SCAN: Multi-Hop Calibration for Mobile Sensor Arrays

Balz Maag, Zimu Zhou, Olga Saukh, Lothar Thiele

Computer Engineering and Networks Laboratory
ETH Zurich, Switzerland

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

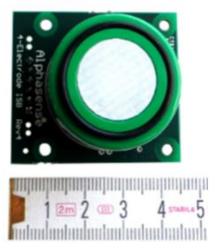
Mobile Air Pollution Monitoring

- Mobile air quality sensors collect vast amounts of measurements with high resolution

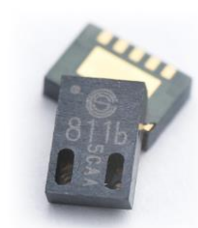
Air pollution measurements collected by Google Streetview cars in Oakland, CA

- Low cost air pollution sensors

SGX Sensortech MiCS-OZ-47 O_3



AlphaSense CO-B4



ams CCS811

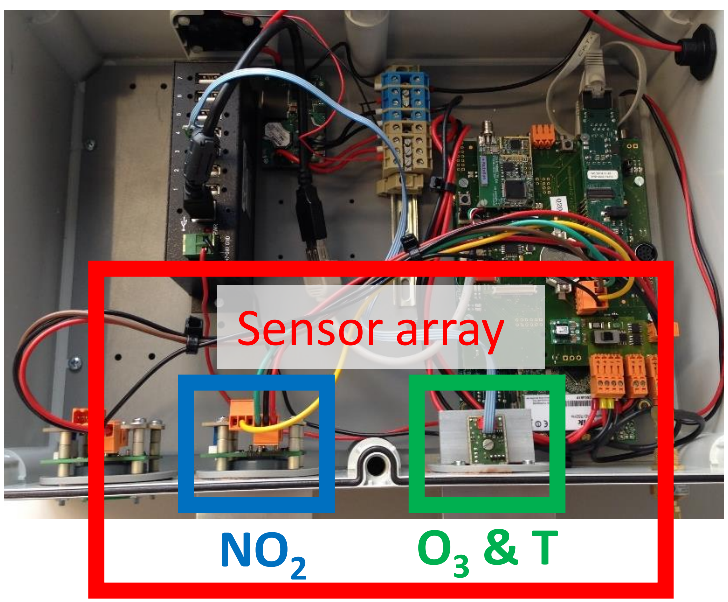
Honeywell Particle Sensor

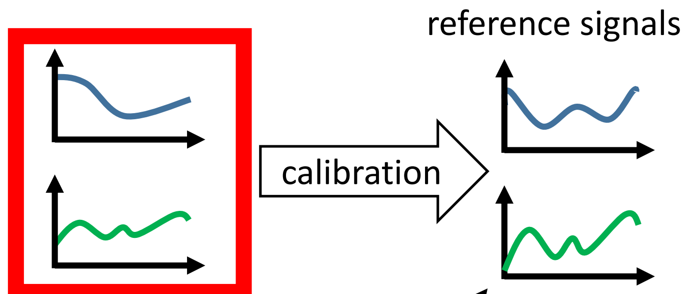
Figaro TGS8100

Reliable long-term monitoring with mobile low-cost sensors is challenging!

Sensor Array Calibration

- **Problem 1**: Low-cost sensors suffer from **low selectivity** and dependencies on meteorological conditions
- **Solution**: Sensor array calibration
  - Compensate for cross-sensitivities by augmenting multiple low-cost sensor and collectively calibrate the signals to references
- State-of-the-art: Multiple Least Squares (MLS) based regression





calibration matrix → $$B \cdot X = Y + N$$ ← calibration error

Real-world example: A sensor array reduces the NO_2 measurement error by a factor of 2 compared to a single NO_2 sensor

Multi-Hop Calibration

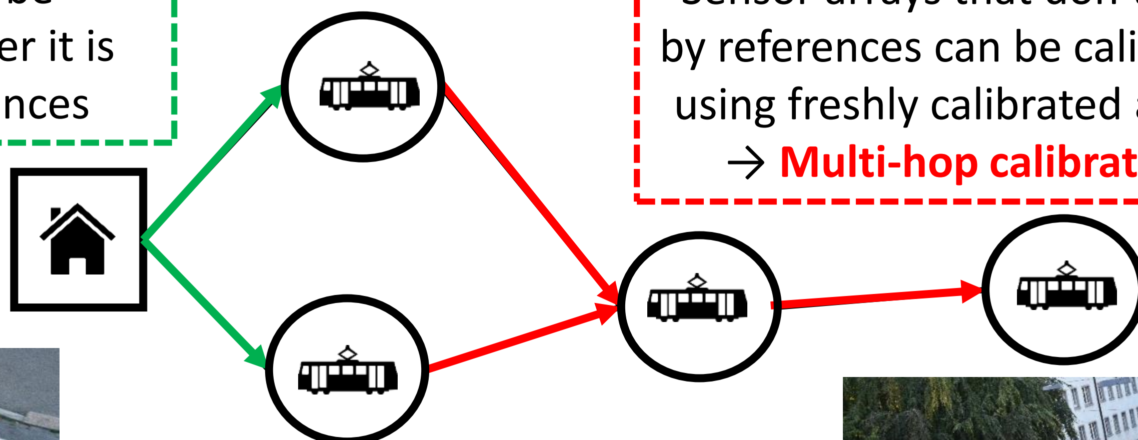
- **Problem 2**: Due to ageing effects and changing environmental conditions sensor arrays need **frequent recalibration**
- **Solution:** Opportunistic calibration during sensor rendez-vous

