Proposed Best Practices for Quantifying, Siting, and Using Gas-Phase Sensors in Partnership with Communities

Presenter: Ashley Collier-Oxandale[1]

Sharad Vikram[2], Michael Ostertag[2], Massimiliano Menarini[2], Evan Coffey[1], Jacob Thorson[1], Jill Johnston[3], William Griswold[2], and Michael Hannigan[1]

[1]University of Colorado Boulder; [2]University of California San Diego, [3]University of Southern California

Air Sensors International Conference, September 12-14, 2018

Overview

- *What can we learn about quantification from deployments in different locations?*

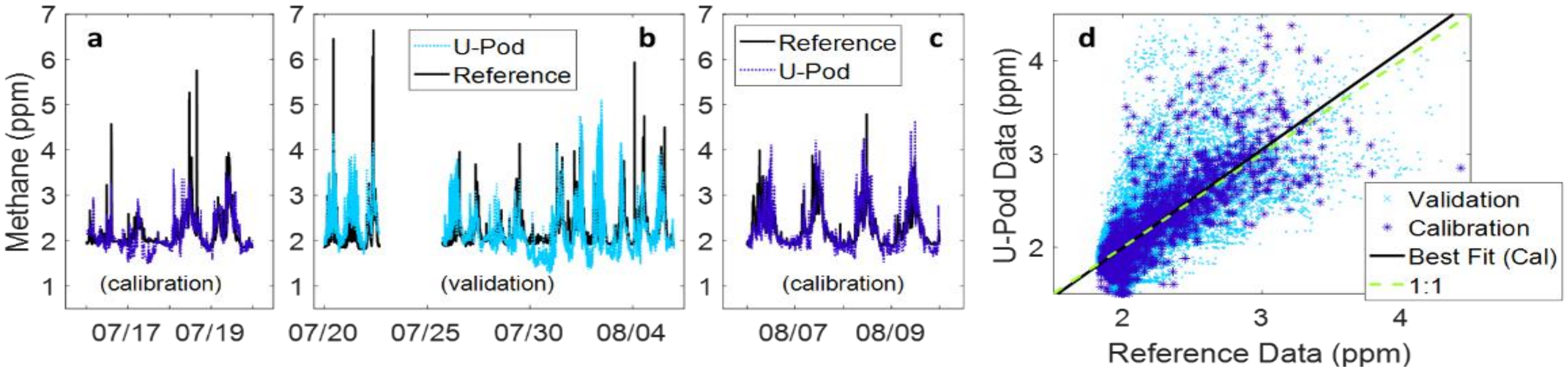
- *How transferable are calibrations between locations?*

- *What other tools, resources, and information (in addition to sensor quantification) could support partnerships with communities around sensor use?*

A Need for Location Specific Calibrations

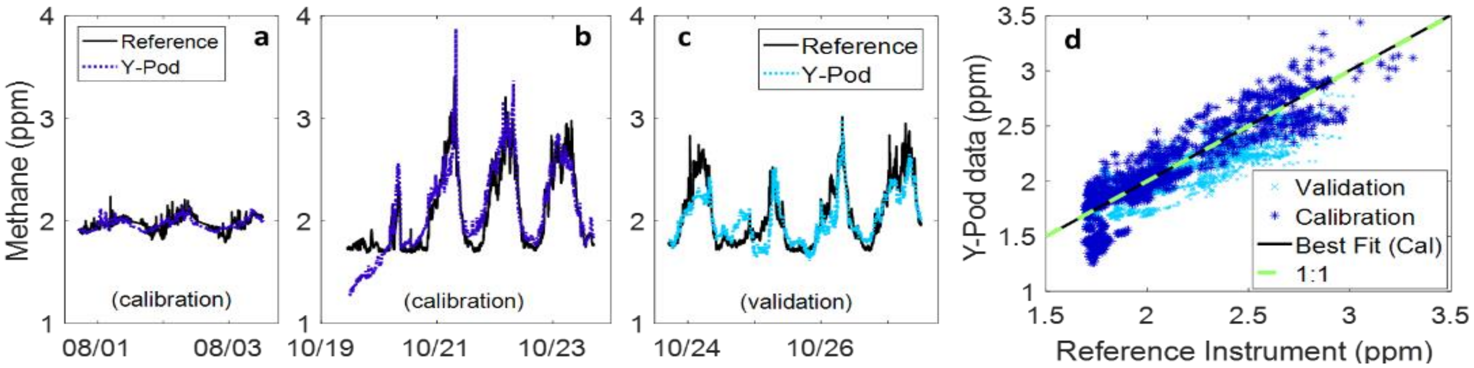
Colorado (rural)

	RMSE (ppm)	R^2
Training	0.27	0.63
Testing	0.38	0.45



Los Angeles (urban)

	RMSE (ppm)	R^2
Training	0.15	0.82
Testing	0.16	0.76

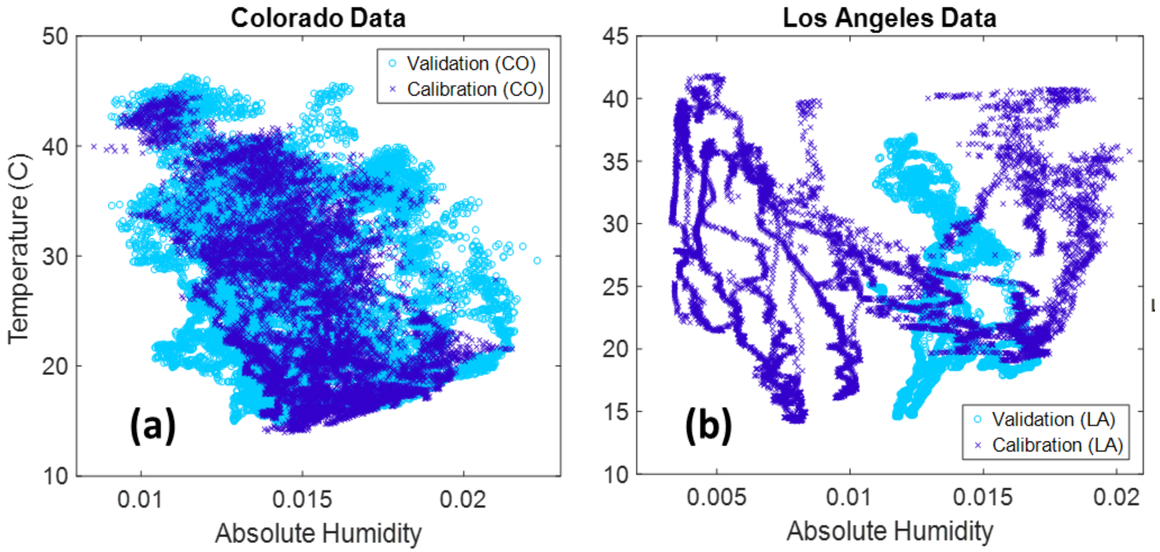


Colorado Model: $R_s/R_0 = p_1 + p_2(C) + T * (p_3 + p_6(C)) + p_4(H) + p_5(T_i)$

Los Angeles Model: $R_s/R_0 = p_1 + p_2(C) + T * (p_3 + p_6(C)) + p_4(H^{-1}) + p_5(T_i) + p_7(T * H^{-1}) + p_8(T_d)$

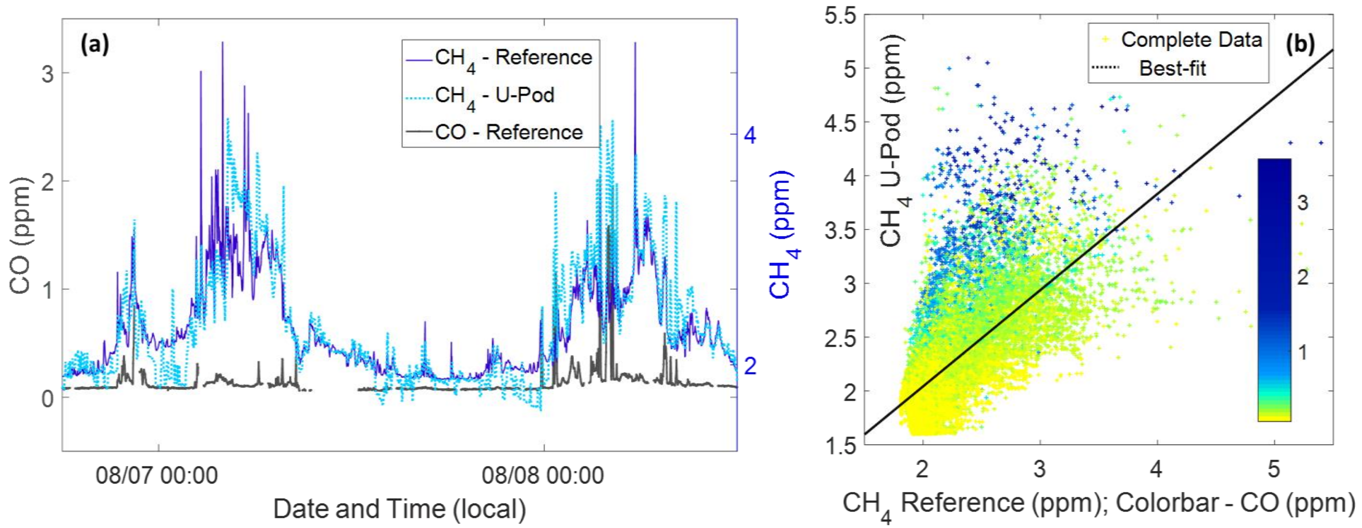
from Collier-Oxandale et al., 2018

What's driving the differences across locations?



