Article

Field Study of Metal Oxide Semiconductor Gas Sensors in Temperature Cycled Operation for Selective VOC Monitoring in Indoor Air

Tobias Baur * , **Johannes Amann, Caroline Schultealbert** and **Andreas Schütze**

Lab for Measurement Technology, Saarland University, Campus A5 1, 66123 Saarbrücken, Germany;
j.amann@lmt.uni-saarland.de (J.A.); c.schultealbert@lmt.uni-saarland.de (C.S.);
schuetze@lmt.uni-saarland.de (A.S.)
* Correspondence: t.baur@lmt.uni-saarland.de

Abstract: More and more metal oxide semiconductor (MOS) gas sensors with digital interfaces are entering the market for indoor air quality (IAQ) monitoring. These sensors are intended to measure volatile organic compounds (VOCs) in indoor air, an important air quality factor. However, their standard operating mode often does not make full use of their true capabilities. More sophisticated operation modes, extensive calibration and advanced data evaluation can significantly improve VOC measurements and, furthermore, achieve selective measurements of single gases or at least types of VOCs. This study provides an overview of the potential and limits of MOS gas sensors for IAQ monitoring using temperature cycled operation (TCO), calibration with randomized exposure and data-based models trained with advanced machine learning. After lab calibration, a commercial digital gas sensor with four different gas-sensitive layers was tested in the field over several weeks. In addition to monitoring normal ambient air, release tests were performed with compounds that were included in the lab calibration, but also with additional VOCs. The tests were accompanied by different analytical systems (GC-MS with Tenax sampling, mobile GC-PID and GC-RCP). The results show quantitative agreement between analytical systems and the MOS gas sensor system. The study shows that MOS sensors are highly suitable for determining the overall VOC concentrations with high temporal resolution and, with some restrictions, also for selective measurements of individual components.

Keywords: MOS; metal oxide semiconductor gas sensor; VOC; volatile organic compounds; IAQ; indoor air quality; randomized gas mixtures; selective; SGP30; quantification

1. Introduction

Air pollution is one of the main environmental concerns in Europe and worldwide with outside and indoor air contributing similarly to the overall burden of disease according to the EU project Healthvent [1]. In recent years, indoor air quality has gained increasing relevance and awareness of its importance is rising [2]. Quality in this context includes many parameters, from temperature to particles to volatile organic compounds (VOCs) and others [3]. With technology becoming cheaper and Internet of Things (IoT) devices being available to a broader public, measurement systems for every parameter are in demand.

For indoor air quality assessment carbon dioxide (CO_2) is the de facto standard because it provides reliable results due to the physical measurement principle. As humans emit a cocktail of VOCs [4–6], and this is mainly responsible for poor air quality in indoor situations, a CO_2 measurement is often referred to as indirect VOC measurement based on the studies of Pettenkofer [7]. However, this approach neglects other VOC sources such as furniture and building materials as well as those coming from human activities like cleaning or cooking [8]. Furthermore, these sensors are relatively large, power-hungry, and expensive compared to metal oxide semiconductor (MOS) gas sensors, especially

in the context of IoT. MOS sensors provide excellent sensitivity and a broad response spectrum covering almost all kinds of VOCs [9,10]. Due to their broad sensitivity spectrum, most commercially available sensors provide a sum signal often designated as total VOC (TVOC) concentration [11]. However, permanent gases like hydrogen (H_2) or carbon monoxide (CO) could also contribute to this sum signal as MOS sensors often show high sensitivity towards these gases [12]. Moreover, chemical sensors can change their chemical properties during operation due to irreversible reactions, so drift is often reported [13]. The latest sensor models offered by different manufacturers are typically smaller than $3 \times 3 \times 1$ mm^3, require less than 10 mW of power and include integrated electronics offering a direct digital interface allowing simple integration in various (IoT) devices. Some of these sensors [11,14] use multiple gas-sensitive layers to provide an even wider response spectrum and allow multisensor evaluation. With more sophisticated data treatment and more complex operation modes, like temperature cycled operation (TCO), it is possible to improve the selectivity of these sensors [15–17]. This was often shown in lab measurements and first studies on inter-laboratory comparisons are available [18], but proof concerning the feasibility of such an approach and its stability in the field is missing. Before a broader public can use the sensors and benefit from their results, the performance needs to be ensured in field studies with comparisons to analytical instrumentation.

We present a study on selective VOC measurements with MOS sensors and their stability in a real-world scenario. A multilayer sensor combined with TCO is used to achieve good selectivity. The capabilities of this low-cost approach for determining the overall VOC concentration independent of interference by ambient humidity, CO and H_2 and also selective quantification of single gases are evaluated. After calibration in the lab, the sensors were installed in an office where several release tests of different substances were conducted to prove the ability to selectively detect and quantify certain VOCs; a method that could also be used as a simple functionality test for the general public. The lab calibration was repeated twice after several weeks of operation each to evaluate the drift of the sensor elements and stability of the model prediction.

2. Materials and Methods

2.1. Experimental Setup

All measurements in this study were performed with sensor hardware designed in-house. The sensor hardware is based on a microcontroller board (Teensy 4.0, Pjrc.com LLC, Sherwood, Oregon, USA), which communicates with an SGP30 sensor (Sensirion AG, Stäfa, Switzerland) via I^{2}C interface. The SGP30 multilayer MOS sensor contains four gas-sensitive layers on a common MEMS micro hotplate [11]. It is possible to digitally program the sensor to set the temperature from 100 °C to 425 °C in 25 °C steps and to synchronously read out the resistance of the four different layers. The commands for temperature control and resistance readout are not described in the sensor datasheet and were provided by Sensirion under a non-disclosure agreement. Our sensor hardware allows us to operate the sensor in TCO mode and readout of the layers' resistances and transfer the data to a PC with a sample rate of 20 Hz. The selected temperature cycle (TC) comprises 10 temperature jumps from high to low temperature [19,20]. Figure 1 shows the TC with 10 steps at 400 °C with a duration of 5 s each are followed by different low-temperature steps, which are set to 100, 125, 150, 175, 200, 275, 300, 325, 350 and 375 °C with a duration of 7 s each resulting in a total duration of the TC of 120 s. The SGP30 sensors were installed in sensor chambers (alumina and polytetrafluoroethylene). The sensor systems including electronics were mounted on a trolly with PC and monitor allowing us to move them from the laboratory to the field test room and back. The trolly carries a flow-regulated micro pump drawing room air through the sensor chambers for the field tests, to ensure similar flow conditions over the sensor in the field as during calibration in the laboratory.
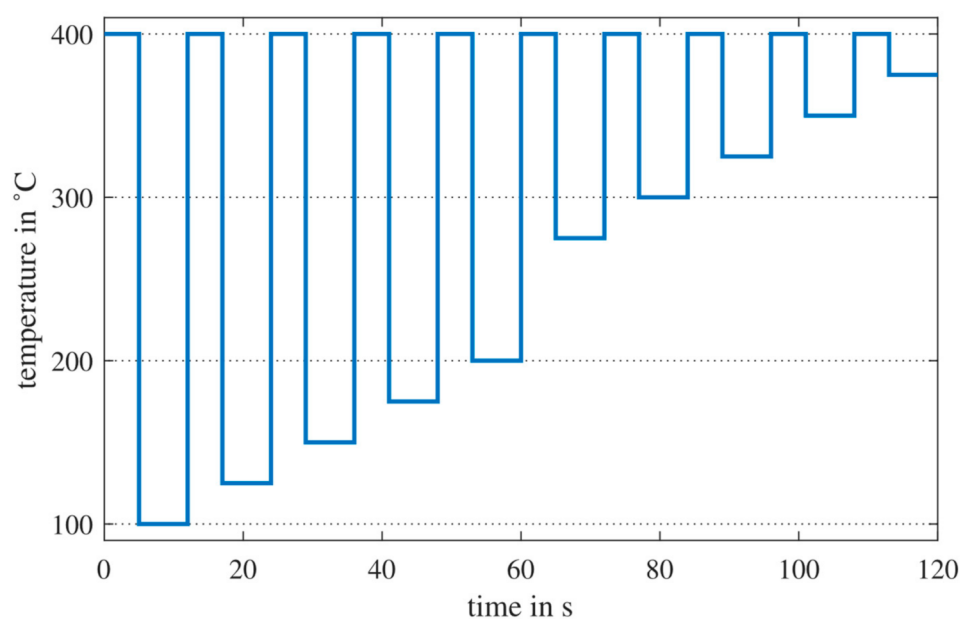
Figure 1. Temperature cycle of the SGP30.

