

THESIS

METHODS TO DETECT AND ANALYSE VOLATILE ORGANIC CARBONS
USING LOW COST REAL-TIME SENSORS

Submitted by

Vatsal Gupta

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2019

Master's Committee:

Advisor: Kenneth Carlson

Ellison Carter

Jay Ham

ABSTRACT

METHODS TO DETECT AND ANALYSE VOLATILE ORGANIC CARBONS USING LOW COST REAL-TIME SENSORS

VOCs are ubiquitous and can be found not only as vapors in the air but also as soil gas and dissolved in ground water. Vapor intrusion occurs when volatile organic compounds from contaminated soil or groundwater migrate upwards toward the ground surface and into overlying buildings or surfaces through gaps and cracks in the ground.

In this thesis I have detailed several statistical analysis techniques and used these techniques on data that I obtained from active real-time soil gas and ground water quality monitoring sensors placed around an abandoned oil and gas well in Longmont, Colorado, to see if there were VOCs still being released from the site.

The main goal of this study was to develop a more precise setup for real-time VOC release monitoring and help regulate fracking sites more efficiently and to analyze the data collected faster and more accurately. Another goal of this study was to bridge the gap between laboratory sampling and real-time on-site testing. From the results, we were able to analyze the movement of the contaminant plume using real time sensing and were also able to identify most of the constituents of the contaminants using in-situ data according to EPA method 18.

ACKNOWLEDGMENTS

I would first like to thank my thesis advisor Dr. Kenneth Carlson of environmental engineering branch at Colorado State University. He consistently allowed this paper to be my own work but steered me in the right the direction whenever he thought I needed it.

I would like to thank the city of Longmont for allowing me to use their data for my research. I would like to thank Ajax Analytics Inc. for providing me with the opportunity to work on this project as well as for providing me with the instruments/sensors to perform data collection for the purpose of this analysis.

I would also like to thank the experts who were involved in the validation survey for this research project: Jason Elkins, Brent Buck and Nikken Cullman. Without their passionate participation and input, the validation survey could not have been successfully conducted.

I would also like to acknowledge Dr. Ellison Carter and Dr. Jay Ham at Colorado State University as the second readers of this thesis, and I am gratefully indebted to them for their very valuable comments on this thesis.

Finally, I must express my very profound gratitude to family and friends for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

TABLE OF CONTENTS

ABSTRACT.....	..ii
ACKNOWLEDGEMENTS.....	..iii
LIST OF TABLES.....	..iv
LIST OF FIGURES.....	..v
LIST OF KEYWORDS.....	..vii
Chapter 1 – Introduction	1
1.1 Background	1
1.2 Problem Statement.....	4
Chapter 2. Literature Review.....	6
2.1 Multiple definitions of the term VOC.....	6
2.2 Sensors to detect VOCs.....	8
2.3 Methods to detect VOCs	19
2.4 Sources of VOCs.....	21
2.5 Thesis objectives	23
Chapter 3 – Methods	24
3.1 Description of the assigned site	24
3.2 Description of the soil gas VOC sensor used	28

3.3 Description of groundwater quality monitoring sensor.....	31
3.4 Data Collection	33
Chapter 4 – Data analysis	34
4.1 Summary of Raw Data	34
4.2 Data Validation	42
4.3 Trend Analysis	43
4.3.1 Linear trend analysis	43
4.3.2 X bar and S charts	47
4.3.3 Mann Kendall and Sen`s slope test	49
4.4 Time Series Analysis	53
4.5 Regression and Correlation	61
4.5.1 Regression of soil gas with ground water data	62
4.5.2 Regression of soil gas with weather data	67
4.6 Chemical compound identification by In – Situ Tests	76
Chapter 5 – Conclusion	78
Chapter 6. – Recommendations for future studies	82
References	85

LIST OF TABLES

TABLE 2.1 Definitions of VOCs and related compounds by various regulatory organizations.....	6
TABLE 2.2 Electro-chemical sensors sensitivity for measurement of toxic compounds in air.....	10
TABLE 2.3 Lower explosion limits of certain VOCs for Pellistor sensors.....	15
TABLE 2.4 Main application of currently commercially available of VOC sensors.....	16
TABLE 2.5 Indoor air quality guidelines and measurement standards.....	20
TABLE 2.6 VOC emission sources from common building products.....	21
TABLE 2.7 Possible VOC reaction products and its sources.....	22
TABLE 3.1 Chemical properties of the soil in the area of study.....	26
TABLE 3.2 Specifications for Aquatroll 400.....	32
TABLE 4.1 Mann Kendall test results for VOC concentrations at SVP 01	50
TABLE 4.2 Sen slope test results for VOC concentrations at SVP 01	50
TABLE 4.3 Mann Kendall test results for VOC concentrations at SVP 02.....	51
TABLE 4.4 Sen slope test results for VOC concentrations at SVP 02	51
TABLE 4.5 Mann Kendall test results for VOC concentrations at SVP 03	52
TABLE 4.6 Sen slope test results for VOC concentrations at SVP 03	52
TABLE 4.7 Pearson Correlation test for VOC concentrations at SVP 01.....	63
TABLE 4.8 Pearson Correlation test for VOC concentrations at SVP 02	64
TABLE 4.9 Pearson Correlation test for VOC concentrations at SVP 03.....	65
TABLE 4.10 ANOVA test results for VOC concentrations at SVP 02 and weather data.....	67
TABLE 4.11 ANOVA model summary for VOC concentrations at SVP 02 and weather data.....	68
TABLE 4.12 ANOVA coefficient for VOC concentrations at SVP 02 and weather data.....	69
TABLE 4.13 ANOVA test results for VOC concentrations at SVP 03 and weather data.....	72
TABLE 4.14 ANOVA model summary for VOC concentrations at SVP 03 and weather data.....	73
TABLE 4.15 ANOVA coefficients for VOC concentrations at SVP 03 and weather data.....	73
TABLE 4.16 In-Situ test result of sample collected onsite in accordance with EPA method 18.....	76

LIST OF FIGURES

FIGURE 2.1 Cox chart vapor pressure plots for VOCs and other common chemicals.....	7
FIGURE 2.2 Commercially available sensors and its detectable concentration ranges of VOC.....	9
FIGURE 2.3 Schematic diagram of a photoionization sensor	12
FIGURE 2.4 Ranges wave numbers characteristic for selected functional groups of VOCs.....	13
FIGURE 3.1 Map of the study area	24
FIGURE 3.2 Soil type of the area of study.....	25
FIGURE 3.3 Map of the monitoring stations.....	27
FIGURE 3.4 Schematic of NDIR sensor.....	29
FIGURE 3.5 Picture of groundwater quality sensor.....	32
FIGURE 4.1 Linear graph of VOC concentration in soil gas at SVP 01 vs time.....	35
FIGURE 4.3 Linear graph of VOC concentration in soil gas at SVP 02 vs time.....	35
FIGURE 4.5 Linear graph of VOC concentration in soil gas at SVP 03 vs time.....	35
FIGURE 4.2 Interval plot of VOC concentration in soil gas at SVP 01	37
FIGURE 4.4 Interval plot of VOC concentration in soil gas at SVP 02	38
FIGURE 4.6 Interval plot of VOC concentration in soil gas at SVP 03	39
FIGURE 4.7 Linear graph of ground water temperature and pH at MW 02 vs time	40
FIGURE 4.8 Linear graph of ground water temperature and pH at MW 03 vs time	41
FIGURE 4.9 Linear graph of ground water temperature and pH at MW 05 vs time	42
FIGURE 4.10 Linear trend plot of VOC concentration in soil gas at SVP 01	44
FIGURE 4.11 Linear trend plot of VOC concentration in soil gas at SVP 02	45
FIGURE 4.12 Linear trend plot of VOC concentration in soil gas at SVP 03	46
FIGURE 4.13 Xbar-S chart of VOC concentration in soil gas at SVP 01	47
FIGURE 4.14 Xbar-S chart of VOC concentration in soil gas at SVP 02	48
FIGURE 4.15 Xbar-S chart of VOC concentration in soil gas at SVP 03	48
FIGURE 4.16 Decomposed data of VOC concentration in soil gas at SVP 01	56
FIGURE 4.17 Seasonal analysis graphs of VOC concentration in soil gas at SVP 01	56

FIGURE 4.18 Decomposed data of VOC concentration in soil gas at SVP 02	57
FIGURE 4.19 Seasonal analysis graphs of VOC concentration in soil gas at SVP 02	57
FIGURE 4.20 Decomposed data of VOC concentration in soil gas at SVP 03	58
FIGURE 4.21 Seasonal analysis graphs of VOC concentration in soil gas at SVP 03	58
FIGURE 4.22 Autocorrelation graphs of VOC concentration in soil gas at SVP 01	59
FIGURE 4.23 Autocorrelation graphs of VOC concentration in soil gas at SVP 02	60
FIGURE 4.24 Autocorrelation graphs of VOC concentration in soil gas at SVP 03	61
FIGURE 4.25 Scatter plot of VOC concentration at SVP 01 and GW conductivity at MW 02	63
FIGURE 4.26 Scatter plot of VOC concentration at SVP 02 and GW conductivity at MW 05	64
FIGURE 4.27 Scatter plot of VOC concentration at SVP 02 and GW temperature at MW 05	65
FIGURE 4.28 Cross correlation of VOC conc. at SVP 02 and GW conductivity at MW 05	66
FIGURE 4.29 Line graph of VOC conc. at SVP 02 and air and soil temperature vs time	70
FIGURE 4.30 Line graph of VOC conc. at SVP 02 and humidity vs time.....	70
FIGURE 4.31 Line graph of VOC conc. at SVP 02 and vapor pressure vs time.....	71
FIGURE 4.32 Line graph of VOC conc. at SVP 03 and air and soil temperature vs time.....	74
FIGURE 4.33 Line graph of VOC conc. at SVP 03 and humidity vs time.....	74
FIGURE 4.34 Line graph of VOC conc. at SVP 03 and vapor pressure vs time	75
FIGURE 4.35 Spatial analysis of the laboratory report	77

LIST OF KEYWORDS

Active Sensors: A sensing device that requires an external source of power to operate

Analog Sensors: A sensing device that produce continuous analog output signal

Digital Sensors: Sensors in which data conversion and data transmission takes place digitally

EC Sensor: Electrochemical sensor

EPA: Environmental protection agency

Ex-Situ Sensing: Measuring the target gas off-site mostly in a laboratory with a gas chromatograph.

In-Situ Sensors: A sensing device that takes a reading of the target gas at the test site

MOS: Metal oxide semiconductor sensors

NDIR: Nondispersive infrared sensors

NMVOC: Non-Methane Volatile Organic Compounds

NPI: National pollution inventory, Australia

ORP: Oxidation reduction potential

Passive Sensors: Sensor that detects and respond to some type of input from the physical environment.

PID: Photoionization detectors

Real-time Sensors: A sensing device that continuously takes a reading of the target gas at the site

Sample based Sensing: Sensors that rely on capturing the gas from the test site at a timed interval

TVOC: Total Volatile Organic Compounds

CHAPTER 1: INTRODUCTION

1.1 Background

Volatile organic compounds, or VOCs are organic chemical compounds whose composition makes it possible for them to evaporate under normal atmospheric conditions of temperature and pressure **(1)**. This is the general definition of VOCs that is used in the scientific literature and is consistent with the definition used for air quality analysis. Though there are a lot of discrepancies on which gases should be measured and considered as VOCs since there is no clear and universally accepted definition of VOC as each agency employs its own definitions and exceptions to the list. The following are some of the definitions of VOCs:

1. EPA definition - Volatile organic compounds means any compound of carbon, excluding carbon monoxide, carbon dioxide, carbonic acid, metallic carbides or carbonates and ammonium carbonate, which participates in atmospheric photochemical reactions.
2. NPI (Australia) VOC definition: Total VOC are defined as any chemical compound based on carbon chains or rings with a vapour pressure greater than 0.01 kPa at 293.15 K (i.e. 20°C), that participate in atmospheric photochemical reactions. (Specifically excluded are: □ carbon monoxide; □ methane; □ acrylamide; □ benzene hexachloro; □ biphenyl; □ chlorophenols; □ n-dibutyl phthalate; □ ethylene glycol; □ di-(2-ethylhexyl) phthalate (DEHP); □ 4,4-methylene bis 2,4 aniline (MOCA); □ Methylenebis; □ Phenol; and □ toluene-2,4-diisocyanate). **(2)**

3. The European Union uses the boiling point, rather than its volatility in its definition of VOCs. A VOC is any organic compound having an initial boiling point less than or equal to 250° C measured at a standard atmospheric pressure of 101.3 kPa. Since the volatility of a compound is generally higher the lower its boiling point temperature, the volatility of organic compounds are sometimes defined and classified by their boiling points.

VOCs are produced both naturally, by plants, animals, microbes, volcanoes and forest fires, and anthropogenically in the form of fuels, paints, refrigerants, cleaning products, adhesives, personal care products and aerosol sprays. The results of an American Geophysical Union study show total global anthropogenic VOC emissions of about 110 Tg/yr in 2012. This estimate is about 10% lower than global VOC inventories developed by other researchers. The study identifies the United States as the largest emitter (21% of the total global VOC), followed by the (former) USSR, China, India, and Japan. Globally, fuel wood combustion and savanna burning were among the largest VOC emission sources, accounting for over 35% of the total global VOC emissions. The production and use of gasoline, refuse disposal activities, and organic chemical and rubber manufacturing were also found to be significant sources of VOC emissions. **(3)**

On the other hand global emissions of BVOCs are estimated to be about 1150 teragrams of carbon per year (TgC year^{-1} , 1 teragram = 10^{12} g), which exceeds those of their anthropogenic counterparts by about a factor of 10. Typically, BVOCs have much shorter atmospheric lifetimes than anthropogenic VOCs due to faster reaction rates with OH. **(4)**

Many VOCs form ground-level ozone by “reacting” with sources of oxygen molecules such as nitrogen oxides (NO_x), and carbon monoxide (CO) in the atmosphere in the presence of sunlight. However, only some VOCs are considered “reactive” enough to be of concern. Some VOCs, such as styrene and limonene, can react with nitrogen oxides or with ozone to produce new oxidation products and secondary aerosols, which can cause sensory irritation symptoms. VOCs contribute to the formation of tropospheric ozone and smog. **(5)**

The ability of organic chemicals to cause health effects varies greatly from those that are highly toxic, to those with no known health effects. As with other pollutants, the extent and nature of the health effect will depend on many factors including level of exposure and length of time exposed. Eye and respiratory tract irritation, headaches, dizziness, visual disorders, and memory impairment are among the immediate symptoms that some people have experienced soon after exposure to some organics. At present, not much is known about what health effects occur from the levels of organics usually found in homes. Many organic compounds are known to cause cancer in animals; some are suspected of causing, or are known to cause, cancer in humans. **(6)**

1.2 Problem statement

Since there is no universally agreed definition hence there is no well-defined detection technique which detects all the VOC compounds of importance. There is also no low-cost sensor currently in the market that can accurately identify and quantify all the VOCs of interest.

Portable Gas Chromatograph enables us to analyze specific VOCs (volatile organic compounds) in soil, in the field with the same quality of results obtained from a laboratory but this instrument is highly expensive and hence it is not feasible to use multiple gas chromatographs in the field for spatial and temporal analysis. This poses a problem to regulatory boards and city officials since they cannot efficiently measure, control, and verify if individuals or organizations are going over the safe permissible limit and affecting others around them. Hence there is a need for analysis techniques using low-cost sensors to better estimate the release of VOCs.

Most industrial emissions of VOCs in the United States are regulated and controlled at the source, either as part of ozone reduction or as a hazardous air pollutant. Other sources of VOCs include traffic, mobile equipment, area sources such as wastewater lagoons, and some natural sources. In a study reported by USA Today in 2008, it was reported that 435 schools across the country had been identified as being potentially exposed to dangerous levels of toxic industrial chemicals. (7) It was also reported that over half of the nation's schools are located "in what the government calls 'vulnerable zones' – areas close to industrial sites that they could be affected by an accident". This widespread awareness of the surrounding potential sources of toxic chemicals, as well as all contaminants of concern, when evaluating the cleanliness and dependability of the ambient air surrounding new construction sites is necessary. (8)

In this thesis I have detailed several statistical analysis techniques and used these techniques on real time data that I obtained from soil gas and ground water quality monitoring sensors placed around an abandoned oil and gas well in Longmont, Colorado, to see if there were VOCs still being released from the site. I performed spatial and temporal analysis on the data obtained from low-cost sensors and was able to identify a contaminated underground plume and over time was able to identify and trace the movement of the plume spatially by trend analysis and event detection techniques. Samples were collected at the site as well and tested in-situ to determine the composition of contaminated plume and collaborate the results of the on-site tests.

CHAPTER 2: LITERATURE REVIEW

2.1 Multiple definitions of the term VOCs

Every regulatory agency has its own definitions of VOCs, some of the various definitions from regulatory agencies for VOCs have been described in the table below.

Table 2.1. Definition of VOCs and its related compounds by various regulatory organizations

Term	Agency/Organization; Report/Publication	Definition
Volatile organic compound (VOC)	ASTM International; <i>ASTM D1356, Standard Terminology Relating to Sampling and Analysis of Atmospheres</i> (ASTM 2005a)	An organic compound with a saturation vapor pressure greater than 40.1×10^{-3} in. H ₂ O (10^{-2} kPa) at 77°F (25°C) (where <i>organic chemical</i> = a carbon-based compound in which the element carbon is attached to other carbon atom(s), hydrogen, oxygen, or other elements in a chain, ring, or three-dimensional structure).
	International Society of Indoor Air Quality and Climate (ISIAQ); "Glossary of the Indoor Air Sciences" (ISIAQ 2006)	Organic compounds with boiling points ranging from a lower limit between 122°F (50°C) and 212°F (100°C) and an upper limit between 464°F (240°C) and 500°F (260°C), where the upper limits represent mostly polar compounds.
	U.S. Environmental Protection Agency (EPA); <i>Code of Federal Regulations</i> (40 CFR 51.100(s)) (GPO 2009)	Any compound of carbon, excluding carbon monoxide, carbon dioxide, carbonic acid, metallic carbides or carbonates, and ammonium carbonate, that participates in atmospheric photochemical reactions (with dozens of exceptions for compounds determined to have negligible photochemical reactivity).
Semi-volatile organic compound (SVOC)	ASTM International; <i>ASTM D1356, Standard Terminology Relating to Sampling and Analysis of Atmospheres</i>	An organic compound with a saturation vapor pressure between 40.1×10^{-3} and 40.1×10^{-9} in. H ₂ O (10^{-2} and 10^{-8} kPa) at 77°F (25°C).
Very volatile organic compound (VVOC)	World Health Organization (WHO)	Compound with boiling point in range from below 32°F (0°C) to between 122°F and 212°F (50°C and 100°C).

(Source: <https://www.epa.gov/indoor-air-quality-iaq/volatile-organic-compounds-impact-indoor-air-quality>)

Some agencies like ASTM consider volatility while other agencies like WHO consider boiling point of compounds for defining VOCs. In general, liquids with higher vapor pressures will have lower boiling points. However, exceptions do exist as described in the example below.

Antoine Equation, an empirical relation between vapor pressure and temperature. Plotting the Antoine equation for a set of liquids lets one visualize the relationship between vapor pressure and temperature very quickly; these plots are called Cox charts. For example, examine the Cox chart below:

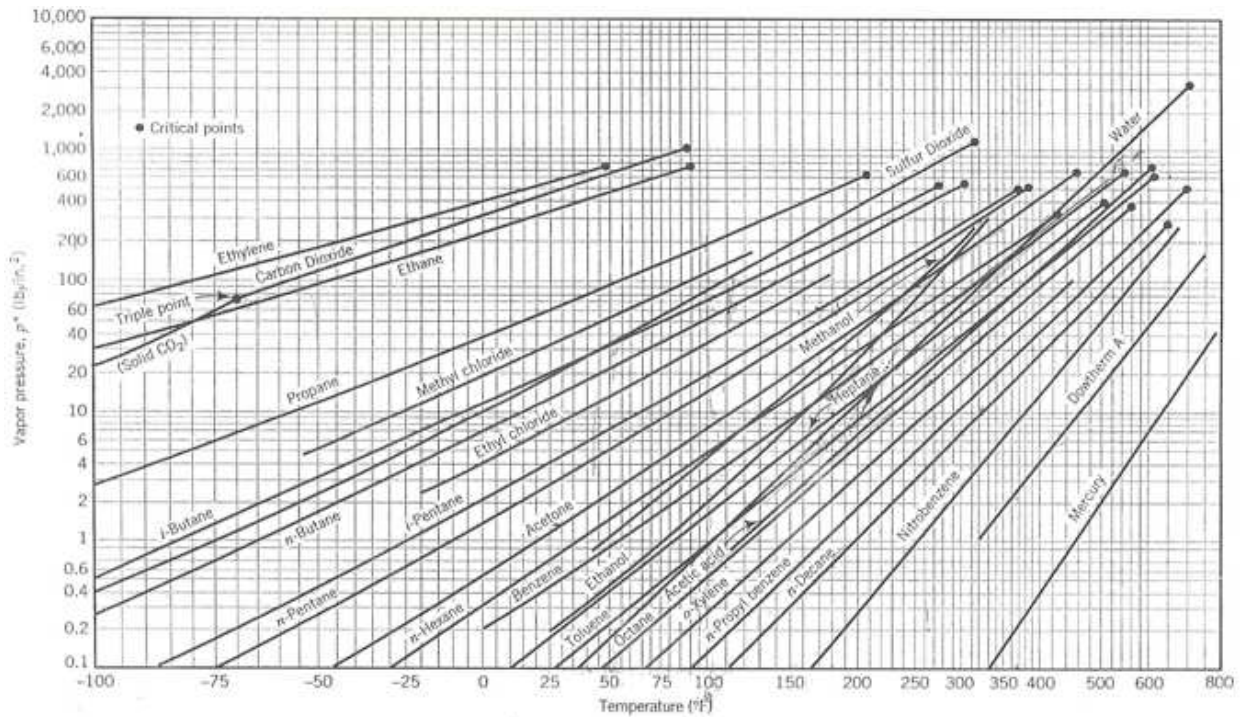


Figure 2.1. Cox chart vapor pressure plots (A.S. Foust et al., Principles of Unit Operations [1960])

Now look at the relationship between ethanol and benzene. At 50 F, benzene has a higher vapor pressure than ethanol. At pressures below ~10 psi, the general trend holds, and benzene has a lower boiling point. However, the lines for benzene and ethanol cross at ~10 psi, so above ~10 psi benzene has a higher boiling point than ethanol.

