Advanced Calibration Strategies for Indoor Air Quality Sensors

Transfer Learning for Addressing Scalability Limitations

Dr. Christian Bur, Saarland University, Germany

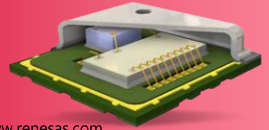
ISOCS Winter School 2023, Bormio, Italy

Lab for Measurement Technology
UNIVERSITÄT DES SAARLANDES

Agenda

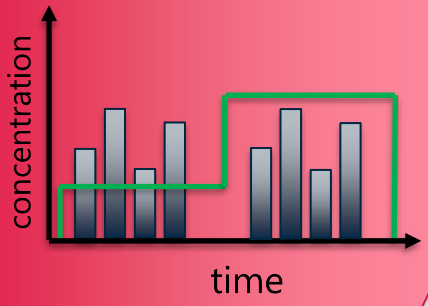
Application

fischerheating.com



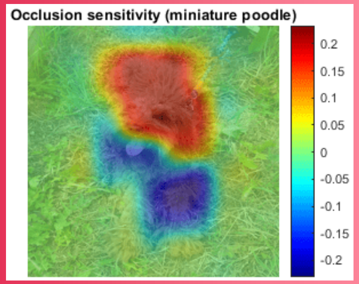
www.renesas.com

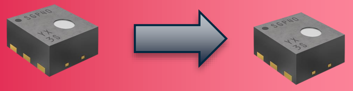
Calibration strategies



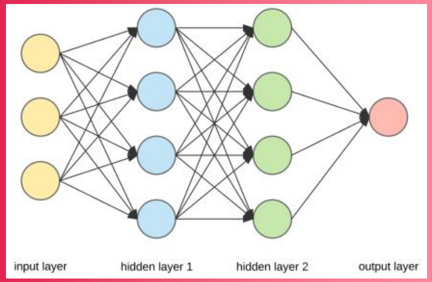
Calibration equipment

Transfer learning Explainable AI





Machine and Deep Learning



Defining the Measurement Task
Identify Target and Interfering Gases

Understanding the Measurement Task

Application specific questions?

- Target substance?
  - Known/unknown?
- Background
  - Static/variable?
  - How many substances?
- Concentration range known?
  - Of target and background
- Online monitoring or random sampling?
- Classification or quantification?
- Costs?

Requirements

- Selectivity
- Sensitivity
- LoD / LoQ
- Stability
- Time resolution
- Sale

Sensor based systems

\+ High sensitivity
\+ Low cost
\+ Allow for continuous monitoring

– Poor selectivity
– Require individual calibration
– Stability is an issue
→ Recalibration, replacement

Analytics

\+ Highest sensitivity
\+ Highly selective
\+ Allow for identification
\+ Good as reference

– Premium-priced
– Often time delayed measurements
– Requires expert knowledge

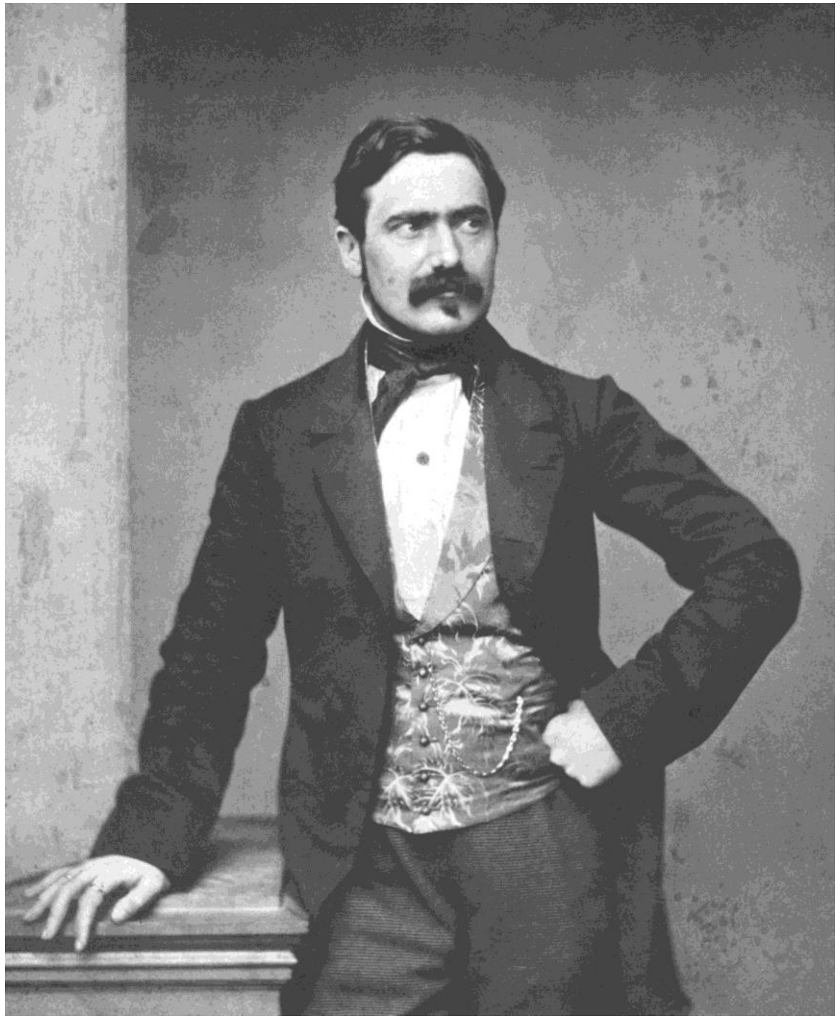
Example: Indoor Air Quality

Max von Pettenkofer (1818 – 1901) wrote 1858:

Above all, the organic compounds, "which betray themselves by the smell when they accumulate", can serve as a benchmark for the air quality, "but unfortunately we have no method of determining it quantitatively". "So, we have no other clue than the carbon dioxide."

CO_2 as indicator or proxy for VOCs emitted by people
(VOC: volatile organic compounds)

Shouldn´t we rather measure VOCs directly?
Can also indicate pollutants from furniture, building materials, cooking, air fresheners, etc.



Max von Pettenkofer, https://en.wikipedia.org/wiki/Max_Joseph_von_Pettenkofer

Monitoring Indoor Air Quality

Air quality is one target in the sustainable development goals set by the United Nations for 2030

- ➢ Volatile organic compounds are on of the main pollutants of concern indoors

Demand for sensors and measuring devices for continuous monitoring indoors

- ➢ Low-cost gas sensors for quantification of the total concentration of volatile organic compounds (TVOC)

https://www.andatechdistribution.com.au/blogs/resources/indoor-air-quality-infographic

Simulating the Complex Indoor Environment

Lab for Measurement Technology

UNIVERSITÄT DES SAARLANDES

Indoor environment is a complex mixture

- Around 200-300 different VOCs were reported
- CO_2, CO, H_2, NO_x plus humidity
- Influence of outdoor air (O_3, NO_x,..)

Dividing VOCs in substance groups

- Suitable for sensors, since the measuring principle is based on reactions with molecules
- Measure also VVOCs in contrast to analytics (TVOC)
- Define a sensor based total VOCs value ($TVOC_{Sens}$)

Testing sensor systems with complex mixtures of VOCs

- Representative environmental mixtures of the individual components

VOC Representatives According to AGÖF, UBA

Lab for Measurement Technology
UNIVERSITÄT DES SAARLANDES

Substance group	Representative 1	Representative 2	Representative 3	Source
Aldehyde (1-150 ppb)	**Formaldehyde**	Acetaldehyde	Hexanal	Food, fragrance additives
Alkane (1-100 ppb)	**n-Hexane**	Undecane	Cyclohexane	From outside: exhaust & fuel
Alcohol (1-200 ppb)	**Ethanol**	2-Propanol	1-Propanol	Cleaner, disinfectant
Aromatics (1-100 ppb)	**Toluene**	Xylene	-	Solvents
Ester (1-75 ppb)	**Ethyl acetate**	n-Butyl acetate	-	Solvents, cooking
Ketone (1-150 ppb)	**Acetone**	Butanone	-	Solvents, human metabolism
Organic Acid (1-100 ppb)	**Acetic acid**	Propiolic acid	Caproic acid	Cleaning products
Terpene (1-100 ppb)	**Limonene**	Alpha-pinene	3-Carene	Fragrance additives

Alkene, Halocarbons, Glycols & Glycol ethers: negligible

AGÖF (2007): Supply of a data base about the occurrence of volatile organic compounds in indoor air

AGÖF (2014): Conflict of Goals between Energy-efficient Buildings and Good Indoor Air Quality - Data Collection of Volatile Organic Compounds in Indoor Air of Residential and Office Buildings

UBA (2010): German Environmental Survey on Children (GerES IV)

AGÖF: Association of Ecological Research Institutes, Germany

UBA: German Environment Agency

Interfering/Background Gases

- Hydrogen: main source are human beings
- CO: mainly due to combustion processes
- Humidity
- Nitrogen Oxides: no valid data available for indoor environments
  - NO_2: annual mean (2020, close to traffic, UBA): 15 ppb
    Richtwert I: 42 ppb; Richtwert II: 140 ppb
  - N_2O: atmospheric concentration ca. 330 ppb
  - Ozone : O_3?
  - Further research needed
- Cyclic siloxanes (D3 – D6): very low concentrations, but relevant as sensor poison

⇒ Separate investigation

Interference	Min conc.	Max. conc.
CO	100 ppb	2000 ppb
Hydrogen	400 ppb	2000 ppb
humidity	20 %	75 %

Metal Oxide Semiconductor (MOS) Sensors

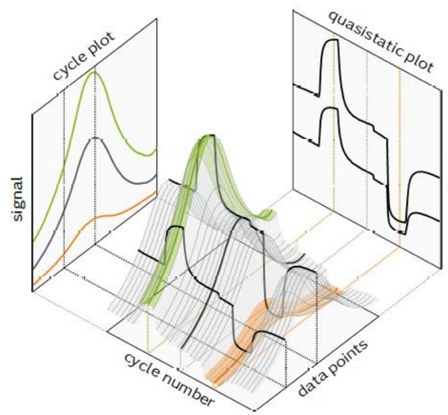
- Gases from the ambient react on the surface
- Reactions are temperature- and material-dependent
  - Lead to a change in resistance
- Resistance depend on oxygen coverage of the surface

SGP40: Sensirion AG, Stäfa, Switzerland

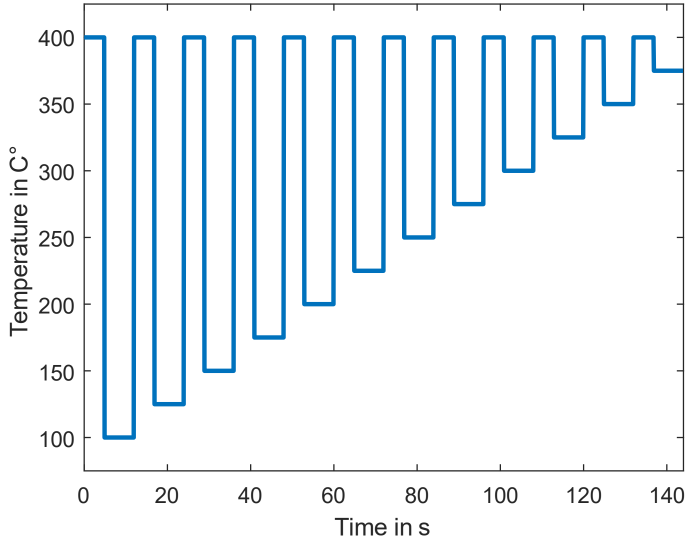
ENS160: Sciosense B.V., Eindhoven, Netherlands

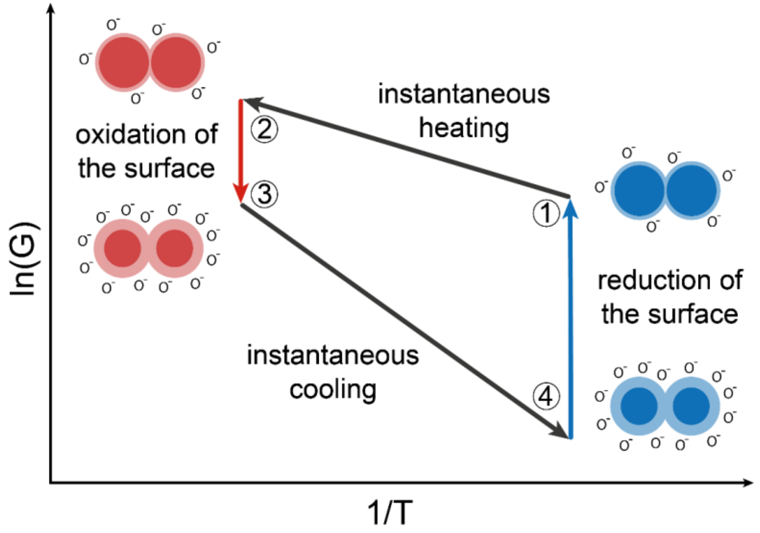
MOS sensors in temperature cycled operation (TCO)

Virtual multi-sensor



Manuel Bastuck, Dissertation, Saarland University and Linköping University, 2019
Shaker Verlag, 2019,ISBN: 978-3-8440-7075-0
http://liu.diva-portal.org/smash/record.jsf?pid=diva2%3A1338901&dswid=-4621



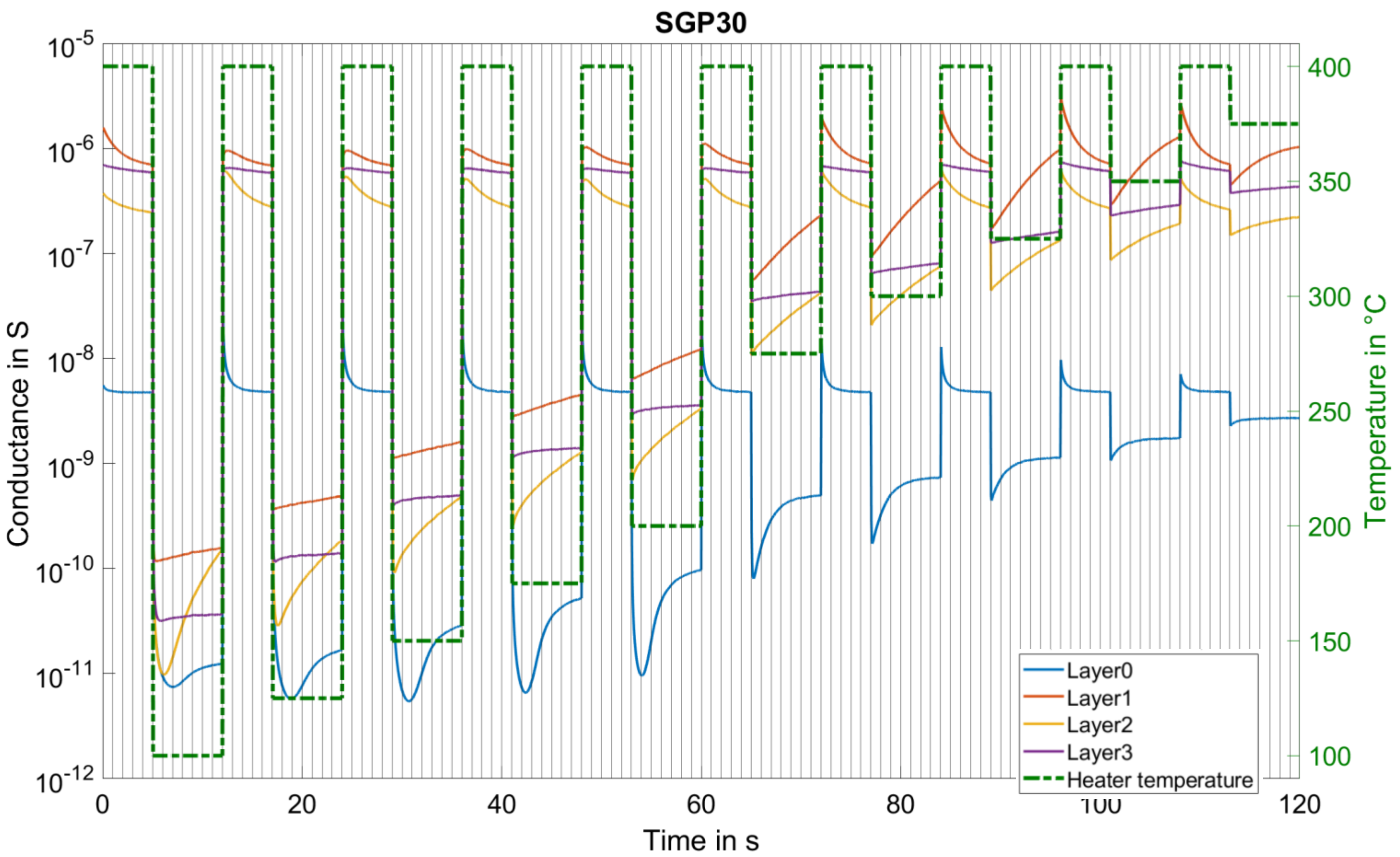


A. Schütze et al.: Environments 2017, *4*, 20; doi: 10.3390/environments4010020

Data Evaluation FESR

Data evaluation

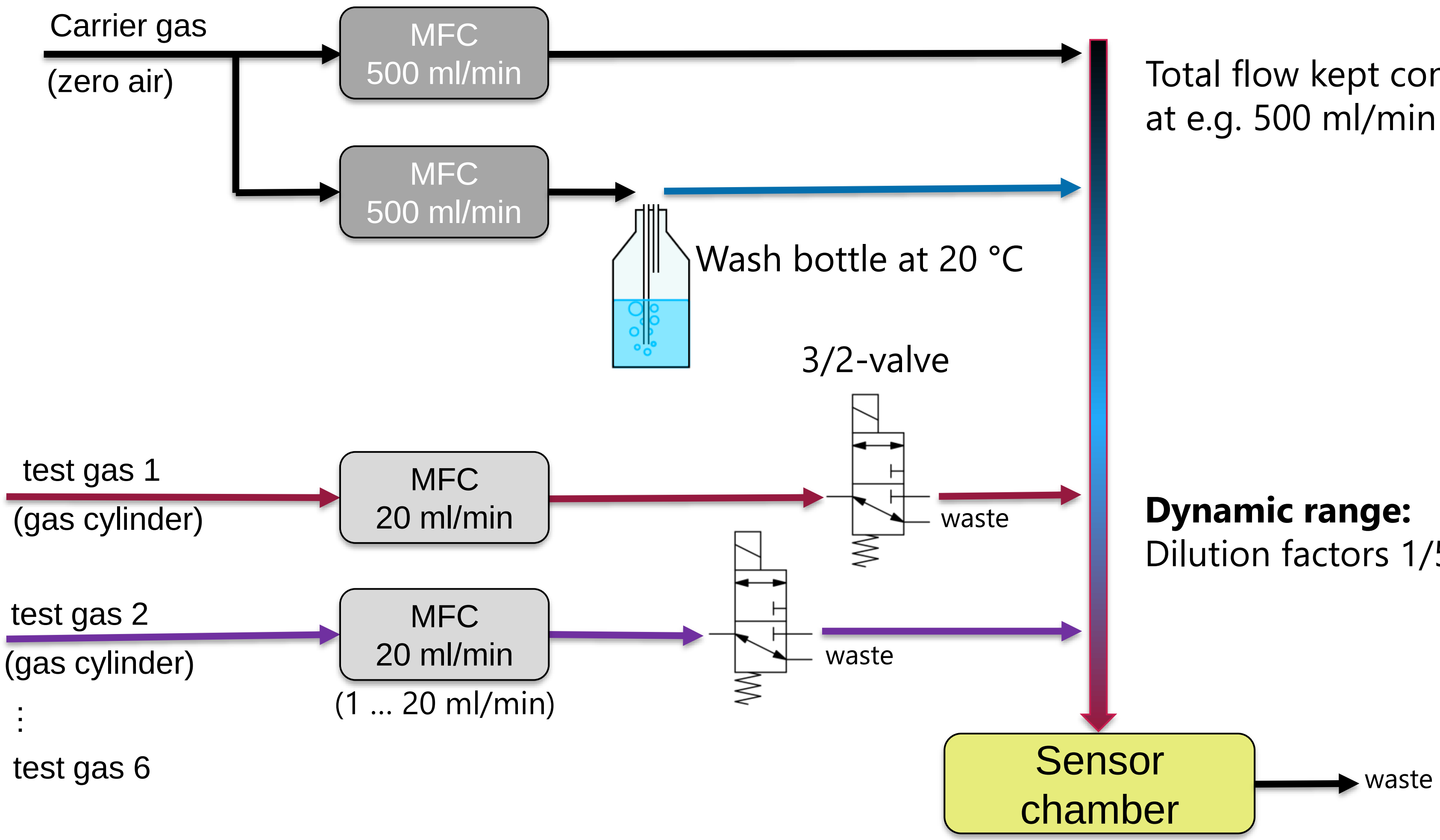
- Machine learning methods: FESR
  - **F**eature **E**xtraction
  - Feature **S**election, and
  - **R**egression
- Divide temperature cycle in 1 s ranges for feature extraction
- Partial Least Squares Regression (PLSR)