The calibration measurements were done with our custom-built gas mixing apparatus (GMA), which is described in detail in [21]. Figure 2 shows a schematic overview of the GMA and the connection to the sensor hardware. The GMA is based on the principle of dynamic mass flow injection of different test gases into a carrier gas flow. The carrier line consists of two 500 mL/min mass flow controllers (MFC), one for dry and one for humidified zero air, for a dynamic humidity setting. Zero air is generated by a GT PLUS 15000 ULTRA-ZERO Air Generator (Schmidlin Labor + Service GmbH & Co. KG, Dettingen, Germany) with different filter steps to remove water vapor, carbon mon-/dioxide, VOCs, nitrogen oxide (NO_x), sulfur oxide (SO_x) and ozone. For humidification, dry zero air is passed through a wash bottle filled with HPLC-grade water followed by a filter to remove particles and droplets. Both the wash bottle and filter are kept at 20 °C (thermostat) to keep the humidity level constant. To achieve reliable low concentrations of the test gases, we use test gas cylinders with concentrations of at least 100 ppm of the target gas in synthetic air with a purity of >99.999% and, if required to achieve low concentrations below 1 ppm, add a predilution step before injection into the carrier flow. For the predilution, the test gas flow from the gas cylinder is diluted once with zero air by two MFCs with 10 or 20 mL/min for the test gas and 500 mL/min for zero air. The test gas from the gas cylinder or the diluted test gas is injected into the carrier gas with another MFC (10/20 mL/min). The GMA includes one test gas line for direct injection and five test gas lines with integrated predilution. The total flow entering the sensor chambers is always kept constant. Therefore, we can dynamically mix six different test gases and humidity in one measurement. In addition to the SGP30 described here, further sensor systems were included in the measurement campaign. To avoid crosstalk between the sensors due to reactions on the sensor elements, the total volume flow of 300 mL/min was divided into four parallel flows using restrictions (1/16", 20 cm), resulting in 75 mL/min per line. The total flow after the sensor chambers is measured with a mass flow meter (MFM) to ensure the tightness of the system.
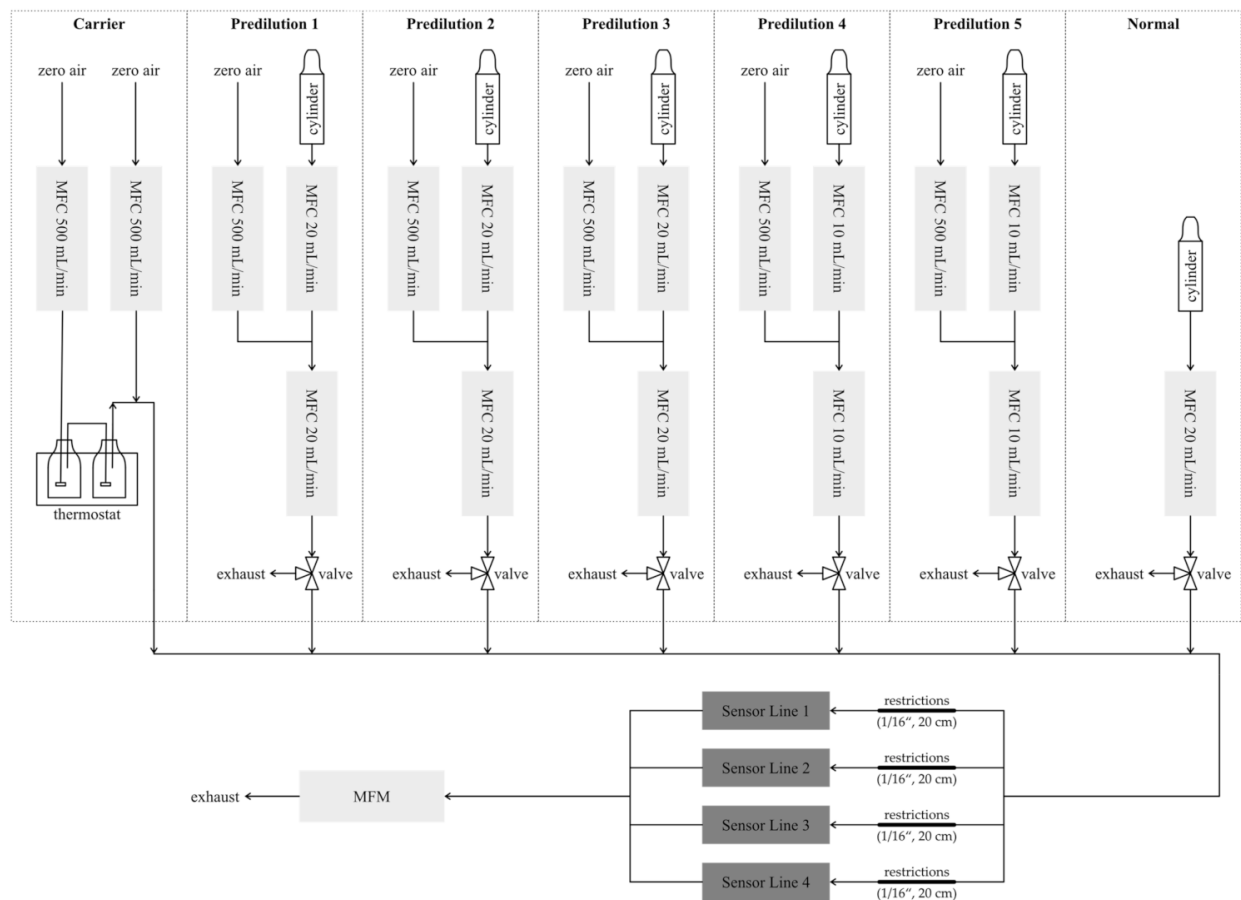
Figure 2. Schematic overview of the calibration setup with the gas mixing apparatus and the sensor systems.

The field tests are accompanied by three different analytical methods. All three systems are based on gas chromatic separation with different detectors.

1. Thermo desorption gas chromatography-mass spectrometry (TD-GC-MS, Markes International Ltd, Llantrisant, Wales, UK, Thermo Fisher Scientific Inc., Waltham, MA, USA), similar to ISO 16000-6. TENAX® tubes were sampled with room air for 10 min at 50 mL/min. This system was used to quantify toluene during the release tests. The TD-GC-MS was calibrated in the same way with known concentrations from the GMA (7 tubes with 50–500 ppb toluene). The calibration was done 3 days before the specific release test. LOQ is smaller than 50 ppb, and the uncertainty is estimated to be 20%, based on the gas cylinder and the sampling method;

2. Peak Performer 1 (Peak Laboratories LLC, Mountain View, CA, USA), a gas chromatograph with a reducing compound photometer as detector (GC-RCP), allows selective quantification of hydrogen with a LOD of 10 ppb and a resolution of 10% of reading or LOD (whichever is higher). The Peak Performer 1 requires nitrogen or another inert carrier gas from a pressure cylinder and provides a time resolution of 3.6 min;

3. Dräger X-pid 9500 (Dräger Safety AG & Co KGaA, Lübeck, Germany), a portable GC-PID offering a broad range of measured gases including acetone (LOQ: 500 ppb, LOD: 170 ppb), toluene (LOQ: 1000 ppb, LOD: 330 ppb), isopropyl alcohol (LOQ: 3000 ppb, LOD: 1000 ppb), and xylene (LOQ: 3000 ppb, LOD: 1000 ppb). The X-pid 9500 requires a daily function test with a test gas cylinder (10 ppm isobutene and 10 ppm toluene). Depending on the selected gases, a single measurement requires 2–3 min.

2.2. Calibration and Recalibration in the GMA

The aim of the calibration is to achieve a reliable mathematical model for the prediction of different VOC, interfering gases, and sum signals, e.g., the sum of all VOCs in indoor air, in our field tests from the multi-dimensional gas sensor data. Analytic studies of VOCs in indoor air show that more than 400 different VOCs representing more than 14 chemical classes can be found in indoor air [22,23]. Studies on other substances besides VOC in indoor air are less diverse. From previous studies, we learned that hydrogen and carbon monoxide, which have a strong influence on MOS sensors, both show large variations in indoor air [24,25]. It is obviously not feasible to include all VOCs or interfering substances in the calibration, therefore a reduced substance list for the calibration strategy is needed. For the calibration, we are restricted to six different gases due to the used GMA. The following criteria were defined to select a reduced list of substances:

1. Divide the list of VOCs found in studies in indoor environments into the most common chemical classes (also named substance types or groups): alcohols, aldehydes, alkanes, alkenes, aromatics, esters, glycols and glycol ethers, halocarbons, ketones, siloxanes, terpenes and organic acids;
2. Sort the chemical classes according to their total concentrations;
3. For each chemical class, select the substance with the highest concentration.

The idea behind this approach is the assumption that most substances of a certain chemical class react similarly on the sensor surface, therefore one single gas could represent each class. However, it is difficult to verify this assumption based on reaction similarity because this would have to be assessed for each sensor model and hundreds of gases. On the other hand, if the assumption is true, it would mean that compounds are difficult to quantify selectively. In addition, the substance occurring with the highest concentration may not be the most reactive on the sensor. This means that substances with lower concentrations can still generate a higher sensor response. The behavior still needs to be investigated in more detail. Table 1 shows the 90th and 95th percentile concentration values determined from the analytical studies for the eight chemical classes with the highest sum concentration.

Table 1. 90th (P90) and 95th (P95) percentile sum concentration in $\mu g/m^3$ and ppb (calculated from the individual substances dominating for each chemical class) for the eight chemical classes with the highest sum concentrations as determined from analytical studies [22,23] in alphabetical order. The substance in parentheses is the representative with the highest concentration for this chemical class.

Chemical Class (Representative)	P90 in $\mu g/m^3$ (ppb)	P95 in $\mu g/m^3$ (ppb)
Alcohols (Ethanol)	320 (~170)	520 (~790)
Aldehydes (Formaldehyde)	340 (~270)	480 (~390)
Alkanes (n-Hexane, n-Heptane)	180 (~50)	350 (~90)
Aromatics (Toluene)	190 (~50)	370 (~90)
Esters (Ethyl acetate)	140 (~30)	280 (~70)
Ketones (Acetone)	250 (~100)	420 (~170)
Terpenes (Limonene, α-Pinene)	170 (~30)	330 (~60)
Organic acid (Acetic acid)	150 (~60)	240 (~100)

The chemical classes with the highest sum concentrations are alcohols, aldehydes, and ketones, followed by alkanes, aromatics, terpenes, and organic acid in similar magnitude. We selected ethanol (alcohols), formaldehyde (aldehyde), acetone (ketones) and toluene (aromatics) as the four VOC representatives for calibration. In addition, we included hydrogen and carbon monoxide as interfering gases for the calibrations.

The calibration strategy is based on randomized gas mixing as described in [26]. The aim of the strategy is to calibrate the sensor with a more realistic measurement including masking effects and other gas interactions. Therefore, statistically distributed gas profiles with unique randomized gas mixtures are measured, and not only single gases with ascending concentrations compared to classical sequential calibration. For the calibration, a randomized gas mixture profile was generated. The distribution is based on Latin Hypercube sampling [27] for each target substance, with the aim to achieve low correlation coefficients between the various target substances. The gas mixture profile was run in the GMA and each gas mixture was kept constant for 20 min or 10 sensor T-cycles at a total flow of 300 mL/min.

We defined concentration ranges based on the analytical studies describing VOC concentrations in empty rooms (background) as well as literature values [28,29] for the interfering gases, cf. Table 2.

Table 2. Background concentration ranges for different substances in the initial calibration and the resulting range for VOC_{sum} calculated from the single VOCs (Acetone, Toluene, Formaldehyde and Ethanol).

Substance	Min.	Max.
Carbon Monoxide	150 ppb	2000 ppb
Hydrogen	400 ppb	2000 ppb
Humidity	25 %RH	70 %RH
Acetone	14 ppb	300 ppb
Toluene	4 ppb	300 ppb
Formaldehyde	1 ppb	400 ppb
Ethanol	4 ppb	300 ppb
VOC_{sum}	300 ppb	1200 ppb

Note that the analytical studies are based on average measurements (sampling time >1 h) according to ISO16000-6 in empty rooms. This results in considerable differences between analytical reference measurements according to the ISO standard and actual real-time measurements with MOS sensors in occupied rooms. Therefore, the concentration ranges are likely to be underestimated because emissions from people in the room as well as during, e.g., cooking and cleaning, are not considered. For the field tests, we performed different release tests to verify the quantification performance and to compare the MOS sensor system with analytical instruments. To cover the higher concentration range during these tests, additional calibration with a larger concentration range for a single gas was added to the calibration scheme, while the range for the remaining gases was kept to the background concentration. The extended calibration schemes for single gases included acetone (14–1000 ppb), toluene (4–1000 ppb), ethanol (4–1000 ppb) and hydrogen (400–4000 ppb) as these four gases were to be used in the release tests.

