Linköping University | Department of Computer and Information Science
Master's thesis, 30 ECTS | Statistics and Machine Learning
2021 | LIU-IDA/LITH-EX-A--2021/001--SE

Quantifying nitrogen oxides and ammonia via frequency modulation in gas sensors
– **DRAFT**

Kvantifiering av kväveoxider och ammoniak via frekvensmodulering i gassensorer

Marcos Freitas Mourão dos Santos

Supervisor : Annika Tillander
Examiner : José M. Peña

External supervisor : Mike Andersson

LINKÖPING UNIVERSITY

Linköpings universitet
SE–581 83 Linköping
+46 13 28 10 00 , www.liu.se

Abstract

The abstract resides in file `Abstract.tex`. Here you should write a short summary of your work.

Acknowledgments

Acknowledgments.tex

Contents

Abstract iii

Acknowledgments iv

Contents v

List of Figures vii

List of Tables viii

Acronyms 1

1 Introduction 2
- 1.1 Motivation . . . 2
- 1.2 Aim . . . 3
- 1.3 Research questions . . . 3

2 Theory 4
- 2.1 Principal Component Regression . . . 4
- 2.2 Partial Least Squares Regression . . . 6
- 2.3 Ridge Regression . . . 6

3 Data 8
- 3.1 Data acquisition . . . 8
- 3.2 Raw data . . . 9
- 3.3 Secondary data . . . 9

4 Method 10

5 Results 11

6 Discussion 12
- 6.1 Results . . . 12
- 6.2 Method . . . 12
- 6.3 The work in a wider context . . . 12

7 Conclusion 13

Bibliography **14**

8 Appendix **16**

8.1 Appendix A: Data acquisition time stamps 16

List of Figures

3.1 Schema of the data acquisition process. 8
3.2 An example of row sensor response 9

8.1 Data acquisition timestamps. 17

List of Tables

3.1 Data acquisition details 9

Acronyms

AC Alternating Current. 3

GBCO Gate Bias Cycled Operation. 3

Hz Hertz. 9

mA miliamperes. 8

NIPALS Nonlinear Iterative Partial Least Squares. 5, 6

PC Principal Component. 5

PCA Principal Components Analysis. 4

PCR Principal Components Regression. 4, 5, 6

PLS Partial Least Squares. 6

PLSR Partial Least Squares Regression. 3, 6

ppm parts per million. 8

SAS Sensor and Actuator Systems. 8

SCR Selective Catalytic Reduction. 2, 3

SiC-FET Silicon Carbide Field Effect Transistor. 3, 8

TCO Temperature Cycled Operation. 3

1 Introduction

1.1 Motivation

Nitric Oxide (NO) and Nitrogen Dioxide (NO_2), commonly referred together as NO_x, are hazardous gases to the environment and to humans. Its main sources are combustion processes in transportation, and industrial processes such as (but not limited to) auto mobiles, trucks, boats, industrial boilers, turbines, etc. [12].

NO_x exposure to humans can cause respiratory illnesses such bronchitis, emphysema and can worsen heart disease [4]. Environmentally, NO_x are deemed precursors of adverse phenomena such as smog, acid rain, and the depletion of ozone (O_3) [1]. It is of high interest, therefore, to reduce NO_x emissions.

One well studied and successful method of reducing emissions is Selective Catalytic Reduction (SCR), which consists in the reduction of NO_x by ammonia (NH_3) into nitrogen gas (N_2) and water (H_2O) [6], both harmless components. The process is based in the following reactions [6]:

- $4\,NH_3 + 4\,NO + O_2 \longrightarrow 4\,N_2 + 6\,H_2O$
- $2\,NH_3 + NO + NO_2 \longrightarrow 2\,N_2 + 3\,H_2O$
- $8\,NH_3 + 6\,NO_2 \longrightarrow 7\,N_2 + 12\,H_2O$

One key element in these reactions, however, is the amount of ammonia dosed into the SCR systems. Ammonia itself is hazardous to humans, causing skin and respiratory irritation, among other illnesses [2]. More importantly, ammonia is one of the main sources of nitrogen pollution and it has direct negative impact on biodiversity via nitrogen deposition in soil and water [8]. Hence it is also desired to keep ammonia emissions to a minimum. Too much ammonia in the SCR catalyst will guarantee NO_x reduction at the expense of undesired ammonia emissions. Concurrently, too little ammonia will impede SCR to occur properly, beating the purpose of the catalyst and as a consequence, undesired NO_x emissions.