A sensor array can be calibrated whenever it is in vicinity of references

Sensor array

Sensor arrays that don't pass by references can be calibrated using freshly calibrated arrays
→ **Multi-hop calibration**

High quality references

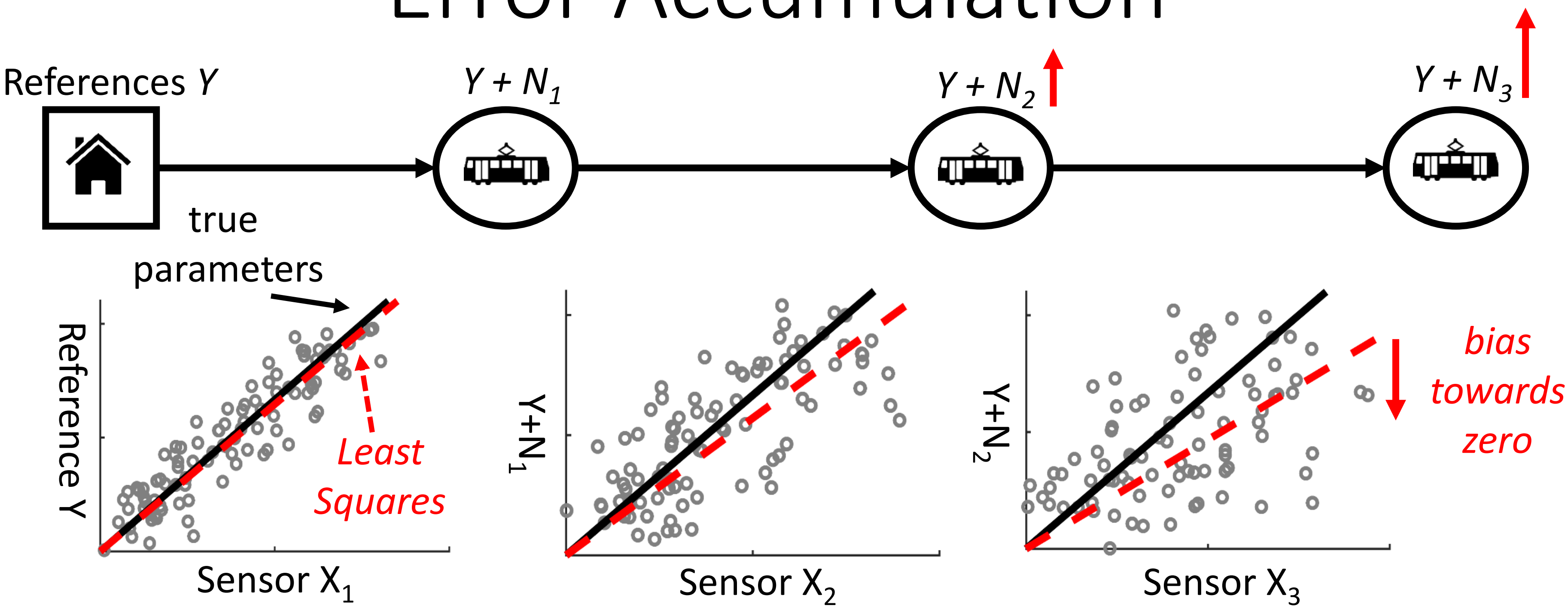


Sensor array on top of streetcar

Rendez-vous between array and references

Rendez-vous between two arrays

Error Accumulation



- Least Squares based regression suffers from error accumulation over multiple hops
  - Due to *regression dilution*, also known as *bias towards zero*
- Saukh et al. [IPSN15]: Geometric Mean Regression (GMR) minimizes error accumulation
- **Problem**: GMR is a single-variable model
  - Not applicable for sensor arrays
  - Poor performance when used to calibrate a cross-sensitive sensor

Our Solution

	One-hop calibration	Multi-hop calibration
Sensor array calibration	Multiple Least Squares	SCAN
Single sensor calibration	Ordinary Least Squares	Geometric Mean Regression

- SCAN: Sensor Array Network Calibration
  - Multi-variable regression model applicable to typical **sensor arrays**
  - Minimizes error accumulation when applied to **multi-hop calibration**

SCAN: Sensor Array Network Calibration

$$\min_{B} tr\left((Y - B \cdot X)(Y - B \cdot X)^T\right)$$
$$\text{subject to} \quad BXX^TB^T = YY^T$$

