Mixed Signals at Large Scales: IoT Sensors for Air Quality Monitoring and Forecasting

IEEE IoT Summit on Sustainable Sensor Systems for IoT

Olga Saukh

Institute of Technical Informatics, TU Graz
Complexity Science Hub Vienna

saukh@tugraz.at

Air Pollution

"Globally air pollution caused an estimated extra 8.8 million premature deaths a year in 2015"

R. Burnett et al., Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter, Academy of Sciences, 2018

"PM2.5 can easily be absorbed by the lungs and prolonged exposure can lead to respiratory impairments, blood diseases and neurodevelopmental disorders"

D. Cory-Slechta et al., The impact of inhaled ambient ultrafine particulate matter on developing brain, Toxicologic Pathology, 2019

"Long-term exposure to air pollutants such as PM2.5 and NO2 increases human susceptibility to SARS-CoV-2 and contributes to higher fatality rates"

Y. Han et al., The effects of outdoor air pollution concentrations and lockdowns on COVID-19 infections in Wuhan and other provincial capitals in China, 2020

Zivin JG, Neidell M. Air pollution's hidden impacts. Science, 2018

ENVIRONMENTAL ECONOMICS

Air pollution's hidden impacts

Exposure can affect labor productivity and human capital

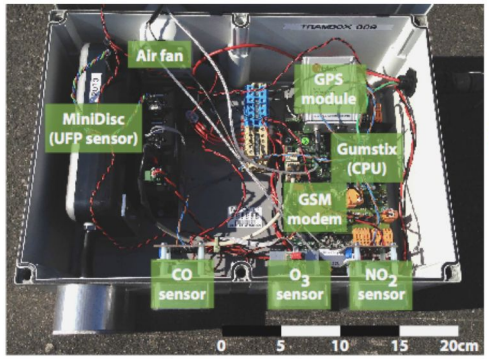
By **Joshua Graff Zivin**[1] *and* **Matthew Neidell**[2]

Nearly every country in the world regulates air pollution. But how much pol-

Our understanding of these less-severe health impacts is hampered by two factors. First, the number of potential channels through which pollution could affect human functioning is large, as pollution can

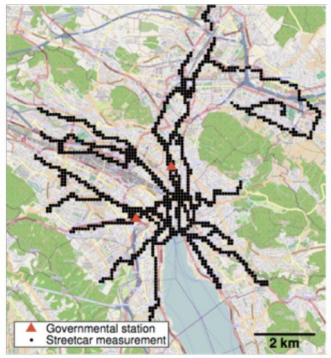
If you can't measure it you can't manage it © P. Drucker

Background

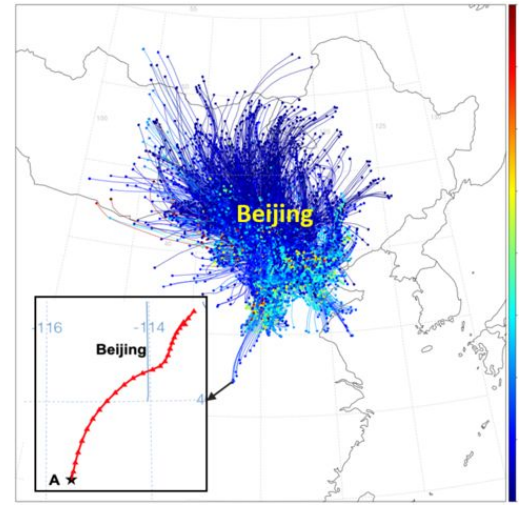


Sensor station

On top of a tram



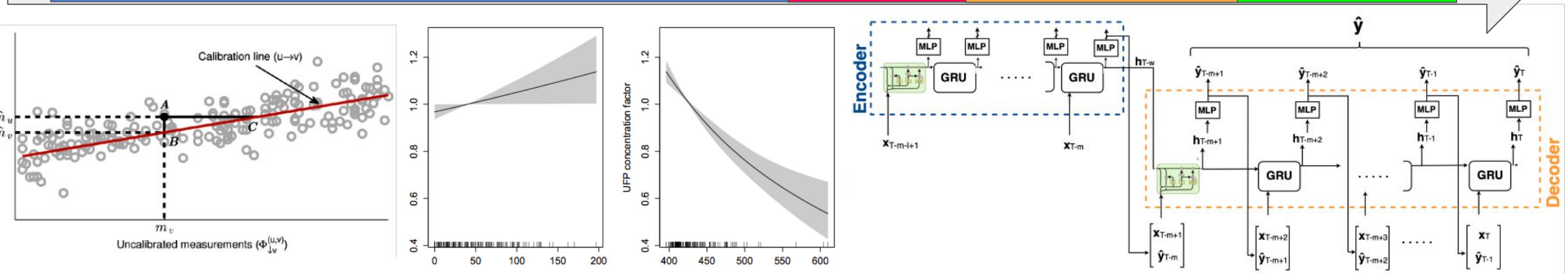
Measurement coverage



OpenSense: Air pollution sensors and high-resolution models

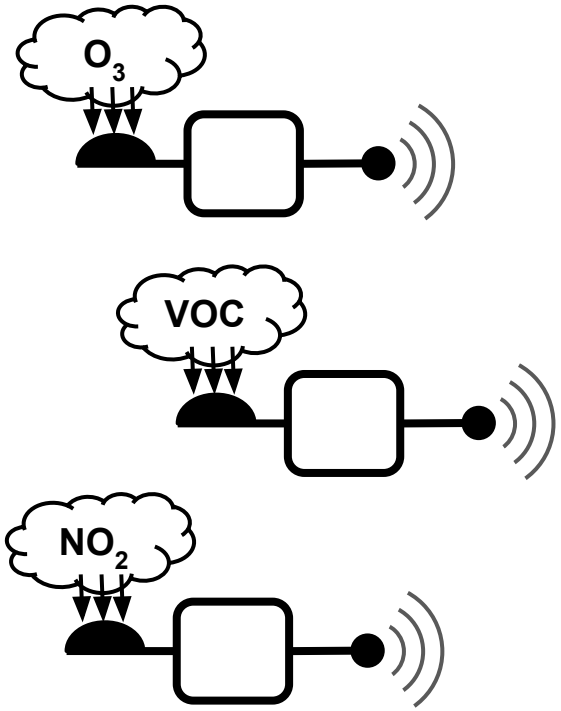
Sensirion AG

Gas sensors and models revisited

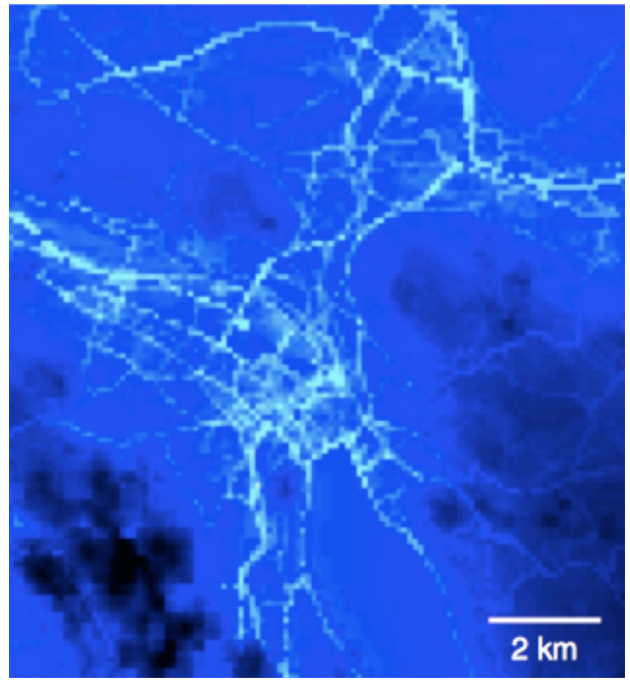


D. Hasenfratz et al., Deriving High-Resolution Urban Air Pollution Maps Using Mobile Sensor Nodes, J PMC 2015
O. Saukh et al., Reducing Multi-Hop Calibration Errors in Mobile Sensor Networks, IPSN 2015