Complex, location specific temperature and humidity effects

Cross-sensitivities and differing mixtures of background pollutants



from Collier-Oxandale et al., 2018

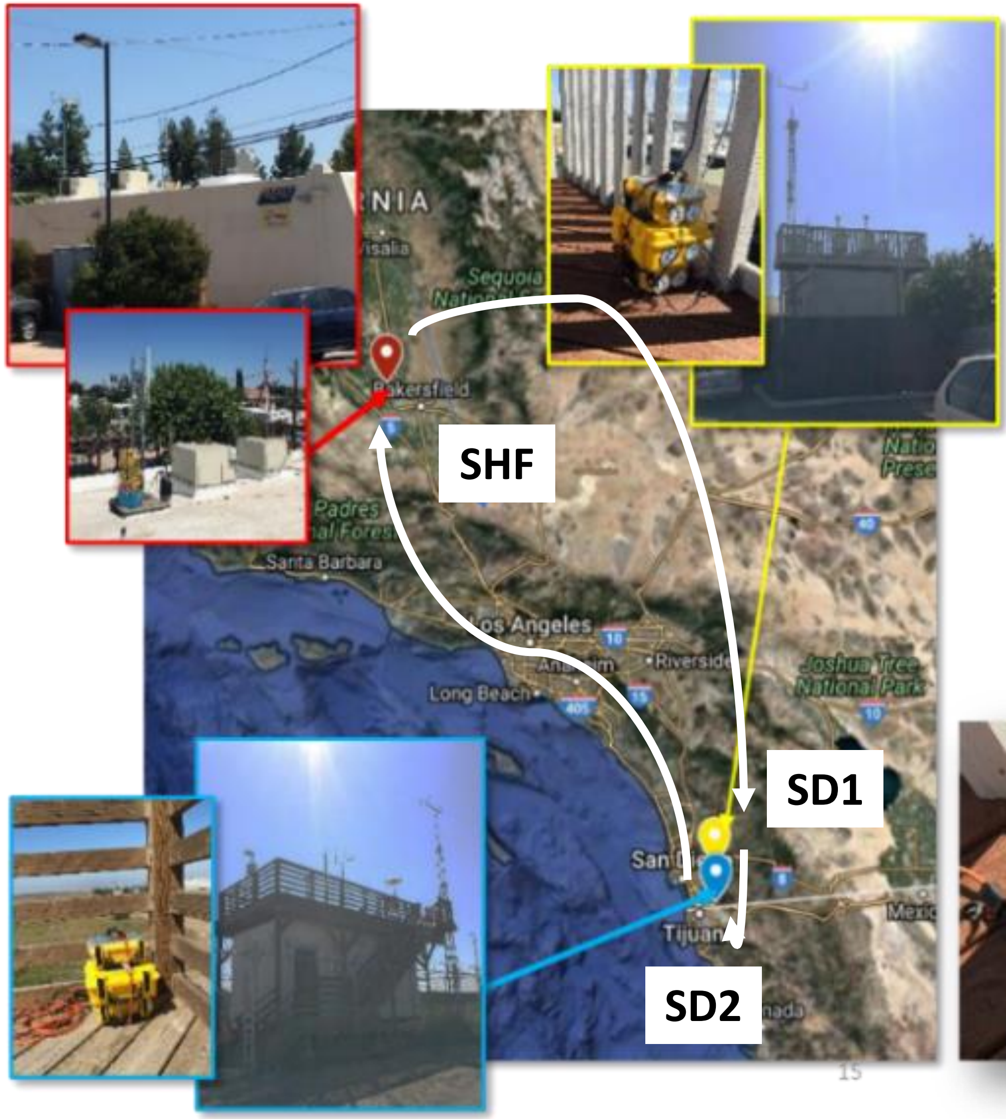
Calibration Transferability

- Three MetaSense sensor systems at three reference sites
  - SD1: El Cajon (urban/suburban site)
  - SD2: Donovan (rural site, near the southern boarder)
  - SHF: Shafter (suburban/rural, near Bakersfield)

Reference Instruments:
- Donovan – NO_2, O_3
- El Cajon – NO_2, O_3, CO
- Shafter DMV – NO_2, O_3, TNMHC, CO_2 (CO_2 – via Licor Analyzer maintained by CU, Boulder)

Tested different quantification models and robustness of models in new locations

Calibration Transferability



- Three MetaSense sensor systems at three reference sites
  - SD1: El Cajon (urban/suburban site)
  - SD2: Donovan (rural site, near the southern boarder)
  - SHF: Shafter (suburban/rural, near Bakersfield)

Reference Instruments:

- Donovan – NO_2, O_3
- El Cajon – NO_2, O_3, CO
- Shafter DMV – NO_2, O_3, TNMHC, CO_2 (CO_2 – via Licor Analyzer maintained by CU, Boulder)

Tested different quantification models and robustness of models in new locations

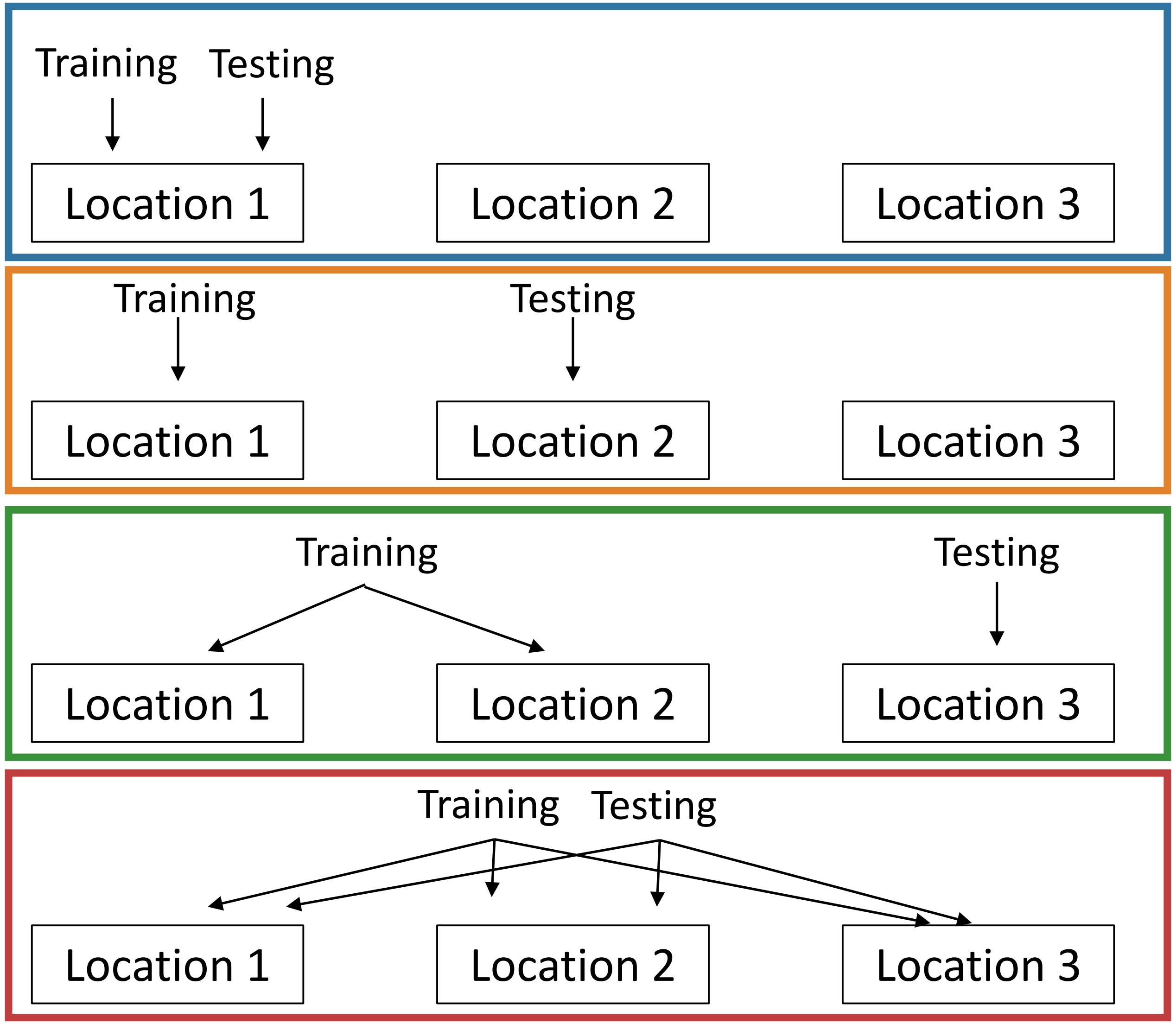
Test Descriptions

Models

- Multiple Linear Regression (MLR)
- Neural Network – 2 layer (NN2)
- Neural Network – 4 layer (NN4)
- Random Forests (RF)

Levels (units – ppb)

- 0: training and testing on one location
- 1: training on 1 location, testing on another
- 2: training on 2 locations, testing on the third
- 3: training and testing across all locations