- SCAN minimizes the **least squares error** with a novel **constraint** on the parameters B
- A **unique and closed-form** expression of the calibration matrix B always exists
- For single sensor calibration, i.e. a single variable regression problem, SCAN and GMR yield the same solution

Calibration Matrix B

- Applicable to calibration matrizes *B* with **general form**
  - For instance, there are not always the same set of references signals available as low-cost sensor signals:

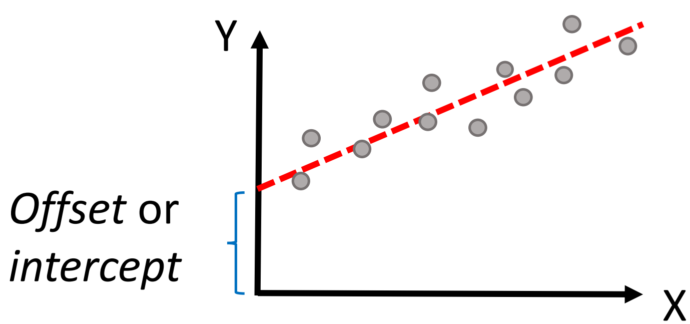
$\in \mathbb{R}^{2 \times 3}$

$$B \cdot X = Y + N$$

NO_2, O_3 and T

NO_2 and O_3

- SCAN can calculate an offset regression term similar to MLS



No Bias Towards Zero

- Contraint is able to completely remove error accumulation if the error components *N* of both arrays exhibit the same covariance matrix

$$B\overbrace{(XX^T+\underbrace{\boxed{NN^T}})}^{\text{Uncalibrated array}}B^T = \overbrace{YY^T+\underbrace{\boxed{NN^T}}}^{\text{Calibrated array}}$$

compensated

  Although in reality the two covariance matrices are usually similar but not perfectly identical, SCAN is still minimizing error accumulation

- More details and proofs are in the paper

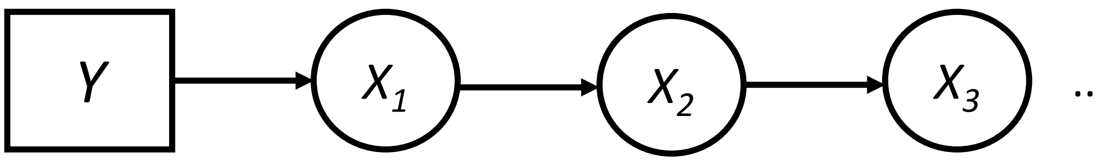
Evaluation: Simulation

- We artificially generate sensor arrays X_i with

$$X_i = B^{-1}(Y_i + N_i)$$

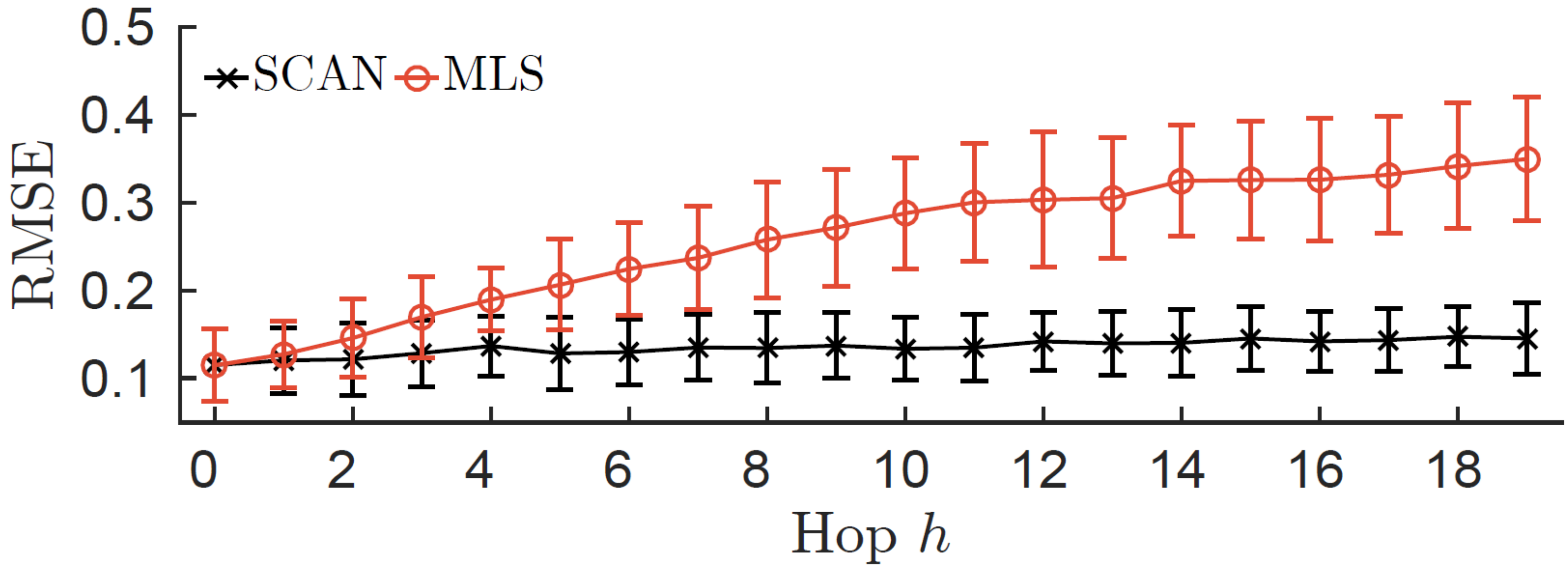
typical calibration matrizes *4 reference signals* *typical error components*

