Supplement of Atmos. Meas. Tech., 18, 4061–4085, 2025
https://doi.org/10.5194/amt-18-4061-2025-supplement

Atmospheric Measurement Techniques
Open Access
EGU

Supplement of

Calibration and performance evaluation of $PM_{2.5}$ and NO_2 air quality sensors for environmental epidemiology

Miriam Chacón-Mateos et al.

Correspondence to: Miriam Chacón-Mateos (miriam.chacon-mateos@ifk.uni-stuttgart.de)

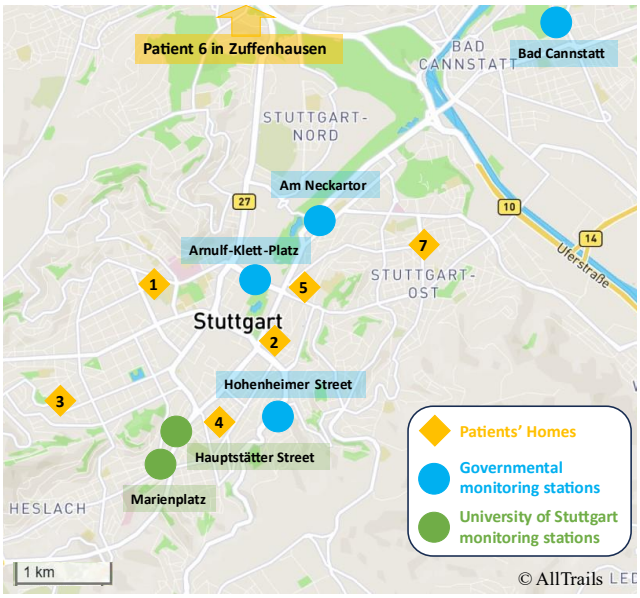

Figure S1. Map of Stuttgart showing the locations of the participants' homes where sensors were deployed (yellow diamonds), the governmental outdoor air quality monitoring stations (blue circles) and the monitoring stations that belong to the University of Stuttgart (green circles).

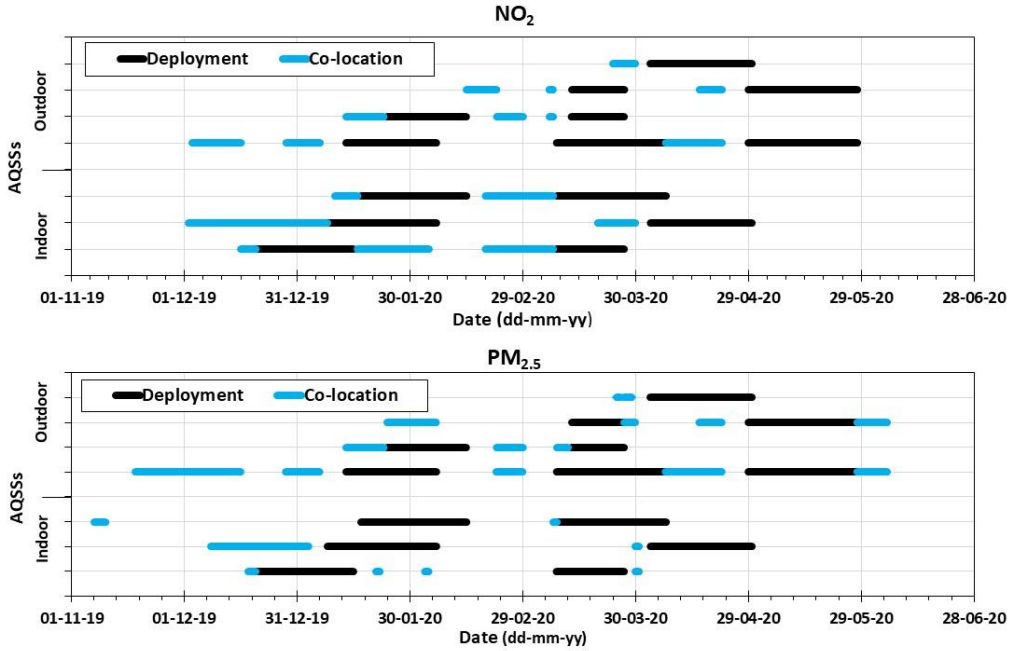

Figure S2. Measurement campaign during the pilot project for NO_2 (top) and $PM_{2.5}$ (bottom) sensors. The blue colour indicates the periods of the co-location and the black colour the periods of the the deployment in the houses of the patients.

Table S1. General information of hyperparameters and grid values used for the tuning of RFR models.

Hyperparameter	Symbol	Function Description	Grid values
Number of trees	n_estimators	Implement trees in the forest	[70, 80, 90, …, 150]
Maximum depth	max_depth	Implement depth of the tree (~levels)	[5, 10, 20, 30, 40, 50]
Minimum number of samples split	min_samples_split	Implement min. number of samples to split the internal nodes	[2, 3, 5, 7, 10]
Minimum number of samples leaf	min_samples_leaf	Implement min. number of samples for a terminal node	[1, 2, 4, 6]
Maximum number of features	max_features	Implement a number of features for the best split	['auto', 'sqrt', 'log2']

Table S2. General information of hyperparameters and grid values used for the tuning of SVR models.

Hyperparameter	Symbol	Function Definition	Grid values
Regularization parameter	C	Minimizes the error and flatness of the function	$[10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 10^{2}]$
Kernel coefficient	γ	Changes kernel shape	$[10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 10^{2}, 10^{3}]$
Error margin	ε	Defines the width of the ε-tube	$[10^{-4}, 10^{-3}, 10^{-2}, 10^{-1},]$
Kernel	-	Implement type of kernel	Linear, polynomial, RBF*
Tolerance	tol (ξ_i)	Penalization of samples outside the tube	By default (10^{-3})

*Radial basis function

Table S3. General information of hyperparameters and grid values used for the tuning of ANN models.

