Quantifying nitrogen oxides and ammonia via frequency modulation in gas sensors

Master Thesis defense seminar

Marcos F Mourão

June 1^{st} 2021

Outline

But first...

A few seconds of gratitude

- Annika Tillander - Internal Supervisor
- Mike Andersson - External Supervisor
- José M. Peña - Examiner
- Oleg Sysoev - Course Leader
- Samia Noreen Butt - Opponent

Thank you!

Problem in a nutshell

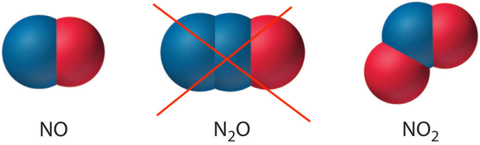
Motivation

[1]Image source: ENVIS Centre on Plants and Pollution

Problem in a nutshell

Motivation

NO_x[1]:

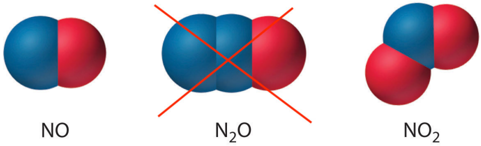


[1] Image source: ENVIS Centre on Plants and Pollution

Problem in a nutshell

Motivation

NO_x[1]:



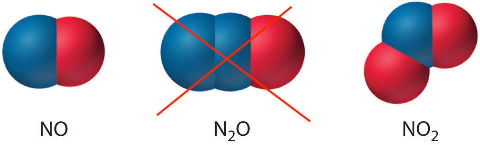
- NO_x are detrimental to the environment and humans.

[1] Image source: ENVIS Centre on Plants and Pollution

Problem in a nutshell

Motivation

NO_x[1]:



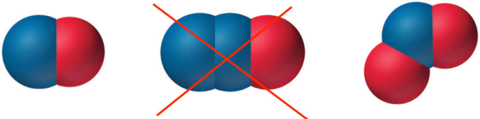
- NO_x are detrimental to the environment and humans.
- NO_x are naturally ocurring in man-made processes. E.g. Combustion.

[1] Image source: ENVIS Centre on Plants and Pollution

Problem in a nutshell

Motivation

NO_x[1]:



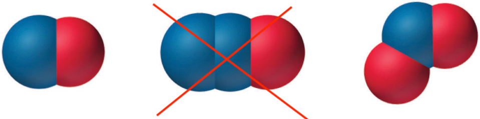
- NO_x are detrimental to the environment and humans.
- NO_x are naturally ocurring in man-made processes. E.g. Combustion.
- Ammonia can "neutralize" NO_x; producing water (H_2O) and nitrogen gas (N_2). Both harmless! - Selective catalytic reduction (SCR).

[1] Image source: ENVIS Centre on Plants and Pollution

Problem in a nutshell

Motivation

NO_x[1]:



NO N_2O NO_2

- NO_x are detrimental to the environment and humans.
- NO_x are naturally ocurring in man-made processes. E.g. Combustion.
- Ammonia can "neutralize" NO_x; producing water (H_2O) and nitrogen gas (N_2). Both harmless! - Selective catalytic reduction (SCR).
- But ammonia is also hazardous to the environment/humans.

[1] Image source: ENVIS Centre on Plants and Pollution

Problem in a nutshell

Motivation

Problem in a nutshell

Motivation

- The dosing of ammonia in the catalyst is key:

Problem in a nutshell

Motivation

- The dosing of ammonia in the catalyst is key:
  - Too much ammonia: NO_x reduction will occur $\rightarrow$ Unnecessary ammonia emissions.

Problem in a nutshell

Motivation

- The dosing of ammonia in the catalyst is key:
  - Too much ammonia: NO_x reduction will occur $\rightarrow$ Unnecessary ammonia emissions.
  - Too little ammonia: NO_x reduction will occur partially/will not occur $\rightarrow$ NO_x emissions.

Problem in a nutshell

Motivation

- The dosing of ammonia in the catalyst is key:
  - Too much ammonia: NO_x reduction will occur $\rightarrow$ Unnecessary ammonia emissions.
  - Too little ammonia: NO_x reduction will occur partially/will not occur $\rightarrow$ NO_x emissions.
- Gas sensors can be used to measure the concentrations of NO_x to aid on ammonia dosing.

Problem in a nutshell

Motivation

- The dosing of ammonia in the catalyst is key:
  - Too much ammonia: NO_x reduction will occur $\rightarrow$ Unnecessary ammonia emissions.
  - Too little ammonia: NO_x reduction will occur partially/will not occur $\rightarrow$ NO_x emissions.
- Gas sensors can be used to measure the concentrations of NO_x to aid on ammonia dosing.
- However, the sensor also responds to ammonia.

Problem in a nutshell

Motivation

- The dosing of ammonia in the catalyst is key:
  - Too much ammonia: NO_x reduction will occur → Unnecessary ammonia emissions.
  - Too little ammonia: NO_x reduction will occur partially/will not occur → NO_x emissions.
- Gas sensors can be used to measure the concentrations of NO_x to aid on ammonia dosing.
- However, the sensor also responds to ammonia.
- Operating the sensor in a cyclic operation (e.g. temperature, gate bias) can enhance selectivity.

Problem in a nutshell

Motivation

- The dosing of ammonia in the catalyst is key:
  - Too much ammonia: NO_x reduction will occur $\rightarrow$ Unnecessary ammonia emissions.
  - Too little ammonia: NO_x reduction will occur partially/will not occur $\rightarrow$ NO_x emissions.
- Gas sensors can be used to measure the concentrations of NO_x to aid on ammonia dosing.
- However, the sensor also responds to ammonia.
- Operating the sensor in a cyclic operation (e.g. temperature, gate bias) can enhance selectivity.
  - Different gasses react differently in different stages of the cycle.

Problem in a nutshell

Motivation

- The dosing of ammonia in the catalyst is key:
  - Too much ammonia: NO_x reduction will occur $\rightarrow$ Unnecessary ammonia emissions.
  - Too little ammonia: NO_x reduction will occur partially/will not occur $\rightarrow$ NO_x emissions.
- Gas sensors can be used to measure the concentrations of NO_x to aid on ammonia dosing.
- However, the sensor also responds to ammonia.
- Operating the sensor in a cyclic operation (e.g. temperature, gate bias) can enhance selectivity.
  - Different gasses react differently in different stages of the cycle.
- Temperature cycling.

Problem in a nutshell
Motivation

- The dosing of ammonia in the catalyst is key:
  - Too much ammonia: NO_x reduction will occur $\rightarrow$ Unnecessary ammonia emissions.
  - Too little ammonia: NO_x reduction will occur partially/will not occur $\rightarrow$ NO_x emissions.
- Gas sensors can be used to measure the concentrations of NO_x to aid on ammonia dosing.
- However, the sensor also responds to ammonia.
- Operating the sensor in a cyclic operation (e.g. temperature, gate bias) can enhance selectivity.
  - Different gasses react differently in different stages of the cycle.
- Temperature cycling.
- **Frequency cycling**.

Problem in a nutshell

Research questions

Problem in a nutshell

Research questions

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?

Problem in a nutshell

Research questions

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
2. Does the quality of fit vary over different prediction models?

