What is the Impact of Common Sources of Error on Air Quality LCS Measurements Performance?

ASIC 2022– 13th May 2022

Sebastian Diez, Stuart Lacy, Pete Edwards

Wolfson Atmospheric Chemistry Laboratories

University of York

Quantification of Utility of Atmospheric Network Technologies

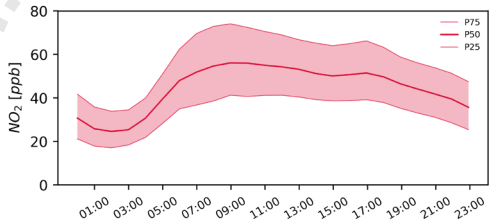
(QUANT, PI Dr Pete Edwards)

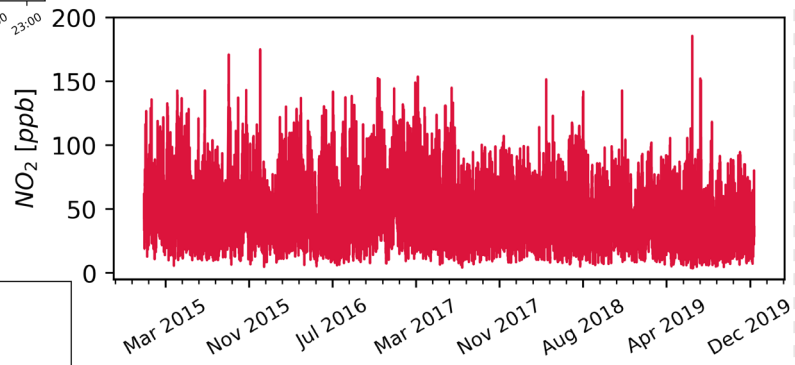
- 3 sites: Manchester, London & York
- ~3 years of data
- 13 brands
- 51 devices
- ~100 gas measurements
- ~200 PM measurements

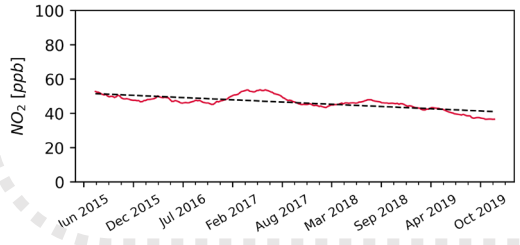
Credit: Dr Nicholas Mardsen

2 Sites' fingerprint (5yr, 1hr data)

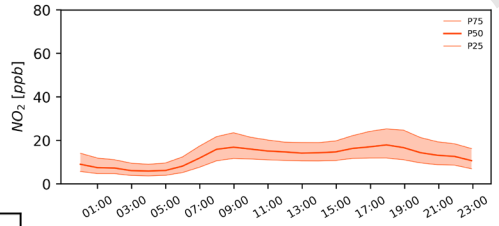
Site 1

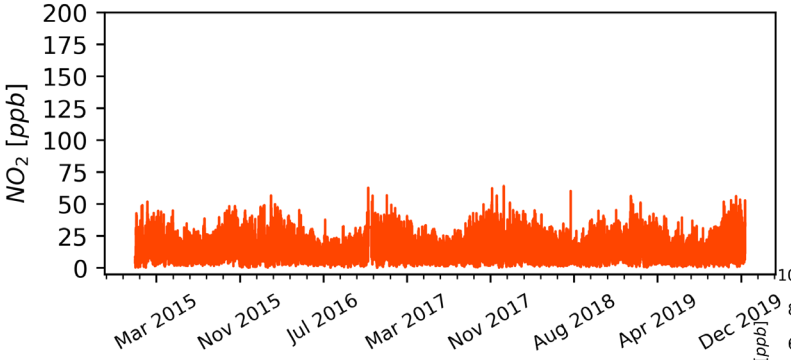


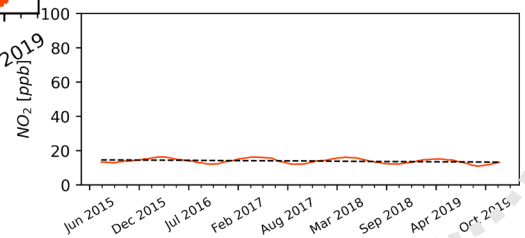




Site 2







Potential instruments to do the job

Instrument 1
R^2 = 0.95
RMSE = 4.5
Time res.: 1hr

Instrument 2
R^2 = 0.90
RMSE = 5.0
Time res.: 24hs

Instrument 3
R^2 = 0.85
RMSE = 5.5
Time res.: 30min

Instrument 4
R^2 = 0.80
RMSE = 5.0
Time res.: 15min

Instrument 5
R^2 = 0.75
RMSE = 4.5
Time res.: 1hr

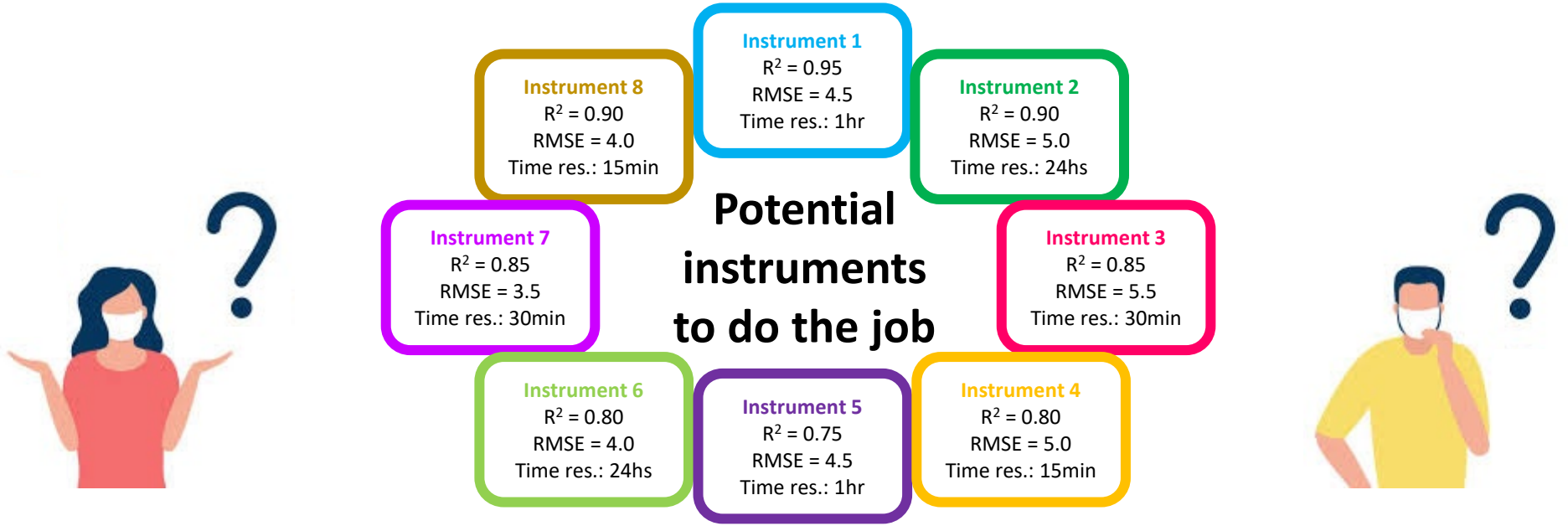
Instrument 6
R^2 = 0.80
RMSE = 4.0
Time res.: 24hs