Most regulatory boards, including the EPA, CARB and NPI, do not consider methane in their list of VOCs. Methane is excluded in air-pollution contexts because it is not harmful. Its low reactivity and thus long life-time in the atmosphere, however, makes it an important greenhouse gas. Non-methane volatile organic compounds (NMVOCs) are a large variety of chemically different compounds, such as benzene, toluene, xylene, formaldehyde, cyclohexane, trichloroethane. NMVOCs are volatile organic compounds, but with methane excluded. An important subset of NMVOCs are the non-methane hydrocarbons (NMHCs). Sometimes NMVOC is also used as a sum parameter for emissions, where all NMVOC emissions are added up per weight into one figure called TVOC (total volatile organic carbon). (9)

2.2 Sensors to detect VOCs

Detecting VOCs is highly dependent on the method used for its detection. This depends on the sensors used for its detection and the VOC's that it is capable of detecting. The most popular real-time sensors currently commercially available for the measurement of volatile organic compounds (VOCs) in outdoor and indoor air are classified into:

- Electrochemical Sensor (EC)
- Metal Oxide Semiconductor (MOS)
- Photoionization Detectors (PID)
- Nondispersive infrared sensors (NDIR)
- Pellistor sensors

The working principles of each of these sensors is described in the following sections and the sensitivity of each of these sensors is shown in the diagram below:

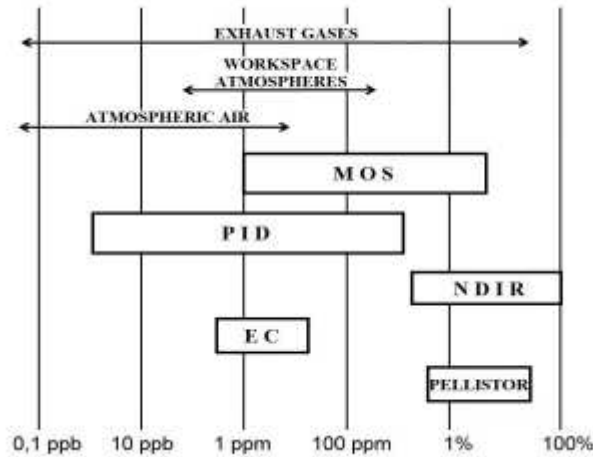


Figure 2.2 Commercially available sensors and its detectable concentration ranges of VOC. (11)

2.2.1 EC—Electrochemical Sensors:

Principles of Operation: Electrochemical sensors usually contain an acid electrolyte, sensing electrode, counter electrode, third reference electrode, and a gas-permeable membrane. As the air diffuses into the cell, certain gases oxidize on the sensor and a voltage differential is produced. The current produced by the chemical reaction is proportional to the concentration level of the reacting gas. The sensing electrode is designed to catalyze a specific reaction. (12)

Advantage: The cost of sensors is comparatively cheaper compared to other sensor types and can be used to detect certain reactive gases.

Disadvantages: While this technology is somewhat specific, other common gases will react at different levels and be detected, resulting in false positives and false alarms. These sensors have a limited lifetime and deplete over a period of time. The depletion rate is primarily determined by the sensor's exposure to the reactant gases. Deciding when to recalibrate these sensors to maintain a specific accuracy can be a problem. On average, most equipment manufacturers using electrochemical sensors recommend recalibration every three months, but this is influenced by the sensor's reactant gas exposure and the required accuracy level. Electrochemical sensors will also degrade when exposed to high humidity conditions.

Table 2.2. Example of detection limits and reactivity with other chemicals other than VOCs of an Electrochemical Sensor by Environment sensor co. (12)

Pollutants	Range (ppm)	Resolution (ppm)	Response Time (s)
ammonia	0-50	0.5	150
carbon monoxide	0-1000	0.5	35
chlorine	0-20	0.1	60
ethylene oxide	0-20	0.1	140
formaldehyde	0-30	0.01	60
glutaraldehyde	0-20	0.01	60
hydrogen sulfide	0-50	0.1	30
nitric oxide	0-1000	0.5	10
nitrogen dioxide	0-20	0.1	35
sulfur dioxide	0-20	0.1	15

As you can see from the table above, this particular sensor is affected by ammonia, CO as well as other common gases, which makes it really unsuitable for accurate readings.

2.2.2 MOS—Metal Oxide Semiconductor

Principles of Operation: MOS sensors consist of a metal oxide semiconductor such as tin dioxide, on sintered alumina ceramic located inside a flame arrestor. Sensitivity to specific gases may be altered by changing the temperature of the sensing element. **(13)**

Advantage: MOS sensors will detect gases at lower ppm levels than NDIR and pellistors.

Disadvantage: These sensors are less gas specific than electrochemical sensors and react to many types of gases, producing many more false positives and false alarms. MOS sensors are also extremely sensitive to humidity, temperature and pressure.

2.2.3 PID—Photoionization Detectors

Principle of Operation: The photoionization sensors ionize (decay into charged particles) neutral molecules of chemical compounds. When diffusing VOCs molecules enter the region of UV lamp impact, they are ionized by photons. Then formed ions are directed between two polarized electrodes. The ions move towards the electrodes in an electric field generated by an electrometer. In this way a current flow is generated, which is then converted into voltage signal. This signal is proportional to concentration of the compounds subjected to ionization. The photoionization sensors utilize electrodeless ultraviolet lamps (wavelength 10–400 nm). Operation of the lamp consists in excitation of the filling gas (most often krypton, xenon, radon)

via the impact of external electromagnetic field. (14) This type of sensor is most frequently applied for measurement of summary concentration of volatile organic compounds.

A schematic of the photoionization sensor is shown below.

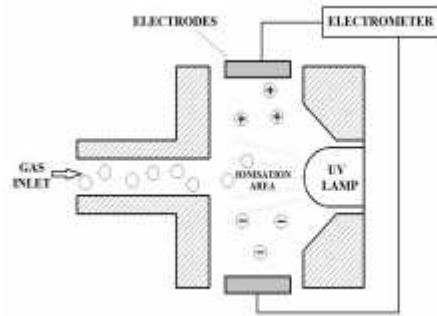


Figure 2.3. Schematic diagram of a PID sensor (14)

Advantage: Moderately low cost, ease of operation and near-instantaneous results coupled with its acceptance by regulation agencies makes this a very suitable option for use in analysis of VOCs. It can also detect VOCs at low ppm level.

Disadvantage: Contamination such as dust particles settling on the sensor lenses and humidity can affect the sensors results and effective working capability.

2.2.4 NDIR—Nondispersive infrared sensors

Principle of Operation: Flammable gases and vapors from the VOCs group are subjected to characteristic absorption of radiation from the infrared range. The ranges of oscillation frequency (wave number) characteristic for selected functional groups of VOCs are presented in the figure below. The principle of operation of this type of sensor consists in arranging a source of infrared radiation along an optical line with a detector. When an analyzed gas appears in a measurement chamber, it absorbs radiation of a particular wavelength and, following the Lambert-Beer law, there is a decrease in radiation reaching the detector, which is converted into electrical signal. Intensity of infrared radiation is diminished as it passes through the measurement cell. This reduction of light intensity is proportional to concentration of the gases or flammable vapors subjected to detection. An important element of the sensor is an optical filter, which passes absorbed light of defined wavelength, thus providing selectivity of particular sensor. Some designs possess additional (reference) chamber, which is filled with non-absorbent gas (typically nitrogen). In this case the signal is generated based on a difference in readings from the detectors of both chambers. (15)

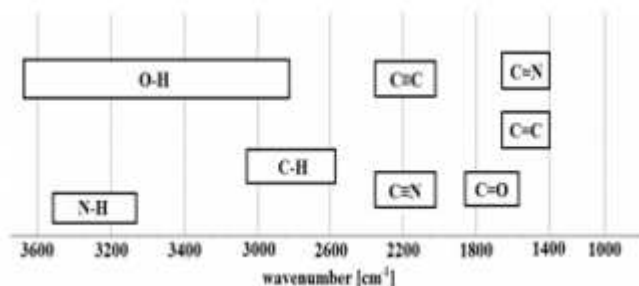


Figure 2.4. Ranges wave number characteristics for selected functional groups of VOCs

Advantage: Longer lifetime when compared to other sensors due to no sensor burn out nor any sensor deterioration with exposure to gases.

Disadvantage: A high detection limit, spectral interference and false readings caused by multiple C-H bonds reduce the effectiveness of this sensor.

2.2.5 PELLISTOR—Thermal sensor.

Principle of Operation: A phenomenon of explosion can be initiated in a mixture of flammable gas and air only within precisely defined concentration range. Lower explosion limit (LEL) determines the minimum concentration of the substance, which can react in a rapid combustion process. Upper explosion limit (UEL) describes the maximum amount of the fuel, at which the mixture contains enough oxidizer to initiate the explosion. The values of LEL and UEL differ for various substances and are usually expressed with respect to air. Concentrations of explosive substances below LEL and above UEL allow for safe operation. Table 2.3. Presents the values of LEL for selected substances from the VOCs group. The principle of operation of this type of sensor consists in diffusion of a mixture of air and particular flammable compound through porous sinter towards porous sensor surface. The porous element contains miniature coil made of platinum wire. Electric current flows through a coil made of platinum wire and heats the pellistor up to a few hundred degrees Celsius. The reaction at the catalytic surface releases heat, which increases temperature of the platinum coil, inducing an increase in its resistance. The pellistor is most commonly implemented as one arm of the Wheatstone bridge, the output of which is the final signal. In the case of temperature changes the output bridge signal is

proportional to heat of reaction. Increase in temperature is a measure of concentration of flammable gas substance. (16)

Advantage: Safety applications of these sensors are mainly interested in % LEL measurements which is provided accurately by this sensor.

Disadvantage: Prone to permanent poisoning by lead, sulphurs, chlorinated or silicone compounds and must be calibrated frequently.

Table 2.3 Lower explosion limits of selected VOCs (16)

VOCs	Lower Explosion Limit (% v/v)	VOCs	Lower Explosion Limit (% v/v)
acetone	2.5	ethyl acetate	2.0
benzene	1.2	styrene	1.0
n-butanol	1.7	toluene	1.1
cyclohexane	1.0	1,3-butadiene	1.4
1,4-dioxane	1.9	n-butane	1.4
ethanol	3.1	methyl chloride	7.6
diethyl ether	1.7	dimethyl ether	2.7
methanol	6.0	ethylene oxide	2.6
n-hexane	1.0	methane	4.4
n-octane	0.8	propane	1.7

The following tables show the applications of the various types of sensors currently available for the measurement of VOCs.

Table 2.4. Applications and compounds identified by the various types of sensors (17)

Sensor Type	Applications	Compounds
MOS	<ul style="list-style-type: none"> - Urban air monitoring - Roadside monitoring - Industrial perimeter measurement - Indoor Air Quality - Smart home & Internet of Things modules - Medical equipment - Fire detection - Ventilation control/air cleaners 	alcohols, aldehydes, aliphatic hydro-carbons, amines, aromatic hydro-carbons (petrol vapors, etc.), carbon oxides, CH ₄ , LPG, ketones, organic acids.
PID	<ul style="list-style-type: none"> - Industrial hygiene & safety monitoring - Confined space entry - Soil contamination and remediation - Hazmat sites and spills - Arson investigation - Low concentration leak detection - EPA method 21 and emissions monitoring 	VOC's with proper ionisation potential (isobutylene, aromatic hydrocarbons)
NDIR	<ul style="list-style-type: none"> - Indoor Air Quality - Combustion process monitoring - Biogas production 	infrared absorbing VOC's (especially methane)
EC	<ul style="list-style-type: none"> - Breathalyzer - Environmental protection - Odorants monitoring in natural gas applications - Urban and industrial area monitoring - Mobile monitoring applications 	ethanol, formaldehyde, mercaptanes
PELLISTOR	<ul style="list-style-type: none"> - Hydrogen and combustible gas leak detectors - Detectors for fuel cells - Explosive atmosphere monitoring 	most combustible gases and vapours (iso-butane, propane, methane)

2.2.6 Other classification methods for sensors:

Other than the classification method mentioned above, the sensors can also be categorized, or sub categorized by the following classifications:

2.2.6.1 Active vs Passive

An active sensor is a sensing device that requires an external source of power to operate while passive sensors simply detect and respond to some type of input from the physical

environment. Active sensors are also widely used in manufacturing and networking environments for example to monitor industrial machines or data center infrastructure, so anomalies can be detected and components can be repaired or replaced before they break and shut everything down.

Examples of other active sensor-based technologies include: scanning electron microscopes, LiDAR, radar, GPS, x-ray, sonar, and infrared. However, as can be the case with some sensors, infrared light sensors exist in both active and passive forms.

2.2.6.2 Real-time vs sample-based sensors

Real-time sensors take a reading of the site continuously while sample-based sensors rely on capturing the gas from the test site at a monthly or annual basis then testing them off site mostly in a laboratory with a gas chromatograph. For the purpose of our analysis we focus on real-time sensors or sample-based sensors with a practically small sample time period such that it can be considered practically real-time for all intents and purposes

2.2.6.3 In situ vs Ex Situ

In situ can refer to where a measurements or remediation of a polluted site is performed in the actual site, contrary to ex situ where samples are examined elsewhere, off site mostly in a laboratory.

2.2.6.4 Digital vs Analog

There are different types of sensors that produce continuous analog output signal and these sensors are considered as analog sensors. This continuous output signal produced by the analog sensors is proportional to the measured substance.

Electronic sensors or electrochemical sensors in which data conversion and data transmission takes place digitally are called as digital sensors. These digital sensors are replacing analog sensors as they can overcome the drawbacks of analog sensors. The digital sensor consists of majorly three components: sensor, cable, and transmitter. In digital sensors, the signal measured is directly converted into digital signal output inside the digital sensor itself. And this digital signal is transmitted through cable digitally.

2.2.6.5 Low-cost vs High Cost

Gas chromatographs cost tens of thousands of dollars to buy and hence they are not feasible to have multiple sensors in the site and are considered high cost sensors. While PID and NDIR sensors cost a few thousand dollars and hence more of these sensors could be used on the site when compared to gas chromatographs. EC, MOS and pellistors are much lower in cost when compared to the other sensors and are hence considered low-cost sensors. Low cost sensors lose specificity, accuracy, maybe reliability but gains spatial coverage (since more sensors could be used at the same cost) and potentially temporal coverage.

2.3 Methods to detect VOCs

There are several guidelines provided by the regulatory boards on how to detect VOCs and which type of sensors to use. Since every sensor has different detection limits and other various limitations, each sensor type provides a different value of VOCs detected.

VOC detection and quantification are highly method dependent. A given sampling and analysis system cannot capture or fully respond to all the VOCs present in any indoor environment or in the test chamber for a given material. The term total is thus misleading as previously explained. Initially Molhave used the term TVOC to describe a specific set of 22 individual compounds, but Molhave and Nielsen warned about misinterpretation of the term. (17) The European Commission (EC 1997a) advocated the inclusion of 67 compounds in the reporting of TVOCs. Other groups report TVOCs as simply the total of what their particular analytical system permits them to measure. In addition, the detectors utilized by any particular analytical system respond differently to individual compounds. The effect of this is that, if lumped together and reported as a single value, the final result will likely have a significant error attached to it. Summing individual peaks is also error prone unless calibration of the system is performed using pure standards for each detected compound. (18)

The different guidelines/standards for detecting and measuring VOCs in indoor air are mentioned for their respective agencies in the table below.

Table 2.5. Indoor Air quality guidelines for VOCs. (Source: <https://www.epa.gov/indoor-air-quality-iaq/technical-overview-volatile-organic-compounds>)

Country; Agency	Document, Details
USA; American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE)	ANSI/ASHRAE Standard 62.1, <i>Ventilation for Acceptable Indoor Air Quality</i> (ASHRAE 2007) Appendix B, Summary of Selected Air Quality Guidelines <ul style="list-style-type: none"> • Table B-1, Comparison of Regulations and Guidelines Pertinent to Indoor Environments • Table B-2, Concentration of Interest for Selected Contaminants Setting target concentrations for TVOCs is not recommended. Setting target concentrations for specific VOCs of concern is preferred.
USA; California Department of Health Services (CDHS)	"Reducing Occupant Exposure to Volatile Organic Compounds (VOCs) from Office Building Construction Materials: Non-Binding Guidelines" (Alevantis 1996) <ul style="list-style-type: none"> • Overview of process and factors related to evaluation of VOC emissions • Appendix E, Survey of Existing Guidelines for VOCs • Section E2, Health Effects and Concentration Guidelines for Selected VOCs <ul style="list-style-type: none"> • Benzene • Formaldehyde • Methylene chloride • Styrene • Tetrachloroethylene • Toluene
USA; California's Office of Environmental Health Hazard Assessment (OEHA)	"OEHA Acute, 8-hour and Chronic Reference Exposure Levels (RELs)" (OEHA 2008) <ul style="list-style-type: none"> • Guideline values for list of ~80 chemicals with non-cancer chronic effects
Canada; Health Canada	<i>Exposure Guidelines for Residential Indoor Air Quality</i> (Health Canada 1987) <ul style="list-style-type: none"> • Guidelines/recommendations established for indoor levels of aldehydes (formaldehyde, acrolein, acetaldehyde), carbon dioxide, carbon monoxide, nitrogen dioxide, ozone, particulates, sulfur dioxide, radon, biological agents, fibrous materials (asbestos, man-made mineral fiber), polycyclic aromatic hydrocarbons, chlorinated hydrocarbons, pesticides, environmental tobacco smoke, lead, consumer products
Australia; National Health and Medical Research Council (NHMRC)	<i>Ambient Air Quality Goals and Interim National Indoor Air Quality Goals</i> (NHMRC)
Finland; Finnish Society of Indoor Air Quality and Climate (FISIAQ)	"Classification of Indoor Climate 2000: Target Values, Design Guidance and Product Requirements" (FISIAQ 2001) <ul style="list-style-type: none"> • TVOC based
Germany; Federal Environmental Agency Indoor Air Hygiene Commission (IRK) and the Working Group of the Health Ministries of the Länder (AOLG)	<i>Guidelines for Indoor Air Quality: Basic Scheme</i> (AOLG 1996) <ul style="list-style-type: none"> • IAQ guidelines set by an ad hoc working group of members of the of Germany Indoor Air Hygiene Commission (Innenraum-lufthygiene-Kommission; IRK) of the Umwelt Bundes Amt (UBA)
Germany; Committee for Health-Related Evaluation of Building Products (AgBB)	<i>Health-related Evaluation Procedure for Volatile Organic Compounds Emissions (VOC and SVOC) from Building Products</i> (AgBB 2008) <ul style="list-style-type: none"> • Procedure for calculation of lowest concentration of interest (LCI) values; LCI values currently established for 178 compounds (AgBB 2008)
Hong Kong; Indoor Air Quality Information Centre (IAQIC)	"Hong Kong Objective" (IAQIC n.d.) <ul style="list-style-type: none"> • Sets objectives for "good" and "excellent" IAQ for offices and public spaces • "Good" IAQ includes objectives for 10 specific VOCs (benzene; carbon tetrachloride; chloroform; 1,2-dichlorobenzene; 1,4-dichlorobenzene; ethylbenzene; tetrachloroethylene; toluene; trichloroethylene; and xylene [σ, m, p-isomers])
Japan; Ministry of Health, Labor and Welfare (MHLW)	<i>Guidelines of Indoor Chemicals</i> (MHLW 2002) <ul style="list-style-type: none"> • Based mainly on long-term exposure (except for formaldehyde)
International; World Health Organization (WHO)	<i>Air Quality Guidelines for Europe</i> , Second Edition (WHO 2000) <ul style="list-style-type: none"> • Lists following "organic air pollutants": acrylonitrile, benzene, butadiene, carbon disulfide, carbon monoxide, 1,2-dichloroethane, dichloromethane, formaldehyde, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), polychlorinated dibenzodioxins (PCDDs), polychlorinated dibenzofurans (PCDFs), styrene, tetrachloroethylene, toluene, trichloroethylene, vinyl chloride
International; WHO	<i>WHO Guidelines for Indoor Air Quality</i> (see WHO [2006] for a report on a working group meeting) [Currently under development]

2.4 Sources of VOCs

VOCs are ubiquitous and are produced both anthropogenically as well as naturally. VOC can be found not only as vapors in the air but also as soil gas and dissolved in ground water.

Vapor intrusion occurs when volatile organic compounds from contaminated soil or groundwater migrate upwards toward the ground surface and into overlying buildings or surfaces through gaps and cracks in foundation slabs or basement walls. The route VOCs take from a subsurface source to the air is referred to as the vapor intrusion pathway. (10)

VOCs are emitted by a variety of common household products as indicated in the table below.

Table 2.6. Examples of secondary VOC emissions from common building products and equipment. (19)

Building Product	Secondary VOCs	Condition*	Literature
Carpet—wool based	Aldehydes, formaldehyde, aldehydes, acids, benzothiazol	Ozone, heat	Knudsen et al. (1999); Wolkoff (1998); Brzezinski et al. (1996); Weschler et al. (1992); Sollinger et al. (1993)
Carpet cushion	Acetic acid	Water, nitrogen	Schaeffer et al. (1996)
Cork	Acetic acid, furfural	Heat	Horn et al. (1998)
Duct lines	C _{6,8,10} aldehydes, fatty acids	Ozone	Morrison et al. (1998)
Furniture coating	Aldehydes, acrylates, isocyanates, styrene, terpenes		Salthammer et al. (1999b)
Linoleum	Aldehydes, unsaturated aldehydes	Water	Jensen et al. (1996); Wolkoff et al. (1995)
Alkyd paint—“natural” paint	C ₃ and C ₅₋₈ aldehydes, fatty acids, terpenes		Rothweiler et al. (1993); Ullrich et al. (1992); Volland and Zölter (1996); Zellweger et al. (1997)
Office equipment/humans	Formaldehyde, aldehydes	Ozone	Leovic et al. (1996); Wolkoff et al. (1992)
Paint—acrylic, latex	Aldehydes, formaldehyde, acet-aldehyde, formic acid	Ozone	Chang and Guo (1998); Knudsen et al. (1999); Young (1992); Reiss et al. (1995a, 1995b)
Primer—water based	Hexanal		Zellweger et al. (1997)
Polyvinyl chloride (PVC)	2-ethylhexanol	Water	Knudsen et al. (1999); Wolkoff (1998); Van der Wal et al. (1997)
Thermal insulation	Aldehydes	Moisture	Van der Wal et al. (1989)

*Clean air is used, if not otherwise stated.

The exact quantity and name of chemical compounds which is present even within the indoor air is hard to estimate in a laboratory setting since there are multiple real-world factors and reactions taking place that influence the quality of the indoor air. There are plenty of opportunities for the VOCs to form and transform into more complex compounds as indicated in the table below.