Data

Data description

Data

Data description

- 3 gases: NO, NO_2, NH_3.

Data

Data description

- 3 gases: NO, NO_2, NH_3.
- 5 possible concentrations each: 5, 10, 20, 40, 80 ppm.

Data

Data description

- 3 gases: NO, NO_2, NH_3.
- 5 possible concentrations each: 5, 10, 20, 40, 80 ppm.
- 125 unique gas mixtures.

Data

Data description

- 3 gases: NO, NO_2, NH_3.
- 5 possible concentrations each: 5, 10, 20, 40, 80 ppm.
- 125 unique gas mixtures.
- Frequency cycle: from 0.05 to 5000 Hz in 60 seconds.

- 3 gases: NO, NO_2, NH_3.
- 5 possible concentrations each: 5, 10, 20, 40, 80 ppm.
- 125 unique gas mixtures.
- Frequency cycle: from 0.05 to 5000 Hz in 60 seconds.
- Sampling rate: 4 Hz, i.e. 4 readings in one second.

- 3 gases: NO, NO_2, NH_3.
- 5 possible concentrations each: 5, 10, 20, 40, 80 ppm.
- 125 unique gas mixtures.
- Frequency cycle: from 0.05 to 5000 Hz in 60 seconds.
- Sampling rate: 4 Hz, i.e. 4 readings in one second.
- Shape-defining features: **Slopes and Averages**.

- 3 gases: NO, NO_2, NH_3.
- 5 possible concentrations each: 5, 10, 20, 40, 80 ppm.
- 125 unique gas mixtures.
- Frequency cycle: from 0.05 to 5000 Hz in 60 seconds.
- Sampling rate: 4 Hz, i.e. 4 readings in one second.
- Shape-defining features: **Slopes and Averages**.
- Two important definitions:

- 3 gases: NO, NO_2, NH_3.
- 5 possible concentrations each: 5, 10, 20, 40, 80 ppm.
- 125 unique gas mixtures.
- Frequency cycle: from 0.05 to 5000 Hz in 60 seconds.
- Sampling rate: 4 Hz, i.e. 4 readings in one second.
- Shape-defining features: **Slopes and Averages**.
- Two important definitions:
  1. Mixture: an **unique** combination of gases, i.e. **no repetitions**.

Data

Data description

- 3 gases: NO, NO_2, NH_3.
- 5 possible concentrations each: 5, 10, 20, 40, 80 ppm.
- 125 unique gas mixtures.
- Frequency cycle: from 0.05 to 5000 Hz in 60 seconds.
- Sampling rate: 4 Hz, i.e. 4 readings in one second.
- Shape-defining features: **Slopes and Averages**.
- Two important definitions:
  1. Mixture: an **unique** combination of gases, i.e. **no repetitions**.
  2. Exposure: a combination of gasses **with repetition**.

Data

Data collection

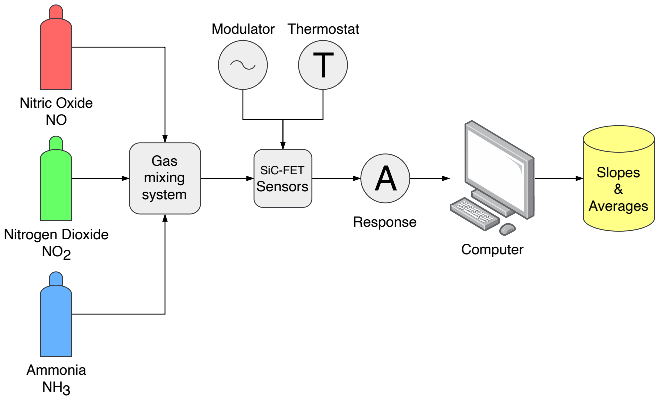


Figure: Schema of the data acquisition process.

Data

Data collection

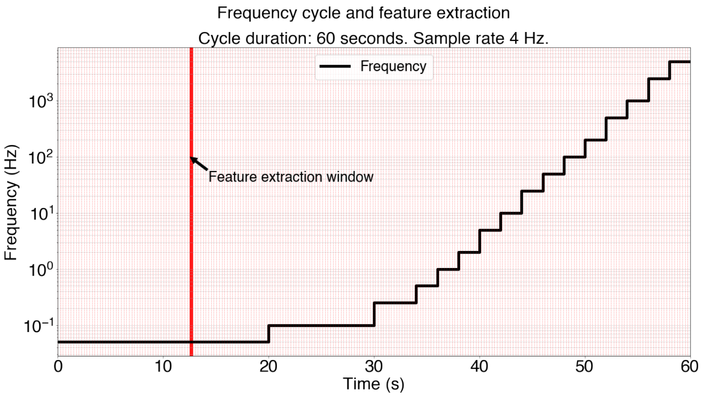


Figure: Feature measurements times per cycle. The width of the red line indicates the duration of one of the feature measurement windows as an example.

Data

Data collection

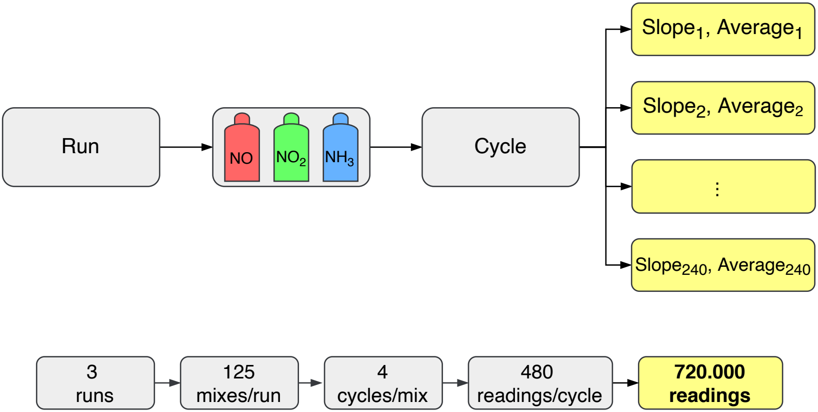


Figure: A visualization of the feature measurement process.

Data

Raw data

Row-based

Table: Sample of raw data.