Instrument 7
R^2 = 0.85
RMSE = 3.5
Time res.: 30min

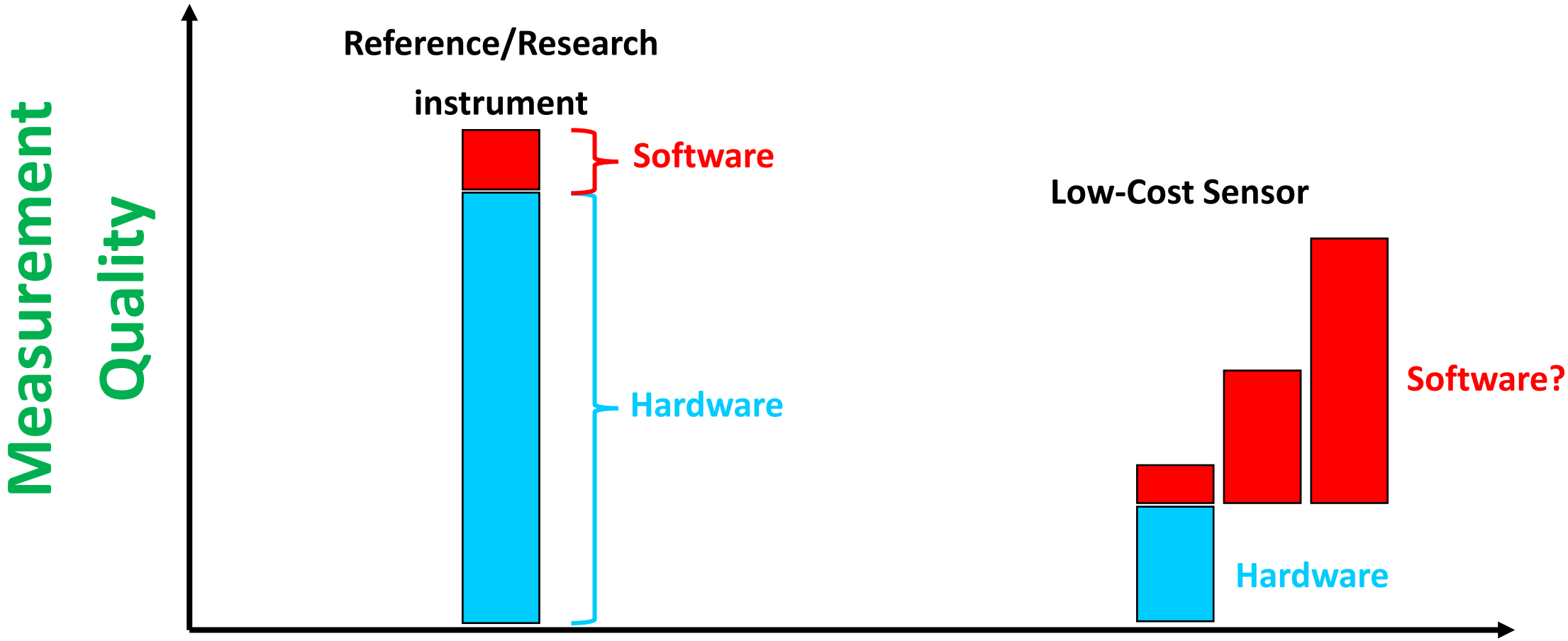
Instrument 8
R^2 = 0.90
RMSE = 4.0
Time res.: 15min

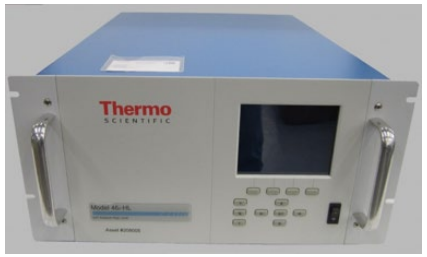
Information content

- Is it possible to know in advance how a LCS will perform?
- How do we now which LCS will perform best for a particular task?
- How the performance indices were obtained? Period? Site?
- And more importantly, are they comparable?

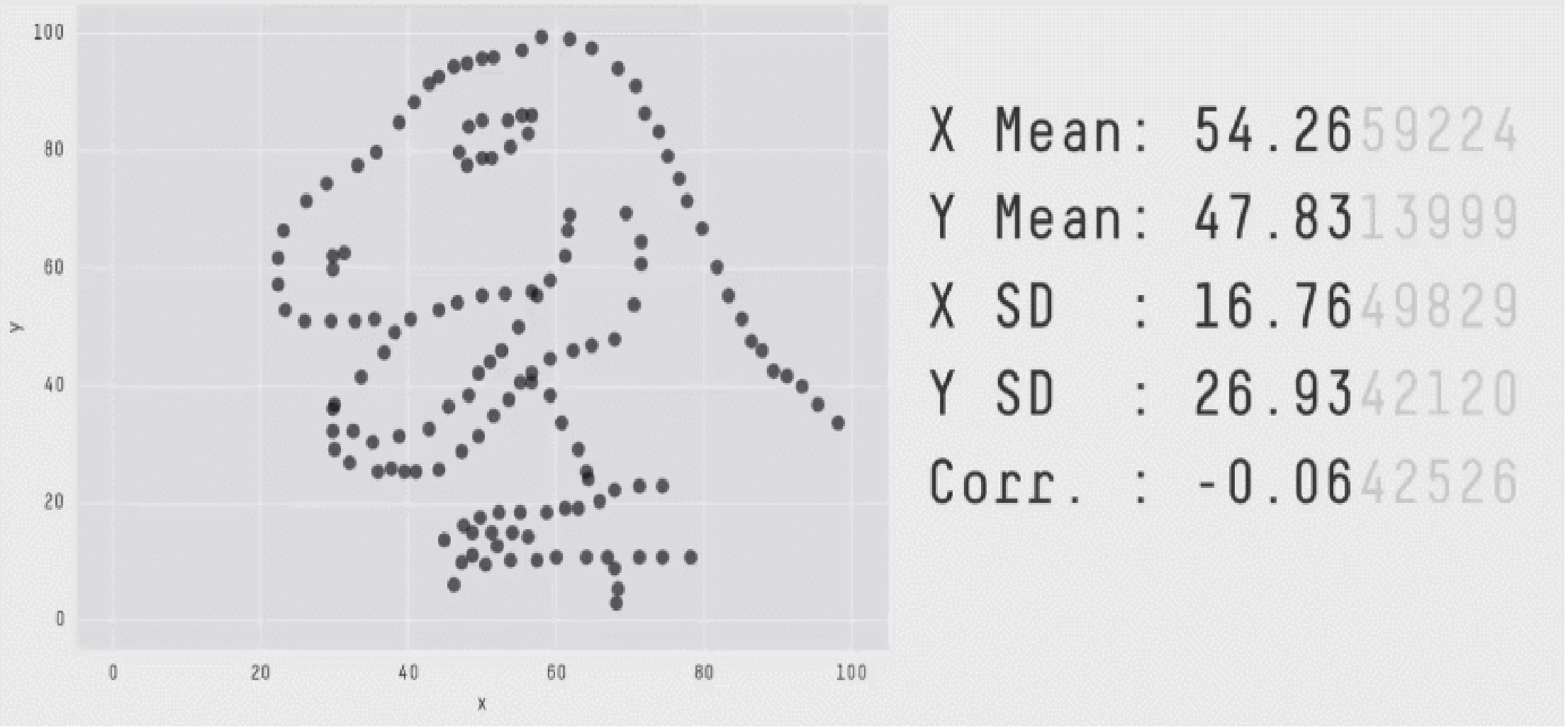


Measurement Quality = f(Hardware, Software)