To test our assumption that a single compound could represent all VOCs of its chemical class we performed additional measurements substituting some gases; in the first test we replaced formaldehyde with acetaldehyde and, in a second, we additionally replaced toluene with benzene for a limited number of gas exposures. At the end of the measurement campaign, we also tested m/p-xylene as another representative for aromatic compounds and limonene as an example of a chemical class (terpenes) not previously included in the calibration. Table 3 gives an overview of the entire measurement campaign with pre-tests, initial calibration, recalibrations, and field test periods.

Table 3. Overview of all performed measurements in the laboratory.

	Measurement Description	Unique Gas Mixtures
Pre-tests	background with acetaldehyde instead of formaldehyde	60
	background with acetaldehyde instead of formaldehyde and benzene instead of toluene	15
Initial calibration	background only	100
	background with modified acetone range: 14–1000 ppb	100
	background with modified toluene range: 4–1000 ppb	100
	background with modified ethanol range: 4–1000 ppb	100
	background with modified hydrogen range: 400–4000 ppb	100
	1st field test period (4 weeks)	
1st Recalibration	background only	100
	background with modified acetone range: 14–1000 ppb	100
	background with modified toluene range: 4–1000 ppb	100
	background with modified ethanol range: 4–1000 ppb	100
	background with modified hydrogen range: 400–4000 ppb	100
	2nd field test period (3 weeks)	
2nd Recalibration	background only without toluene *	100
	background only with m/p-xylene instead of toluene	50
	background only with limonene instead of toluene	50

* The measurement was performed without toluene due to the delayed delivery of a test gas cylinder.

2.3. Field and Release Tests

The field tests were performed in a regular office in our building (Figure 3). The office has a floor area of 3.5 m × 6.3 m and a height of 2.8 m, thus a total volume of 61.8 m^3 (Room 2.30 in [30]). The room contains one door to a long corridor and, on the opposite side, one window. The furnishing includes one wall cabinet, three desks, three office chairs and two shelves. The flooring is carpet and the walls are wallpapered and painted. Due to the age of the furnishing, flooring and wall coverings of over 20 years, we did not expect high VOC emissions in this office. After the field tests, VOC analysis according to ISO 16000-6 and very volatile organic compound (VVOC) analyses, evaluated by a certified laboratory, obtained a TVOC concentration of 130 μg/m^3 in the room. The substances with the highest concentrations were n-hexadecane (25 μg/m^3) and acetic acid (9 μg/m^3). Two VVOCs were reported at the highest concentrations: 2-propanol with 66 μg/m^3 and ethanol with 21 μg/m^3. Probably due to the current COVID-19 pandemic, an increased use of disinfectants based on 2-propanol and ethanol contributed to this result. Figure 3 shows a schematic top view of the room indicating the locations of the measurement trolly, the location for the release tests and a fan to ensure continuous air circulation in the room.

VOC release tests were performed via evaporation of a certain volume of the target compound liquid at the location marked in Figure 3. The expected increase in concentration during the evaporation can be estimated with Equation (1):

$$\Delta C_{target} = V_{target,gas}/V_{room} \tag{1}$$

V_{room} is the volume of the room and $V_{target,gas}$ is the volume of the VOC after evaporation. $V_{target,gas}$ can be calculated with Equation (2), where n is the amount of substance, R the gas constant, T the room temperature, p the pressure, M the molar mass, m the released VOC mass, ρ the density and $V_{target, liquid}$ the volume of the VOC in liquid form:

$$V_{target, gas} = n \cdot R \cdot T \cdot p^{-1} \text{ with } n = m/M \text{ and } m = \rho/V_{target, liquid} \tag{2}$$

Note that the increase of the concentration according to Equation (1) is a theoretical value assuming homogeneous distribution in a sealed room without air exchange. The release tests have an estimated uncertainty of 10%, due to the accuracy of the pipette and the handling of the liquid (e.g., evaporation during the process).
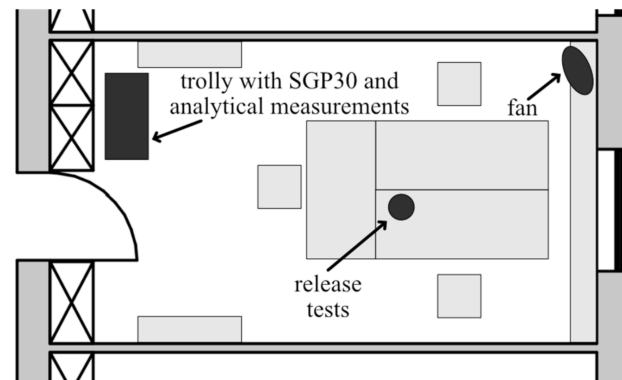

Figure 3. Schematic top view of the field test room, a standard office in our building. The locations of the trolly containing sensor systems and analytical measurements, release test and a fan are indicated (modified from [30]).

Furthermore, hydrogen was released at the same location from a pressure cylinder with a concentration of 2000 ppm in the air at a constant rate of 500 mL/min controlled by a mass flow controller for different durations. The estimated uncertainty is 4%, due to the dominating accuracy of the used gas cylinder compared to the accuracy of the MFC and time measurement. To be more comparable to the analytical studies, the field tests were performed without human presence as much as possible. However, the room had to be entered briefly for ventilation after release tests as well as to allow operation of the analytical systems or to collect samples. Analytical measurements were performed at the same location as the sensor measurements, cf. Figure 3.

2.4. Data Evaluation

Data evaluation of the gas sensor data is performed with the open-source software DAV^{3}E [31]. Figure 4 shows the flowchart of the data evaluation. The data evaluation is divided into two parts. The first part (left) is the calculation of the initial regression model (IRM) with feature selection and hyperparameter optimization. The second part (right) is the calculation of a drift compensated regression model (DCRM) with an additional recalibration dataset.

Both parts of data evaluation start with data preprocessing and feature extraction. We excluded the first four and the last temperature cycle in each gas exposure in the datasets to ensure stable gas mixtures, thus each tested gas mixture yields five patterns for data evaluation. The raw signal of the SGP30 is the sensor resistance of each layer. Based on our model concept for MOX gas sensors in TCO [19,20], the optimal signal for data evaluation is the logarithmic sensor conductance. Therefore, the preprocessed data is the common logarithm of the reciprocal sensor resistance. In the feature extraction, we divide each cycle into 120 equidistant segments. For each segment, mean and slope is calculated resulting in 240 features for each gas-sensitive layer of the SGP30 and a total of 960 for the sensor with 4 layers. Since, in some cases, the measurement range of the SGP30 is exceeded at low temperatures the features of those segments are excluded.
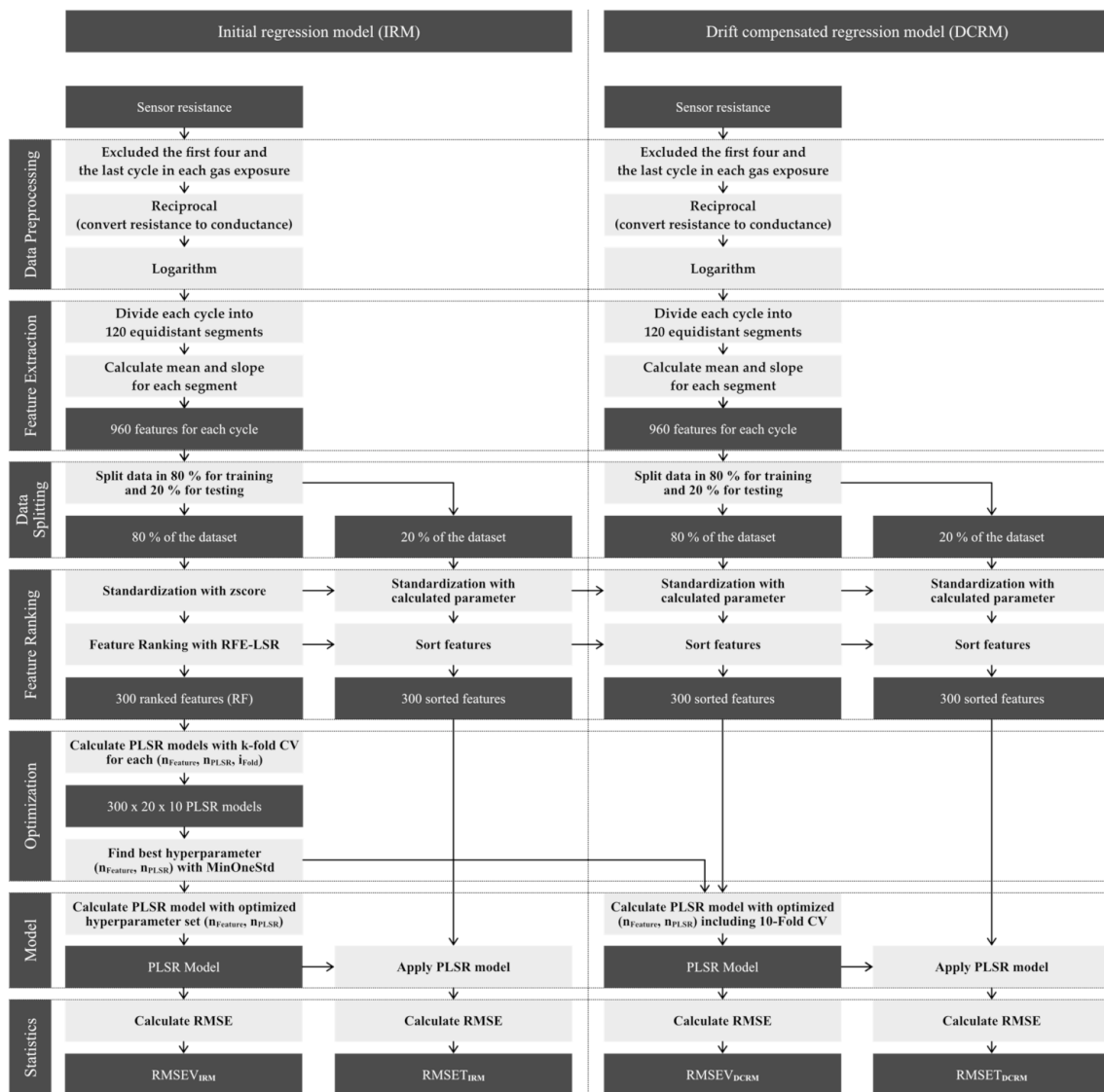
Figure 4. Flowchart of the data evaluation for an initial and drift-compensated regression model.