Index	Exposure nr	Cycle nr	Sample nr	NO [ppm]	NO2 [ppm]	NH3 [ppm]	Freq [Hz]	Slope sensor 1 [μA/s]	Slope sensor 2 [μA/s]	Average sensor 1 [μA]	Average sensor 2 [μA]	Sensor temperature [C]
0	1	1	1	10	5	20	0.05	-18.855169	-22.588416	32.926184	27.961554	274.994683
1	1	1	2	10	5	20	0.05	-28.289268	-28.185027	25.853867	20.915297	274.980487
2	1	1	3	10	5	20	0.05	-0.390916	-0.482129	25.756138	20.794765	274.985895
3	1	1	4	10	5	20	0.05	-0.234549	-0.156366	25.697501	20.755673	275.020372
4	1	1	5	10	5	20	0.05	-0.143336	-0.247580	25.661667	20.693778	275.014964
							⋮					
100000	105	1	161	5	5	40	5.0	-38.366212	-48.495271	30.241896	24.821197	275.021724
100001	105	1	162	5	5	40	5.0	6.619507	8.521964	31.896773	26.951688	274.999415
100002	105	1	163	5	5	40	5.0	-1.941549	6.580416	31.411386	28.596792	275.011584
100003	105	1	164	5	5	40	5.0	27.401023	22.012900	38.261641	34.100017	275.009894
100004	105	1	165	5	5	40	5.0	-27.016623	-28.439121	31.507486	26.990236	275.014400
							⋮					
359995	375	4	236	20	80	5	5000.0	-0.136821	-0.158538	34.129879	30.345597	275.002007
359996	375	4	237	20	80	5	5000.0	0.010859	0.010859	34.132593	30.348312	274.986797
359997	375	4	238	20	80	5	5000.0	-0.043435	0.030405	34.121734	30.355913	274.979811
359998	375	4	239	20	80	5	5000.0	-0.117275	-0.026061	34.092416	30.349398	274.984543
359999	375	4	240	20	80	5	5000.0	0.073840	0.039092	34.110876	30.359171	274.998063

Data

Preprocessing

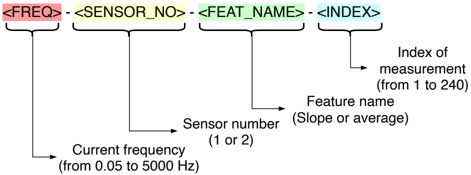


Figure: Feature naming convention.

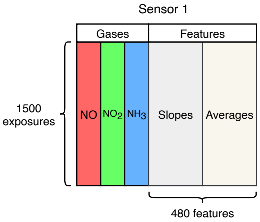


Figure: Pre-processed data structure.

Data

Preprocessed data

Column-based.

Table: Sample of pre-processed data.

Index	EXPOSURE	NO	NO2	NH3	0.05-1-slope-0	0.05-1-slope-1	...	5000.0-1-slope-239	0.05-1-avg-0	0.05-1-avg-1	0.05-1-avg-2	...	5000.0-1-avg-238	5000.0-1-avg-239
0	1.0	10.0	5.0	20.0	-18.855169	-28.289268	...	0.019546	32.926184	25.853867	25.756138	...	35.840135	35.845021
1	1.0	10.0	5.0	20.0	-28.979886	-9.251672	...	-0.056466	28.600050	26.287132	26.225237	...	35.884113	35.869996
2	1.0	10.0	5.0	20.0	-25.431240	-12.874158	...	-0.052122	29.512187	26.293647	26.238267	...	35.913432	35.900401
3	1.0	10.0	5.0	20.0	-30.126572	-8.196200	...	-0.156366	28.368758	26.319708	26.254555	...	35.939493	35.900401
4	2.0	20.0	40.0	40.0	-19.506695	-27.051368	...	-0.078183	33.180279	26.417437	26.303420	...	35.685397	35.665852
							⋮							
700	176.0	40.0	20.0	40.0	-21.011721	-25.822155	...	-0.071668	31.458621	25.003082	24.902639	...	34.554999	34.537082
701	176.0	40.0	20.0	40.0	-27.505265	-10.847911	...	0.086870	27.660766	24.948788	24.918927	...	34.504506	34.526224
702	176.0	40.0	20.0	40.0	-27.516124	-10.750182	...	-0.097729	27.647193	24.959647	24.928700	...	34.531653	34.507221
703	176.0	40.0	20.0	40.0	-27.364102	-10.875058	...	0.086870	27.666195	24.947431	24.935215	...	34.537082	34.558800
704	177.0	80.0	40.0	40.0	-20.794546	-26.195696	...	0.041263	31.640505	25.091581	25.088324	...	34.695078	34.705393
							⋮							
1495	374.0	80.0	80.0	40.0	-27.937445	-10.891346	...	-0.097729	27.166692	24.443855	24.392276	...	34.151596	34.127164
1496	375.0	20.0	80.0	5.0	-24.358394	-22.933723	...	-0.008687	30.315735	24.582305	24.530726	...	34.134765	34.132593
1497	375.0	20.0	80.0	5.0	-28.862612	-9.827186	...	-0.112931	26.916940	24.460144	24.410736	...	34.159740	34.131507
1498	375.0	20.0	80.0	5.0	-25.839531	-12.780772	...	-0.021718	27.671625	24.476432	24.430282	...	34.143452	34.138023
1499	375.0	20.0	80.0	5.0	-28.002598	-10.645937	...	0.073840	27.137373	24.475889	24.424853	...	34.092416	34.110876

Data

Feature averaging

Data

Feature averaging

- 1500 exposures - hard to visualize.

Data

Feature averaging

- 1500 exposures - hard to visualize.
- Averaging all **exposures** of a particular **mixture**.

Data

Feature averaging

- 1500 exposures - hard to visualize.
- Averaging all **exposures** of a particular **mixture**.
- 125 mixtures - easier (or less hard) to visualize

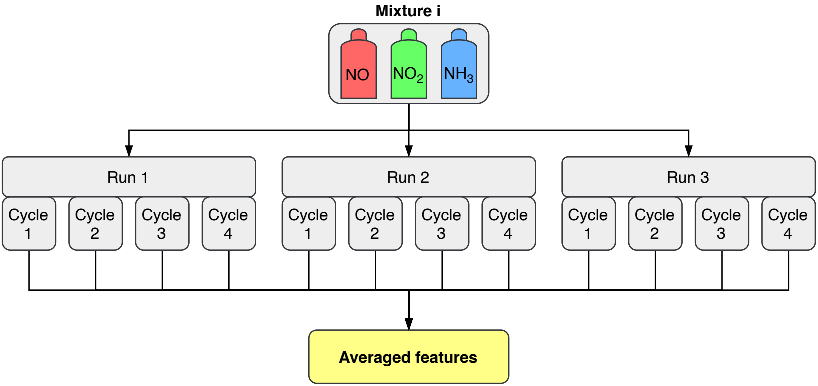


Figure: A visualization of the feature averaging process.

Data

Slopes

In each plot, each line represents a gas **mixture**.

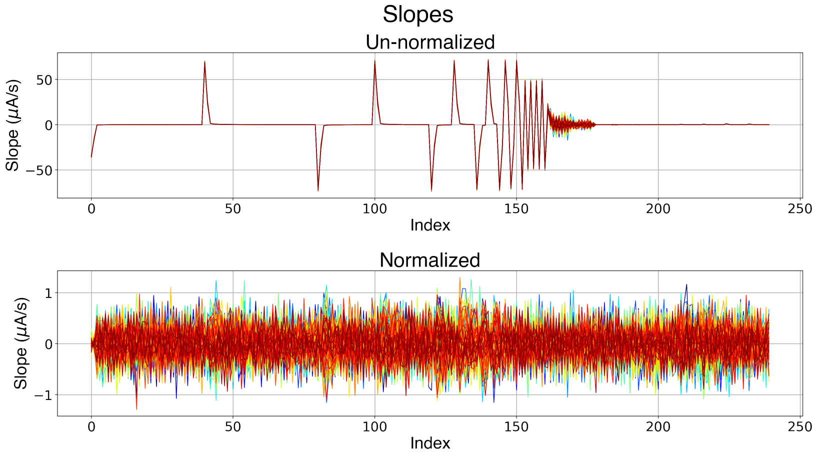


Figure: Slope features, un-normalized and normalized.