Raw Data → Feature Extraction Mean & Slope → Feature Selection RFE-LSR → Quantification PLS regression

Gas Mixing Apparatus
for complex lab calibration

Standard Gas Mixing Apparatus



Total flow kept constant at e.g. 500 ml/min

Dynamic range:
Dilution factors 1/500 ... 20/500

But…

For trace-level detection in complex atmospheres, this setup is not feasible

- Dynamic range is limited
- Number of test gases is too small
- Trace level (ppb range) generation is questionable

Example:

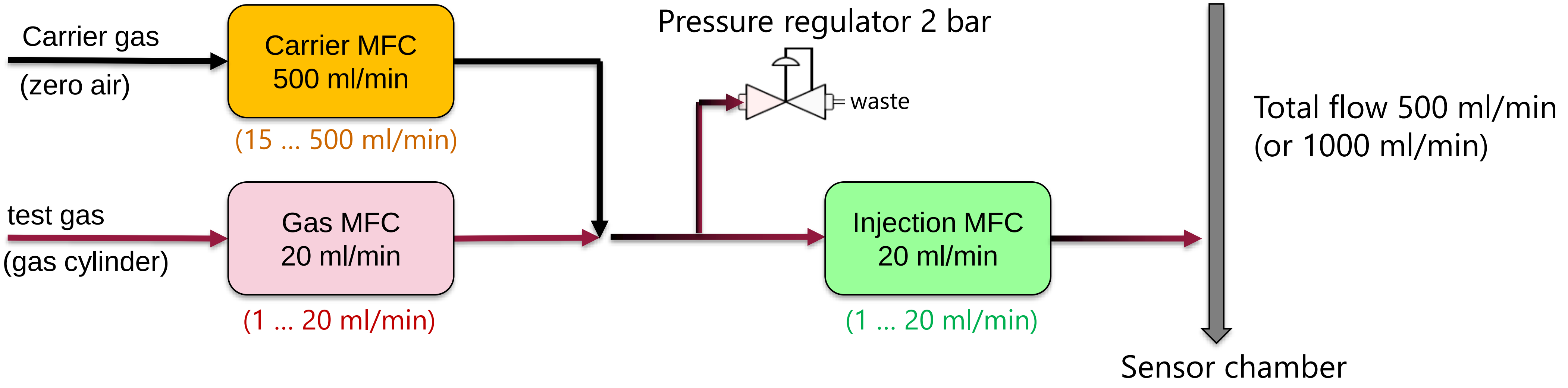
Gas cylinder with purity of 5.0

=> means a purity of 99.999 %

Up to 10 ppm contaminations like CO, H_2, and other VOCs in the cylinder

With simple dilution between 20 ppb and 400 ppb unknown substances in test gas mixture

Pre-Dilution Line (2-stage-dilution)



Dynamic range:

Max. dilution: 1/500 * 1/500 = 1/250.000 = 4 x 10^**-6**

Min. dilution: 20/500 = 1/25 = 4 x 10^**-2**

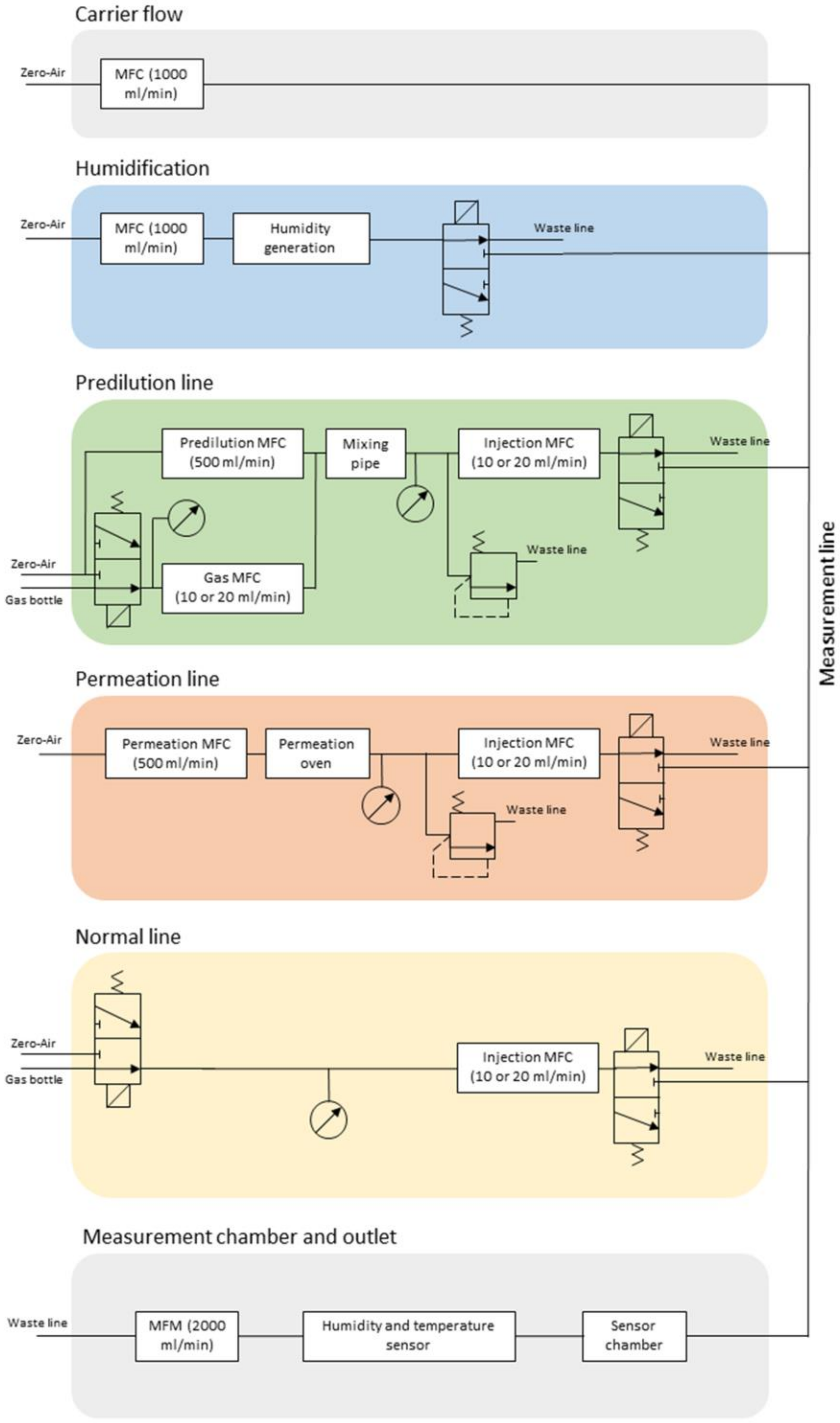
Dynamic range over 4 magnitudes
With the same test gas cylinder

Dilution of test gas AND contaminations in the test gas bottle

=> Choose a higher bottle concentration (typ. 200 ppm) and a high dilution factor

Latest GMA at U Saarland

- Up to 18 individual test gas lines
  - Currently 14 pre-dilution lines
  - Possibility to add permeation lines (permeation tubes)
- Humidification
  - Via temperature controlled wash bottles
- Fully automated
  - Set concentrations individually from sub-ppb up to several ppm
- Monitoring of pressure and flow level in each line
- Validation by analytics regularly
- Round Robin Tests (inter-lab testing)



Dennis Arendes et al., 15. Dresdner Sensor-Symposium, 2021, https://www.ama-science.org/proceedings/details/4133
Dennis Arendes et al., 16. Dresdner Sensor-Symposium, 2022, https://www.ama-science.org/proceedings/details/4282

Photos of Norm-GMA

Calibration Strategies
Sequential vs. randomized

Lab Calibration

- Define target gas(es) and concentration range(s)
- Define background mixture (interfering gases incl. conc. range)
  - Major components and relevant for your sensor technology
  - Don’t forget humidity

Lab calibration is just a simplification of the real world

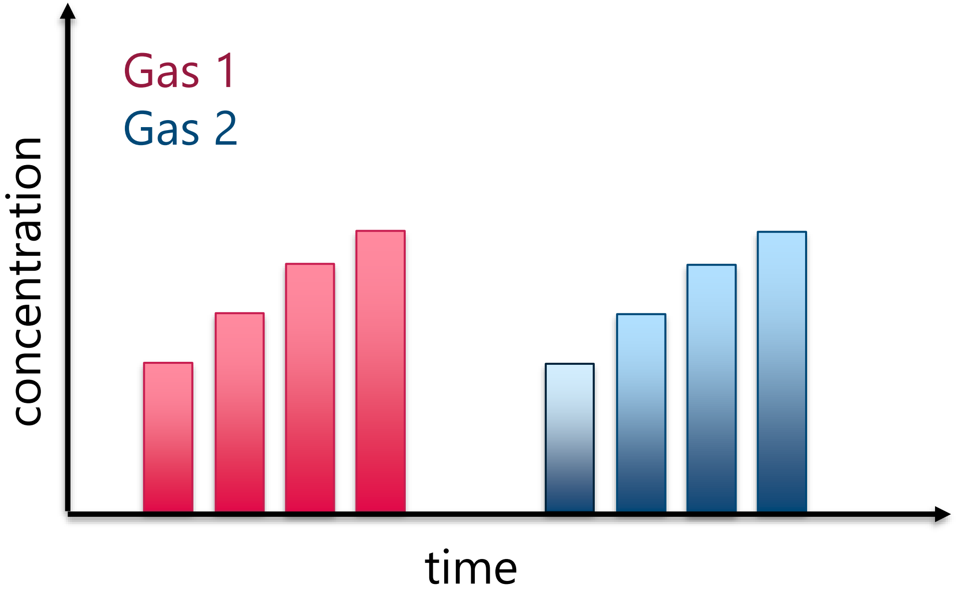
And always a trade-off between complexity/time and accuracy

Sequential Procedures

Iteratively vary concentration of two gases

Problems

- No interaction between the gases
- Quantization errors
- Systematic profile can lead to overfitting of the ML model
- Drift and memory effects are hard to detect



Sequential Procedures

Lab for Measurement Technology | UNIVERSITÄT DES SAARLANDES

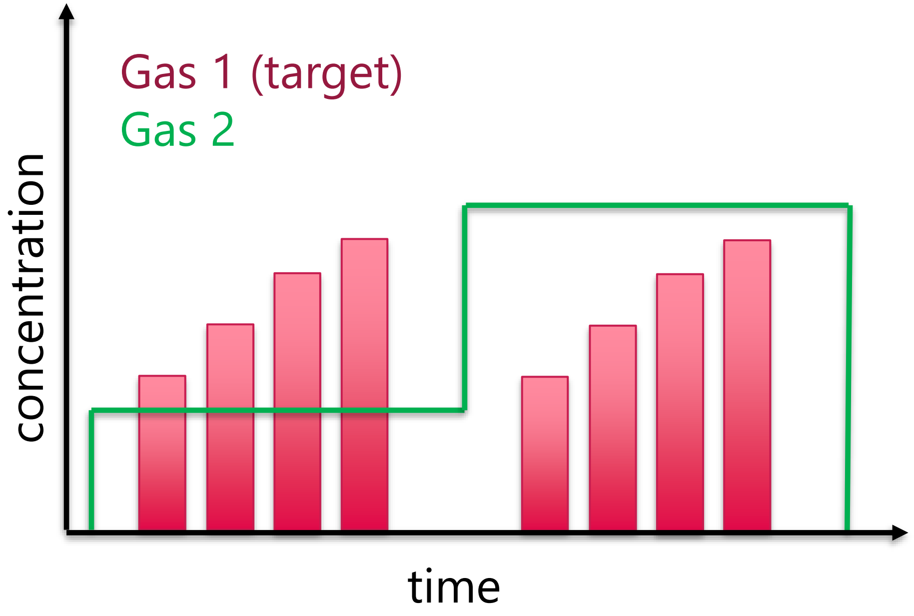
Iteratively vary the concentration of background substances and target gases

Problems

- Leads to long calibration times

  Example: 10 conc. of target gas, 5 background gases (4 conc. each)
  $\Rightarrow$ 10.240 gas exposures
  Assume 10 min per exposure with 10 min pause in between
  $\Rightarrow$ **142 days** of calibration!

- Systematic profile can lead to overfitting of the ML model
- Drift and memory effects are hard to detect

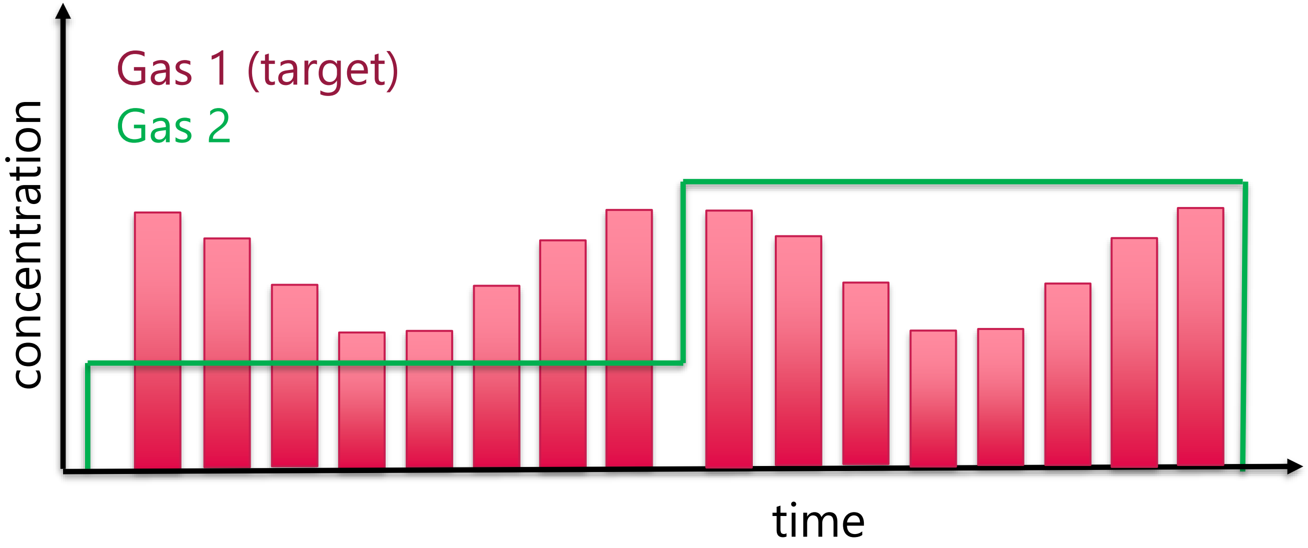


Sequential Procedures

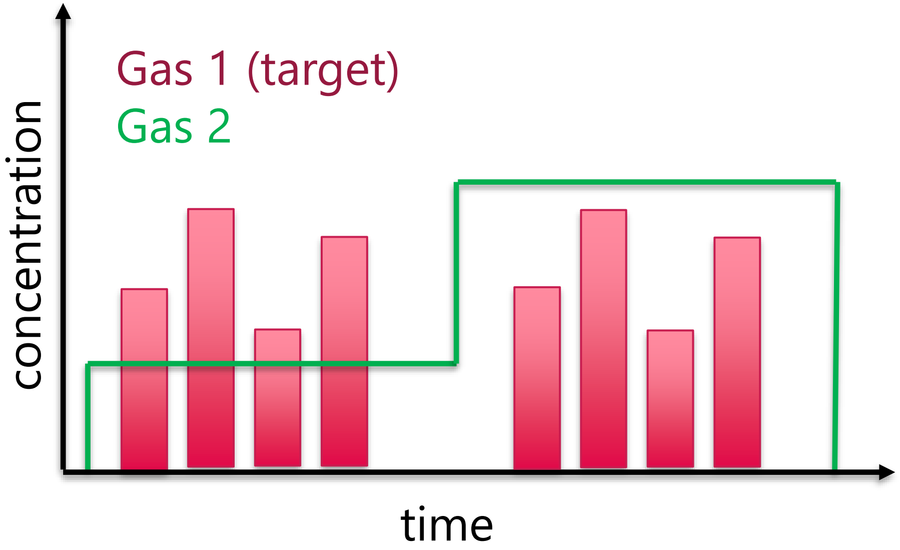
Also common: concentration ramp down and up

Compare corresponding concentrations

But: calibration time doubles



Better: pseudo-random order

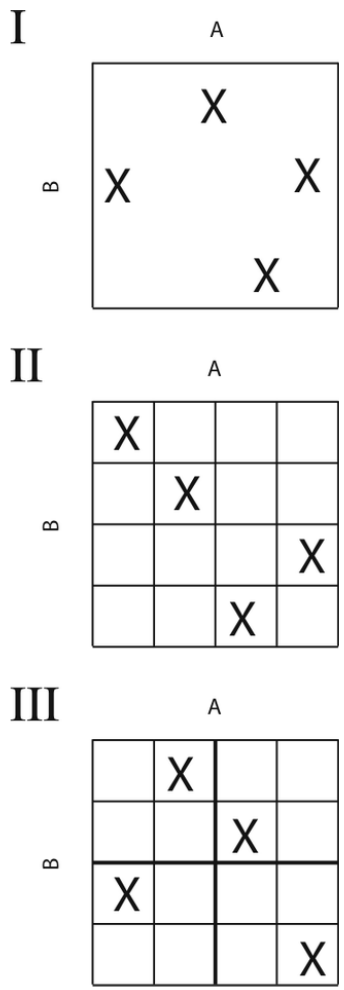


Randomized Calibration

Only define concentration ranges and the number of individual exposures

Let an algorithm choose specific concentrations

- Random sampling
- Random effects
  - Random variations of variables drawn from defined distributions
- Latin Hypercube sampling
  - Each sampling space dimension is roughly evenly sample
- Orthogonal sampling
  - Optimized for minimum correlation between dimensions (gases)



Wikipedia.org

Comparison Sequential vs. Random Calibration

Sequential setup:

- 3 levels of RH
- 6 gases
- 4 conc. per test gas
- CO and H2 at atmospheric conc.