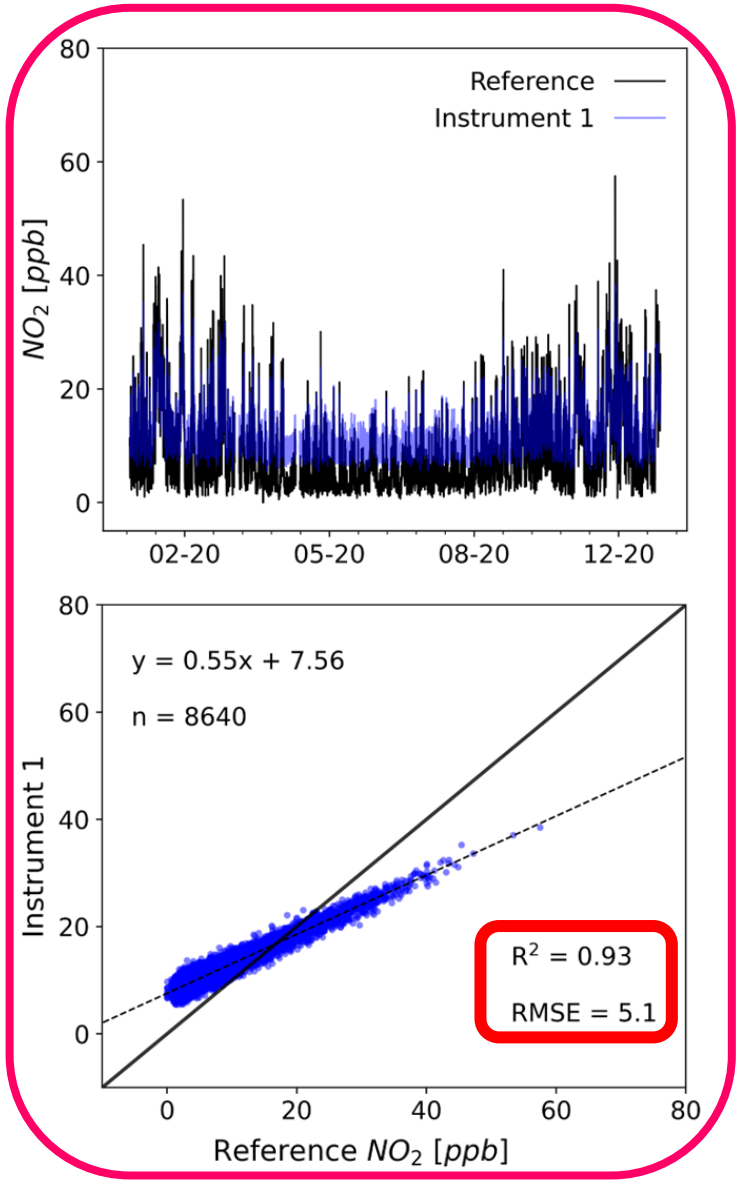


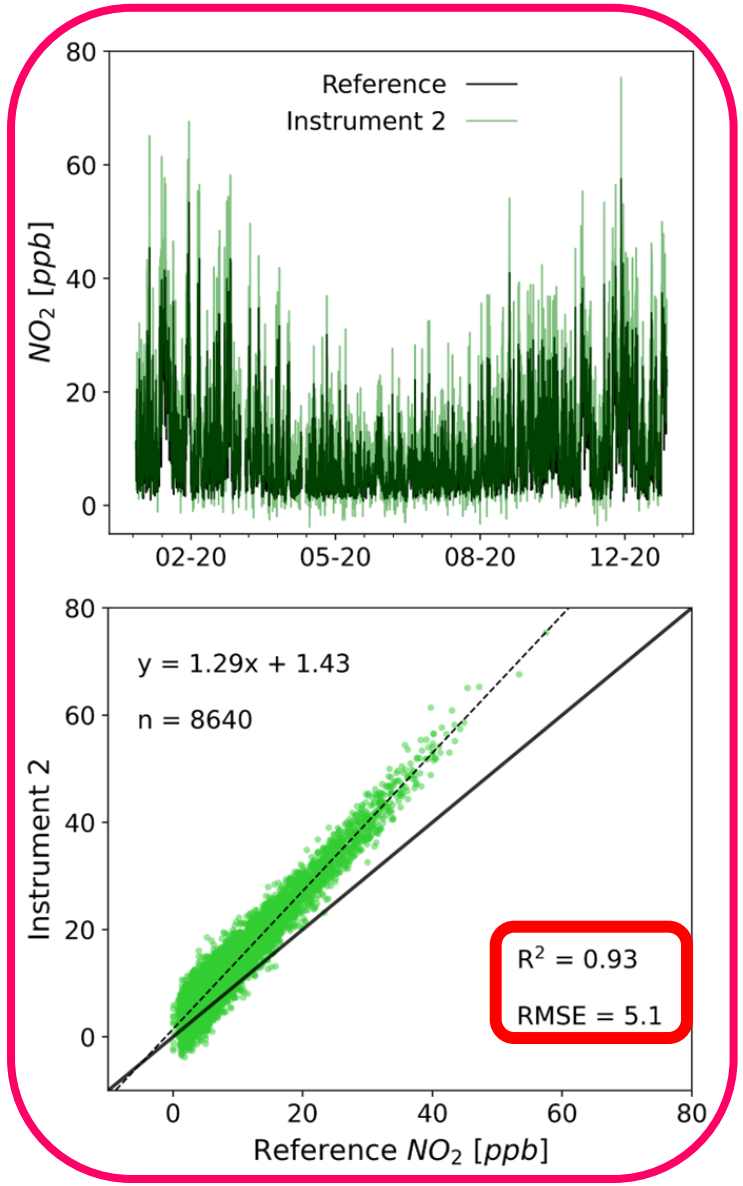
Single value metrics: The Datasaurus

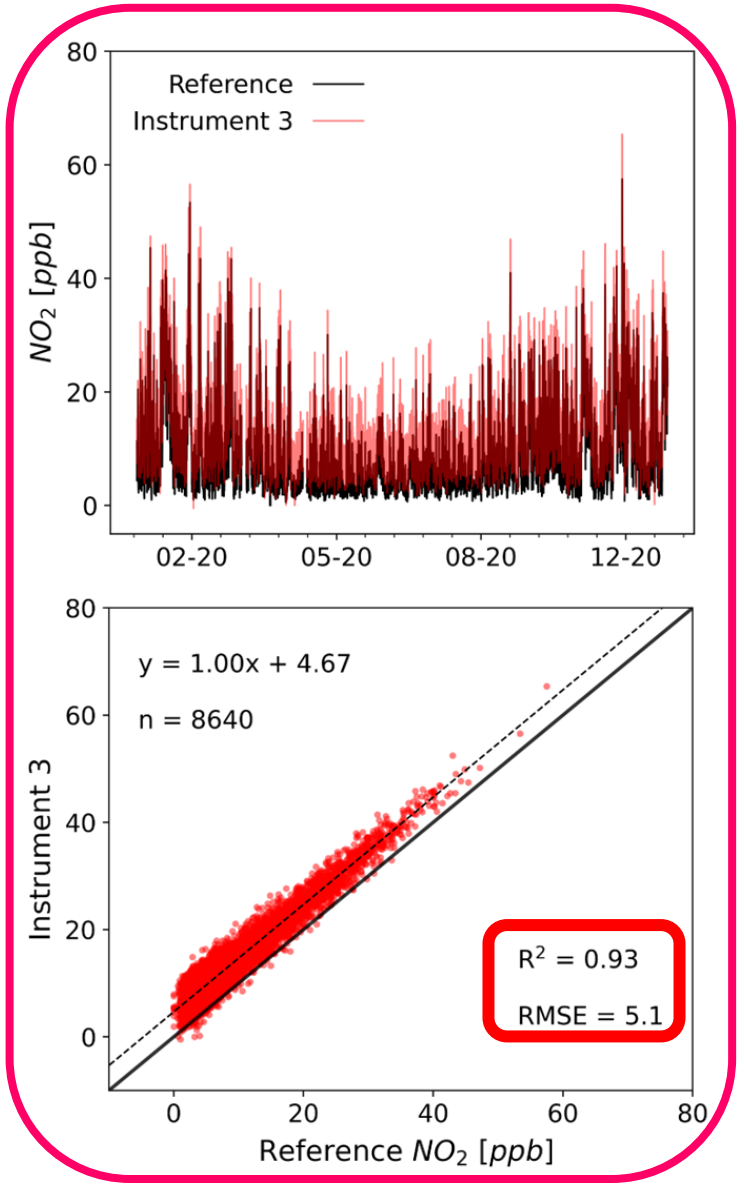


Same Stats, Different Graphs:
Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing
(Matejka & Fitzmaurice, 2017)

Single value metrics: different error structure, same R^2 and RMSE