IoT Sensors → Predictions



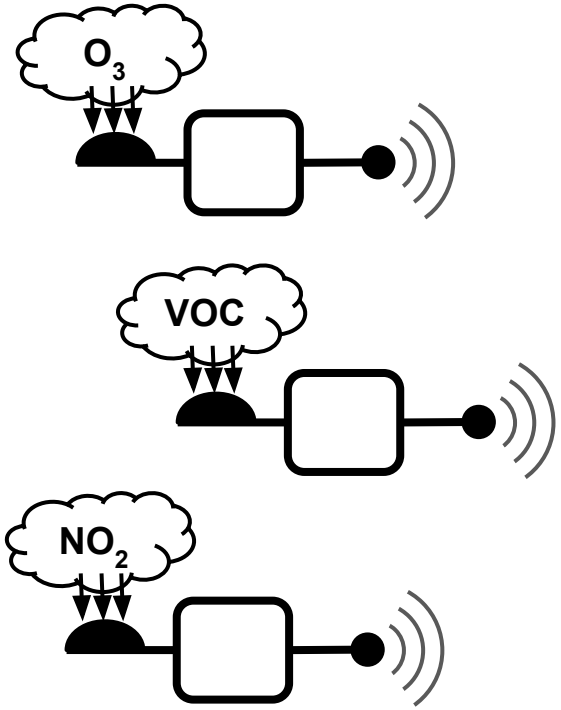
1. "Are measured data relevant?"
   - ⇒ Depends on sensor characteristics
   - ⇒ Data quality / calibration
2. "What can we learn from the data?"
   - ⇒ Spatiotemporal predictions



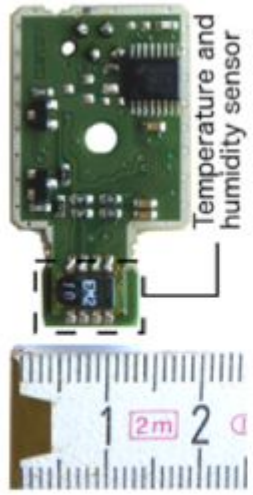
Energy consumption
Portability
Real-timiness
Cross-sensitivity
Response time
Accuracy
Robustness
Lifetime / Stability
Resolution
Short- / long-term
Size / Cost

1 “Are measured data relevant?” → Weak signals

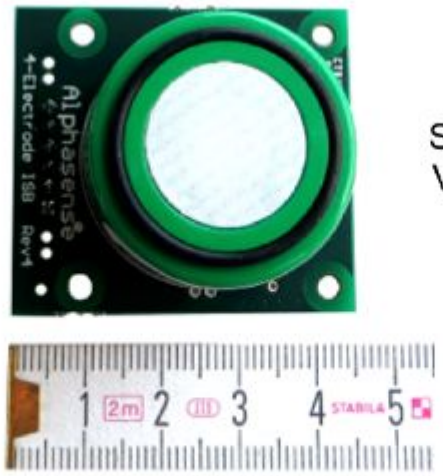
→ Monitoring air quality = chemical composition of the ambient air



SGX MiCS-OZ-47
ozone sensor



Temperature and
humidity sensors



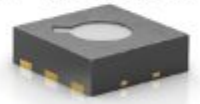
Alphasense CO-B4
carbon monoxide sensor

SGX MICS-5524 NH3 / H2 /
VOC / CO / C2H6OH / CH4

Alphasense NO2-B4
nitrogen dioxide sensor

AMS CCS811 VOC

Bosch BME680
T / RH / Pressure / VOC



Sensirion SGP30
VOC / eCO2

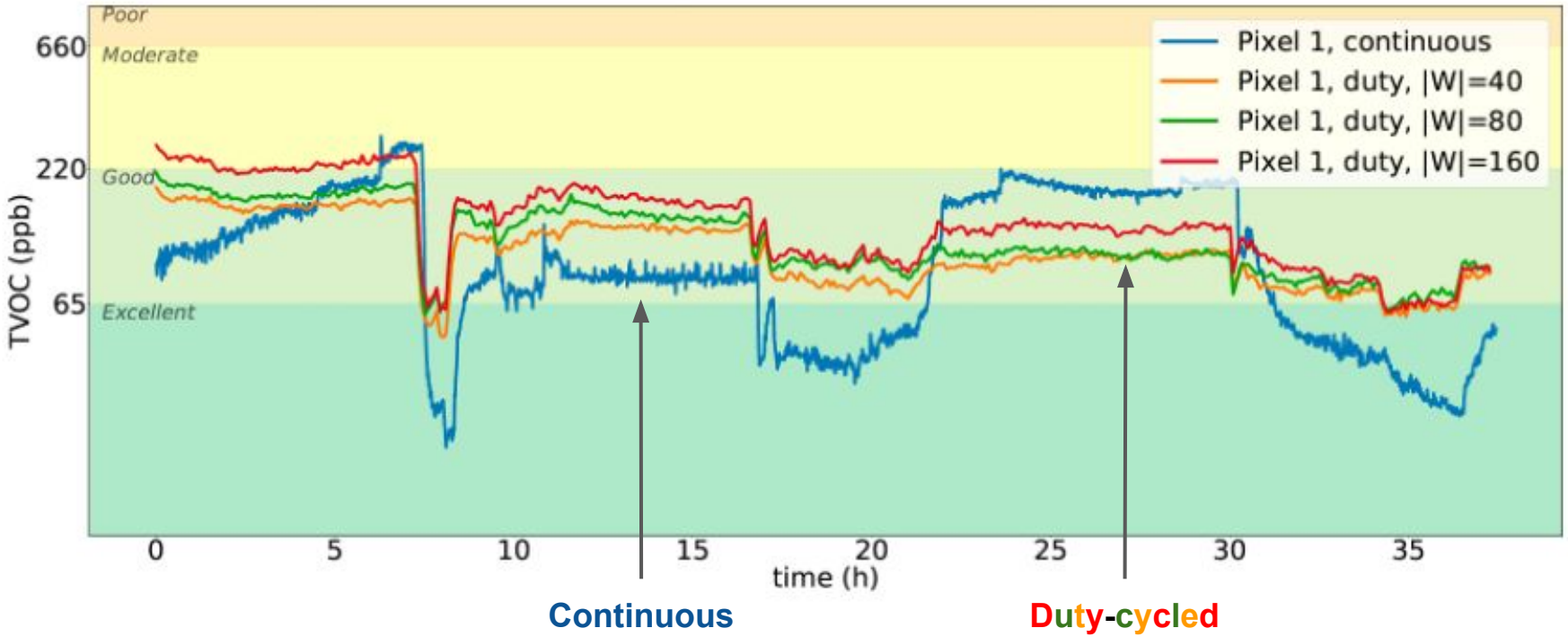
→ Miniaturized sensors lack specificity (=selectivity), drift, susceptible to T/H variations