To monitor gasses concentrations, chemical sensors are deployed, one of which is the Silicon Carbide Field Effect Transistor (SiC-FET). The identification and quantification of gasses is normally achieved through multiple sensor in so called sensor arrays. Ideally each sensor in the array needs to have different responses to different compounds [3]. The deployment of multiple sensors, on the other hand, proves itself cumbersome due to the increased chances of failure, and decalibration of the system should one or multiple sensors be replaced [3].

One solution to this problem is the cycled operation of one single sensor, referred as virtual multi-sensor [3]. By cycling the working point parameters of the sensor, different substances react differently in the sensor surface, which in turn produces different responses. Temperature Cycled Operation (TCO), Gate Bias Cycled Operation (GBCO), and the combination of the two have been proven to increase selectivity of SiC-FET sensors [3].

TCO, in contrast with a constant temperature evaluation, produces unique transient sensor responses, i.e. each gas mixture yields a slightly different sensor output. This unique gas signature increases selectivity [5]. Additionally, the high temperatures reached in these cycles help in the cleansing of the sensor surface, preparing it for the new mixtures to come.

Frequency modulation tries to achieve the same goal: avoid steady state responses in exchange of unique signatures that could help identify/quantify the gasses at hand. It consists on operating the sensor in Alternating Current (AC). One then can regulate the frequency of this operation and create cycles of different frequencies, similar to what is done in TCO. This is equivalent to GBCO, but with more frequency changes and achieving overall higher frequencies.

The main question is: given these set of unique sensor responses, how one can quantify the gasses that produced them? The answer lies in multivariate regression techniques. Partial Least Squares Regression (PLSR) has been used in chemometrics extensively and it has been proven to be good at this task [3] [13]. Other multivariate regression methods, naturally, can also be used. This is the aim of this thesis work, which is shown in the following section.

1.2 Aim

The aim of this thesis is to investigate different regression methods, namely: PLSR, Ridge Regression and (neural nets XXXX - TENTATIVE), and their fit to correctly quantify gas mixtures such NO_x and Ammonia subjected to sensor frequency modulation.

1.3 Research questions

1. Is it possible to achieve acceptable prediction levels for NO_x and Ammonia using frequency modulation?
2. Which method yields best predictions of gas concentrations?

2 Theory

Comment to teacher and opponents: I feel I should rewrite this section contextualized to the problem. Since i do not have the data yet, (although is far from perfect) I thought it was best to keep it "pure" and not long. What do you think? It is also kinda "hard" to write these math heavy derivations by paraphrasing (and citing ,of course) the original authors. Should I just transcribe the math down using quotes ""?

It is often the case that sensor data points are acquired in quick succession, which in turn leads to highly correlated features [3], which can result in high variance [7]. It is desired, then, to apply some feature selection before using the data in prediction models.

TODO: Add section on Linear Regressin to better introduce all other methods

TODO: Standardize notation. It is all over the place.

TODO: Give more details

TODO: complete PLSR and Ridge section

TODO: Perhaps do a separate, in depth section on Principal Components

2.1 Principal Component Regression

The idea behind Principal Components Regression (PCR) is first to reveal more simple underlying structures in data [11] via Principal Components Analysis (PCA) and then performing linear regression on them. PCA aims to find linear combinations of the input variables in such a way that a few of those new, derived variables can explain most of the variability in the system [9].

