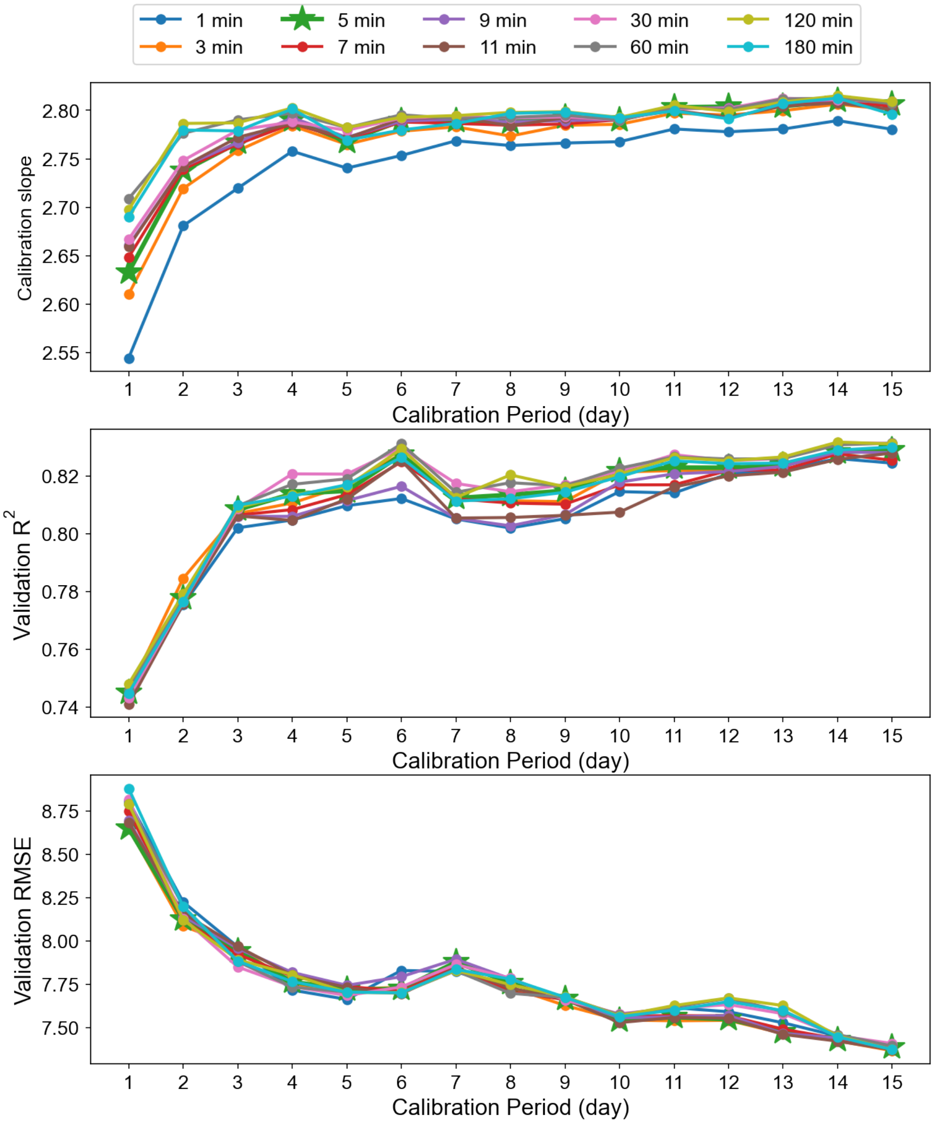

Figure S9. Calibration slope median, the R^2 median of the validation set, and the RMSE median of the validation set under different time averaging for a given calibration period for the MAS1 O_3 sensor.

