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
Agenda

Application

Indoor Air Quality

www.renesas.com

fischerheating.com

Calibration

strategies

concentration

time

Transfer learning

Explainable AI

Oclusion sensitivity (miniature poodle)

Machine and Deep Learning

Calibration equipment

input layer

hidden layer 1

hidden layer 2

output layer
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
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."

CO₂ 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.
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 one 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)

Simulating the Complex Indoor Environment

Indoor environment is a complex mixture

- Around 200-300 different VOCs were reported
- \( \text{CO}_2, \text{CO}, \text{H}_2, \text{NO}_x \) plus humidity
- Influence of outdoor air (\( \text{O}_3, \text{NO}_x \ldots \))

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\text{Sens})

Testing sensor systems with complex mixtures of VOCs

- Representative environmental mixtures of the individual components
## VOC Representatives According to AGÖF, UBA

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

Alkene, Halocarbons, Glycols & Glycol ethers: negligible

AGÖF (2007): Supply of a data base about the occurrence of volatile organic compounds in indoor air
UBA (2010): German Environmental Survey on Children (GerES IV)
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
  - NO₂: annual mean (2020, close to traffic, UBA): 15 ppb
    Richtwert I: 42 ppb; Richtwert II: 140 ppb
  - N₂O: atmospheric concentration ca. 330 ppb
  - Ozone : O₃?
  - Further research needed
- Cyclic siloxanes (D3 – D6): very low concentrations, but relevant as sensor poison
  ⇒ Separate investigation

<table>
<thead>
<tr>
<th>Interference</th>
<th>Min conc.</th>
<th>Max. conc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>100 ppb</td>
<td>2000 ppb</td>
</tr>
<tr>
<td>Hydrogen</td>
<td>400 ppb</td>
<td>2000 ppb</td>
</tr>
<tr>
<td>humidity</td>
<td>20 %</td>
<td>75 %</td>
</tr>
</tbody>
</table>
MOS sensors in temperature cycled operation (TCO)

Virtual multi-sensor

➢ 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

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

Manuel Bastuck, Dissertation, Saarland University and Linköping University, 2019
Shaker Verlag, 2019.ISBN: 978-3-8440-7075-0
Data Evaluation FESR

**Data evaluation**

- Machine learning methods: FESR
  - Feature Extraction
  - Feature Selection, and
  - Regression
- Divide temperature cycle in 1 s ranges for feature extraction
- Partial Least Squares Regression (PLSR)

![Diagram of data evaluation process]

**Steps:**
- Raw Data
- Feature Extraction Mean & Slope
- Feature Selection RFE-LSR
- Quantification PLS regression
Gas Mixing Apparatus
for complex lab calibration
Standard Gas Mixing Apparatus

Carrier gas (zero air) → MFC 500 ml/min

MFC 500 ml/min → MFC 20 ml/min

Wash bottle at 20 °C

3/2-valve

Sensor chamber

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

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

Test gas 1 (gas cylinder) → MFC 20 ml/min

Test gas 2 (gas cylinder) → MFC 20 ml/min (1 … 20 ml/min)

Test gas 6 → waste

...
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\textsubscript{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: \( \frac{1}{500} \times \frac{1}{500} = \frac{1}{250,000} = 4 \times 10^{-6} \)

Min. dilution: \( \frac{20}{500} = \frac{1}{25} = 4 \times 10^{-2} \)

Dilution of test gas AND contaminations in the test gas bottle

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

### Diagram

- **Carrier gas** (zero air)
  - Carrier MFC 500 ml/min
    - (15 ... 500 ml/min)
  - Gas MFC 20 ml/min
    - (1 ... 20 ml/min)
  - Total flow 500 ml/min (or 1000 ml/min)
  - Pressure regulator 2 bar

- **Test gas** (gas cylinder)
  - Carrier MFC 500 ml/min
    - (15 ... 500 ml/min)
  - Gas MFC 20 ml/min
    - (1 ... 20 ml/min)
  - Injection MFC 20 ml/min
    - (1 ... 20 ml/min)
  - Sensor chamber