Data

Averages

In each plot, each line represents a gas **mixture**.

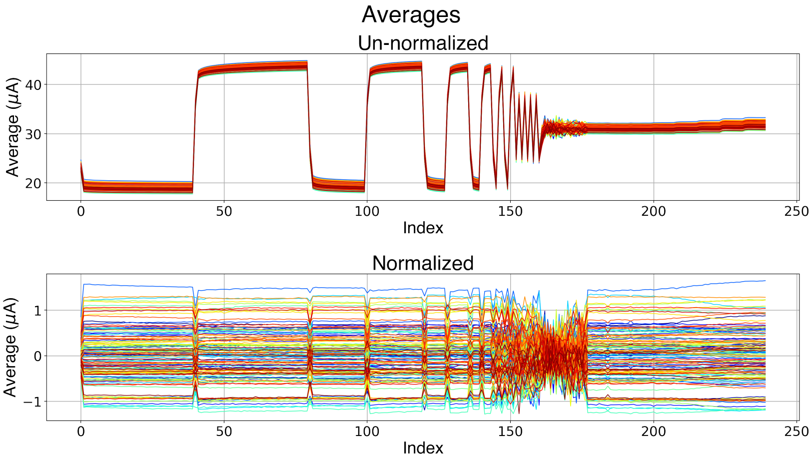


Figure: Average features, un-normalized and normalized.

Data

Feature averaging

Data

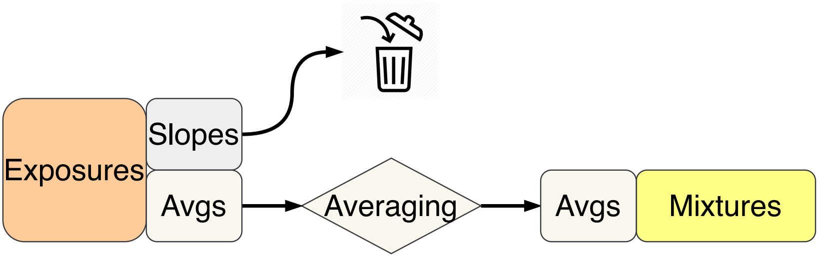
Feature averaging

- Slope features don't change between mixtures (down to measuring noise).

Data

Feature averaging

- Slope features don't change between mixtures (down to measuring noise).
- Analysis will be **also** run on averaged averages only.



Data

Feature averaging

Index	UNIQUE MIXTURE	NO	NO2	NH3	0.05-1-avg-0	0.05-1-avg-1	0.05-1-avg-2	...	5000.0-1-avg-238	5000.0-1-avg-239
0	0	5.0	5.0	5.0	28.983749	25.442410	25.383750	...	35.162932	35.152458
1	1	5.0	5.0	10.0	28.538652	24.933269	24.879247	...	34.622460	34.623591
2	2	5.0	5.0	20.0	29.038925	25.245278	25.181935	...	35.025637	35.023103
3	3	5.0	5.0	40.0	28.698684	25.057399	24.980686	...	34.575699	34.576649
4	4	5.0	5.0	80.0	28.738748	25.289980	25.229714	...	34.860040	34.854340
							⋮			
70	70	20.0	80.0	5.0	28.142217	24.646824	24.596195	...	34.208650	34.213536
71	71	20.0	80.0	10.0	28.615026	24.952453	24.893228	...	34.511972	34.497900
72	72	20.0	80.0	20.0	28.432463	24.705665	24.649538	...	34.317554	34.313437
73	73	20.0	80.0	40.0	28.327675	24.725143	24.685825	...	34.213989	34.197429
74	74	20.0	80.0	80.0	28.611836	25.056652	24.993128	...	34.592507	34.593548
							⋮			
120	120	80.0	80.0	5.0	28.548244	25.157684	25.103051	...	34.742313	34.749349
121	121	80.0	80.0	10.0	28.630183	25.045884	25.015773	...	34.678857	34.675690
122	122	80.0	80.0	20.0	28.420835	24.737087	24.687996	...	34.354338	34.347416
123	123	80.0	80.0	40.0	28.457189	24.743263	24.682929	...	34.327938	34.319839
124	124	80.0	80.0	80.0	28.615161	25.093255	25.046698	...	34.743829	34.734124

Methods

Chosen models

Methods

Chosen models

- Multivariate multiple regression problem
  - Multiple predictors - slopes and averages.
  - Multiple responses - gas concentrations.

Methods

Chosen models

- Multivariate multiple regression problem
  - Multiple predictors - slopes and averages.
  - Multiple responses - gas concentrations.
- Proposed models:
  1. Ordinary Least Squares Regression (OLS).
  2. Principal Components Regression (PCR).
  3. Partial Least Squares Regression (PLSR).
  4. Ridge regression.

Methods

Chosen models

- Multivariate multiple regression problem
  - Multiple predictors - slopes and averages.
  - Multiple responses - gas concentrations.
- Proposed models:
  1. Ordinary Least Squares Regression (OLS).
  2. Principal Components Regression (PCR).
  3. Partial Least Squares Regression (PLSR).
  4. Ridge regression.
- Why?

Methods

Chosen models

- Multivariate multiple regression problem
  - Multiple predictors - slopes and averages.
  - Multiple responses - gas concentrations.
- Proposed models:
  1. Ordinary Least Squares Regression (OLS).
  2. Principal Components Regression (PCR).
  3. Partial Least Squares Regression (PLSR).
  4. Ridge regression.
- Why?
  - Natural progression of model choice: simple $\rightarrow$ complex

Methods

Chosen models

- Multivariate multiple regression problem
  - Multiple predictors - slopes and averages.
  - Multiple responses - gas concentrations.
- Proposed models:
  1. Ordinary Least Squares Regression (OLS).
  2. Principal Components Regression (PCR).
  3. Partial Least Squares Regression (PLSR).
  4. Ridge regression.
- Why?
  - Natural progression of model choice: simple $\rightarrow$ complex
  - PLSR as a starting point: shown to work well in this field/problem.

Methods

Chosen models

- Multivariate multiple regression problem
  - Multiple predictors - slopes and averages.
  - Multiple responses - gas concentrations.
- Proposed models:
  1. Ordinary Least Squares Regression (OLS).
  2. Principal Components Regression (PCR).
  3. Partial Least Squares Regression (PLSR).
  4. Ridge regression.
- Why?
  - Natural progression of model choice: simple $\rightarrow$ complex
  - PLSR as a starting point: shown to work well in this field/problem.
  - Sensor data collected in short succession: possibly high correlated features $\rightarrow$ dimensionality reduction or shrinkage to alleviate it.

Methods

Chosen models

- Multivariate multiple regression problem
  - Multiple predictors - slopes and averages.
  - Multiple responses - gas concentrations.
- Proposed models:
  1. Ordinary Least Squares Regression (OLS).
  2. Principal Components Regression (PCR).
  3. Partial Least Squares Regression (PLSR).
  4. Ridge regression.
- Why?
  - Natural progression of model choice: simple $\rightarrow$ complex
  - PLSR as a starting point: shown to work well in this field/problem.
  - Sensor data collected in short succession: possibly high correlated features $\rightarrow$ dimensionality reduction or shrinkage to alleviate it.
- Pros: backed by literature, simple to implement.