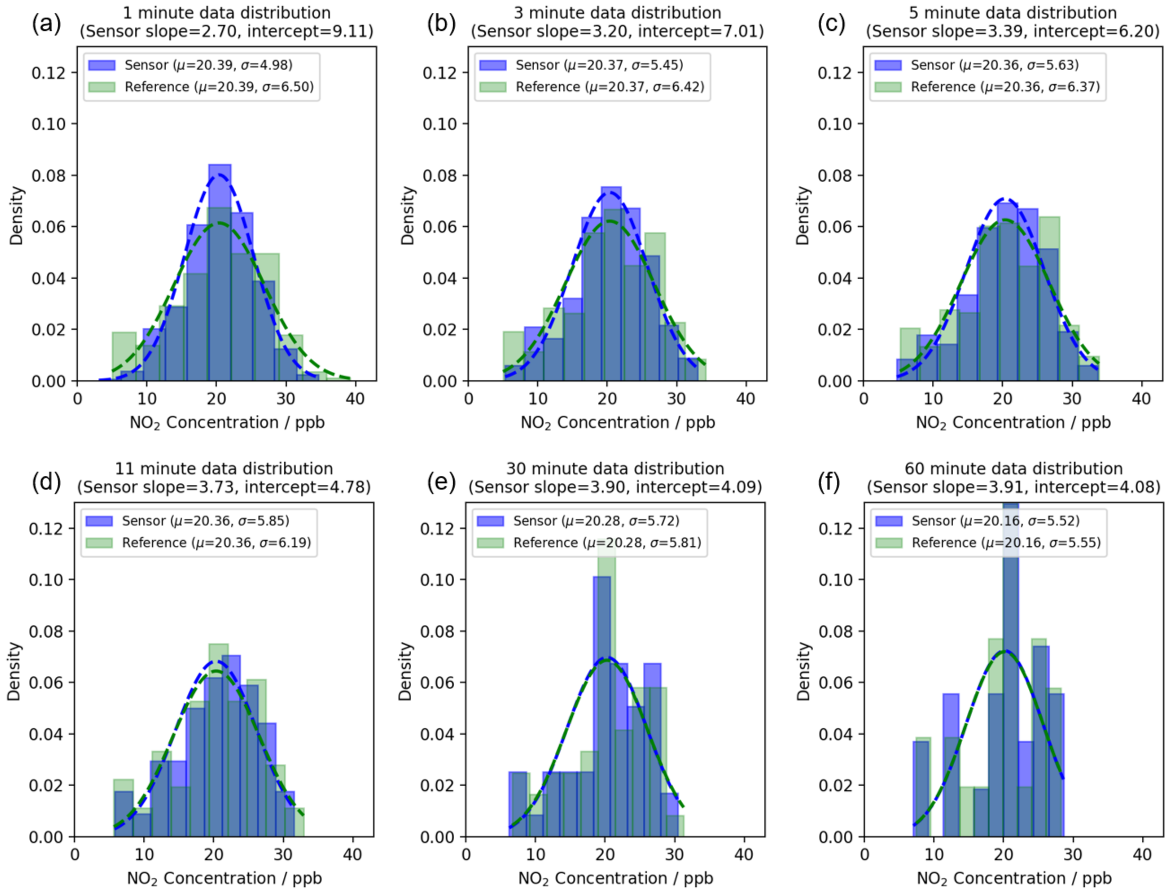

Figure S10. Comparison of the distribution between reference data and the MAS1 NO_2 sensor data after calibration using varying time averaging. The purple and green dashed lines represent the normal distribution of the sensor and reference fitting, respectively. μ and σ represent the mean and standard deviation of the normal distribution, respectively.

Text S1. Potential causes of sensor calibration coefficient variation from mathematical perspective

To delve deeper into the enhanced calibration performance with increased time averaging, the principles of linear regression are pivotal (Marill 2004). In calibrating sensor data via linear regression (Reference = a × Sensor_raw + b + ε), it is posited that the dependent variable Y (Reference) consists of the linear portion (a × Sensor_raw + b) and a residual component (ε) adhering to a normal distribution (Tripepi et al. 2008). The residuals in this model should conform to the white noise criteria (Kulperger 1998; Rahmatullah Imon 2009), signifying their independence, identical distribution, and lack of autocorrelation. These residuals should also be uncorrelated with both the independent variable X (Sensor_raw) and dependent variable Y, maintaining a zero mean and constant variance, indicative of 'homoscedasticity' or its absence, 'heteroscedasticity'.

Analyzing the residuals is essential for grasping sensor data nuances and validating the model's calibration coefficients(Law and Jackson 2017). Figure S11 displays the residual plot in the X and Y axes, where residuals display a random, homoscedastic distribution over X values but turn heteroscedastic with a strong linear tie to *Y* in minute-level data. This pattern suggests overlooked influential factors in the calibration model (Tripepi et al. 2008), which significantly interact with *Y*, affecting the residuals systematically.