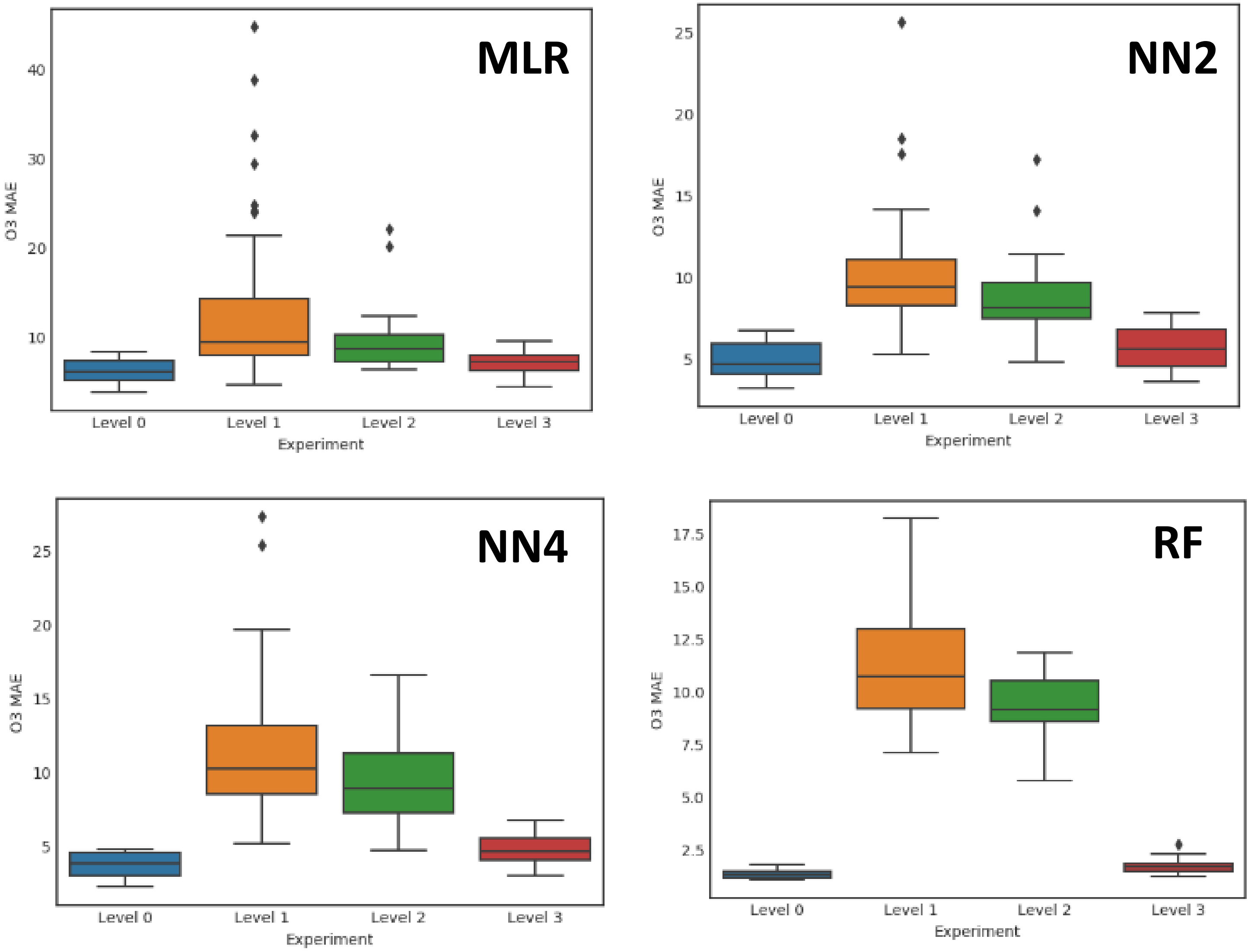
Transferability Results – O_3 (electrochemical)

Models

- Multiple Linear Regression (MLR)
- Neural Network – 2 layer (NN2)
- Neural Network – 4 layer (NN4)
- Random Forests (RF)

Levels (units – ppb)

- 0: training and testing on one location
- 1: training on 1 location, testing on remaining 2
- 2: training on 2 locations, testing on the 3rd
- 3: training and testing across all locations



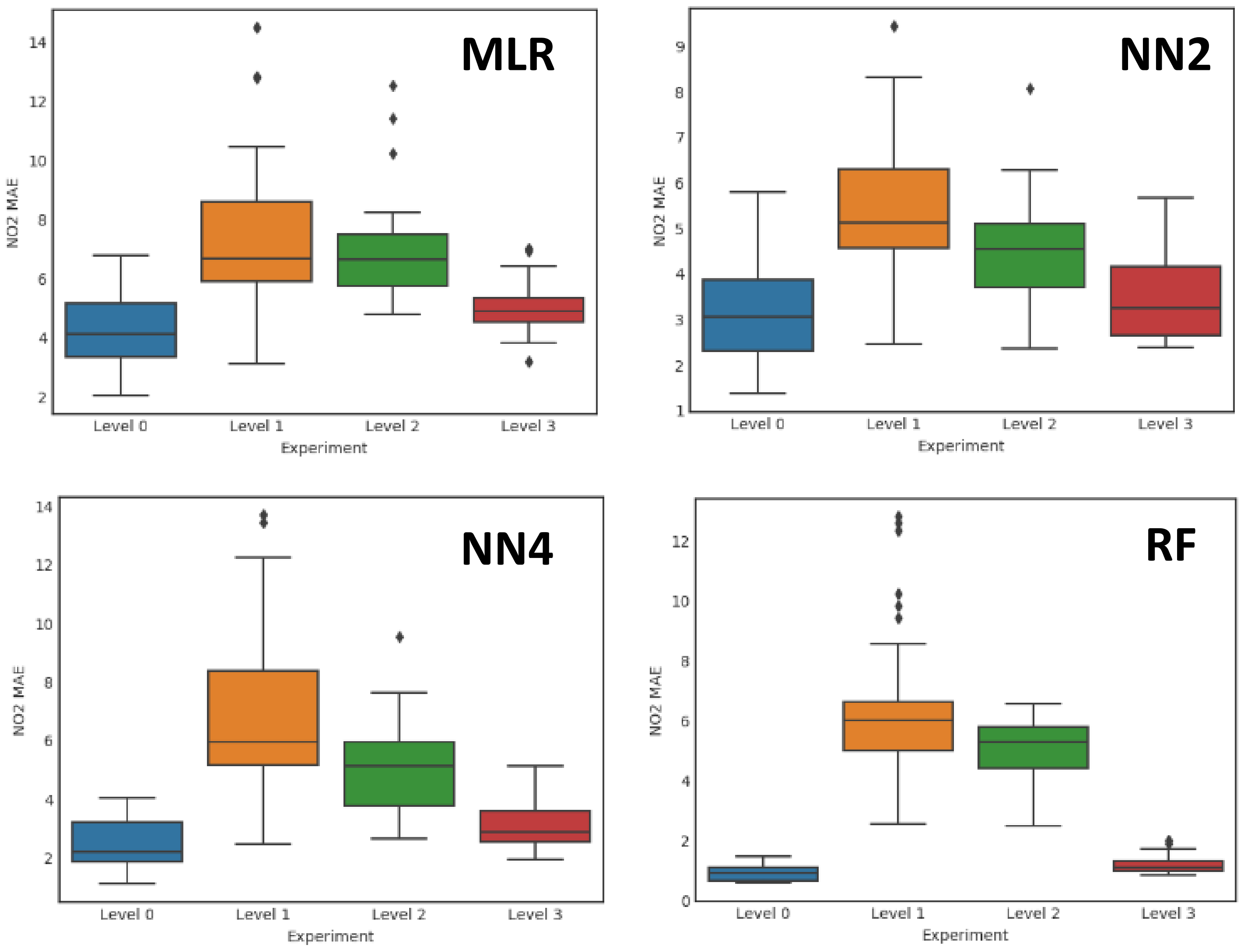
Transferability Results – NO2 (electrochemical)

Models

- Multiple Linear Regression (MLR)
- Neural Network – 2 layer (NN2)
- Neural Network – 4 layer (NN4)
- Random Forests (RF)

Levels (units – ppb)

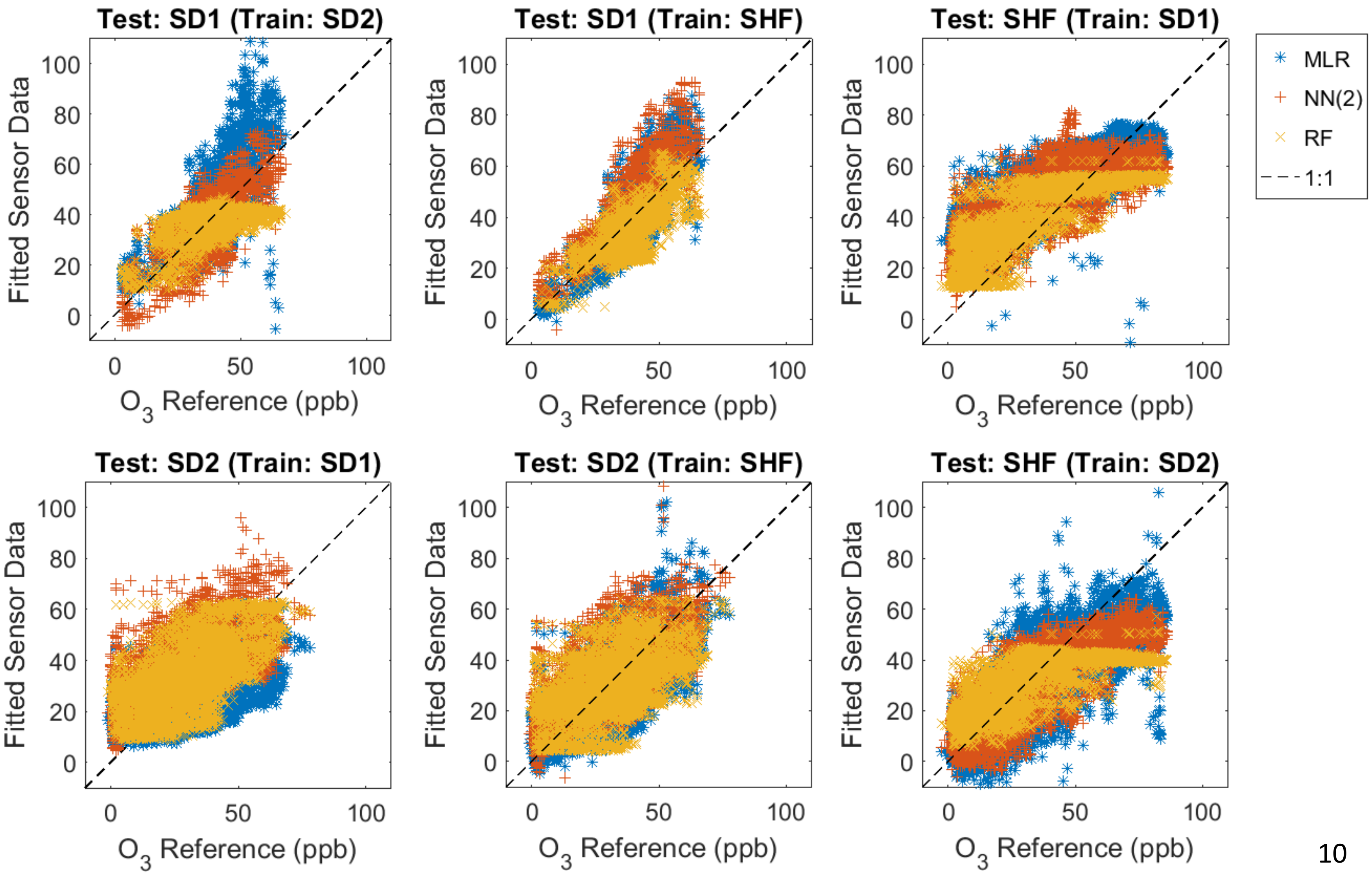
- 0: training and testing on one location
- 1: training on 1 location, testing on remaining 2
- 2: training on 2 locations, testing on the 3rd
- 3: training and testing across all locations