For the initial calibration, the dataset is split into trainings (80%) and testing (20%). Dimensionality reduction is performed by feature selection. In the feature selection, the 300 highest ranked features are selected with feature ranking. Feature ranking is done by (ordinary) least squares regression (LSR) with recursive feature elimination (RFE) to determine the relative weights or the importance of all features. The features are sorted according to their linear coefficients. A flowchart of the feature ranking can be found in the Appendix A (Figure A1). In the next step, we use partial least square regression (PLSR) as a learning algorithm for the regression model. For hyperparameter optimization (number of PLSR components n_{PLSR} and number of the features $n_{feature}$), 10-fold group-based cross-validation [32] is performed, where the folds are determined based on gas exposures and not on individual temperature cycles. This ensures that complete gas exposures are used as validation data, i.e., the validation does not only check for overfitting but also for the ability of the model to correctly interpolate between various gas mixtures. A flowchart of the learning algorithm with k-fold cross-validation and hyperparameter optimization can also be found in the Appendix A (Figure A1). Iteratively, for each combination of (n_{PLSR}, $n_{feature}$, i_{fold}) a PLSR model is calculated with 1 to 20 PLSR components, 1 to 300 features and 10 folds. The root mean square error of validation for the initial regression model ($RMSEV_{IRM}$) is calculated as the mean over all folds for each

combination (n_{PLSR}, $n_{feature}$). The optimal combination of PLSR components and features is determined from the resulting $RMSEV_{IRM}$ matrix with a dimension of 20×300 (number of PLSR components $\times$ selected features). Therefore, we defined a criterion to find a stable and good model with a small number of dimensions: MinOneStd [26]. MinOneStd searches the absolute minimum of the matrix and adds the standard deviation as the threshold. The combination with the minimum product of number of features, times PLSR components, where the $RMSEV_{IRM}$ is smaller than this threshold, is selected as the optimal combination. With this optimal combination, 20% holdout of the dataset is tested to determine the root-mean-square error of testing ($RMSET_{IRM}$).

To compensate for the drift of the sensor, a regression model is calculated with the additional recalibration dataset (initial calibration and only background of 1st recalibration), but without new feature ranking and hyperparameter optimization. The data preprocessing and feature extraction are the same as for the initial calibration. The data is also split in the training (80%) and testing dataset (20%) for statistics. Features are sorted with the trained feature ranking from the initial calibration. The PLSR model is trained with an optimized hyperparameter from the initial calibration and the resulting $RMSEV_{DCRM}$ is calculated. With the new regression model, 20% holdout of the dataset is tested to determine the $RMSET_{DCRM}$.

3. Results

Results for hydrogen calibration and field tests including a comparison to the analytical instrument were recently published [25]; in this contribution, we focus primarily on selective VOC quantification and the overall VOC concentration, VOC_{sum}.

3.1. Calibration and Recalibration

For the generation of the prediction models for different targets, we used the dataset of the initial calibration and first recalibration (background only, i.e., without higher concentration exposures). One sensitive layer of the SGP30 gas sensor shows a small drift of the raw signal (logarithmic resistance) over time. Therefore, to compensate for this, but also other drift effects which are not as obvious, a part of the first recalibration after four weeks was included in the calibration data to optimize the model for drift stability as previously reported [33]. Figure 5 shows different prediction models for VOC_{sum} proving this approach using extended calibration to compensate drift: (a) and (b) trained with the initial calibration dataset only and (c) trained with extended calibration set (initial calibration combined with background only of the 1st recalibration). Figure 5a shows a stable and linear VOC_{sum} prediction model for training data and test data, i.e., 20 % holdout of the calibration dataset. Prediction of the 1st recalibration dataset reveals good linear correlation, but with an offset of approx. 200 ppb and a somewhat larger RMSET, Figure 5b. By extending the training dataset with the first part of the 1st recalibration (background only, i.e., only low concentrations) the model in Figure 5c is obtained. It yields comparable prediction results as the initial calibration, Figure 5a, also for the additional gas exposure with higher concentrations from the 1st recalibration and for the 2nd recalibration. Compared to Figure 5b, the offset between the training and testing data is eliminated, only the RMSET is approx. doubled. Thus, the extended calibration provides a stable model for the VOC_{sum} prediction for the total duration of this study, i.e., at least 11 weeks. The prediction models of the other target gases reveal similar results. Figure 6 provides an overview of the RMSE of all prediction models for the 10-fold validation and 20% holdout testing for the initial and the drift compensated PLSR model. The smallest RMSE values are achieved for acetone with approx. 10 ppb followed by formaldehyde, ethanol, and toluene with 20–35 ppb for validation and testing in the initial calibration. The RMSEs for the models of hydrogen and VOC_{sum} are in the range of 30–40 ppb. The worst prediction is obtained for carbon monoxide with an RMSE of approximately 80 ppb, because no sensitive layer of the SGP30 shows a high sensitivity to carbon monoxide. The drift compensated model compared to the initial PLSR model shows similar RMSE values for acetone, toluene,

hydrogen, and VOC_{SUM}. Ethanol, formaldehyde, and carbon monoxide show slightly higher RMSE values. For formaldehyde, it can be probably be explained with the gas cylinder change between the initial and the 1st re-calibration. The formaldehyde cylinders have a large systematic uncertainty of nearly 20% and in previous investigations [34] we saw the same behavior. Compared to the tested target ranges for the single VOCs we achieved a dynamic range [26] between 10 to 20 even for the low background level with 300–400 ppb; the highest dynamic range (>100) is achieved for hydrogen with an RMSET of approx. 35 ppb for concentrations up to 4000 ppb. In Table 4 the $RMSET_{DCRM}$ and the estimated accuracy and precision of the GMA are shown. The accuracy and precision depend on the MFC opening settings during the measurement. Therefore, the ranges—in percent of the set concentration and in parts per billion (ppb)—are shown. The GMA accuracy is dominated by the gas pressure cylinders. The $RMSE_{DCRM}$ is larger compared to the expected precision of the GMA. This indicates that the uncertainty of the models is due to cross-sensitivity to the other gases or other sensor effects, but not from the GMA.
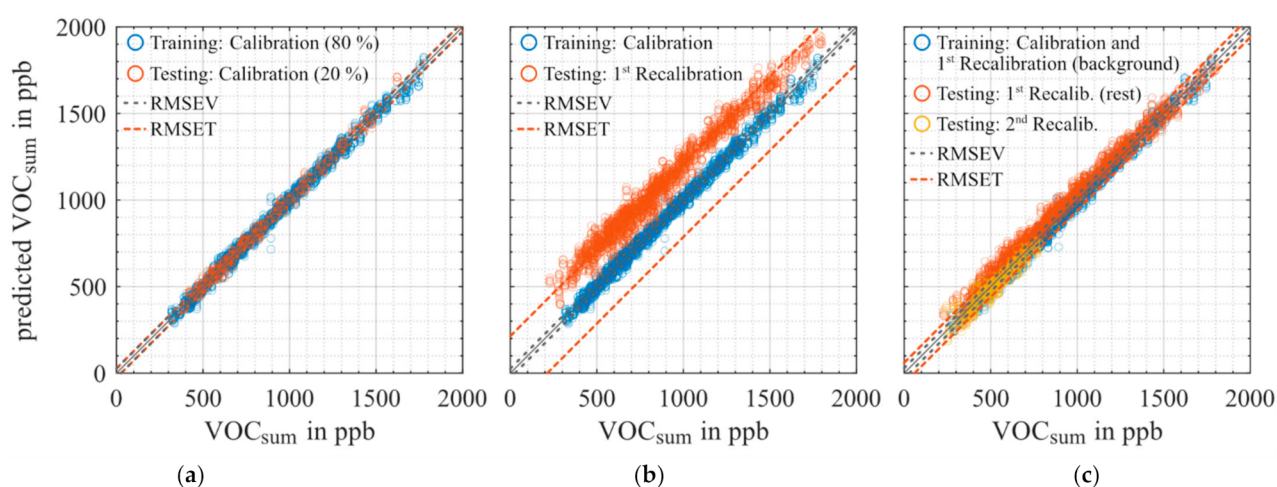

Figure 5. PLSR model for quantification of VOC_{sum} for (**a**) training and testing with data from the initial calibration (initial regression model), (**b**) training with initial calibration, testing with 1st recalibration, (**c**) the drift compensated regression model (training with initial calibration plus background only of the 1st recalibration), testing with extended range data from 1st recalibration as well as 2nd recalibration. Dashed lines indicate the root-mean-square-error of validation (RMSEV) based on the training data set and of testing (RMSET), respectively.
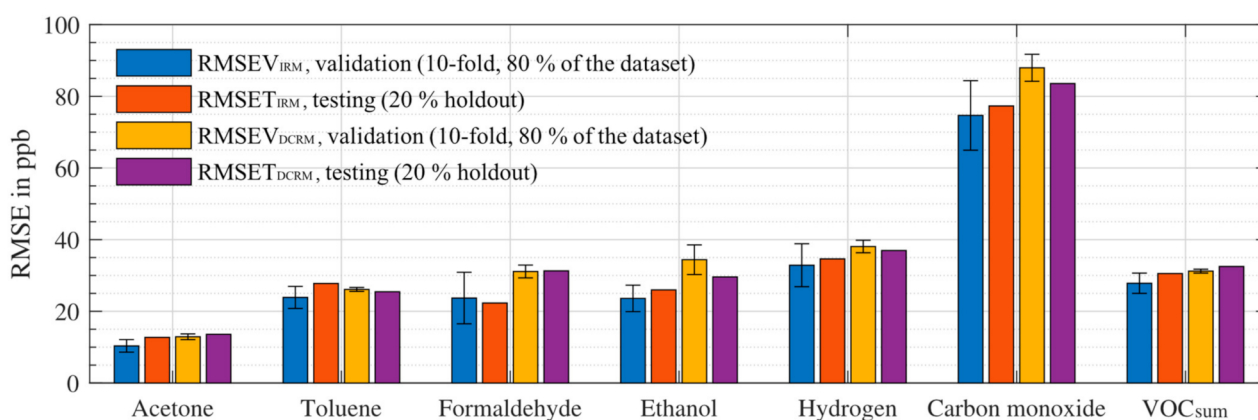

Figure 6. RMSE of the models for different target VOCs for the initial ($RMSEV_{IRM}$, $RMSET_{IRM}$) and the drift compensated regression model ($RMSEV_{DCRM}$, $RMSET_{DCRM}$). For each model RMSEV for 10-fold validation during training (error bars indicate the standard deviation of the RMSEV for the different folds) and testing (20% holdout).

Table 4. RMSET$_{DCRM}$ for different prediction models and the accuracy and precision of the calibration measurements. The accuracy and precision depend on the MFC settings during the measurement. The accuracy is dominated by the used gas pressure cylinders, which have an accuracy of between 3 and 20%.