Methods

Chosen models

- Multivariate multiple regression problem
  - Multiple predictors - slopes and averages.
  - Multiple responses - gas concentrations.
- Proposed models:
  1. Ordinary Least Squares Regression (OLS).
  2. Principal Components Regression (PCR).
  3. Partial Least Squares Regression (PLSR).
  4. Ridge regression.
- Why?
  - Natural progression of model choice: simple $\rightarrow$ complex
  - PLSR as a starting point: shown to work well in this field/problem.
  - Sensor data collected in short succession: possibly high correlated features $\rightarrow$ dimensionality reduction or shrinkage to alleviate it.
- Pros: backed by literature, simple to implement.
- Cons: only linear models.

Methods

OLS

Methods

OLS

- Relation between predictors (X) and responses (Y):

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E} \tag{1}$$

Methods

OLS

- Relation between predictors (X) and responses (Y):

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \tag{1}$$

- The difference between the response and the prediction is the residual. The Residual Sum of Squares is defined as:

$$\mathrm{RSS}(\mathbf{B}) = \mathrm{Tr}[(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})^{\intercal}(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})] \tag{2}$$

Methods

OLS

- Relation between predictors (X) and responses (Y):

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E} \tag{1}$$

- The difference between the response and the prediction is the residual. The Residual Sum of Squares is defined as:

$$\mathrm{RSS}(\mathbf{B}) = \mathrm{Tr}[(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})^\intercal(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})] \tag{2}$$

- The estimate for coefficients $\hat{\mathbf{B}}$ is chosen to minimize the RSS:

$$\widehat{\mathbf{B}}^{\mathrm{OLS}} = \arg\min_{\mathbf{B}} \mathrm{RSS}(\mathbf{B}) \tag{3}$$

$$\widehat{\mathbf{B}}^{\mathrm{OLS}} = (\mathbf{X}^\intercal\mathbf{X})^{-1}\mathbf{X}^\intercal\mathbf{Y} \tag{4}$$

Methods

OLS

- Relation between predictors (X) and responses (Y):

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E} \tag{1}$$

- The difference between the response and the prediction is the residual. The Residual Sum of Squares is defined as:

$$\mathrm{RSS}(\mathbf{B}) = \mathrm{Tr}[(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})^\intercal(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})] \tag{2}$$

- The estimate for coefficients $\hat{\mathbf{B}}$ is chosen to minimize the RSS:

$$\widehat{\mathbf{B}}^{\mathrm{OLS}} = \arg\min_{\mathbf{B}} \mathrm{RSS}(\mathbf{B}) \tag{3}$$

$$\widehat{\mathbf{B}}^{\mathrm{OLS}} = (\mathbf{X}^\intercal\mathbf{X})^{-1}\mathbf{X}^\intercal\mathbf{Y} \tag{4}$$

- In this problem, features are sampled in quick succession (4 Hz). This will lead to highly correlated features $\rightarrow$ high variance of

Methods

PCR

- The objective of PCA is to find a matrix $\mathbf{P}$ such that the linear transformation

$$\mathbf{T} = \mathbf{X}\mathbf{P}^{\intercal} \tag{5}$$

yields new variables that are uncorrelated and arranged in decreasing order of variance. The linear relation between response and PC scores is then:

$$\mathbf{Y} = \mathbf{T}\mathbf{B} + \mathbf{E} \tag{6}$$

Methods

PCR

- The objective of PCA is to find a matrix $\mathbf{P}$ such that the linear transformation

$$\mathbf{T} = \mathbf{X}\mathbf{P}^\intercal \tag{5}$$

yields new variables that are uncorrelated and arranged in decreasing order of variance. The linear relation between response and PC scores is then:

$$\mathbf{Y} = \mathbf{T}\mathbf{B} + \mathbf{E} \tag{6}$$

- And the regression coefficients are found analogously to OLS:

$$\widehat{\mathbf{B}}^{\text{PCR}} = (\mathbf{T}^\intercal\mathbf{T})^{-1}\mathbf{T}^\intercal\mathbf{Y} \tag{7}$$

Methods

PCR

- The objective of PCA is to find a matrix $\mathbf{P}$ such that the linear transformation

$$\mathbf{T} = \mathbf{X}\mathbf{P}^{\mathsf{T}} \tag{5}$$

yields new variables that are uncorrelated and arranged in decreasing order of variance. The linear relation between response and PC scores is then:

$$\mathbf{Y} = \mathbf{T}\mathbf{B} + \mathbf{E} \tag{6}$$

- And the regression coefficients are found analogously to OLS:

$$\widehat{\mathbf{B}}^{\text{PCR}} = (\mathbf{T}^{\mathsf{T}}\mathbf{T})^{-\mathbf{1}}\mathbf{T}^{\mathsf{T}}\mathbf{Y} \tag{7}$$

- While this method successfully best explains the variability in $\mathbf{X}$, the responses $\mathbf{Y}$ are not taken into account.

Methods

PLSR

Methods

PLSR

- Supervision of the identification of components.

Methods

PLSR

- Supervision of the identification of components.
- Linear transformations for both $\mathbf{X}$ and $\mathbf{Y}$, analogous to PCA.

$$\mathbf{W} = \mathbf{X}\mathbf{L}^{\intercal} \tag{8}$$

$$\mathbf{U} = \mathbf{Y}\mathbf{Q}^{\intercal} \tag{9}$$

- Supervision of the identification of components.
- Linear transformations for both **X** and **Y**, analogous to PCA.

$$\mathbf{W} = \mathbf{X}\mathbf{L}^{\intercal} \tag{8}$$

$$\mathbf{U} = \mathbf{Y}\mathbf{Q}^{\intercal} \tag{9}$$

- The relation between responses and PLS scores is:

$$\mathbf{Y} = \mathbf{W}\mathbf{B} + \mathbf{E} \tag{10}$$

- Supervision of the identification of components.
- Linear transformations for both $\mathbf{X}$ and $\mathbf{Y}$, analogous to PCA.

$$\mathbf{W} = \mathbf{X}\mathbf{L}^\intercal \tag{8}$$

$$\mathbf{U} = \mathbf{Y}\mathbf{Q}^\intercal \tag{9}$$

- The relation between responses and PLS scores is:

$$\mathbf{Y} = \mathbf{W}\mathbf{B} + \mathbf{E} \tag{10}$$

- Again, analogous to OLS, the coefficients are:

$$\widehat{\mathbf{B}}^{\mathrm{PLSR}} = (\mathbf{W}^\intercal\mathbf{W})^{-1}\mathbf{W}^\intercal\mathbf{Y} \tag{11}$$

Methods

Ridge

Methods

Ridge

- Shrinkage of coefficients. Coefficients tend to, but never reach, zero.

Methods

Ridge

- Shrinkage of coefficients. Coefficients tend to, but never reach, zero.
- Addition of a regularization term.

Methods

Ridge

- Shrinkage of coefficients. Coefficients tend to, but never reach, zero.
- Addition of a regularization term.
- λ controls the regularization strength.