Hyperparameter	Function Definition	Element selected / Grid values
Learning rate	Find the global minimum of MSE* derivative	$[3\cdot10^{-3}, 2\cdot10^{-3}, 1\cdot10^{-3}]$
Number of neurons	Implement number of neurons	[0, 1, 2, …, 100]
Number of hidden layers	Implement number of hidden layers	[1, 2]
		Manual Tuned
Kernel initializer	Set initial random weights	Fixed (uniform)
Bias initializer	Set initial value of biases	Fixed (zeros)
Momentum	Update gradient descent (avoid being stuck in local minimum)	By default (0.9)
Input dimension	Defines number of features	4
Epochs	Implement the number of passes through the training dataset	[700, 1000]
Batch size	Implement the number of training points for one iteration	Fixed (20)
Activation function	Characterizes nonlinear patterns	Sigmoid
Loss function	Helps for optimization of model	Mean absolute error (MAE)
Optimizer	Optimization of weights	Stochastic gradient descent

*Mean squared error

Table S4. Statistical indicators for the performance evaluation.

Performance metrics	Short name	Mathematical formula	Ideal value	Reference literature
Standard deviation	σ	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$	-	(WMO, 2024)
Coefficient of determination	R^2	$1 - \frac{\sum_{i=1}^{n}(RM_i - \bar{M})^2}{(RM_i - \overline{RM})^2}$	1	(WMO, 2024)
Pearson correlation coefficient	r	$\frac{\sum_{i=1}^{n}(M_i - \bar{M})(RM_i - \overline{RM})}{\sqrt{\sum_{i=1}^{n}(M_i - \bar{M})^2 \sum_{i=1}^{n}(RM_i - \overline{RM})^2}}$	1	(Zimmerman et al., 2018; Penza, 2020)
Mean bias error	MBE	$\bar{M} - \overline{RM}$	0	(Zimmerman et al., 2018; Penza, 2020)
Root-mean-square error	RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - RM_i)^2}$	0	(Zimmerman et al., 2018; Penza, 2020)
Centred root-mean-square error	CRMSE	$\mathrm{CRMSE} = \sqrt{\mathrm{RMSE}^2 - \mathrm{MBE}^2}$	0	(Zimmerman et al., 2018; Penza, 2020)
Mean absolute error	MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert M_i - RM_i\rvert$	0	(Zimmerman et al., 2018; Penza, 2020)
Model Efficiency	MEF	$\mathrm{MEF} = 1 - \left(\frac{\mathrm{RMSE}}{\sigma_{\mathrm{ref}}}\right)^2$	1	(Santos and Fernández-Olmo, 2016; Paas et al., 2017)
Fractional Bias	FB	$\mathrm{FB} = \frac{2(\bar{M} - \overline{RM})}{\bar{M} + \overline{RM}}$	0	(Paas et al., 2017; Penza, 2020)

Note: x_i refers to a value measured by the sensor or the reference instrument, M_i refers to a value measured by the sensor at time i, RM_i refers to a value measured by reference instrument at time i, n refers to the total number of observations.

Calculation of the relative expanded uncertainty (REU)

The following equations are based on the DIN CEN/TS 17660-1 (2021), the DIN CEN/TS 17660-2 (2025) and the guidelines VDI 4202 Part 1 (2018) for gases and VDI 4202 Part 3 (2019) for particulate matter for automated measuring systems for air quality monitoring. These documents are based on the methodology described in JCGM (2008).

To calculate the REU of the sensor Equations S1 and S2 are used.

$$u_s^2(y_i) = \frac{RSS}{(n-2)} - u^2(x_i) + [a + (b-1)x_i]^2 \quad \textbf{Eq. S1}$$

$$RSS = \sum(y_i - a - b\,x_i)^2 \quad \textbf{Eq. S2}$$

Where $u_s^2(y_i)$ is the relative uncertainty of the NO_2 or the $PM_{2.5}$ sensor determined by comparison with the reference instrument for a concentration level y_i measured by the sensor, *RSS* is the sum of the residuals of the orthogonal regression, *n* is the number of measurements, $u(x_i)$ is the random uncertainty of the reference instrument, *a* and *b* are the intercept and slope of the orthogonal regression, respectively, and x_i is the i^{th} measured value of reference instrument. Equation S2 can only be used if $(y_i - a - b\,x_i)^2$= const. For further information, the reader is referred to the VDI 4202-1:2018 and VDI 4202 Part 3 (2019). Note that in these documents the last term of Equation S1 is $[a + (b-1)L]^2$, where L corresponds to the limit value, representing the systematic measurement error of the sensor at the level of the limit.

The random uncertainty of the reference instrument $u(x_i)$ is calculated with Equation S3. The uncertainty corresponding uncertainty of the reference instrument $u_{bs,RM}$ between the two measuring systems $x_{i,1}$ and $x_{i,2}$ shall be calculated following Equation S4. A list of values for different reference-grade instruments is listed in Table J.1 of the CEN/TS 17660-1:2021.

$$u(x_i) = \frac{u_{bs,RM}}{\sqrt{2}} \quad \textbf{Eq. S3}$$

$$u_{bs,RM}^2 = \frac{\sum_{i=1}^{n}\left(x_{i,1} - x_{i,2}\right)^2}{2 \times n} \quad \textbf{Eq. S4}$$

The combined uncertainty $u_{c,field}^2(y_i)$ is calculated using Equation S5:

$$u_{c,field}^2(y_i) = \frac{u_s^2(y_i)}{y_i^2} \quad \textbf{Eq. S5}$$

The expanded uncertainty $U_{field}(y_i)$ is finally obtained by multiplying the combined standard uncertainty $u_{c,field}(y_i)$ by a coverage factor *k* which is in this case 2 as in Eq. S6:

$$U_{field}(y_i) = k \times u_{c,field}(y_i) = 2 \times \sqrt{u_{c,field}^2(y_i)} \quad \textbf{Eq. S6}$$

To calculate the parameters of the orthogonal regression (Eq. S7), Eq. 8-14 are been used as demonstrated in Annex D of the VDI 4202 Part 1 (2018) considering the sum of squares (SS):

$$y = a + b\,x \quad \textbf{Eq. S7}$$

$$a = \bar{y} - b\,\bar{x} \quad \textbf{Eq. S8}$$

$$b = \frac{S_{yy} - S_{xx} + \sqrt{\left(S_{yy} - S_{xx}\right)^2 + 4\left(S_{xy}\right)^2}}{2\,S_{xy}} \quad \textbf{Eq. S9}$$

with

$$S_{xx} = \sum(x_i - \bar{x})^2 \quad \textbf{Eq. S10}$$

$$S_{yy} = \sum(y_i - \bar{y})^2 \quad \textbf{Eq. S11}$$

$$S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) \quad \textbf{Eq. S12}$$

$$\bar{x} = \frac{1}{n}\sum x_i \quad \textbf{Eq. S13}$$

$$\bar{y} = \frac{1}{n}\sum y_i \quad \textbf{Eq. S14}$$

Table S5. General information of the linear correction of indoor and outdoor $PM_{2.5}$ sensors.