- Sensor arrays X_i form a calibration path



  - X_1 is calibrated to reference Y
  - X_2 to virtual reference provided by calibrated array X_1
- 500 samples for training, 500 samples for testing

Results: Simulation



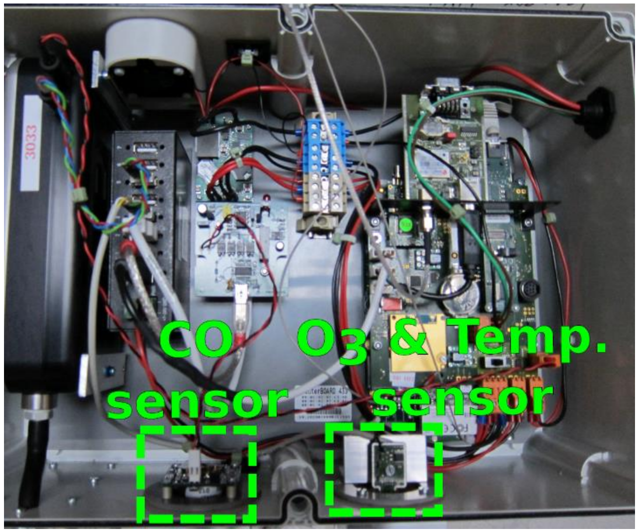
- Over a 20 hop calibration path SCAN clearly outperforms multiple least-squares
- MLS only performs better in the first hop

Comparison with Other Methods

Method	Hop = 0	Hop = 5	Hop = 19
SCAN	0.12	**0.12**	**0.14**
Multiple least squares (MLS)	**0.11**	0.21	0.34
Geometric mean regression (GMR)	0.4	0.54	0.54
Total least squares (TLS)	0.14	0.62	4.2
Neural Networks	0.19	>100	>100
GMR Generalization by Draper	0.13	0.24	0.34
GMR Generalization by Tofallis	0.14	0.25	0.35

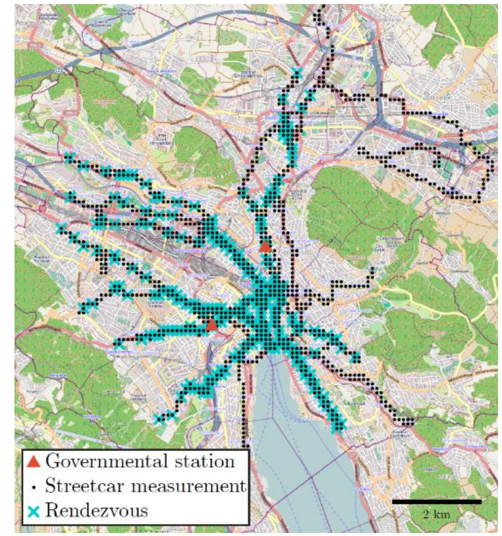
- As expected GMR performs poorly when applied to cross-sensitive sensors
- Over multiple hops SCAN outperforms all other methods

Evaluation: Streetcar Deployment



Sensors

Box on streetcar



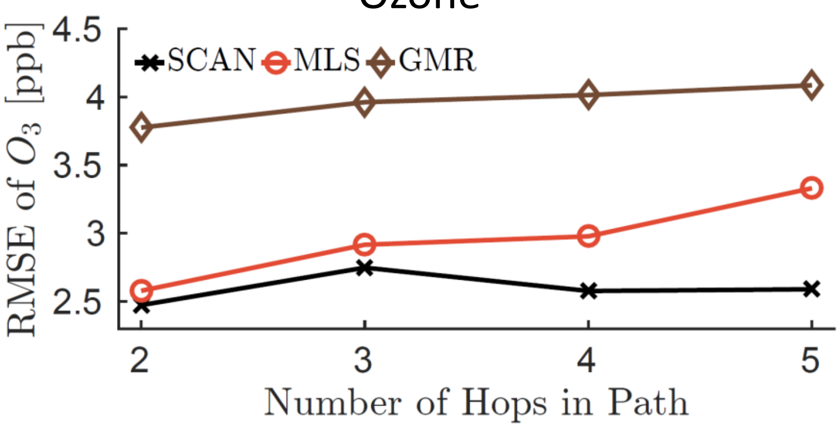
Measurement coverage

- 7 sensor arrays deployed on streetcars in the city of Zürich, Switzerland
  - Ozone (O3), carbon monoxide (CO) and temperature sensor
  - Gas sensors suffer from substantial meteorological dependencies
- Reference measurements provided by two governmental stations
- 56 Mio. measurements recorded between 03/2014 and 03/2016
  - Trained on at least 200 within 4 weeks
  - Tested on data from consecutive 4 weeks