D. Hasenfratz et al., Deriving High-Resolution Urban Air Pollution Maps Using Mobile Sensor Nodes, J PMC 2015

Predictive Sensing

→ Duty-cycling changes sensor's sensitivity

- Chemical reactions are not instantaneous and need time to run to completion

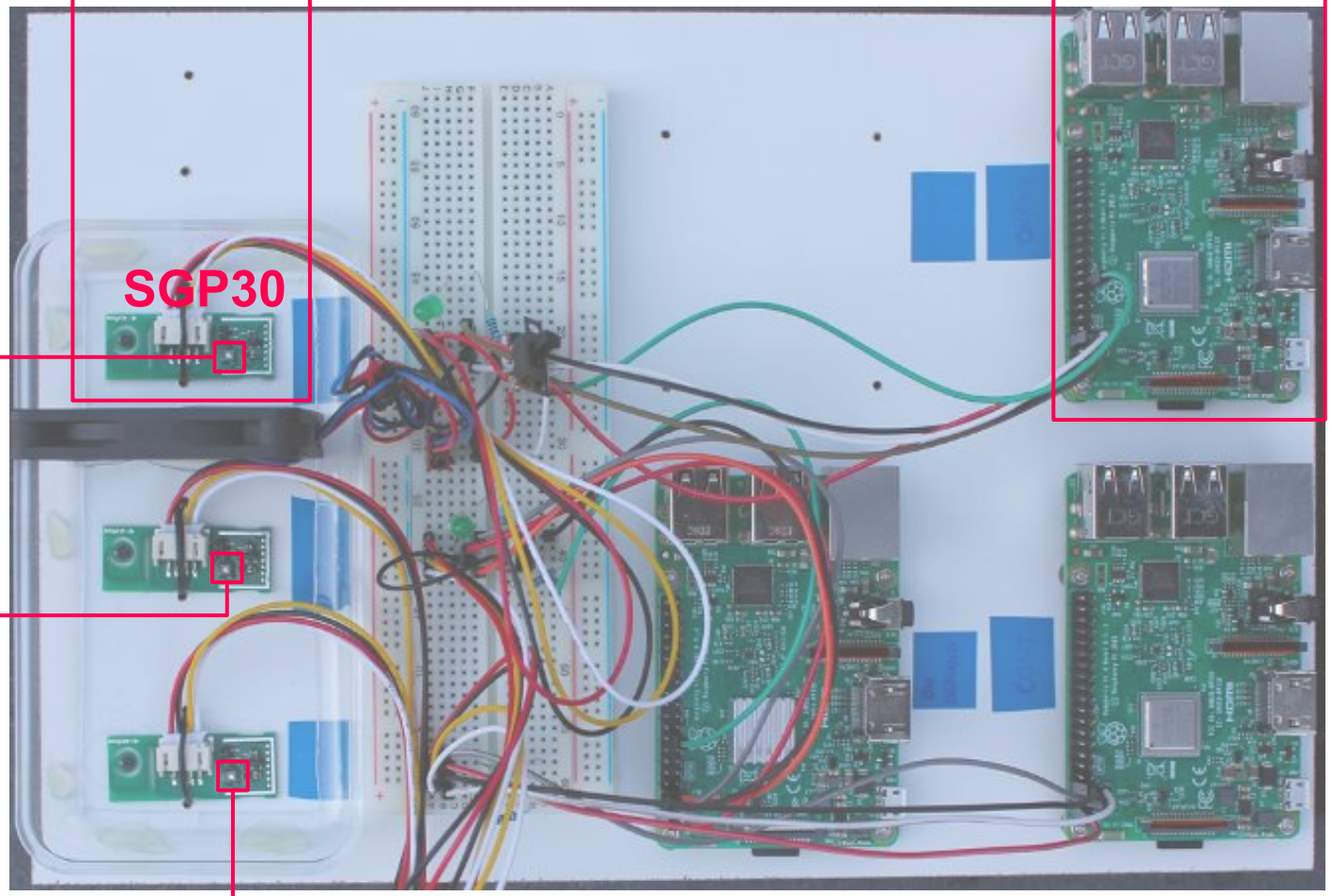


Sensirion's SGP30

⇒ Can we compensate for sensor sensitivity change due to duty-cycling?

M.-P. Gherman, Y. Cheng, A. Gomez, O. Saukh,
Compensating Altered Sensitivity of Duty-cycling MOX Gas Sensors by Machine Learning, IEEE SECON 2021

Measurement Setup



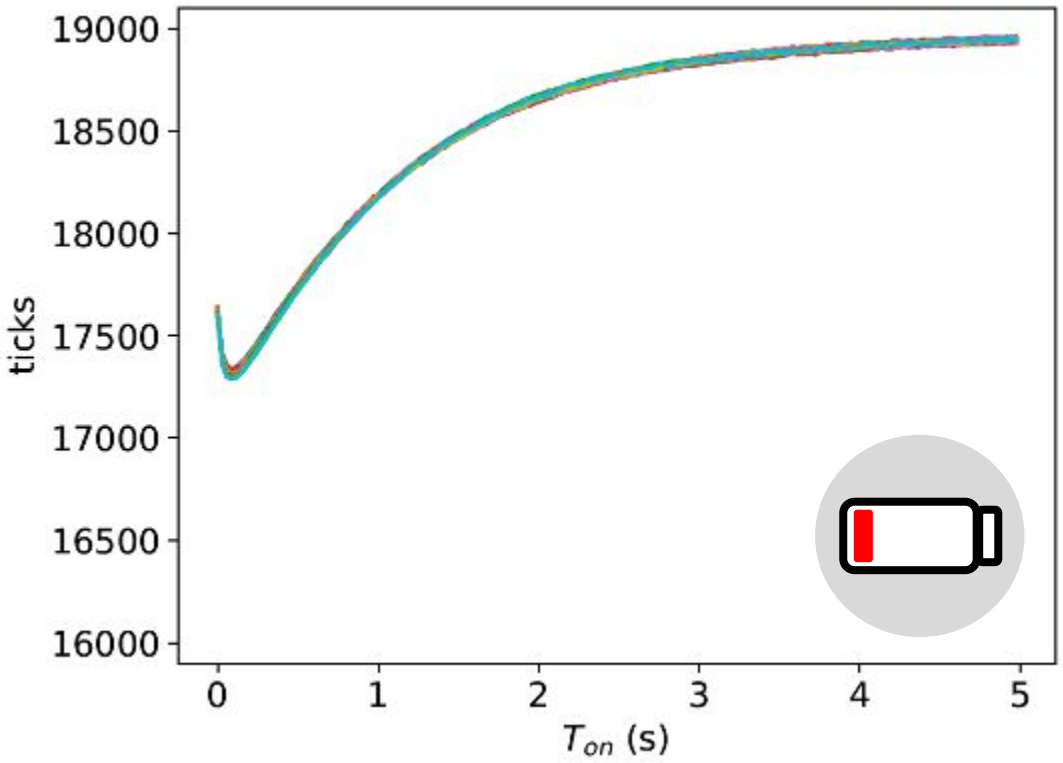
Continuous (86 mW)