Level 1 Tests

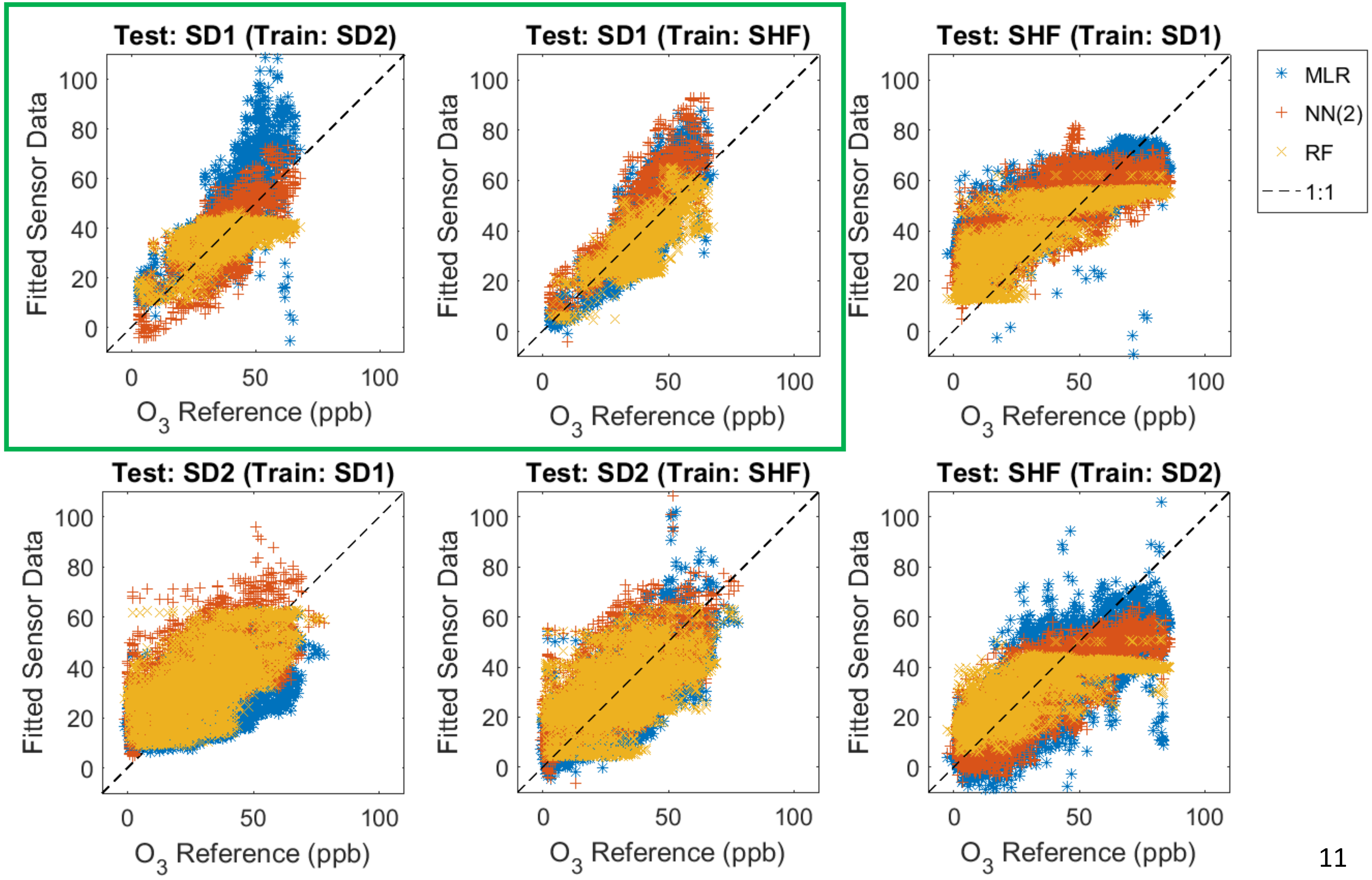
Comparing a calibration within the same city to calibrating in another part of the state

Observations:



Comparing a calibration within the same city to calibrating in another part of the state

Level 1 Tests



Observations:

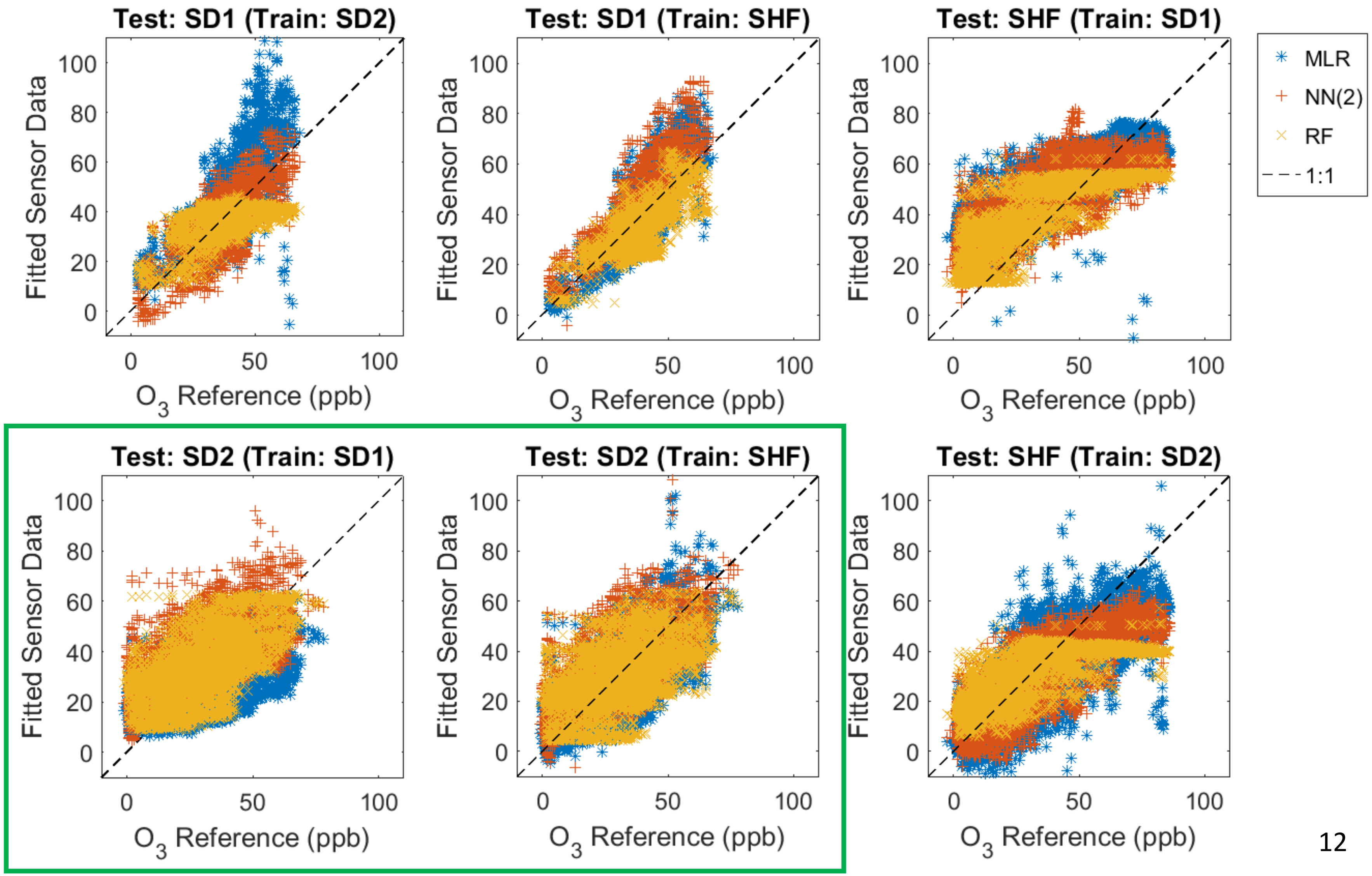
- *Over-predictions (particularly in MLR)*

Level 1 Tests

Comparing a calibration within the same city to calibrating in another part of the state

Observations:

- Over-predictions (particularly in MLR)
- *Relatively large variance*

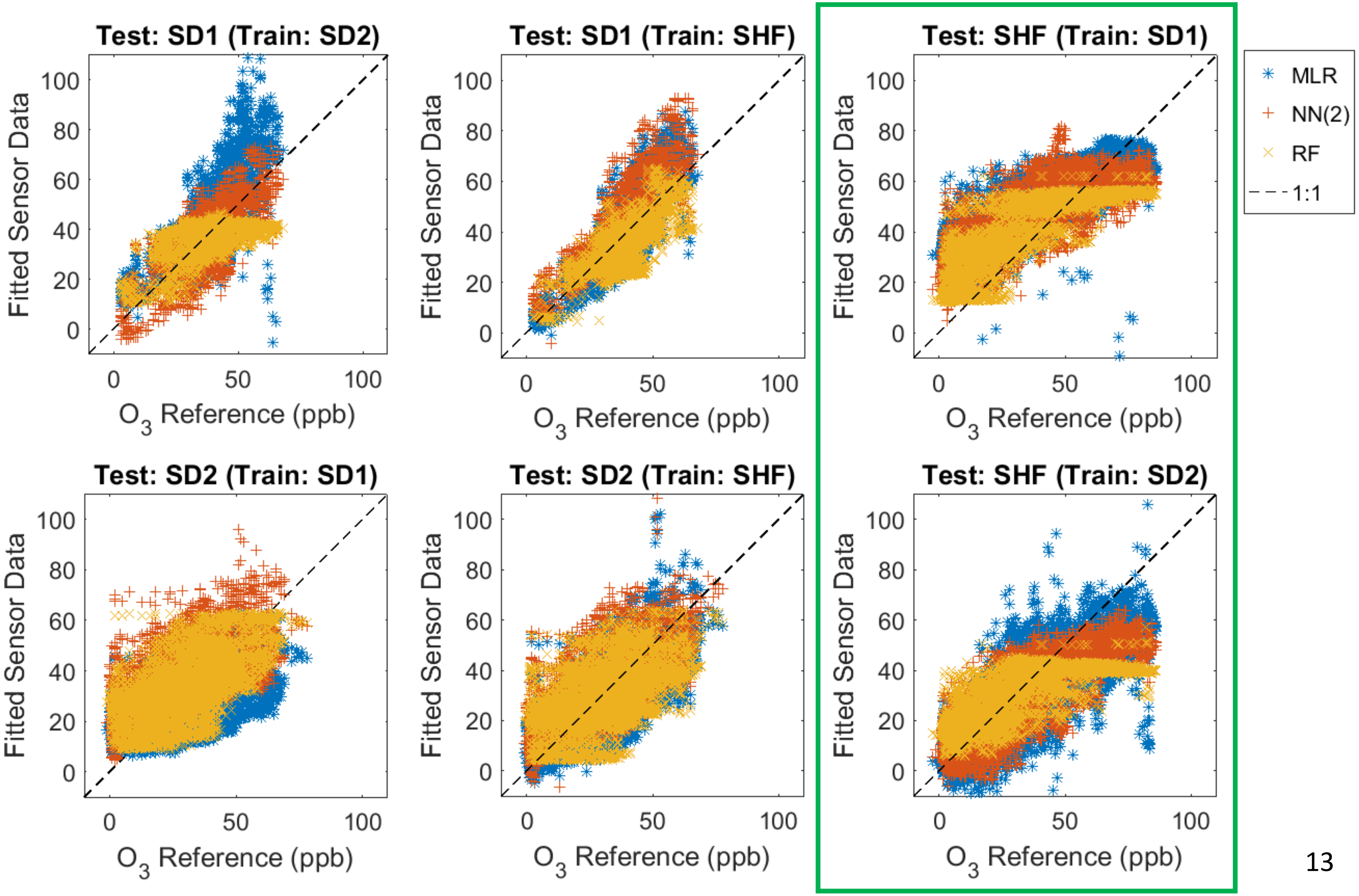


Level 1 Tests

Comparing a calibration within the same city to calibrating in another part of the state

Observations:

- Over-predictions (particularly in MLR)
- Relatively large variance
- *Underpredictions at high levels*

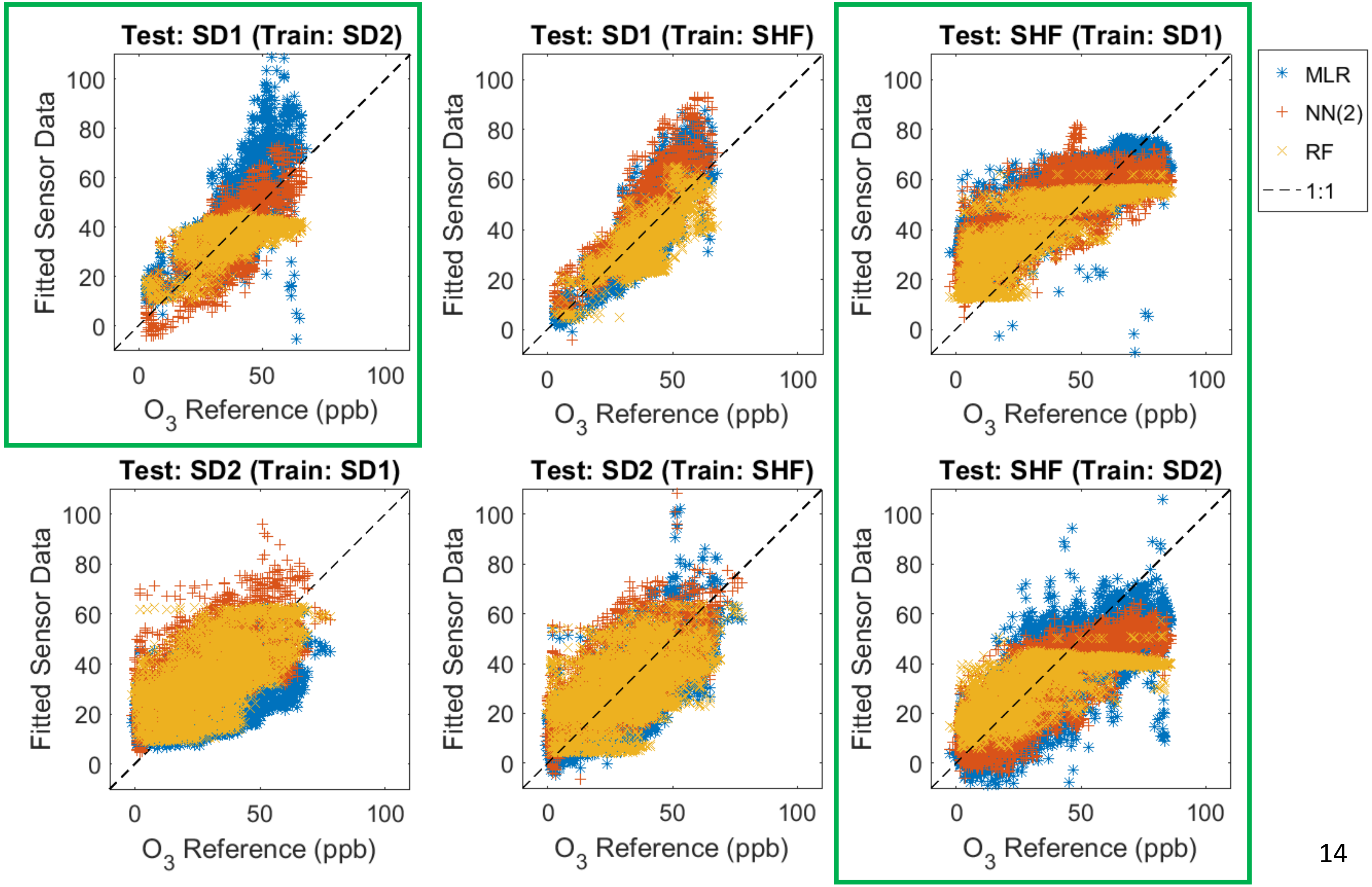


Level 1 Tests

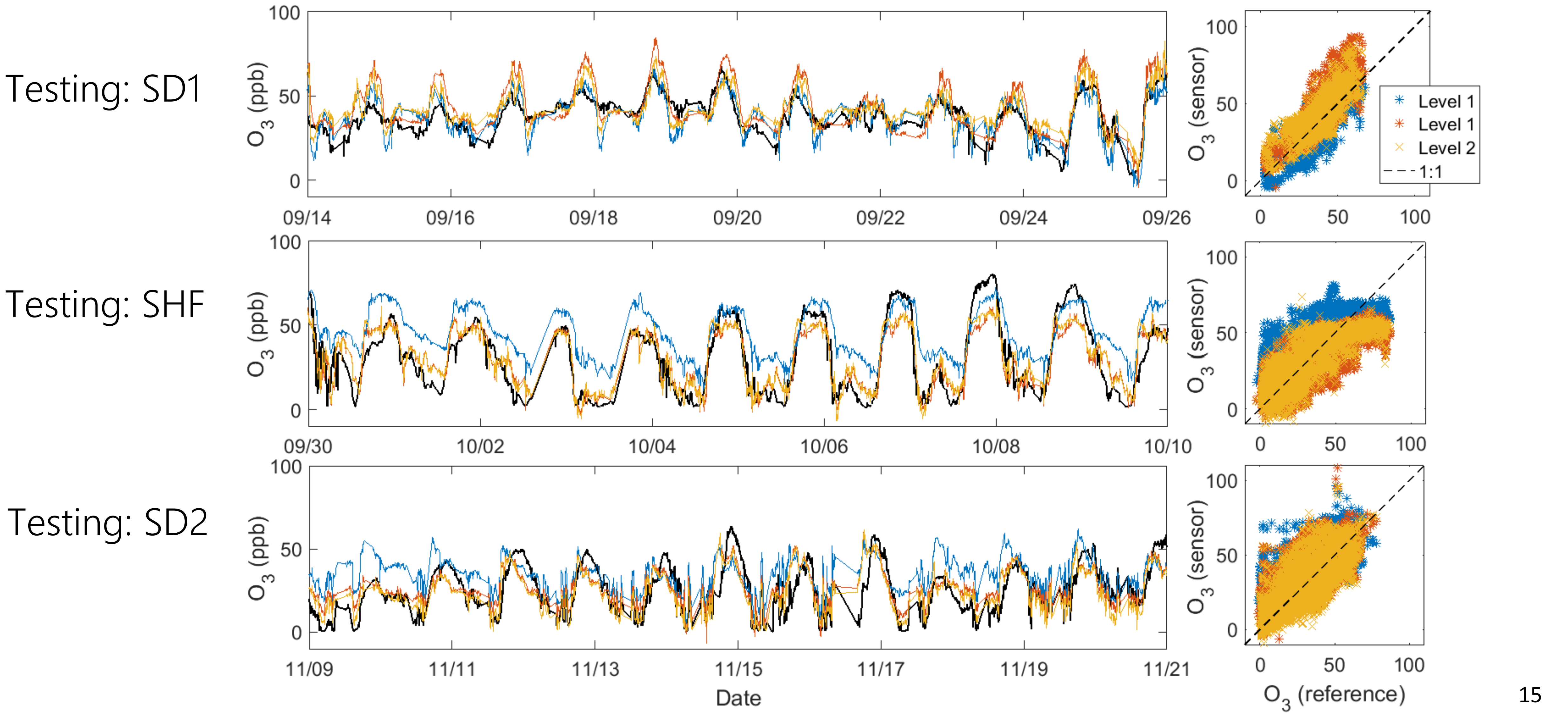
Comparing a calibration within the same city to calibrating in another part of the state

