**Target audience of paper: scientists wanting to reproduce paper, professionals interested in gas analysis**

**Include only:**

* **Everything necessary to reproduce experiment**
  + **Describe own work extensively**
  + **Describe others’ work as far as my own knowledge and integrity allows**
* **Value of experiment: competition, value proposition, customers (all blood-diagnosed disease specialists, other markets?), key partners and contributions (QCL OEM, MATLAB, any company whose products are used)**

**Intro to gas analysis and chemometrics**

**Experimental setup**

Describe system on the basis of figure

Use Adonis’ paper, Zhe’s thesis, questionspaper

**Data calibration**

Wavenumber calibration – (compare with Adonis’) sound system calibration

H2O & CO2 calibration

**Concentration and compound determination**

Acetone and other molecules determination (what, the stuff Adonis did?)

~~Finding p-values + Explanation of p-values~~

~~Determine compounds with high intensity in regions of low p value~~

~~Determine compounds without need for p-values (and explain pros/cons)~~

~~Determine concentration from compounds~~

Remove CO2 in order to better estimate underlying ethyl-alcohol (check Adonis’ Analytical Chemistry paper for theory/understanding. Did this, Adonis’ paper uses the removal, but doesn’t state if or why it should improve concentration detection.)

**Clustering and classification**

Got 80% accuracy on sorting healthy vs. asthmatic using SVM

**Appendix**

List of molecules: Standard, chosen by p-region, total list of all molecules processed.

Info on the samples processed: Measurements were taken by [WHOM?] of children in the [SOPHIA KINDERZIEKENHUIS] in ROTTERDAM. 70 breath samples of healthy children and 70 breath samples of asthmatic children were used. Most children had two samples taken which are treated as separate measurements, so the measurements originate from about 70 children total.

Data as found in between algorithms such as molecule lists, wavenumber vector range and spacing, and regions of low p-values. Check separate paragraphs for what input/output they use or reference to.

**AT THE END OF EACH PARAGRAPH, ADD A SECTION OF WHAT INPUT IS USED AND WHAT PARAMETERS WERE CHOSEN IN PRACTICE, AND SHOW THE RESULTS.**

Every chapter individually readable, also works for consecutive readability. Requires page flipping if only results are wanted.

**OR SHOULD PARAMETERS AND RESULTS BE PUT IN A COMPLETE SEPARATE CHAPTER?**

Actual process more easily available, but requires page flipping to relate explanation to process.

**OR SHOULD CHOSEN PARAMETERS BE PUT INBETWEEN EXPLANATION?**

Makes both explanation of method and of actual process harder to read/less available. Seems to be preferred for specialised papers of short-medium length, only read by specialists.

**Discussion**

Cause of noise (in rawest signal): Thermal fluctuations expand laser box(, causing different wavelengths to be stable and emitted, seen in the signal as intensity fluctuations? [source: Adonis]). Piëzo hysteresis causes wavenumber (horizontal) shift and stretch. Mode-hopping is caused by piëzo + grating allowing short periods were multiple wavelengths are supported [source: Adonis]. Mode-hopping is also caused by random gas fluctuations?

In order to achieve higher accuracy with the QCL setup the intensity fluctuations should be minimized. The main cause X could be replaced by …. The wavenumber shifting and stretching caused by the hysteresis of the piëzo can be fixed using a similar method as for the wavenumber calibration.

Optional/what else was done during this project

Get NEC as in Florian’s paper

CO2H2O concentration determination

Acetone concentration determination (Done by Adonis, not me)

Helped Adonis filter molecules from organic molecule paper: Compound\_filter.m was used to find which of 199 molecules of paper found by adonis has absorption in the 1020-1100 cm-1 wavenumber region. These were 69 molecules. Can be found in email correspondence of 26/11/2014 “I'm working from home.” What was done with these molecules eventually? Used as standard molecules, or added to large collection of 456 molecules?

Learned R: reading and editing of Erasmus people their exhaled.r script

Learned Python: used for machine learning data ordering and CO2 removal script