⇒ 72 gas exposures

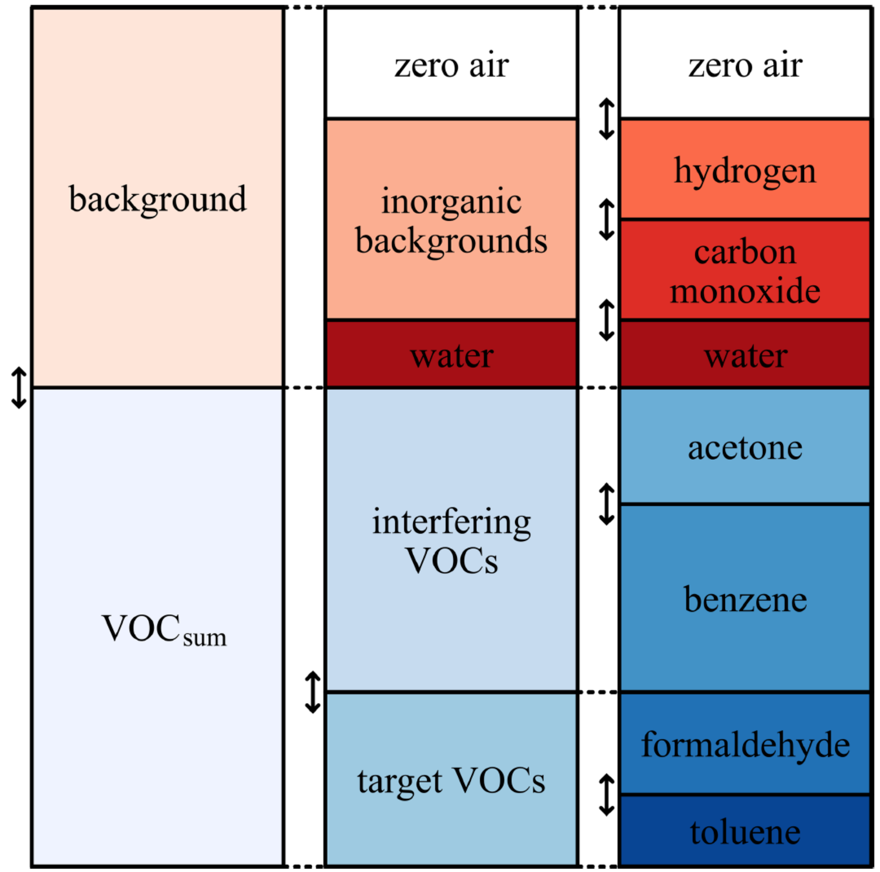
Gas	Concentration (ppb)			
acetone	250	500	750	1000
benzene	250	500	750	1000
carbon monoxide	150	300	450	600
formaldehyde	40	80	120	160
hydrogen	500	750	1000	1250
toluene	5	25	45	65

Randomized setup:

- Define concentration ranges
- For comparison:

  Limit concentration range to fit to the sequential measurement

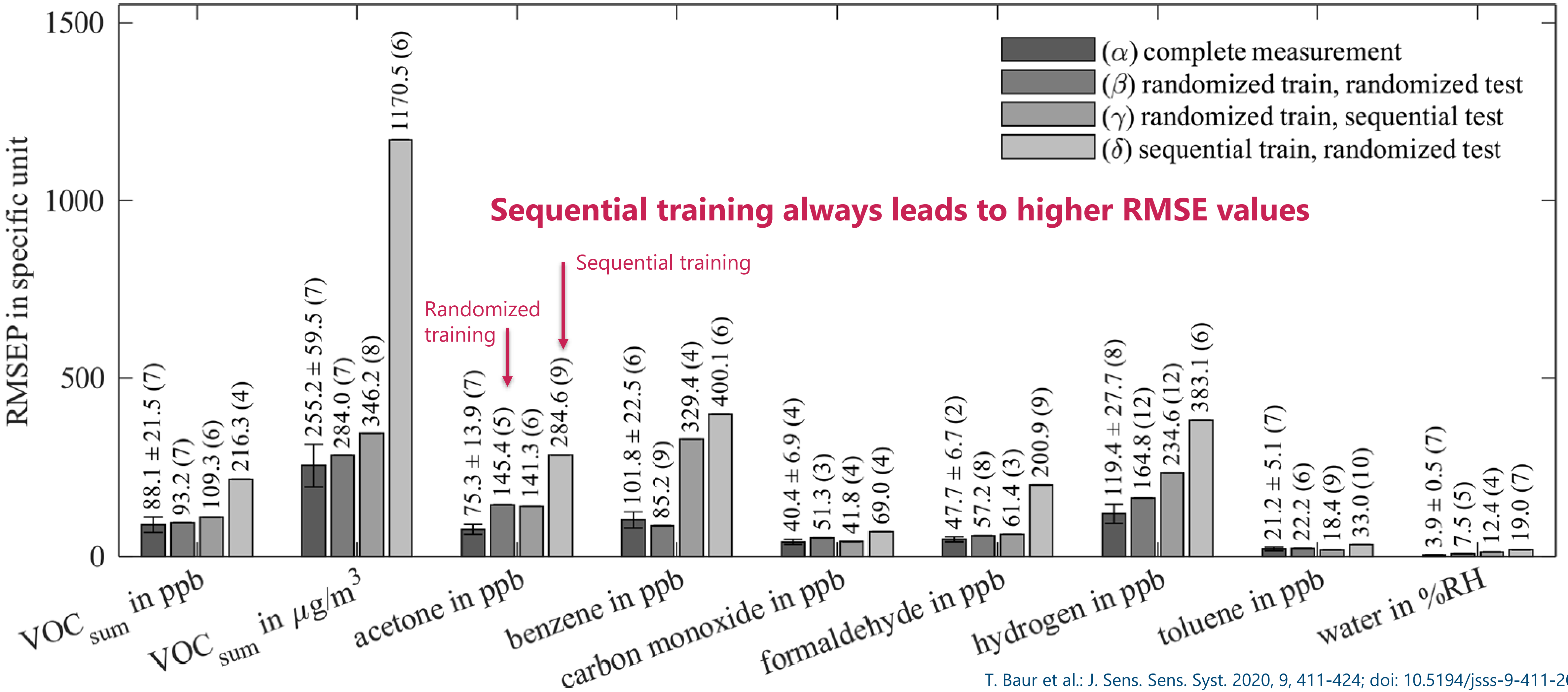
	Concentration range
hydrogen	301–2499 ppb
carbon monoxide	101–1995 ppb
humidity	25–75 %RH
VOC_{sum} in µg/m^3	21–4902 µg/m^3
VOC_{sum} in ppb	6–2312 ppb
acetone	0–1846 ppb
benzene	0–1180 ppb
formaldehyde	0–723 ppb
toluene	0–245 ppb



T. Baur et al.: J. Sens. Sens. Syst. 2020, 9, 411-424; doi: 10.5194/jsss-9-411-2020

Comparison Sequential vs. Random Calibration

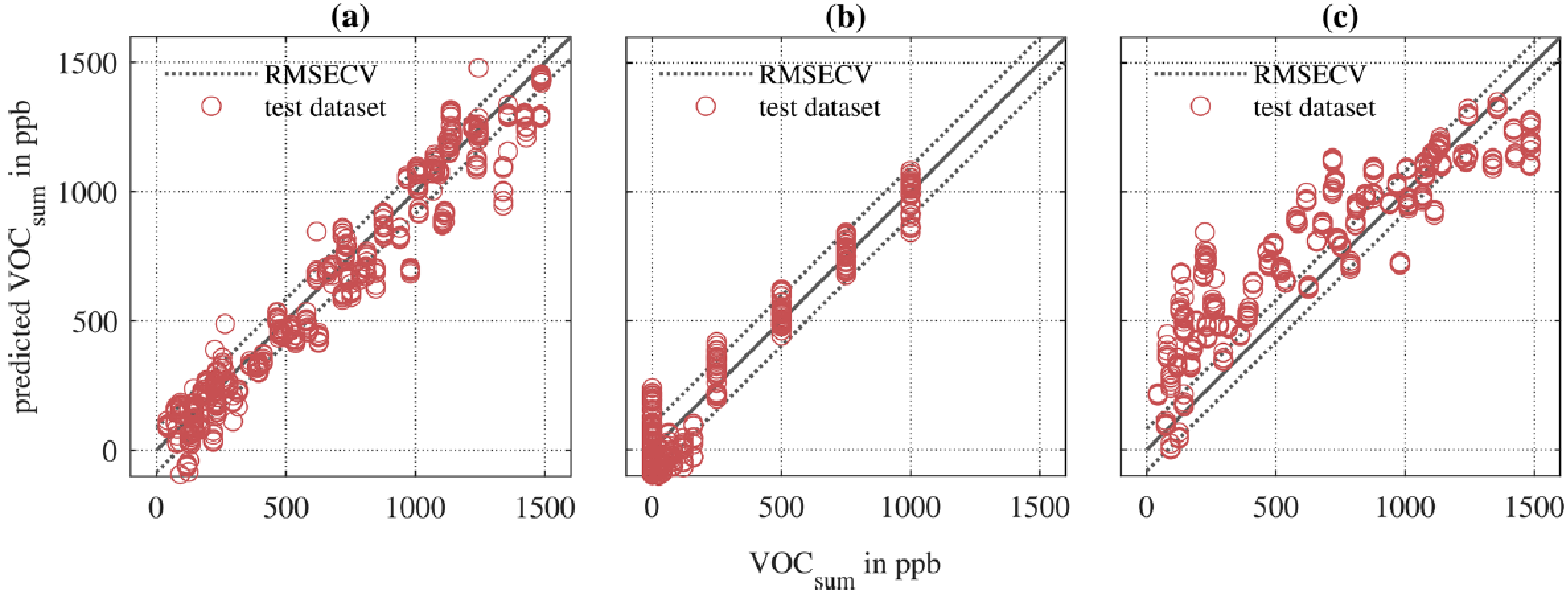
Lab for Measurement Technology
UNIVERSITÄT DES SAARLANDES



T. Baur et al.: J. Sens. Sens. Syst. 2020, 9, 411-424; doi: 10.5194/jsss-9-411-2020

Comparison Sequential vs. Random Calibration

Train:	random	random	sequential
Test:	random	sequential	random



T. Baur et al.: J. Sens. Sens. Syst. 2020, 9, 411-424; doi: 10.5194/jsss-9-411-2020

Sequential vs. Randomized Calibration Strategy

Sequential

+ Check general sensitivity
+ Reveals time constants
+ Easy to interpret
+ Intuitive

– Huge simplification
– Quantization errors
– Systematic approach overfitting, memory effects, …
– Only for a small number of gases

Randomized

+ No correlation between exposures
+ Statistically valid
+ Good for high amounts of substances
+ Automatic approach
+ Ideal to build machine learning models

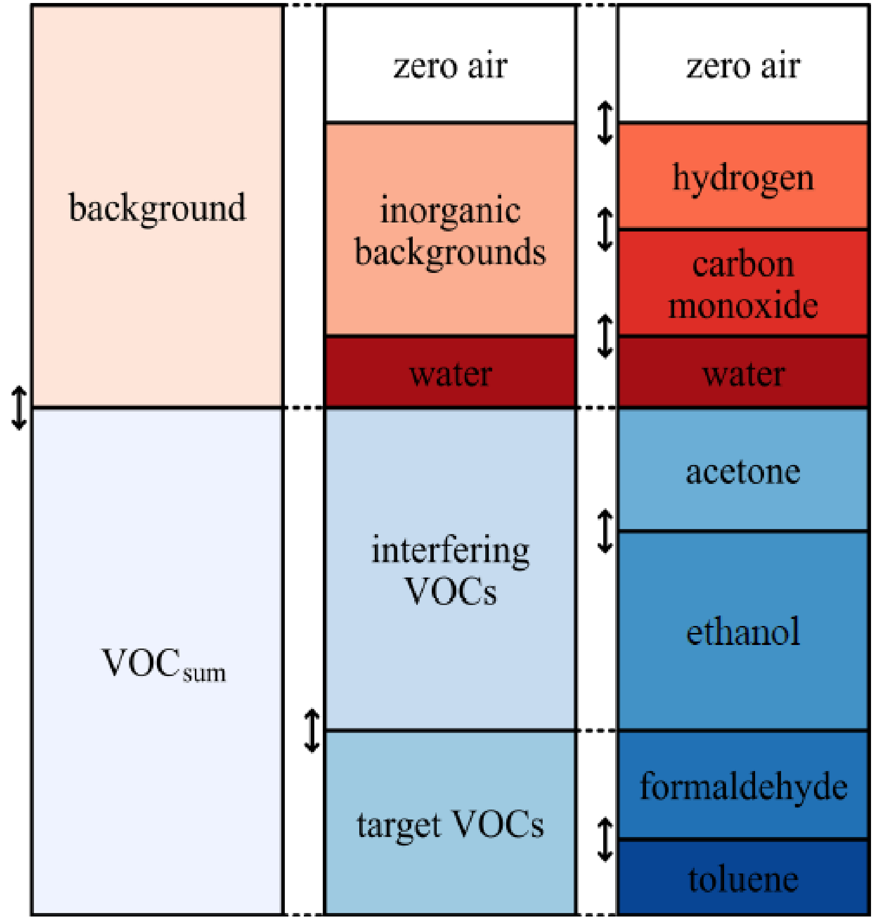
– Cannot be interpreted manually
– Required complex gas mixing systems
– Requires well-annotated data

**A representative and comprehensive data base
is the key for
building robust machine learning models**

Machine Learning Approaches
FESR vs. CNN

Dataset

- Randomized profile with **500 unique gas mixtures (UGMs)**
- Sensor: SGP30, Sensirion: **4** gas-sensitive layers
- Temperature cycled operation
  - Cycle length: 144 seconds at 10 Hz
  - ⇒ 1440 sample points per T-cycle
- Feature Extraction
  - Raw signal is divided in equidistant segments (1 second)
    - Extract slope and mean as features
    - **288** features per T-cycle and sensor
- Gas exposure time: 20 min
  - ⇒ about **10 T-cycles (sample points) per UGM**



T. Baur et al.: J. Sens. Sens. Syst. 2020, 9, 411-424; doi: 10.5194/jsss-9-411-2020

Feature Table

Lab for Measurement Technology | UNIVERSITÄT DES SAARLANDES

4 x 2x144 features

Concentrations of gases

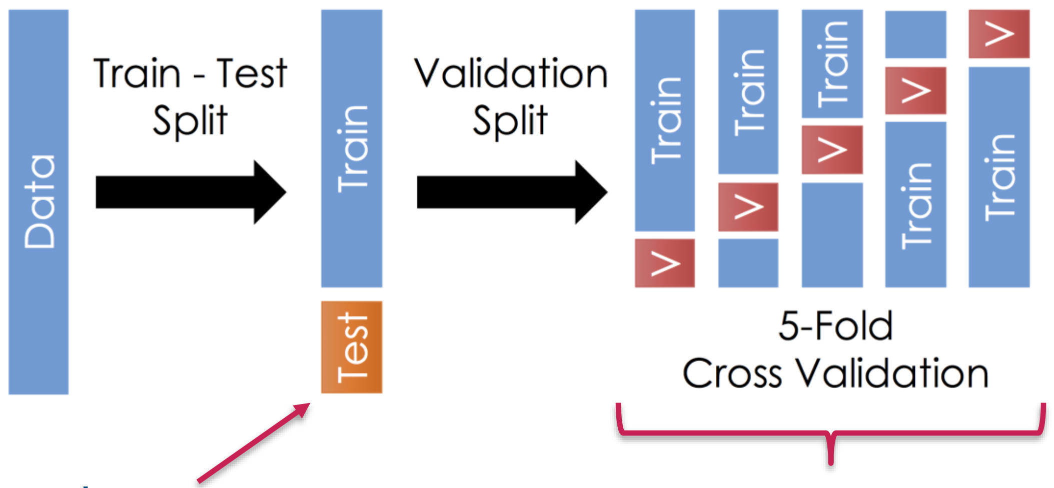
900 UGM with 10 T-cycles each

		sensor 0					sensor 1					sensor 2					sensor 3					annotation gas 1	annotation gas 2	...
		F1	F2	F3	...	F288	F1	F2	F3	...	F288	F1	F2	F3	...	F288	F1	F2	F3	...	F288			
UGM 1	sample 1																					20 ppb	75 ppb	
	sample 2																					20 ppb	75 ppb	
	sample 3																					20 ppb	75 ppb	
	…																					20 ppb	75 ppb	
	sample 10																					20 ppb	75 ppb	
UGM 2	sample 1																					50 ppb	110 ppb	
	sample 2																					50 ppb	110 ppb	
	sample 3																					50 ppb	110 ppb	
	…																					50 ppb	110 ppb	
	sample 10																					50 ppb	110 ppb	
UGM 3	sample 1																					30 ppb	150 ppb	
	sample 2																					30 ppb	150 ppb	
	sample 3																					30 ppb	150 ppb	
	…																					30 ppb	150 ppb	
	sample 10																					30 ppb	150 ppb	
⋮																								

UGM: unique gas mixture

Modell Building – Prepare Dataset

- Split dataset into training and testing (80:20) using hold out
- Split training into training and validation (LOOCV, k-fold)
- Take entire gas exposures out (not observations/T-cycles!)



ds100.org

Unknown data

Check for generalizability

Used for building the model

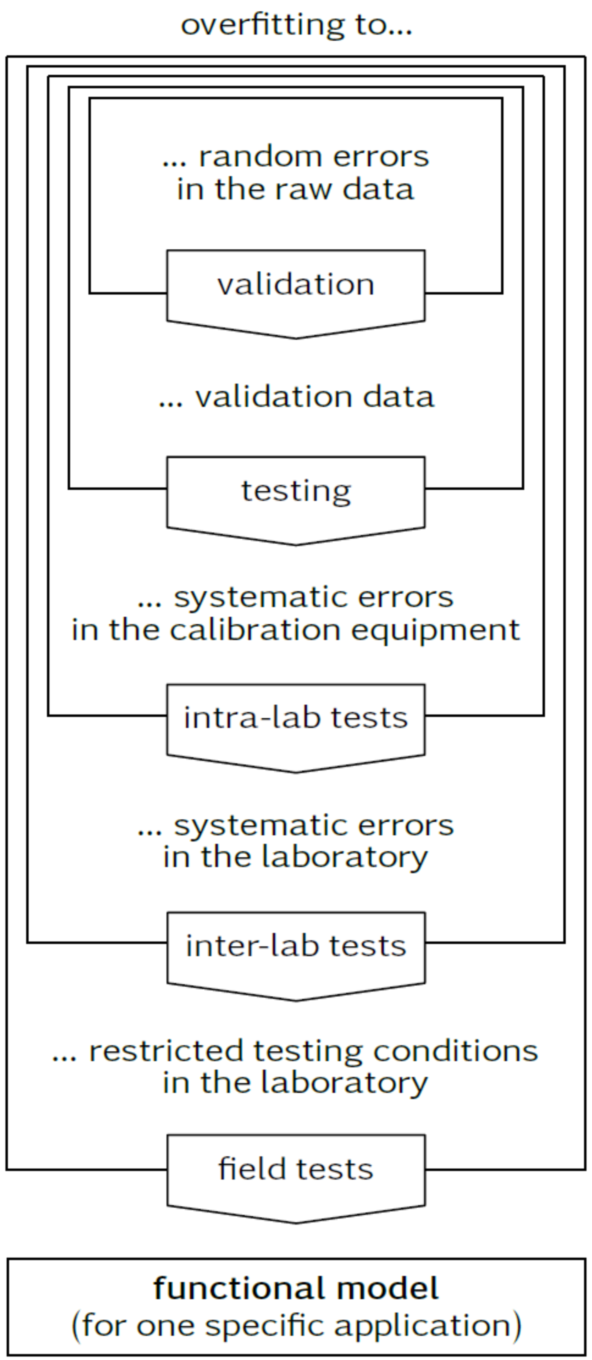
Hyperparameter tuning, avoid overfitting

Common Causes for Overfitting

Overfitting due to...

- Random errors in the raw data ⇒ Validation
- Insufficient validation data ⇒ Testing
- Systematic errors in the calibration equipment
- Systematic errors in the lab ⇒ Intra-lab tests
- Restricted testing conditions ⇒ Field tests

non-ideal design of experiment



Manuel Bastuck, Dissertation, Saarland University and Linköping University, 2019
Shaker Verlag, 2019,ISBN: 978-3-8440-7075-0
http://liu.diva-portal.org/smash/record.jsf?pid=diva2%3A1338901&dswid=-4621

Standard Machine Learning Approach

Lab for Measurement Technology | UNIVERSITÄT DES SAARLANDES

sensor → Temperature modulation
(virtual) sensor array

→ raw data → Pre-processing

→ Feature extraction → Standardization → Feature pre-selection

→ Feature selection ↔ Dimension Reduction ↔ Model building
Cross-validation

↕ testing

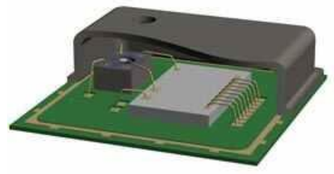
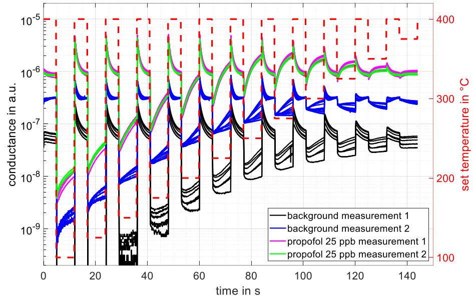
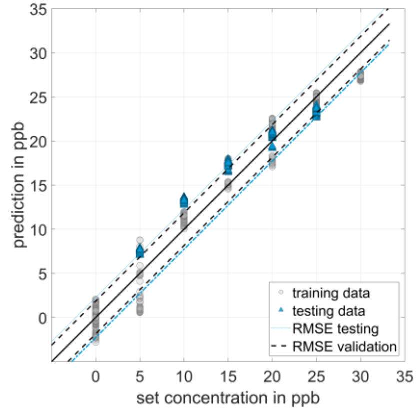
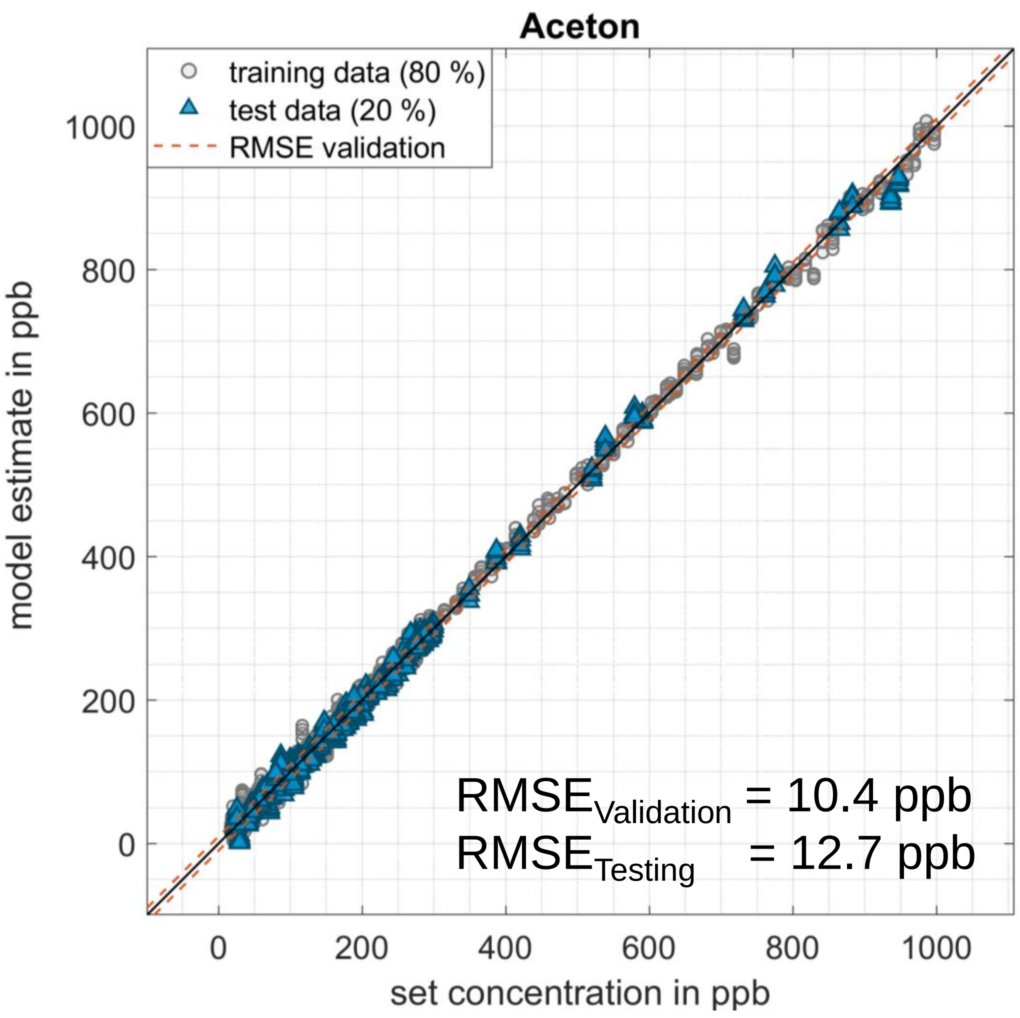