Table 2.7 Possible reaction products in indoor air with potential emission sources and reactants (20)

Reactants	Products	Possible Sources
x-Pinene	Pinene oxide, pinaldehyde	Wood, wood-based products
Limonene	Limonene oxide, carvone, formaldehyde	Wood, coating systems
Oleic acid	Heptanal, octanal, nonanal, decanal, 2-decanal	Linoleum, eco-lacquers, nitrocellulose
Linolenic acid	2-Pentenal, 2-hexenal, 3-hexenal, 2-heptenal, 2,4-heptadienal, 1-penten-3-one	Lacquers, alkyl resins
Linoleic acid	Hexanal, heptanal, 2-heptenal, octanal, 2-octenal, 2-nonenal, 2-decanal, 2,4-nonadienal, 2,4-decadienal	
Hemicelluloses	Furfural, acetic acid	Cork
5-Methyl-N-Phenyl-2-Pyrrolidone (PHMP)	Benzaldehyde, acetone, benzil	Ultraviolet-cured coatings
1-hydroxycyclohexyl phenyl ketone (HCPK)	Benzaldehyde, cyclohexanone, benzil	Ultraviolet-cured coatings
2-Ethyl-hexyl acetate	Acetic acid, 2-ethyl-a-hexanol	Solvent
Zn-2-ethylhexanoate	2-Ethyl-a-hexanoic acid	Stabilizers
n-Butylacrylate	n-Butanol	Acrylate coatings
Di(2-ethylhexyl)-phthalate (DEHP)	2-Ethyl-q-hexanol	Plasticiser
Di-n-butyl-phthalate (DBP)	n-Butanol	Plasticiser
Di-iso-butyl-phthalate (DIBP)	2-Butanol	Plasticiser
Tris (2-chloroisopropyl) phosphate (TCPP)	1-Chloro-2-propanol, 2-chloro-1-propanol	Flame retardant
Tris(2,3-dichloropropyl) phosphate(TDCPP)	1,2-Dichloropropane, 1,2-dichloropropanol	Flame retardant
Tris (2-chloroethyl) phosphite (TCEP)	2-Chloro-ethanol	Flame retardant
Tert-Butyl peroxyphthalate (TBPP)	1-Bromo-2-propanol, 2-bromo-1-propanol	Flame retardant
Tris(2,4-di-t-butylphenyl)phosphite (TDBPP)	2,3-Dibromo-1-propanol	Flame retardant
Styrene + cis-1,3-butadiene	4-Phenylcyclohexene (4-PC)	Styrene-butadiene rubber
Cis-1,3-butadiene + trans-1,3-butadiene	4-Vinyl-cyclohexene (4-VCH)	Styrene-butadiene rubber
2-Chloro-1,3-butadiene	1-Chloro-4-(1-chlorovinyl)-cyclohexene 1-Chloro-5-(1-chlorovinyl)-cyclohexene	Rubber
Zn-diethylthiocarbamate	Carbon disulfide, diethylamine	Vulcanization accelerator
Azodicarbonamide	Semicarbazide	Foaming agent
Adipic acid + 1,4-butanediol	1,6-Dioxo-cyclododecane-7,12-dione	Adhesive
Dimethylaminoethanol + formic acid	Dimethylformamide	"Green" paint
1-Tryptophane	o-Aminoacetophenone	Casein products
2,3,4,6 Tetrachlorophenol	2,3,4,6 Tetrachloroisole	Application of pentachlorophenol (PCP)
T4MDD	MIBK, 3,5-dimethyl-1-hexyne-3-ol	Water-based paint
AIRN	Tetramethyl succinonitrile	Flexible polyurethane foam

For the purpose of this analysis we are concerned with VOCs released from oil and gas sites such as benzene, toluene, ethyl benzene, xylene, dichloroethane, trichloroethane, trichloroethene, tetrachloroethane, styrene, chlorobenzene, carbon tetrachloride and other such VOCs.

2.5 Thesis objectives

There were multiple objectives for this project. The first objective was to see if the abandoned fracking site was still releasing VOCs from the soil into the atmosphere. We used low-cost, real-time sensors to ascertain if the site was releasing VOCs. Once we were able to ascertain without a doubt that the site was releasing VOCs, we used a combination of real-time sensors and laboratory tests to determine the quantity and composition of the VOCs being released. We placed multiple sensors around the site and used spatial and temporal analysis techniques to determine the movement of the VOC and the contaminant plume underground. We also measured other environmental factors such as ground water quality and weather conditions such as temperature, humidity, winds, atmospheric pressure, etc. and used regression analysis to see how these environmental factors affect the release of VOCs. We used multiple analysis techniques to see which of these techniques were suitable and try to determine a methodology to use multiple low-cost sensors instead of a single high cost sensor to get more valuable data. The method of analysis and analysis techniques used are mentioned and explained in the following sections.

CHAPTER 3: METHODS

3.1 Description of the assigned site

There are currently tens of thousands of wells where massive hydraulic fracturing is performed routinely producing natural gas and condensate in the Denver Basin of central Colorado. This field is known as the Wattenberg gas field and it covers more than 2,000 square miles between the cities of Denver and Greeley as indicated in yellow boundary in the map below. This field was discovered in the 1970s and has had fracking being conducted regularly since then. The Wattenberg gas field lies in a 8hr Ozone Nonattainment Area. A Nonattainment Area is defined as any area that does not meet the national primary or secondary ambient air quality standard which in this case is ozone levels.

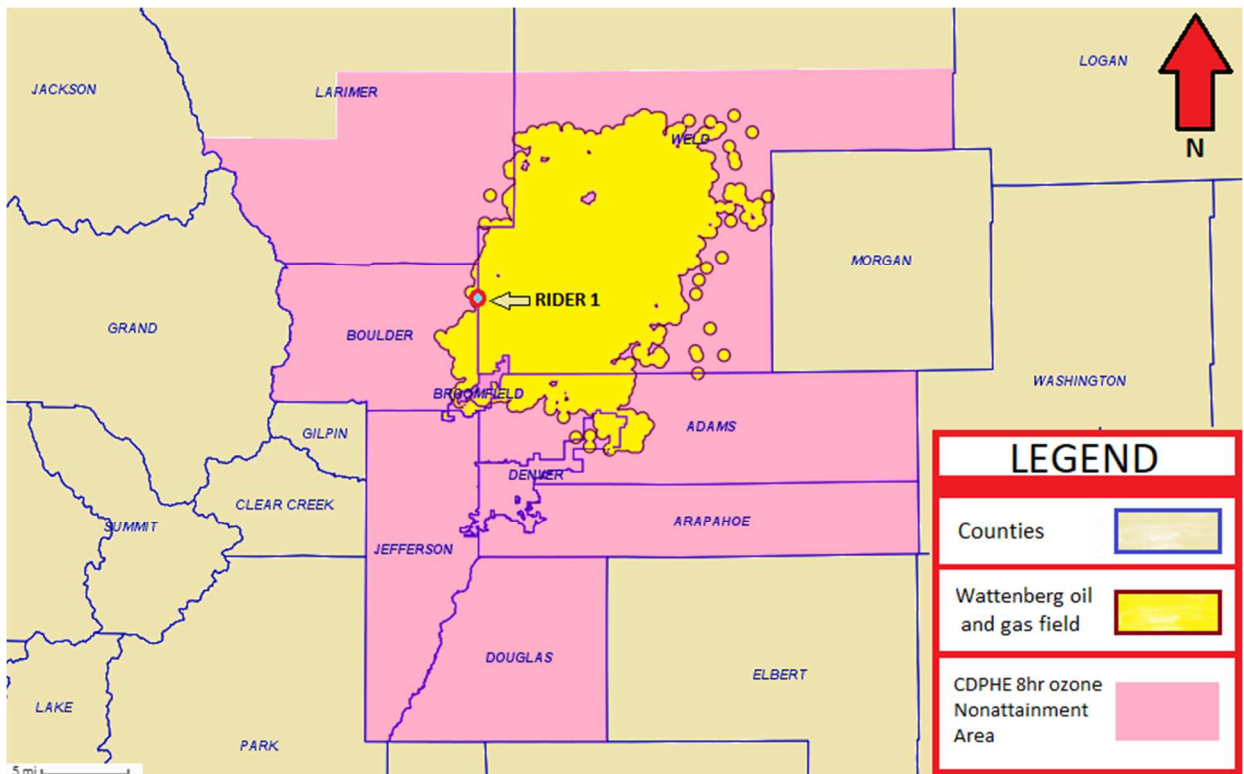


Figure 3.1. Map of the study area (Source https://cogccmap.state.co.us/cogcc_gis_online/)

We were assigned a plugged and abandoned well named ‘Rider-1’ (as indicated in Figure 3.1) by the city of Longmont to investigate if it was still releasing VOCs into the atmosphere and ground water. The site lies within the west edge of the Wattenberg gas field and was operational between the August 1998 to October 2016. The well is located in Boulder county at latitude 40.1791 and longitude -105.05879 and was at least 10,000 feet deep. The soil type of the site is shown in Figure 3.2 and the chemical properties of the soil are indicated in table 3.1.



Figure 3.2 Soil type of the area of study (Source: <https://websoilsurvey.nrcs.usda.gov/>)

There are 2 types of soils found in the region, namely Colby silty clay loam (CsB) to the north and Weld loam (WIB) to the south. After a depth of a foot the soil is completely alkaline with a moderately high cation-exchange capacity.

Table 3.1 Chemical properties of the soil in the area of study

(Source: <https://websoilsurvey.nrcs.usda.gov/app/>)

Soil name	Depth	Cation-exchange capacity	Soil reaction	Calcium carbonate	Salinity	Sodium adsorption ratio
	In	meq/100g	pH	%	mmhos/cm	
CsB— Colby silty clay loam, wet	0-12	10-20	6.6-7.3	5-10	0	0
	12-40	5.0-20	7.4-8.4	5-10	0	0
	40-60	10-25	7.4-9.0	5-10	0.0-2.0	0
WIB— Weld loam	0-8	18-21	6.6-7.8	0-2	0.1-1.0	0
	8-12	29-37	6.6-7.8	0-2	0.1-1.0	0
	12-15	27-34	7.4-7.8	0-2	0.1-1.0	0
	15-28	18-21	7.9-8.4	2-10	0.1-2.0	0-2
	28-60	21-26	7.9-9.0	5-14	0.1-2.0	0-5

The City of Longmont suspected that there was a contaminated plume in the soil but were unsure of the location, composition and extent of the plume and had asked us to investigate it in more detail. To find out the location, extent and quantity of VOCs being released into the atmosphere and to trace the movement of the plume in the ground, we placed 3 soil gas VOC sensors and 3 ground water quality monitoring sensors. We used laboratory analysis using gas chromatograph to figure out the composition of the VOCs being released.

A map of the site is shown below, the larger circles indicate the location of the soil gas sensors and the smaller blue dots indicate the locations of the ground water monitoring stations. The grey box near the top left corner with the white pipelines, used to be the Rider 1 site.

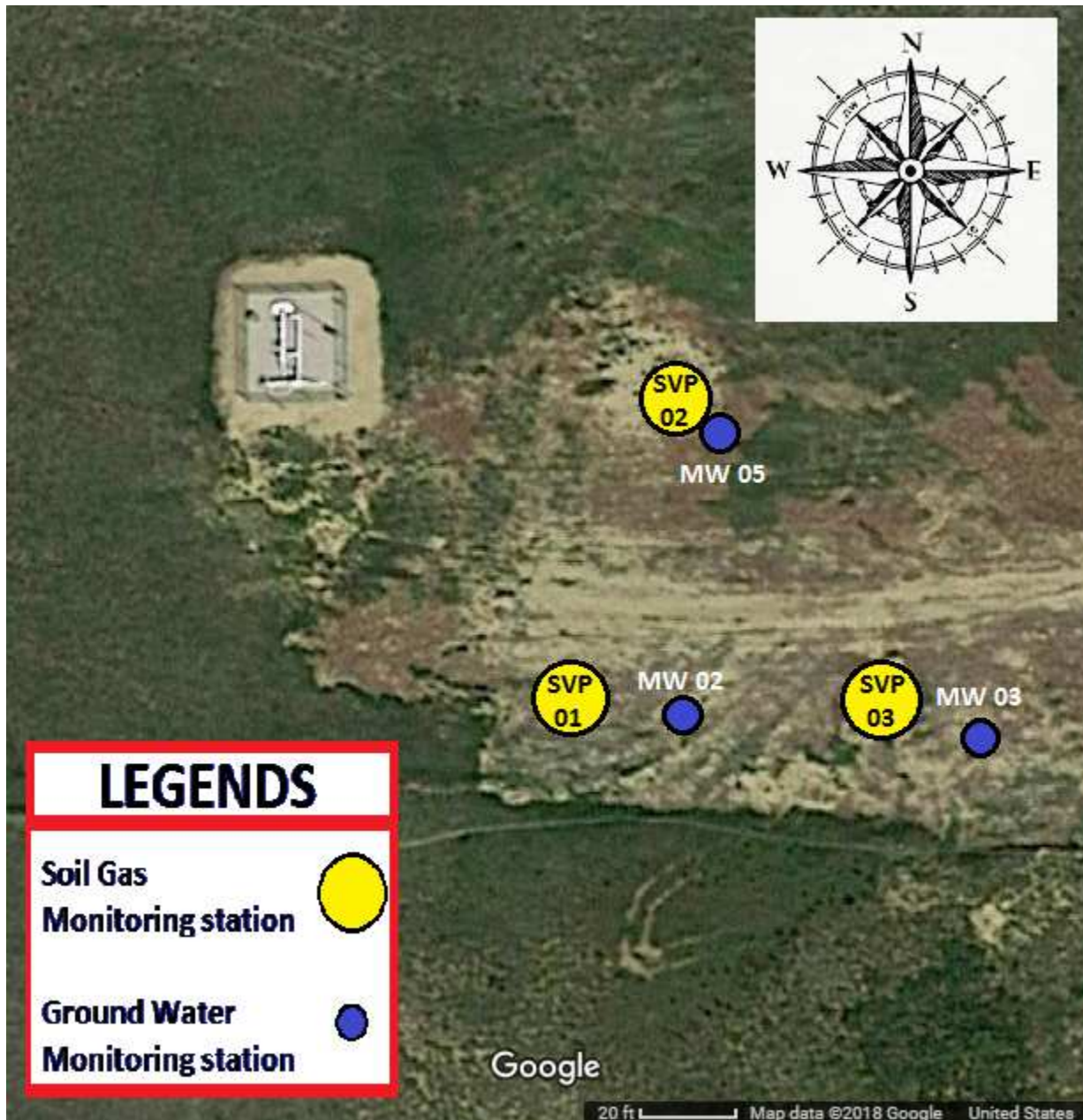


Figure 3.3. Map of the monitoring stations

We chose these sites based on the land allotted to us by the city of Longmont and availability previously dug areas for underground pipes for soil gas monitoring and deployment of groundwater sensors.

3.2 Description of soil gas VOC sensor used

We chose to use NDIR sensors for our analysis since infrared gas detection is a well-developed measurement technology. Infrared gas analyzers had a reputation for being complicated, cumbersome, and expensive. However, recent technical advancements, including the availability of powerful amplifiers and associated electronic components, have opened a new frontier for infrared gas analysis. The gases to be detected are often corrosive and reactive with most sensor types, the sensor itself is directly exposed to the gas, often causing the sensor to drift or die prematurely. The main advantage of IR instruments is that the detector does not directly interact with the gas (or gases) to be detected.

We used 3 NDIR sensors from Metrologics LLC, capable of pumping soil gas, sensing it and transmitting the data and placed them around the abandoned fracking site. A schematic of the NDIR sensor used is shown in figure 3.4. These sensors pumped soil gas from 6-10 feet under the ground using an in-built pump and air pipes attached to the sensor and extending all the way underground through larger pipelines. The air was pumped at 1 hour intervals into the sensor and tested in the NDIR chamber and the results of the tests were sent wirelessly to a cloud based server.

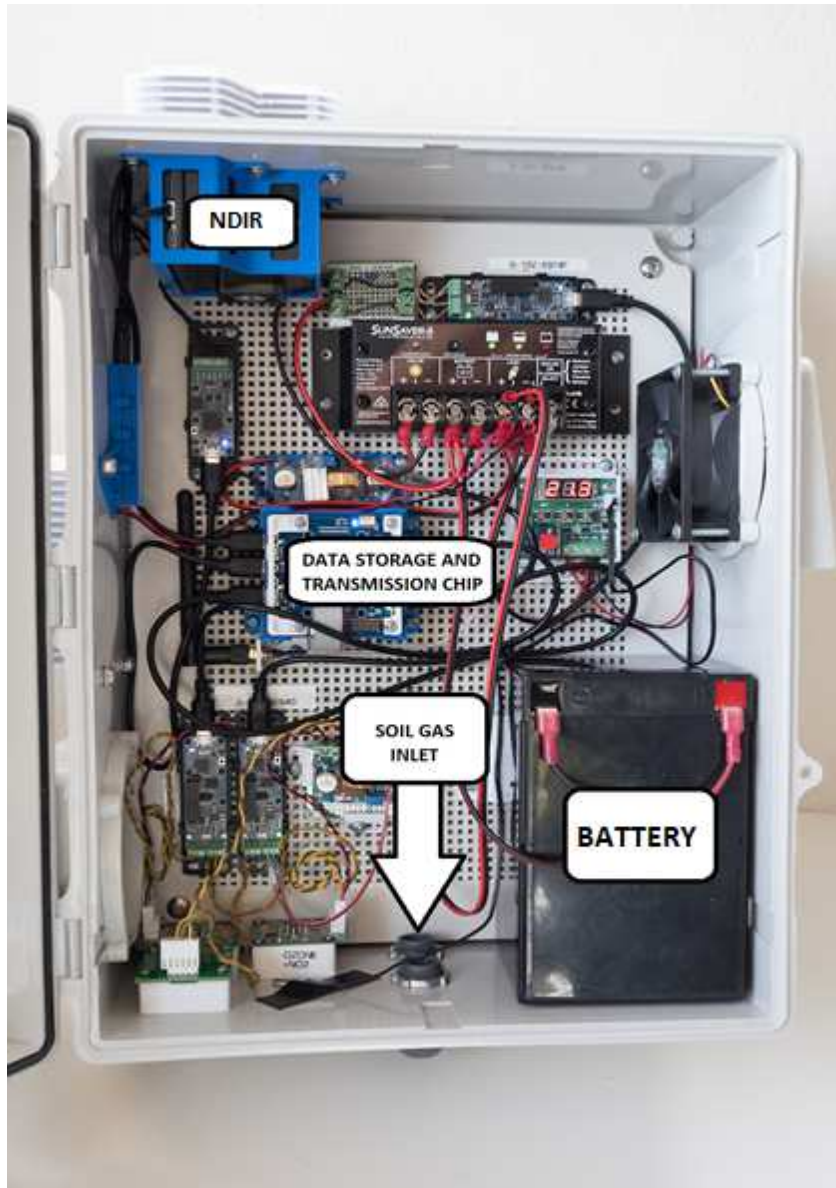


Figure 3.4. Schematic of NDIR soil gas sensor

The device consists of a battery with enough capacity to be active for 3 months and support all the activities. Most of the battery power was used to power the air pumps followed by the actual NDIR sensor and hence we chose 1-hour intervals for pumping to conserve the battery life of the sensor. The device also houses the NDIR sensing chamber and a data storage and transmission chips. The NDIR sensors were capable of detecting carbon and hydrogen bonds but

could not detect the exact type and quantity of VOC being released. We used the data received from the sensors to determine the change in quantity of VOCs being released over time and space and the laboratory results to determine the exact type and quantity of VOCs being released.

Following is a list of common gases that are detected by this detector:

1. Alkanes or saturated hydrocarbons such as methane, ethane, propane, butane, pentane, hexane, and heptane, etc.
2. Cycloalkanes such as cyclopropane, cyclohexane, methyl cyclohexane, etc.
3. Alkenes or unsaturated hydrocarbons such as ethylene, propylene, butene, pentene, hexene, octene, etc. Acetylene has absorption at 3.1 microns which is not detectable.
4. Cycloalkenes such as cyclohexene and pinene.
5. Aromatics such as benzene, toluene, and xylene.
6. Alcohols such as methanol, ethanol, propanol, and allyl alcohol.
7. Amines such as dimethyl amine, trimethyl amine, butanamine, cyclopropanamine, and pyridines.
8. Ethers such as dimethyl ether, ethyl ether, n-propyl ether, methylvinyl ether, vinyl ether, ethylene oxide, tetrahydrofuran, furan, and 1,4-dioxane.
9. Ketones such as acetone, methyl ethyl ketone, pentanone, methyl isobutyl ketone and heptanone.

10. Aldehydes that have a central wavelength mostly at the 3.55 micron region and generally have a weak detection signal at 3.4 microns.

Though the results from these sensors do not accurately represent the exact quantity or speciate the VOCs being tested, instead they provide us with the change in concentration of VOCs temporally and spatially. We used laboratory analysis to speciate the VOCs and get the exact composition and interpolated these values with the real-time analysis. Samples of the soil gas in were collected in Tedlar bags before and after setting up the NDIR sensors and tested these samples in a laboratory using a gas chromatograph in accordance with EPA's method 18.

3.3 Description of ground water quality monitoring sensor

There was no information available on the groundwater quality of the area, but the groundwater table was at a height of about 11 feet below the ground from historical data obtained from USGS.

We setup 3 ground water monitoring sensors from In-Situ Inc. and placed them around the abandoned fracking site at depths of about 10-12 feet. The sensors were capable of detecting groundwater pH, conductivity, ORP, DO and temperature. The specifications of the sensor and its measurement capabilities are shown in table 3.2. These sensors were also capable of transmitting the data to the cloud-based server. A picture of the sensor used is shown figure 3.5.



Figure 3.5. In-Situ AQUATROLL 400 sensor

Table 3.2. Specifications for Aquatroll 400

STANDARD SENSORS	ACCURACY	RANGE	RESOLUTION	SENSOR TYPE	RESPONSE TIME	UNITS OF MEASURE	METHODOLOGY
LEVEL, DEPTH, PRESSURE	Typical $\pm 0.1\%$ FS @ 15° C; $\pm 0.3\%$ FS max. from 0 to 50° C	76 m (250 ft); absolute (non-vented)	$\pm 0.01\%$ FS or better	Fixed	Instantaneous in thermal equilibrium	Pressure: psi, kPa, bar, mbar, mmHg Level: mm, cm, m, in., ft	Piezoresistive; ceramic
CONDUCTIVITY	Typical $\pm 0.5\% + 1$ $\mu\text{S/cm}$; $\pm 1\%$ max.	5 to 100,000 $\mu\text{S/cm}$	0.1 $\mu\text{S/cm}$	Fixed	Instantaneous in thermal equilibrium	Actual conductivity ($\mu\text{S/cm}$, mS/cm) Specific conductivity ($\mu\text{S/cm}$, mS/cm) Salinity (PSU) Total dissolved solids (ppt, ppm) Resistivity (Ohms-cm) Density (g/cm ³)	Std. Methods 2510 EPA 120.1
DISSOLVED OXYGEN OPTICAL RDO-X CAP	± 0.1 mg/L $\pm 2\%$ of reading	0 to 20 mg/L 20 to 60 mg/L Full operating range: 0 to 60 mg/L	0.01 mg/L	Fixed with replaceable RDO Sensor Cap. Supports Classic, Fast, and RDO-X caps. Ships with RDO-X cap.	T90: <45 sec. T95: <60 sec.	mg/L, % saturation, ppm, ppO ₂	EPA-approved In-Situ Methods 1002-8-2009 1003-8-2009 1004-8-2009
Interferences: Alcohols >5%; hydrogen peroxide >3%; sodium hypochlorite (commercial bleach) >3%; gaseous sulfur dioxide; gaseous chlorine. Organic solvents and certain petroleum-based hydrocarbons may swell the sensing element and destroy it. Examples include, but are not limited to, acetone, chloroform, methylene chloride, and BTEX compounds.							
ORP	± 5.0 mV	± 1400 mV	0.1 mV	Replaceable pH/ORP combo sensor	<15 sec.	mV	Std. Methods 2580
pH	± 0.1 pH unit	0 to 14 pH units	0.01 pH unit	Replaceable pH/ORP combo sensor	<15 sec., pH 7 to pH 4	pH units, mV	Std. Methods 4500-H+ EPA 150.2
TEMPERATURE*	± 0.1 ° C	-5 to 50° C (23 to 122° F)	0.01° C or better	Fixed	T90: <120 sec.	Celsius, Fahrenheit	EPA 170.1

3.4 Data Collection

The area of study consisted of 3 ground-water monitoring sites (namely MW 02, MW 03, MW 05) and 3 soil-vapor gas monitoring sites (namely SVP 01, SVP 02, SVP 03). We conducted hourly measurements of VOCs using NDIR sensors pumped from a depth of about 6-7 feet. Ground water quality parameters such as pH, conductivity, ORP and temperature every quarter hour from a depth of about 10-11 feet for the period of 23rd of October 2017 till the 15th of May 2018. Along with that we also performed base line sampling before the sensors were placed in the monitoring wells and performed in-situ tests on the gas captured from these wells after the real time sensors were removed from the wells.