Methods

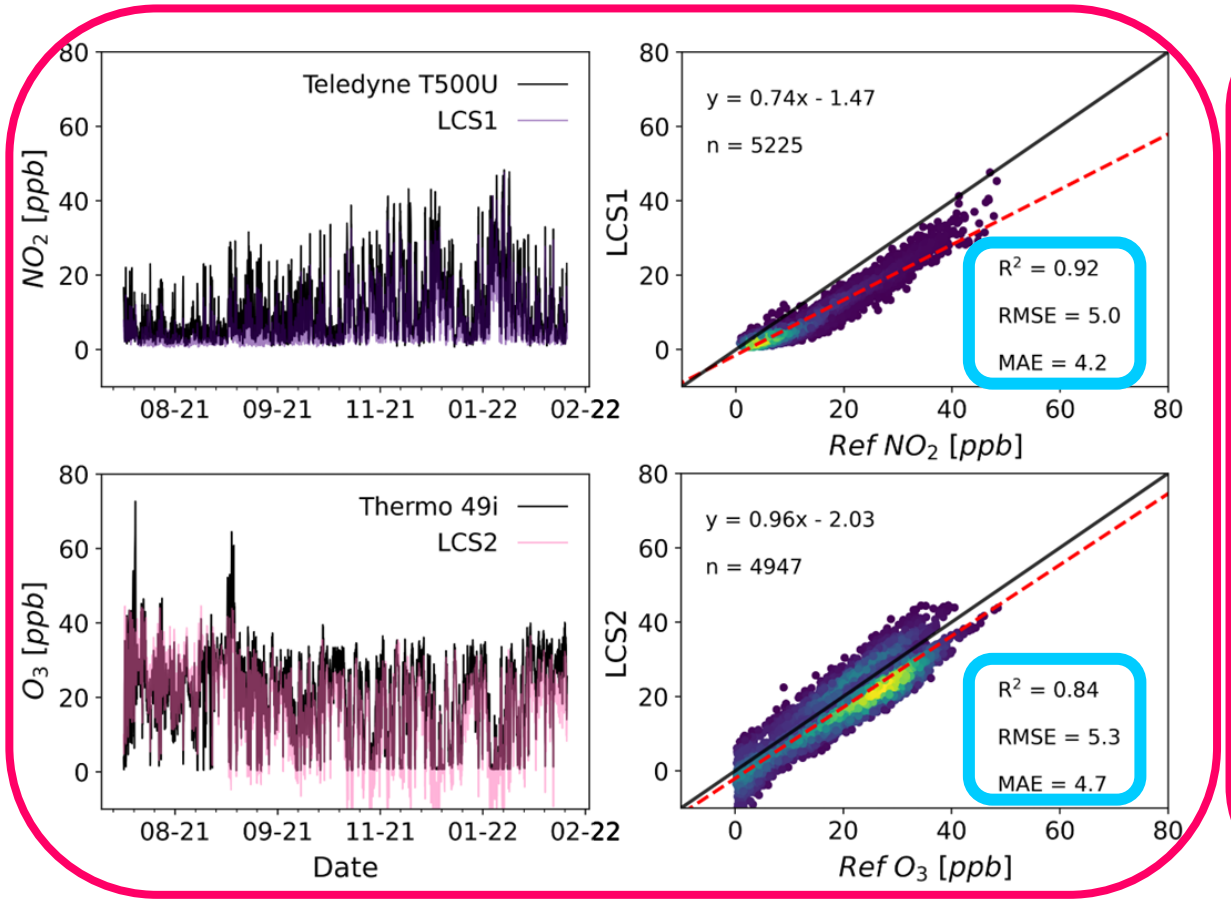
“Real-world” measurements:

- Duplicate QUANT sensors: PM2.5, NO_2 and O_3
- Duplicate Ref grade instruments: NO_2 and O_3

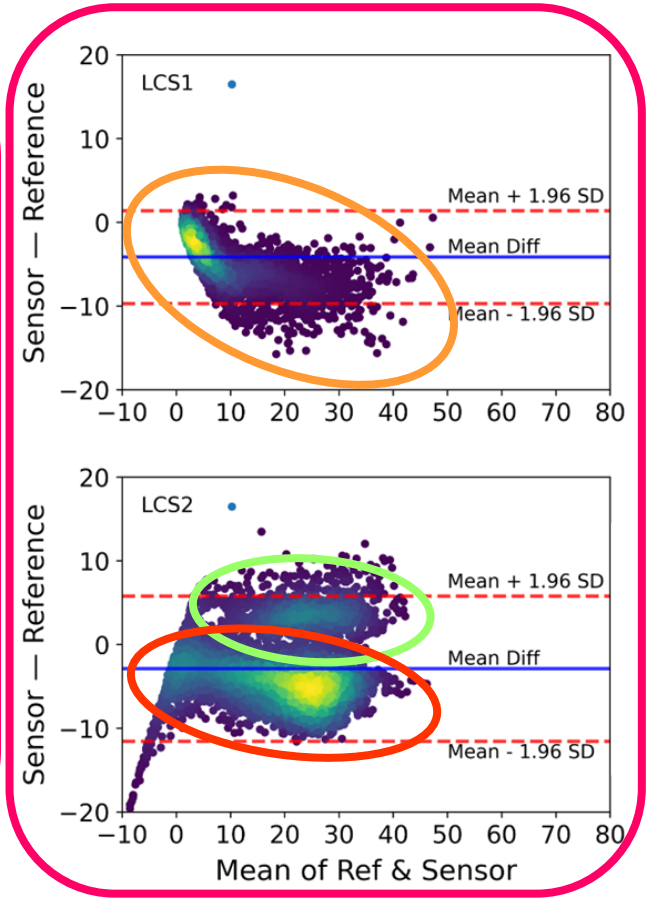
Tools:

- Single-value metrics
- Time-series plots
- Regression plots
- Bland-Altman plots
- Relative Expanded Uncertainty plots