Consequently, the predictive capability of the calibration model may be compromised since it fails to encapsulate these crucial variables' effects on *Y*. However, averaging *X* and *Y* over time tends to normalize the residuals' homoscedasticity in the hourly data along the *Y* axis, possibly due to mitigating heteroscedasticity-inducing elements during time averaging (Long and Ervin 2000). This leads to a more homoscedastic arrangement in *Y* and a mitigated impact on the *X*-*Y* regression relationship. While further analysis is needed to pinpoint the factors affecting residual behavior, it's clear that extended time averaging of sensor data facilitates calibration coefficients nearing the ideal solution, highlighting the importance of appropriate time averaging in achieving optimal calibration.

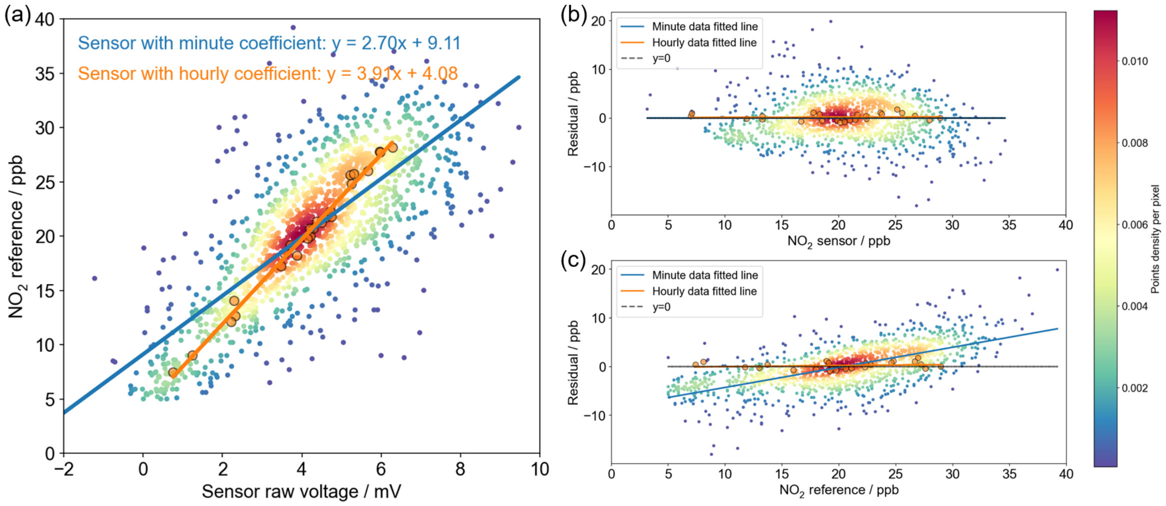

Figure S11. (a) Scatter plots of the reference and sensor_raw data of MAS1's NO_2 sensor at hourly and minute-level time averaging, respectively. (b) Distribution of sensor residuals (Reference - Sensor) in the sensor direction. (c) Distribution of sensor residuals in the reference direction. The red and blue lines in (b) and (c) are the fitted trends of the residuals.

References

Kulperger, RJ. 1998. "A Regression Residual Process." In , edited by B Szyszkowicz, 741–57. https://doi.org/10.1016/B978-044450083-0/50049-6.

Law, M, and D Jackson. 2017. "Residual Plots for Linear Regression Models with Censored Outcome Data: A Refined Method for Visualizing Residual Uncertainty." *COMMUNICATIONS IN STATISTICS-SIMULATION AND COMPUTATION* 46 (4): 3159–71. https://doi.org/10.1080/03610918.2015.1076470.

Long, JS, and LH Ervin. 2000. "Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model." *AMERICAN STATISTICIAN* 54 (3): 217–24. https://doi.org/10.2307/2685594.

Marill, KA. 2004. "Advanced Statistics: Linear Regression, Part I: Simple Linear Regression." *ACADEMIC EMERGENCY MEDICINE* 11 (1): 87–93. https://doi.org/10.1197/S1069-6563(03)00600-6.

Rahmatullah Imon, AHM. 2009. "Deletion Residuals in the Detection of Heterogeneity of Variances in Linear Regression." *JOURNAL OF APPLIED STATISTICS* 36 (3): 347–58. https://doi.org/10.1080/02664760802466237.

Tripepi, G., K.J. Jager, F.W. Dekker, and C. Zoccali. 2008. “Linear and Logistic Regression Analysis.” *Kidney International* 73 (7): 806–10. https://doi.org/10.1038/sj.ki.5002787.