Duty-cycle (1.6%)
→ T_{ON}=5s, T_{OFF}=295s

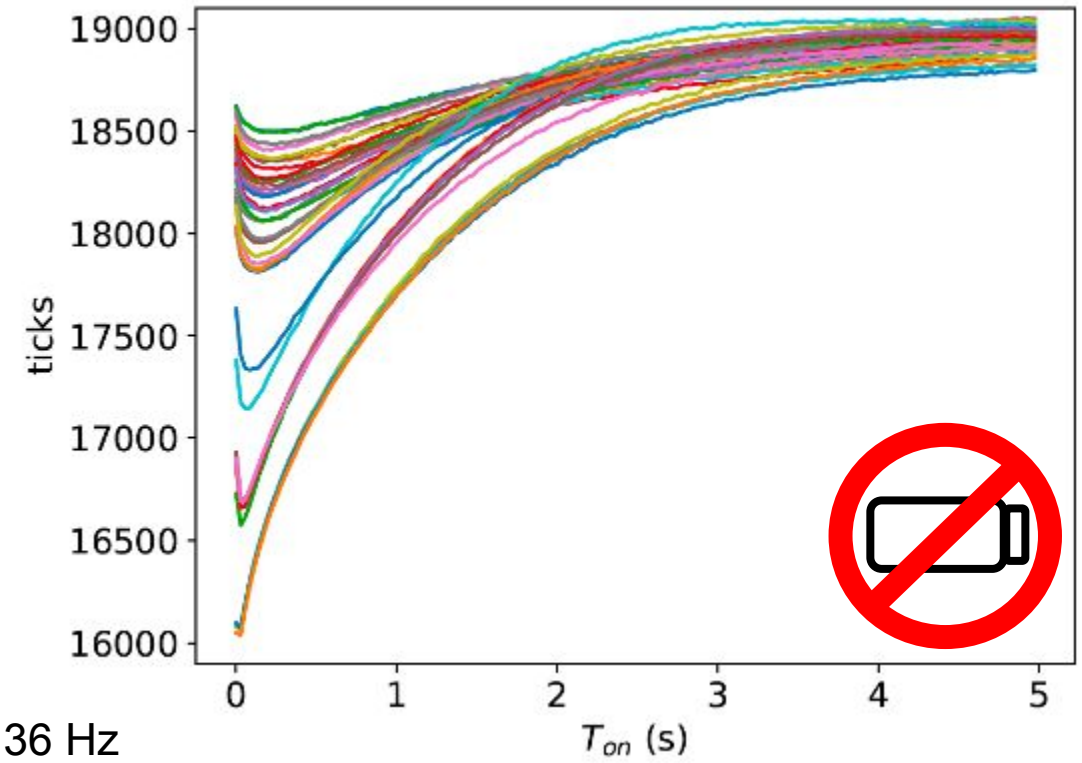
On-demand (1.6% on average)
→ T_{ON}=5s, T_{OFF} = RandomChoice {25s, 40s, 55s, 85s, 115s, 295s, 595s, 1195s}

Transient Responses

Duty-cycle (=regular sampling)



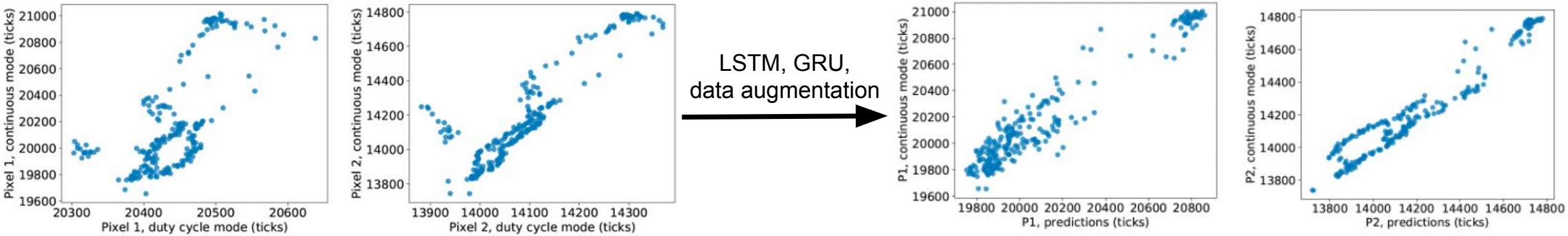
On-demand (=irregular sampling)



36 Hz

⇒ Transient response depends on the T_{OFF} history and on the gas concentration

Low-power Gas Sensing



⇒ 98.5% energy saved, tVOC: 3% (23ppb) error, 14KB model size, 2.5ms on RasPi

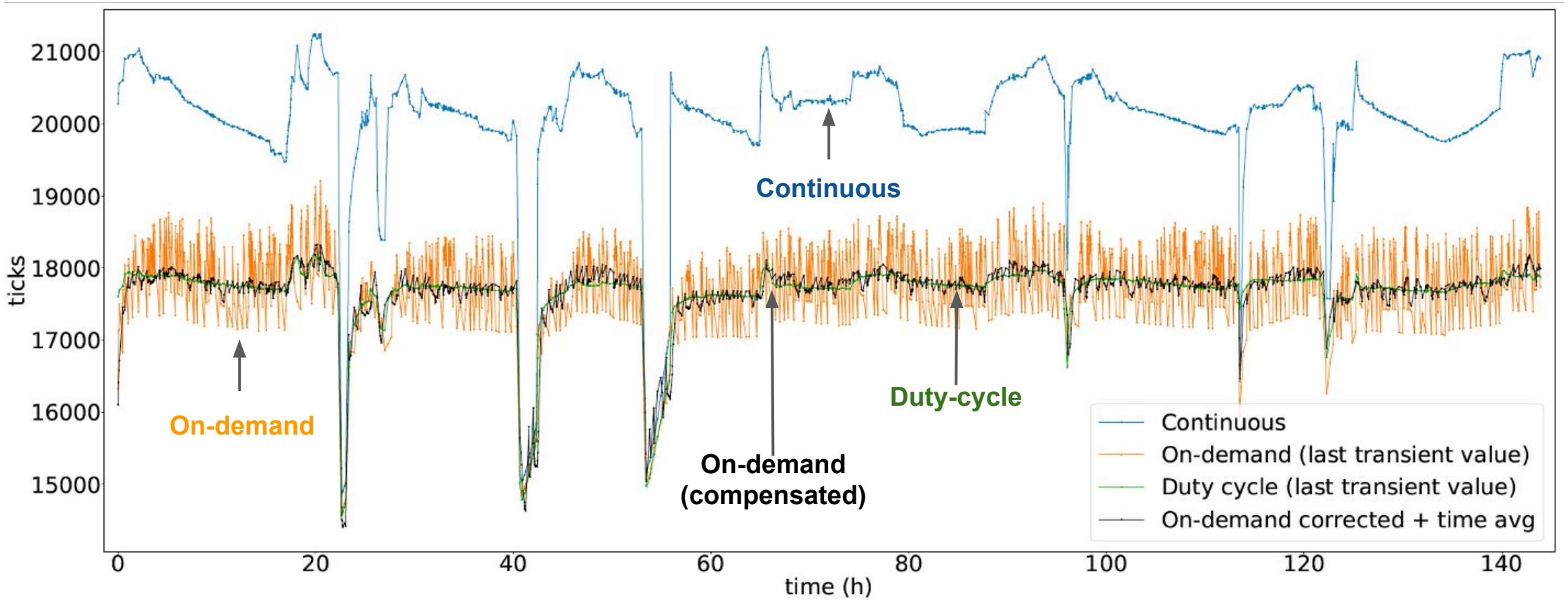
→ Today: sensing layer is optimized for selectivity. Compensation (T / H / drift) on top.

→ Models can derive selectivity given sufficient data. A stable signal is needed.