Patient ID	P5	P1	P2	P3	P4		P6	P7
Indoor AQSS	**B01**	**B02**	**B04**	**B04**		**B01**	**B02**	**B03***
				Training data				
Correction factor	0.97	1.15	0.89	2.34		1.06	1.01	0.93
Offset	0.20	-0.14	0.42	-1.38		1.74	3.24	2.33
$PM_{2.5}$ range (µg m^{-3})	5 - 114	0 - 151	2 - 84	1 - 110		3 - 179	3 - 134	4 - 25
Averaging time (min)	1	1	1	1		1	1	30
Number of data points	180	555	314	69		88	88	433
				Testing data				
Averaging period (min)	1	1	1	1		1	1	30
$PM_{2.5}$ range (µg m^{-3})	1 - 110	0 - 54	1 - 64	11 - 117		0 - 115	1 - 128	5 - 21
R^2	0.99	0.82	0.97	1.00		0.98	0.98	0.90
Outdoor AQSS	**-**	**B03**	**B05**	**B03**	**B05**	**B06**	**B08**	**B06**
				Training data				
Correction factor	-	0.83	0.37	0.93	0.96	0.55	0.86	0.79
Offset	-	4.66	8.96	2.33	1.04	1.15	1.40	2.92
Temperature range (°C)	-	-0.5 - 11.3	-1.9 - 7.3	3.0 - 27.9	3.3 - 18.1	0.4 - 16.5	2 - 25	11.5 - 26.1
Relative humidity range (%)	-	51.5 - 78.6	47.7 - 73.9	15.1 - 60.2	24.4 - 82.9	40.8 - 77.4	19 - 65	12.7 - 68.0
$PM_{2.5}$ range (µg m^{-3})	-	2 - 123	6 - 35	4 - 25	1 – 11	1 - 39	2 - 32	5 - 21
Averaging time (min)	-	30	30	30	30	30	10	30
Number of data points	-	406	219	433	356	626	263	294
				Testing data				
Averaging period (min)	-	30	30	30	30	30	10	30
Temperature range (°C)	-	0 .0 - 15.4	4.2 - 16.6	3.4 - 16.9	6.6 - 19.8	2.6 - 27.2	1 - 21	11 - 30.6
Relative humidity range (%)	-	45.0 - 78.4	32.6 - 71.7	21.2 - 77.7	38.4 - 72.4	18.7 - 71.3	13 - 57	19.4 - 76.5
$PM_{2.5}$ range (µg m^{-3})	-	2 - 31	3 - 25	1 - 11	1 - 6	2 - 32	9 - 19	2 - 9

*Note that B03 is an outdoor AQSS that was calibrated outdoor but used indoors in the house of patient P7.

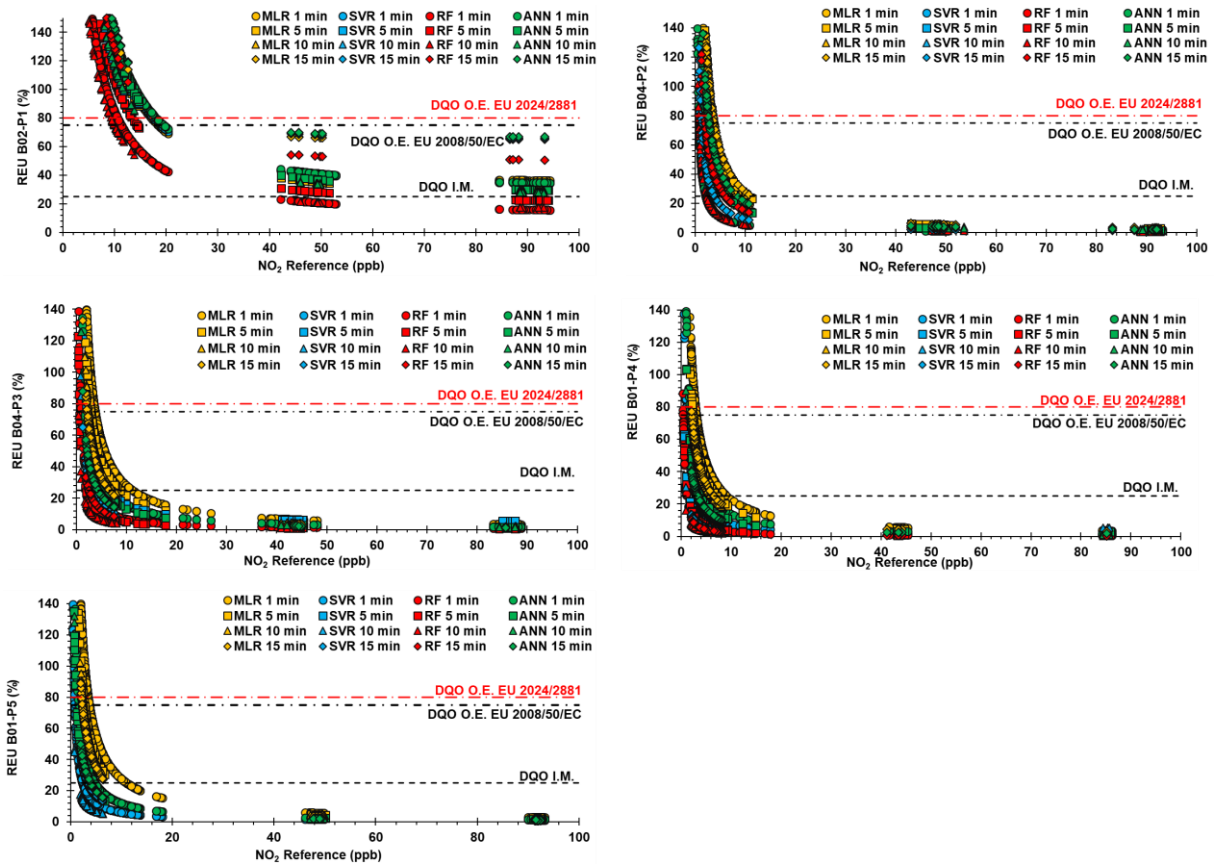

Figure S3. Relative expanded uncertainties of indoor NO_2 sensors for the tested models and different averaging times against reference concentrations. The dashed line indicates the DQO for indicative measurements while the dash-dot lines represent the DQOs for objective estimation (black for EU Directive 2008/50/EC and red for EU Directive 2024/2881). The DQO for short-term indicative measurements is the same in both Directives.