Substance	RMSE$_{DCRM}$ in ppb	GMA Accuracy [1] in % (ppb)	GMA Precision [1] in % (ppb)
Acetone	13.6	5.0–6.5 (1–50)	0.7–4.1 (1–6)
Toluene	25.5	2.1–2.8 (0.1–23)	0.5–1.8 (0.1–10)
Formaldehyde	31.3	20.0–20.3 (0.3–82)	0.6–3.5 (0.1–10)
Ethanol	29.6	3.1–4.7 (0.2–35)	0.6–3.5 (0.2–17)
VOC$_{sum}$	32.5	1.8–15 (8–93)	0.6–1.8 (2–19)
Carbon Monoxide	83.6	2.2–4.2 (6–49)	0.9–3.7 (6–20)
Hydrogen	37.0	2.1–3.9 (16–85)	0.5–3.3 (14–23)

[1] Accuracy and precision are calculated by the propagation of uncertainty from the accuracy and precision of the MFCs and gas cylinder used for the calibration measurements.

3.2. Field Tests

During the time in the field, we performed 17 release tests, mostly by evaporation of VOCs, but also using test gas bottles and MFCs as well as burning a tea candle. Table 5 provides an overview of all release tests giving the start time, substance, type of release and the idealized concentration increase in the room calculated using Equation (1). A complete list of all events, including persons entering the room, ventilation etc. is given in Table A1.

Table 5. Overview of calibration, recalibration, and all release tests. The complete list of all events is given in Table A1.

Release	Event	Time	Substance (Type of Release)	Released Amount of Substance (Approx. Increase in Room Conc.)
			Pre-tests and Initial calibration	
1	11	06 October, 17:42	Hydrogen (MFC, gas cylinder)	2000 ppm @ 500 mL/min for 62 min (~1 ppm ± 4%)
2	12	07 October, 16:01	Hydrogen (MFC, gas cylinder)	2000 ppm @ 500 mL/min for 124 min (~2 ppm ± 4%)
3	16	13 October, 15:00	Toluene (MFC, gas cylinder)	100 ppm @ 500 mL/min for 497 min (~300 ppb ± 10%)
5	22	16 October, 14:50	Acetone (evaporation) Toluene (evaporation)	0.114 mL (~600 ppb ± 10%) 0.164 mL (~600 ppb ± 10%)
6	23	16 October, 18:00	Acetone (evaporation) Toluene (evaporation)	0.114 mL (~600 ppb ± 10%) 0.164 mL (~600 ppb ± 10%)
			1st Recalibration	
7	28	02 November, 16:50	Toluene (evaporation)	0.164 mL (~600 ppb ± 10%)
9	32	04 November, 16:22	Acetone (evaporation)	0.114 mL (~600 ppb ± 10%)
10	34	05 November, 15:10	Acetone (evaporation) Toluene (evaporation)	0.114 mL (~600 ppb ± 10%) 0.164 mL (~600 ppb ± 10%)
11	36	06 November, 10:03	Limonene (evaporation)	0.251 mL (~600 ppb ± 10%)
12	39	09 November, 18:00	Ethanol (evaporation)	0.1 mL (~664 ± 10%)
13	41	10 November, 14:30	Isopropyl alcohol (evaporation)	0.12 mL (~600 ppb ± 10%)
14	43	11 November, 15:49	m/p-Xylene (evaporation)	0.189 mL (~600 ppb ± 10%)
15	45	12 November, 15:08	Toluene (evaporation) m/p-Xylene (evaporation)	0.164 mL (~600 ppb ± 10%) 0.189 mL (~600 ppb ± 10%)

Table 5. *Cont.*

Release	Event	Time	Substance (Type of Release)	Released Amount of Substance (Approx. Increase in Room Conc.)
16	47	13 November, 14:30	Acetone (evaporation) Toluene (evaporation) Ethanol (evaporation)	0.114 mL (~600 ppb $\pm$ 10%) 0.164 mL (~600 ppb $\pm$ 10%) 0.1 mL (~664 ppb $\pm$ 10%)
17	50	16 November, 17:06	Hydrogen (MFC, gas cylinder)	2000 ppm @ 500 mL/min for 134 min (~2 ppm $\pm$ 4%)
18	52	17 November, 18:24	Ethanol (evaporation)	0.1 mL (~664 ppb $\pm$ 10%)
19	54	19 November, 12:02	Carbon monoxide etc. (tea candle)	4 h burn time
			2nd Recalibration	

The presented signals are based on the drift compensated PLSR model (DCRM). Note, that with this model, we are using the future to predict the past for the first field tests (release #1–#6). The release tests in the second field test (release #7–#19) were all conducted after the 1st recalibration.

Figure 7 shows results recorded during release tests for toluene (release test #7), acetone (#9), ethanol (#10), and the simultaneous release of all three (#16). In a sealed room with homogeneous distribution, the release of 0.164 mL toluene should lead to an increase of the toluene concentration of approx. 600 ppb $\pm$ 10 %. Since the amount released and the homogeneous distribution in the room may vary, there may be deviations in the level of the expected concentration. With the start of the toluene release, the MOS sensor model for toluene indicates a quick increase from nearly zero to 620 ppb. After full evaporation of the toluene, the model prediction slowly decreases again over several hours. The X-pid 9500 shows a similar course of the toluene signal as the MOS sensor model, but ~150 ppb higher; the increase vs. the base level before release is approx. 700 ppb. Note that the manufacturer gives a limit of quantification (LOQ) for toluene of 1000 ppb. The model predictions for the other target gases show only small changes with the onset of the evaporation and nearly constant results afterward. Only the VOC$_{sum}$ model indicates an increase of approximately 600 ppb, thus a consistent prediction. Note that calculating the sum of the four individual VOC model predictions (dashed line) yields a similar increase with a small offset of approx. 50 ppb. No statement can be made about the true absolute concentration since these releases were not accompanied by any analytical reference for this concentration range. However, the MOS sensor model and X-pid 9500 show similar signals in the same order of magnitude of the expected concentration for the release tests.

Similarly, the release of acetone and ethanol show an increase in the corresponding prediction models. The acetone model with a higher base level of approximately 120 ppb indicates an increase of 350 ppb to a peak value of 570 ppb. The same increase can be observed in the VOC$_{sum}$ model as well as the calculated sum of the individual VOC signals. Again, the other model does not show a reaction and remains nearly constant, except for carbon monoxide and hydrogen. Carbon monoxide shows an increase of nearly 150 ppb after the start of evaporation and hydrogen increases during the acetone signal decrease. Note that, during the field tests, the hydrogen signal shows more variations than all other models [25]. Similar to toluene, the acetone signal of the X-pid 9500 shows a higher increase of the concentration but confirms the course vs. time.
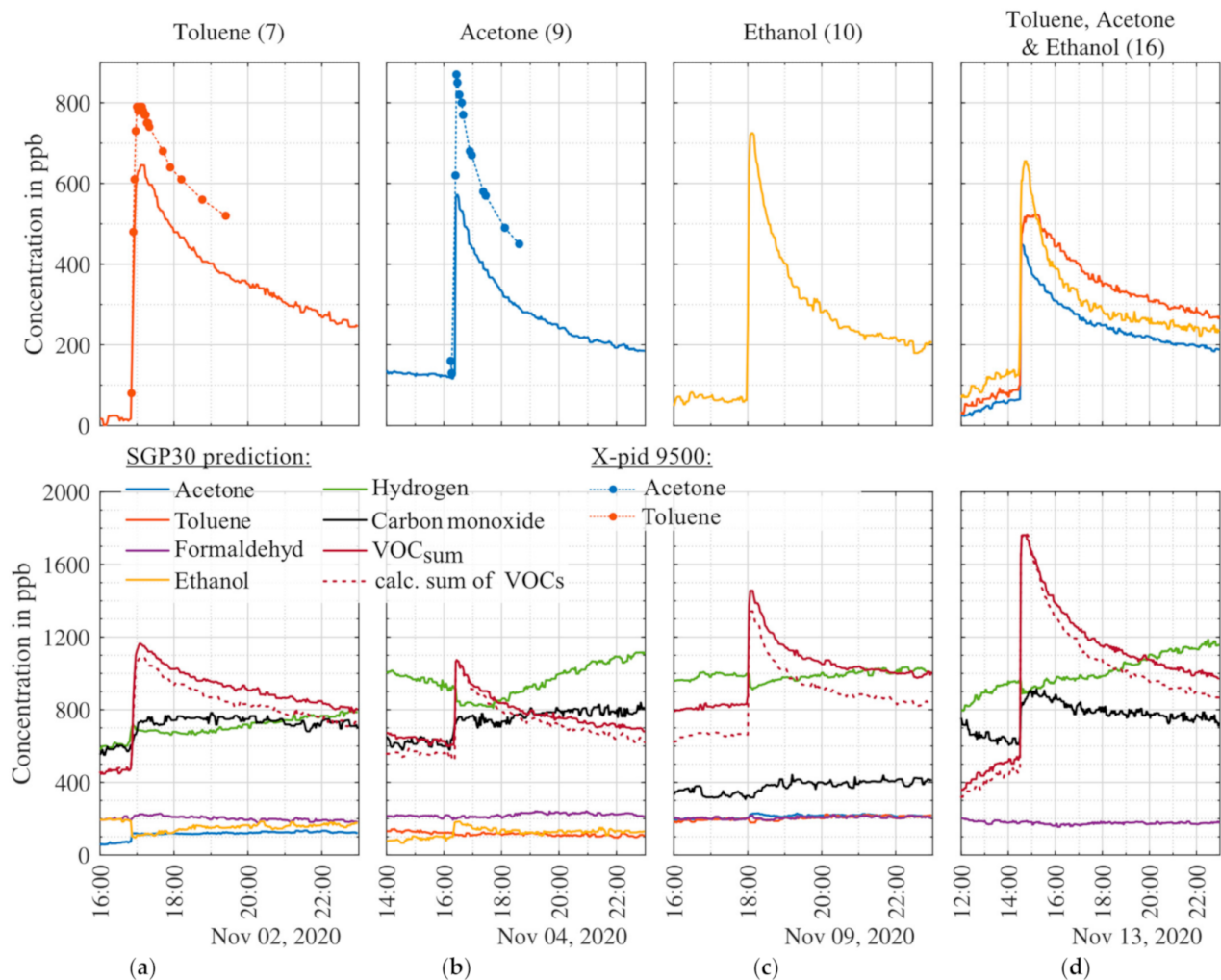
Figure 7. Results recorded during four release tests with toluene (**a**), acetone (**b**), ethanol (**c**), and the simultaneous release of toluene, acetone, and ethanol (**d**). The upper graphs show the PLSR prediction of the MOS sensor model for the released gases and, if available, reference data (dots); all other signals are shown in the lower graphs. The signals are smoothed over five points (10 min). Numbers in parenthesis behind the released substances refer to the release tests, cf. Table 4.

The ethanol release test shows an increase of 660 ppb (expected 664 ppb $\pm$ 10%). At the start of the evaporation, the hydrogen signal decreases by nearly 100 ppb, while all other single target signals remain constant. The VOC_{sum} signal increases from 830 ppb to 1455 ppb, corresponding to an increase of 625 ppb, again very similar to the ethanol signal itself. The sum of the four single VOC signals is lower with an offset of approx. 180 ppb.