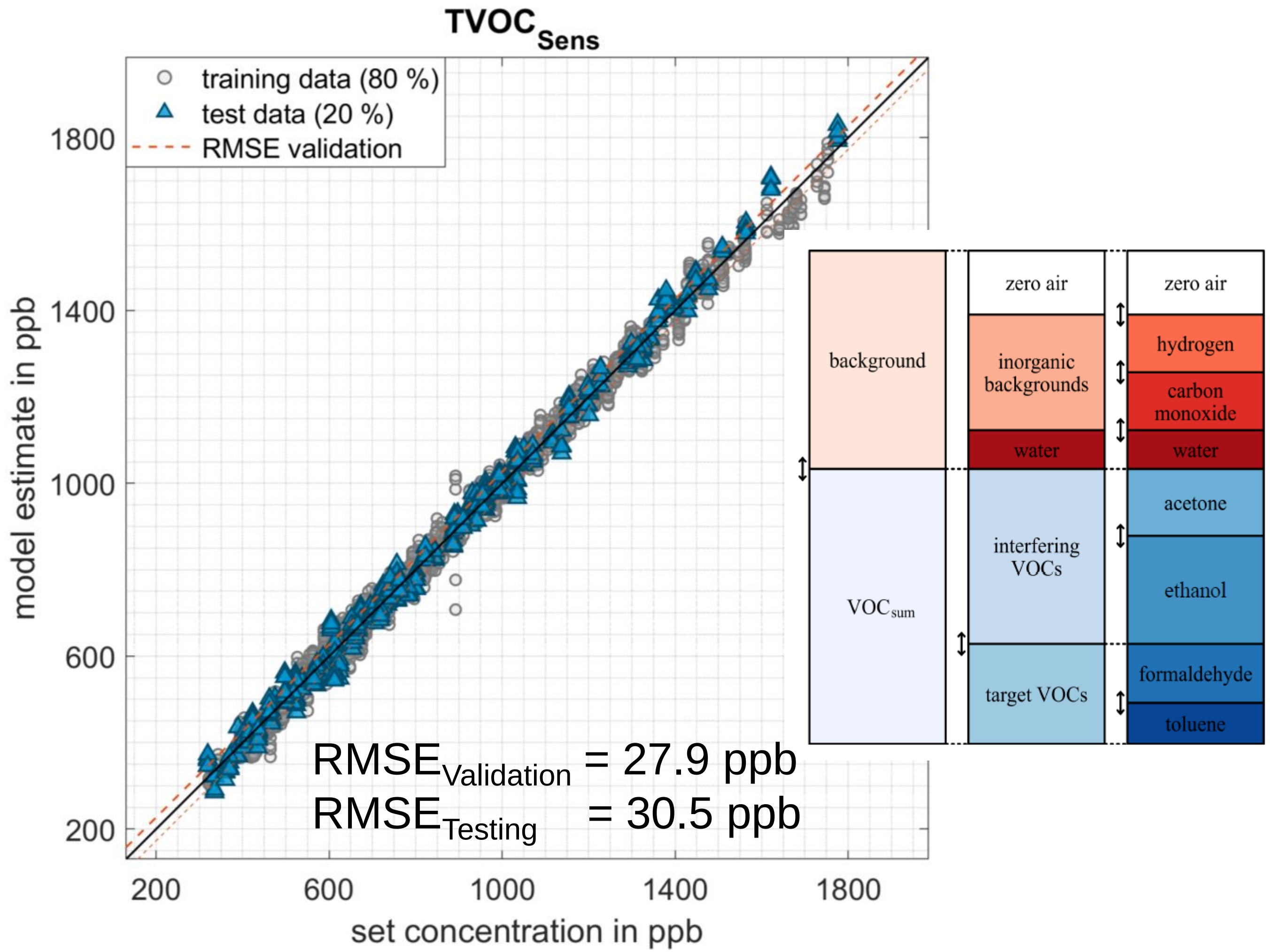
Photo: air quality sensor
Renesas www.renesas.com





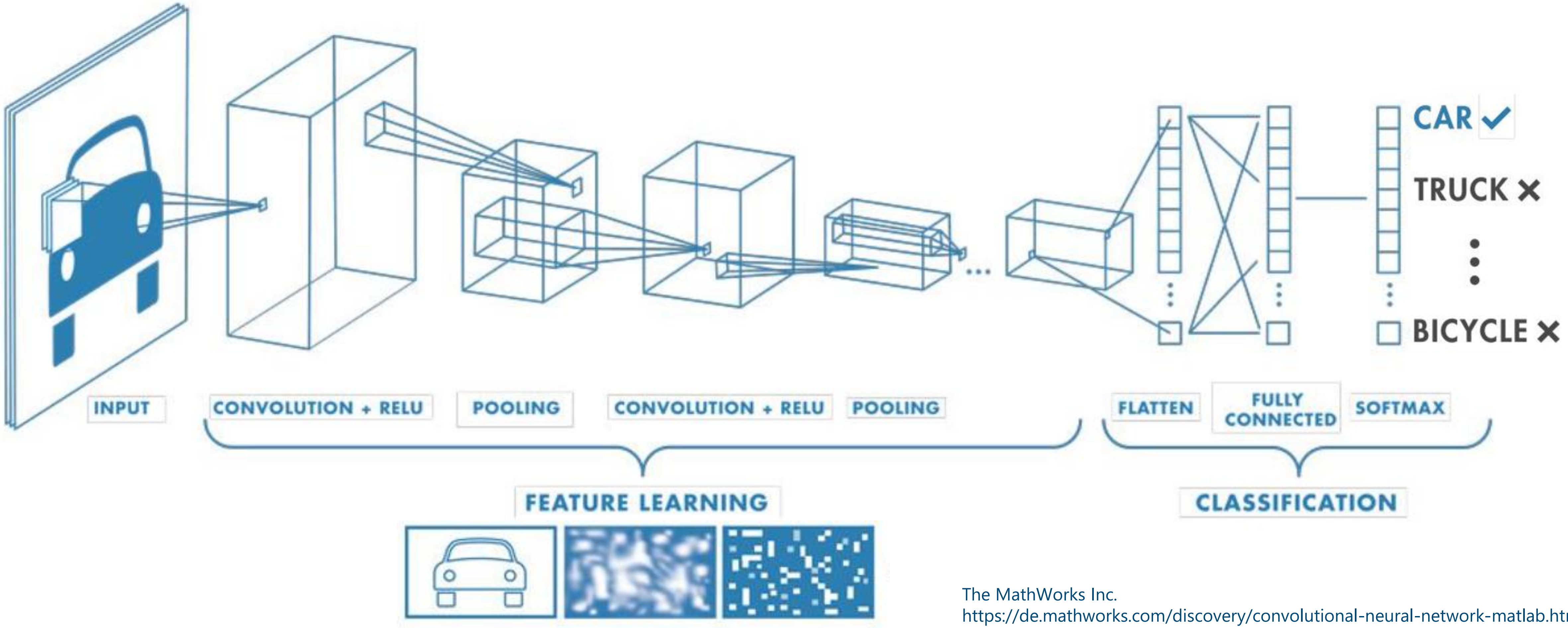
Results Partial Least Squares Regression (PLSR)





T. Baur et al.: Atmosphere 2021, 12(5), 647; doi: 10.3390/atmos12050647

Convolution Neural Networks – Image Processing



The MathWorks Inc.
https://de.mathworks.com/discovery/convolutional-neural-network-matlab.html

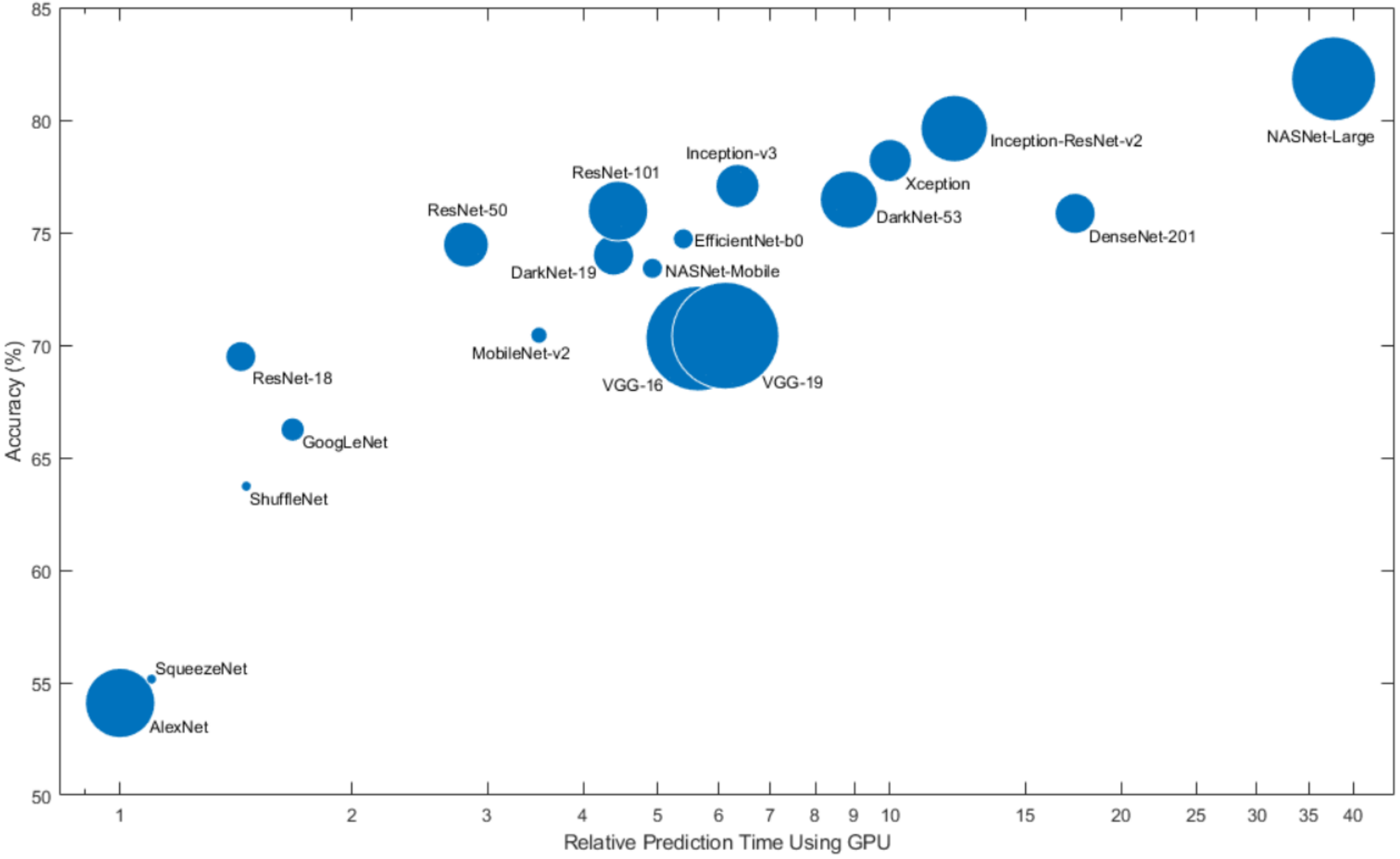
Famous Convolution Neural Networks

IMAGENET

Annual competition

1.2 Mio images

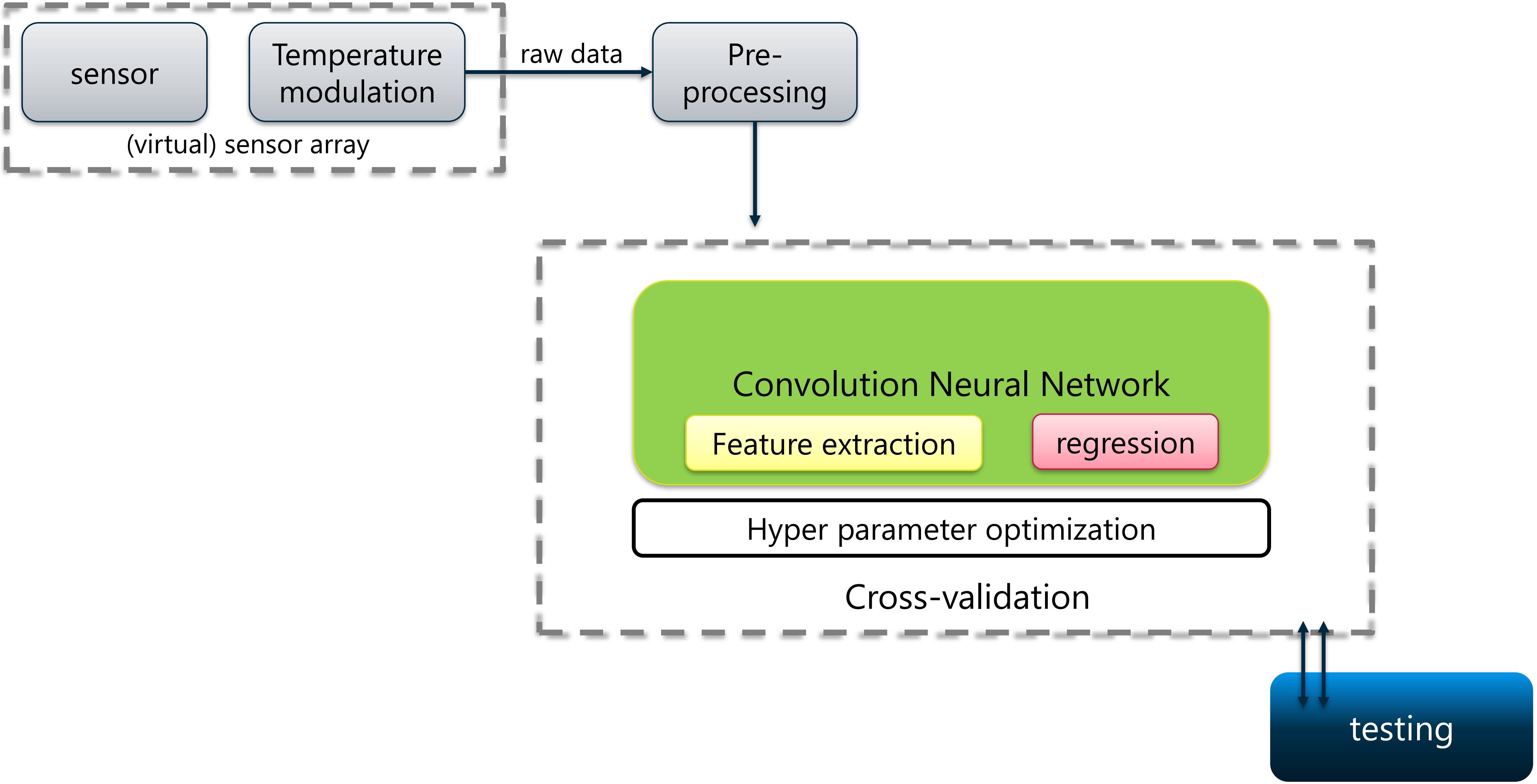
1000 categories



The MathWorks Inc.
https://de.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html

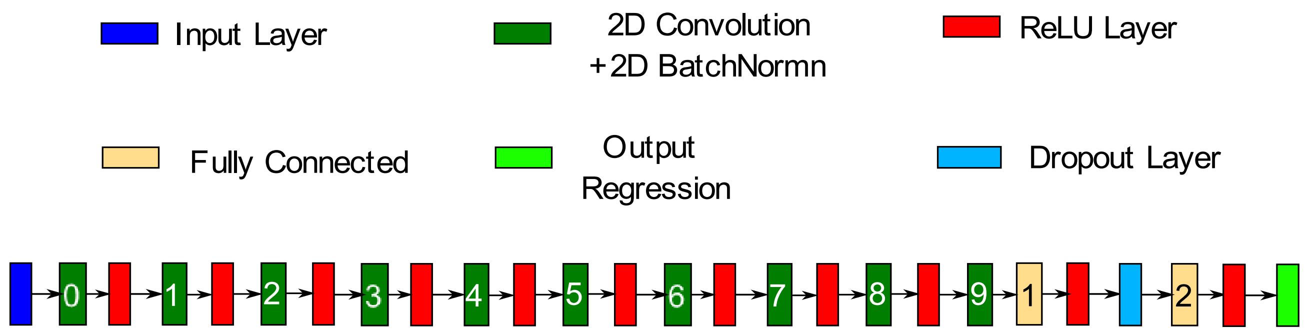
ImageNet
https://www.image-net.org/

Deep Learning Approach for Gas Sensing



Designing a TCO-CNN

10-layer deep convolutional neural network



Input layer:

[number of gas sensitive layers] x [sample points per cycle]

e.g. SGP40, Sensirion, Switzerland: **4 x 1440**

Output layer:

Regression layer with mean squared error as loss function

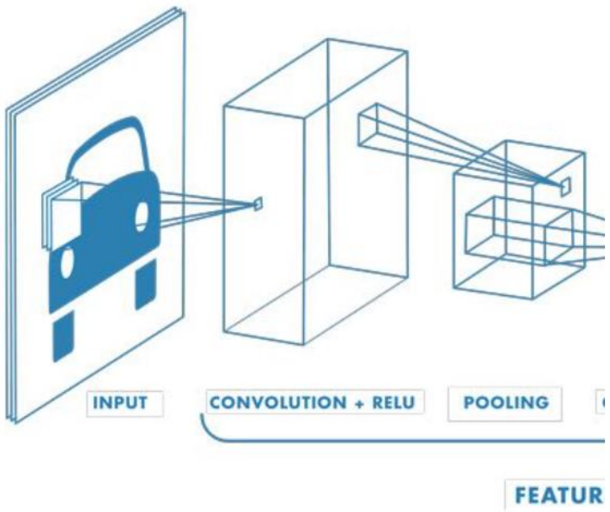
	Filter Size	Striding	# Filter	Output
2DConv #0	1 x 96	1 x 15	153	4 x 80 x 153
2DConv #1	1 x 1	1 x 1	153	4 x 80 x 153
2DConv #2	1 x 2	1 x 2	153	4 x 40 x 153
2DConv #3	1 x 1	1 x 1	153	4 x 40 x 153
2DConv #4	1 x 2	1 x 2	306	4 x 20 x 306
2DConv #5	1 x 1	1 x 1	306	4 x 20 x 306
2DConv #6	1 x 2	1 x 2	459	4 x 10 x 459
2DConv #7	1 x 1	1 x 1	459	4 x 10 x 459
2DConv #8	1 x 2	1 x 2	612	4 x 5 x 612
2DConv #9	1 x 1	1 x 1	612	4 x 5 x 612
Fully Connected #1	1x 12240	1x12240	1280	1 x 1280
Fully Connected #2	1 x 1280	1 x 1280	1	1

Y. Robin et al., Atmosphere 2021, 12(11), 1487, DOI 10.3390/atmos12111487
Y. Robin et al., 15. Dresdner Sensor-Symposium, 2021, https://www.ama-science.org/proceedings/details/4119
Y. Robin et al., IEEE I2MTC 2021, May 17-20, 2021

Neural Architecture Search (NAS)