Methods

Ridge

- Shrinkage of coefficients. Coefficients tend to, but never reach, zero.
- Addition of a regularization term.
- λ controls the regularization strength.
- The RSS for multioutput ridge regression is:

$$\mathrm{RSS}^{\mathrm{Ridge}}(\widehat{\mathbf{B}}, \boldsymbol{\lambda}) = \mathrm{Tr}[(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})^{\intercal}(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})] + \mathrm{Tr}[\widehat{\mathbf{B}}^{\intercal}\widehat{\mathbf{B}} + \boldsymbol{\lambda}\mathbf{I}] \tag{12}$$

Methods

Ridge

- Shrinkage of coefficients. Coefficients tend to, but never reach, zero.
- Addition of a regularization term.
- λ controls the regularization strength.
- The RSS for multioutput ridge regression is:

$$\mathrm{RSS}^{\mathrm{Ridge}}(\widehat{\mathbf{B}}, \boldsymbol{\lambda}) = \mathrm{Tr}[(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})^{\intercal}(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})] + \mathrm{Tr}[\widehat{\mathbf{B}}^{\intercal}\widehat{\mathbf{B}} + \boldsymbol{\lambda}\mathbf{I}] \tag{12}$$

- Similar to before, the estimate for $\mathbf{B}$ is found by minimizing the RSS.

$$\widehat{\mathbf{B}}^{\mathrm{Ridge}} = \arg\min_{\mathbf{B}} \mathrm{RSS}^{\mathrm{Ridge}}(\mathbf{B}) \tag{13}$$

Methods

Ridge

- Shrinkage of coefficients. Coefficients tend to, but never reach, zero.
- Addition of a regularization term.
- λ controls the regularization strength.
- The RSS for multioutput ridge regression is:

$$\mathrm{RSS}^{\mathrm{Ridge}}(\widehat{\mathbf{B}}, \boldsymbol{\lambda}) = \mathrm{Tr}[(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})^{\intercal}(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})] + \mathrm{Tr}[\widehat{\mathbf{B}}^{\intercal}\widehat{\mathbf{B}} + \boldsymbol{\lambda}\mathbf{I}] \tag{12}$$

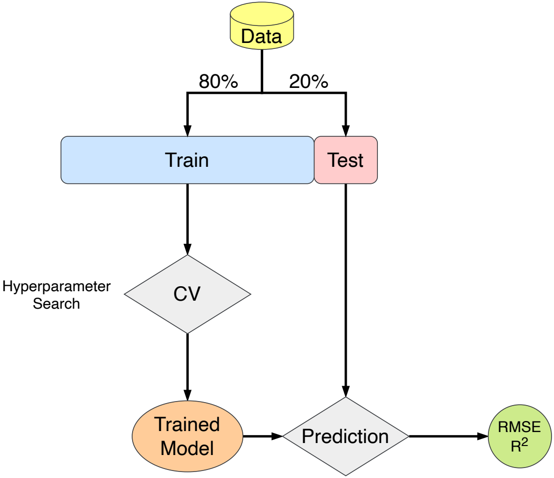
- Similar to before, the estimate for $\mathbf{B}$ is found by minimizing the RSS.

$$\widehat{\mathbf{B}}^{\mathrm{Ridge}} = \arg\min_{\mathbf{B}} \mathrm{RSS}^{\mathrm{Ridge}}(\mathbf{B}) \tag{13}$$

- Solving for $\widehat{\mathbf{B}}$ yields:

$$\widehat{\mathbf{B}}^{\mathrm{Ridge}} = (\mathbf{X}^{\intercal}\mathbf{X} + \boldsymbol{\lambda}\mathbf{I})^{-\mathbf{1}}\mathbf{X}^{\intercal}\mathbf{Y} \tag{14}$$

Methods



Data
80%
20%
Train
Test
Hyperparameter
Search
CV
Trained
Model
Prediction
RMSE
R2

Results

OLS

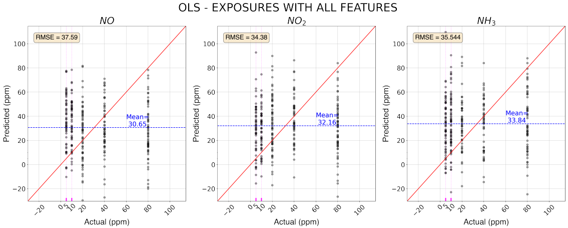


Figure: Actual vs. Predicted plot for **exposures**.

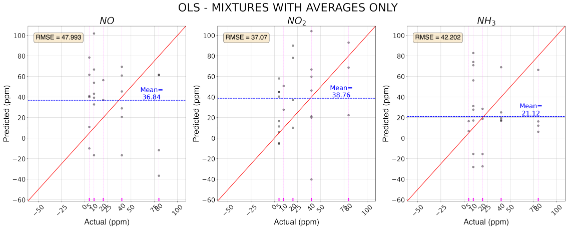


Figure: Actual vs. Predicted plot for **mixtures** - averaged average features.

Results

PCR

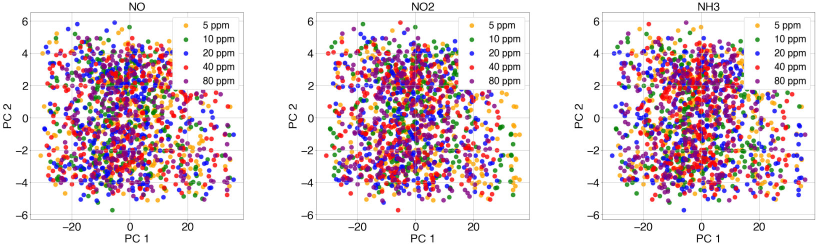


Figure: PCA for slopes and averages through **exposures**

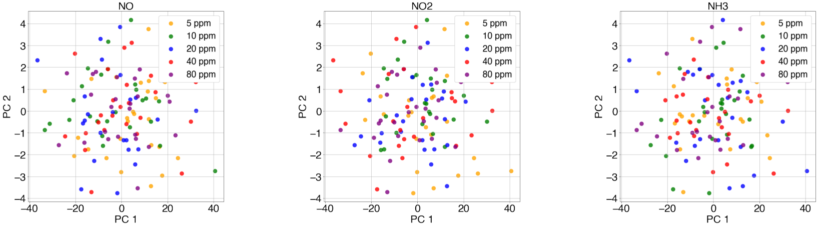


Figure: PCA for averaged average features through **mixtures**

Results

PCR

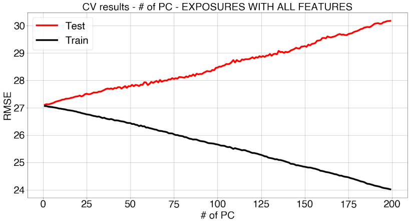


Figure: CV results for **exposures**.

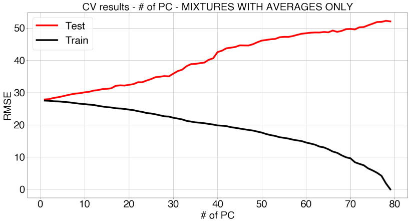


Figure: CV results for **mixtures**.