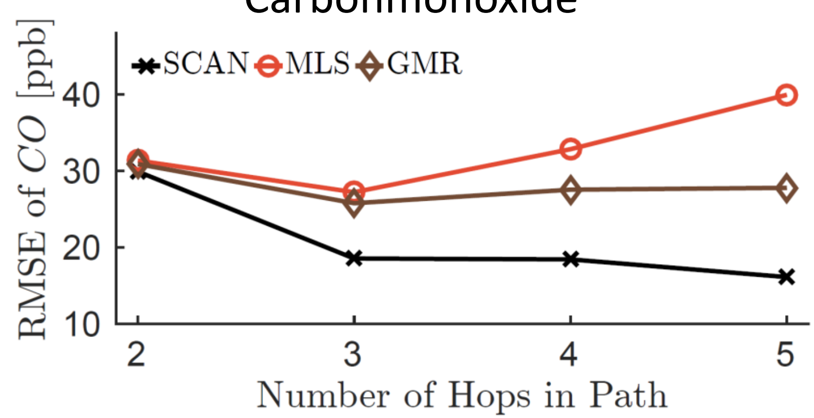
Results: Streetcar Deployment

- Network-wide calibration
  - At most three streetcars pass by reference stations
  - Multi-hop calibration allows to calibrate all seven streetcars in 94% of all months
- SCAN outperforms GMR by up to 42% and MLS by up to 60% over 5 hops

Ozone



Carbonmonoxide



Conclusion

- Multi-hop sensor array calibration is a powerful tool to provide reliable data over long time periods in mobile low-cost sensor deployments
- We propose SCAN, novel constrained multi-variable linear regression technique that
  - calibrates sensor arrays and
  - reduces error accumulation over multiple hops
- Evaluation based on two different datasets show that SCAN is able to outperform different calibration techniques

Thank you!
Questions?

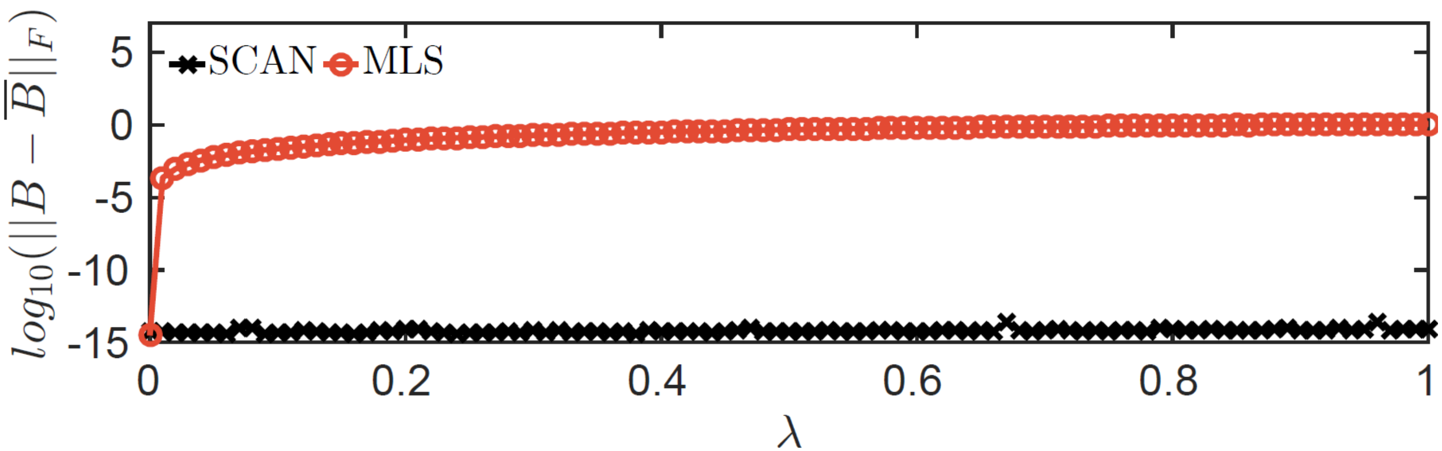
Streetcar data is publicly available:

www.opensense.ethz.ch

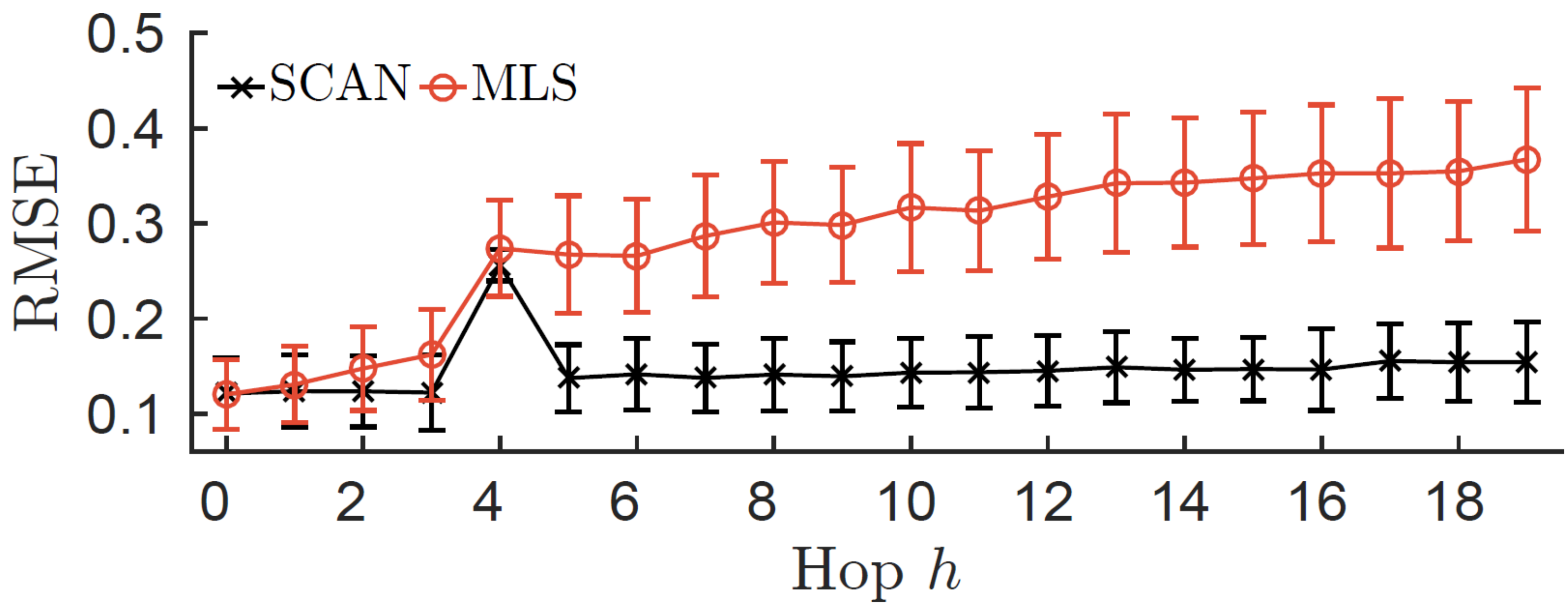
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Backup: No Bias Towards Zero

- If calibration error is uncorrelated, i.e. $NN^T = \sigma I$ SCAN is able to recover the true underlying calibration matrix B
- Simulation:

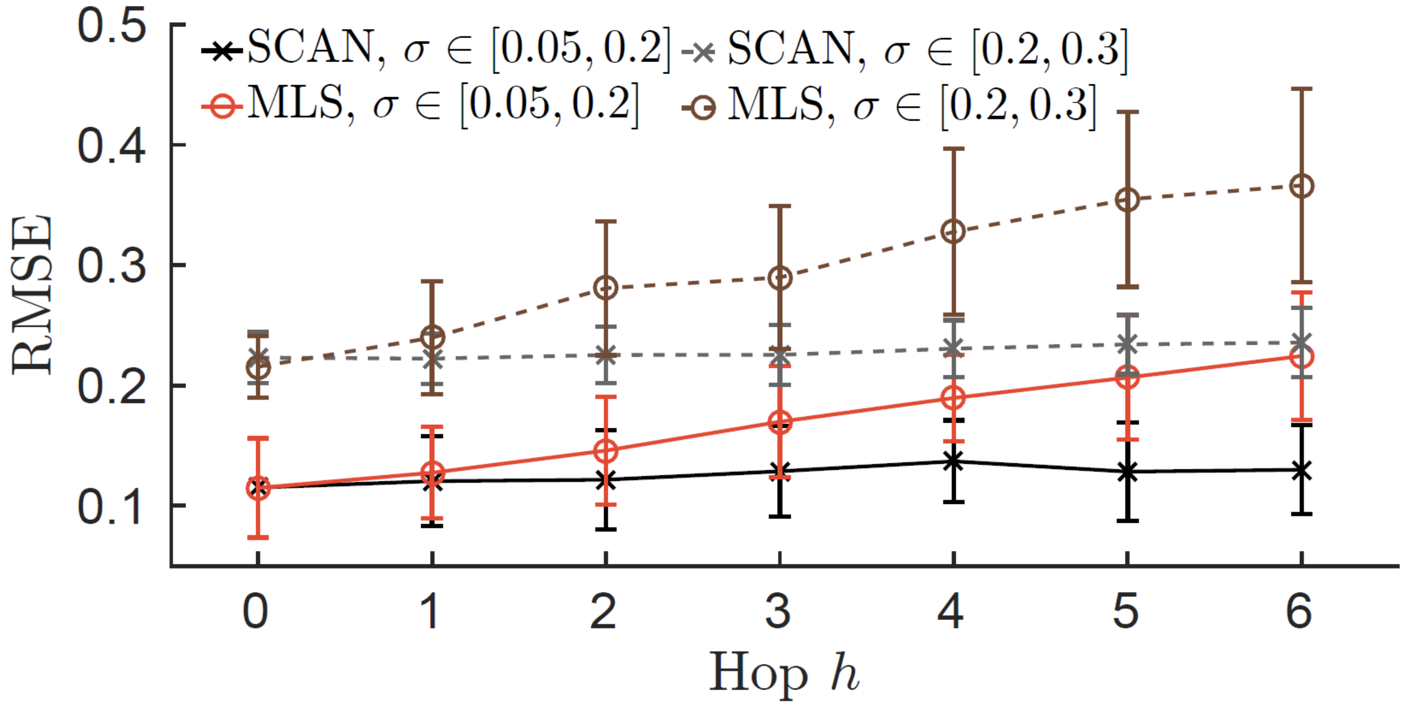


Backup: Increased noise variance



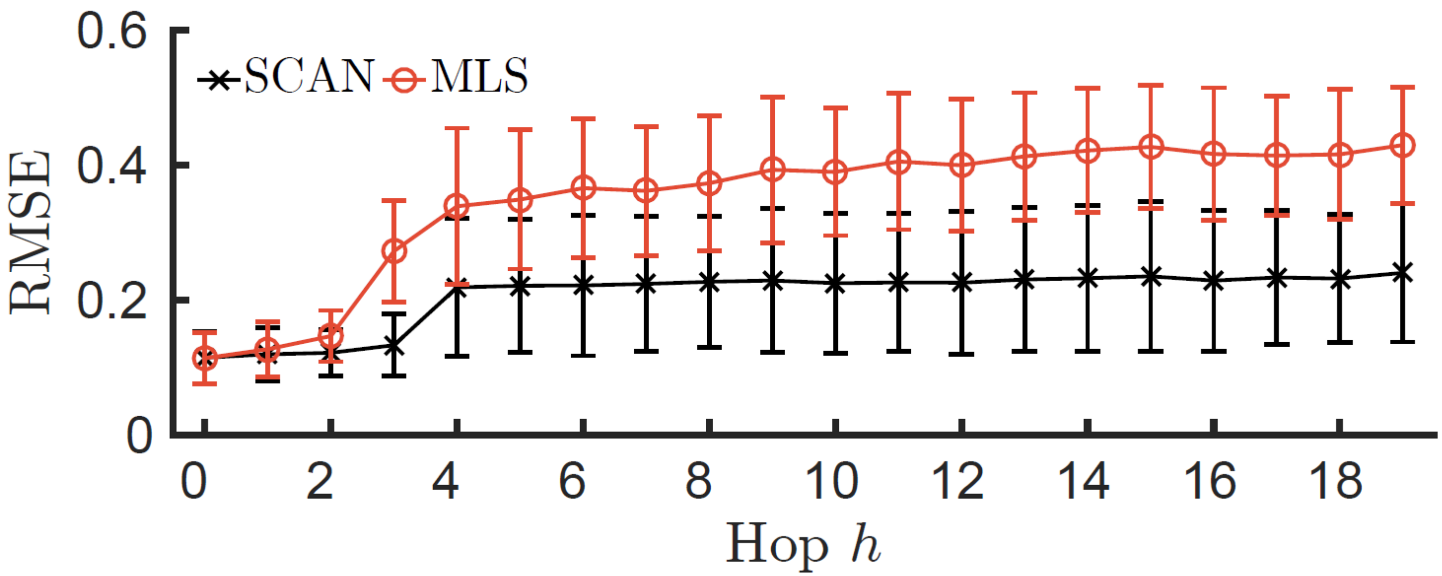
Increased noise in measurements (2x standard deviation) of node 4 has no effect on nodes after hop 5