Hyperparameter optimization

- ➢ Learning rate
- ➢ Number of filters
- ➢ Kernel size
- ➢ Stride size
- ➢ Drop out rate
- ➢ Number of neurons (fully connected layers)



The MathWorks Inc.
https://de.mathworks.com/discovery/convolutional-neural-network-matlab.html

Bayesian optimization search to find smallest RMSE

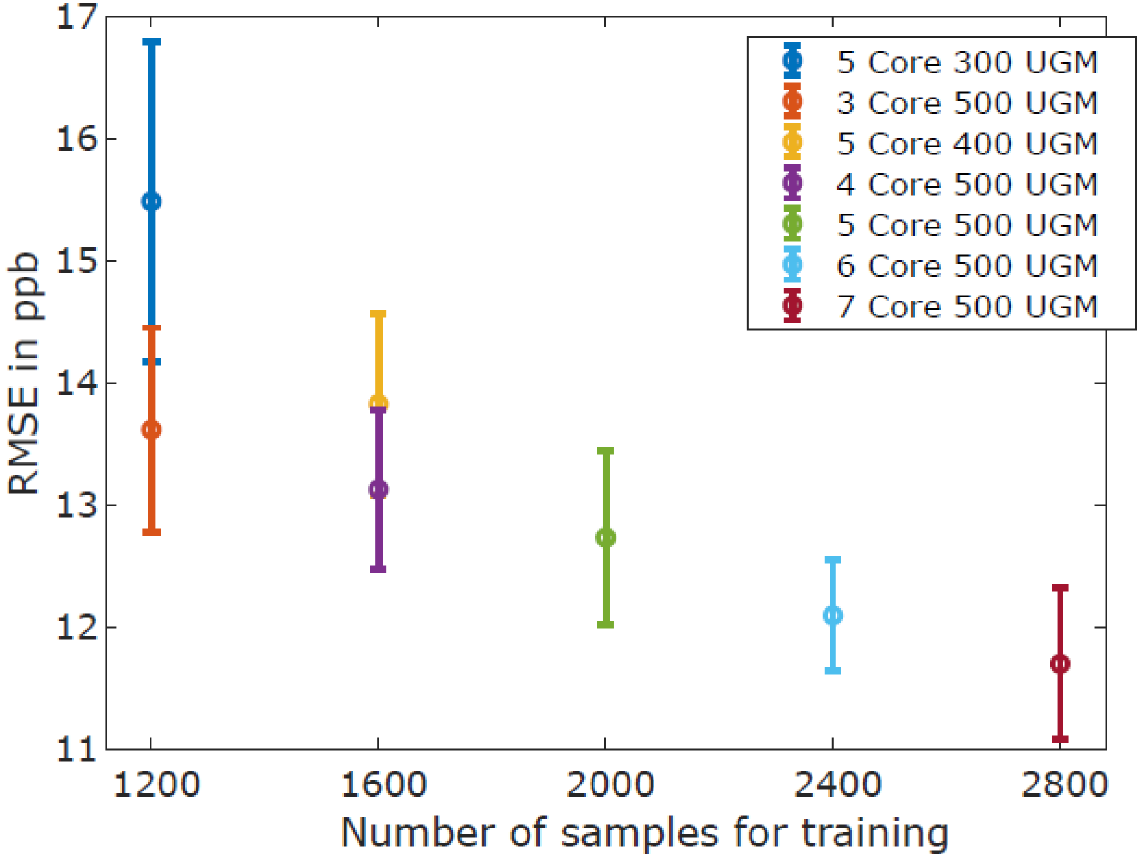
Initial Learning Rate (Log Scale)	Number of Filters (First Two Layers)	Kernel Size (First Two Layers)	Stride Size (First Layer)	Dropout	Number of Neurons (FC)
1×10^{-4}–9×10^{-3}	60–240	40–80	15–45	30–50%	1000–2500

Y. Robin et al.: Atmosphere 2021, 12(11), 1487, DOI 10.3390/atmos12111487

How Many Samples are Needed?

- Sample points (T-cycles) within an exposure (UGM)
- Number of (unique) gas exposures (UGMs)

		sensor 0	sensor 1	sensor 2	sensor 3	annotation
		s1 … sN	s1 … sN	s1 … sN	s1 … sN	
UGM 1	sample 1					20 ppb
	sample 2					20 ppb
	sample 3					20 ppb
	…					20 ppb
	sample 10					20 ppb
UGM 2	sample 1					50 ppb
	sample 2					50 ppb
	sample 3					50 ppb
	…					50 ppb
	sample 10					50 ppb
UGM 3	sample 1					30 ppb
	sample 2					30 ppb
	sample 3					30 ppb
	…					30 ppb
	sample 10					30 ppb



Results are given for TCOCNN

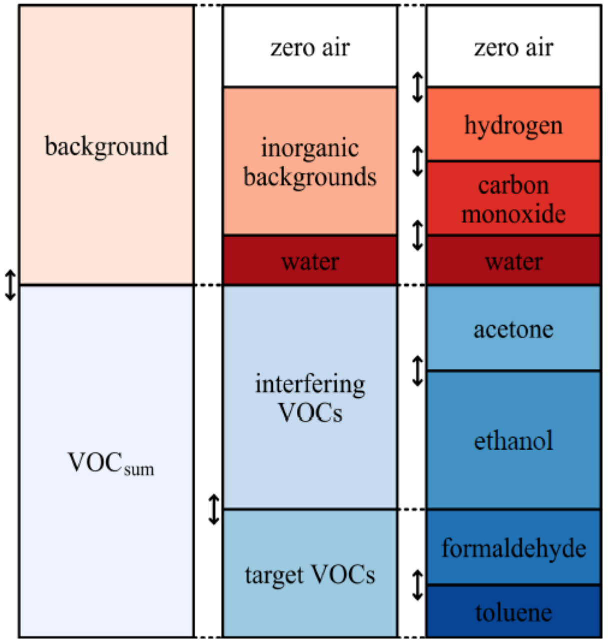
Y. Robin et al.: Atmosphere 2021, 12(11), 1487, DOI 10.3390/atmos12111487

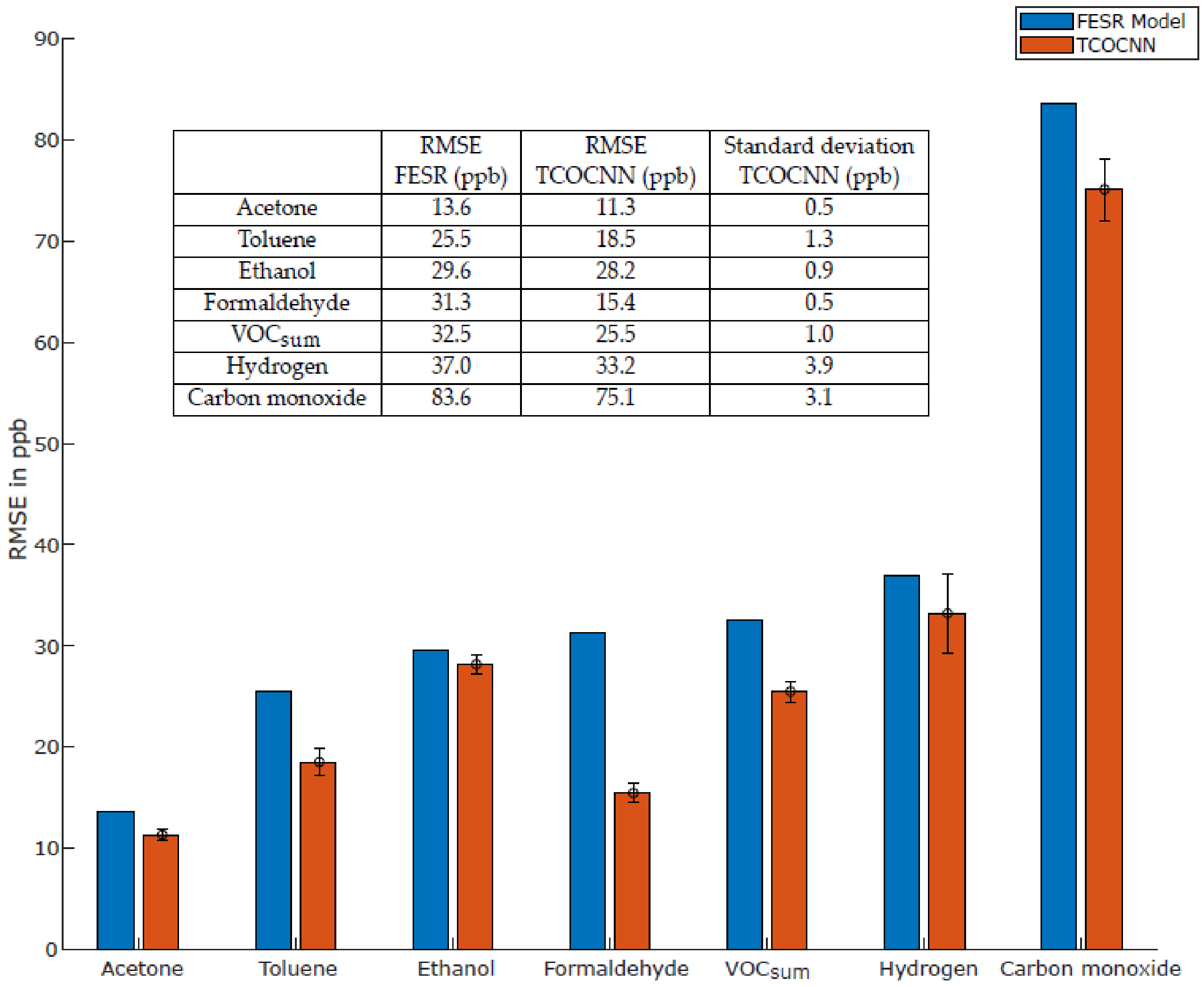
Comparison FESR and TCOCNN

Lab for Measurement Technology

UNIVERSITÄT DES SAARLANDES

- Hyperparameters are optimized for each gas individually
- TCOCNN outperforms FESR





	RMSE FESR (ppb)	RMSE TCOCNN (ppb)	Standard deviation TCOCNN (ppb)
Acetone	13.6	11.3	0.5
Toluene	25.5	18.5	1.3
Ethanol	29.6	28.2	0.9
Formaldehyde	31.3	15.4	0.5
VOC_{sum}	32.5	25.5	1.0
Hydrogen	37.0	33.2	3.9
Carbon monoxide	83.6	75.1	3.1

Y. Robin et al.: Atmosphere 2021, 12(11), 1487, DOI 10.3390/atmos12111487

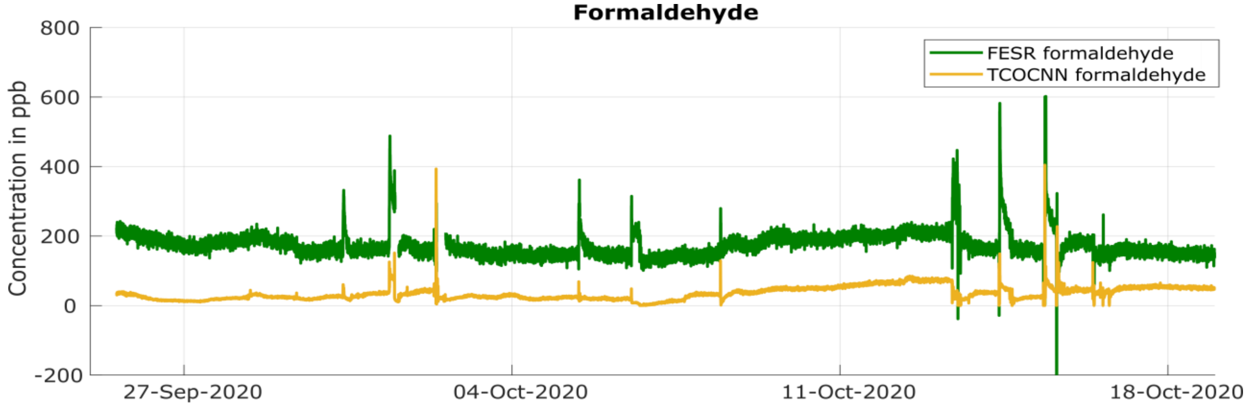
Results Field Tests

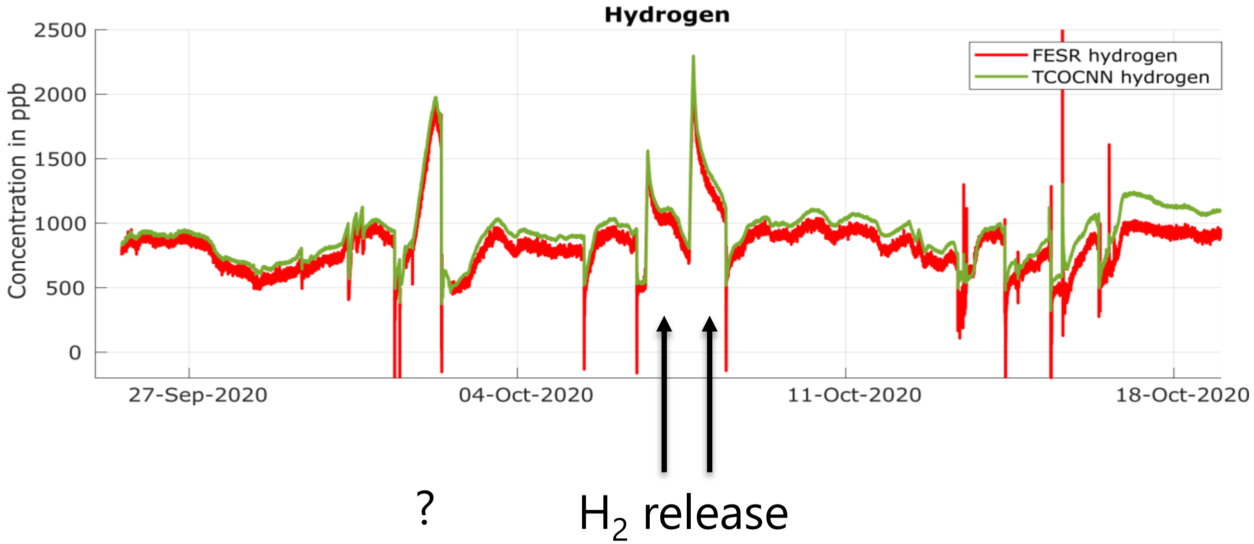
Lehrstuhl für Messtechnik

UNIVERSITÄT DES SAARLANDES

General findings compared to FESR methods:

- Baseline of the TCOCNN more realistic
  - Formaldehyde < 80 ppb (limit by WHO)
  - Hydrogen in atmosphere ~ 500 ppb
  - No calibrated reference values available therefore no absolute statement is possible
- TCOCNN is always above 0 ppb
- TCOCNN less noise



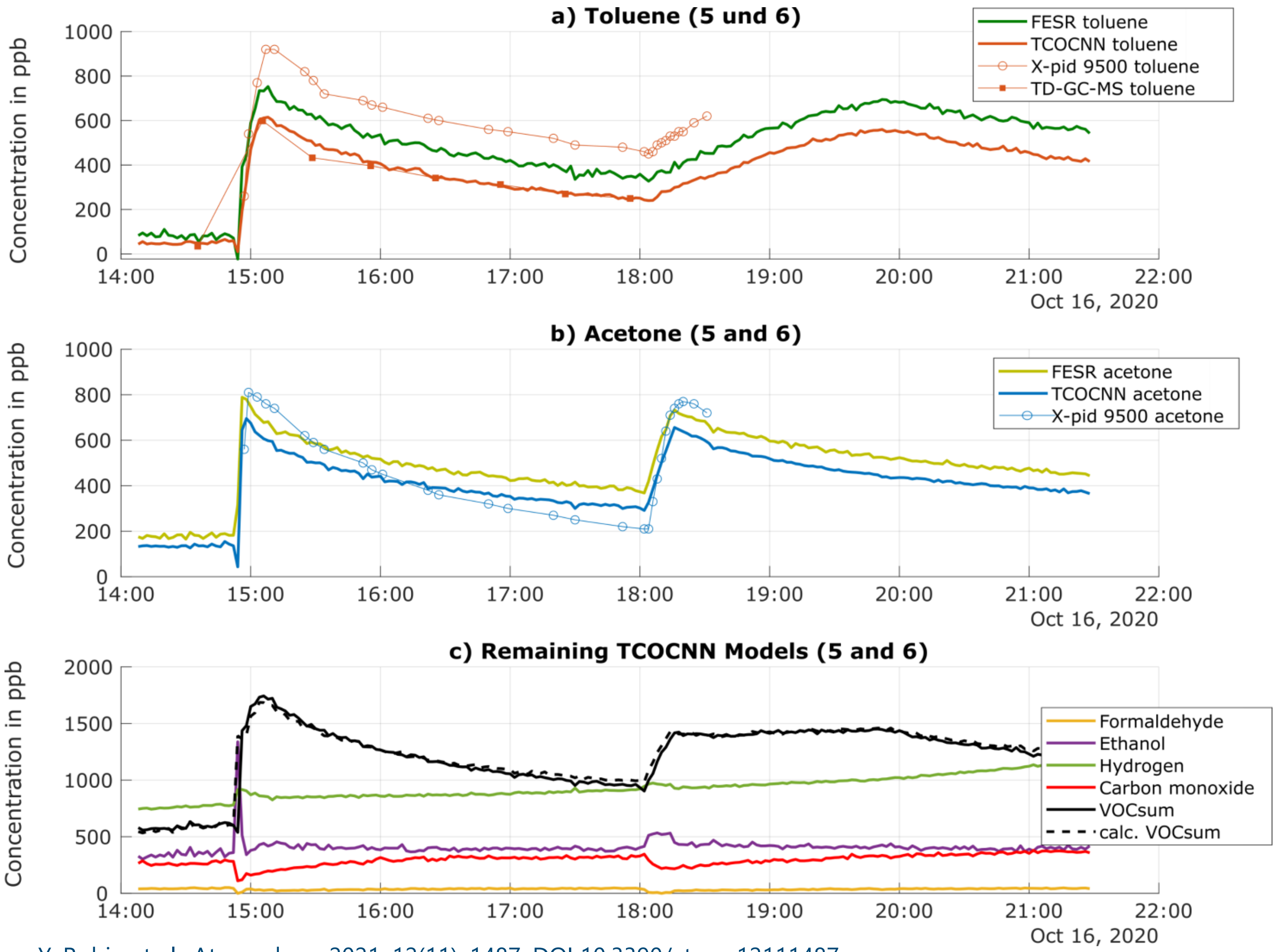


Y. Robin et al.: Atmosphere 2021, 12(11), 1487, DOI 10.3390/atmos12111487

Release Tests – Acetone and Toluene

Lehrstuhl für Messtechnik

UNIVERSITÄT DES SAARLANDES

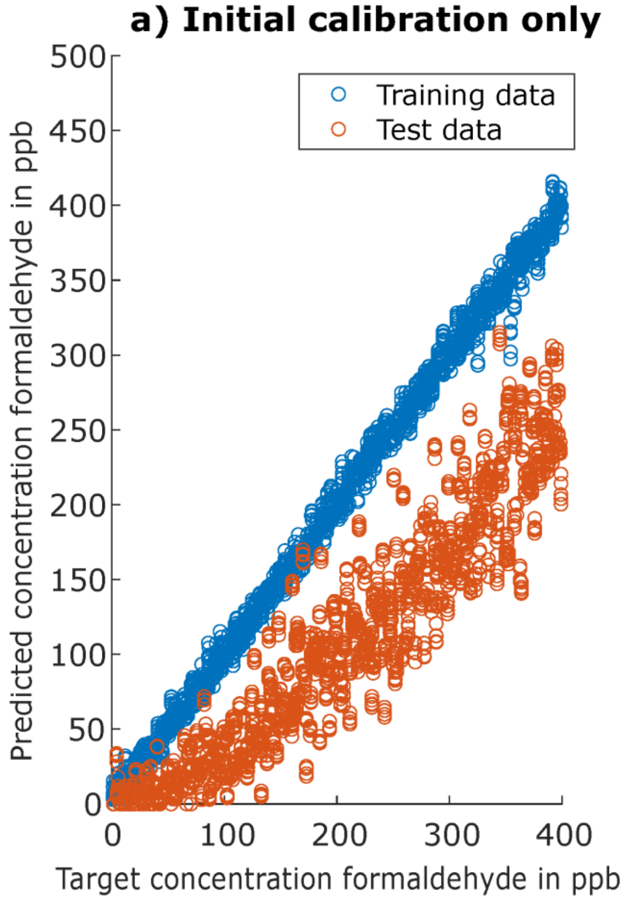


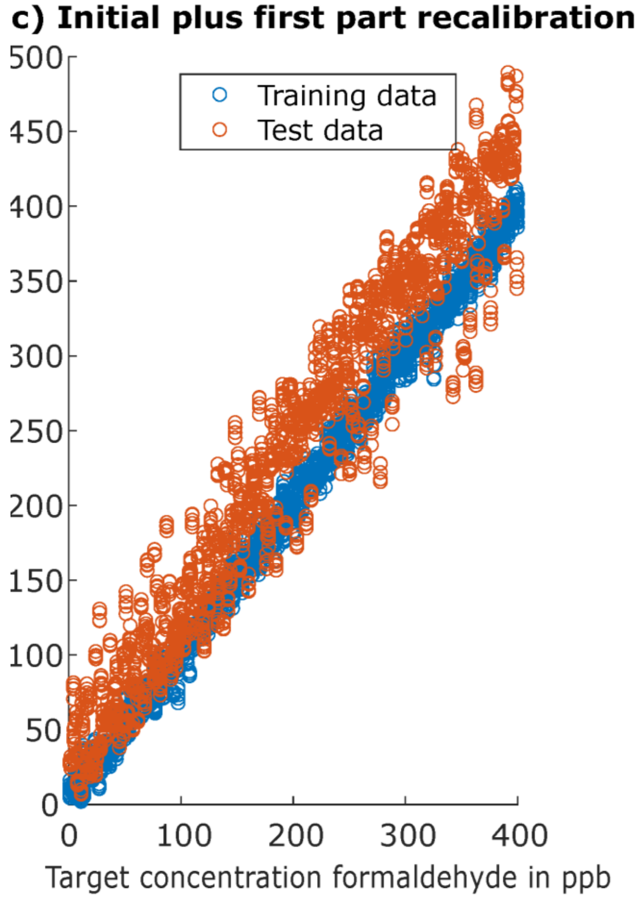
Y. Robin et al.: Atmosphere 2021, 12(11), 1487, DOI 10.3390/atmos12111487