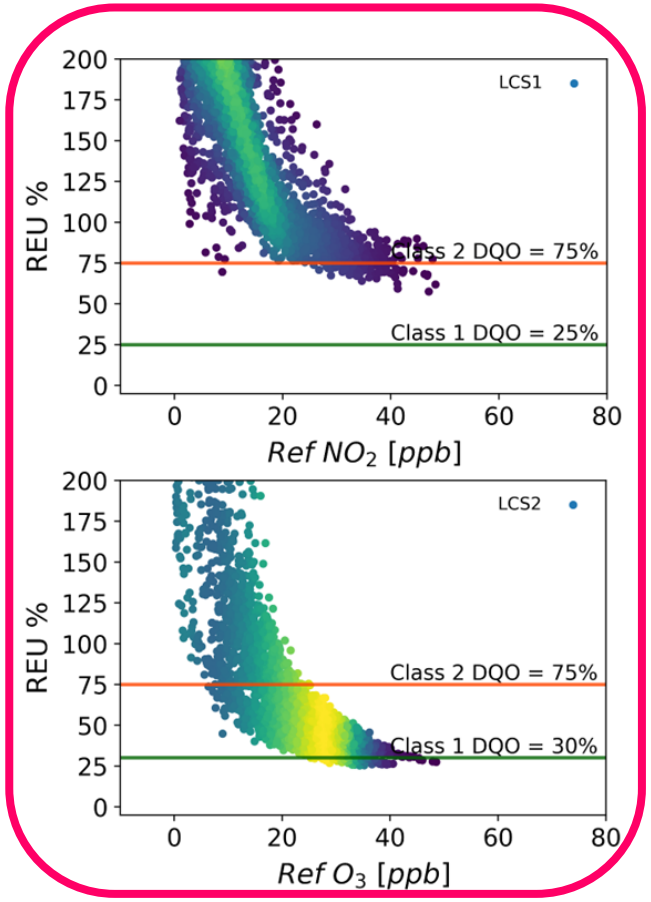
“Real-world” instruments: LCS NO_2 and O_3