In release test #16 we tested the simultaneous evaporation of all three substances: toluene (~600 ppb $\pm$ 10%), acetone (~600 ppb $\pm$ 10%) and ethanol (~664 ppb $\pm$ 10%). The toluene model shows an increase of 380 ppb, acetone of 430 ppb, and ethanol of 530 ppb. All three VOC models yield consistently lower concentrations compared to the individual release tests. This might be because during calibration only one gas at a time had higher concentrations and, thus, the models have to extrapolate the prediction beyond the calibrated range. The VOC_{sum} model prediction as well as the sum of the four single VOC models shows similar increases.

In Figure 8 two release tests with hydrogen and two with acetone and toluene are shown. The hydrogen releases were designed to yield an increase of approx. 2 ppm in the room. Because the hydrogen molecule is very small and has a high diffusion constant, we expect a somewhat faster diffusion out of the room and, thus, a smaller peak. The first hydrogen release (2) yields an increase of the model prediction of 1440 ppb, the second (17) of 1500 ppb. The second release was also monitored with the hydrogen measurement

system (GC-RCP). A high correlation between analytical and MOS sensor model prediction can be observed. Compared to the MOS model prediction the GC-RCP indicates a nearly identical increase of 1490 ppb, but with a constant offset of 150 ppb. Other signals show minimal changes except for Carbon monoxide wherein both releases a small change can be observed. The ethanol model shows an inverse effect during the second release, but no reaction during the first release.
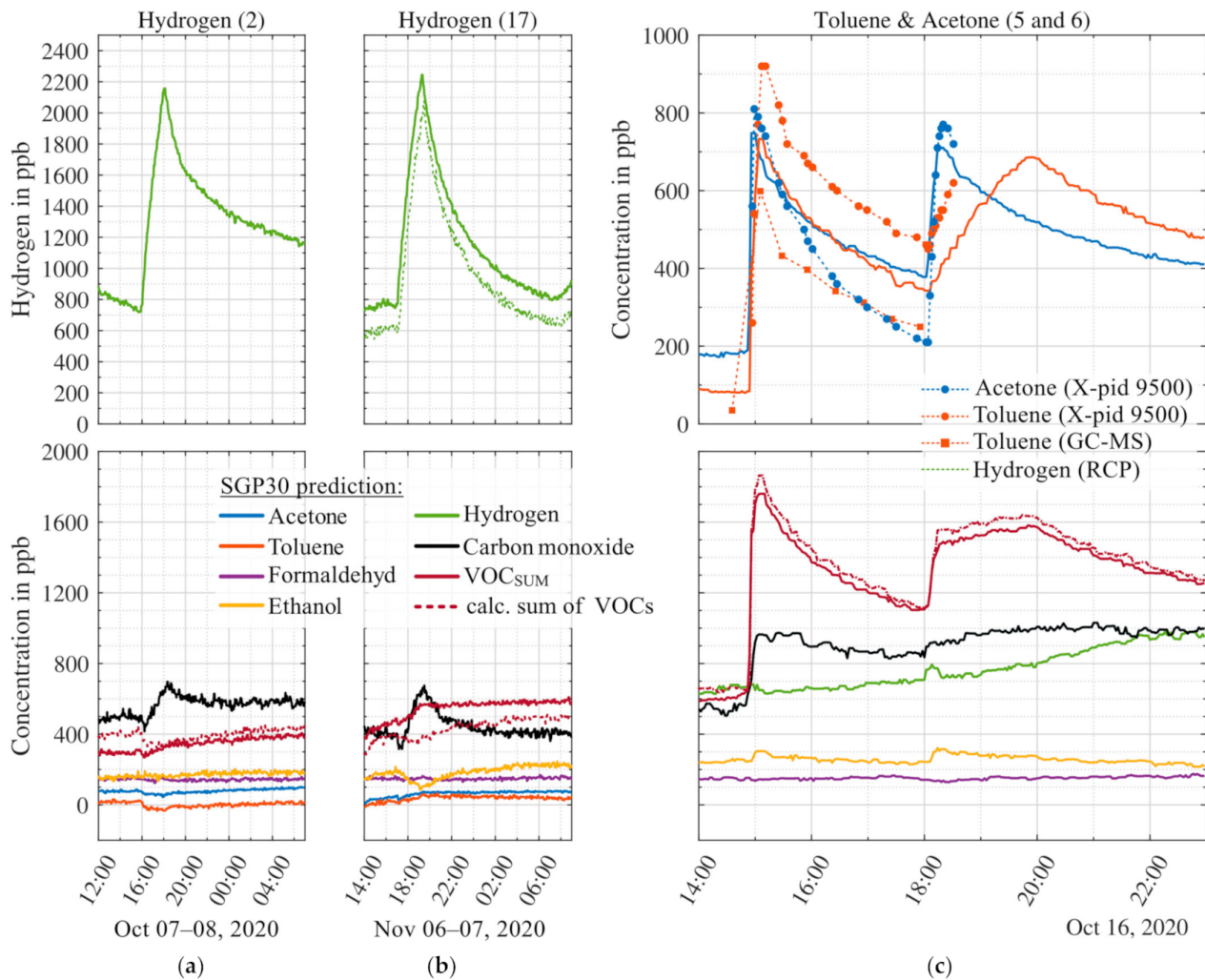

Figure 8. Results recorded during three release tests with hydrogen (**a**,**b**) and the simultaneous release of toluene and acetone (**c**). The upper graphs show the PLSR prediction of the MOS sensor model for the released gases and, if available, reference data (dots indicate distinct sampling times); all other signals are shown in the lower graphs. The signals are smoothed over five points (10 min). Numbers in parenthesis behind the released substances refer to the release tests, cf. Table 4.

Figure 8c shows two release tests with acetone (~600 ppb ± 10%) and toluene (~600 ppb ± 10%). The first toluene peak shows an increase of 600 ppb above the baseline level, similar to the result shown in Figure 7 for the same amount of substance released. During the first release in Figure 8c samples were taken with Tenax sampling tubes for analysis by GC-MS in addition to the X-pid 9500 measurements. The X-pid 9500 indicates an increase of 920 ppb, thus slightly higher compared to release #7, cf. Figure 7. The GC-MS analysis, on the other hand, yields an increase of 560 ppb. Thus, comparing the MOS sensor with X-pid 9500 and GC-MS, the toluene concentration predicted by the model is much closer to the GC-MS. Note that the GC-MS and the MOS sensor are calibrated with gas

mixtures from the GMA with the same gas cylinders, and therefore, the accuracy of the gas cylinder (which brings along the highest uncertainty) has no influence on the comparison.

All three signals show the same time temporal development. The first acetone release in Figure 8c yields an increase of 570 ppb, again comparable to release #9 in Figure 7. The X-pid 9500 again yields a higher absolute acetone signal. The second release shows the same trends as before, only the toluene evaporation is slower in comparison to the first release. The reason for the different evaporation and diffusion speed can be a lower ambient temperature, because of the experiment being performed later in the day. The increase of both signals is nearly the same as during the first release. The VOC$_{sum}$ model prediction also indicates an increase due to the release of acetone and toluene and corresponds to the sum of the four single VOC signals. Other than as observed during the release of the triple mixture (toluene, acetone, and ethanol), cf. Figure 7, the model signals during the release of the double mixture (toluene and acetone) are higher and comparable to the release tests with single gases.

3.3. Uncalibrated Substances

In Section 2.2 we described the general idea of the calibration scheme based on the selection of representatives for different chemical classes to simplify the VOC composition. One assumption is that the substances of a certain class react similarly on MOS sensor surfaces yielding a similar response patterns in the TCO and thus all VOCs of a type can be represented by one specific compound. In order to test this assumption, additional substances of chemical classes previously included and also not included in the calibration were tested. Figure 9 shows release tests with two substances not included in the calibration: m/p-xylene (aromatic) and isopropyl alcohol. The chemical class aromatics was represented in the calibration by toluene. Indeed, the m/p-xylene release, Figure 9b, results in an increase of the toluene signal, i.e., the MOS sensor model trained for toluene. The corresponding toluene signal indicates an increase of only 460 ppb, compared to 630 ppb for the same amount of toluene #7. In addition, the carbon monoxide signal shows a slight increase and the same trend as the toluene signal. Thus, for m/p-xylene and toluene as two aromatic compounds our assumption is confirmed, but with different response factors and an additional interference with the carbon monoxide signal. Note that this approach is similar to quantifying unknown substances with the response factor of toluene in GC-MS analysis (ISO16000-6). The simultaneous release of toluene and m/p-xylene (#15) results in an increase of the MOS sensor model of 910 ppb.