M.-P. Gherman, Y. Cheng, A. Gomez, O. Saukh,
Compensating Altered Sensitivity of Duty-cycling MOX Gas Sensors by Machine Learning, IEEE SECON 2021

On-demand Gas Sensing

⇒ T_{OFF} between 5s and 20min, tVOC accuracy improved by 79% compared to no compensation



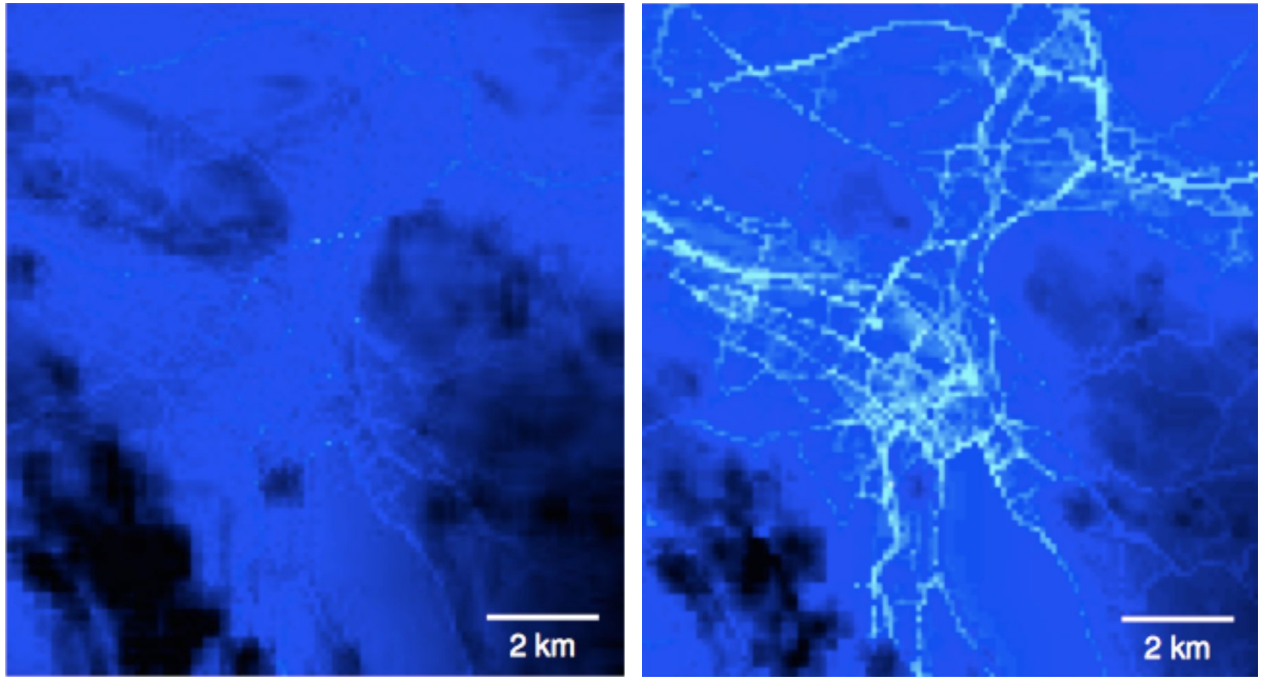
M.-P. Gherman, Y. Cheng, A. Gomez, O. Saukh, Towards on-demand gas sensing, DCOSS 2021

2 “What can we learn from the data?”

→ Spatiotemporal predictions
- Accurate
- High resolution
- Transferable
- Stable / long-term

→ Control of local air quality levels
- Policies, schedules, air purification, …

→ Understanding of reduction potential



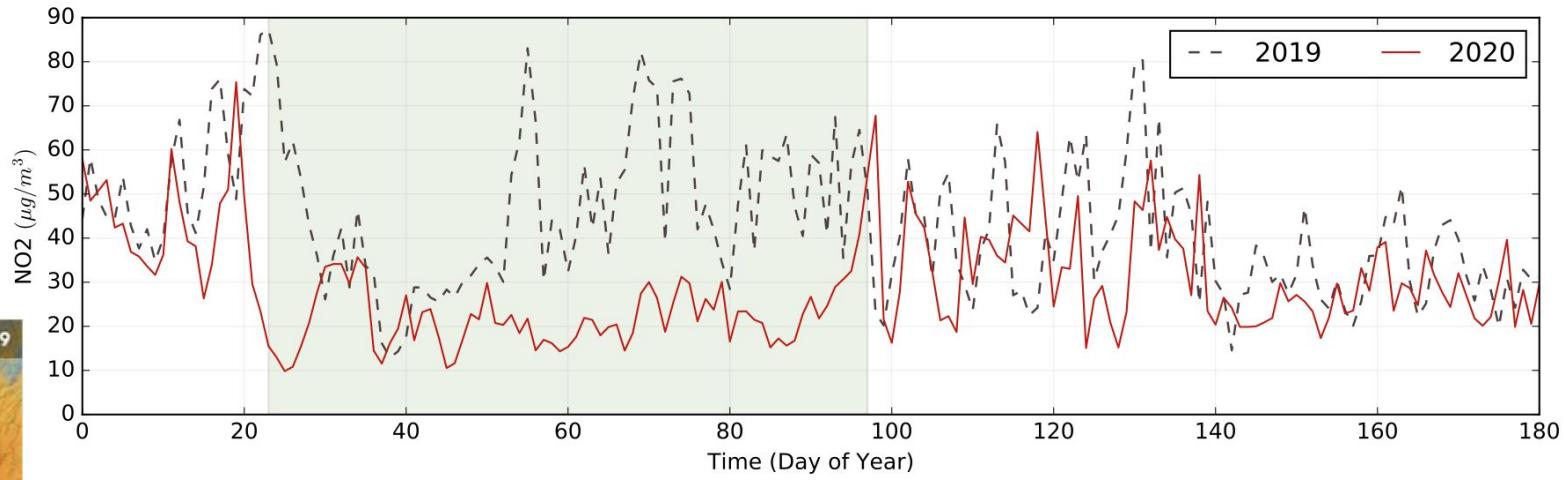
D. Hasenfratz et al., Deriving High-Resolution Urban Air Pollution Maps Using Mobile Sensor Nodes, J PMC 2015

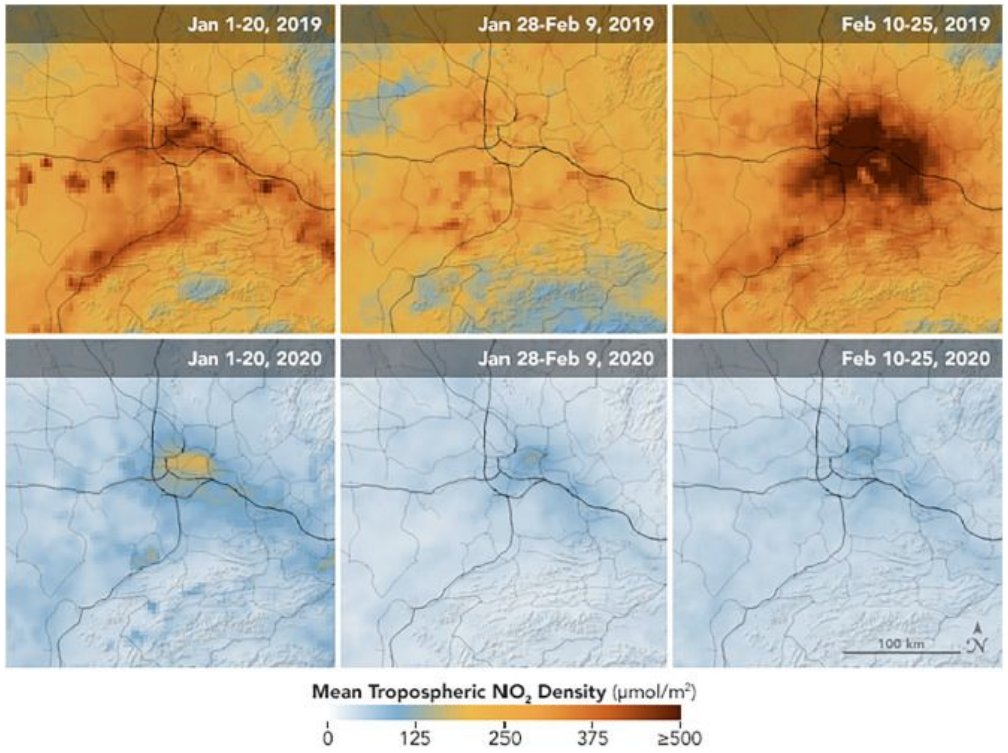
COVID-19 Example

⇒ How much air quality improvement was due to COVID-19 lockdowns?