We were getting very high values in SVP01 so we interchanged the NDIR sensors at that site with the ones at SVP02 and SVP03 at different instances to check for sensor-drift and got the same high readings at SVP01. We calibrated the water quality sensors but the NDIR sensors came precalibrated and were not calibrated during the duration of the testing period.

After enough data was collected we analyzed the data using various analysis techniques as detailed in the following section. I wanted to see the change in concentration over time and hence I used trend analysis and time series analysis on the data. I also wanted to correlate the soil gas data with the other environmental factors such as ground water and climate data hence I used correlation and regression analysis techniques on the data.

CHAPTER 4: DATA ANALYSIS

Any observed data representing a physical phenomenon can be broadly classified as being either deterministic or non-deterministic. Deterministic data are those that can be described by an explicit mathematical relationship. There are many physical phenomena in practice which produce data that can be represented with reasonable accuracy by explicit mathematical relationships. However, there are many other physical phenomena which produce data that are not deterministic. For example, the acoustic pressure generated by air rushing through a pipe, or the electrical output of a noise generator, the height of waves in a confused sea represent data which cannot be described by explicit mathematical relationships. There is no way to predict an exact value at a future instant of time. These data are random in character and must be described in terms of probability statements and statistical averages rather than explicit equations. The same statistics can be used to describe VOC data released into the atmosphere from the soil gas.

4.1 Summary of Raw Data:

4.1.1 Soil Gas Data

The soil gas samples were pumped from a depth of about 6 feet and measured hourly on site in an infra-red sensor which was set to detect methane and other chain hydrocarbon. Other VOCs are detected by the NDIR sensor as well and hence the results are represented as a surrogate for total percentage of VOC. The raw total VOC detected in the soil gas in each site is shown in figures 4.1 - 4.3.

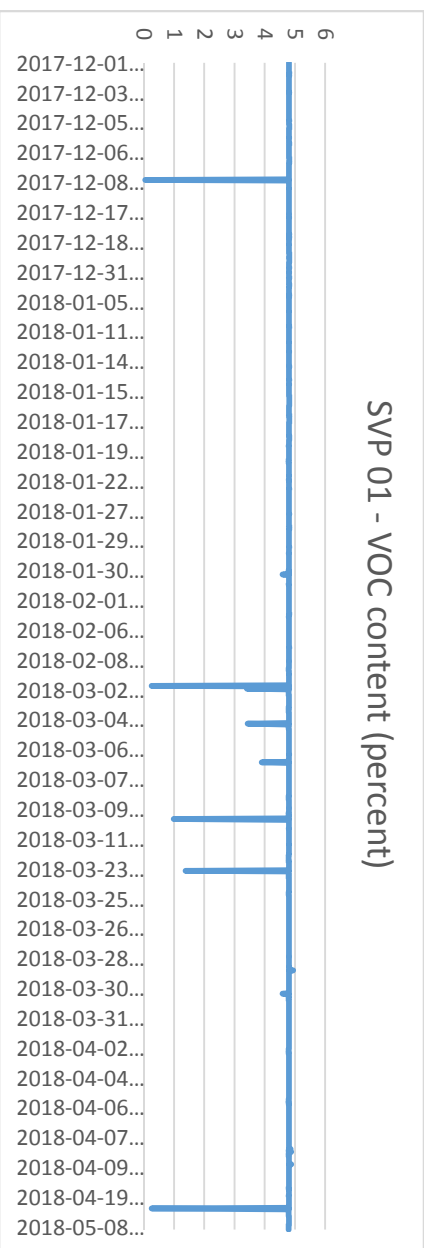


Figure 4.1 Linear graph of VOC concentration in soil gas at SVP 01 vs time

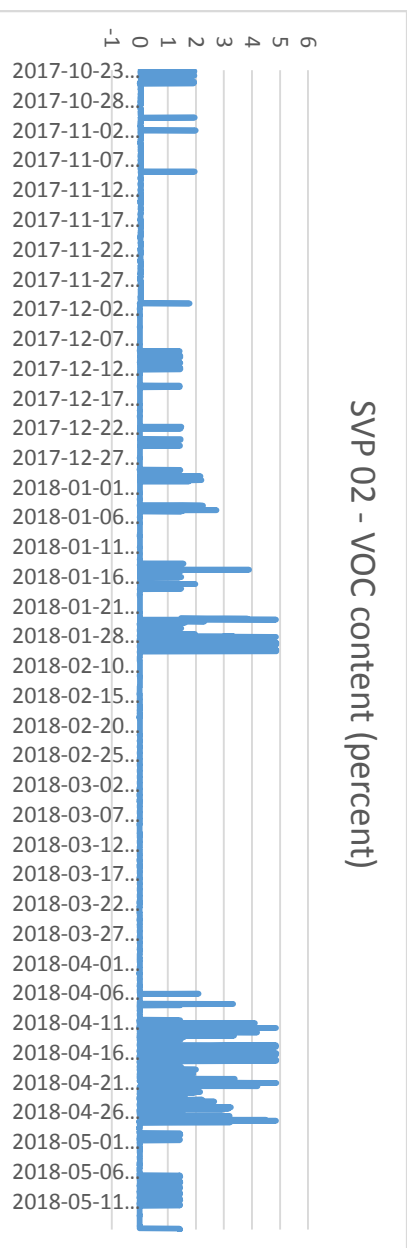


Figure 4.2. Linear graph of VOC concentration in soil gas at SVP 02

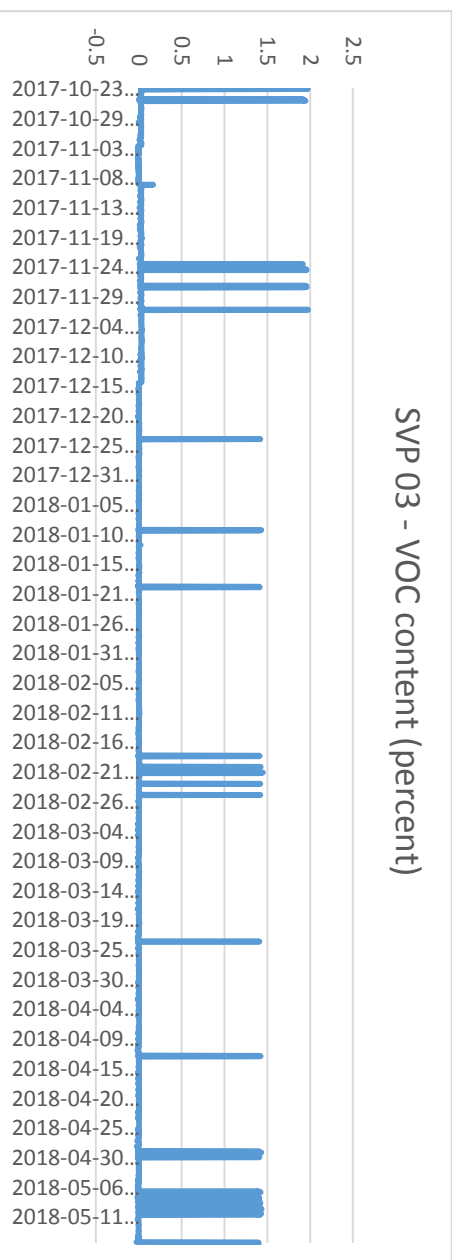


Figure 4.3. Linear graph of VOC concentration in soil gas at SVP 03

4.1.1.1 SVP 01

The soil gas sensor at SVP 01 had a high and pretty constant VOC content of about 4.78% as can be seen in the raw data graph and interval plot below. There were a few drops in the magnitude of values, but it is predominantly high and this monitoring station was most likely the closest survey point to the source of the pollutant. The graph below shows quantity of TVOCs detected (In our case single chain hydrocarbons) vs time in hours. The extremely high quantity of TVOCs detected is because the NDIR sensors used for this analysis is calibrated for methane, other hydrocarbons with higher number of single chain hydrocarbon will provide higher values which cause the sensors to provide a false reading. Hence these readings on its own do not provide accurate quantitative measurements of the VOCs present but since we take near real-time readings and with multiple sensors, we can get valuable insight from the comparative values between the sensors. Also, these results coupled with the in-situ tests results can help predict the more exact quantitative values of the contaminant over time.

An interval plot shows a 95% confidence interval for the mean of each group. An interval plot works best when the sample size is at least 10 for each group. Usually, the larger the sample size, the smaller and more precise the confidence interval. The interval plot below shows that the mean of VOC values at SVP 01 is 4.78% with a statistical 95% confidence the mean of being between 4.764 to 4.798 %.

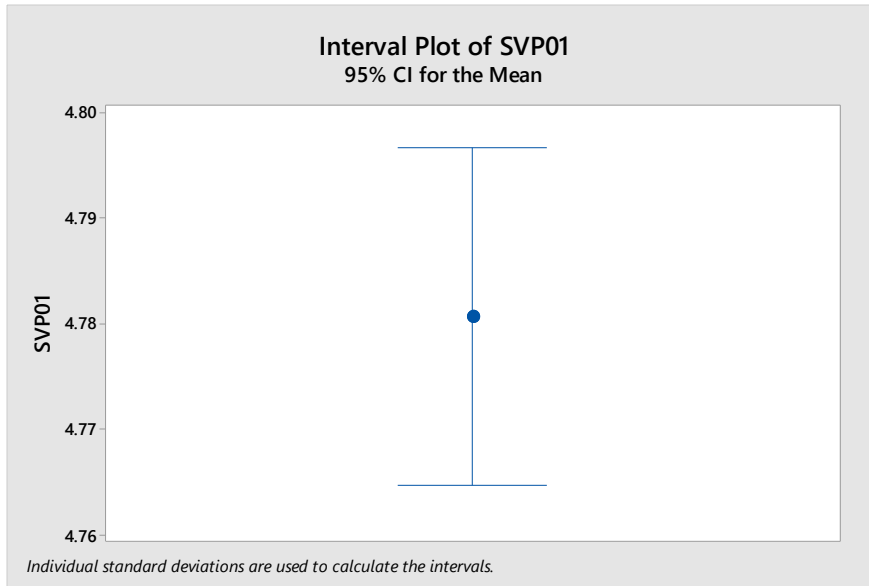


Figure 4.2 Interval plot of VOC concentration at SVP 01

4.1.1.2 SVP 02

The VOC sensor at SVP 02 has a lot of spikes ranging from 1.5 – 5 % but the background or most common values is around 0.075 % as can be seen from the interval plot and raw data graph below. There is a quiet period from the mid of January till the middle of April when no spikes are observed, this is most likely due to sub surface water freezing due to the weather conditions in Colorado during that period. We start seeing the spikes again when the weather and ground water temperature increases.

From the interval plot below, we can say with 95% confidence that the mean concentration of VOC in the soil gas at SVP 02 is between 0.059 to 0.084 %.

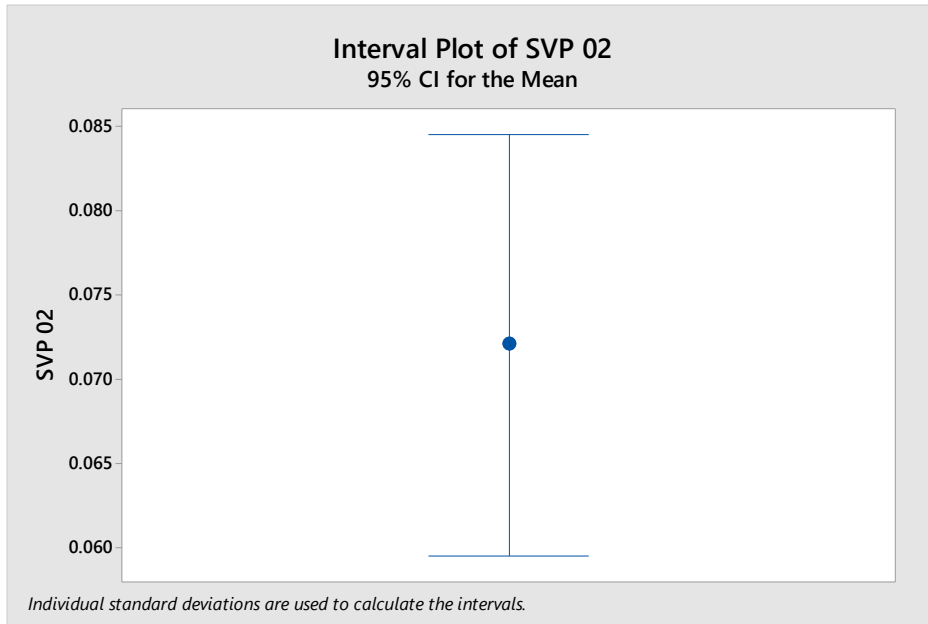


Figure 4.4 Interval plot for VOC concentration at SVP 02

4.1.1.3 SVP 03

SVP 03 follows a similar pattern to SVP 02 but with a few key differences. The spikes do not go as high as SVP 02 but instead max out at about 1.5 – 2 %. There are also few spikes during the SVP 02's quiet period of mid Jan to mid Apr. These spikes occur usually on warmer days. I hypothesize that the reason for the spikes in SVP 03 during the quiet period is due to the soil type. SVP 03 has a sandy layer of soil below while SVP 02 has clay soil covering and below it. Hence SVP 03 spikes on hotter weather while SVP 02 spike when the ground water and resulting sub surface heated up.

From the box plot below, we can say with 95% confidence that the mean of the VOC concentration in soil gas at SVP 03 is between 0.031 to 0.043 %.

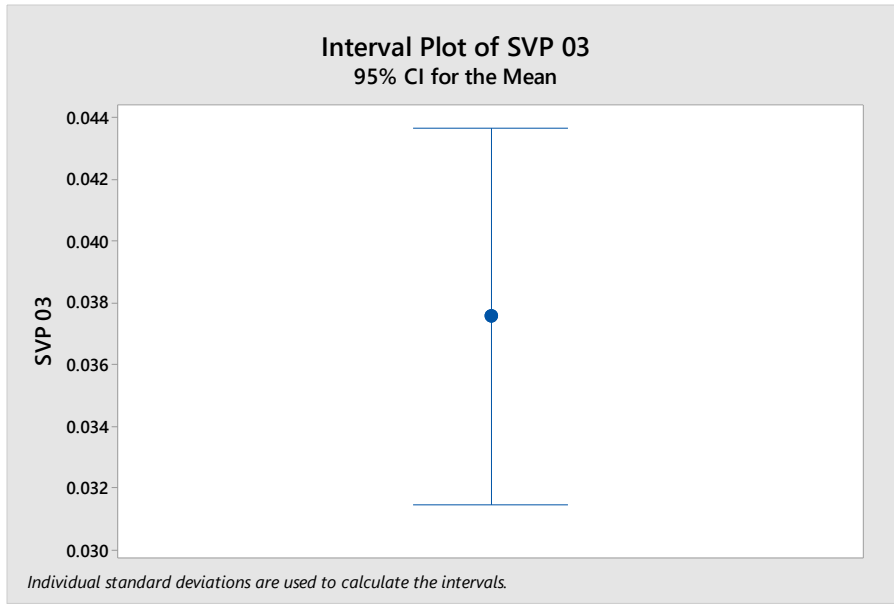


Figure 4.6 Interval plot of VOC concentration at SVP 03

4.1.2 Ground water Data

The Ground water data was collected from 3 sites at depths of about 10 – 12 feet. There were a lot of errors and issues with the ORP and DO data hence I considered only ground water temperature, pH and conductivity data for the purpose of my analysis. We used the company In-Situ Inc.'s water quality sondes Aqua Troll 400 to detect, log and transmit our data.

4.1.2.1 MW02

This was the closest ground water monitoring site to SVP 01. As can be seen from the graphs below the temperature constantly drops till the middle of April after which it starts to rise again. The pH constantly keeps increasing during the same period indicating that the pollutant level may be increasing in the ground water.

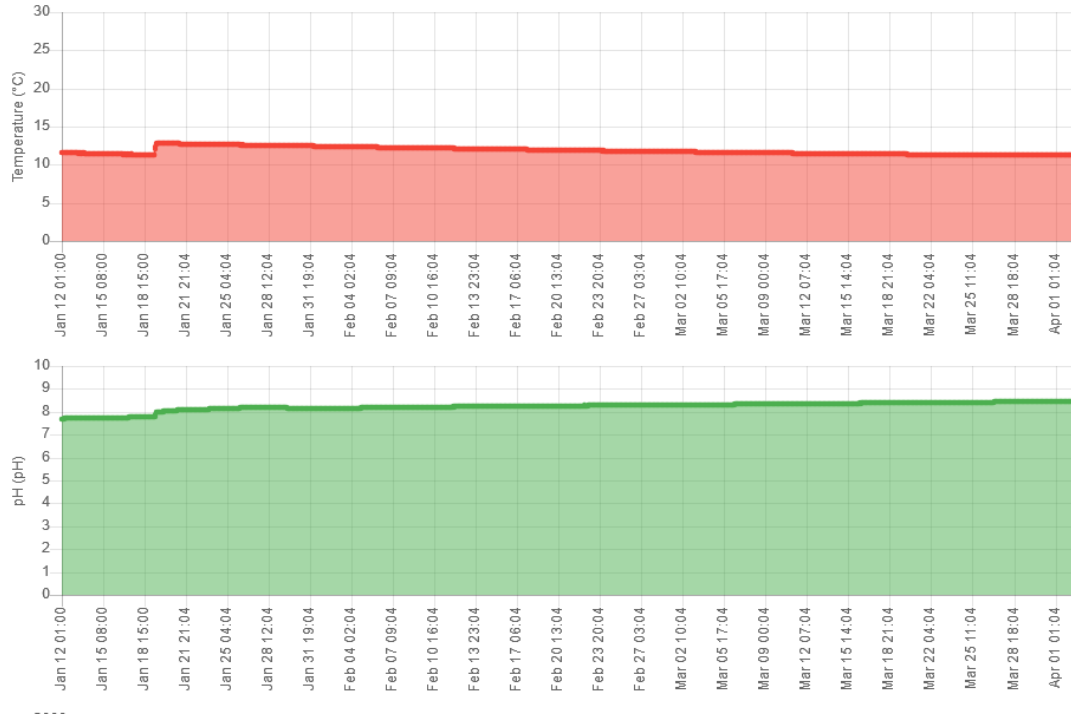


Figure 4.7. Graph of ground water temperature and pH vs time at MW 02

4.1.2.2 MW03

MW03 is the closest ground water monitoring site to SVP 03. Similar to MW02, we can see that the ground water temperature constantly keeps dropping till the middle of April after which it starts to rise again, and the pH is also constantly rising for the period of the study. The conductivity varies during the period and both the conductivity and pH indicate a rise in the pollutant levels in the ground water.

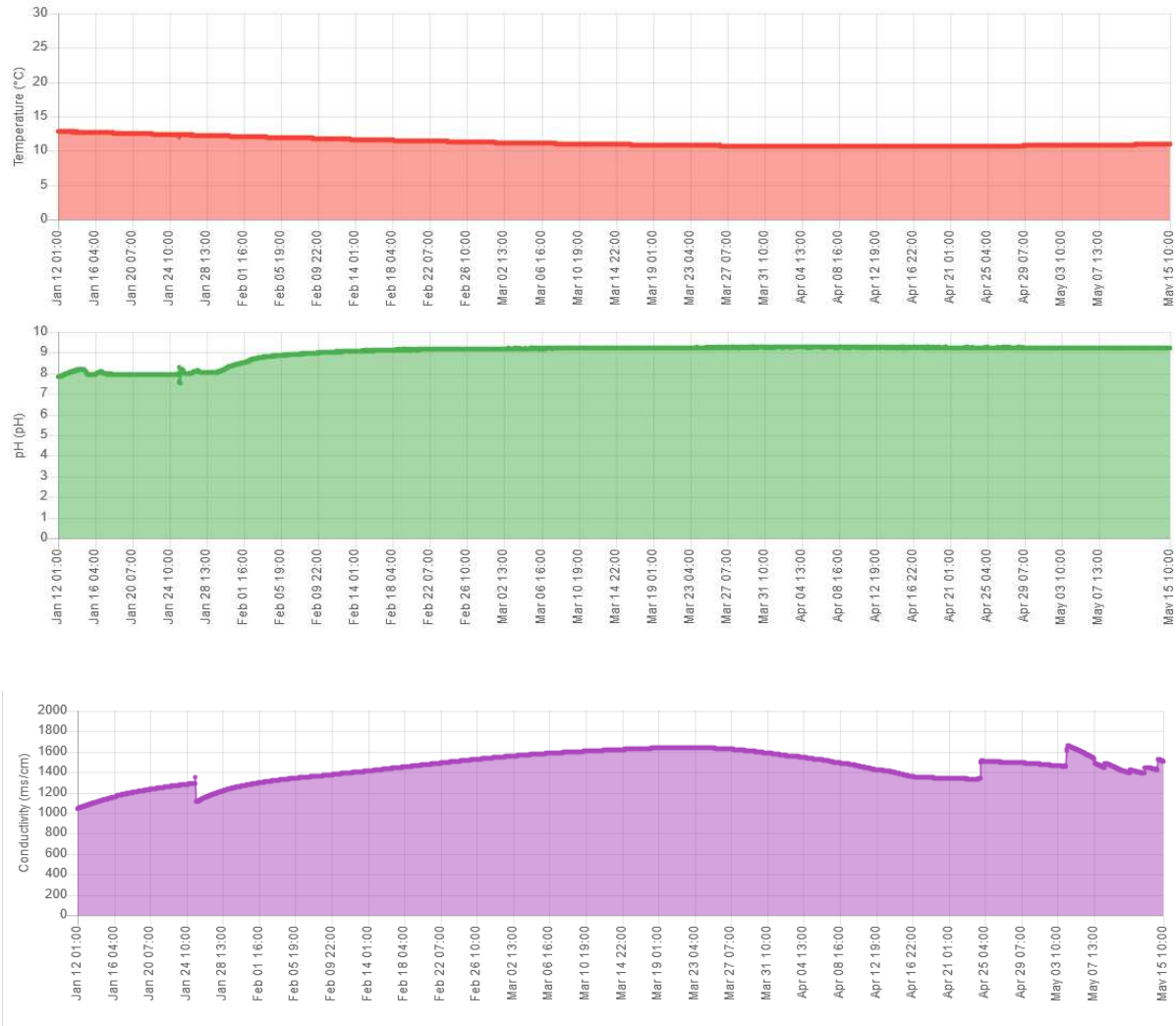


Figure 4.8. Graph of ground water temperature, pH and conductivity at MW 03

4.1.2.3 MW05

MW05 is the closest ground water monitoring station to SVP 02 and follows the similar falling and rising temperature and rising pH levels to that of the other two monitoring stations as shown in the graph below.