For the chemical class alcohol, only ethanol was contained in the calibration. A release test with isopropyl alcohol was performed to check the reaction of the various model to this second alcohol. While the X-pid 9500 confirms the release, Figure 9d, the MOS sensor models for ethanol and all other targets stay constant, although we observe a reaction to isopropyl alcohol in the raw sensor data. This means that the sensor does react to isopropyl alcohol but that the models, especially the model for ethanol, compensates for this reaction. Thus, the assumption of similar reaction patterns is not valid in this case and ethanol is not suitable to represent the chemical class of alcohol, at least not alone.
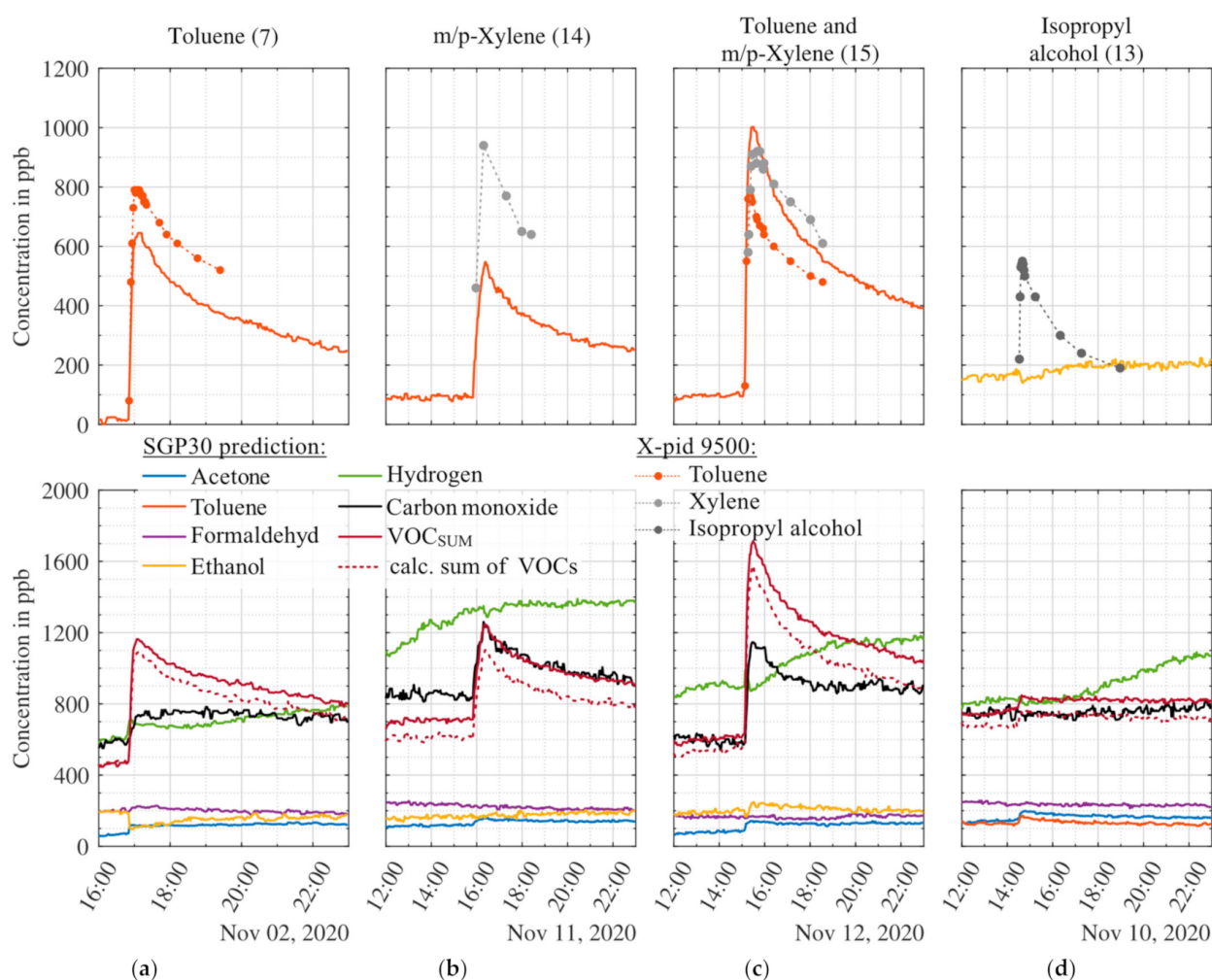
Figure 9. Results recorded during four release tests with toluene (**a**), m/p-xylene (**b**), simultaneous release of toluene and m/p-xylene (**c**), and isopropyl alcohol (**d**). The upper graphs show the PLSR prediction of the MOS sensor model for the same types of VOC and, if available, reference data (dots); all other signals are shown in the lower graphs. The signals are smoothed over five points (10 min). Numbers in parenthesis behind the released substances refer to the release tests, cf. Table 4.

4. Discussion

In this study, a SGP30 sensor in TCO was successfully calibrated for VOC quantification, both for the overall sum and selective signals, using a randomized calibration scheme in the laboratory. The randomized calibration scheme was based on our previous study [26] with an improved randomized gas mixture generation based on Latin Hypercube sampling. The calculated models yield low RMSE values for different VOC targets based on the lab measurements. The performance of the models is similar to those achieved previously with other MOS sensor types (AS-MLV and AS-MLV-P2, ScioSense B.V., Eindhoven, The Netherlands) [26]. Both studies are not completely identical due to some different gases being used but the results indicate that the SGP30 achieves lower RMSE values for all gases except carbon monoxide. This can be attributed to the higher information obtained from the four different gas-sensitive layers of the SGP30, all of which show only low sensitivity to carbon monoxide. For VOC measurements in indoor air, this might be beneficial, because large variations of the carbon monoxide concentration are possible in room air. Sensor drift, which was especially obvious for one layer off the SGP30, could be effectively eliminated from the models by extended calibration based on two GMA measurements spread over a period of several weeks.

The field tests show quantitative and repeatable results for VOC release tests which were performed by the evaporation of different substances. This method has proven to also be a reliable option for simple verification of the sensor performance. Release tests with substances included in the calibration (toluene, acetone, ethanol, and hydrogen) show concentration increases close to theoretically expected values. Analytical measurements with GC-MS, GC-PID and GC-RCP show the same temporal course during the release tests. Absolute concentrations obtained from the MOS sensor model prediction and the analytical systems are similar but reveal some offsets, also between the different analytical systems. However, these offsets are not higher than normally expected for trace gas measurements even using high-cost lab analysis. Compared to GC-MS, the X-pid 9500 provides better temporal resolution but has a high LOQ, higher than the concentrations tested here. For an exact time-resolved quantification, further analytical measurements with optimized sampling methods for the GC-MS or other analytical measurement systems, like PTR-MS, are required. The difference between the MOS sensor toluene model and the GC-MS, which is the gold standard in VOC analysis, is small (<100 ppb) and similar to the RMSE value determined during calibration. One reason can be that the GC-MS and the MOS sensor are calibrated with gas mixtures from the GMA and the same gas cylinders. Therefore, the accuracy of the gas cylinder (with the highest uncertainty) has no influence on the comparison.

During the release test of hydrogen, the GC-RCP consistently indicated approx. 150 ppb lower concentrations than the hydrogen model of the SGP30. In fact, during ventilation of the room, the GC-RCP indicated a concentration of less than 500 ppb, i.e., below the atmospheric background, indicating that the GC-RCP is underestimating the actual hydrogen concentration. The difference of the two systems is within the error range of the two systems; the RCP has an accuracy of 10% and the MOS sensor system at least 3–5% (uncertainty of the gas mixtures for calibration; model, stability and drift of the system are not considered).

We tested the assumption that substances of the same chemical class react similarly on the sensor surface and can therefore be represented by one single compound. Release tests with m/p-xylene and a laboratory test with benzene indeed showed a reaction of the toluene sensor model, indicating that this model does indeed represent all aromatic compounds, although with different response factors. On the other hand, this means that selectively measuring individual aromatics independent of each other needs further investigation and will at least require more comprehensive calibration. The second chemical class-tested was alcohols where calibration was based on ethanol and a release test was performed with isopropyl alcohol. However, other than for the three aromatics, the ethanol model does not respond to isopropyl alcohol. In the raw sensor signals, a reaction towards isopropyl alcohol was observed but the gases obviously have different reaction processes leading to different sensor response patterns. The approach with a single representative is not valid for this type of VOC, which means that at least two alcohols will be needed for a valid calibration.

5. Conclusions

In this study, we have demonstrated that using MOS gas sensor systems can provide quantitative and selective results not only in the laboratory but also in field measurements as demonstrated by release tests accompanied by analytical measurements. TCO dynamic operation, randomized calibration, and optimized model training are suggested as necessary and practical tools for achieving this performance with commercially available sensor elements. We were able to successfully demonstrate that the sensor can measure calibrated substances in real-time selectively and quantitatively while being released in a room. Also, further investigations about the metrological accuracy or precision and long-term stability of the sensor system are required. Two contrary behaviors concerning the approach of detecting VOC by type were observed, so further work on this approach is required to simplify calibration for complex environments. Even more important for industrial appli-

cation of the demonstrated elaborate calibration, as presented in this manuscript, is the optimization of the model stability without the need for a 2nd calibration after some time in the field. While the approach using extended calibration yields excellent results, this is a very inefficient approach, at least for the calibration of high volumes. Therefore, a study and optimization of long-term stable features and models is necessary. Also, the transfer of the feature selection and of full evaluation models between sensors of the same type should be investigated.
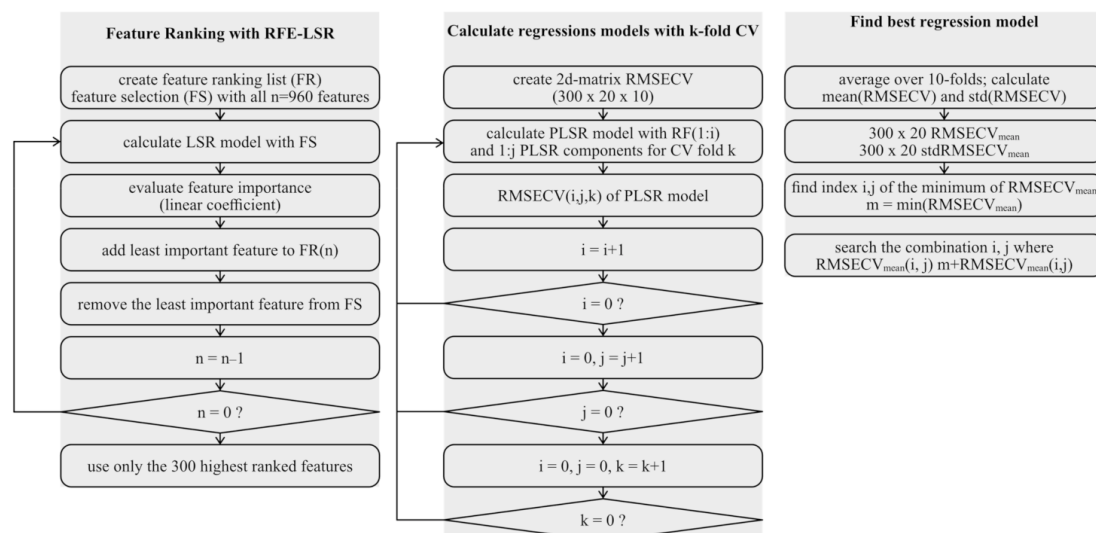
Appendix A

Figure A1. Flowchart of different data evaluation steps. Indices i: number of features, j: number of PLSR components and k: current fold during 10-fold cross validation.

Table A1. Overview of all release tests and events during the field tests.