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

$$\mathbf{T} = \mathbf{XP} \tag{2.1}$$

yields new variables $(t_1, t_2, ..., t_m)$ that are uncorrelated and arranged in decreasing order of variance . $\mathbf{T}$ is named scores and $\mathbf{P}$ Principal Component (PC) of $\mathbf{X}$ [10]. Since the matrix $\mathbf{T}$ is ordered, it follows that most of the variance on the data $\mathbf{X}$ is captured by the first *k-th* PC [10]. This approximation of $\mathbf{X}$ is defined in Equation 2.2:

$$\mathbf{T_{|k}} = \mathbf{XP_{|k}} \tag{2.2}$$

Finally, PCR is simply performing linear regression on $\mathbf{T_{|k}}$ instead of $\mathbf{X}$:

$$\mathbf{y} = \mathbf{T_{|k}}\beta + \epsilon \tag{2.3}$$

And the regression coefficients just as in linear regression:

$$\hat{\beta}^{\text{PCR}} = (\mathbf{T_{|k}^T T_{|k}})^{-1}\mathbf{T_{|k}^T y} \tag{2.4}$$

The Nonlinear Iterative Partial Least Squares (NIPALS) algorithm can also compute Principal Component and its scores [10] [14].

Algorithm 1: Nonlinear Iterative Partial Least Squares (NIPALS)

Result: First k Principal Components
i = 1;
$\mathbf{X_i} = \mathbf{X}$;
while $i < k$ **do**
  repeat
    Choose $\mathbf{t_i}$ as any column of $\mathbf{X_i}$;
    Compute loadings $\mathbf{p_i} = \frac{\mathbf{X_i^T t_i}}{\mathbf{t_i^T t_i}}$;
    Let $\mathbf{p_i} = \frac{\mathbf{p_i}}{\sqrt{\mathbf{p_i^T p_i}}}$;
    Compute scores $\mathbf{t_i} = \frac{\mathbf{X_i p_i}}{\mathbf{p_i^T p_i}}$;
  until *Until* $\mathbf{t_i}$ *converges*;
  $\mathbf{X_{i+1}} = \mathbf{X_i} - \mathbf{t_i p_i^T}$;
  i += 1;
end

2.2 Partial Least Squares Regression

PLSR, much like PCR, also tries to reduce dimensionality via linear combinations of the inputs. In this technique, however, also takes into account the dependent variables $\mathbf{y}$. One key advantage of PLSR is that it seeks axes with most variance (like PCR) and high correlation with the dependent variables [7].

The main idea can be described as decomposing both the design matrix $\mathbf{X}$ and response matrix $\mathbf{Y}$ as follows [10], similarly to what was done in Section 2.1:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathbf{T}} \tag{2.5}$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^{\mathbf{T}} \tag{2.6}$$

Instead of simply running NIPALS on $\mathbf{X}$ and $\mathbf{Y}$ separately. PLSR uses information from $\mathbf{Y}$ to decompose $\mathbf{X}$ and *vice-versa* [10].

Algorithm 2: Partial Least Squares (PLS)

Result: First k Partial Least Squares directions
i = 1;
$\mathbf{X_i} = \mathbf{X}$;
$\mathbf{Y_i} = \mathbf{Y}$;
while $i < k$ **do**
  repeat
    Compute loading of $\mathbf{X_i}$ based on score of $\mathbf{Y_i}$: $\mathbf{p_i} = \frac{\mathbf{X_i^T u_i}}{\|\mathbf{X_i^T u_i}}\|$;
    Compute score of $\mathbf{X_i}$: $\mathbf{t_i} = \mathbf{X}\mathbf{p_i}$;
    Compute loading of $\mathbf{y_i}$ based on score of $\mathbf{X_i}$: $\mathbf{q_i} = \frac{\mathbf{Y_i^T t_i}}{\|\mathbf{Y_i^T t_i}}\|$;
    Compute score of $\mathbf{Y_i}$: $\mathbf{u_i} = \mathbf{Y}\mathbf{q_i}$;
  until *Until* $\mathbf{t_i}$ *converges*;
  $\mathbf{X_{i+1}} = \mathbf{X_i} - \mathbf{t_i}\mathbf{p_i^T}$;
  $\mathbf{Y_{i+1}} = \mathbf{Y_i} - \mathbf{u_i}\mathbf{q_i^T}$;
  i += 1;
end

After finding the k partial least squares directions from the algorithm above, the score matrices $\mathbf{T}$ and $\mathbf{U}$ are found.

