How’s the Air Out There? Using a National Air Quality Database to Introduce First Year Students to the Fundamentals of Data Analysis

David Hall and Jessica D’eon (Corresponding Author)

19/03/2021

# Abstract

# Introduction

Whether we like it or not we’re living in an age of data, and the world of chemistry is no exception. From big-data atmospheric chemistry in climate-change models to machine-learning organic synthesis, every domain of chemistry is increasingly relying upon data-driven science. Consequently, undergraduate chemistry curriculums need to adapt to better prepare the next-generation of chemists. An oft-overlooked aspect of this is how exactly data (measurements, signals, etc.) are transformed into information (trends, correlation) and finally into knowledge. Moreover, the explicit teaching of these concepts is often neglected resulting in creasing student frustration. Motivated by this, and the need to transfer to a virtual laboratory environment as a result of Covid-19 social distancing restrictions, we sough to develop a new, remove learning compatible, experiment.

Real data is often imperfect and permeated with outliers and the fingerprints of gross experimental errors; data soon to be collected by undergraduate chemist is no different. However, acquiring sufficient data for analysis if often stressful for undergraduate students given time- and equipment-constraints in the teaching laboratory. To compensate for this, we chose to integrate actual measurements of atmospheric chemicals

Prominent atmospheric pollutants are structurally simple, and undergo reaction schemes comparable to those covered in introductory chemistry lectures. Ozone (O3) and nitrogen dioxide (NO2) are two choice candidates for analysis by undergraduate students. They are structurally simple molecules, and undergo reaction schemes comparable to those covered in introductory chemistry lectures. Notable of these compounds is their interdependent diurnal cycles. The relationship between O3~ and NO2 is so intimate, the term “odd-oxygen” (Ox) is used to express the sum of these two compounds, although the between O3 and NO2 can vary with environmental and anthorpogenic influences. Fortunately, with hourly measurements of O3 and NO2 since 1975, Environment and Climate Change Canada (ECCC) National Airborne Pollutants Surveillance (NAPS) program, there is no shortage of data to be analyzed.

Our Air quality lab described herein, is the result of our efforts. In this new experiment, first year students are introduced to fundamental data analysis concepts as they explore some of the chemistry of atmospheric chemical pollutants.

# Experimental Overview And Pedagogical Goals

This 3 hr data-analysis laboratory exercise uses publicly available data and open-source code (described in the Supplementary information), and has been run successfully in in the one-semester general chemistry course entitled “Chemistry: Phyiscal Principles” at the University of Toronto since Summer 2020. This course is most often conduceted in the first-term of the first-year of life-sciences/chemistry students. This exercise is the first-lab of five and is designed as much as an tutorial on data-analysis and microsoft excel as it is to explore atmospheric chemistry. It is devided into three parts:

Part A consist of pre-laboratory exercecies featuring explanaoarotry videos and material instrudocuing gas phase chemistry questions. The aim of this part is to situate students in the chemistry that forms the basis of their data analysis.

Part B

# Results and Pedagogical Outcomes

# Conclusions

# Supporting Information

# Author Information

# Acknowledgements

# References