Observations:

- Over-predictions (particularly in MLR)
- Relatively large variance
- Underpredictions at high levels
- *MLR and NN(2) seem more robust across new locations*

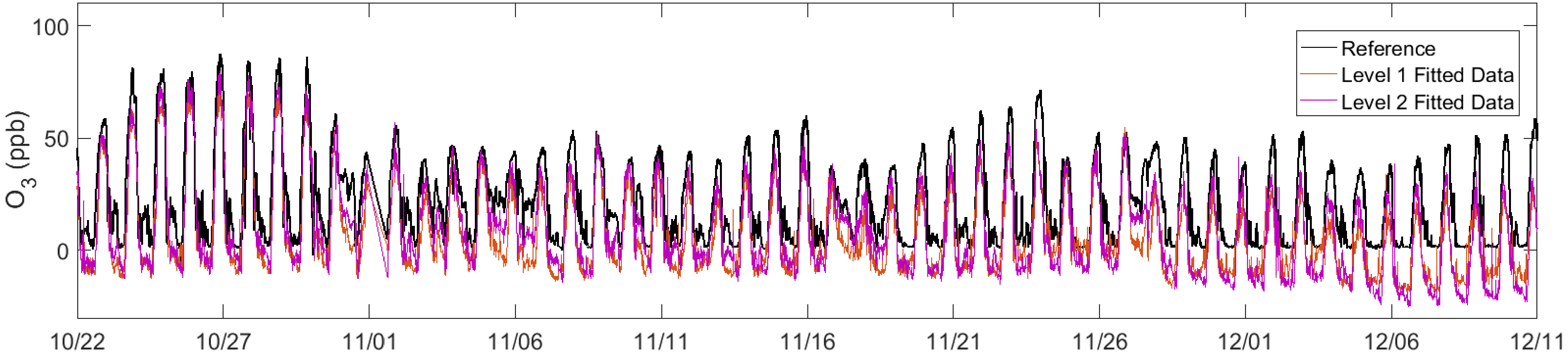


Adding a Second Calibration Location (Level 2)



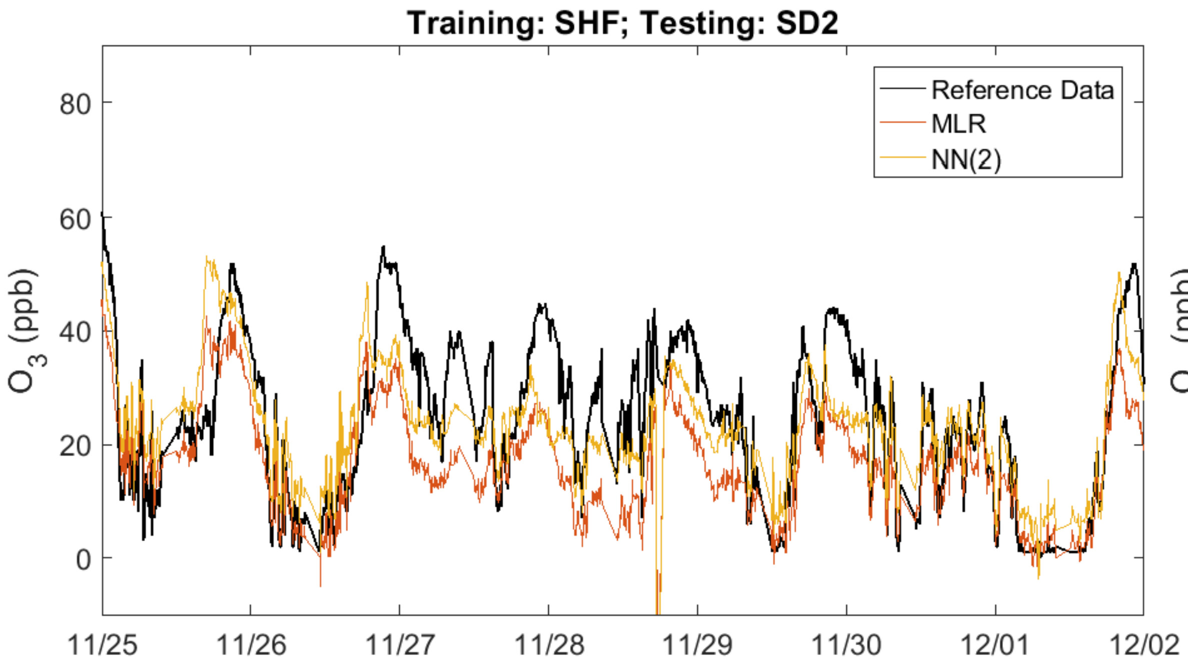
The Potential for Transferability

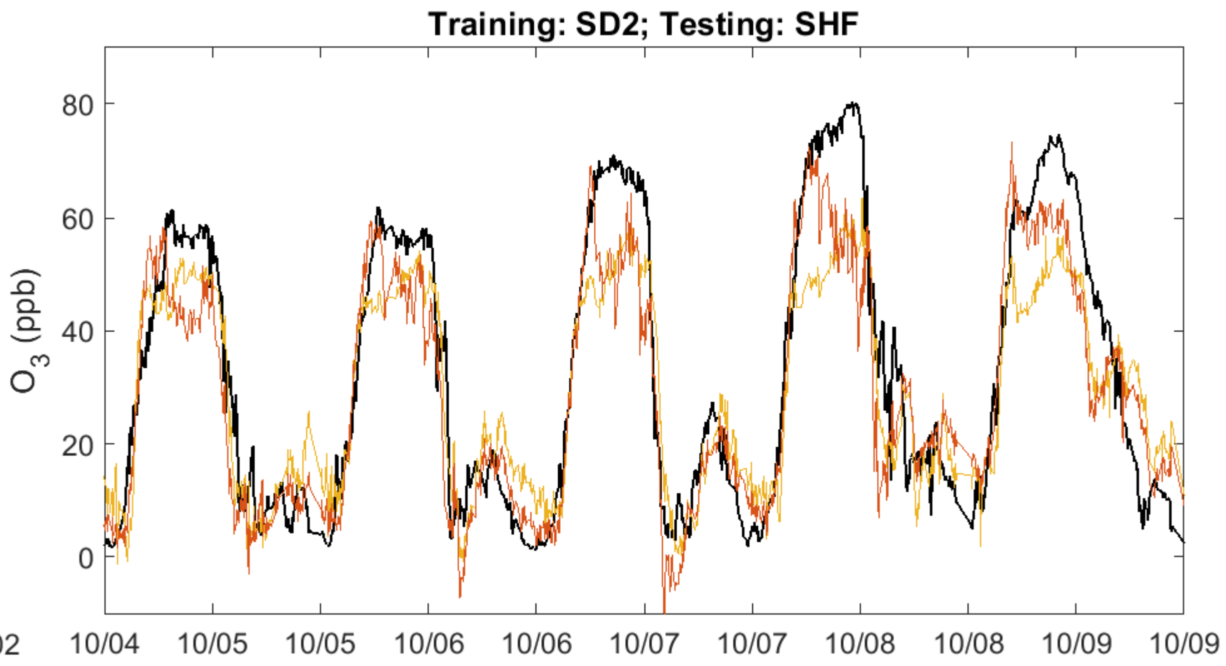
- Model: Neural Net (2-layer); Testing: Shafter (SHF)
- Level 1: R^2 = 0.76; RMSE = 18.1 ppb; Mean Bias = 15.3 ppb
- Level 2: R^2 = 0.82; RMSE = 15.9 ppb; Mean Bias = 13.3 ppb



The Potential for Transferability

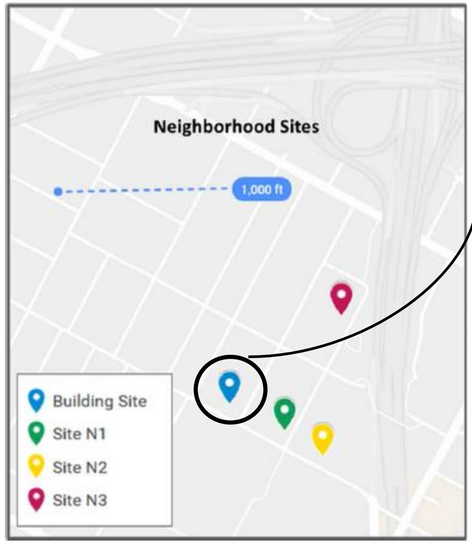
- Level 1 (training on one location and testing on a second location, using one site in San Diego and the Shafter site -> models indicative of overall ozone trends (differing)