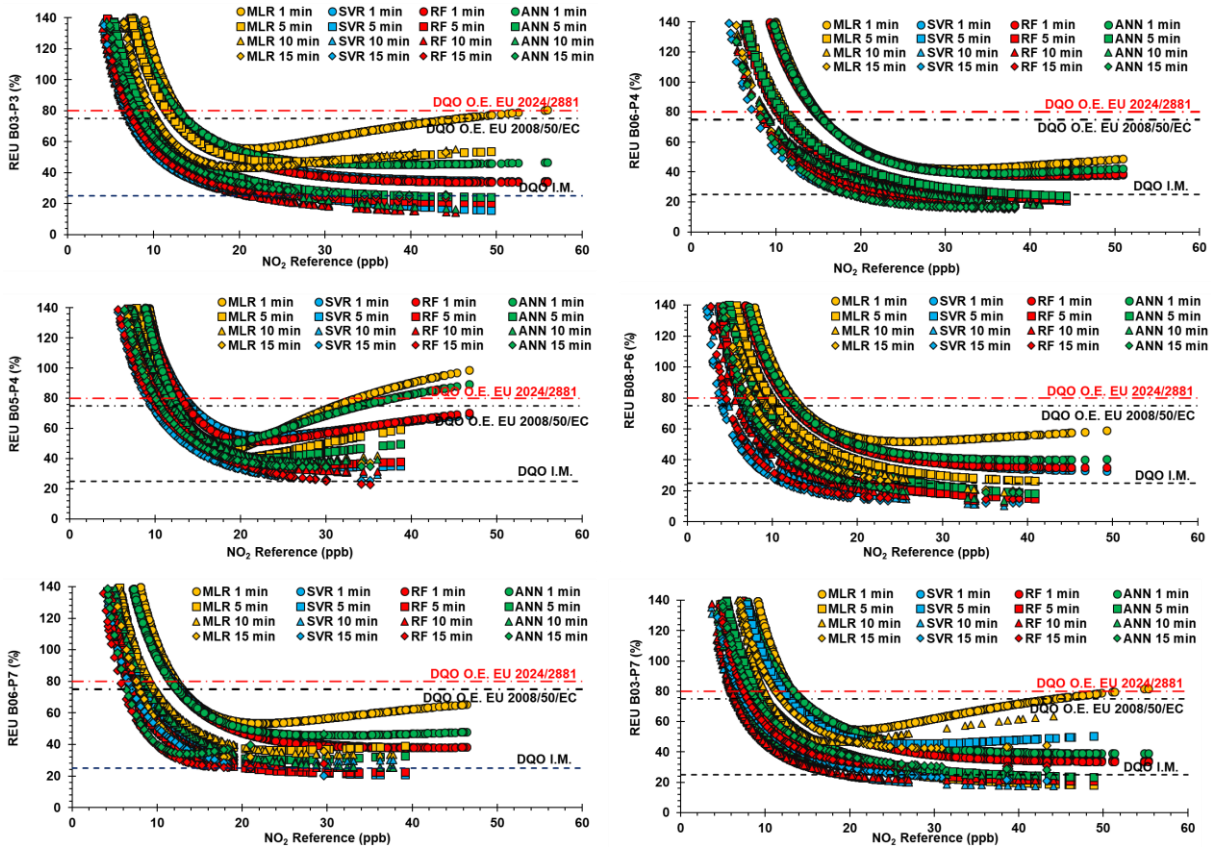

Figure S4. Relative expanded uncertainties of outdoor NO_2 sensors for the tested models for different averaging times against reference NO_2 concentrations. The dashed line indicates the DQO for indicative measurements while the dash-dot lines represent the DQOs for objective estimation (black for EU Directive 2008/50/EC and red for EU Directive 2024/2881). B03-P7 is an outdoor AQSS that was used for indoor measurements as part of an experiment to test calibration methods for indoor measurements. The DQO for short-term indicative measurements is the same in both Directives.

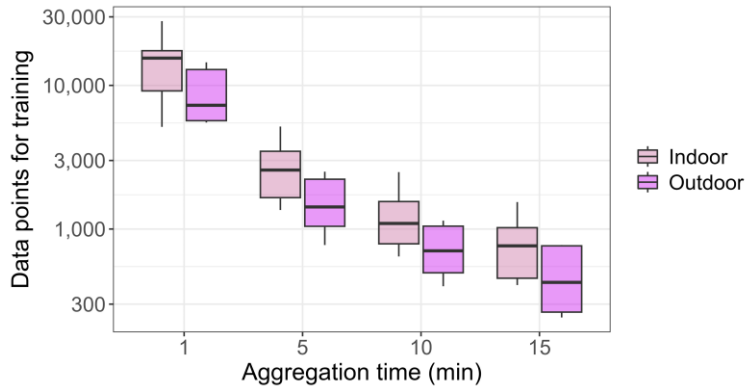

Figure S5. Boxplots of data points used for the training of NO_2 calibration models for the different time aggregations.