2.3 Ridge Regression

Another option is shrink regression coefficients via a penalty term. As stated on [7], ”ridge coefficients minimize a penalized sum of squares”, as shown on Equation 2.7 and Equation 2.8.

$$\hat{\beta}^{\text{ridge}} = \text{argmin}_\beta \left\{ \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\} \tag{2.7}$$

Where $\lambda \geq 0$ is a parameter that controls strength of the penalization. This could also be written in matrix form:

$$\hat{\beta}^{\text{ridge}} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y} \tag{2.8}$$

3 Data

3.1 Data acquisition

The data was acquire at the Sensor and Actuator Systems (SAS) laboratory at Linköping University. The experiment — as shown on Figure 3.1 — consisted of exposing different gas combinations to the SiC-FET sensor under a certain frequency cycle and recording its response, measured in miliamperes (mA). The is then used to extract secondary features, namely average and slope values from certain regions of the frequency cycle.

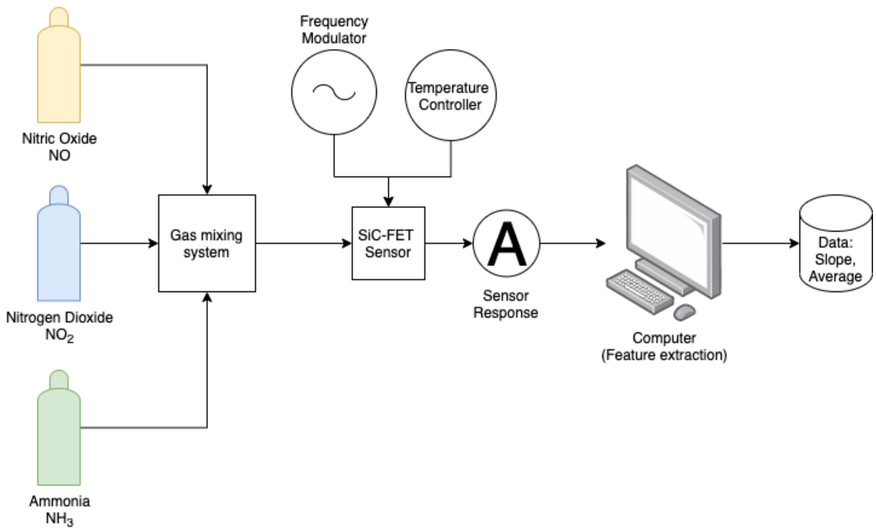

Figure 3.1: Schema of the data acquisition process.

In more detail, NO, NO_2 and NH_3 had five possible concentration values each: 10, 20, 40, 80 and 160 parts per million (ppm). The experiment was designed to encompass all possible combinations of these gasses, which totals to 125 different gas mixtures. Each feature was submitted to the same frequency cycle five times. The cycle consists of 16 unique frequencies: 0.05, 0.1, 0.25, 0.5, 1, 2, 5, 10, 25, 50, 100, 200, 500, 1000, 2500 and 5000 Hertz (Hz). A typical raw sensor response for frequency modulation experiments is shown on Figure 3.1.