**Dynamic range over 4 magnitudes**

With the same test gas cylinder
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)
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

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)
  ⇒ 10.240 gas exposures
  Assume 10 min per exposure with 10 min pause in between
  ⇒ **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)
Comparison Sequential vs. Random Calibration

Sequential setup:
- 3 levels of RH
- 6 gases
- 4 conc. per test gas
- CO and H₂ at atmospheric conc.
  \[\Rightarrow 72 \text{ gas exposures}\]

Randomized setup:
- Define concentration ranges
- For comparison:
  Limit concentration range to fit to the sequential measurement

<table>
<thead>
<tr>
<th>Gas</th>
<th>Concentration (ppb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>acetone</td>
<td>250 500 750 1000</td>
</tr>
<tr>
<td>benzene</td>
<td>250 500 750 1000</td>
</tr>
<tr>
<td>carbon monoxide</td>
<td>150 300 450 600</td>
</tr>
<tr>
<td>formaldehyde</td>
<td>40 80 120 160</td>
</tr>
<tr>
<td>hydrogen</td>
<td>500 750 1000 1250</td>
</tr>
<tr>
<td>toluene</td>
<td>5 25 45 65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Concentration range</th>
</tr>
</thead>
<tbody>
<tr>
<td>hydrogen</td>
</tr>
<tr>
<td>carbon monoxide</td>
</tr>
<tr>
<td>humidity</td>
</tr>
<tr>
<td>(\text{VOC}_{\text{sum}}) in (\mu\text{g/m}^3)</td>
</tr>
<tr>
<td>(\text{VOC}_{\text{sum}}) in ppb</td>
</tr>
<tr>
<td>acetone</td>
</tr>
<tr>
<td>benzene</td>
</tr>
<tr>
<td>formaldehyde</td>
</tr>
<tr>
<td>toluene</td>
</tr>
</tbody>
</table>

Comparison Sequential vs. Random Calibration

Sequential training always leads to higher RMSE values.

- (α) complete measurement
- (β) randomized train, randomized test
- (γ) randomized train, sequential test
- (δ) sequential train, randomized test

## Comparison Sequential vs. Random Calibration

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>random</td>
<td>random</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>sequential</td>
</tr>
<tr>
<td></td>
<td>sequential</td>
<td>random</td>
</tr>
</tbody>
</table>

### Graphs:

#### (a)
- **Predicted VOC$_{sum}$ in ppb vs. VOC$_{sum}$ in ppb for Train:**
  - RMSECV
  - Test dataset

#### (b)
- **Predicted VOC$_{sum}$ in ppb vs. VOC$_{sum}$ in ppb for Test:**
  - RMSECV
  - Test dataset

#### (c)
- **Predicted VOC$_{sum}$ in ppb vs. VOC$_{sum}$ in ppb for Test:**
  - RMSECV
  - Test dataset
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**
# Feature Table

4 x 2x144 features

<table>
<thead>
<tr>
<th>sensor 0</th>
<th>sensor 1</th>
<th>sensor 2</th>
<th>sensor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 F2 F3 ... F288</td>
<td>F1 F2 F3 ... F288</td>
<td>F1 F2 F3 ... F288</td>
<td>F1 F2 F3 ... F288</td>
</tr>
</tbody>
</table>

Concentrations of gases

<table>
<thead>
<tr>
<th>annotation</th>
<th>gas 1</th>
<th>annotation</th>
<th>gas 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>gas 1</td>
<td>20 ppb</td>
<td>gas 2</td>
<td>75 ppb</td>
</tr>
<tr>
<td>gas 2</td>
<td>20 ppb</td>
<td>gas 3</td>
<td>75 ppb</td>
</tr>
<tr>
<td>...</td>
<td>20 ppb</td>
<td>...</td>
<td>75 ppb</td>
</tr>
</tbody>
</table>

UGM: unique gas mixture

900 UGM with 10 T-cycles each
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!)

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**
Standard Machine Learning Approach

sensor → Temperature modulation → raw data → Pre-processing

Feature extraction → Standardization → Feature pre-selection

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

Sensor array (virtual)