Results

PCR

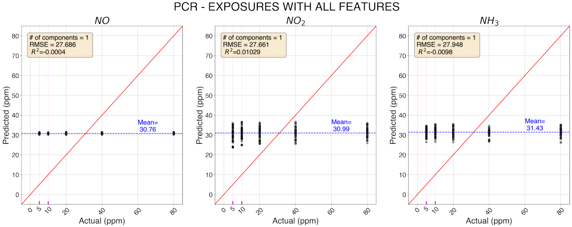


Figure: Actual vs. pred for **exposures**.

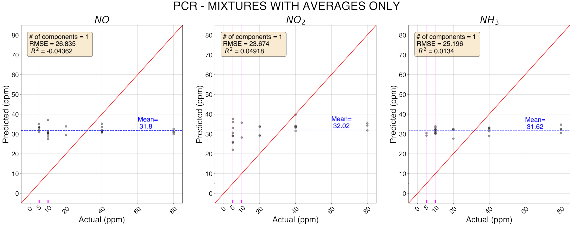


Figure: Actual vs. pred for **mixtures**.

Results

PLSR

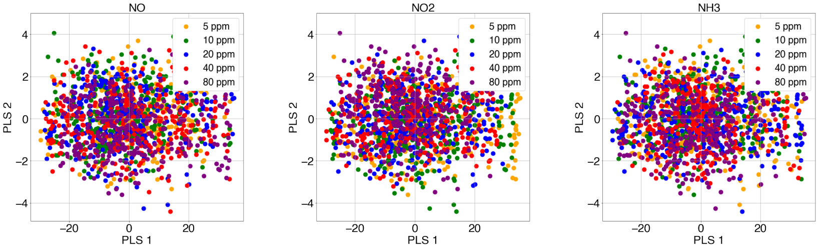


Figure: PLS scores for **exposures**.

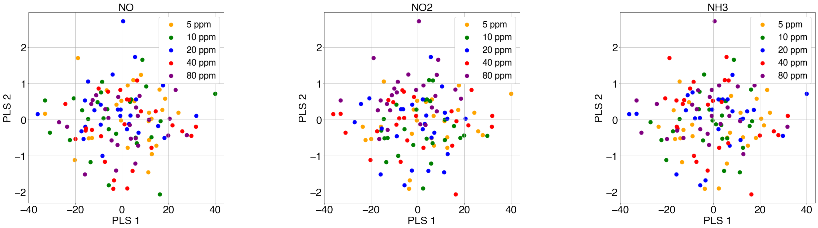


Figure: PLS scores for **mixtures**.

Results

PLSR

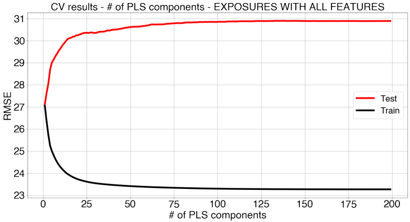


Figure: CV results for **exposures**.

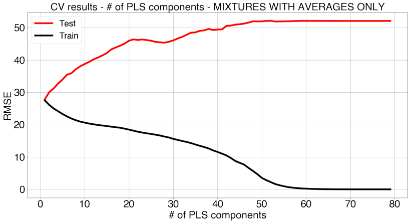


Figure: CV results for **mixtures**.

Results

PLSR

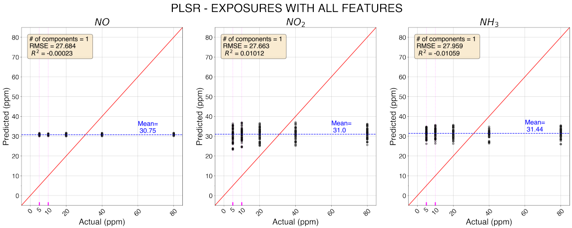


Figure: Actual vs. pred for **exposures**.

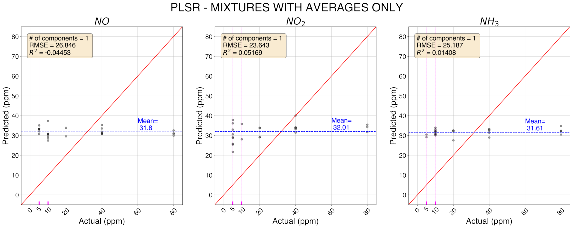


Figure: Actual vs. pred for **mixtures**.

Results

Ridge regression

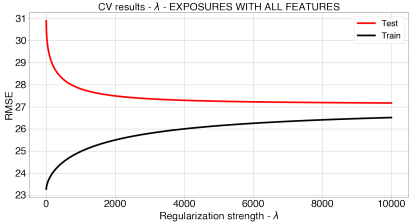


Figure: CV results for **exposures**.

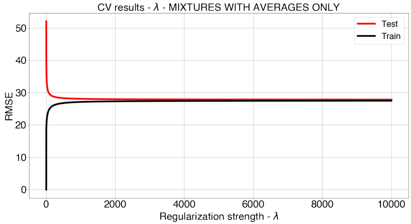


Figure: CV results for **mixtures**.

Results

Ridge regression

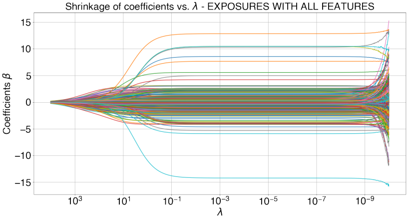


Figure: Shrinkage of coefficients for **exposures**.

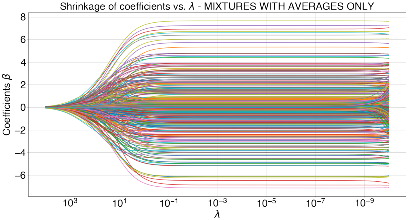


Figure: Shrinkage of coefficients for **mixtures**.

Results

Ridge regression

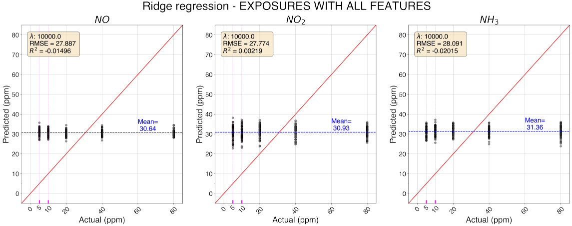


Figure: Actual vs. pred for **exposures**.

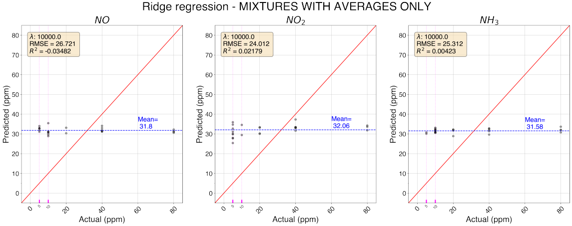


Figure: Actual vs. pred for **mixtures**.

Discussion

Discussion

- All models fail on predicting gas concentrations.

Discussion

- All models fail on predicting gas concentrations.
- Slopes are not informative.

Discussion

- All models fail on predicting gas concentrations.
- Slopes are not informative.
- Averages might be informative but there is no linear order in it.