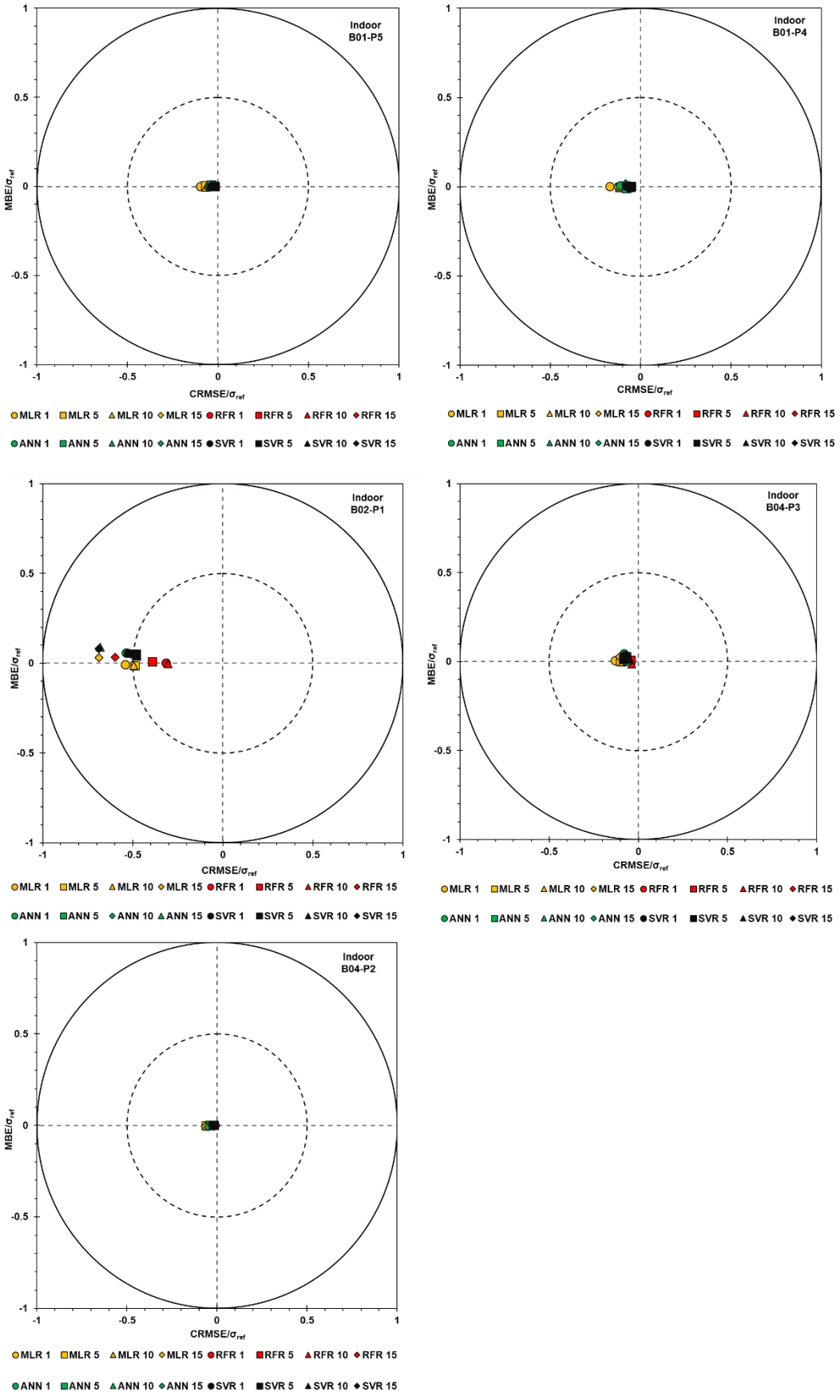

Figure S6. Target diagrams for indoor NO_2 sensors for the tested models and different averaging times.