NO2 measured in 2019 and 2020 in Wuhan, China

J. Einsidler, Y. Cheng, F. Papst, O. Saukh, Transferable Models to Understand the Impact of Lockdown Measures on Local Air Quality, UrbCom, 2021





Studies assessed reduction of NO2 based on satellite imagery collected by NASA and ESA (no weather compensation)

S. Muhammad, X. Long, and M. Salman, COVID-19 pandemic and environmental pollution: A blessing in disguise?, Science of The Total Environment, vol. 728, p. 138820, 2020

M. Bauwens, S. Compernolle, andet al., Impact of coronavirus outbreak on NO2 pollution assessed using TROPOMI and OMI observations, Geophysical Research Letters, vol. 47, no. 11, Jun. 2020

Methodology

1

Train on pre-lockdown data and predict the lockdown period (→ weather dependency)

⇒ *pre-LD model*

→

2

Compare to actual measurements during lockdown

→

3

Use the difference as an estimate of lockdown-related reduction

Datasets: China: Beijing / Wuhan, Switzerland: Zurich / St. Gallen, Lower Austria

→ Long-Term Air Pollution Models

- Classical Dispersion Models are widely used (theory driven)
- Generalized Additive Models (data driven)

→ Short-Term Air Pollution Models

- Focused on a few hours to a few days ahead prediction, state-of-the-art performance by DNNs

Generalized Additive Models / Model Selection

→ Developed in 1986 by Hastie & Tibshirani

→ Extension of standard linear regression models

→ Impact of a predictive variable captured through smooth function

$$ln(\mathrm{UFP}) = a + s_1(A_1) + s_2(A_2) + \cdots + e,$$

where a is the intercept, $s_{1..n}(\cdot)$ are spline functions, $A_{1..n}$ are explanatory variables, and e is the error term

Variable name	Abbr.	Switzerland NO2	Switzerland PM10	Beijing NO2	Beijing PM2.5	Wuhan NO2	Wuhan PM2.5
Wind speed	WS	4	5	32	32	5	-
Wind dir. X	WD_x	6	8	18	21	8	10
Wind dir. Y	WD_y	3	4	7	7	11	11
Precipitation	P	-	-	-	-	-	-
Temperature	T	2	5	1	-	-	-
Rel. humidity	RH	2	3	21	9	11	9
Month	M	-	-	-	-	-	1
Day of year	DY	2	-	24	3	4	8
Dew point	DP	0	-	7	1	1	-
PCA	PCA	7	-	26	21	4	9
Weekday	D	1	2	-	2	1	-

TABLE II: #stations where the corresponding explanatory variable was chosen by the model selection algorithm.

J. Einsidler, Y. Cheng, F. Papst, O. Saukh, Transferable Models to Understand the Impact of Lockdown Measures on Local Air Quality, UrbCom, 2021

COVID-19 Example

NO2 reduction in Beijing / Wuhan over the lockdown periods:

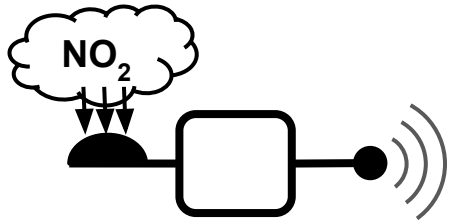
→ Beijing: -28.1%

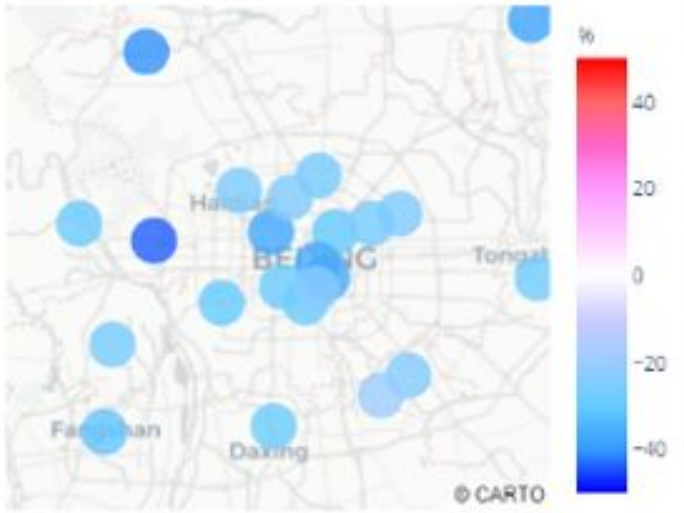
→ Wuhan: -52.8%

Comparison to estimates based on satellite imagery:

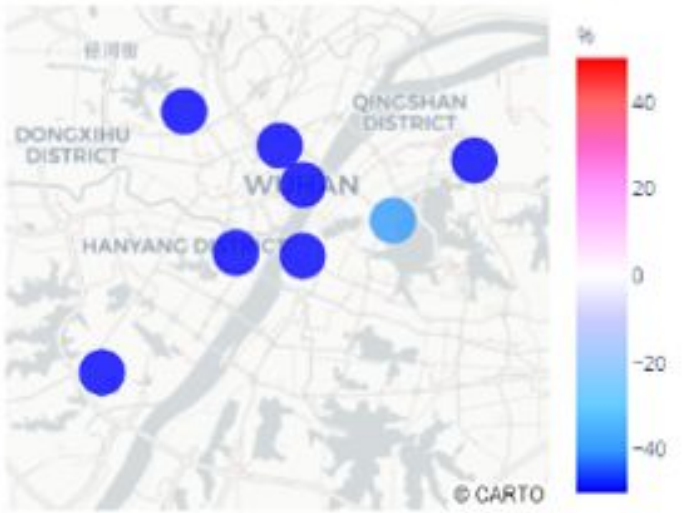
Estimators	Beijing Urban %	Beijing Suburban %	Beijing Rural %	Beijing Road %	Wuhan %
pre-LD	-33.3	-25.8	-37.4	-33.1	-64.1
TROPOMI [11]	-25±10	-25±10	-25±10	-25±10	-43±14
OMI [11]	-33±10	-33±10	-33±10	-33±10	-57±14

TABLE IV: Estimated NO2 reduction in Beijing and Wuhan compared to [11] as measured by satellite imagery.