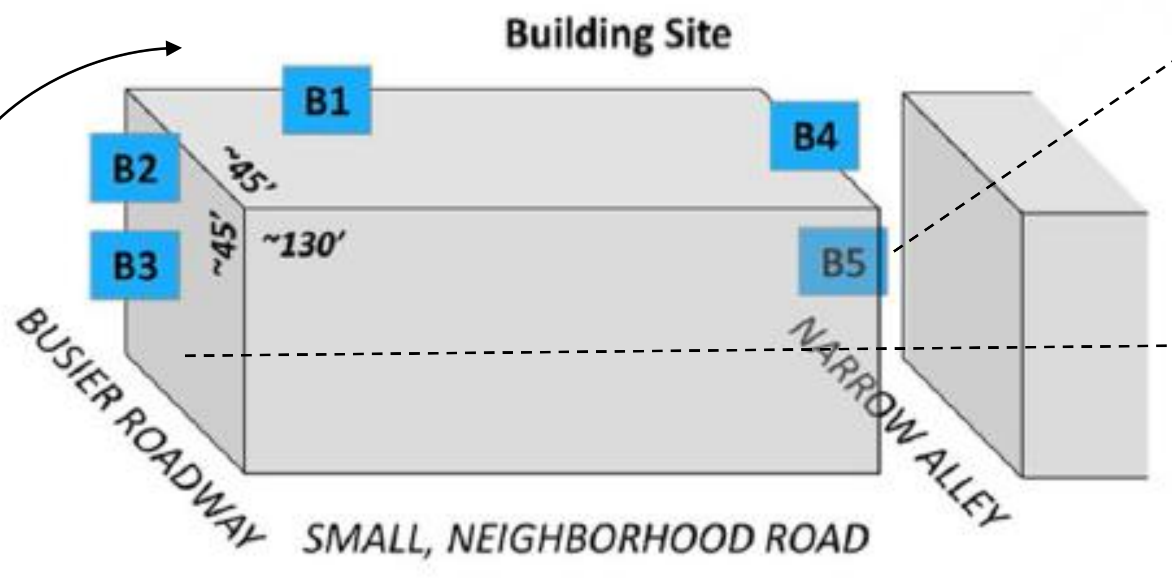




Sensor Deployments with Communities

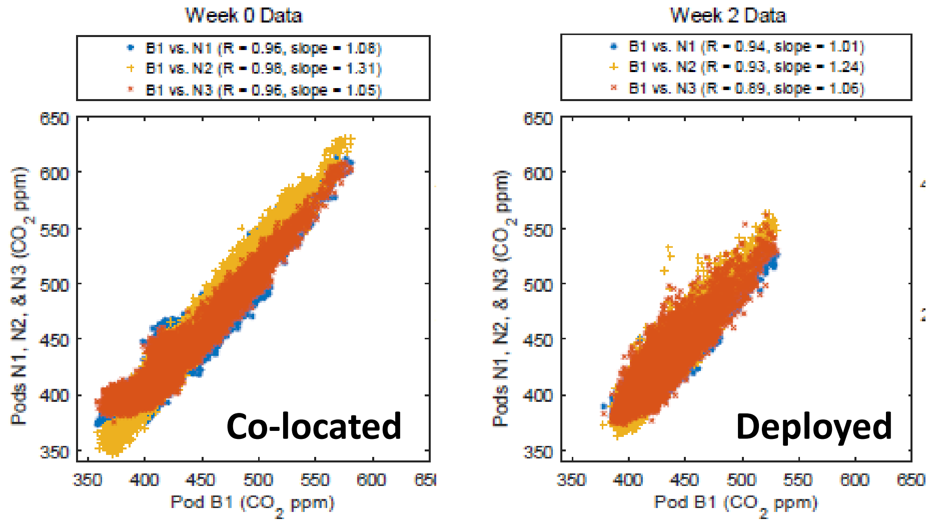
- Tools/resources such as MOUs valuable
- Need for more guidelines/best practices
  - Comparing building and neighborhood-scale variability
  - Five sensor systems, around a building; three additional systems on nearby buildings



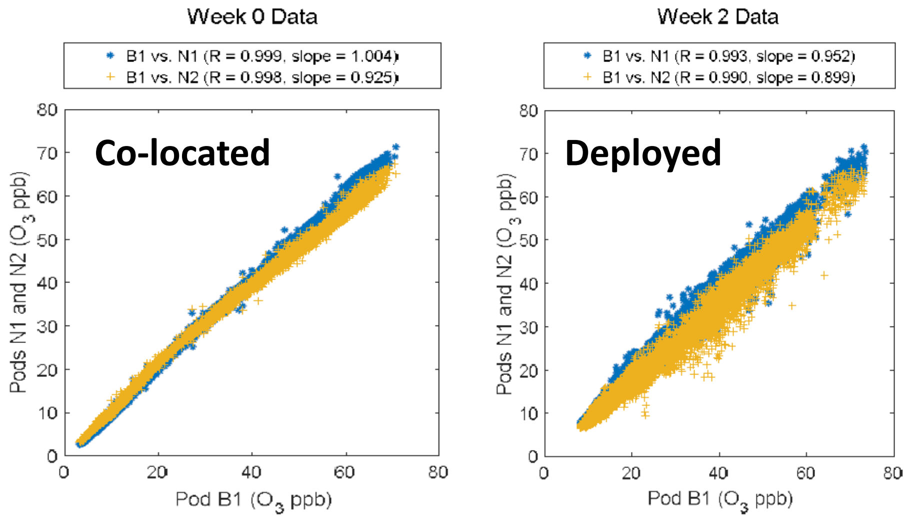


Building-Scale Variability – Observations

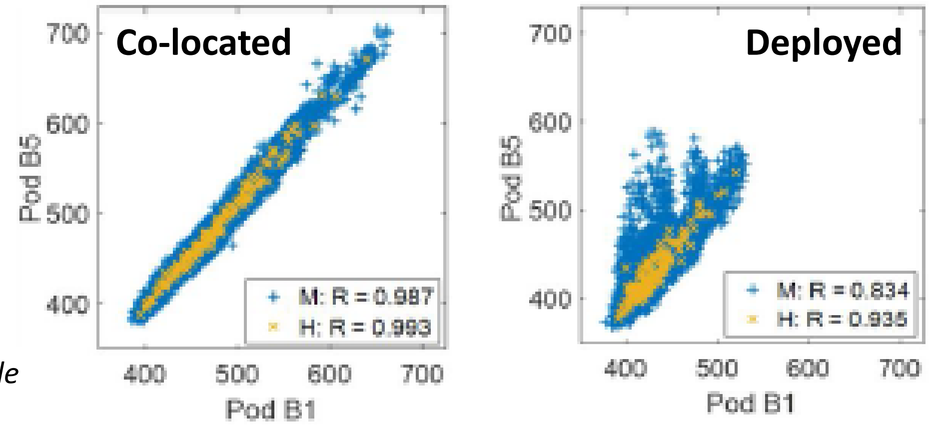
Across neighborhood sites (CO_2)



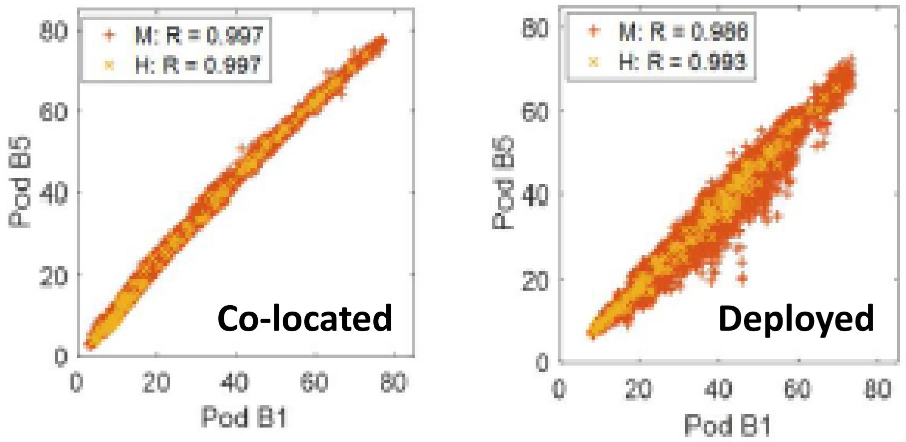
Across neighborhood sites (O_3)



Across paired building sites (CO_2)



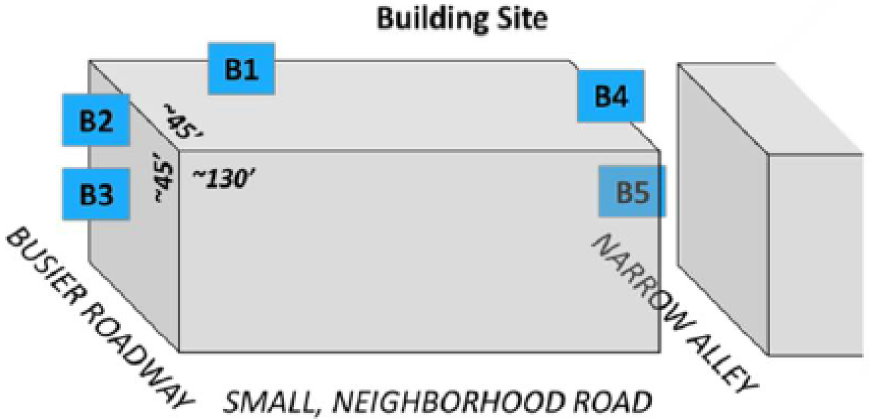
Across paired building sites (O_3)