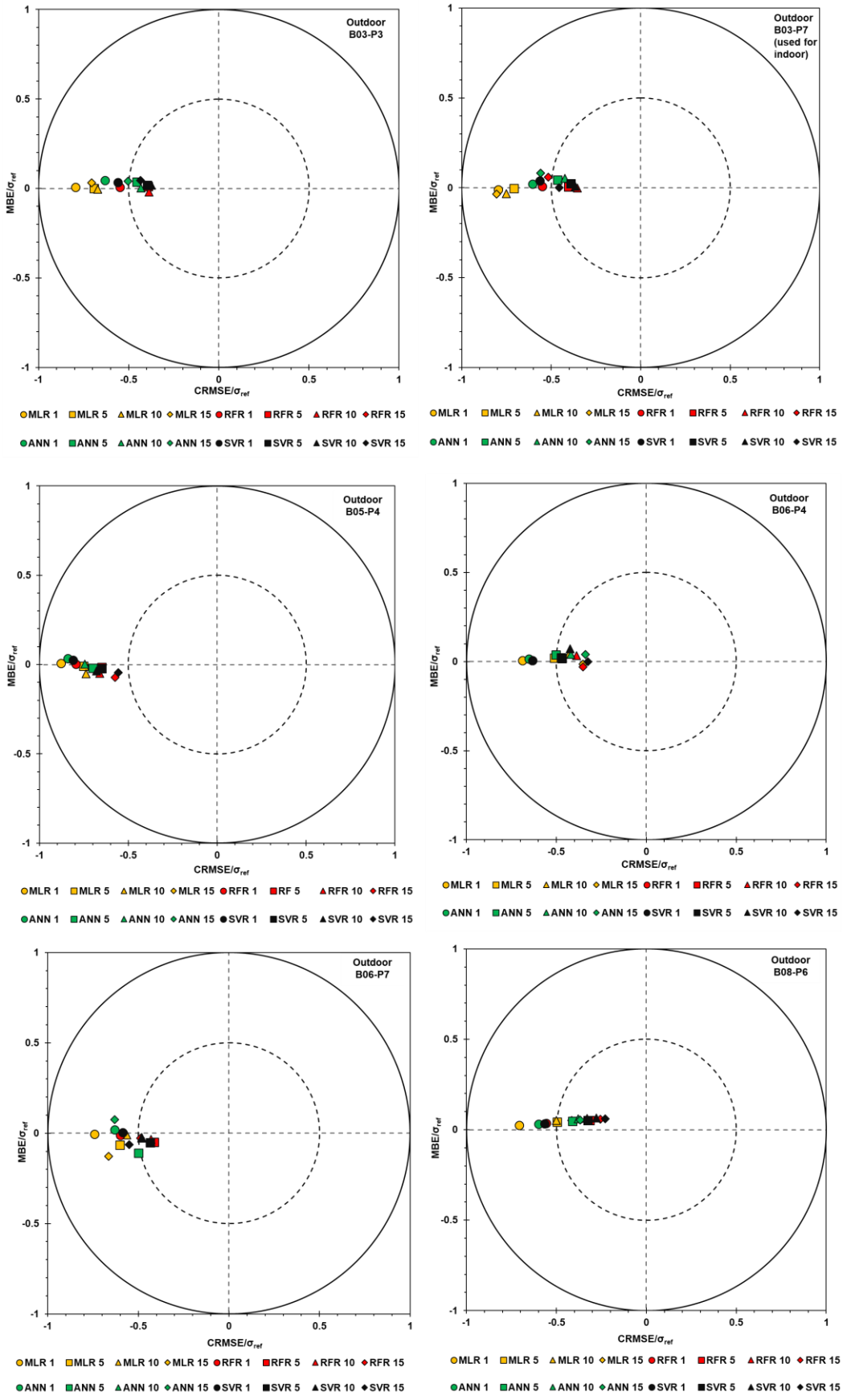

Figure S7. Target diagrams for outdoor NO_2 sensors for the tested models and different time aggregations. B03-P7 is an outdoor AQSS that was used for indoor as part of an experiment to test calibration methods for indoor measurements.

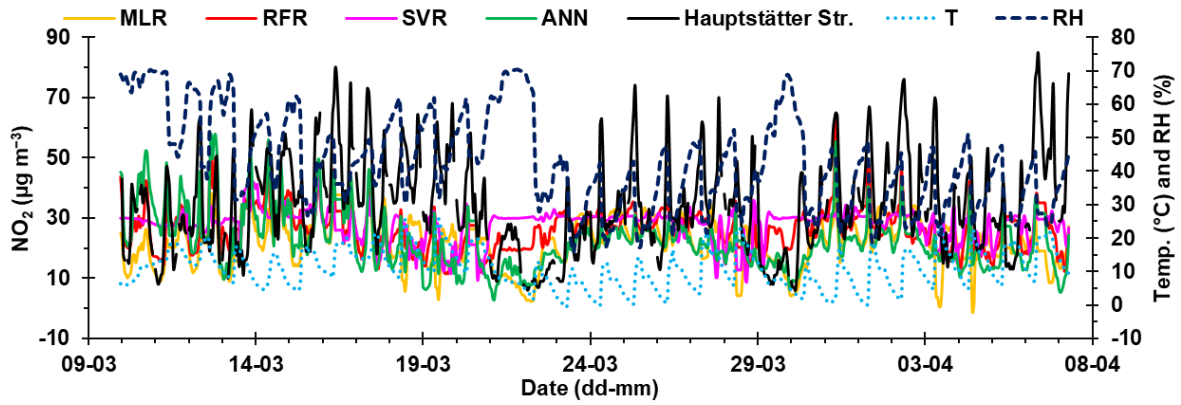

Figure S8. Time series of hourly outdoor NO_2 concentrations calculated using four sensor models (MLR, RFR, SVR and ANN) during deployment in the house of patient P3, NO_2 concentrations measured at the monitoring station in Hauptstätter Street (traffic station) and T and RH measurements of the AQSS.

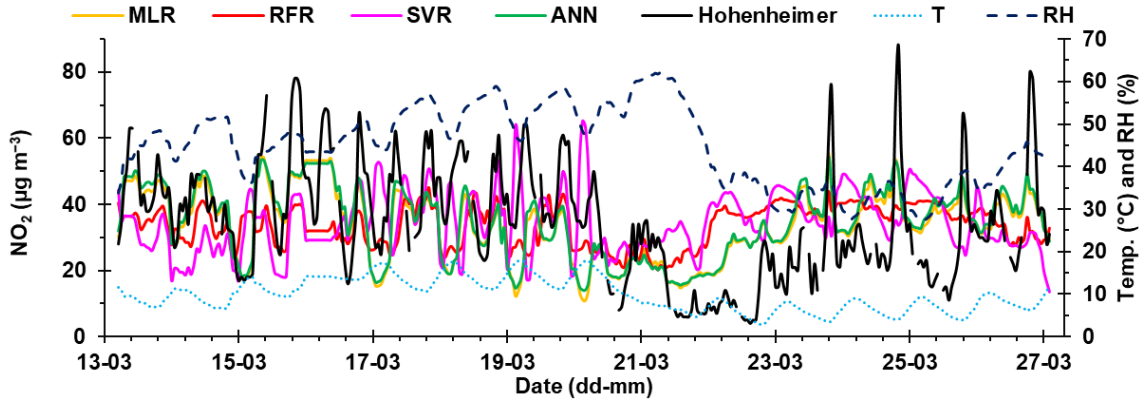

Figure S9. Time series of hourly outdoor NO_2 concentrations predicted with four sensor models (MLR, RFR, SVR and ANN) during deployment in the house of patient P4 (garden), NO_2 concentrations measured at the monitoring station in Stuttgart Hohenheimer Street (traffic station) and T and RH measurements of the AQSS.

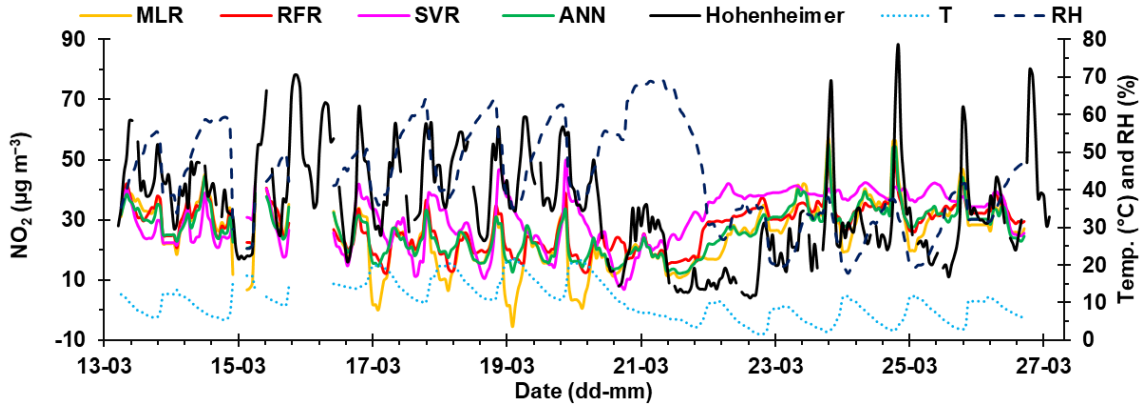

Figure S10. Time series of hourly outdoor NO_2 concentrations calculated using four sensor models (MLR, RFR, SVR and ANN) during deployment in the house of patient P4 (street), NO_2 concentrations measured at the monitoring station in Stuttgart Hohenheimer Street (traffic station) and T and RH measurements of the AQSS.