Discussion

- All models fail on predicting gas concentrations.
- Slopes are not informative.
- Averages might be informative but there is no linear order in it.
- Predictions are centered approximately around the target mean $\bar{y} = 31$ ppm.

Discussion

- All models fail on predicting gas concentrations.
- Slopes are not informative.
- Averages might be informative but there is no linear order in it.
- Predictions are centered approximately around the target mean $\bar{y} = 31$ ppm.
- Results show that less complex, under-fitting models are preferred.

Discussion

Attempts at finding order

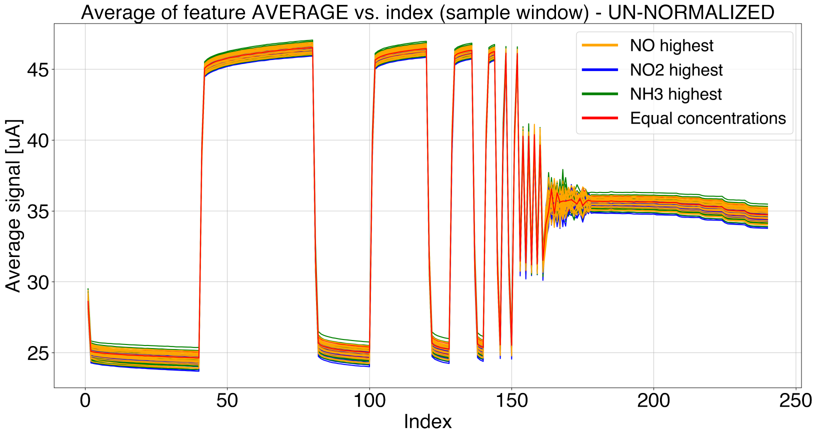


Figure: Averaged sensor average divided by predominant gas. Each line corresponds to a unique **mixture**.

Discussion

Attempts at finding order

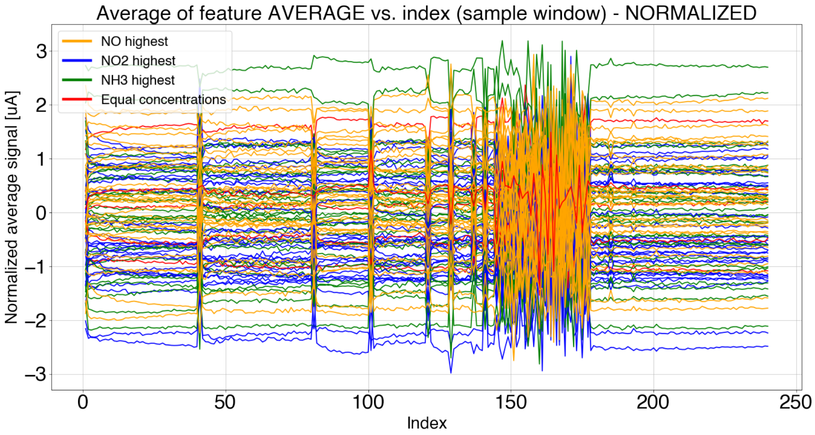


Figure: Normalized averaged sensor average divided by predominant gas. Each line corresponds to a unique **mixture**.

Discussion

Attempts at finding order

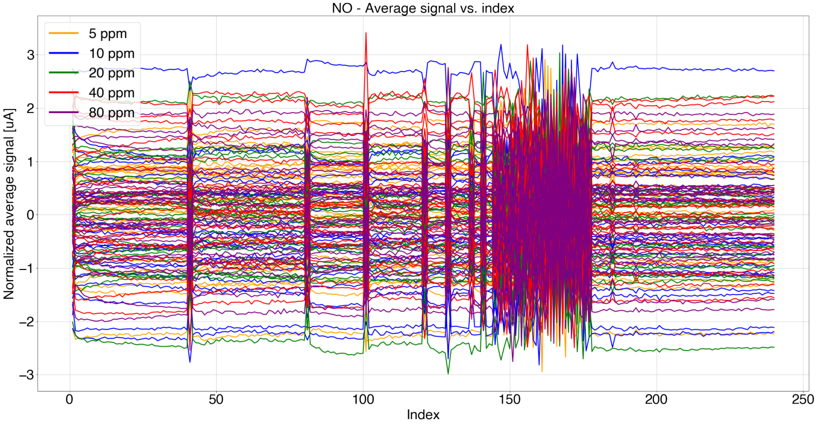


Figure: Normalized sensor averaged for NO. Each line corresponds to a unique **mixture**. The levels are the NO concentrations of the mixture.

Conclusion

and future work

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**
   - To answer this conclusively, different approaches need to be experimented with:

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
    - **Possibly not.**
    - To answer this conclusively, different approaches need to be experimented with:
        1.1 Different shape-defining features? E.g. differences.

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**
   - To answer this conclusively, different approaches need to be experimented with:
     1.1 Different shape-defining features? E.g. differences.
     1.2 Different wave shape? E.g. Triangular waves.

Conclusion
and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**
   - To answer this conclusively, different approaches need to be experimented with:
     1.1 Different shape-defining features? E.g. differences.
     1.2 Different wave shape? E.g. Triangular waves.
     1.3 Different models? E.g. Non-parametric models.

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**
   - To answer this conclusively, different approaches need to be experimented with:
     - 1.1 Different shape-defining features? E.g. differences.
     - 1.2 Different wave shape? E.g. Triangular waves.
     - 1.3 Different models? E.g. Non-parametric models.
     - 1.4 Different measurement window? e.g. lower sampling rate

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**
   - To answer this conclusively, different approaches need to be experimented with:
     1.1 Different shape-defining features? E.g. differences.
     1.2 Different wave shape? E.g. Triangular waves.
     1.3 Different models? E.g. Non-parametric models.
     1.4 Different measurement window? e.g. lower sampling rate
2. Does the quality-of-fit vary over different prediction models?

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**
   - To answer this conclusively, different approaches need to be experimented with:
     1.1 Different shape-defining features? E.g. differences.
     1.2 Different wave shape? E.g. Triangular waves.
     1.3 Different models? E.g. Non-parametric models.
     1.4 Different measurement window? e.g. lower sampling rate
2. Does the quality-of-fit vary over different prediction models?
   - **No.**

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**
   - To answer this conclusively, different approaches need to be experimented with:
     1.1 Different shape-defining features? E.g. differences.
     1.2 Different wave shape? E.g. Triangular waves.
     1.3 Different models? E.g. Non-parametric models.
     1.4 Different measurement window? e.g. lower sampling rate
2. Does the quality-of-fit vary over different prediction models?
   - **No.**
   - No point in selecting which model is "less bad".

Conclusion

and future work

1. Can frequency modulation be used to simultaneously quantify NO_x and Ammonia concentrations?
   - **Possibly not.**
   - To answer this conclusively, different approaches need to be experimented with:
     1.1 Different shape-defining features? E.g. differences.
     1.2 Different wave shape? E.g. Triangular waves.
     1.3 Different models? E.g. Non-parametric models.
     1.4 Different measurement window? e.g. lower sampling rate
2. Does the quality-of-fit vary over different prediction models?
   - **No.**
   - No point in selecting which model is "less bad".

However, this work succeeded in pointing what does *not* work. Arguably, this *is as important as* know what works.

Thank you!