Time series & regression plots

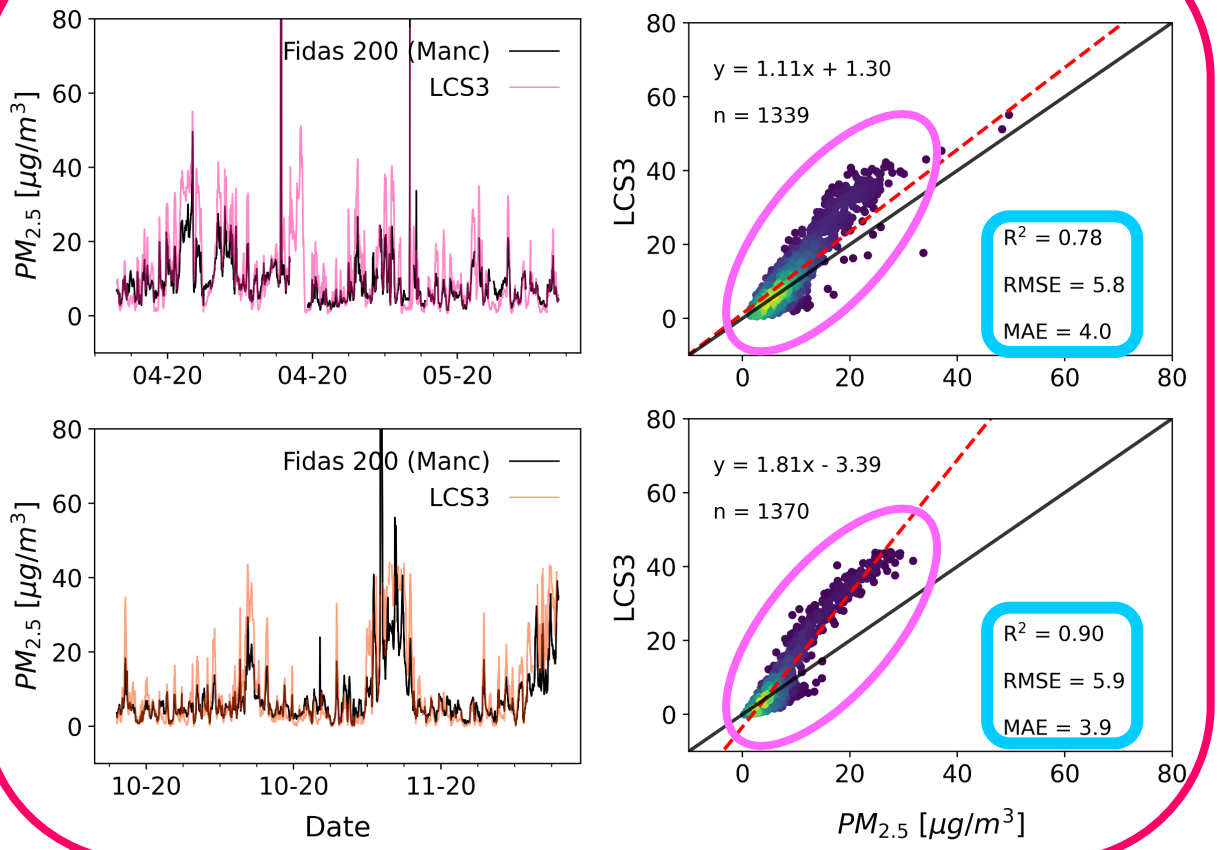


Bland-Altman plots

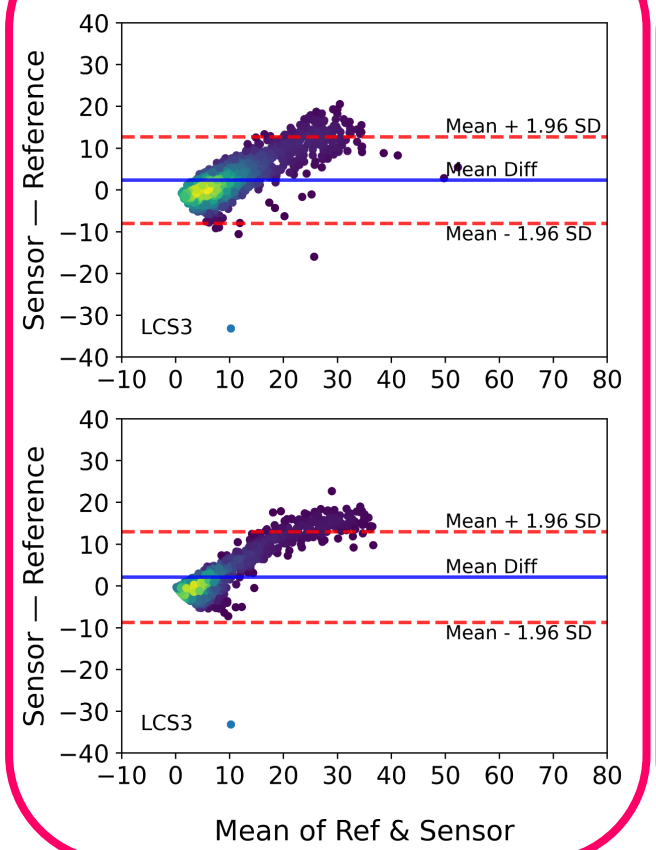


REU plots

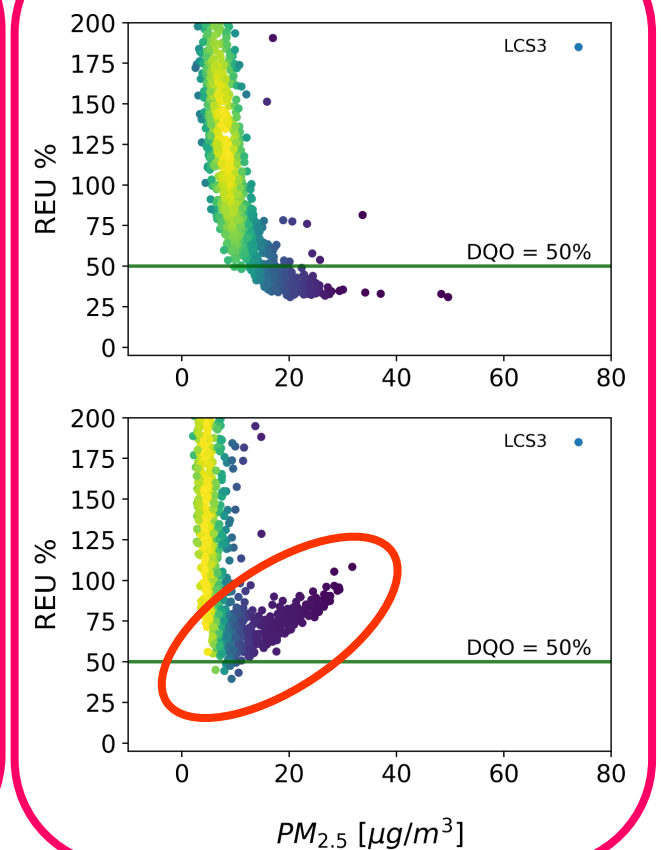
“Real-world” instruments: LCS $PM_{2.5}$