(c) Beijing



(d) Wuhan

J. Einsidler, Y. Cheng, F. Papst, O. Saukh,
Transferable Models to Understand the Impact of Lockdown Measures on Local Air Quality, UrbCom, 2021

COVID-19 Example

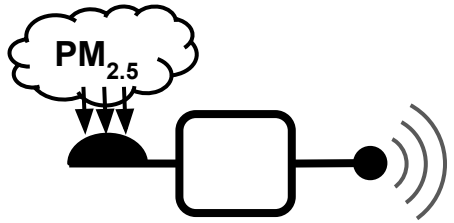
PM2.5 reduction in Beijing / Wuhan over the lockdown periods:

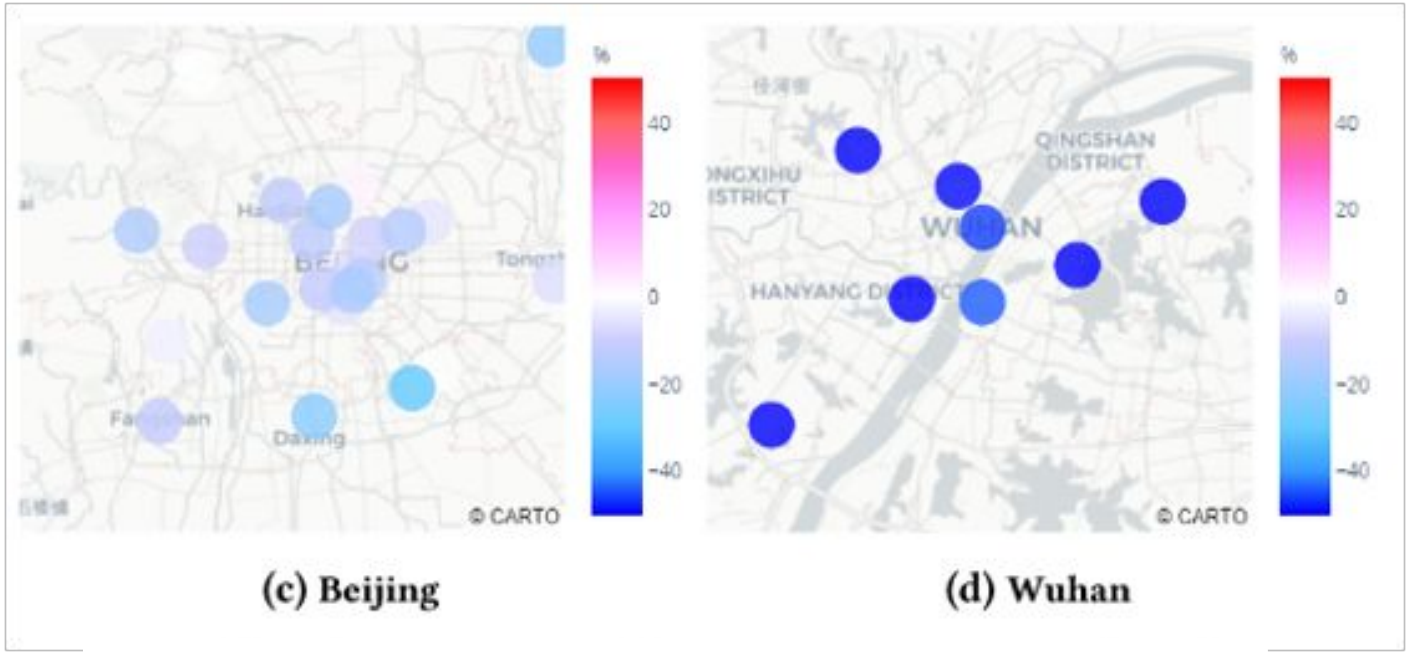
→ Beijing: -10.8%

→ Wuhan: -50.0%

⇒ Comparatively low reduction in Beijing due to increased oxidation capacity

⇒ Pollution transfer





J. Einsidler, Y. Cheng, F. Papst, O. Saukh,
Transferable Models to Understand the Impact of Lockdown Measures on Local Air Quality, UrbCom, 2021

Pollution Transfer

⇒ Can modelling pollution transfer between cities improve predictions?

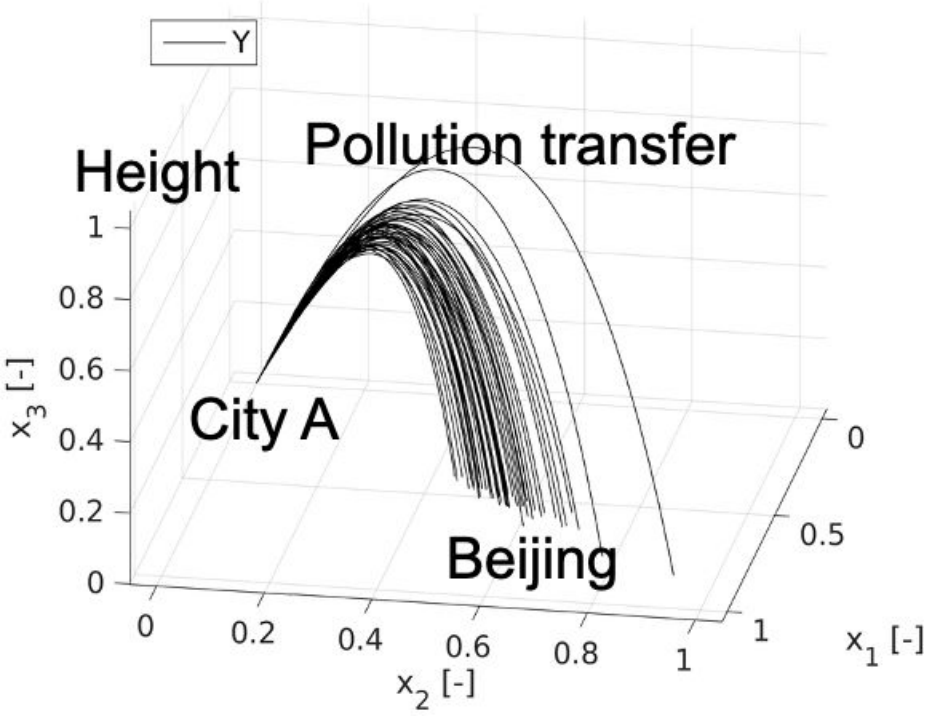


Image source: Rushing Towards Oxygen Cylinders, 2021

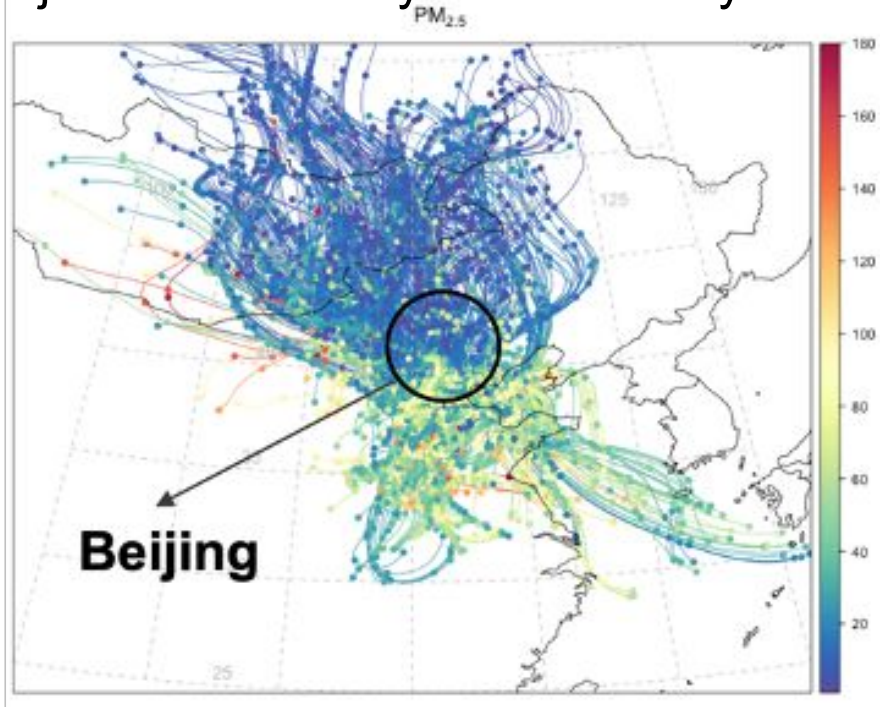
→ Inter-city pollution transfer exists in clusters of cities