Backup: Noise Independence



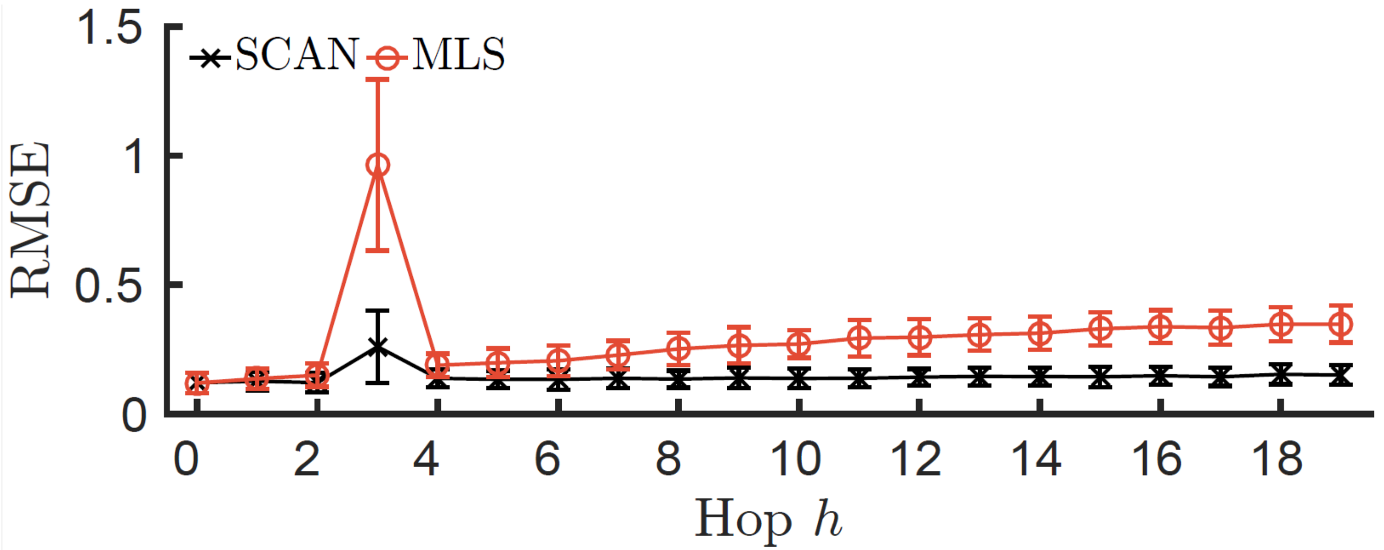
Increased noise in measurements has only an effect on the average calibration error but not on the error accumulation

Backup: Measurement range 1/2



- 3x smaller measurement range of rendezvous between array 3 and 4
- Measurement range = interval defined by the smallest and highest absolute value of the phenomena measurements
- General problem for any method due to missing training data

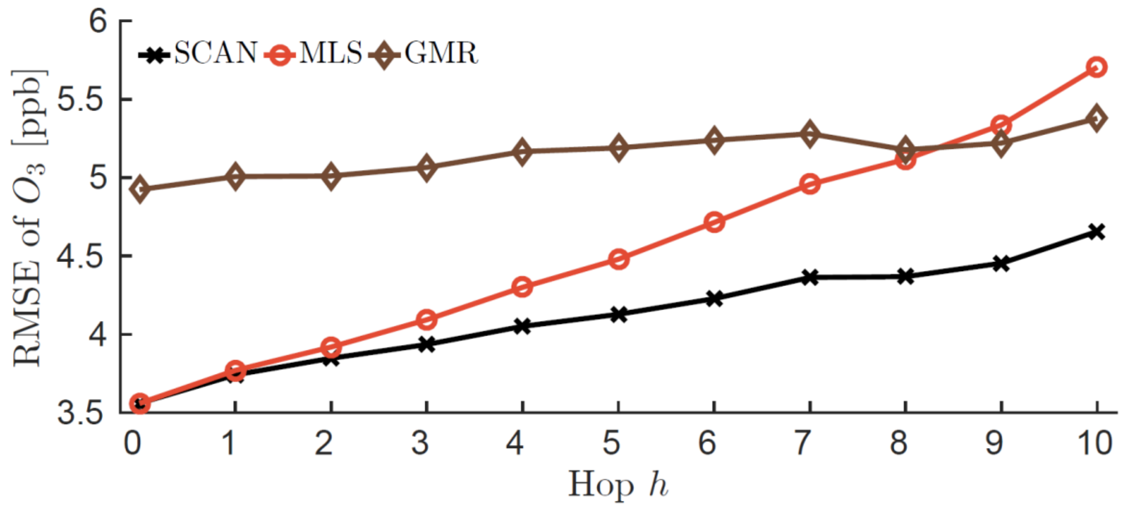
Backup: Measurement Range 2/2



- 3x higher measurement range of rendezvous between array 3 and 4
- Only affects array 4, because calibration paramters are only vaild for the measurement range of all other arrays

Backup: Metaloxide Sensors

- Calibration path with 11 sensor arrays consisting of
  - 2 different metaloxide low-cost gas sensors
  - 1 temperature sensor
- Deployed next to high quality reference station in Switzerland
- Calibrated to ozone (O3) provided by reference station
- 10 Mio. samples recorded between 07/2015 and 07/2016
  - 200 samples for training within two weeks
  - 200 samples for testing within consecutive two weeks



Over all hops SCAN achieves a 16% to 38% smaller error than GMR and an up to 23% smaller error than MLS

Backup: Typical Calibration Error

- Based on metal-oxide sensors
- Variance of calibration error shows strong similarity to diagonal matrix