Time series & regression plots

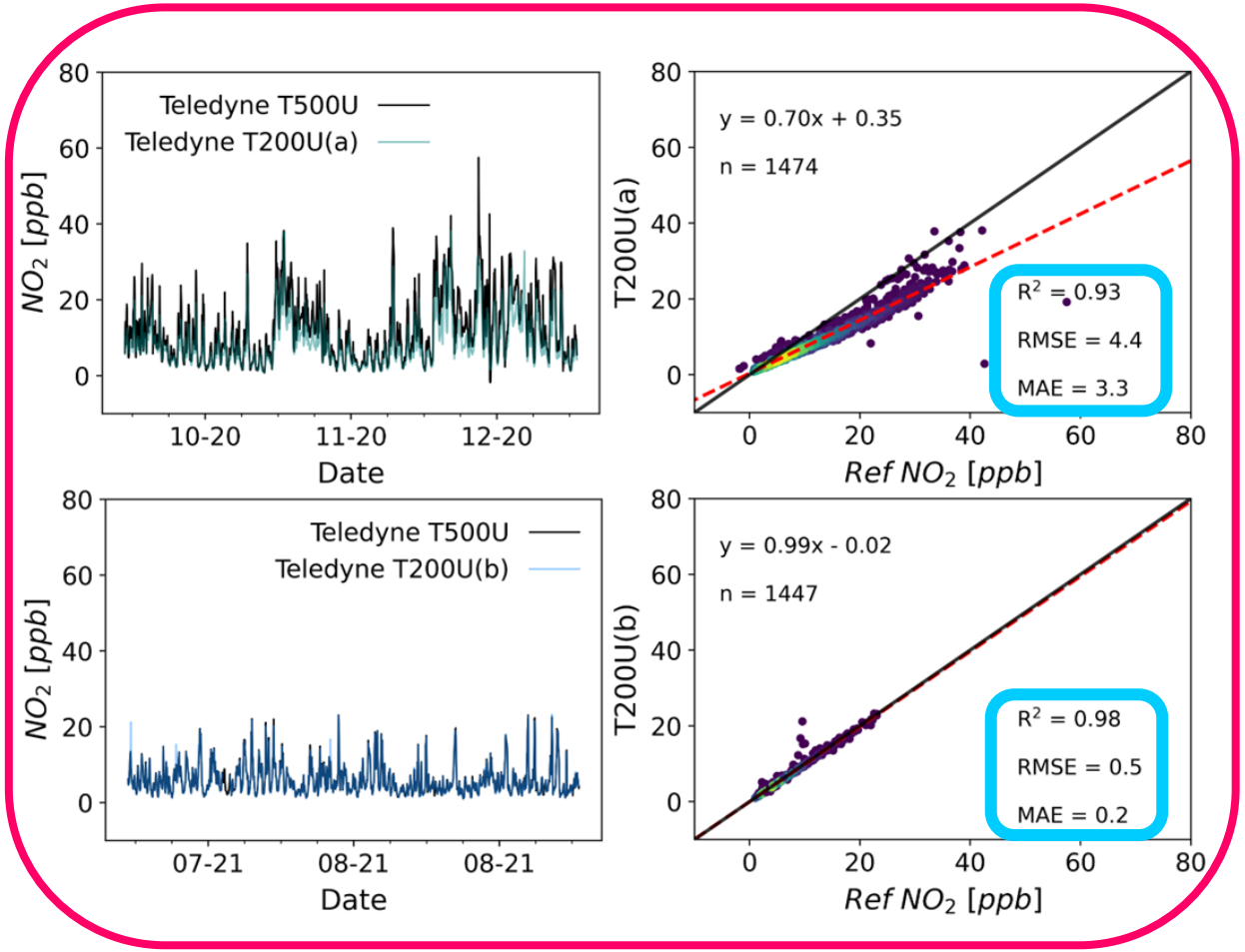


Bland-Altman plots

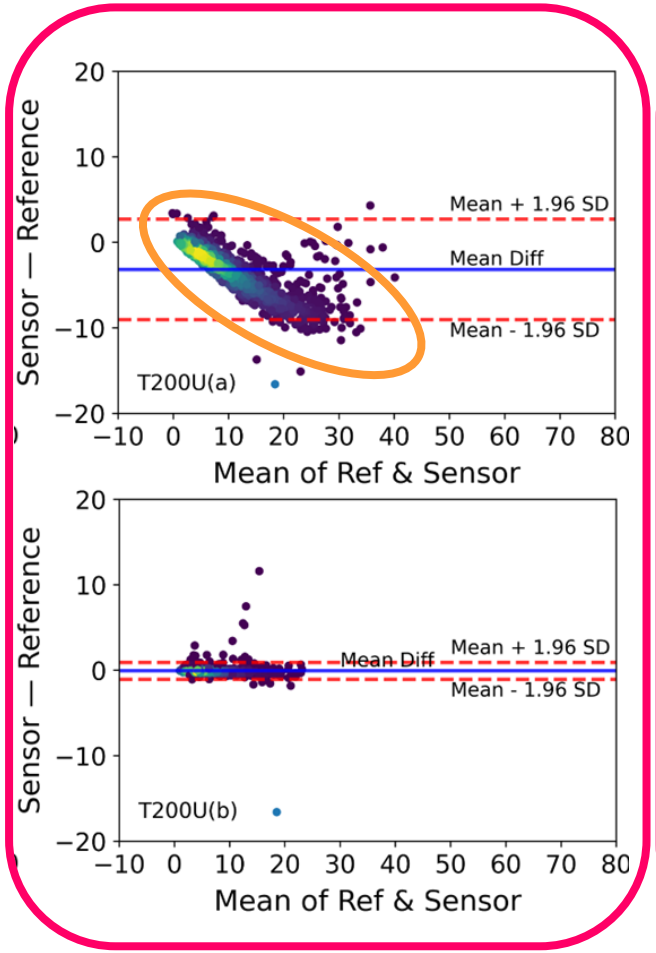


REU plots

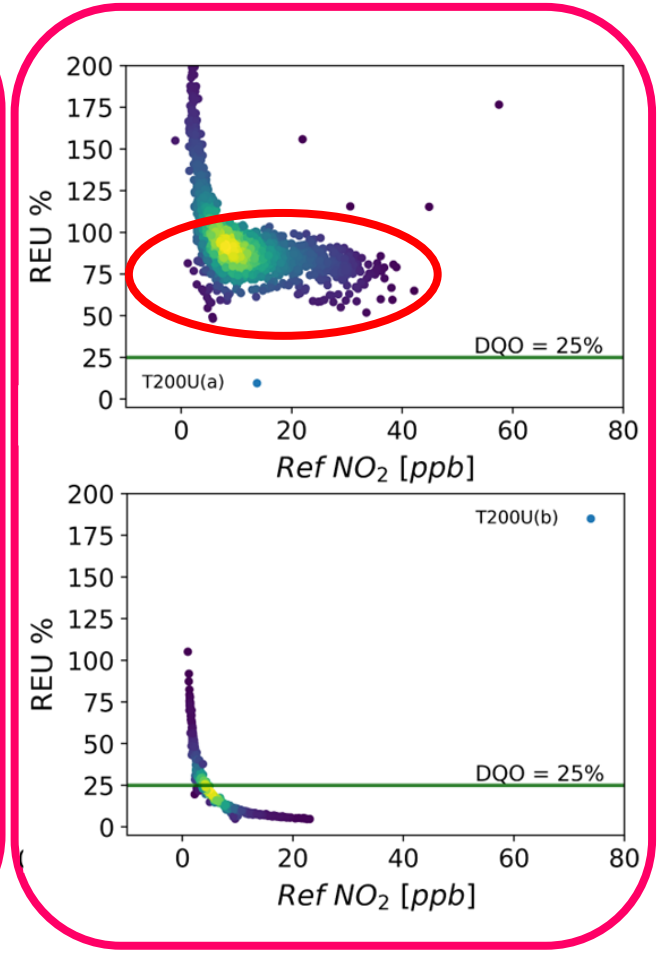
“Real-world” instruments: NO_2 Reference Grade



Time series & regression plots



Bland-Altman plots



REU plots

Take-home messages

- End users need to **clearly identify** the data requirements. What is the question you want to ask?
- **Single-value metrics** convey very useful information…but take proper care as they **can be misleading**
- **Plot your data!!!...in as many ways as possible**
- **Visualizing the data** enables the user to make **more informed decisions** about the LCS capabilities in the target range
- As **error sources** can **change** significantly for different **sites/seasons**, the LCS should ideally be tested **relevant environments**

Gracias!

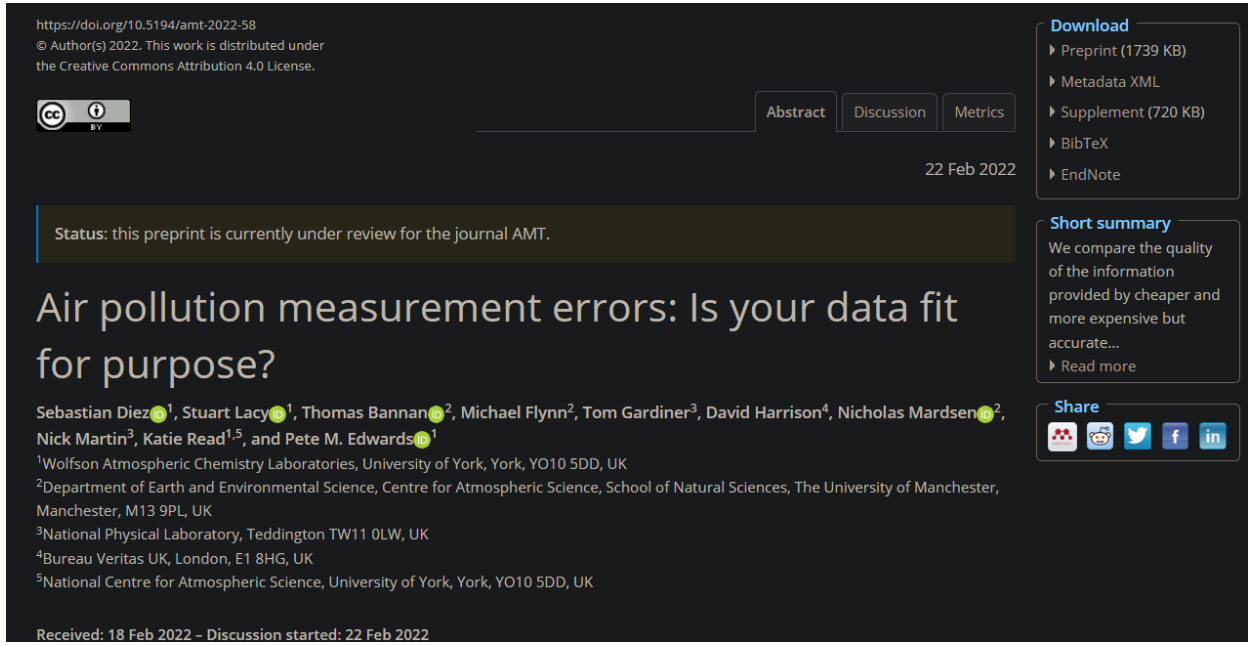
Sebastian Diez*

sebastian.diez@york.ac.uk

Pete Edwards

pete.edwards@york.ac.uk

Extras



https://doi.org/10.5194/amt-2022-58
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Abstract | Discussion | Metrics

22 Feb 2022

Status: this preprint is currently under review for the journal AMT.

Air pollution measurement errors: Is your data fit for purpose?

Sebastian Diez[1], Stuart Lacy[1], Thomas Bannan[2], Michael Flynn[2], Tom Gardiner[3], David Harrison[4], Nicholas Mardsen[2], Nick Martin[3], Katie Read[1,5], and Pete M. Edwards[1]

[1]Wolfson Atmospheric Chemistry Laboratories, University of York, York, YO10 5DD, UK
[2]Department of Earth and Environmental Science, Centre for Atmospheric Science, School of Natural Sciences, The University of Manchester, Manchester, M13 9PL, UK
[3]National Physical Laboratory, Teddington TW11 0LW, UK
[4]Bureau Veritas UK, London, E1 8HG, UK
[5]National Centre for Atmospheric Science, University of York, York, YO10 5DD, UK