- Toluene
  - General prediction of TCOCNN, FESR & GC-PID (X-pid 9500, Dräger) similar
  - Peak close to the expected value (600 ppb)
  - Highest match between TCOCNN & TD-GC-MS
- Acetone
  - General prediction of TCOCNN, FESR & GC-PID (X-pid 9500, Dräger) similar
  - Peak close to the expected value (600 ppb)
- The other TCOCNN Models are not influenced by release tests

Drift Compensation

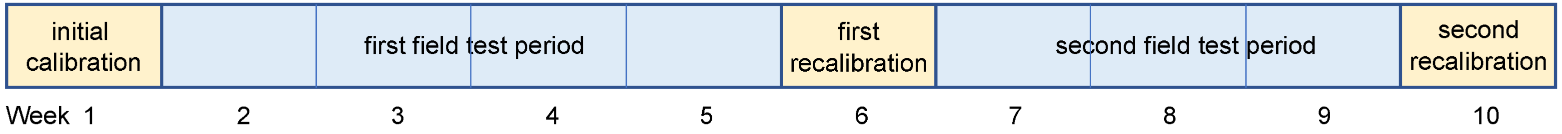
- Combine data from two calibration runs
- Include drift behavior in model building
- “global approach”





Y. Robin et al.: Atmosphere 2021, 12(11), 1487, DOI 10.3390/atmos12111487

Overview over measurement campaign



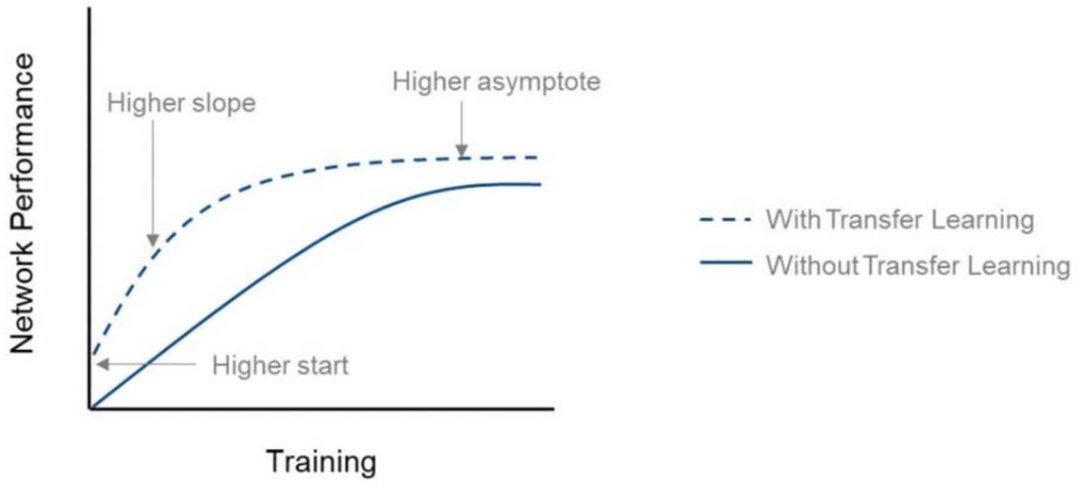
Transfer Learning
for calibration transfer
and calibration time reduction

Challenges

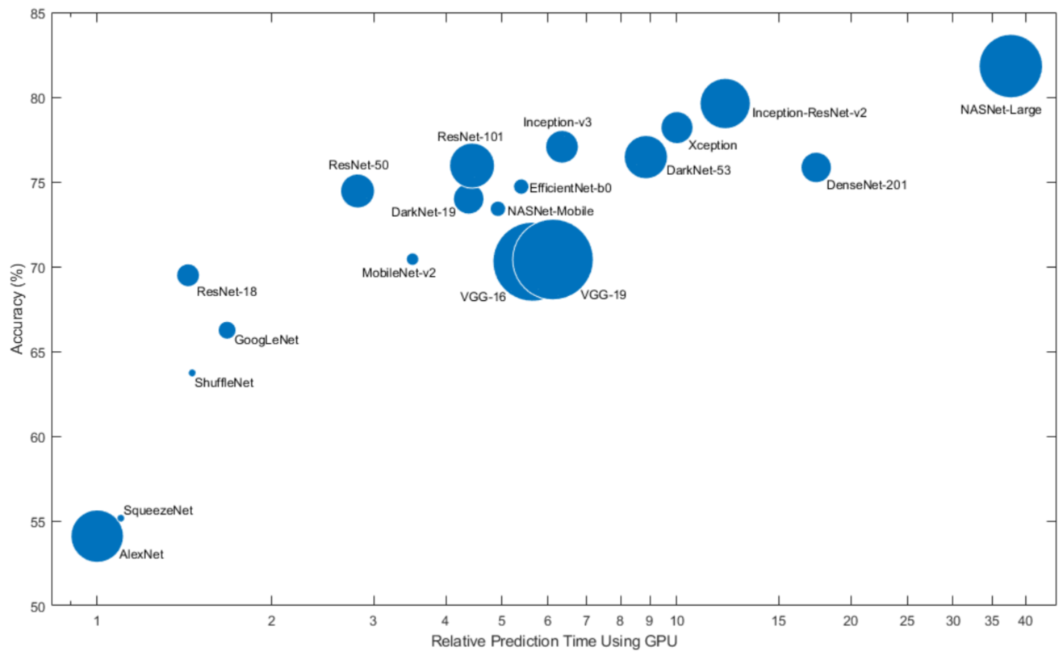
- Individual calibration needed
  - Variations in micro-heater
  - Variations in sensing layer
- Aging/poisoning of the sensors: need for recalibration
  - Changing of the sensing layer
  - Siloxane poisoning (growing glass layer on top of the sensing layer)
- Replacement of sensors in the field

Transfer Learning in Image Processing

- Use a pretrained network
- Replace input and output layer (if needed)
- Continue to train with less but new samples



Vergleich der Netzleistung (Genauigkeit) beim Trainieren von Grund auf und beim Transfer Learning.



The MathWorks Inc.
https://de.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html
https://de.mathworks.com/discovery/transfer-learning.html

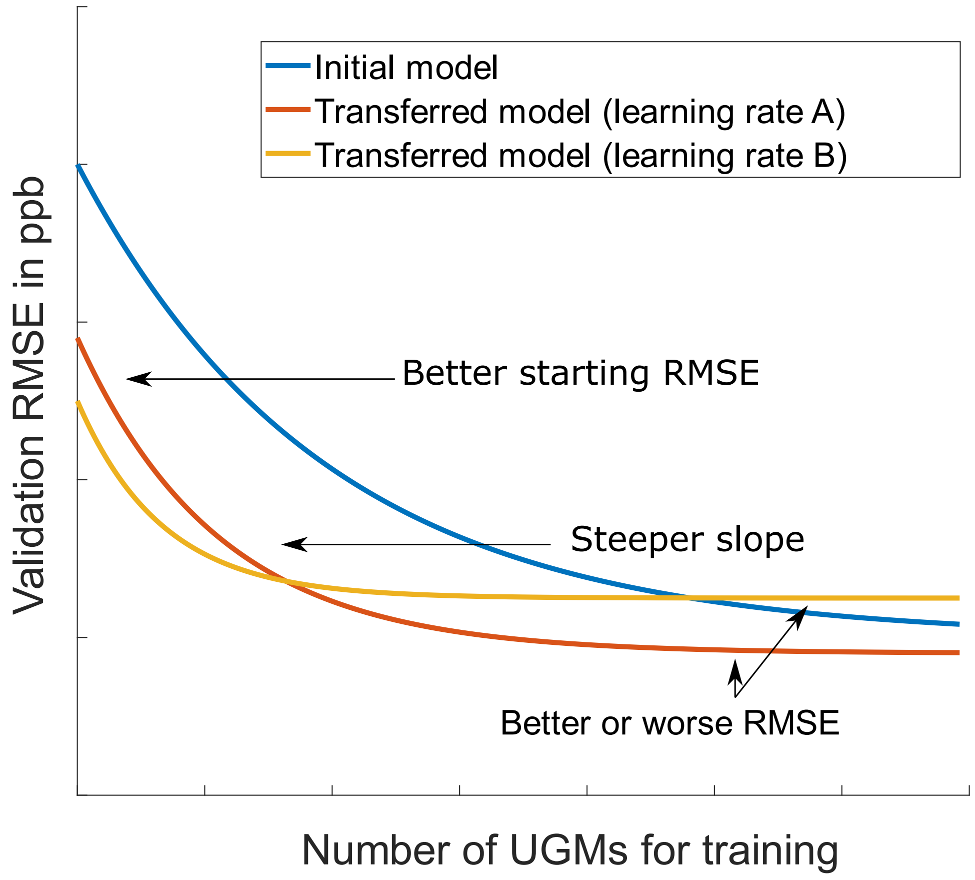
Transfer Learning TCOCNN

Concept of transfer learning

- Reduce calibration time with process insights from previous calibrations
  - Pre-trained model as a starting point
    (Instead of randomly chosen weights)

- Reduce calibration time to reach sufficient result

- Reach sometimes even better results



Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614
Y. Robin et al., ISOEN 2022, Aveiro, Portogal, doi: 10.1109/ISOEN54820.2022.9789596, Best Paper Award

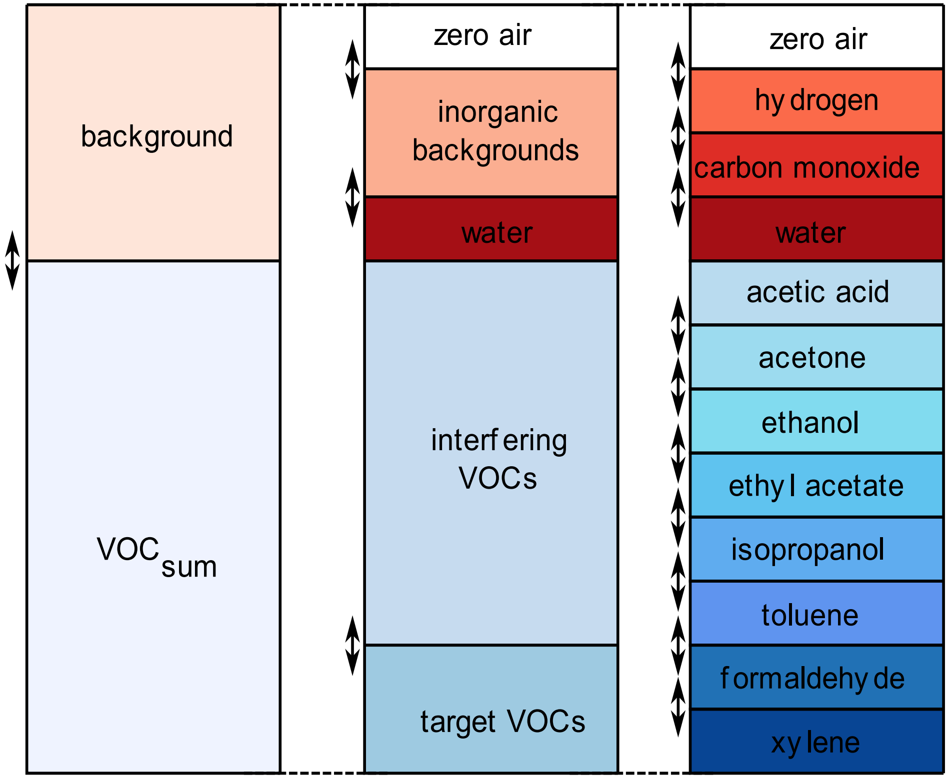
UGM: unique gas mixtures

Extended Dataset

Lab for Measurement Technology | UNIVERSITÄT DES SAARLANDES

Randomized gas mixtures

- 8 VOCs (formaldehyde, acetone, acetic acid, ethanol, toluene, **xylene**, isopropanol, ethyl acetate)
- 2 interfering gases (hydrogen and carbon monoxide) & relative humidity (RH)
- In total **900 unique gas mixtures** (UGMs), ~14 days
- Split into training set (700 UGMs) and test set (200 UGMs)



Sensors

- Several sensors of same type, i.e. SGP40, Sensirion
  - Sensor A
  - Sensor B
  - Sensor C

Sensor A and Sensor B: from same batch

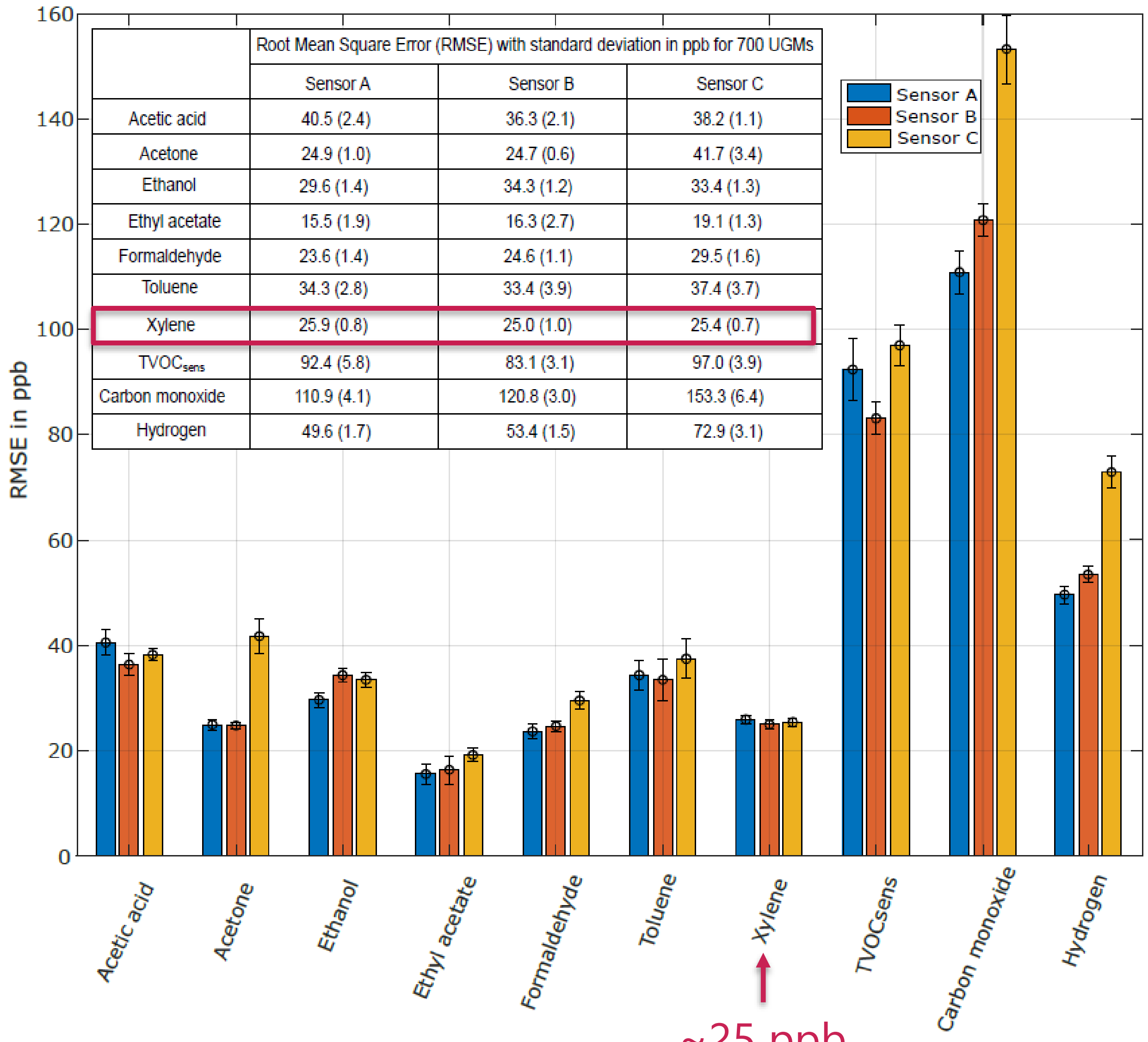
Sensor C: different batch

Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614
Y. Robin et al., ISOEN 2022, Aveiro, Portogal, doi: 10.1109/ISOEN54820.2022.9789596, Best Paper Award

Results Training

Training TCOCNNs – reference

- ➢ Train a TCOCNN for each gas and each of the three sensors A, B, C
- ➢ Hyperparameters are optimized for each gas but only for Sensor A
- ➢ Every model is trained 10 times
- ➢ For xylene: RMSE of **~25 ppb**



	Root Mean Square Error (RMSE) with standard deviation in ppb for 700 UGMs		
	Sensor A	Sensor B	Sensor C
Acetic acid	40.5 (2.4)	36.3 (2.1)	38.2 (1.1)
Acetone	24.9 (1.0)	24.7 (0.6)	41.7 (3.4)
Ethanol	29.6 (1.4)	34.3 (1.2)	33.4 (1.3)
Ethyl acetate	15.5 (1.9)	16.3 (2.7)	19.1 (1.3)
Formaldehyde	23.6 (1.4)	24.6 (1.1)	29.5 (1.6)
Toluene	34.3 (2.8)	33.4 (3.9)	37.4 (3.7)
Xylene	25.9 (0.8)	25.0 (1.0)	25.4 (0.7)
$TVOC_{sens}$	92.4 (5.8)	83.1 (3.1)	97.0 (3.9)
Carbon monoxide	110.9 (4.1)	120.8 (3.0)	153.3 (6.4)
Hydrogen	49.6 (1.7)	53.4 (1.5)	72.9 (3.1)

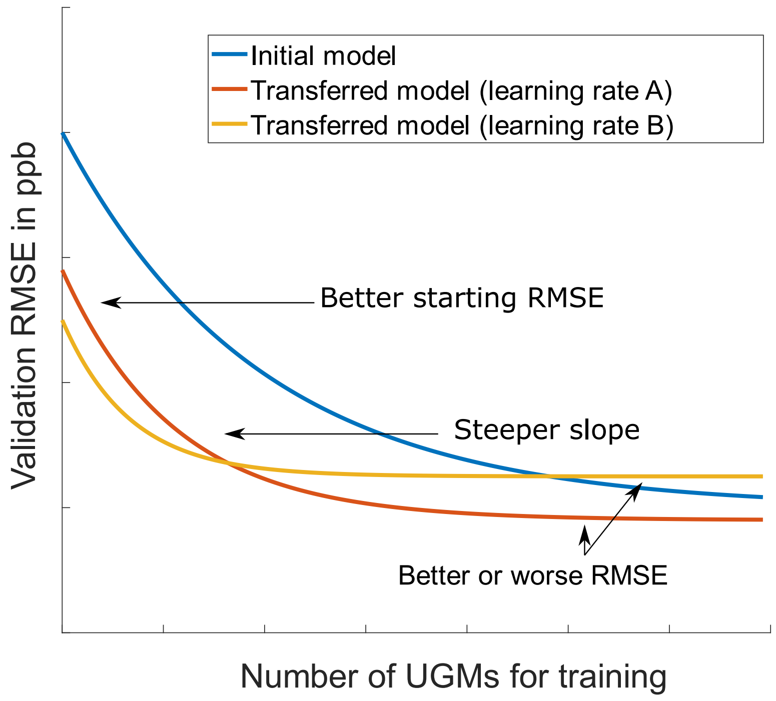
Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614

Transfer Learning TCOCNN

Lab for Measurement Technology
UNIVERSITÄT DES SAARLANDES

Validation approach:

- Use trained model of Sensor A with all 700 UGMs and transfer it to Sensor B or Sensor C
- Retrain the model (continue learning) with a few UGMs measured with Sensor B or C
- Study/minimize the number of necessary calibration samples (UGMs) from Sensor B or C



Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614

Results Transfer Learning

Individual training on 700 UGMs

- For all three sensors RMSE ~25 ppb

Without transfer learning

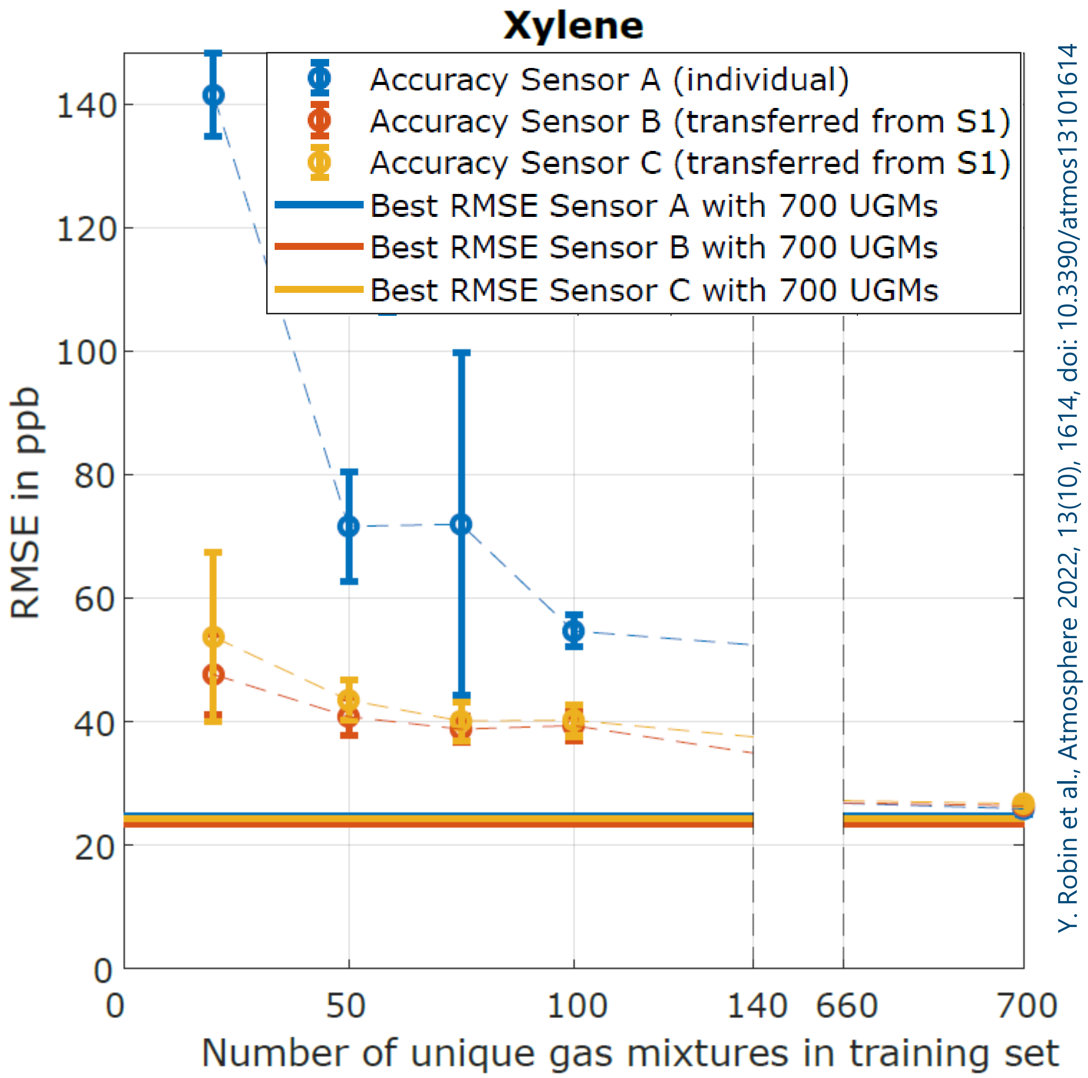
Simply apply model from Sensor A

- on test data Sensor B: ~74 ppb
- on test data Sensor C: ~103 ppb

Transfer learning

Use trained model from Sensor A plus

- 20 UGMs (i.e. 3 %) Sensor B/C: ~47 ppb / 55 ppb
- 100 UGMs (i.e. 14 %) Sensor B/C: ~40 ppb / 40 ppb



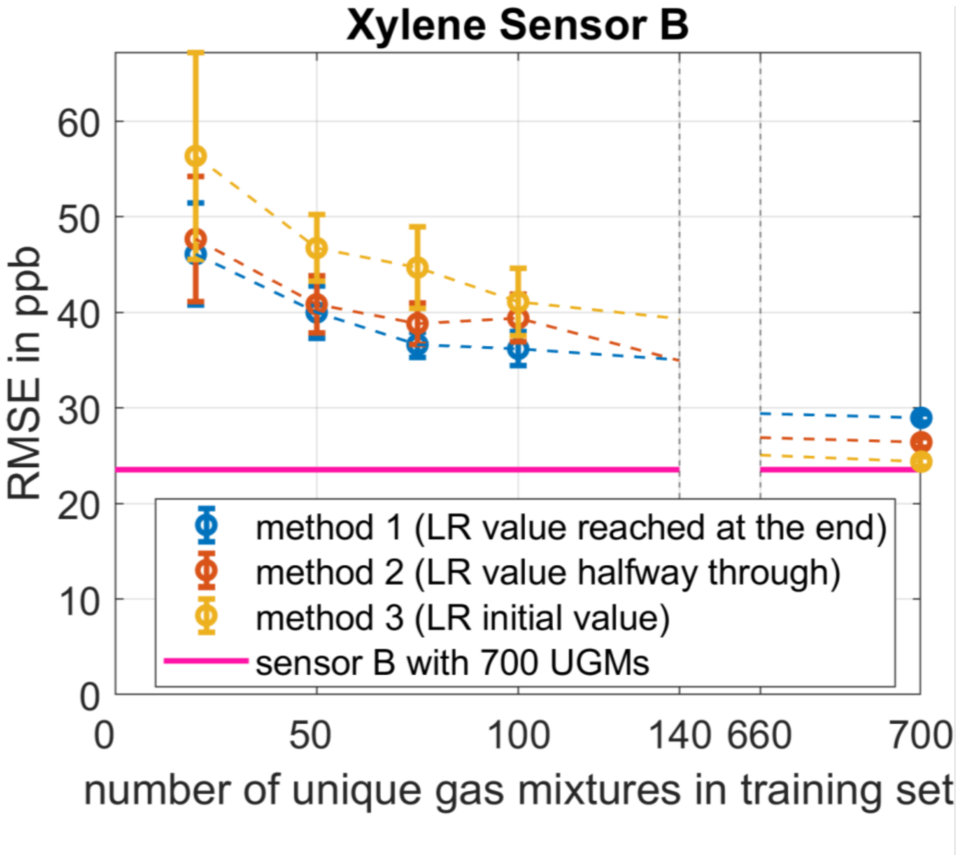
Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614

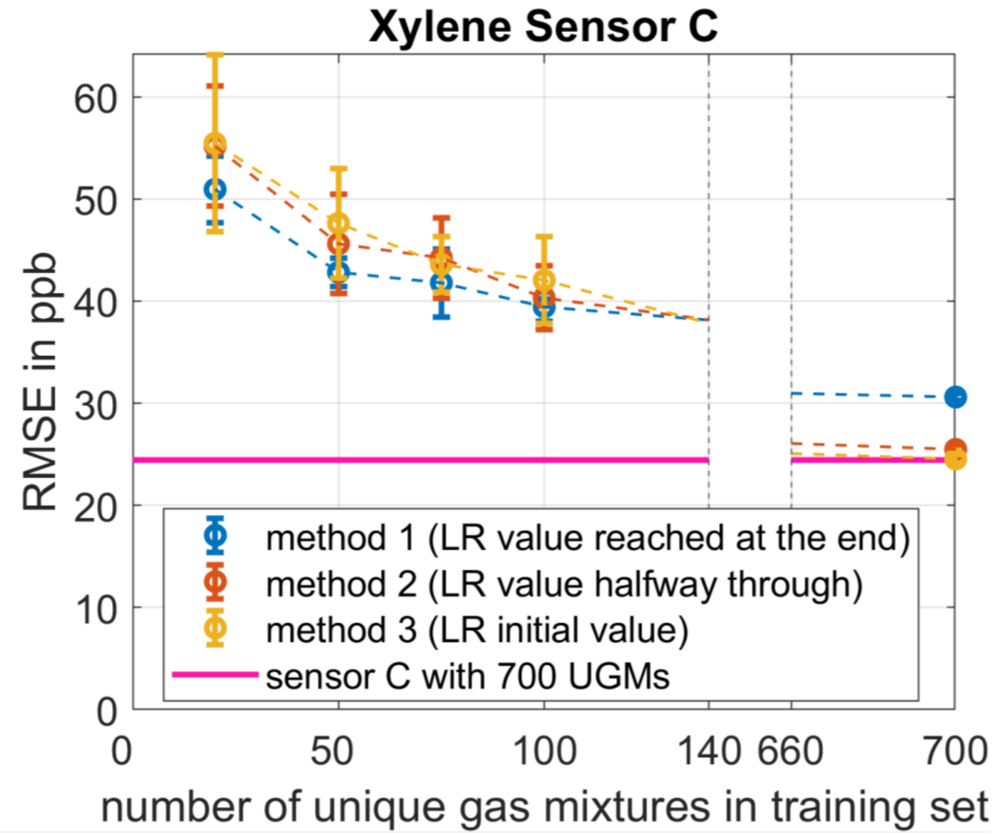
Parameters in Transfer Learning

- Learning rate
  - Hyperparameter for transfer learning
- Layers of the CNN to be adapted
  - Adapt every layer or keep feature extraction constant?
- Number of UGMs for transfer learning
  - How many calibration samples are needed when using transfer learning?
- Which UGMs to choose (influence of sampling)
  - Impact of the chosen UGMS for transfer learning

Influence of the Learning Rate

- The learning rate determines the performance of transfer learning
- A smaller learning rate seems better (method 1, blue)



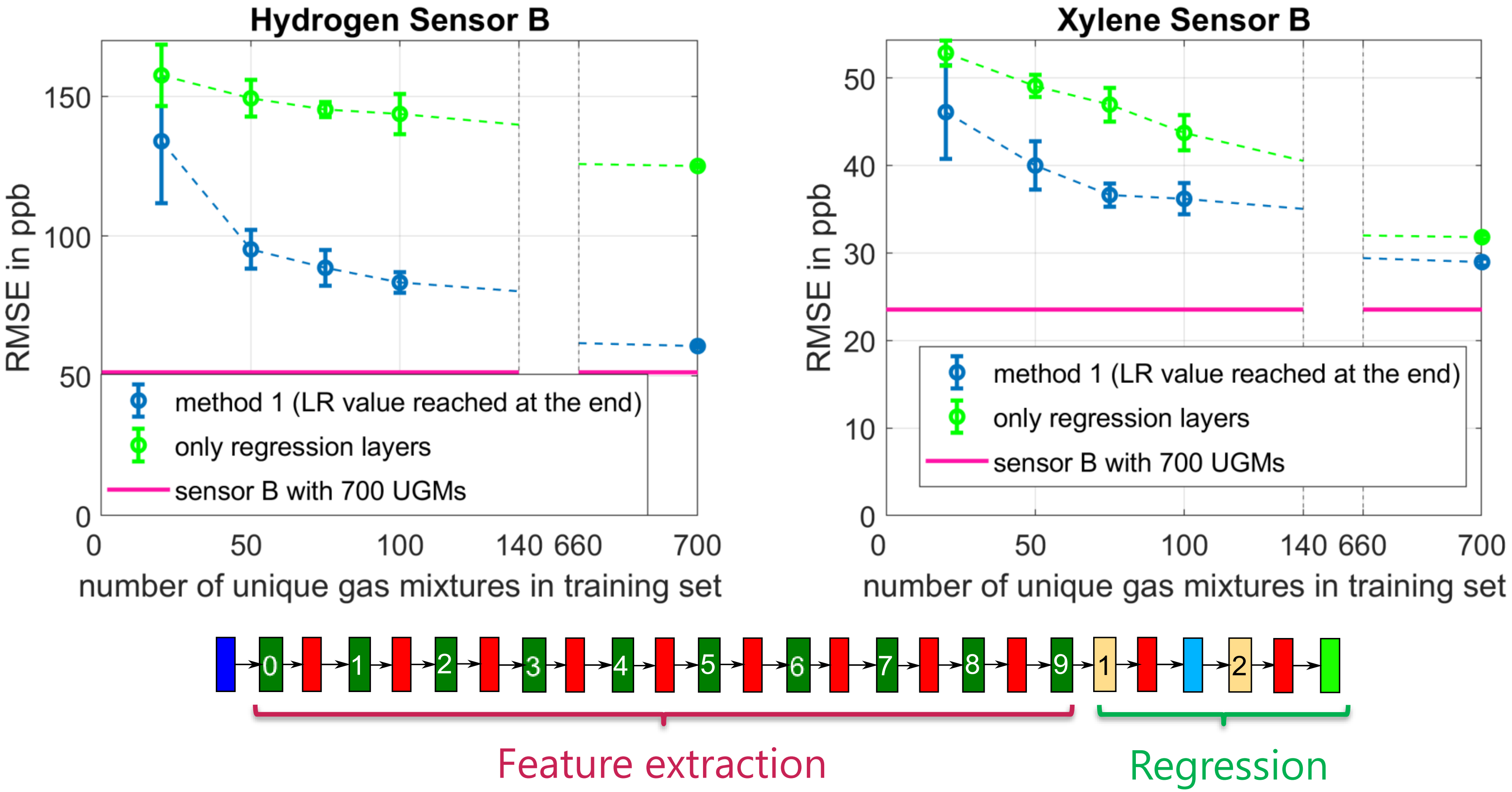


Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614

Which Layers of the TCOCNN to Adapt?

Lab for Measurement Technology | UNIVERSITÄT DES SAARLANDES

Transfer learning can be applied to all or just a subset of layers



Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614

Random Subsampling – Influence of UGMs?

- Sensor A

  is trained on 700 UGMs

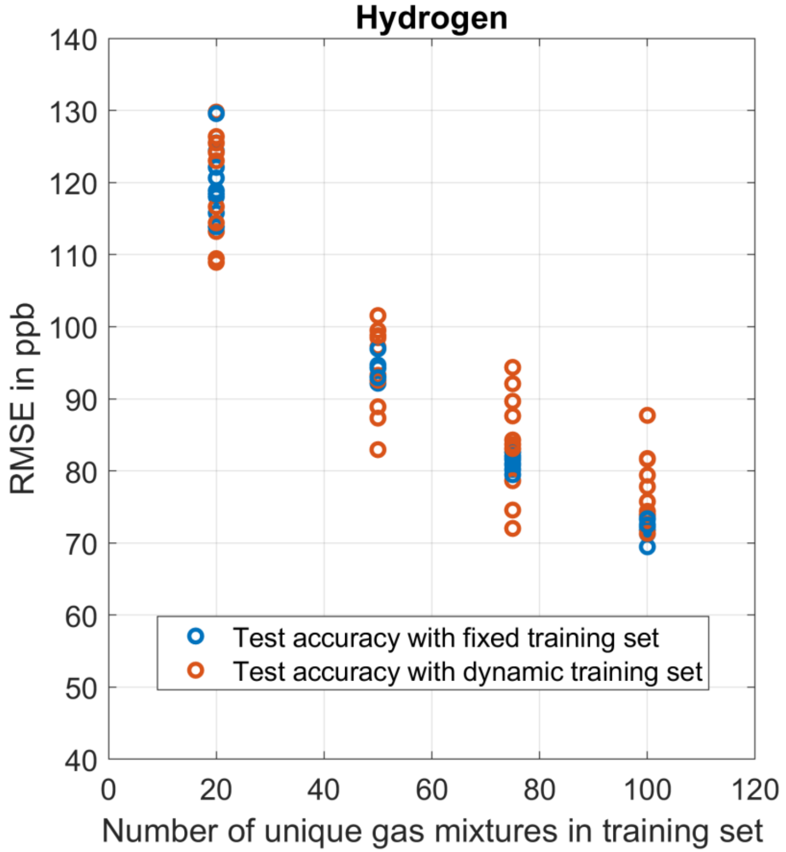
- Sensor B (transfer learning)