→ Combining air quality measurements from all cities in the cluster improves predictions

Air Flow Trajectory Data from HYSPLIT

→ Physics-based pollution propagation analysis tool

- An air flow trajectory is a sequence of sample points from the underlying route of the moving air flow
- Aggregated air flow trajectories over a year colored by measured PM2.5 values



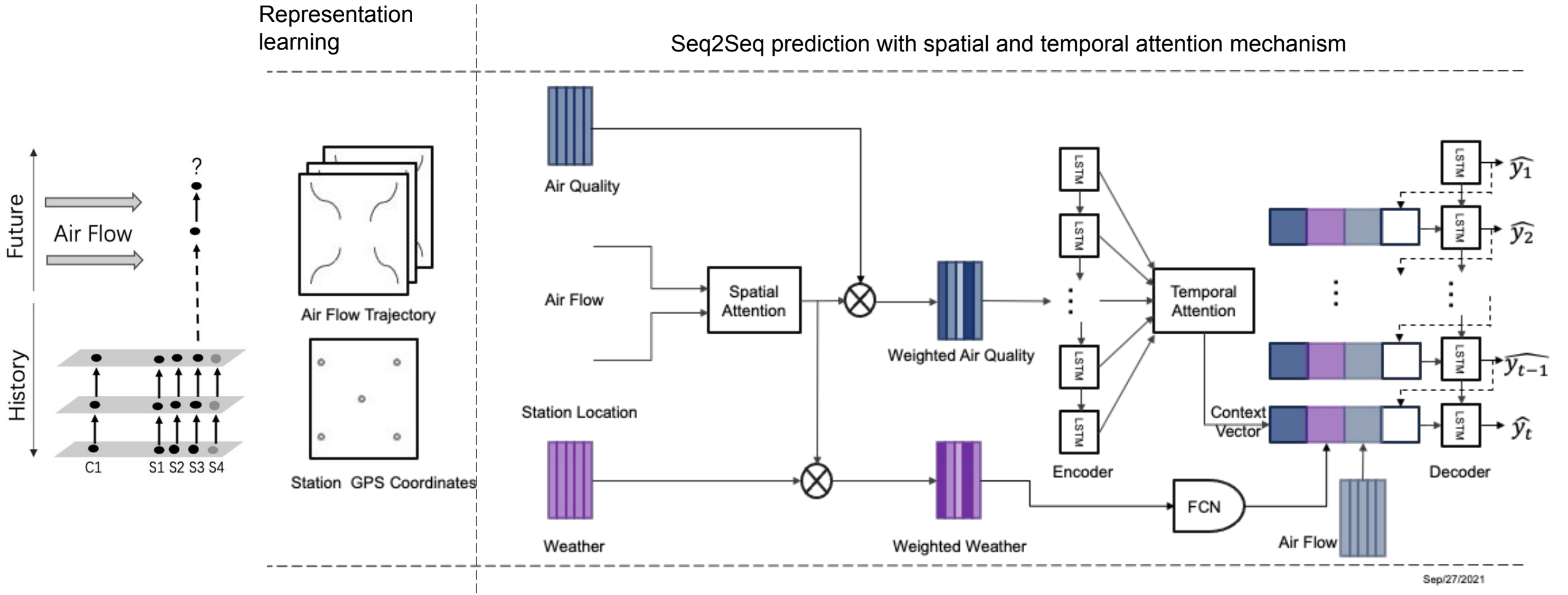
Given:

- Air quality data (X)
- Weather data (W)
- Air Flow Trajectory data (J)
- Datetime Features (D)

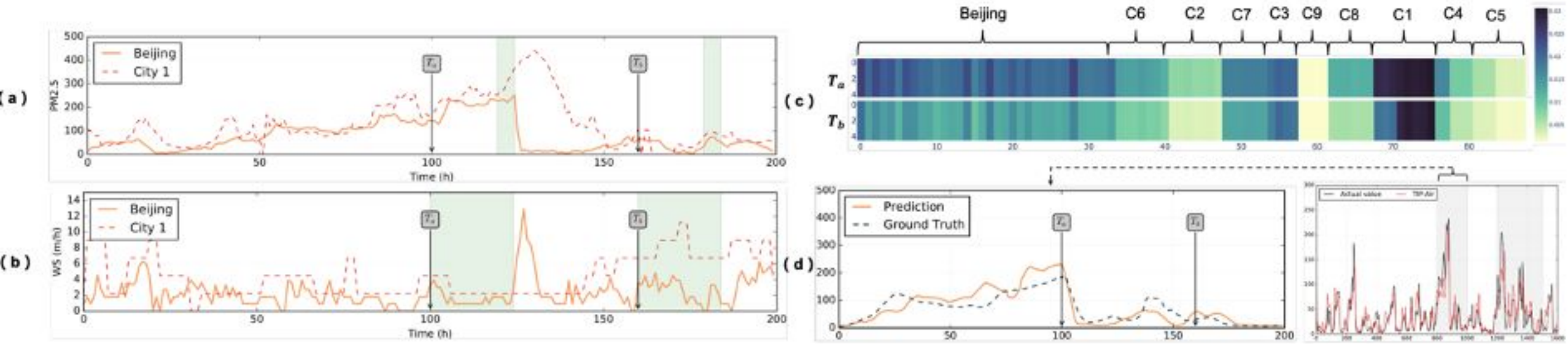
we aim to predict the air quality of future hours:

$$\hat{y} = \mathcal{F}(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{D}, \boldsymbol{J})$$

TIP-Air Framework



Results and Insights



State-of-the-art includes only data in the target city

⇒ TIP-Air works best among all methods and reduces MAE by 9.6% to 22.4%

⇒ TIP-Air behaves best for sudden change prediction

	1-6h	7-12h	13-18h	19-24h
Naïve	14.87	26.00	32.21	35.45
LSTM	14.17	25.88	32.67	37.04
Seq2Seq	14.13	23.99	30.14	33.61
GeoMAN	14.03	19.42	22.95	24.23
MGED-Net	13.44	18.05	20.95	21.91
TIP-Air	**12.15**	**15.34**	**16.25**	**17.21**

Lessons Learned

→ Data processing in air quality contexts increasingly leverages deep learning

- IoT sensor density and data relevance are important
- New methods improve performance of air pollution models
- Especially for short-term pollution predictions

→ Increasingly more processing is done on the edge

- Resource constraints: Energy, compute, bandwidth
- Predictive / on-demand sensing has high potential

→ Robustness of models is an issue

- Models trained on past data may not work on newer data
- Spatiotemporal model transfer is difficult:
  - Policies change, economy evolves, human's awareness increases / behavior changes

→ No framework for all areas and sensors …

Thank You!

Olga Saukh

saukh@tugraz.at

http://www.olgasaukh.com

TU Graz

COMPLEXITY SCIENCE HUB VIENNA

Collaborators:

Yun Cheng ○ Johanna Einsiedler ○ Markus-Philipp Gherman ○ Andres Gomez
David Hasenfraz ○ Christoph Hueglin ○ Balz Maag ○ Franz Papst ○ Naomi Stricker
Lothar Thiele ○ Christoph Waser ○ Zimu Zhou

Questions?