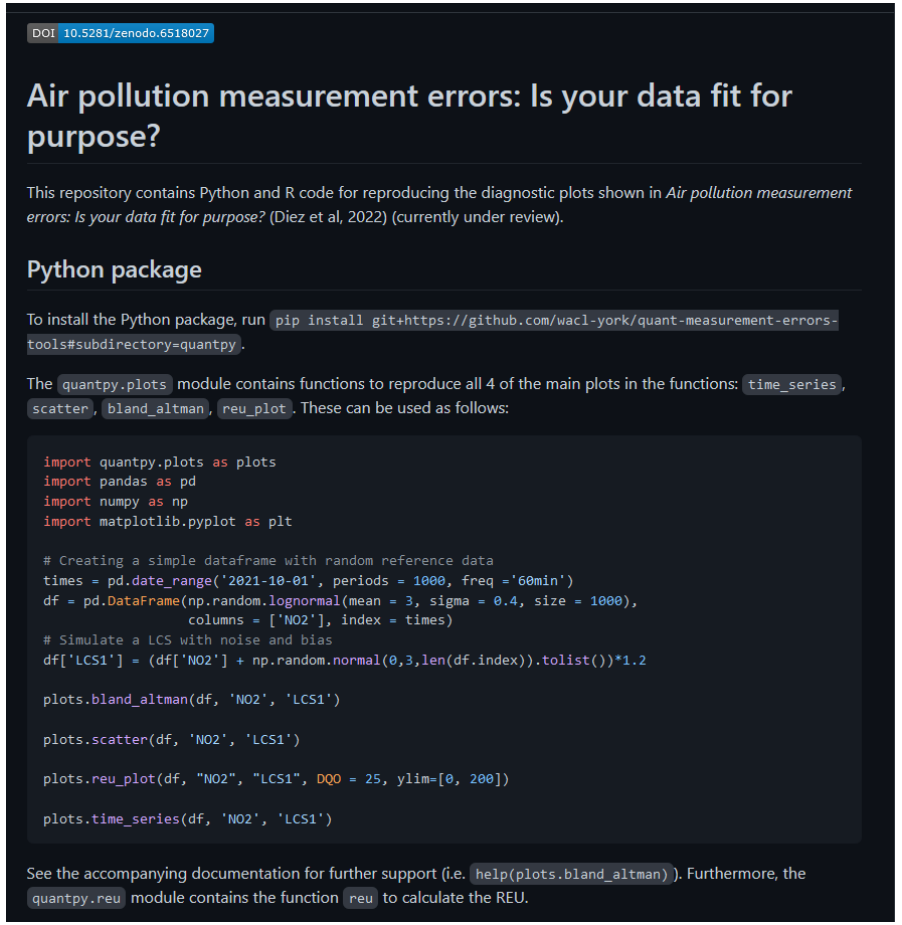
Received: 18 Feb 2022 – Discussion started: 22 Feb 2022

Download
- Preprint (1739 KB)
- Metadata XML
- Supplement (720 KB)
- BibTeX
- EndNote

Short summary
We compare the quality of the information provided by cheaper and more expensive but accurate...
- Read more

Share

https://amt.copernicus.org/preprints/amt-2022-58/



DOI 10.5281/zenodo.6518027

Air pollution measurement errors: Is your data fit for purpose?

This repository contains Python and R code for reproducing the diagnostic plots shown in *Air pollution measurement errors: Is your data fit for purpose?* (Diez et al, 2022) (currently under review).

Python package

To install the Python package, run `pip install git+https://github.com/wacl-york/quant-measurement-errors-tools#subdirectory=quantpy`.

The `quantpy.plots` module contains functions to reproduce all 4 of the main plots in the functions: `time_series`, `scatter`, `bland_altman`, `reu_plot`. These can be used as follows:

```
import quantpy.plots as plots
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Creating a simple dataframe with random reference data
times = pd.date_range('2021-10-01', periods = 1000, freq ='60min')
df = pd.DataFrame(np.random.lognormal(mean = 3, sigma = 0.4, size = 1000),
                  columns = ['NO2'], index = times)
# Simulate a LCS with noise and bias
df['LCS1'] = (df['NO2'] + np.random.normal(0,3,len(df.index)).tolist())*1.2

plots.bland_altman(df, 'NO2', 'LCS1')

plots.scatter(df, 'NO2', 'LCS1')

plots.reu_plot(df, "NO2", "LCS1", DQO = 25, ylim=[0, 200])

plots.time_series(df, 'NO2', 'LCS1')
```

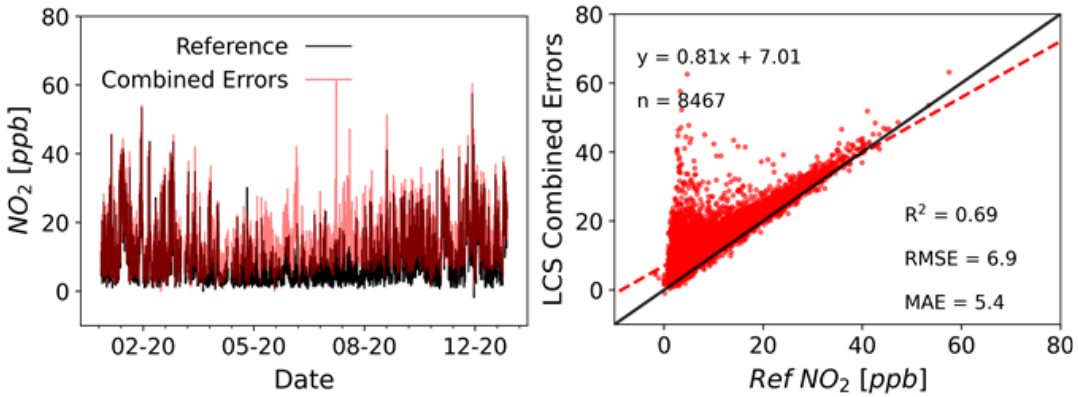
See the accompanying documentation for further support (i.e. `help(plots.bland_altman)`). Furthermore, the `quantpy.reu` module contains the function `reu` to calculate the REU.

https://github.com/wacl-york/quant-air-pollution-measurement-errors

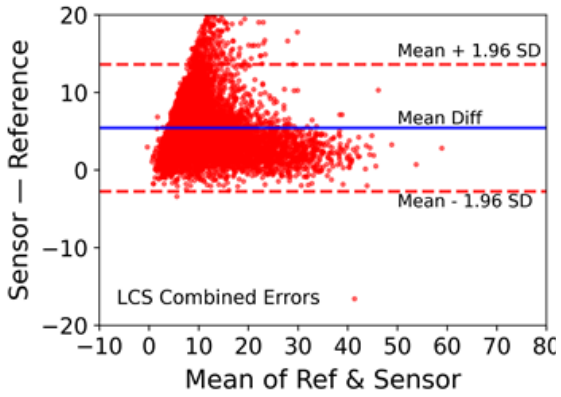
Simulated instruments: NO_2

Linear combination of arbitrary interferences:

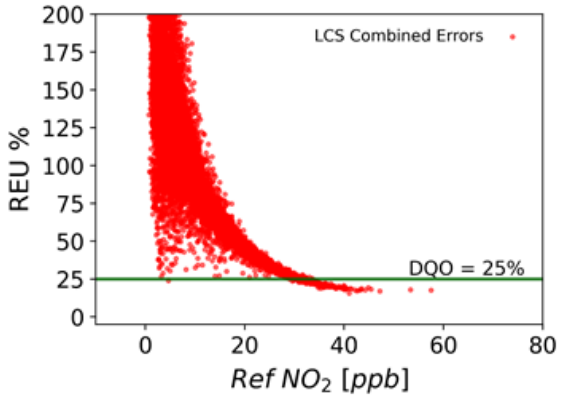
- temperature
- ozone
- electrical noise



Time series & regression plots



Bland-Altman plots



REU plots

Session Q&A Discussion

Please submit your questions for the session speakers through Whova – on your mobile or desktop device.

Make sure to note WHOM your question should be addressed to.

ASIC

Thank you for joining this session.

Please head over to the Exhibit Hall to enjoy a light, final-day lunch! During the final 20 minutes of the lunch break we will host another raffle with prizes supplied by exhibitors, so be sure to stick around!

Please note, the Exhibit Hall will close at 1:10 PM.

Don’t forget about this afternoon’s activities!

We will have a Champagne Toast in Ballroom DE at 3:30 PM.

And at 4:30 PM, please join us for the International Connections Happy Hour, located at the Hilton Hotel, 168 South Los Robles Ave.

ASIC