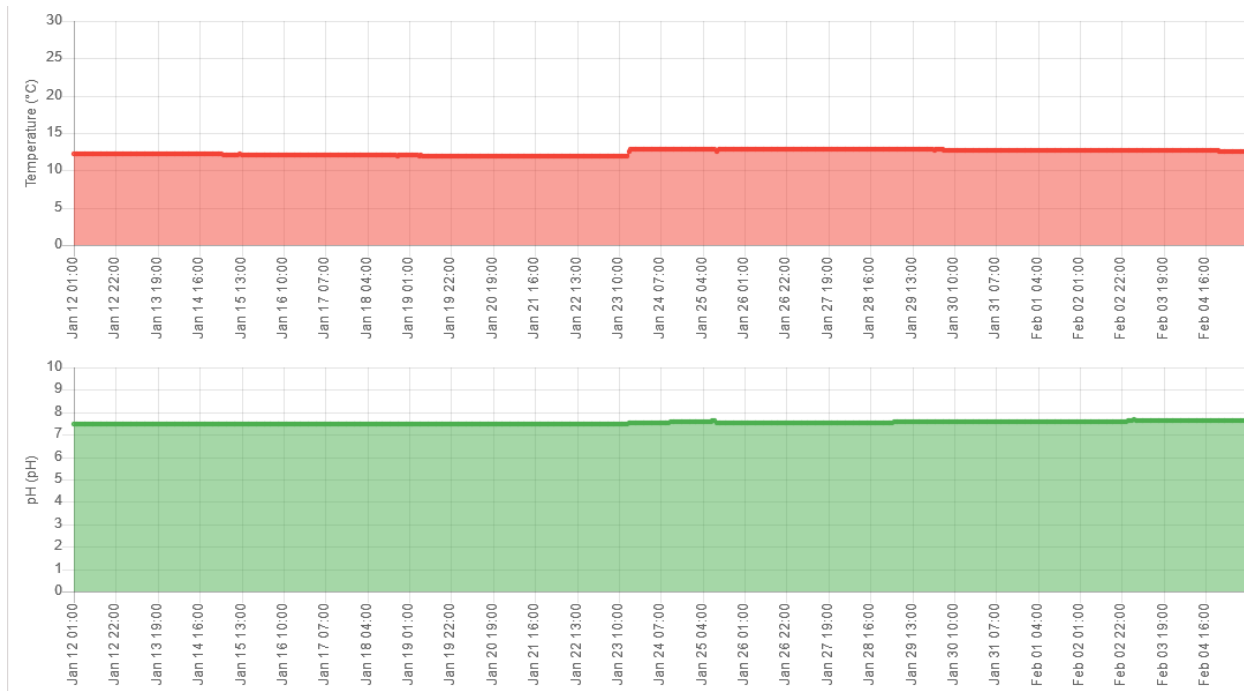


Figure 4.9. Graph of ground water temperature and pH at MW 05

4.2 Data Validation

The Raw data undergoes a process of data validation before any analysis is performed on it. During this process all blank and obviously erroneous values are removed for example pH values below 3 and negative values from the soil gas sensors. All the quarter hour ground water data is converted to hourly data by taking averages. This data along with the weather data is synced with hourly soil gas data and verified if all the points and the number of data match for all the columns.

4.3 Trend Analysis

Trend analysis is a statistical procedure performed to evaluate hypothesized linear and nonlinear relationships between two quantitative variables. Typically, it is implemented either as an analysis of variance for quantitative variables or as a regression analysis. It is commonly used in situations when data have been collected over time or at different levels of a variable; especially when a single independent variable, or factor, has been manipulated to observe its effects on a dependent variable, or response variable (such as in experimental studies). The means of a dependent variable are observed across conditions, levels, or points of the manipulated independent variable to statistically determine the form, shape, or trend of such relationship. Data over time which allows the use of statistical data to demonstrate deviations and conformance to established limits.

4.3.1 Linear trend analysis

Linear trend forecasting is used to impose a line of best fit to time series historical data. It is a simplistic forecasting technique that can be used to predict variation and is an example of a time series forecasting model. Linear trends show steady, straight-line increases or decreases where the trend-line can go up or down and the angle may be steep or shallow. Linear trend estimation expresses data as a linear function of time. There are various ways to do so, but the most usual choice is a least-squares fit. This method minimizes the sum of the squared errors in the data series. Given a set of points in time and data values observed for those points in time, values of a and b are chosen so that S is minimized. Here $Y_t = at + b$ is the trend line, so the sum of

squared deviations from the trend line is what is being minimized. This can always be done in closed form since this is a case of simple linear regression. (36)

4.3.1.1 Linear trend analysis of SVP01

From the linear trend plot below we can see that there is a positive coefficient of 2.3×10^{-5} associated with 't' time indicating an increase in the quantity of VOCs at SVP 01 over time. Which implies that there might be an overall slight increase in the concentrations of VOCs in SVP 01 over the period of the study.

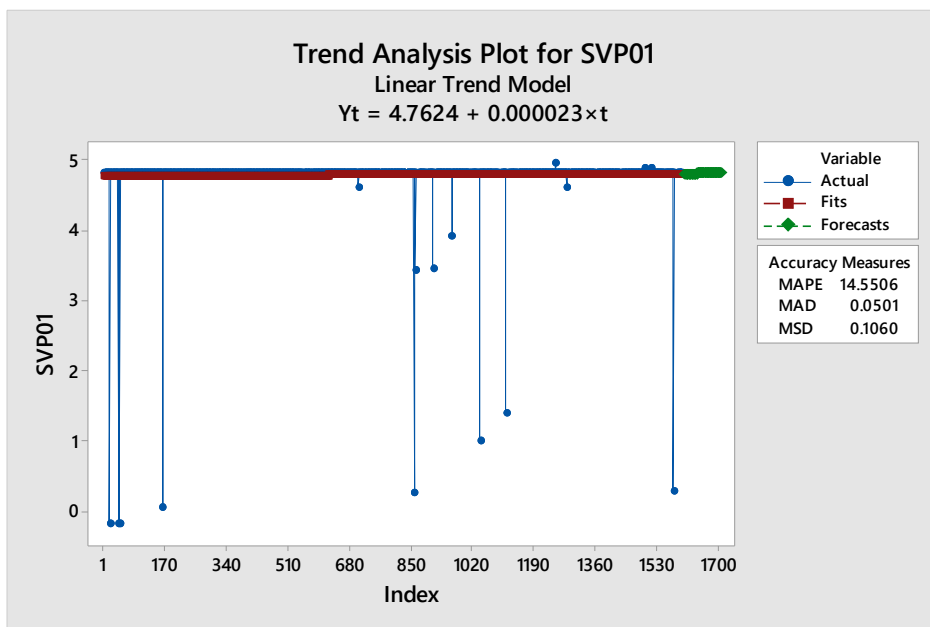


Figure 4.10. Linear trend plot of VOCs at SVP 01

4.3.1.2 Linear trend analysis of SVP02

From the linear trend plot below we can see that there is a positive coefficient of 1.6×10^{-5} associated with 't' time indicating an increase in the quantity of VOCs at SVP 02 over time. Which implies that there might be an overall slight increase in the concentrations of VOCs in SVP 02 over the period of the study.

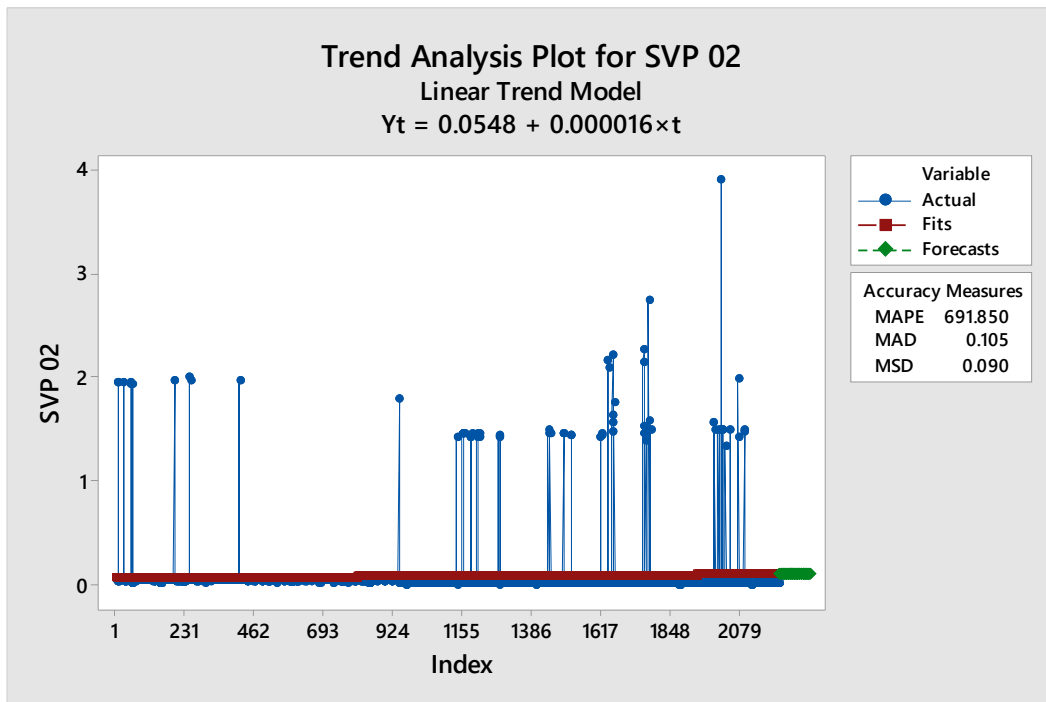


Figure 4.11. Linear trend plot of VOC concentration at SVP 02

4.3.1.3 Linear trend analysis of SVP03

From the linear trend plot below we can see that there is a positive coefficient of 2×10^{-6} associated with 't' time indicating an increase in the quantity of VOCs at SVP 03 over time. Which implies that there might be an overall slight increase in the concentrations of VOCs in SVP 02 over the period of the study.

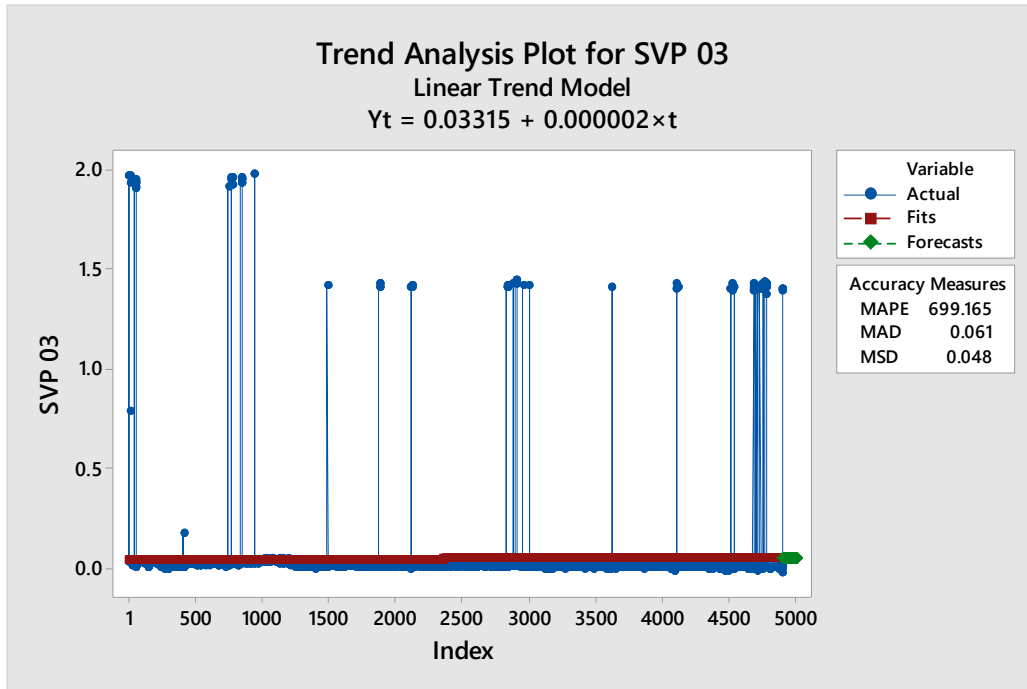


Figure 4.12. Linear trend plot of VOC concentration at SVP 03

4.3.2 X bar and S charts

Xbar-S charts are often used collectively to plot the process mean (Xbar) and subgroup standard deviations (S) over time for continuous data. Xbar-S Chart is used to monitor the mean and variation of a process when you have continuous data and subgroup sizes of 9 or more. This control chart is used to monitor change over time of a day/week/month or season. (37) For my analysis I used my subgroup size as 24 so that I could get daily mean and standard deviation values. The assumptions of using X-Bar S charts are that they need to normally distributed and the number of subgroups should be equal. I made sure that the sub group sizes were equal but the data we have is not normally distributed.

4.3.2.1 X bar and S charts of SVP01

The daily means are indicated in by the blue points in the upper chart below and UCL and LCL stand for upper and lower control limits respectively. The red points in the chart below indicate the daily standard deviation. The average of the daily mean (\bar{x}) of VOC concentration at SVP 01 is 4.735 % and it has a low average daily standard deviation of 0.186.

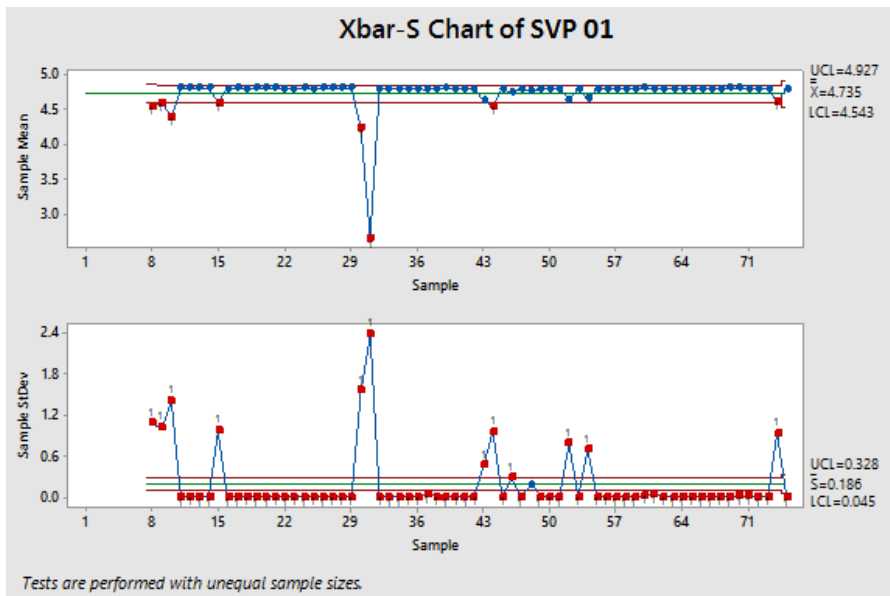


Figure 4.13. Xbar-S chart of VOC concentration at SVP 01

4.3.2.2 X bar and S charts of SVP02

According to the Xbar-S chart below, the average of the daily mean (\bar{x}) VOC concentration at SVP 02 is 0.072 % and average daily standard deviation of 0.148. This technique can be used to determine the release event dates.

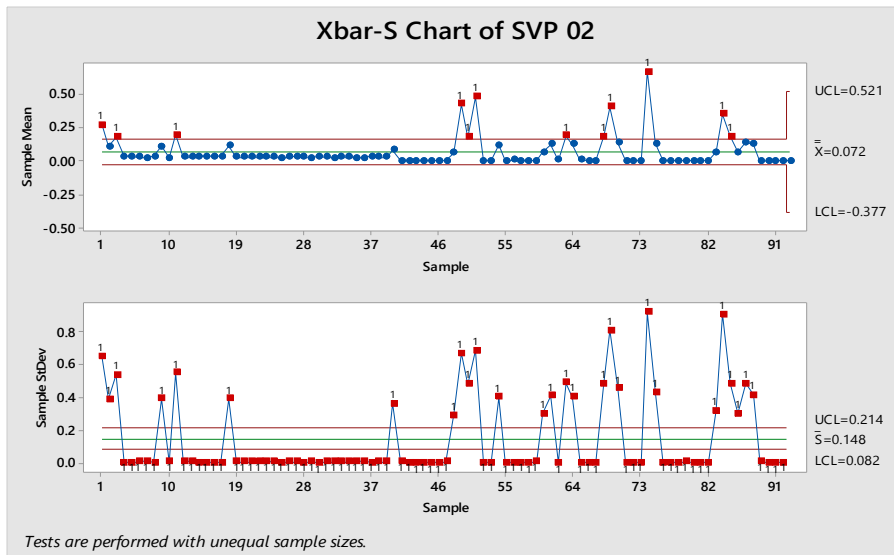


Figure 4.14. Xbar-S chart of VOC concentration at SVP 02

4.3.2.3 X bar and S charts of SVP03

According to the Xbar-S chart below, the average of the daily mean (\bar{x}) of VOC concentration at SVP 03 is 0.0376% and low average daily standard deviation of 0.0664.

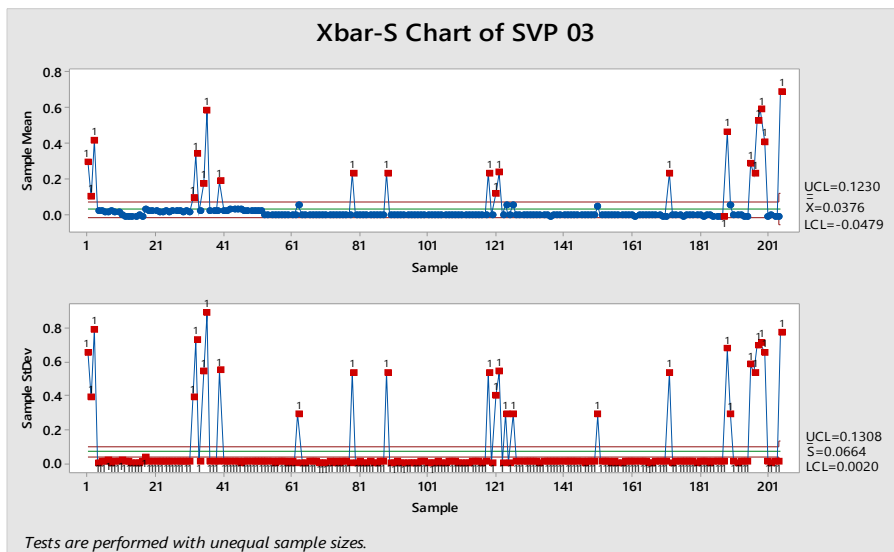


Figure 4.15. Xbar-S chart of VOC concentration at SVP 03

4.3.3 Mann-Kendall and Sen's slope test.

The Mann-Kendall (M-K) Test is a simple test for trend. Mann-Kendall is a non-parametric test and as such, it is not dependent upon irregularly spaced monitoring periods. Mann-Kendall assesses whether a time-ordered data set exhibits an increasing or decreasing trend, within a predetermined level of significance. **(38)**

The purpose of the Mann-Kendall (MK) test is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. **(39)** A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear. The MK test can be used in place of a parametric linear regression analysis, which can be used to test if the slope of the estimated linear regression line is different from zero. The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test, that is, the MK test is a non-parametric (distribution-free) test.

If a significant trend is found, the rate of change can be calculated using the Sen Slope estimator. **(40)** For linear trend, the slope is usually estimated by computing the least squares estimate using linear regression. However, it is only valid when there is no serial correlation, and the method is very sensitive to outliers. A more robust method was developed by Sen. This test computes both the slope (i.e. linear rate of change) and intercept according to Sen's method.

The following analysis was performed in R programming language. The Mann-Kendall and Sen Slope test assumes that the sample values very taken independently which is true in our case. The results of the Mann-Kendall tests may not be conclusive since the monotonic rise or fall in the readings might be caused by sensor-drift.

4.3.3.1 Slope tests on SVP01:

Both Mann- Kendall and the Sen Slope estimator tests were performed and the results for the VOC concentrations at SVP 01 is shown in the tables below. The negative tau value in the Mann-Kendall test below indicates a falling trend may be present for VOC concentrations at SVP01.

Table 4.1 Mann-Kendall test results for VOC concentrations at SVP 01

summary(MannKendall(svp01_data))
Score = -14858 , Var(Score) = 11045518
denominator = 96064.23
tau = -0.155, 2-sided pvalue =7.8105e-06

The negative z value for Sen`s slope test indicates that there might be a falling trend for VOC concentrations at SVP01 with a slope value of near zero.

Table 4.2 Sen Slope test results for VOC concentrations at SVP 01

sens.slope(svp01_data)
z = -4.4703, n = 503, p-value = 7.81e-06
alternative hypothesis: true z is not equal to 0
95 percent confidence interval:

0 0
Sen's slope
0

4.3.3.2 Slope tests on SVP02:

The positive tau value for Mann-Kendall test in the table below indicates a rising trend may be present for VOC concentrations at SVP02.

Table 4.3 Mann-Kendall test results for VOC concentrations at SVP 02

summary(MannKendall(svp02_data))
Score = -6641 , Var(Score) = 161497
denominator = 12997
tau = 0.611, 2-sided pvalue =< 2.22e-12

The positive z value for Sen's slope test indicates a rising trend may be present for VOC concentrations at SVP02 with an estimated positive slope value of 3.088235e-06.

Table 4.4 Sen slope test results for VOC concentrations at SVP 02

sens.slope(na.omit(svp02_data))
z = 56.23, n = 5326, p-value < 2.2e-16
alternative hypothesis: true z is not equal to 0
95 percent confidence interval:
3.33750e-06 2.84509e-06
Sen's slope
3.088235e-06

4.3.3.3 Slope tests on SVP03:

The positive tau value for Mann-Kendall test indicates that a rising trend may be present for VOC concentrations at SVP03.

Table 4.5 Mann-Kendall test results for VOC concentrations at SVP 03

summary(MannKendall(svp03_data))
Score = -1340793 , Var(Score) = 4071198720
denominator = 5143188
tau = 0.261, 2-sided pvalue =< 2.22e-16

The positive z value for Sen's slope test indicates a rising trend may be present for VOC concentrations at SVP02 with a slope value of 2.690374e-06.

Table 4.6 Sen slope test results for VOC concentrations at SVP 03

sens.slope(na.omit(svp03_data))
z = -21.014, n = 3393, p-value < 2.2e-16
alternative hypothesis: true z is not equal to 0
95 percent confidence interval:
2.062162e-06 3.212761e-06
Sen's slope
2.690374e-06

4.4. Time series analysis

Temporal statistical analysis enables us to examine and model the behavior of a variable in a data set over time (e.g., to determine whether and how concentrations are changing over time). The behavior of a variable in a data set over time can be modeled as a function of previous data points of the same series, with or without extraneous, random influences.

Time series analysis accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for. Time series is defined as an ordered sequence of values of a variable at equally spaced time intervals. The applications of time series models are two-fold. Firstly, to obtain an understanding of the underlying forces and structure that produced the observed data. Secondly, to fit a model and proceed to forecasting, monitoring or even feedback and feedforward control.

4.4.1 Decomposition of time - series

The decomposition of time series is a statistical task that deconstructs a time series into several components, each representing one of the underlying categories of patterns. This is an important technique for all types of time series analysis, especially for seasonal adjustment. It seeks to construct, from an observed time series, a number of component series (that could be used to reconstruct the original by additions or multiplications) where each of these has a certain characteristic or type of behavior. For example, time series are usually decomposed into:

- The trend component at time t , which reflects the long-term progression of the series (secular variation). A trend exists when there is a persistent increasing or decreasing direction in the data. The trend component does not have to be linear.
- The cyclical component at time t , which reflects repeated but non-periodic fluctuations. The duration of these fluctuations is usually of at least two years.
- The seasonal component at time t , reflecting seasonality (seasonal variation). A seasonal pattern exists when a time series is influenced by seasonal factors. Seasonality occurs over a fixed and known period (e.g., the quarter of the year, the month, or day of the week).
- The irregular component (or "noise") at time t , which describes random, irregular influences. It represents the residuals or remainder of the time series after the other components have been removed.