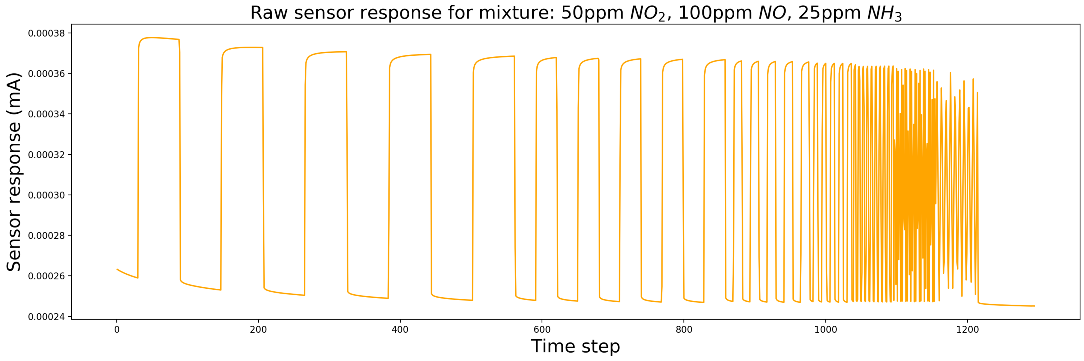

Figure 3.2: An example of row sensor response

For each frequency in each cycle, two slope and two average features were extracted. Finally, all 125 gas mixtures were subjected to the experiment three times, each time at a different temperature. Table 3.1 summarizes the data acquisition details.

Parameter	**Value**
Factors (gases)	3
Levels (concentrations)	5
Frequencies	16
Features per frequency	4 (2 slopes and 2 averages)
Features per cycle	64
Number of cycles	5
Data points per mixture	320
Number of mixtures	125
Datapoints per experiment	40.000
Number of experiments	3
Total data points	120.000

Table 3.1: Data acquisition details

For specific timestamps and measurement durations, the reader is referred to Appendix 8.1.

3.2 Raw data

TODO: add data itself. (Fingers crossed it will be this week.)

3.3 Secondary data

See above

4 Method

5 Results

6 Discussion

6.1 Results

6.2 Method

6.3 The work in a wider context

7 Conclusion

Bibliography

[1] R. Alberto Bernabeo, K. Webster, and M. Onofri. “Health and Environmental Impacts of Nox: An Ultra- Low Level of Nox (Oxides of Nitrogen) Achievable with A New Technology.” In: *Global Journal of Engineering Sciences* 3 (), pp. 2–7. DOI: 10.33552/gjes.2019.02.000540.

[2] ASTDR. “Sheet for ammonia published by the Agency for Toxic Substance and Disease Registry (ASTDR).” In: 2672 (2004), pp. 1–18. URL: https://www.atsdr.cdc.gov/MHMI/mmg126.pdf%5C%0Ahttps://www.atsdr.cdc.gov/mmg/mmg.asp?id=7&tid=2#bookmark02.

[3] Manuel Bastuck. “Improving the performance of gas sensor systems with advanced data evaluation, operation, and calibration methods.” PhD thesis. Jan. 2019, p. 267.

[4] Thirupathi Boningari and Panagiotis G. Smirniotis. “Impact of nitrogen oxides on the environment and human health: Mn-based materials for the NOx abatement.” In: *Current Opinion in Chemical Engineering* 13.x (2016), pp. 133–141. ISSN: 22113398. DOI: 10.1016/j.coche.2016.09.004. URL: http://dx.doi.org/10.1016/j.coche.2016.09.004.

[5] Christian Bur, Manuel Bastuck, Anita Lloyd Spetz, Mike Andersson, and Andreas Schütze. “Selectivity enhancement of SiC-FET gas sensors by combining temperature and gate bias cycled operation using multivariate statistics.” In: *Sensors and Actuators B: Chemical* 193 (2014), pp. 931–940. ISSN: 0925-4005. DOI: https://doi.org/10.1016/j.snb.2013.12.030. URL: https://www.sciencedirect.com/science/article/pii/S0925400513015037.

[6] Pio Forzatti. “Present status and perspectives in de-NOx SCR catalysis.” In: *Applied Catalysis A: General* 222.1 (2001). Celebration Issue, pp. 221–236. ISSN: 0926-860X. DOI: https://doi.org/10.1016/S0926-860X(01)00832-8. URL: https://www.sciencedirect.com/science/article/pii/S0926860X01008328.

[7] Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. *The elements of statistical learning*. Vol. 1. 10. Springer series in statistics New York, 2001.