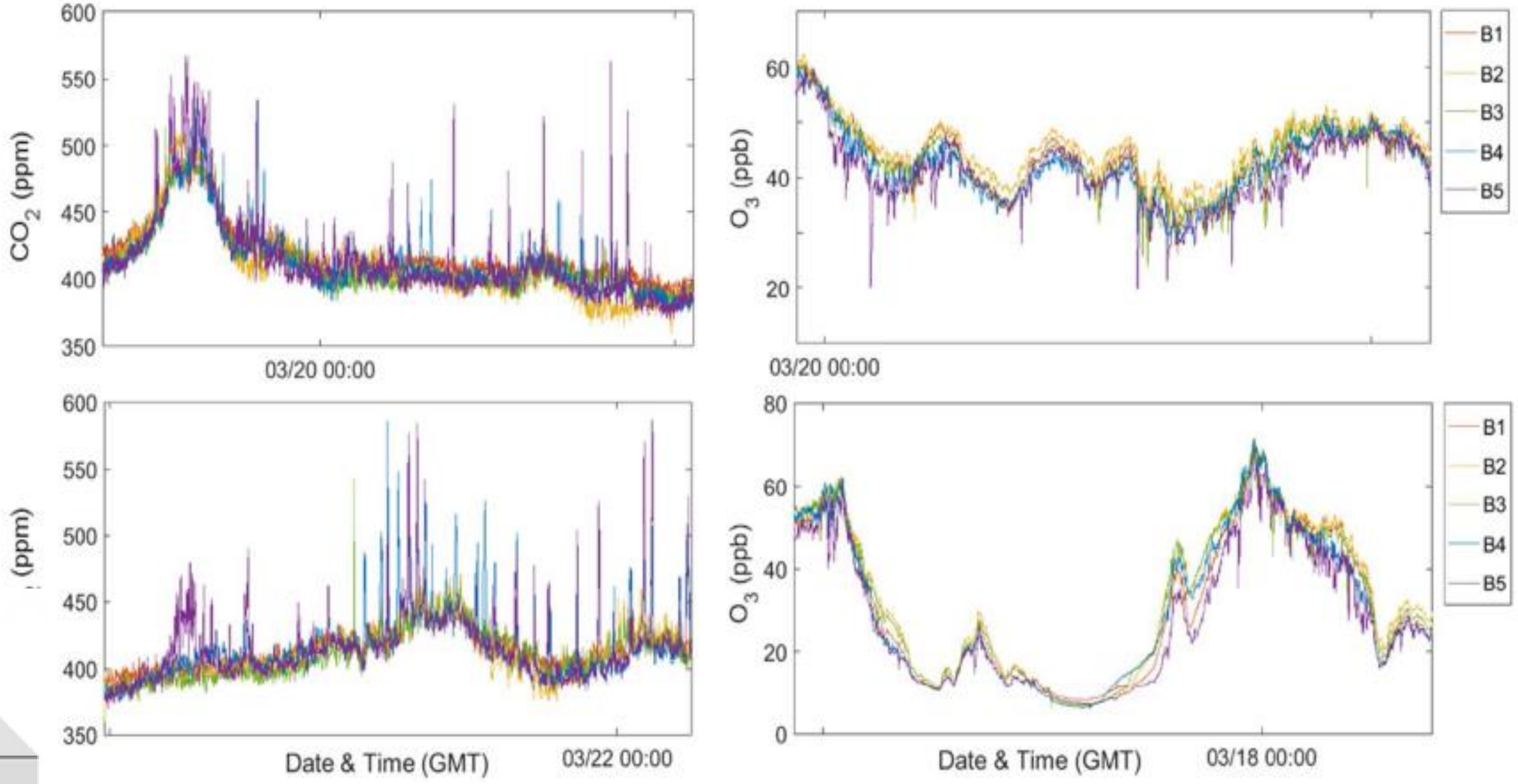


from Collier-Oxandale et al., 2018

Potential Implications for Siting

- Variability primarily driven by short-term enhancements (or depletions)
- Likely nearby emission sources

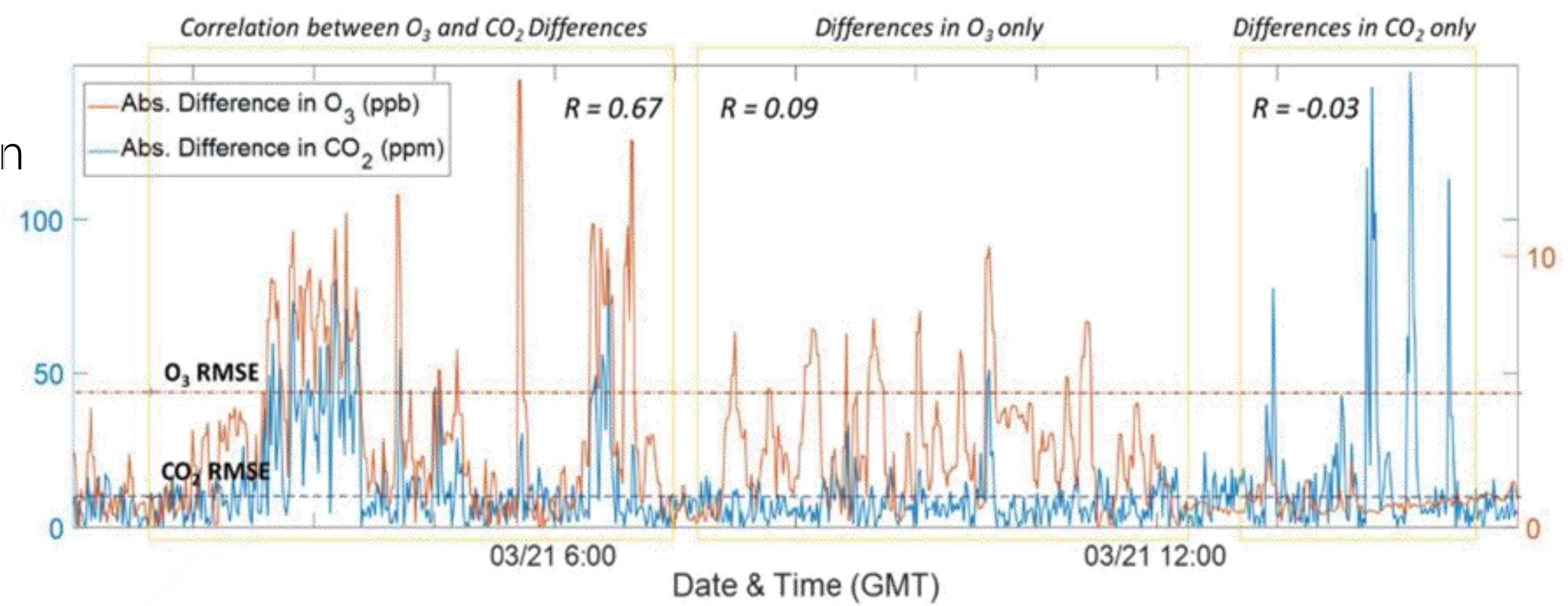


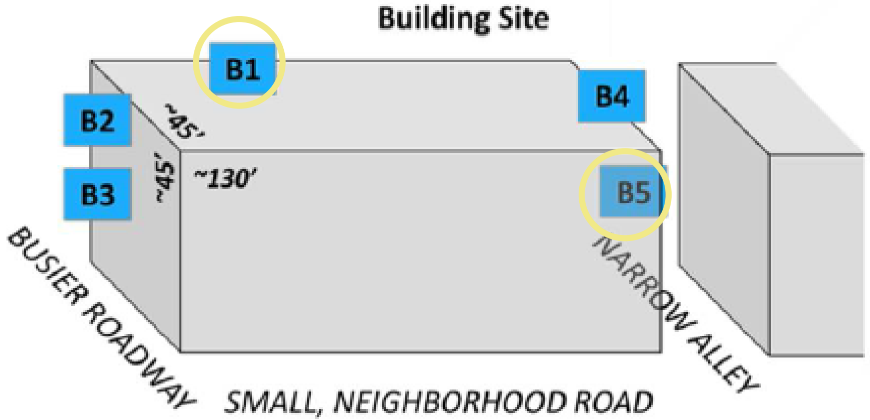


from Collier-Oxandale et al., 2018

Potential Implications for Siting

- Difference between sites B1 and B5
- Data suggests multiple sources responsible





from Collier-Oxandale et al., 2018

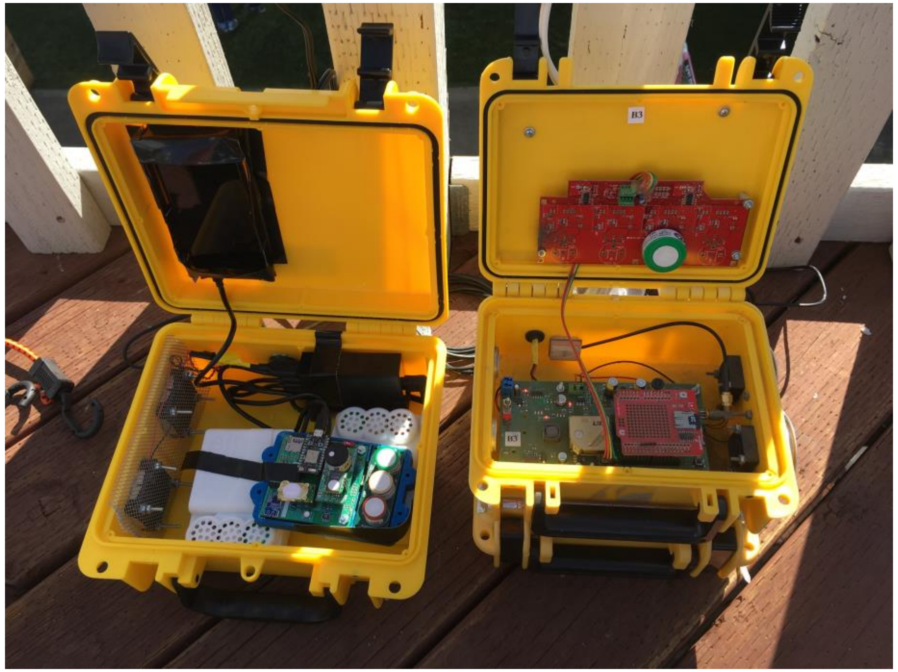
Conclusions

Regarding calibrations

- Training in 2 locations vs 1 can improve models, though at the cost of more effort
- MLR and NNs seem to provide more robust results, though at the cost of a higher uncertainty
- NEED: Methods for evaluating trustworthiness of data

Regarding siting

- Siting seems to be more important for high-time resolution data or near-source, neighborhood-scale studies
- NEED: More case studies and examples would help build broader situationally-specific knowledge



Best-practices can improve data quality thus better support community-based research

Acknowledgements

Co-authors: Sharad Vikram, William G. Griswold, Massimiliano Menarini, Michael Ostertag, Jill Johnston, Jacob Thorson, and Michael Hannigan

Hannigan Lab (CU Boulder, current and former): Joanna Gordon Casey, Evan Coffey, Kyle Karber, Ricardo Piedrahita, Jake Thorson, Nicholas Masson, Kira Sadighi, Drew Meyers, David John Pfotenhauer, Sarah Toth, and many others

MetaSense Project Partners (UC San Diego): William G. Griswold, Tajana Rosing, Sanjoy Dasgupta, Kevin Patrick, Massimiliano Menarini, Christine S. Chan, Sharad Vikram, and Michael Ostertag

Partners at USC: Jill Johnston and Wendy Gutschow

Community and other Collaborators: Nicole Wong (Redeemer Community Partnership), and Sandy Navarro, Esperanza Community Housing, William Flores, Bhavna Shamasunder (Occidental College), Sue Chiang & Michael Lucia (Center for Environmental Health) and Gus Aguirre, Hannah Halliday & the Native Trailer Team (Penn State)

Regulatory partners: San Diego Air Pollution Control District, San Joaquin Valley Air District & California Air Resources Board, South Coast Air Quality Management District, and Colorado Department Public Health and the Environment

Funding: MetaSense Project, NSF grant CNS-1446912, NSF-SRN AirWaterGas CBET: 1240584,