More extensive decompositions might also include long-run cycles, holiday effects, day of week effects and so on. Here, we'll only consider trend and seasonal decompositions. One of the main objectives for a decomposition is to estimate seasonal effects that can be used to create and present seasonally adjusted values. A seasonally adjusted value removes the seasonal effect from a value so that trends can be seen more clearly. To fit a model that weights all observations equally to determine the best regression fit, perform a decomposition analysis. Use when your series exhibits a seasonal pattern, with or without a trend. **(41)**

Singular spectrum analysis (SSA) is a nonparametric spectral estimation method. It combines elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamical systems and signal processing. Its roots lie in the classical Karhunen (1946)–Loève (1945, 1978) spectral decomposition of time series and random fields and in the Mañé (1981)–Takens (1981) embedding theorem. SSA can be an aid in the decomposition of time series into a sum of components, each having a meaningful interpretation. **(42)**

In the original formulation of SSA it was assumed that the time series under analysis has a deterministic component (such as a trend and/or a seasonal) with noise superimposed and that the deterministic component can be successfully extracted from the noise.

4.4.1.1 Decomposition of SVP01 time series

The graph below shows the detrended and seasonally adjusted time series data. Since there is not much of a variation in VOC concentrations at SVP01 hence there is not much of a variation in the graphs below.

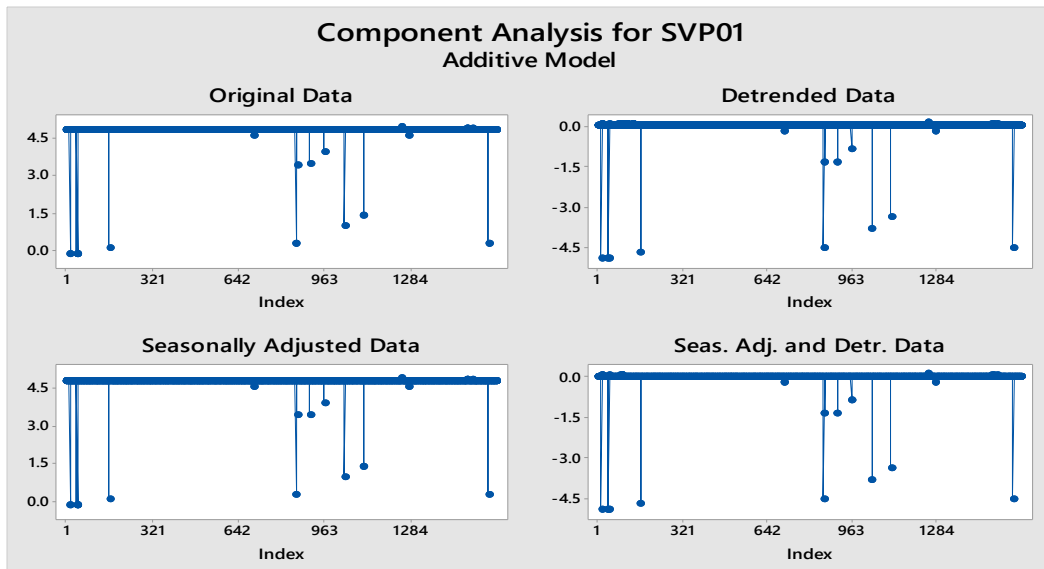


Figure 4.16. Decomposed data of VOC concentrations at SVP 01

The graph below shows percent variation during the hours of the day which is useful in understanding when the most change occurs in the time series. For VOC concentrations in soil gas at SVP01 the most change occurs in the evening hours.

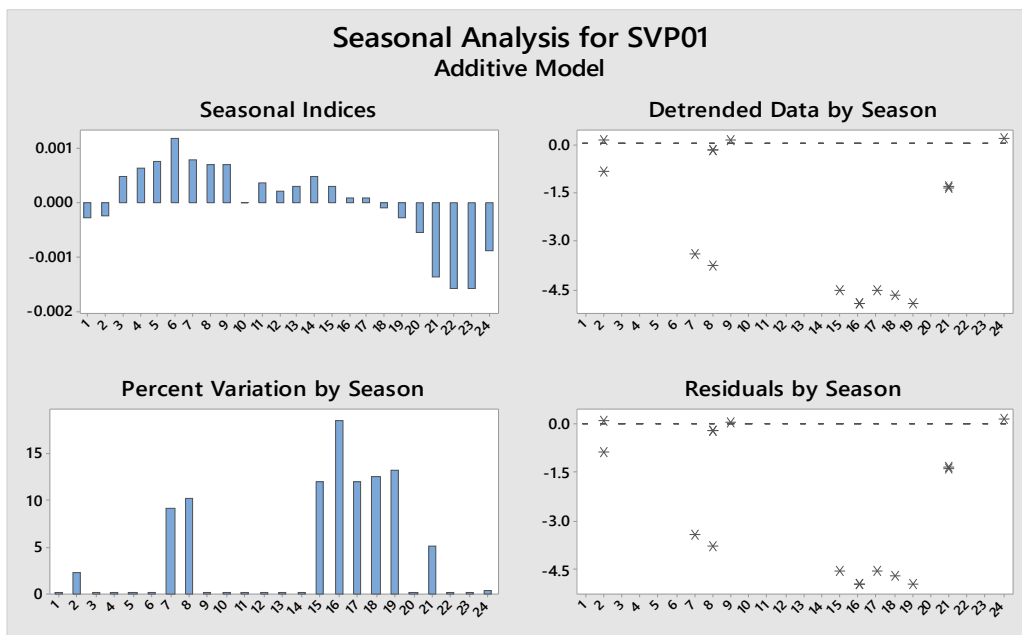


Figure 4.17. Seasonal analysis graphs of VOC concentrations at SVP 01

4.4.1.2 Decomposition of SVP02 time series

The detrended graphs for VOC concentrations at SVP 02 are shown below.

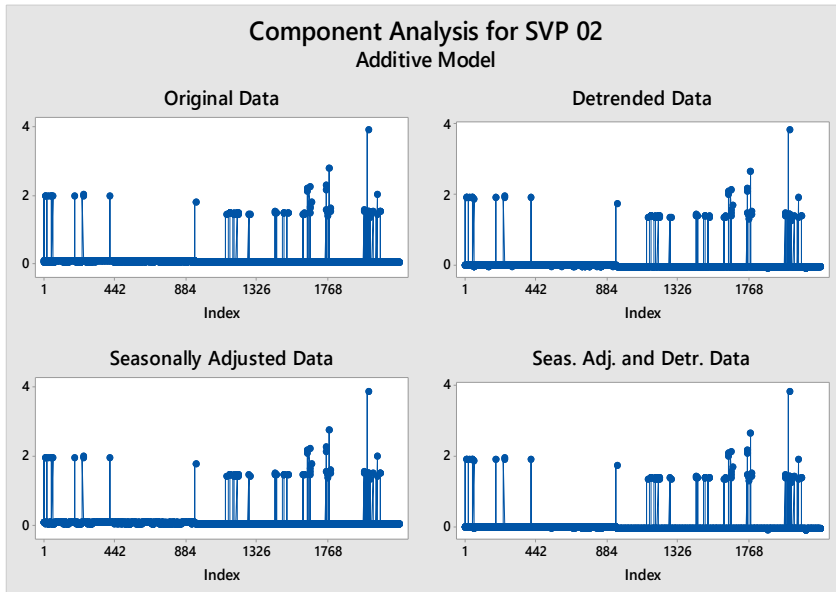


Figure 4.18. Decomposed data of VOC concentrations at SVP 02

The VOC concentration in the soil gas at SVP02 has maximum changes occurring during the morning and some in the afternoon hours as indicated in the graph below.

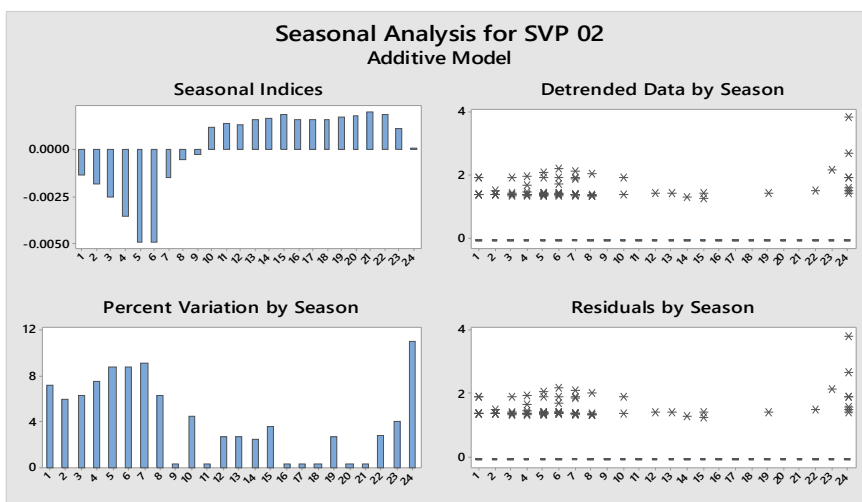


Figure 4.19. Seasonal analysis graphs of VOC concentrations at SVP 02

4.4.1.3 Decomposition of SVP03 time series

The detrended graphs for VOC concentrations at SVP 03 are shown below.

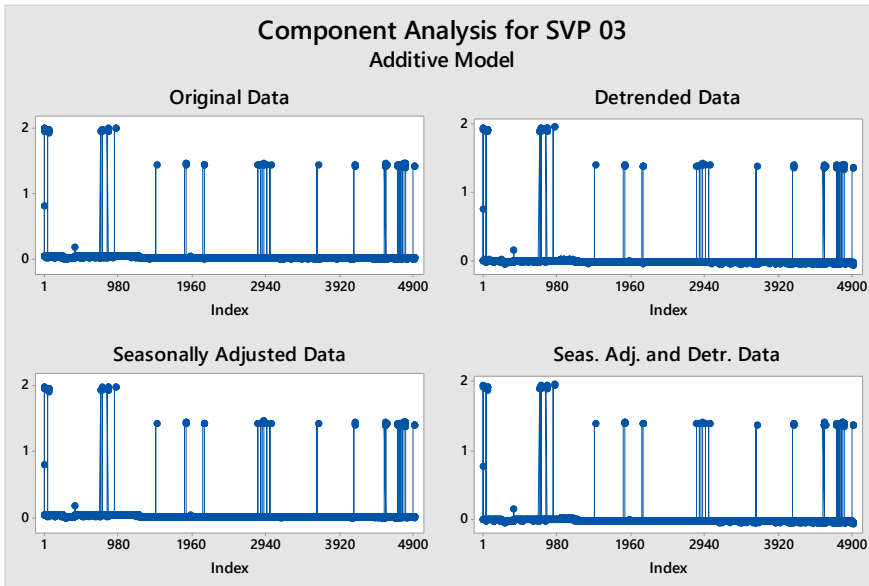


Figure 4.20. Decomposed data of VOC concentrations at SVP 03

The maximum change for VOC concentrations in soil gas at SVP 03 occurs during the morning hours.

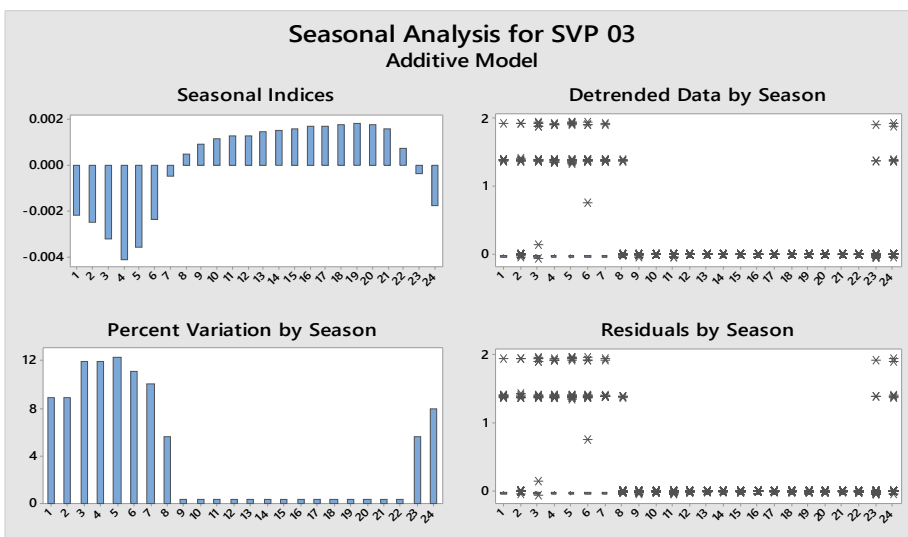


Figure 4.21. Seasonal analysis graphs of VOC concentrations at SVP 03

4.4.2 Autocorrelation

To measure how well observations at different points of time correlate with each other and look for a seasonal pattern, perform an autocorrelation analysis. The autocorrelation function can be used for the following two purposes namely to detect non-randomness in data and to identify an appropriate time series model if the data are not random. (43)

4.4.2.1 Autocorrelation of SVP01 time series

From the graph below it is evident that VOC concentrations at SVP 01 seem to relate to itself the most in 1-10 lag points i.e. 1-10 hours with 1 and 3 being the highest.

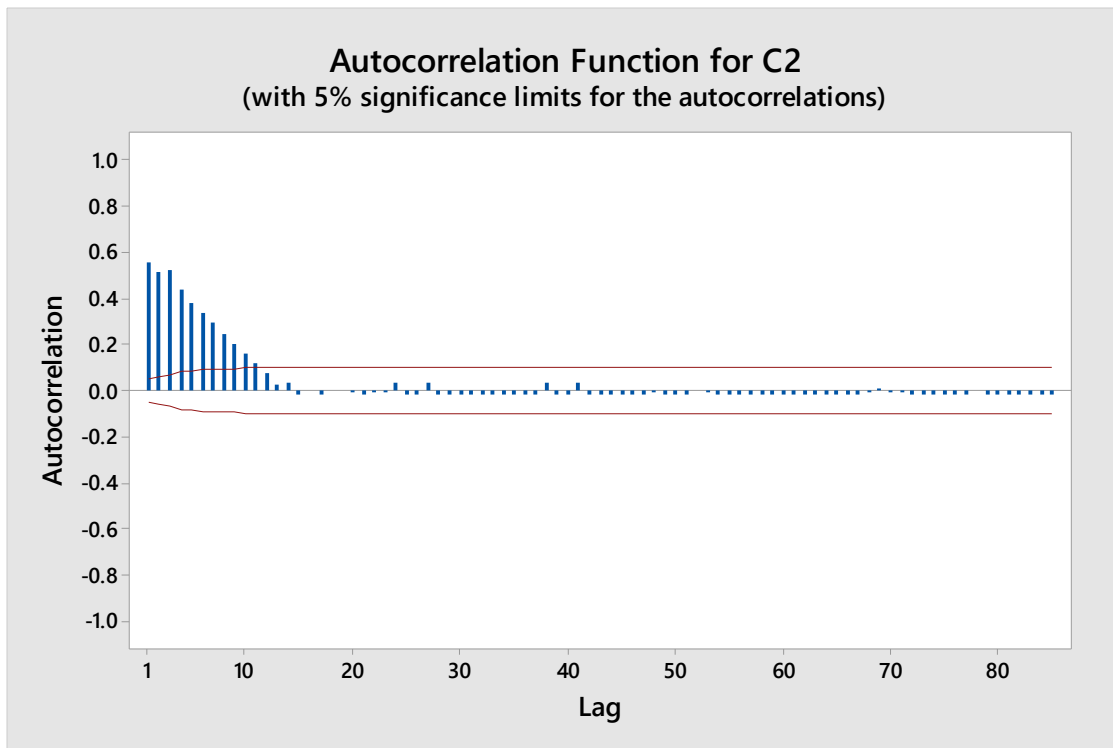


Figure 4.22. Autocorrelation graphs of VOC concentrations at SVP 01

4.4.2.2 Autocorrelation of SVP02 time series

From the graph below it is evident that VOC concentrations at SVP 02 seem to relate to itself the most in 1-5 hours. There is also a high correlation with VOC concentrations with itself at 24 hrs and comparatively high on 48 hrs, indicating a daily correlation.

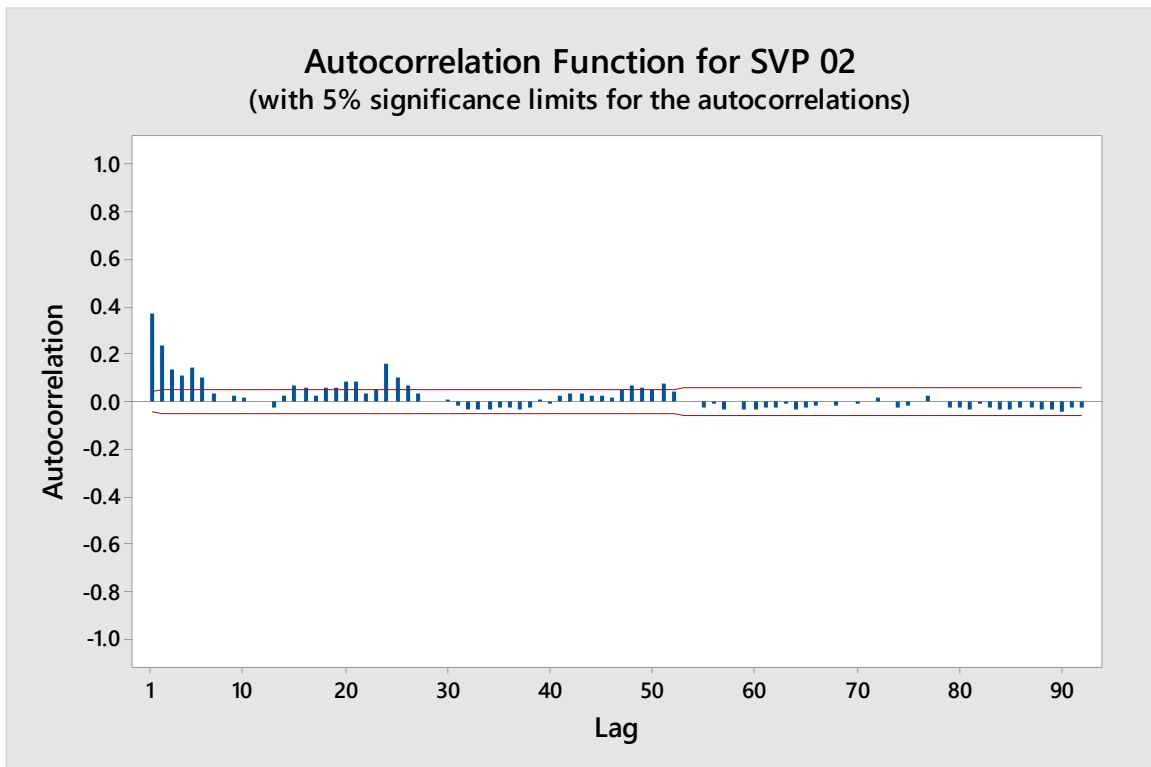


Figure 4.23. Autocorrelation graphs of VOC concentrations at SVP 02

4.4.2.3 Autocorrelation of SVP03 time series

From the graph below it is evident that VOC concentrations SVP 03 seem to relate to itself the most in 1-5 hours. There is also a high correlation with 24, 48, 72 and comparatively high on 96 hours

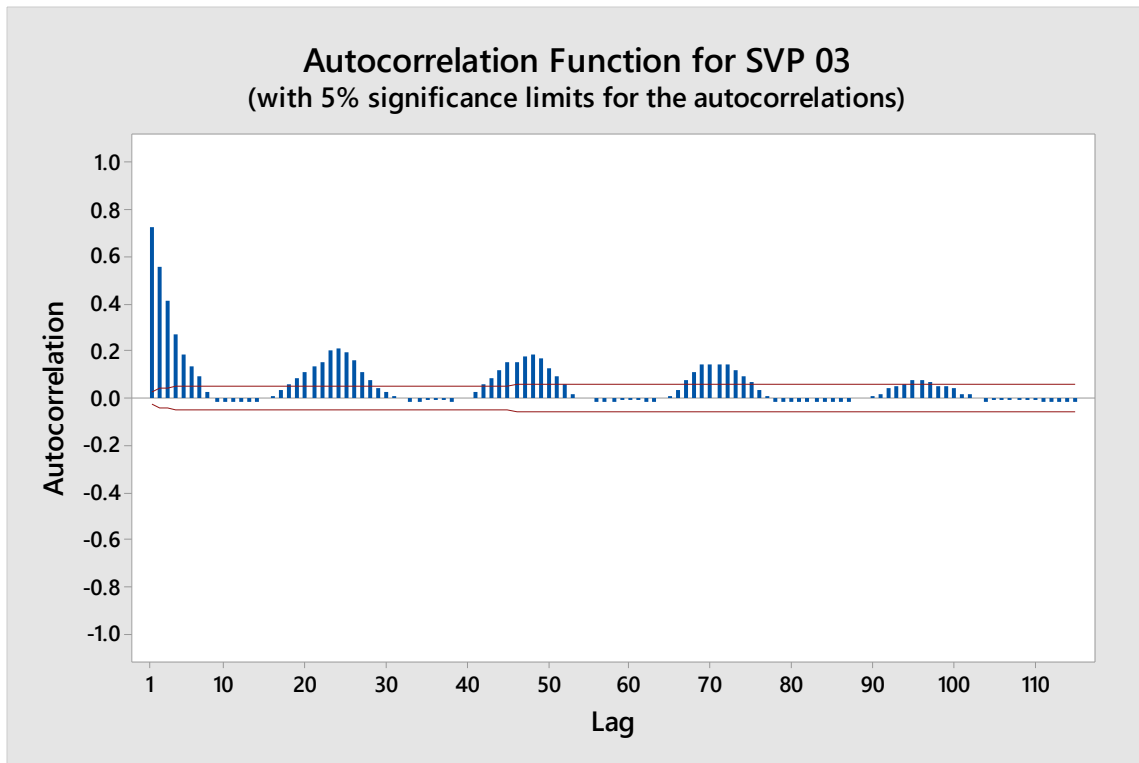


Figure 4.24. Autocorrelation graphs of VOC concentrations at SVP 03

4.5 Regression and Correlation

Regression analysis is a way of mathematically sorting out which of the variables factors has an impact on the data of concern. It answers the questions: Which factors matter most? Which can we ignore? How do those factors interact with each other? And, perhaps most importantly, how certain are we about all of these factors? In regression analysis, those factors are called variables. You have your dependent variable, the main factor that you're trying to understand or predict (Soil gas data in our case). And then you have your independent variables, the factors you suspect have an impact on your dependent variable (the weather and ground water quality data). It will also give you a slew of statistics (including a p-value and a correlation

coefficient) to tell you how accurate your model is. **(45)** Most elementary statistical analysis involve making scatter plots and performing linear regression.

4.5.1 Regression of soil gas with ground water data

The ground water quality data was collected from 3 sensors placed in about 12 ft deep wells namely MW02, MW03 and MW05. The water temperature and conductivity data were used for the purpose of the following correlation analysis. Correlation is a technique for investigating the relationship between two quantitative, continuous variables.