ISOCS Winter School 2023 – Advanced Calibration Strategies for IAQ Sensors

17.01.2023
Results Partial Least Squares Regression (PLSR)

Aceton

RMSE_{Validation} = 10.4 ppb  
RMSE_{Testing} = 12.7 ppb

TVOC_{Sens}

RMSE_{Validation} = 27.9 ppb  
RMSE_{Testing} = 30.5 ppb

Convolution Neural Networks – Image Processing

The MathWorks Inc.
Famous Convolution Neural Networks

ImageNet

Annual competition
1.2 Mio images
1000 categories

ImageNet
https://www.image-net.org/
Deep Learning Approach for Gas Sensing

(sensor) (virtual) sensor array

Temperature modulation

raw data

Pre-processing

Convolution Neural Network

Feature extraction

regression

Hyper parameter optimization

Cross-validation

testing
Designing a TCO-CNN

10-layer deep convolutional neural network

Input layer:
[number of gas sensitive layers] × [sample points per cycle]
e.g. SGP40, Sensirion, Switzerland: 4 × 1440

Output layer:
Regression layer with mean squared error as loss function

<table>
<thead>
<tr>
<th>Layer #</th>
<th>Filter Size</th>
<th>Striding</th>
<th># Filter</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DConv #0</td>
<td>1 x 96</td>
<td>1 x 15</td>
<td>153</td>
<td>4 x 80 x 153</td>
</tr>
<tr>
<td>2DConv #1</td>
<td>1 x 1</td>
<td>1 x 1</td>
<td>153</td>
<td>4 x 80 x 153</td>
</tr>
<tr>
<td>2DConv #2</td>
<td>1 x 2</td>
<td>1 x 2</td>
<td>153</td>
<td>4 x 40 x 153</td>
</tr>
<tr>
<td>2DConv #3</td>
<td>1 x 1</td>
<td>1 x 1</td>
<td>153</td>
<td>4 x 40 x 153</td>
</tr>
<tr>
<td>2DConv #4</td>
<td>1 x 2</td>
<td>1 x 2</td>
<td>306</td>
<td>4 x 20 x 306</td>
</tr>
<tr>
<td>2DConv #5</td>
<td>1 x 1</td>
<td>1 x 1</td>
<td>153</td>
<td>4 x 20 x 306</td>
</tr>
<tr>
<td>2DConv #6</td>
<td>1 x 2</td>
<td>1 x 2</td>
<td>459</td>
<td>4 x 10 x 459</td>
</tr>
<tr>
<td>2DConv #7</td>
<td>1 x 1</td>
<td>1 x 1</td>
<td>459</td>
<td>4 x 10 x 459</td>
</tr>
<tr>
<td>2DConv #8</td>
<td>1 x 2</td>
<td>1 x 2</td>
<td>612</td>
<td>4 x 5 x 612</td>
</tr>
<tr>
<td>2DConv #9</td>
<td>1 x 1</td>
<td>1 x 1</td>
<td>612</td>
<td>4 x 5 x 612</td>
</tr>
<tr>
<td>Fully Connected #1</td>
<td>1 x 12240</td>
<td>1 x 12240</td>
<td>1280</td>
<td>1 x 1280</td>
</tr>
<tr>
<td>Fully Connected #2</td>
<td>1 x 1280</td>
<td>1 x 1280</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Neural Architecture Search (NAS)

Hyperparameter optimization

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

Bayesian optimization search to find smallest RMSE

<table>
<thead>
<tr>
<th>Initial Learning Rate (Log Scale)</th>
<th>Number of Filters (First Two Layers)</th>
<th>Kernel Size (First Two Layers)</th>
<th>Stride Size (First Layer)</th>
<th>Dropout</th>
<th>Number of Neurons (FC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 \times 10^{-4} - 9 \times 10^{-3}$</td>
<td>60–240</td>
<td>40–80</td>
<td>15–45</td>
<td>30–50%</td>
<td>1000–2500</td>
</tr>
</tbody>
</table>