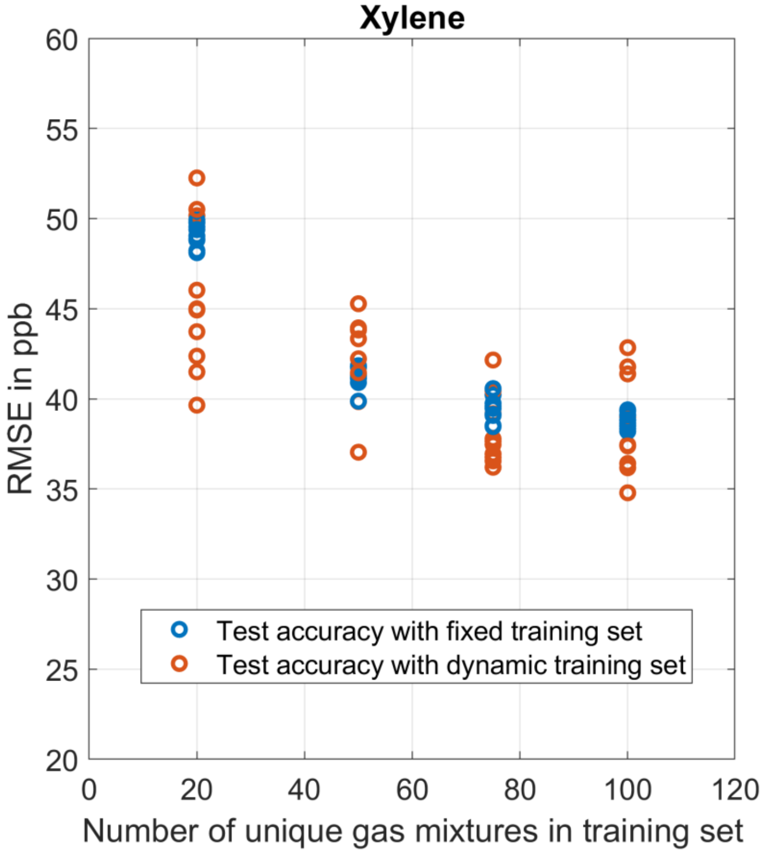
  20 to 100 UGMs

  randomly picked

  with 10 iterations

The choice of UGMs for transfer learning is crucial

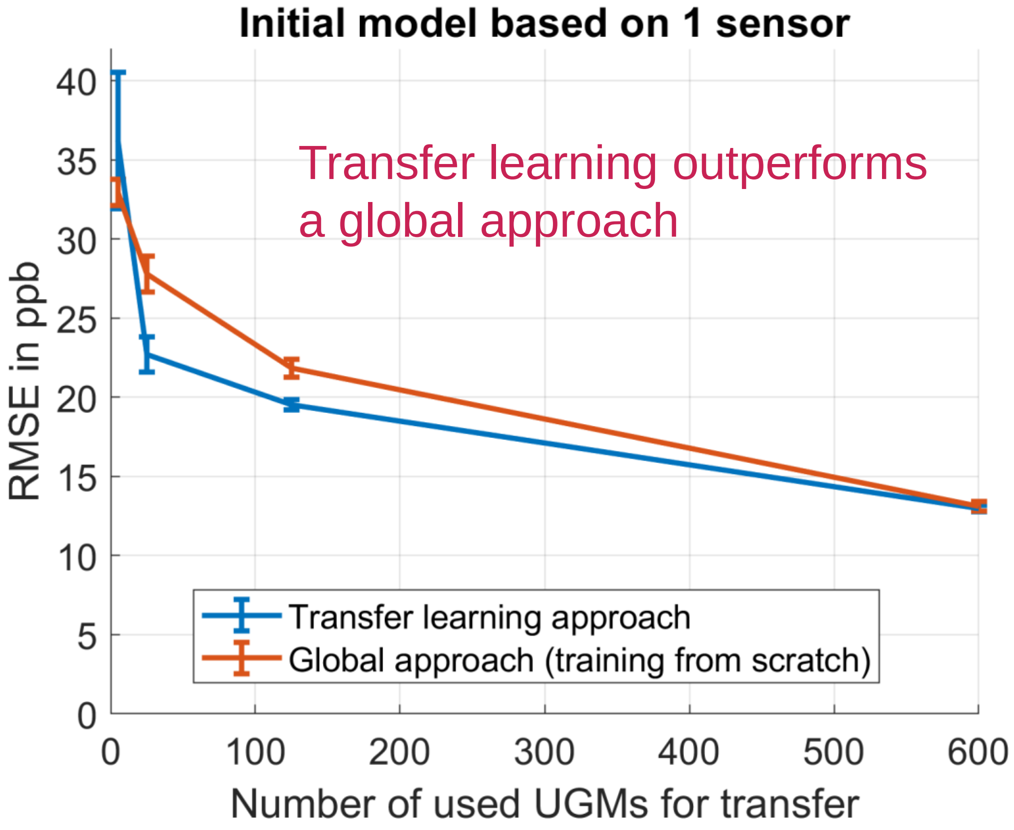


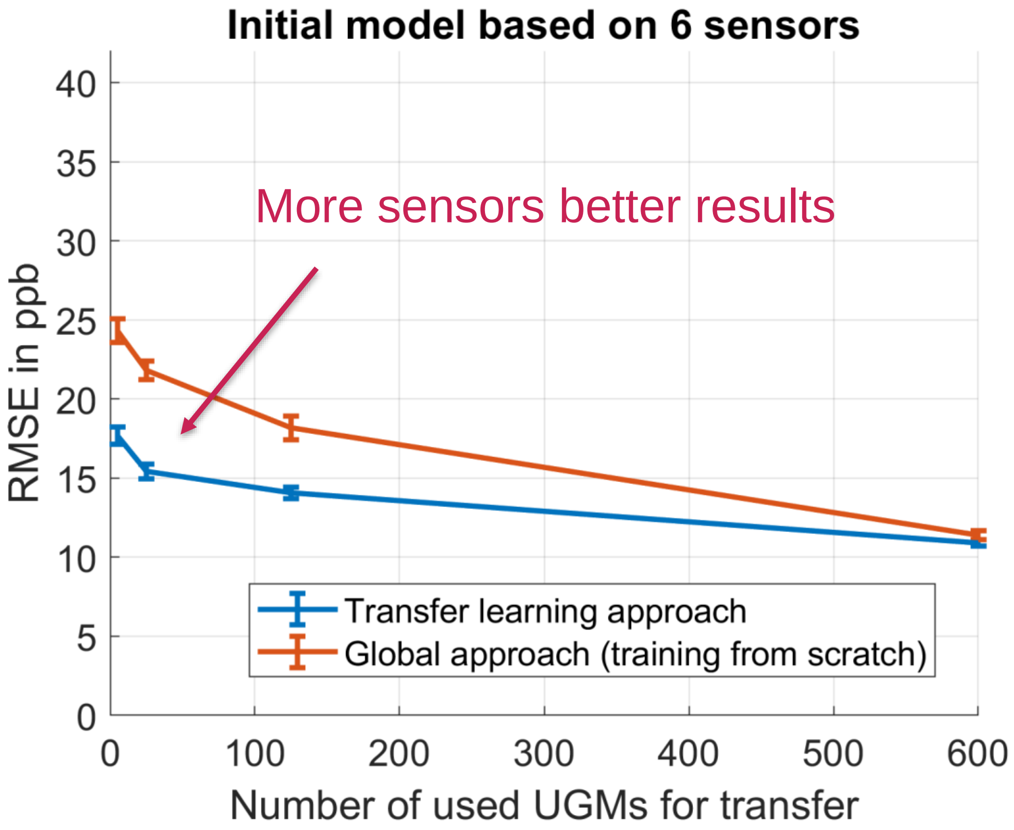


Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614

Global Approach vs. Transfer Learning

- Global Approach: use data from all sensors and train from scratch
- Transfer Learning: initial model trained on 1 (6) sensors, transferred to new





Acetone

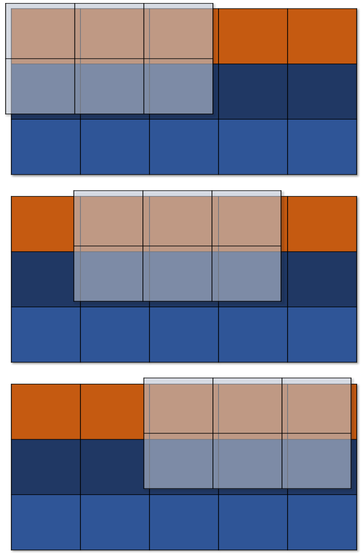
Outlook: Explainable AI

Explainable Machine Learning Algorithms

Occlusion Map:

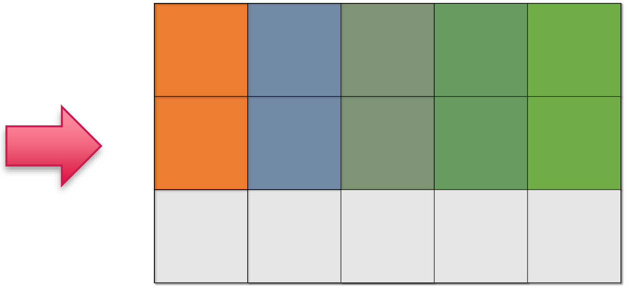
- Method to determine the most important parts in an image for the trained task
- Works on single instances
  - Occlude certain areas and calculate the difference in performance

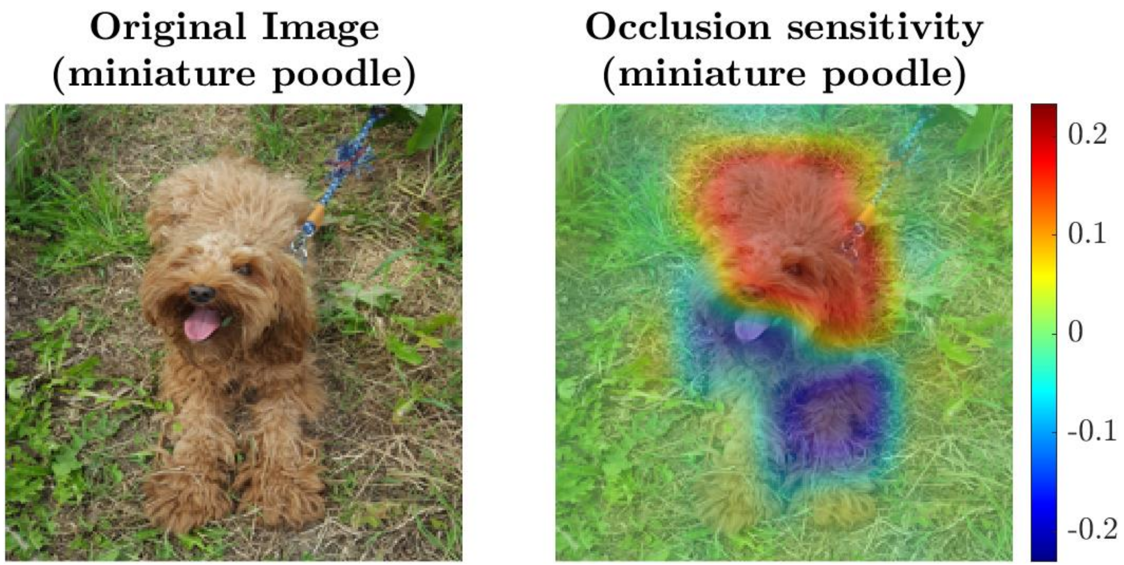
Cover area:



Jannis Morsch Master Arbeit

Calculate occlusion map:



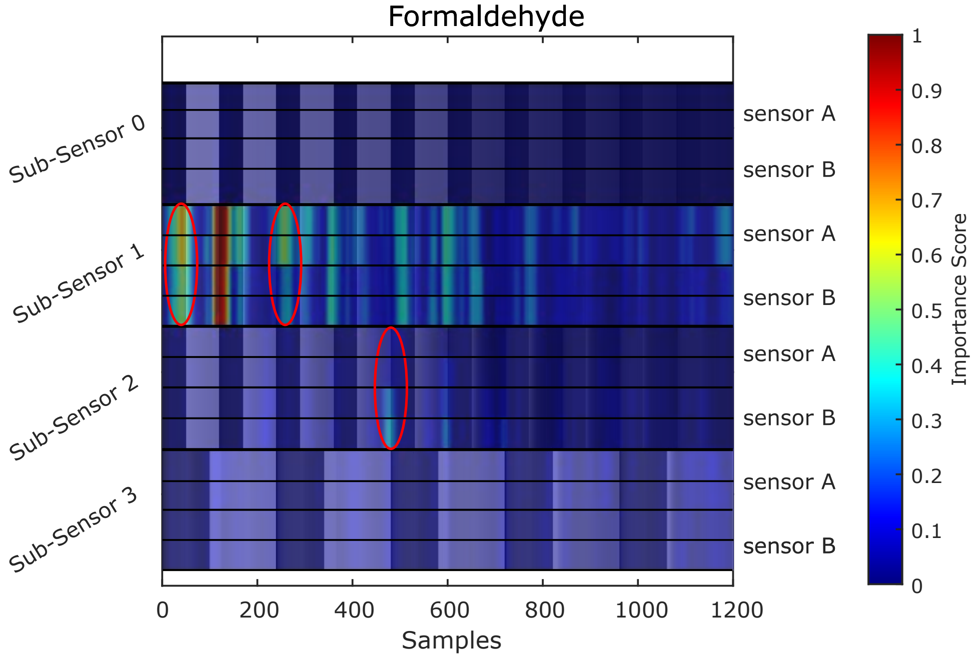


The MathWorks Inc.
https://de.mathworks.com/help/deeplearning/ug/understand-network-predictions-using-occlusion.html?searchHighlight=occlusion&s_tid=srchtitle_occlusion_4

Explainable Machine Learning Algorithms

Occlusion Map:

- Example Formaldehyde (two different sensors)
- Red resembles important area
- Differences between sensors visible
  - Can be used to improve transfer



Extraction of the most important areas:

- Occlusion map indeed highlights the most important areas
- Can be used to optimize the TCO
  - Formaldehyde: 50 % TC reduction

	Training data set	Mean RMSE in ppb ± standard deviation
Sensor A	All data	**15.8 ± 0.3**
	w/o most important 7 %	23.8 ± 1.0
	Only most important 7 %	19.3 ± 1.0
Sensor B	All data	**18.8 ± 0.6**
	w/o most important 7 %	26.3 ± 0.5
	Only most important 7 %	19.9 ± 1.0

Y. Robin et al., *submitted to I2MTC 2023*

Conclusion

Take Home Messages

Lab for Measurement Technology

UNIVERSITÄT DES SAARLANDES

- ✓ Understand your measuring task
- ✓ Use representative, comprehensive, and well annotated dataset
- ✓ Check your calibration in field measurements
- ✓ Deep Learning can outperform classic ML approaches
- ✓ Transfer Learning as an effective way to reduce calibration time

Team at LMT

- Yannick Robin, PhD student (deep learning and transfer learning)
- Johannes Amann, PhD student (indoor air quality)
- Dennis Arendes, future PhD student (gas mixing apparatus)
- Julian Joppich, PhD student (food quality, reducing food waste)
- Henrik Lensch, PhD student (hardware, data management)
- Oliver Brieger, PhD student (MOS sensor as GC detector)
- Wolfhard Reimringer, PhD student (electronics, mini-GC)
- My Sa Marschibois, lab engineer (GC-MS, cultural heritage, general support)
- Dr. Christian Bur, postdoc (bio-medical applications, breath analysis)
- Prof. Dr. Andreas Schütze

www.lmt.uni-saarland.de

Dr.-Ing. Christian Bur

Saarland University
Lab for Measurement Technology
Campus A5 1 | 66123 Saarbrücken | Germany
mail: c.bur@lmt.uni-saarland.de
phone: +49 681 / 302-2256

Literature – Temperature Cycled Operation

Lab for Measurement Technology

UNIVERSITÄT DES SAARLANDES

Temperature Cycled Operation (TCO)

- A. Schütze et al.: „Highly Sensitive and Selective VOC Sensor Systems Based on Semiconductor Gas Sensors: How to? ", Environments 2017, *4*, 20; doi: 10.3390/environments4010020
- A. Schütze et al. „Dynamic operation of semiconductor sensors, in: Raivo Jaaniso and Ooi Kiang Tan (eds.): Semiconductor Gas Sensors, Woodhead Publishing, 2nd Edition, 2020, Pages 385-412, ISBN: 9789176850039, doi: 10.1016/B978-0-08-102559-8.00012-4
- T. Baur et al., „Optimierung des temperaturzyklischen Betriebs von Halbleitergassensoren ", tm - Technisches Messen (2015), 82 (4), 187-195, doi: 10.1515/teme-2014-0007
- T. Baur et al. Novel method for the detection of short trace gas pulses with metal oxide semiconductor gas sensors, J. Sens. Sens. Syst. (2018) 7, 411-419
- C. Schultealbert et al, « Facile Quantification and Identification Techniques for Reducing Gases over a Wide Concentration Range Using a MOS Sensor in Temperature-Cycled Operation" MDPI Sensors (2018) 18:744 doi:; 10.3390/s18030744
- C. Schultealbert et al., « A novel approach towards calibrated measurement of trace gases using metal oxide semiconductor sensors" Sensors and Actuators B: Chemical (2017), 239, pp 390-396 doi: 10.1016/j.snb.2016.08.002
- C. Bur et al. Discrimination and quantification of volatile organic compounds in the ppb-range with gas sensitive SiC-FETs using multivariate statistics, Sensors and Actuators B: Chemical 214 (2015), doi: 10.1016/j.snb.2015.03.016
- C. Bur et al. „ Selectivity Enhancement of SiC-FET Gas Sensors by Combining Temperature and Gate Bias Cycled Operation Using Multivariate Statistics," Sensors and Actuators B: Chemical 193 (2014), pp. 931-940, doi: 10.1016/j.snb.2013.12.030
- C. Bur et al. "Increasing the Selectivity of Pt-Gate SiC Field Effect Gas Sensors by Dynamic Temperature Modulation", Increasing the Selectivity of Pt-Gate SiC Field Effect Gas Sensors by Dynamic Temperature Modulation," doi: 10.1109/JSEN.2011.2179645

Literature – Deep Learning

TCOCNN

- Y. Robin et al., " High-Performance VOC Quantification for IAQ Monitoring Using Advanced Sensor Systems and Deep Learning," Atmosphere (2021) 12(11), 1487, DOI 10.3390/atmos12111487

Transfer Learning

- Y. Robin et al. " Deep Learning Based Calibration Time Reduction for MOS Gas Sensors with Transfer Learning" Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614
- Y. Robin et al. "Transfer Learning to Significantly Reduce the Calibration Time of MOS Gas Sensors," ISOEN2022 - International Symposium on Olfaction and Electronic Nose, Aveiro, Portugal, doi: 10.1109/ISOEN54820.2022.9789596

Explainable AI

- Y. Robin et al. Insight in Dynamically Operated Gas Sensor Arrays with Shapley Values for Data Segments" MNE EUROSENSORS 2022, Poster T3-P2-WeA_0, Leuven, BE, Sep. 19-23. 2022

Literature – GMA and Inter-lab test

Gas Mixing Apparatus

- N. Helwig et al. „Gas mixing apparatus for automated gas sensor characterization ", Meas. Sci. Technol. 25 (2014) 055903, doi: 10.1088/0957-0233/25/5/055903
- M. Leidinger et al. "Characterization and calibration of gas sensor systems at ppb level - a versatile test gas generation system" IOP Measurement Science and Technology 29 (2018), 015901 (10pp), doi: 10.1088/1361-6501/aa91da
- D. Arendes et al. „Modular design of a gas mixing apparatus for complex trace gas mixtures "15. Dresdner Sensor-Symposium, 6. - 8. Dezember 2021, P13.1 - Modular design of a gas mixing apparatus for complex trace gas mixtures · AMA Science (ama-science.org)
- D. Arendes et al. "Qualification of a Gas Mixing Apparatus for Complex Trace Gas Mixtures" 16. Dresdner Sensor-Symposium, Posterbeitrag, https://www.ama-science.org/proceedings/details/4282

Validation

- T. Sauerwald, „Highly sensitive benzene detection with metal oxide semiconductor gas sensors - an inter-laboratory comparison ", J. Sens. Sens. Syst. (2018), 7, 235-243, doi: 10.5194/jsss-7-235-2018
- M. Bastuck et al., « Comparison of ppb-level gas measurements with a metal-oxide semiconductor gas sensor in two independent laboratories," Sensors and Actuators: B. Chemical (2018) 273, 1037-1046, doi: 10.1016/j.snb.2018.06.097
- L. Spinelle et al. "Review of Portable and Low-Cost Sensors for the Ambient Air Monitoring of Benzene and Other Volatile Organic Compounds, MDPI Sensors 2017, 17(7), 1520, DOI 10.3390/s17071520

Literature – Signal Processing

Lab for Measurement Technology

UNIVERSITÄT DES SAARLANDES

Signal Processing

- Manuel Bastuck, Dissertation, Saarland University and Linköping University, 2019.
  http://liu.diva-portal.org/smash/record.jsf?pid=diva2%3A1338901&dswid=-3162

T. Baur, et al. , " Random gas mixtures for efficient gas sensor calibration",J. Sens. Sens. Syst. (2020) 9, 411-424, DOI 10.5194/jsss-9-411-2020.

Toolbox

- M. Bastuck et al., "DAV³E - a MATLAB toolbox for multivariate sensor data evaluation," J. Sens. Sens. Syst. (2018), 7, 489-506, DOI: 10.5194/jsss-7-489-2018
- https://github.com/lmtUds/dav3e-beta
- T. Schneider et al., « Industrial condition monitoring with smart sensors using automated feature extraction and selection," IOP Meas. Sci. Technol. (2018) 29 094002, 2018, doi: 10.1088/1361-6501/aad1d4

Sensor-Hardware

- C. Fuchs et al. „Concept and realization of a modular and versatile platform for metal oxide semiconductor gas sensors" tm - Technisches Messen, vol. 89, no. 12, 2022, pp. 859-874, doi: 10.1515/teme-2022-0046

Literature – Indoor Air Quality

Indoor Air Quality

- A. Schütze et al. "Indoor Air Quality" n: Eduard Llobet (ed.): Advanced Nanomaterials for Inexpensive Gas Microsensors, Elsevier, 2020, Pages 209-234, ISBN: 978-0-12-814827-3, doi: 10.1016/B978-0-12-814827-3.00011-6
- T. Baur et al, " Field Study of Metal Oxide Semiconductor Gas Sensors in Temperature Cycled Operation for Selective VOC Monitoring in Indoor Air," Atmosphere (2021) 12, 647, DOI 10.3390/atmos12050647
- C. Schultealbert et al, "Measuring Hydrogen in Indoor Air with a Selective Metal Oxide Semiconductor Sensor" Atmosphere (2021) 12(3), 366, DOI 10.3390/atmos12030366
- M. Leidinger et al., "Selective detection of hazardous VOCs for indoor air quality applications using a virtual gas sensor array,"
- C. Bur et al. "Detecting Volatile Organic Compounds in the ppb Range with Gas Sensitive Platinum gate SiC-Field Effect Transistors," IEEE Sensors Journal 14 (9), pp. 3221 - 3228 (2014), doi: 10.1109/JSEN.2014.2326693
- J. Amann et al., 16. Dresdner Sensor-Symposium, 2022, https://www.ama-science.org/proceedings/details/4287
- J. Amann et al. Air Sensors International Conference, May 11-13, 2022, https://asic.aqrc.ucdavis.edu/2022-program-topics

Poisoning of MOS Sensors

- C. Schultealbert et al., " Siloxane treatment of metal oxide semiconductor gas sensors in temperature-cycled operation – sensitivity and selectivity" J. Sens. Sens. Syst. (2020) 9, 283-292, DOI 10.5194/jsss-9-283-2020
- Caroline Schultealbert, Dissertation, Saarland University, 2022. Shaker Verlag, 2022,ISBN: 978-3-8440-8443-6 https://publikationen.sulb.uni-saarland.de/handle/20.500.11880/32398
- C. Schultealbert et al. « Erkennung und Kompensation von Vergiftung durch Siloxane auf Halbleitergassensoren im temperaturzyklischen Betrieb (Identification and compensation of siloxane poisoning in metal oxide semiconductor gas sensors in temperature cycled operation), tm - Technisches Messen, 87(S1), S120–S125, doi: 10.1515/teme-2020-0041
- C. Schultealbert et al, « Siloxane treatment of metal oxide semiconductor gas sensors in temperature-cycled operation – sensitivity and selectivity," J. Sens. Sens. Syst. (2020) 9, 283-292, DOI 10.5194/jsss-9-283-2020

Literature – HW & Design of Exeriment

Sensor-Hardware

- C. Fuchs et al. „Concept and realization of a modular and versatile platform for metal oxide semiconductor gas sensors" tm - Technisches Messen, vol. 89, no. 12, 2022, pp. 859-874, doi: 10.1515/teme-2022-0046

Check list for machine learning with industrial data

- C. Schnur et al., „Checkliste – Mess- und Datenplanung für das maschinelle Lernen in der Montage doi: 10.5281/zenodo.6943476