4.5.1.1 Pearson Correlation test

Pearson's correlation coefficient (r) is a measure of the strength of the association between the two variables. The first step in studying the relationship between two continuous variables is to draw a scatter plot of the variables to check for linearity. The correlation coefficient should not be calculated if the relationship is not linear. For correlation only purposes, it does not really matter on which axis the variables are plotted. However, conventionally, the independent (or explanatory) variable is plotted on the x-axis (horizontally) and the dependent (or response) variable is plotted on the y-axis (vertically).

The assumptions of Pearson correlation test are that the data should be normal, homoscedastic, linear, continuous and paired. We do not meet most of the assumption except continuous and paired but our data is not normal or linear.

4.5.1.1.1 Pearson correlation test for SVP01

The p value of MW02 is the only one below 0.05, hence SVP01 is statistically correlated to MW02 which makes sense since it is the closest water well to it.

Table 4.7. Pearson correlation for VOC concentrations at SVP 01 and ground water quality data

Pearson correlation			
SVP 01	MW02	MW03	MW05
Pearson correlation	-0.095	0.027	-0.038
P-Value	0.004	0.409	0.257

The scatter plot of VOC concentrations at SVP01 and MW02 ground water conductivity is shown below and it shows the points that follow the fitted line and the outliers.

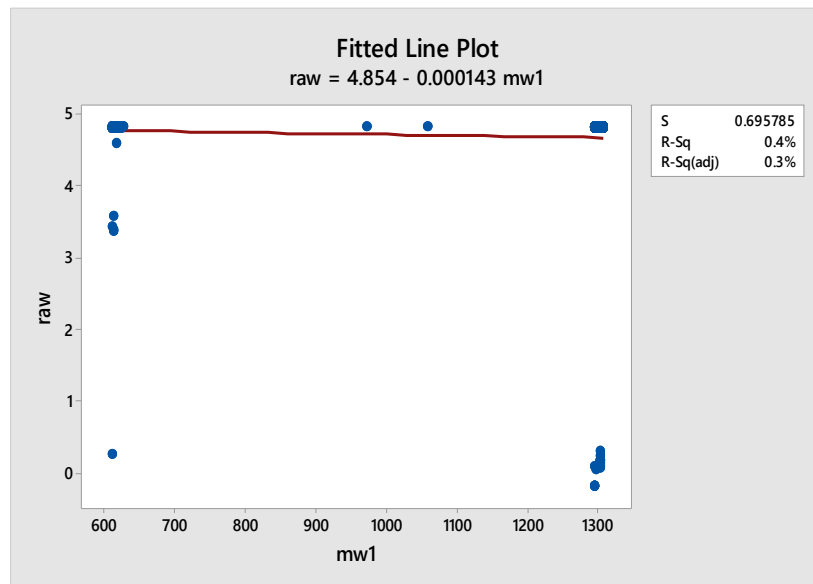


Figure 4.25. Scatter plot of groundwater conductivity at MW02 and VOC concentrations at SVP01

4.5.1.1.2 Pearson correlation test for SVP02

The p value of MW02, MW03 and MW05 are all below 0.05, hence SVP02 is at least slightly correlated to all three ground water wells. It is the most correlated to MW05 since that has the highest Pearson’s correlation constant followed by MW02 then MW03 which are the closest water well sites to them.

Table 4.8. Pearson correlation for VOC concentrations at SVP 02 and ground water quality data

Pearson correlation			
SVP 02	MW02	MW03	MW05
Pearson correlation	0.184	0.150	-0.249
P-Value	0.000	0.000	0.00

The scatter plot of ground water conductivity and temperature at MW05 vs SVP02 are shown below. We can see that conductivity has higher R-squared values indicating that ground water conductivity is a more important factor than ground water temperature.

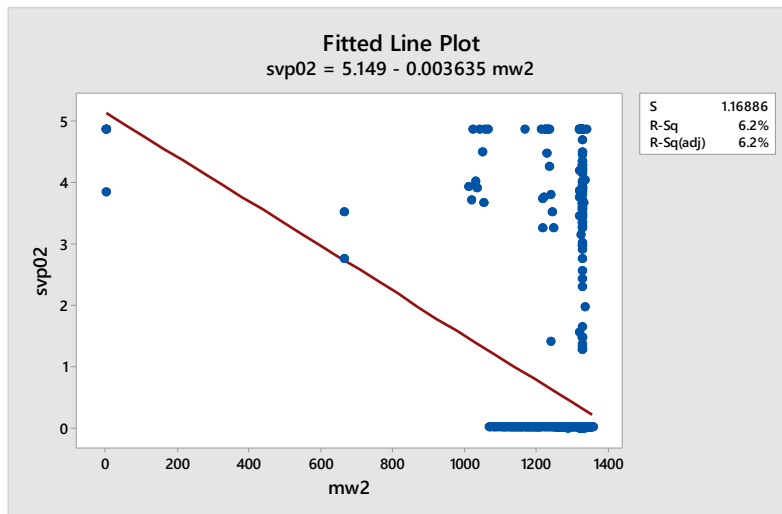


Figure 4.26. Scatter plot of ground water conductivity at MW 05 and VOC concentrations at SVP 02

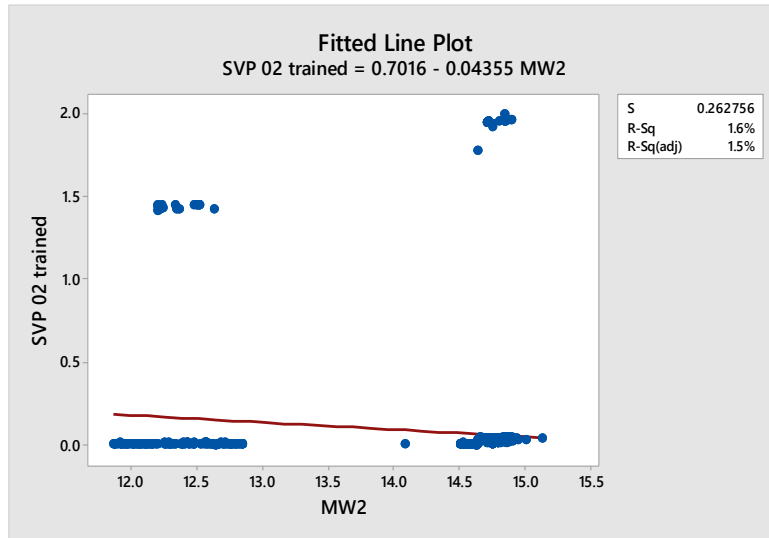


Figure 4.27. Scatter plot of groundwater temperature at MW05 and VOC concentrations at SVP02

4.5.1.1.3 Pearson correlation test for SVP03

The p value of MW02, MW03 and MW05 are all above 0.05, hence SVP03 is not statistically correlated to any of the ground water wells

Table 4.9. Pearson correlation for VOC concentrations at SVP 03 and ground water quality data

Pearson correlation			
SVP 02	MW02	MW03	MW05
Pearson correlation	-0.027	0.005	0.021
P-Value	0.283	0.839	0.402

4.5.2.2 Cross Correlation test

Cross correlation function is used to determine whether there is a relationship between two time series. Usually, a correlation is significant when the absolute value is greater than $\frac{2}{\sqrt{n - |k|}}$, where n is the number of observations and k is the lag. This calculation is a rule of thumb procedure based on large-sample normal approximation. Since approximately 95% of a normal population is within 2 standard deviations of the mean, a test that rejects the hypothesis that the population cross correlation of lag k equals zero when $|r_{xy}(k)|$ is greater than $2/\sqrt{n - |k|}$ has a significance level (α) of approximately 5%. (45)

A cross correlation test performed on GW conductivity of MW05 and VOC concentration at SVP02 in the graph below shows that the conductivity of the ground water is correlated to the soil gas VOC content the most at 24-48 lags i.e. hours.

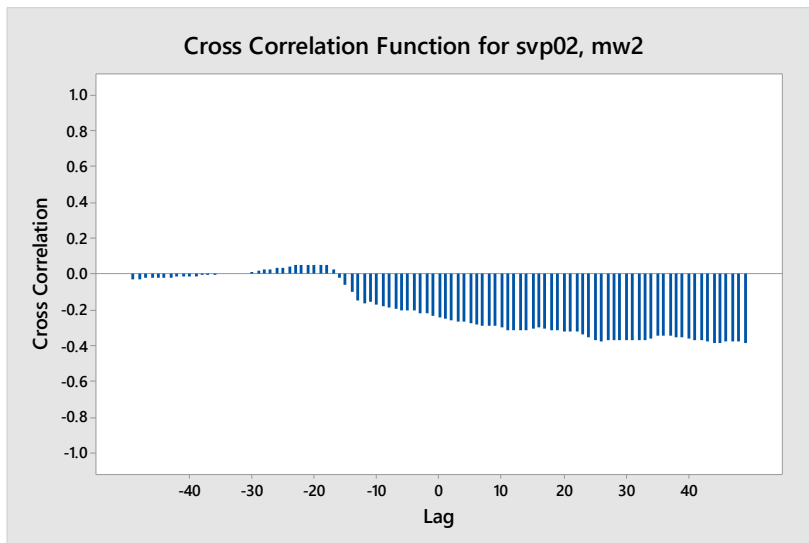


Figure 4.28. Cross correlation of groundwater conductivity at MW05 with VOC concentrations at SVP02

4.5.2 Regression of soil gas with weather data

The weather and evapotranspiration data were imported from northernwater.org website for the Longmont area from a weather monitoring station that was closest to the site of interest. The data was converted to hourly averages to coincide with the soil gas data collected. Since SVP01 does not vary as much as the other two soil gas monitoring sites, we performed regression analysis with only the other two soil gas monitoring sites i.e. SVP02 and SVP03. The one-way analysis of variance (ANOVA) is used to determine whether there are any statistically significant differences between the means of three or more independent (unrelated) groups. The one-way ANOVA compares the means between the groups you are interested in and determines whether any of those means are statistically significantly different from each other. **(46)**

The basic assumptions of linear regression is that the data sets need to have a linear relationship, multivariate normality, no or little multicollinearity, no auto-correlation and homoscedasticity

4.5.2.1 Regression Analysis: SVP 02 with weather data

From the p values of the ANOVA test performed of SVP02 and weather factors below it is evident solar radiation, wind speeds and catchment have statistically the least correlation to the soil gas content.

Table 4.10. ANOVA results for VOC concentrations at SVP 02 and the weather data

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	12	65.379	5.4482	6.9	0
Ave Air Temp (°F)	1	8.401	8.4011	10.64	0.001

Ave Soil Temp (°F)	1	6.955	6.9554	8.81	0.003
Rel Humidity (%)	1	1.676	1.6756	2.12	0.146
Vapor Pressure (kPa)	1	21.621	21.621	27.39	0
kp EvapPan	1	13.281	13.2807	16.82	0
Rain (TB) (in)	1	7.057	7.0571	8.94	0.003
Catch (WB) (in)	1	2.053	2.0526	2.6	0.107
Precip (WB) (in)	1	5.071	5.0706	6.42	0.012
Solar Rad Tot (cal/cm2)	1	0.006	0.0057	0.01	0.932
Rso Clear Sky (cal/cm2)	1	0.583	0.5825	0.74	0.391
Ave Wind Speed (3m) (mph)	1	1.795	1.7955	2.27	0.132
Ave Wind Speed (2m) (mph)	1	1.767	1.767	2.24	0.135
Error		556	438.967	0.7895	
Total		568	504.345		

From the summary table below the R squared value for the analysis is low, we cannot use this equation to calculate an exact quantity of VOCs in soil gas using just the weather data but we can use the equation to see if the factors are directly or inversely related to soil gas depending on the +/- sign in front of the factor. The magnitude of the coefficient in front of the factor can also help us judge how important that factor is when compared to the other factors. So from the equation below we can say that air temperature, soil temperature, humidity, evapotranspiration are positively correlated to the soil gas content while rain, precipitation and vapor pressure are inversely correlated to the VOC at SVP02 soil gas content.

Table 4.11. ANOVA Model summary for VOC concentrations at SVP 02 and weather data

Model Summary			
S	R-sq	R-sq(adj)	R-sq(pred)
0.88854	12.96%	11.08%	7.52%

Regression Equation
SVP 02 = -9.35 + 0.03110 Ave Air Temp (°F) + 0.0333 Ave Soil Temp (°F)
+ 0.00893 Rel Humidity (%) - 2.208 Vapor Pressure (kPa) + 9.95 kp EvapPan
- 14.13 Rain (TB) (in) - 0.1370 Catch (WB) (in) - 13.84 Precip (WB) (in)
- 0.00025 Solar Rad Tot (cal/cm2) - 0.00208 Rso Clear Sky (cal/cm2)
+ 13.88 Ave Wind Speed (3m) (mph) - 15.0 Ave Wind Speed (2m) (mph)

From the table below we can see that coefficients for vapor pressure, precipitation, and evaporation are pretty high, by an order of magnitude, when compared to other factors.

Table 4.12. ANOVA model coefficients for VOC concentrations at SVP 02 and weather data

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-9.35	1.98	-4.71	0	
Ave Air Temp (°F)	0.0311	0.00953	3.26	0.001	11.29
Ave Soil Temp (°F)	0.0333	0.0112	2.97	0.003	5.1
Rel Humidity (%)	0.0893	0.00613	1.46	0.146	20.31
Vapor Pressure (kPa)	-2.208	0.422	-5.23	0	10.01
kp EvapPan	9.95	2.43	4.1	0	26.88
Rain (TB) (in)	14.13	4.73	2.99	0.003	4.95
Catch (WB) (in)	-0.137	0.085	-1.61	0.0107	3.43
Precip (WB) (in)	-13.84	5.46	-2.53	0.012	5.16
Solar Rad Tot (cal/cm2)	-0.00025	0.00291	-0.09	0.932	4.13
Rso Clear Sky (cal/cm2)	-0.00208	0.00242	-0.86	0.391	3.06
Ave Wind Speed (3m) (mph)	13.88	9.2		0.132	1746211
Ave Wind Speed (2m) (mph)	-15	10	-1.5	0.015	1746261

The positive correlation between air and soil temperature and SVP02 soil gas VOC content can be seen in the graph below.

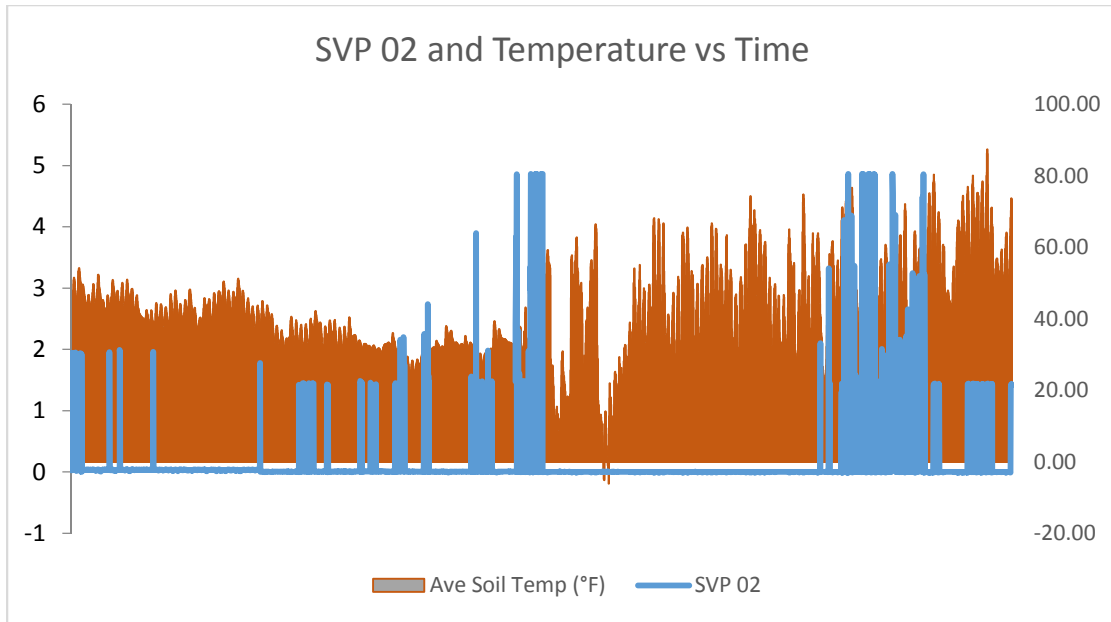


Figure 4.29. Linear graph of VOC concentration at SVP 02 and air and soil temperatures vs time

The positive correlation between humidity in the air and SVP02 soil gas VOC content can be seen in the graph below.

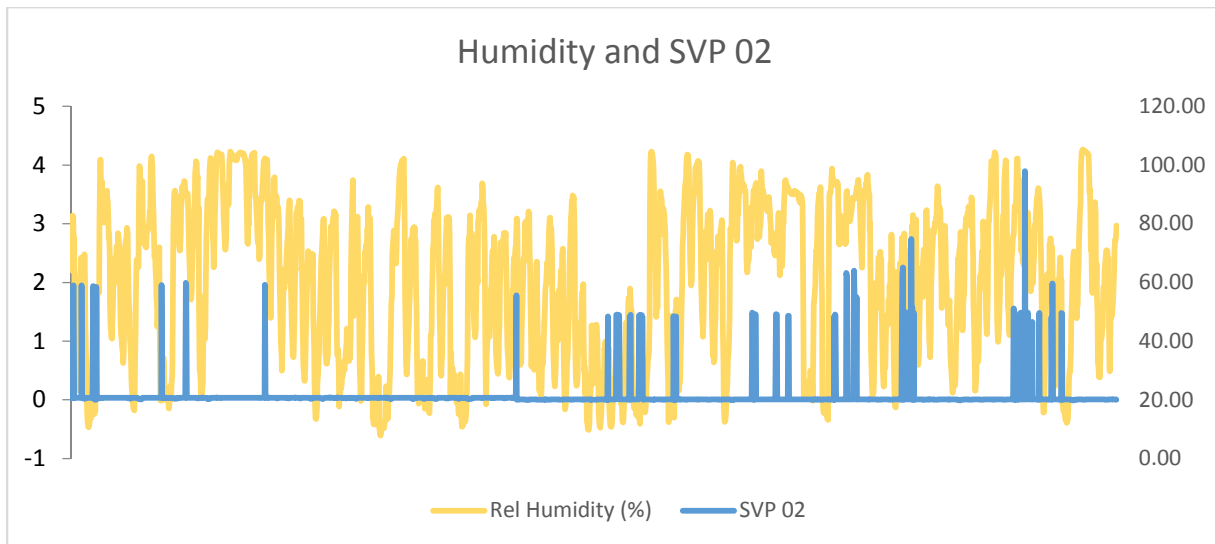


Figure 4.30. Linear graph of VOC concentrations at SVP 02 and humidity vs time

The negative correlation between vapor pressure and SVP02 soil gas VOC content can be seen in the graph below.

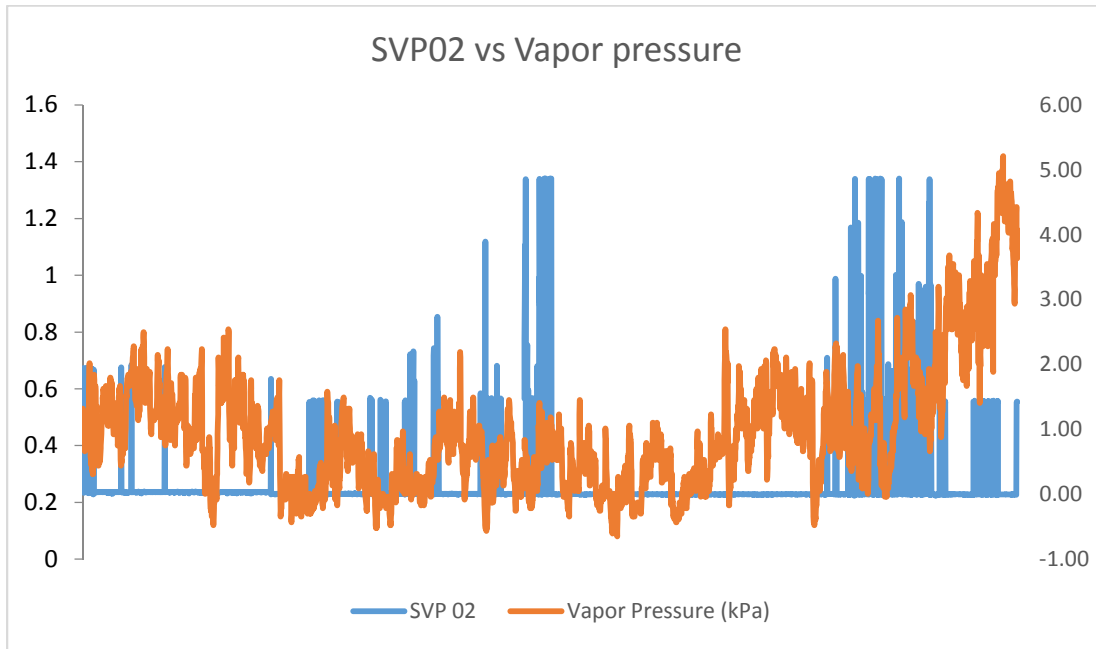


Figure 4.31. Linear graph of VOC concentrations at SVP 02 and vapor pressure vs time

4.5.2.2 Regression Analysis: SVP 03 with weather data

From the p values of the ANOVA test performed with VOC concentrations at SVP03 and weather factors below it is evident that air temperature, solar radiation, wind speeds, rain and precipitation have statistically the least correlation to the soil gas content.

Table 4.13. ANOVA results for VOC concentrations at SVP 03 and weather data

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F- Value	P - Value
Regression	12	1.4809	0.123409	4.86	0
Ave Air Temp (°F)	1	0.0005	0.00049	0.02	0.89

Ave Soil Temp (°F)	1	0.254	0.253967	10	0.002
Rel Humidity (%)	1	0.019	0.018976	0.75	0.388
Vapor Pressure (kPa)	1	0.0116	0.011584	0.46	0.5
kp EvapPan	1	0.0238	0.02376	0.94	0.334
Rain (TB) (in)	1	0.0001	0.00012	0	0.945
Catch (WB) (in)	1	0.023	0.022956	0.9	0.342
Precip (WB) (in)	1	0.005	0.004958	0.2	0.659
Solar Rad Tot (cal/cm2)	1	0.0016	0.001632	0.06	0.8
Rso Clear Sky (cal/cm2)	1	0.1519	0.151902	5.98	0.015
Ave Wind Speed (3m) (mph)	1	0.0175	0.017535	0.69	0.406
Ave Wind Speed (2m) (mph)	1	0.0176	0.017582	0.69	0.406
Error		556	14.121	0.025398	
Total		568	15.6019		

Similar to SVP02, from the table below the R squared value for the analysis is low, we cannot use this equation to calculate an exact quantity of VOCs in soil gas using just the weather data, but we can use the equation to see if the factors are directly or inversely related to soil gas depending on the +/- sign in front of the factor. The magnitude of the coefficient in front of the factor can also help us judge how important that factor is when compared to the other factors. So, from the equation below we can say that soil temperature, and humidity are positively correlated to the soil gas content while vapor pressure is inversely correlated to SVP03 soil gas content.