[8] Susan Guthrie, Sarah Giles, Fay Dunkerley, Hadeel Tabaqchali, Amelia Harshfield, Becky Ioppolo, and Catriona Manville. *Impact of ammonia emissions from agriculture on biodiversity: An evidence synthesis*. Santa Monica, CA: RAND Corporation, 2018. DOI: 10.7249/RR2695.

[9] Richard Arnold Johnson, Dean W Wichern, et al. *Applied multivariate statistical analysis*. Vol. 5. 8. Prentice hall Upper Saddle River, NJ, 2002.

[10] Kee Siong Ng. "A simple explanation of partial least squares." In: *The Australian National University, Canberra* (2013).

[11] Jonathon Shlens. "A tutorial on principal component analysis." In: *arXiv preprint arXiv:1404.1100* (2014).

[12] USEPA. *Nitrogen Oxides Control Regulations*. https://www3.epa.gov/region1/airquality/nox.html. Accessed 2021-02-09. 2019.

[13] Svante Wold, Michael Sjöström, and Lennart Eriksson. "PLS-regression: a basic tool of chemometrics." In: *Chemometrics and Intelligent Laboratory Systems* 58.2 (2001). PLS Methods, pp. 109–130. ISSN: 0169-7439. DOI: https://doi.org/10.1016/S0169-7439(01)00155-1. URL: https://www.sciencedirect.com/science/article/pii/S0169743901001551.

[14] Kevin Wright. *The NIPALS algorithm*. https://cran.r-project.org/web/packages/nipals/vignettes/nipals_algorithm.html. Accessed: 2021-03-12.

8 Appendix

8.1 Appendix A: Data acquisition time stamps

Frequency (Hz)	Duration (s)	Feature	Start time (s)	End time (s)
0.05	20	Slope	0,0	0,4
		Average	9,6	10,0
		Slope	10,0	10,4
		Average	19,6	20,0
0.1	10	Slope	20,0	20,4
		Average	24,6	25,0
		Slope	25,0	25,4
		Average	29,6	30,0
0.25	4	Slope	30,0	30,4
		Average	31,6	32,0
		Slope	32,0	32,4
		Average	33,6	34,0
0.5	2	Slope	34,0	34,4
		Average	34,6	35,0
		Slope	35,0	35,4
		Average	35,6	36,0
1.0	2	Slope	36,0	36,4
		Average	36,6	37,0
		Slope	37,0	37,4
		Average	37,6	38,0
2.0	2	Slope	38,0	38,4
		Average	38,6	39,0
		Slope	39,0	39,4
		Average	39,6	40,0
5.0	2	Slope	40,0	40,4
		Average	40,6	41,0
		Slope	41,0	41,4
		Average	41,6	42,0
10.0	2	Slope	42,0	42,4
		Average	42,6	43,0
		Slope	43,0	43,4
		Average	43,6	44,0
25.0	2	Slope	44,0	44,4
		Average	44,6	45,0
		Slope	45,0	45,4
		Average	45,6	46,0
50.0	2	Slope	46,0	46,4
		Average	46,6	47,0
		Slope	47,0	47,4
		Average	47,6	48,0
100.0	2	Slope	48,0	48,4
		Average	48,6	49,0
		Slope	49,0	49,4
		Average	49,6	50,0
200.0	2	Slope	50,0	50,4
		Average	50,6	51,0
		Slope	51,0	51,4
		Average	51,6	52,0
500.0	2	Slope	52,0	52,4
		Average	52,6	53,0
		Slope	53,0	53,4
		Average	53,6	54,0
1000.0	2	Slope	54,0	54,4
		Average	54,6	55,0
		Slope	55,0	55,4
		Average	55,6	56,0
2500.0	2	Slope	56,0	56,4
		Average	56,6	57,0
		Slope	57,0	57,4
		Average	57,6	58,0
5000.0	2	Slope	58,0	58,4
		Average	58,6	59,0
		Slope	59,0	59,4
		Average	59,6	60,0

Figure 8.1: Data acquisition timestamps.