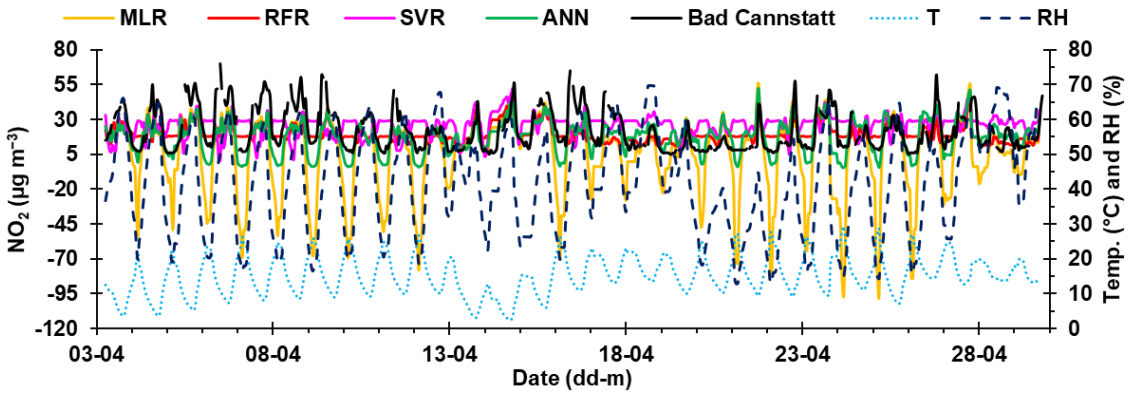

Figure S11. Time series of hourly outdoor NO_2 concentrations calculated using four sensor models (MLR, RFR, SVR and ANN) during deployment in the house of patient P6, NO_2 concentrations measured at the monitoring station in Bad Cannstatt (urban station) and T and RH measurements of the AQSS.

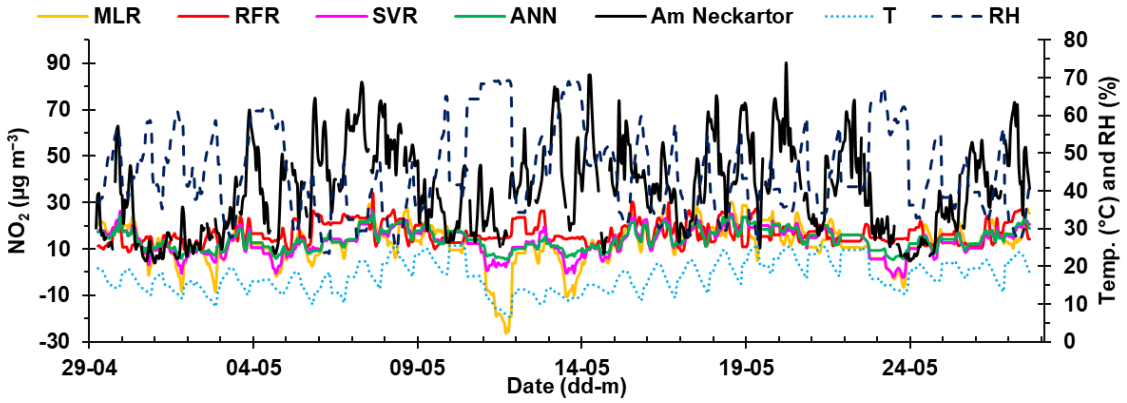

Figure S12. Time series of hourly outdoor NO_2 concentrations calculated using four sensor models (MLR, RFR, SVR and ANN) during deployment in the house of patient P7, NO_2 concentrations measured at the monitoring station in Neckartor (traffic station) and T and RH measurements of the AQSS.

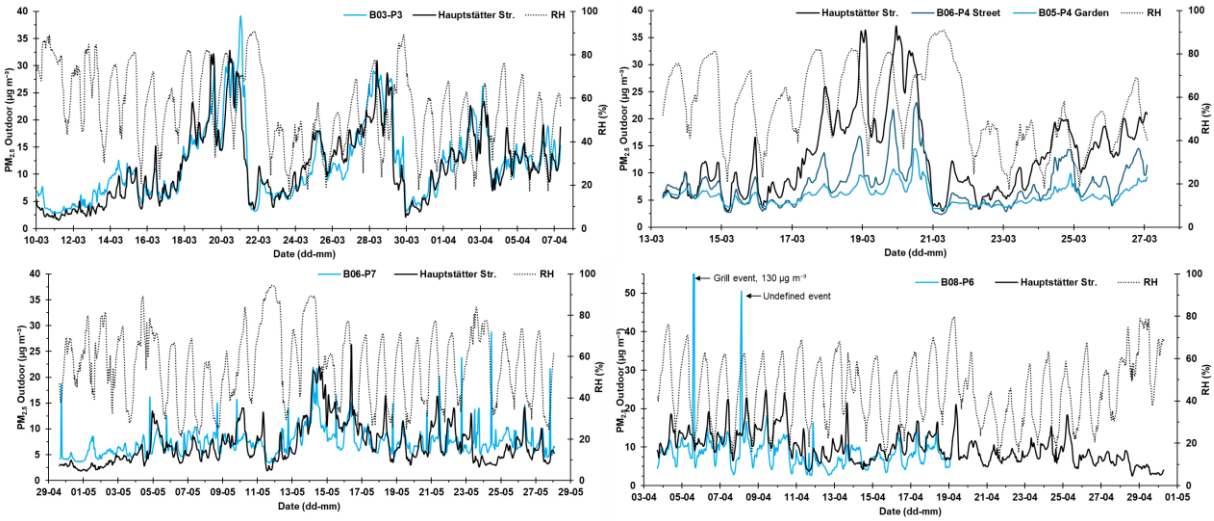

Figure S13. Hourly $PM_{2.5}$ concentration and RH of outdoor sensors during deployment near the houses and the reference station (Stuttgart Hauptstätter Str.). The monitoring station is located from 1 to 6 km away from the houses. Sensor B08-P6 stopped working on the 19-04-2020.