Number	Time	Type of Event
1	29 September, 09:35–09:48	Door opened
2	30 September, 09:24–09:58	Window opened
3	01 October, 09:10–09:30	Window opened
4	01 October, 11:47–12:05	Door and window opened
5	01 October, 18:30–02 October, 06:30	No specifiable event
6	2 October, 09:00–09:30	Door and window opened
7	02 October, 14:00–05 October, 10:00	Days without events and human presence
8	05 October, 10:10–10:30	Door and window opened
9	05 October and 05 October	Several short periods of human presence
10	06 October, 13:08–16:51	Door and window opened
11	06 October, 17:42–18:44	Release test: 1 ppm H_2
12	07 October, 16:01–18:05	Release test: 2 ppm H_2
13	08 October, 10:46–11:00	Door and window opened
14	08 October to 13 October	Days without events and human presence
15	13 October, 09:25–14:00	Door and window opened, human presence
16	13 October, 15:00	Release test: toluene
17	14 October, 09:30–10:05	Door and window opened
18	08 October to 13 October	Sporadic human presence, no specifiable events
19	15 October, 09:00–09:30	Door and window opened
20	15 October, 15:00	Release test: acetone
21	16 October, 09:40–10:10	Door and window opened
22	16 October, 14:50	Release test: acetone and toluene
23	16 October, 18:00	Release test: acetone and toluene
24	17 October to 19 October	Days without events and human presence
25	29 October, 12:55–13:10	Door and window opened
26	29 October to 02 November	Days without events and human presence
27	02 November, 12:40–12:55	Door and window opened
28	02 November, 16:50	Release test: toluene
29	03 November, 10:55–11:10	Door and window opened
30	03 November, 15:30	Release test: acetone followed by defect of the pump, human presence during fixing
31	04 November, 09:00–09:15	Door and window opened
32	04 November, 16:22	Release test: acetone
33	05 November, 09:26–09:41	Door and window opened
34	05 November, 15:10	Release test: acetone and toluene. Unidentified event due to construction inside the building
35	05 November, 18:30–18:50	Door and window opened
36	06 November, 10:03	Release test: limonene
37	06 November to 09 November	Days without events and human presence
38	09 November, 12:21–13:01	Door and window opened

Table A1. *Cont.*

Number	Time	Type of Event
39	09 November, 18:00	Release test: ethanol
40	10 November, 09:10–09:25	Door and window opened
41	10 November, 14:30	Release test: isopropyl alcohol
42	11 November, 09:28–09:48	Door and window opened
43	11 November, 15:49	Release test: m/p-xylene
44	12 November, 09:15–09:30	Door and window opened
45	12 November, 15:08	Release test: toluene and m/p-xylene
46	13 November, 09:28–11:06	Door and window opened
47	13 November, 14:30	Release test: acetone, toluene, and ethanol
48	13 November–16 November	Days without events and human presence
49	16 November, 11:55–12:20	Door and window opened
50	16 November, 17:06–19:20	Release test: 2 ppm H_2
51	17 November, 09:54–10:24	Door and window opened
52	17 November, 18:24	Release test: ethanol
53	18 November, 09:36–09:56	Door and window opened
54	19 November, 12:02–16:02	Release test: carbon monoxide (tea candle)

References

1. Asikainen, A.; Carrer, P.; Kephalopoulos, S.; Fernandes, E.D.O.; Wargocki, P.; Hänninen, O. Reducing burden of disease from residential indoor air exposures in Europe (HEALTHVENT project). *Environ. Health* **2016**, *15*, S35. [CrossRef] [PubMed]
2. Settimo, G.; Manigrasso, M.; Avino, P. Indoor Air Quality: A Focus on the European Legislation and State-of-the-Art Research in Italy. *Atmosphere* **2020**, *11*, 370. [CrossRef]
3. Spaul, W.A. Building-related factors to consider in indoor air quality evaluations. *J. Allergy Clin. Immunol.* **1994**, *94*, 385–389. [CrossRef] [PubMed]
4. Herberger, S.; Herold, M.; Ulmer, H.; Burdack-Freitag, A.; Mayer, F. Detection of human effluents by a MOS gas sensor in correlation to VOC quantification by GC/MS. *Build. Environ.* **2010**, *45*, 2430–2439. [CrossRef]
5. Buszewski, B.; Kesy, M.; Ligor, T.; Amann, A. Human exhaled air analytics: Biomarkers of diseases. *Biomed. Chromatogr.* **2007**, *21*, 553–566. [CrossRef]
6. Veres, P.R.; Faber, P.; Drewnick, F.; Lelieveld, J.; Williams, J. Anthropogenic sources of VOC in a football stadium: Assessing human emissions in the atmosphere. *Atmos. Env.* **2013**, *77*, 1052–1059. [CrossRef]
7. Pettenkofer, M. Über den Luftwechsel in Wohngebäuden; Literarisch-Artistische Anstalt der J.G. *Cotta'schen Buchhandlung*. 1858. Available online: https://opacplus.bsb-muenchen.de/title/BV013009721 (accessed on 17 May 2021).
8. Liu, Y.; Misztal, P.K.; Xiong, J.; Tian, Y.; Arata, C.; Weber, R.J.; Nazaroff, W.W.; Goldstein, A.H. Characterizing sources and emissions of volatile organic compounds in a northern California residence using space- and time-resolved measurements. *Indoor Air* **2019**, *29*, 630–644. [CrossRef]
9. Madou, M.J.; Morrison, S.R. *Chemical Sensing with Solid State Devices*; Academic Press, Inc.: San Diego, CA, USA, 1989; ISBN 978-0-12-464965-1.
10. Bârsan, N.; Weimar, U. Understanding the fundamental principles of metal oxide based gas sensors; the example of CO sensing with SnO 2 sensors in the presence of humidity. *J. Phys. Condens. Matter* **2003**, *15*, 813–839. [CrossRef]
11. Rüffer, D.; Hoehne, F.; Bühler, J. New digital metal-oxide (MOx) sensor platform. *Sensors* **2018**, *18*, 1052. [CrossRef]
12. Schultealbert, C.; Baur, T.; Schütze, A.; Sauerwald, T. Facile Quantification and Identification Techniques for Reducing Gases over a Wide Concentration Range Using a MOS Sensor in Temperature-Cycled Operation. *Sensors* **2018**, *18*, 744. [CrossRef]
13. Korotcenkov, G.; Cho, B.K. Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey). *Sens. Actuators B* **2011**, *156*, 527–538. [CrossRef]
14. ScioSense ENS160 Datasheet. Available online: https://www.sciosense.com/wp-content/uploads/documents/SC-001224-DS-1-ENS160-Datasheet-Rev-0.95.pdf (accessed on 4 April 2021).
15. Schütze, A.; Baur, T.; Leidinger, M.; Reimringer, W.; Jung, R.; Conrad, T.; Sauerwald, T. Highly Sensitive and Selective VOC Sensor Systems Based on Semiconductor Gas Sensors: How to? *Environments* **2017**, *4*, 20. [CrossRef]
16. Kohl, D.; Kelleter, J.; Petig, H. Detection of Fires by Gas Sensors. *Sens. Updat.* **2001**, *9*, 161–223. [CrossRef]

17. Leidinger, M.; Sauerwald, T.; Reimringer, W.; Ventura, G.; Schütze, A. Selective detection of hazardous VOCs for indoor air quality applications using a virtual gas sensor array. *J. Sens. Sens. Syst.* **2014**, *3*, 253–263. [CrossRef]

18. Bastuck, M.; Baur, T.; Richter, M.; Mull, B.; Schütze, A.; Sauerwald, T. Comparison of ppb-level gas measurements with a metal-oxide semiconductor gas sensor in two independent laboratories. *Sens. Actuators B* **2018**, *273*, 1037–1046. [CrossRef]

19. Baur, T.; Schütze, A.; Sauerwald, T. Optimierung des temperaturzyklischen Betriebs von Halbleitergassensoren. *Tech. Mess.* **2015**, *82*, 187–195. [CrossRef]

20. Schultealbert, C.; Baur, T.; Schütze, A.; Böttcher, S.; Sauerwald, T. A novel approach towards calibrated measurement of trace gases using metal oxide semiconductor sensors. *Sens. Actuators B* **2017**, *239*, 390–396. [CrossRef]

21. Leidinger, M.; Schultealbert, C.; Neu, J.; Schütze, A.; Sauerwald, T. Characterization and calibration of gas sensor systems at ppb level–a versatile test gas generation system. *Meas. Sci. Technol.* **2018**, *29*, 015901. [CrossRef]

22. Baur, T.; Bastuck, M.; Schultealbert, C.; Sauerwald, T.; Schütze, A. Random gas mixtures for efficient gas sensor calibration. *J. Sens. Sens. Syst.* **2020**, *9*, 411–424. [CrossRef]

23. Hofmann, H.; Plieninger, P. Bereitstellung einer Datenbank zum Vorkommen von flüchtigen organischen Verbindungen in der Raumluft. *WaBoLu Hefte* **2008**, *5*, 161.

24. Hofmann, H.; Erdmann, G.; Müller, A. Zielkonflikt energieeffiziente Bauweise und gute Raumluftqualität–Datenerhebung für flüchtige organische Verbindungen in der Innenraumluft von Wohn- und Bürogebäuden (Lösungswege); 2014. Available online: https://www.agoef.de/forschung/fue-ll-voc-datenerhebung/abschlussbericht.html (accessed on 17 May 2021).

25. Traynor, G.W.; Apte, M.G.; Carruthers, A.R.; Dillworth, J.F.; Grimsrud, D.T.; Gundel, L.A. Indoor air pollution due to emissions from wood-burning stoves. *Environ. Sci. Technol.* **1987**, *21*, 691–697. [CrossRef] [PubMed]

26. Schultealbert, C.; Amann, J.; Baur, T.; Schütze, A. Measuring Hydrogen in Indoor Air with a Selective Metal Oxide Semiconductor Sensor. *Atmosphere* **2021**, *12*, 366. [CrossRef]

27. WHO. *WHO Regional Office for Europe WHO Guidelines for Indoor Air Quality: Selected Pollutants*; World Health Organization. Regional Office for Europe: Copenhagen, Denmark, 2010; Volume 9, ISBN 978-92-890-0213-4.

28. Schultealbert, C.; Baur, T.; Schütze, A.; Sauerwald, T. Investigating the role of hydrogen in the calibration of MOS gas sensors for indoor air quality monitoring. In Proceedings of the Indoor Air, Philadelphia, PA, USA, 22–27 July 2018.

29. Loh, W.-L. On Latin Hypercube Sampling. *Ann. Stat.* **1996**, *24*, 2058–2080. [CrossRef]

30. Floor Plan of Building A5 1 at Saarland University. Available online: https://www.uni-saarland.de/fileadmin/upload/footer/grundriss/SBC-13_00-002.pdf (accessed on 4 April 2021).

31. Bastuck, M.; Baur, T.; Schütze, A. DAV^{3}E–a MATLAB toolbox for multivariate sensor data evaluation. *J. Sens. Sens. Syst.* **2018**, *7*, 489–506. [CrossRef]

32. Youssef, S.; Zimmer, C.; Szielasko, K.; Schütze, A. Automatic feature extraction of periodic time signals using 3MA-X8 method. *Tm Tech. Mess.* **2019**, *86*, 267–277. [CrossRef]

33. Bur, C.; Engel, M.; Horras, S.; Schütze, A. Drift compensation of virtual multisensor systems based on extended calibration, IMCS 2014-the 15th International Meeting on Chemical Sensors. In Proceedings of the IMCS 2014—The 15th International Meeting on Chemical Sensors, Buenos Aires, Argentina, 16–19 March 2014.

34. Robin, Y.; Goodarzi, P.; Baur, T.; Schultealbert, C.; Schütze, A.; Schneider, T. Machine Learning based calibration time reduction for Gas Sensors in Temperature Cycled Operation. In Proceedings of the IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Glasgow, Scotland, 17–20 May 2021.