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)

Results are given for TCOCNN.
Comparison FESR and TCOCNN

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

<table>
<thead>
<tr>
<th>Gas</th>
<th>RMSE FESR (ppb)</th>
<th>RMSE TCOCNN (ppb)</th>
<th>Standard deviation TCOCNN (ppb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetone</td>
<td>13.6</td>
<td>11.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Toluene</td>
<td>25.5</td>
<td>18.5</td>
<td>1.3</td>
</tr>
<tr>
<td>Ethanol</td>
<td>29.6</td>
<td>28.2</td>
<td>0.9</td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>31.3</td>
<td>15.4</td>
<td>0.5</td>
</tr>
<tr>
<td>VOC&lt;sub&gt;sum&lt;/sub&gt;</td>
<td>32.5</td>
<td>25.3</td>
<td>1.0</td>
</tr>
<tr>
<td>Hydrogen</td>
<td>37.0</td>
<td>33.2</td>
<td>3.9</td>
</tr>
<tr>
<td>Carbon monoxide</td>
<td>83.6</td>
<td>75.1</td>
<td>3.1</td>
</tr>
</tbody>
</table>
Results Field Tests

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

- **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

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

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

The MathWorks Inc.
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, Portugal, doi: 10.1109/ISOEN54820.2022.9789596, Best Paper Award
Extended Dataset

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 from same batch
  - Sensor B different batch

Y. Robin et al., Atmosphere 2022, 13(10), 1614, doi: 10.3390/atmos13101614
Y. Robin et al., ISOEN 2022, Aveiro, Portugal, doi: 10.1109/ISOEN54820.2022.9789596, Best Paper Award
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 $\sim25$ ppb
Transfer Learning TCOCNN

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
Results Transfer Learning

**Individual training** on 700 UGMs
- For all three sensors RMSE $\sim 25$ ppb

**Without transfer learning**
Simply apply model from Sensor A
- on test data Sensor B: $\sim 74$ ppb
- on test data Sensor C: $\sim 103$ ppb

**Transfer learning**
Use trained model from Sensor A plus
- 20 UGMs (i.e. 3 %) Sensor B/C: $\sim 47$ ppb / 55 ppb
- 100 UGMs (i.e. 14 %) Sensor B/C: $\sim 40$ 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?

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

**Feature extraction**

**Regression**

![Graphs showing RMSE in ppb vs. number of unique gas mixtures in training set for Hydrogen Sensor B and Xylene Sensor B.](image)
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

![Graphs showing comparison between Global Approach and Transfer Learning for Acetone](image)

- Transfer learning outperforms a global approach
- More sensors better results
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

---

The MathWorks Inc.
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

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Training data set</th>
<th>Mean RMSE in ppb ± standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>All data</td>
<td>15.8 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>w/o most important 7 %</td>
<td>23.8 ± 1.0</td>
</tr>
<tr>
<td></td>
<td>Only most important 7 %</td>
<td>19.3 ± 1.0</td>
</tr>
<tr>
<td>B</td>
<td>All data</td>
<td>18.8 ± 0.6</td>
</tr>
<tr>
<td></td>
<td>w/o most important 7 %</td>
<td>26.3 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>Only most important 7 %</td>
<td>19.9 ± 1.0</td>
</tr>
</tbody>
</table>

Y. Robin et al., submitted to I2MTC 2023
Conclusion
Take Home Messages

✓ 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
Temperature Cycled Operation (TCO)

Literature – Deep Learning

TCOCNN

Transfer Learning

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


Validation

Literature – Signal Processing

**Signal Processing**


**Toolbox**


  [https://github.com/lmtUds/dav3e-beta](https://github.com/lmtUds/dav3e-beta)


**Sensor-Hardware**

Literature – Indoor Air Quality

Indoor Air Quality


• M. Leidinger et al., “Selective detection of hazardous VOCs for indoor air quality applications using a virtual gas sensor array,”


Poisoning of MOS Sensors


• 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

Sensor-Hardware


Check list for machine learning with industrial data