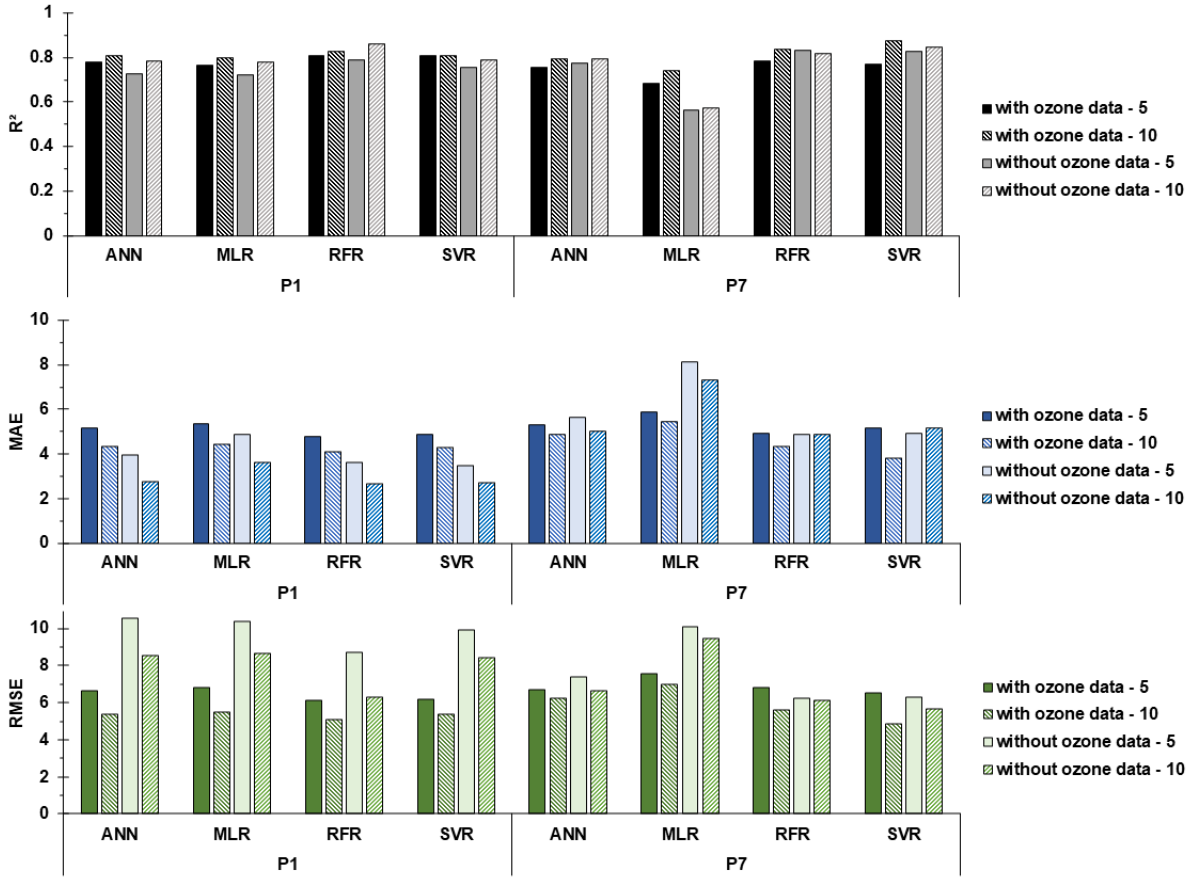

Figure S14. Comparison of the validation results of the R^2, RMSE and MAE for outdoor NO_2 sensors before deployment in the houses of the patients P1 and P7 with and without additional ozone data as input for the regression and ML models for 5- and 10-min averages. MAE and RMSE in ppb.

References

CEN/TS 17660-1: Air quality - Performance evaluation of air quality sensor systems - Part 1: Gaseous pollutants in ambient air, 2021.

CEN/TS 17660-2: Air quality - Performance evaluation of air quality sensor systems - Part 2: Particulate matter in ambient air, 2025.

JCGM: Evaluation of measurement data – guide to the expression of uncertainty in measurement, 2008.

Paas, B., Stienen, J., Vorländer, M., and Schneider, C.: Modelling of Urban Near-Road Atmospheric PM Concentrations Using an Artificial Neural Network Approach with Acoustic Data Input, Environments, 4, 26, https://doi.org/10.3390/environments4020026, 2017.

Penza, M.: Low-cost sensors for outdoor air quality monitoring, in: Advanced Nanomaterials for Inexpensive Gas Microsensors, Elsevier, 235–288, https://doi.org/10.1016/B978-0-12-814827-3.00012-8, 2020.

Santos, G. and Fernández-Olmo, I.: A proposed methodology for the assessment of arsenic, nickel, cadmium and lead levels in ambient air, Sci. Total Environ., 554-555, 155–166, https://doi.org/10.1016/j.scitotenv.2016.02.182, 2016.

VDI 4202 Part 1: Automated measuring systems for air quality monitoring. Performance test, declaration of suitability, and certification of point-related measuring systems for gaseous air pollutants, 2018.

VDI 4202 Part 3: Automated measuring systems for air quality monitoring. Performance test, declaration of suitability, and certification of point-related measuring systems of mass concentration for particulate air pollutants, 2019.

WMO: Integrating Low-cost Sensor Systems and Networks to Enhance Air Quality Applications, Switzerland, GAW Report No. 293, 2024.

Zimmerman, N., Presto, A. A., Kumar, S. P. N., Gu, J., Hauryliuk, A., Robinson, E. S., and Robinson, A. L.: A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring, Atmos. Meas. Tech., 11, 291–313, https://doi.org/10.5194/amt-11-291-2018, 2018.