Table 4.14. ANOVA model summary for VOC concentrations at SVP 03 and weather data

Model Summary			
S	R-sq	R-sq(adj)	R-sq(pred)
0.159366	9.49%	7.54%	5.11%

Regression Equation
SVP 03 = -0.027 - 0.00024 Ave Air Temp (°F) + 0.00636 Ave Soil Temp (°F)
+ 0.00095 Rel Humidity (%) - 0.0511 Vapor Pressure (kPa) - 0.421 kp EvapPan
+ 0.058 Rain (TB) (in) + 0.0145 Catch (WB) (in) - 0.433 Precip (WB) (in)
+ 0.000132 Solar Rad Tot (cal/cm2) - 0.001063 Rso Clear Sky (cal/cm2)
+ 1.37 Ave Wind Speed (3m) (mph) - 1.49 Ave Wind Speed (2m) (mph)

Table 4.15. ANOVA model coefficients for VOC concentrations at SVP 03 and weather data

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-0.027	0.356	-0.08	0.939	
Ave Air Temp (°F)	-0.00024	0.00171	-0.14	0.89	11.29
Ave Soil Temp (°F)	0.00636	0.00201	3.16	0.002	5.1
Rel Humidity (%)	0.00095	0.0011	0.86	0.388	20.31
Vapor Pressure (kPa)	-0.0511	0.0757	-0.68	0.5	10.01
kp EvapPan	-0.421	0.435	-0.97	0.334	26.88
Rain (TB) (in)	0.058	0.848	0.07	0.945	4.95
Catch (WB) (in)	0.0145	0.0152	0.95	0.342	3.43
Precip (WB) (in)	-0.433	0.98	-0.44	0.659	5.16
Solar Rad Tot (cal/cm2)	0.000132	0.000522	0.23	0.8	4.13
Rso Clear Sky (cal/cm2)	-0.00106	0.000435	-2.45	0.015	3.06
Ave Wind Speed (3m) (mph)	1.37	1.65	0.83	0.406	1746211
Ave Wind Speed (2m) (mph)	-1.49	1.79	-0.83	0.406	1746261

The correlation between soil and air temperature and SVP03 soil gas VOC content can be seen visually in the graph below.

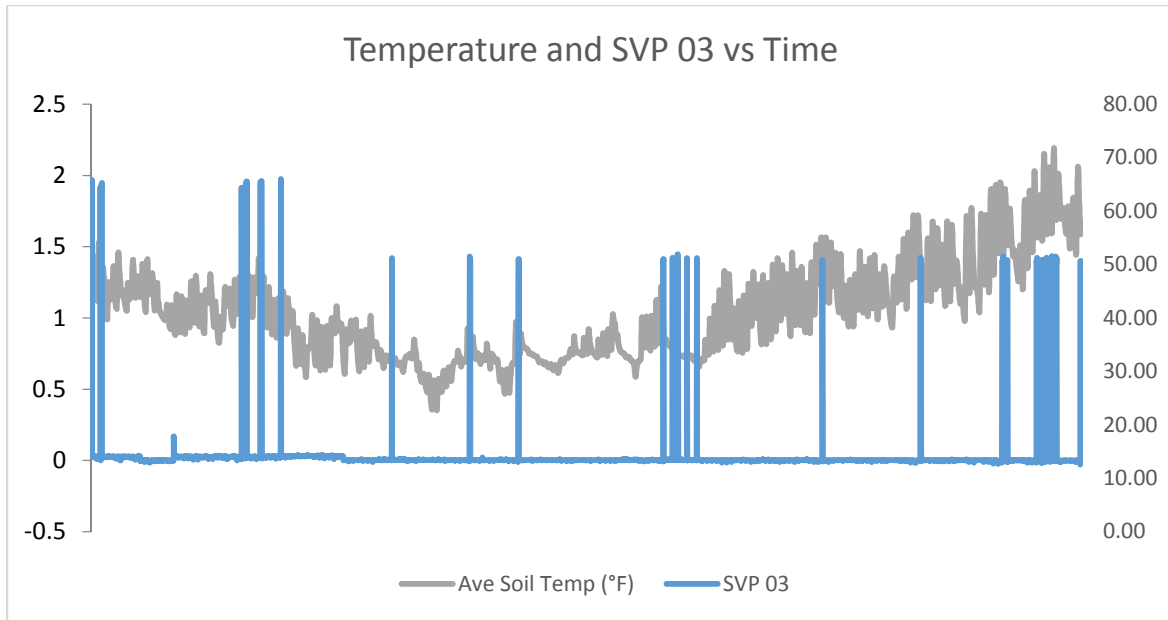


Figure 4.32. Linear graph of VOC concentrations at SVP 03 and air and soil temperature vs time

The correlation between humidity in the air and SVP03 soil gas VOC content can be seen in the graph below.

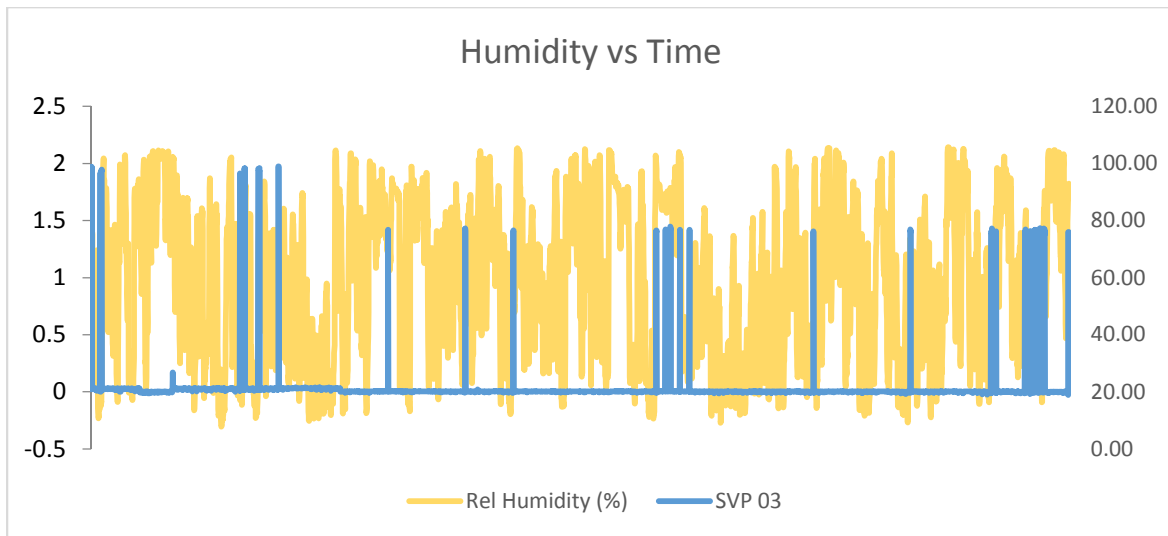


Figure 4.33. Linear graph of VOC concentrations at SVP 03 and humidity vs time

The correlation between vapor pressure and precipitation and SVP03 soil gas VOC content can be seen in the graph below.

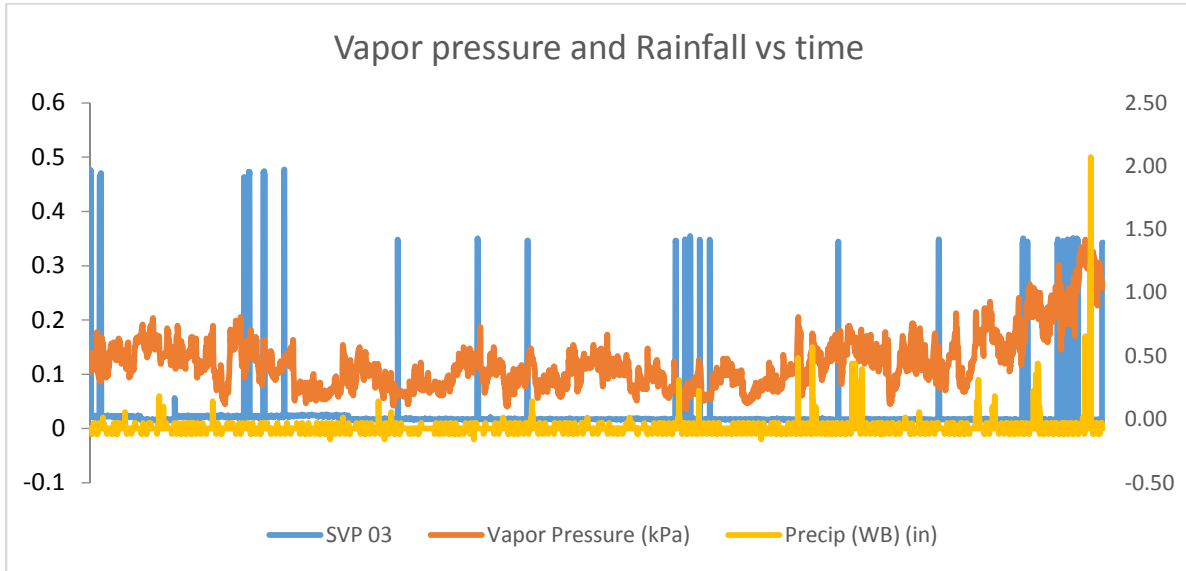


Figure 4.34. Linear graph of VOC concentrations at SVP 03 and vapor pressure vs time

4.6 Chemical compound identification by In-Situ tests

The following test was performed by using Tedlar bags in accordance with EPA method 18. According to the results of these tests, tetrachloroethene (PCE) and trichloroethene (TCE) were found in the test site while toluene was found upstream towards the south of the site.

Table 4.16. Laboratory test report of samples collected onsite in accordance with EPA method 18

VOC Measured	SVP01	SVP02	SVP03	South of Site	Reporting Limits (ug/m3)	CAS
1,1,1-Trichloroethane	ND	ND	ND	ND	11	71-55-6
1,1,2-Trichloroethane	ND	ND	ND	ND	93.4	79-00-5
1,1-Dichloroethane	ND	ND	ND	ND	9.8	75-34-3
1,1-Dichloroethene	ND	ND	ND	ND	17	75-35-4
1,2-Dichloroethane	ND	ND	ND	ND	13.4	107-06-2
1,2-Dichloropropane	ND	ND	ND	ND	20.8	78-87-5
1,4-Dichlorobenzene	ND	ND	ND	ND	34.6	106-46-7
Benzene	ND	ND	ND	ND	28.4	71-43-2
Carbon Tetrachloride	ND	ND	ND	ND	33.2	56-23-5
Chlorobenzene	ND	ND	ND	ND	22.8	108-90-7
Ethylbenzene	ND	ND	ND	ND	22.4	100-41-4
m,p-Xylene	ND	ND	ND	ND	49	179601-23-1

o-Xylene	ND	ND	ND	ND	29.6	95-47-6
Styrene	ND	ND	ND	ND	29.2	100-42-5
Tetrachloroethene	25	ND	ND	ND	13.6	127-18-4
Toluene	ND	ND	ND	12.2	11.2	108-88-3
Trichloroethene	ND	39.0	53.6	ND	47	79-01

The statistical analysis of the soil gas shows that there may be a rising trend in the pollutant levels in the test site with a constant high VOC release at SVP01 and the multiple spikes at SVP02 and SVP03 indicate that the pollutant may be spreading and getting released under these sites as well. I propose that the pollutant might be travelling north from SVP01 to SVP02 as indicated in the diagram below.

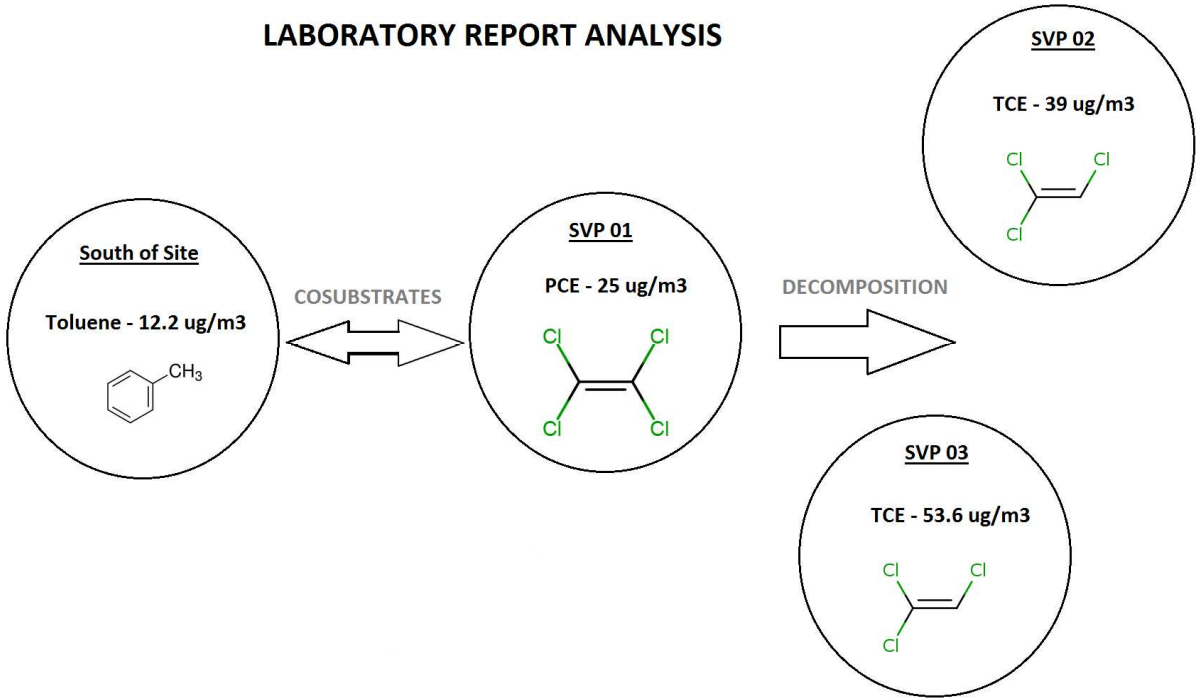


Figure 3.36. Spatial analysis of the laboratory report

CHAPTER 5. CONCLUSIONS

Multiple analysis techniques were applied on the data and the results of these tests and their meaning will be discussed in the following section.

Preliminary data analysis was conducted such as the mean were calculated using box interval plots and simple linear plots were created to visually see and analyze the data. We noticed that there is a period between middle of January till the middle of April when the soil gas sensor at SVP 02 has no high emissions of VOCs while the soil gas sensor at SVP 03 has intermittent high emission phases of VOCs which could be explained by the soil type of the region. SVP 03 has high sandy soil content while SVP 02 has a high clay soil type content.

One of the trend analysis techniques used was linear trend analysis as shown in section 3.3.1. These trend analyses have an equation associated with it and the sign and magnitude of the coefficient associated with time indicates if it is a rising or falling trend. According to the linear trend analysis, there may be a rising trend associated with VOC concentrations at all the soil gas monitoring stations. However linear trend analysis is not the most accurate way to detect rising trends hence we used other tests to verify the trends such as the Mann-Kendall and Sen's slope test as shown in section 3.3. These analyses show that VOC concentrations at SVP 01 may be falling while VOC concentrations at SVP 02 and SVP 03 may be rising. Mann-Kendall and Sen's Slope test are non-parametric and have better p-values associated with them and are hence a better option to linear trend analysis.

Another method of analysis is to break the results into cyclic periods and see how they vary over similar periods of time. This can be done in multiple ways such as over seasons (i.e. summer, fall, winter and spring), or weekly, or daily. Since we have hourly data and the period of study isn't sufficiently large to study seasonal changes, I divided the data into 24-hour chunks to see how the data varied over a period of a day. One way to do this is by using Xbar and S charts as shown in section 3.3.2. These charts are useful to calculate daily means and standard deviations. We can use this to figure out which days had the most deviations to the normal and use the metadata to figure out why that was the case. Another way to perform this analysis is using time-series decomposition as explained in section 3.4.1. From this analysis we can see that VOC concentrations at SVP 01 vary the most at night with a decreasing concentration while VOC concentrations at SVP 02 and SVP 03 vary the most during the morning hours with a decreasing concentration while they have a slightly increasing concentrations all through the day and start decreasing again through the night. Autocorrelation was also performed on data as shown in section 3.4.2 to see how closely the data relates to itself in cyclic periods.

Another important part of this analysis was to correlate other environmental data such as the weather data and ground water quality data with the VOC concentration data at the soil gas stations. The Pearson correlation test was performed for the ground water quality data and soil gas data and it was found that the ground water monitoring station which was the closest to the soil gas monitoring station was most correlated to it (except in the case of SVP 03 which had no statistically correlated data with any of the ground water monitoring stations). I also plotted scatter plots to see how closely related the ground water quality data was to the soil gas data. We also conducted cross correlation tests between the ground water quality and VOC concentrations

at SVP 02 and found out that the soil gas data is most related to the ground water quality data from 24-48-hour lags as shown in section 3.5.2.2. Analysis of variance test was also performed on the soil gas data and weather data. Factors such as average air temperature, soil temperature, relative humidity, vapor pressure, precipitation, solar radiation and wind speeds were considered for regression.

From the lab reports of the samples collected at the site we see that tetrachloroethene (PCE) was predominantly found in SVP 01 and trichloroethane (TCE) was predominantly found in SVP 02 and SVP 03 while toluene was found towards the south of the site. From the real-time readings though we were able to statistically show that the contaminant plume near SVP 01 is spreading towards the other sites. Toluene towards the south of the site indicates that other aromatic hydrocarbons are also present near the contaminant plume.

We were able to use low-cost sensors and data analysis techniques to ascertain the spatial and temporal movement of the contaminant plume underground which was the main goal of this thesis. We were also able to correlate some of the environmental factors and make some inferences on its impact on the release of VOCs from the soil.

Some of the key learnings and inferences from the data analysis performed are listed below.

- We can extract some of the information just by the visual data such as a rising or falling trend of the ground water temperature and pH but other values such as

VOC concentrations at the soil gas station are harder to predict visually. Hence further statistical analysis is needed to predict and analyze these trends.

- Time Series analysis gives us more in-depth data such as the average values at a particular time of day as well as deviations at each hour. This data is useful to predict which time of day has maximum variations and the magnitude of change at an average hourly basis.
- From Auto correlation it was found that all the VOC concentrations are at comparatively most related to themselves at 24-hour periods. Hence VOC release is correlated to the time of the day.
- We got low R squared values from correlating just ground water quality data and soil gas data indicating that more factors were required to be considered other than the ground water temperature, pH and conductivity.
- Cross correlation test of ground water quality and soil gas data indicates that the changes in ground water quality data seems to affect the soil gas release the most after 24 to 48 hours
- The VOC release seem to be positively correlated to humidity, air, ground water and soil temperatures while they seem to be negatively correlated to the vapor pressure, precipitation and to a certain extent to ground water conductivity.
- Multiple low-cost sensors coupled with laboratory testing and data analysis techniques can provide better spatial and temporal analysis results and provide us with more useful key insights.

CHAPTER 6. RECOMMENDATIONS FOR FUTURE STUDIES

There are a lot of gaps in my study due to lack of resources and technical knowledge. For instance, I was only able to use NDIR sensors for real time sensing which is capable of sensing single chain hydrocarbons such as methane, which restricts me from detecting multi chain hydrocarbons and gives me false readings when multiple single chain hydrocarbons are present in the same chemical compound. The EPA method 21 recommends PID sensors for real-time VOC testing and I would recommend PID sensors be used for more accurate results.

Due to untimely calibration of the ground water sensors which caused a lot of the conductivity, oxidation reduction potential (ORP) and dissolved oxygen (DO) data to be inaccurate and hence unusable for the analysis. Ground water conductivity showed the highest correlation with VOC concentration at SVP 03 with the data that was usable. ORP and ground water temperature have a good correlation with the VOC concentrations as well and should be studied in a more in-depth study with cross-correlation in mind to explain what reactions in the ground water lead to the delayed release of VOCs into the soil gas.

The factors such as soil type and its relation to the weather conditions and release of VOC were not a focus of this study but were found to be a major factor affecting it. Further analysis should be done in this field to detect the release of VOC from varying soil type. Other factors studied such as humidity, air, ground water and soil temperatures, vapor pressure,

precipitation and ground water conductivity need to be studied in more detail to provide better relation coefficients with higher R-squared values associated with the equations.

We collected the soil gas data over a period of 8 months consisting of 3 seasons namely fall, winter and spring. There was a tremendous change in magnitude and frequency of the emissions of VOC between the soil gas station during these seasons and I would recommend doing these tests over longer periods which would help give better regression equations and better decomposed graphs since more cyclic periods can be extrapolated and analyzed. I would recommend at least a 3-year study on a site with corresponding soil, ground water and weather data.

The main goal of this study was to develop a more precise setup for real-time VOC release monitoring and help regulate fracking sites more efficiently and productively and to analyze the data collected faster and more accurately and keep it open and easily understandable to the public so that they comprehend the effects of the fracking sites without inaccurate prejudice. Another goal of this study was to bridge the gap between laboratory sampling and real-time on-site testing. We were able to analyze the movement of the contaminant plume using real time sensing and were also able to identify most of the constituents of the contaminants using in-situ data. However, we did not conduct any real-time testing on the south side of our site and hence were not able to perform any statistical analysis on how toluene was reacting with the main contaminant plume (i.e. PCE). Many studies indicate that toluene is one of several co-substrates able to support the co-metabolism of PCE and TCE by soil microbial communities.

Hence it would have been interesting to see how real time changes in toluene would have reacted with the changes in the contaminant plume.

References

1. VANA, M., PEKA'REK, J. (2018, April). Long-term Trends of Air Quality at the Czech EMEP and GAW Station Kosetice Observatory. In *EGU General Assembly Conference Abstracts* (Vol. 20, p. 2751)
2. DEPARTMENT OF THE ENVIRONMENT, WATER, HERITAGE AND THE ARTS, AUSTRALIAN GOVERNMENT. (2009). *NPI Volatile Organic Compound Definition and Information. Version 2.7.*
3. SINDELAROVA, K., GRANIER, C., BOUARAR, I. (2014). Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmospheric Chemistry and Physics*, 14(17), 9317-9341.
4. GUENTHER, A., HEWITT, C.N., ERICKSON, D., FALL, R. (1995). A global model of natural volatile organic compound emissions. *Journal of Geophysical Research: Atmospheres*, 100(D5), 8873-8892.
5. ATKINSON, R. (2000). Atmospheric chemistry of VOCs and NOx. *Atmospheric environment*, 34(12-14), 2063-2101.
6. KAMPA, M., CASTANAS, E. (2008). Human health effects of air pollution. *Environmental pollution*, 151(2), 362-367.
7. MENDELL, M.J., HEATH, G.A. (2005). Do indoor pollutants and thermal conditions in schools influence student performance? A critical review of the literature. In *Indoor Air*.
8. SOFOUGLU, S.C., ASLAN, G., INAL, F. (2011). An assessment of indoor air concentrations and health risks of volatile organic compounds in three primary schools. *International journal of hygiene and environmental health*, 214(1), 36-46.

9. MOLHAVE, L. (1991). Volatile organic compounds, indoor air quality and health. *Indoor Air*, 1(4), 357-376.
10. ABREU, L. SCHUVER, H. (2012). Conceptual Model Scenarios for the Vapor Intrusion Pathway. *US EPA*
11. SPINELLE, L., GERBOLES, M., KOK, G., SAUERWALD, T. (2015). Sensitivity of VOC sensors for air quality monitoring within the EURAMET key-VOC project. In *Proceedings of the Fourth EuNetAir Scientific Meeting, Linköping, Sweden* (pp. 3-5).
12. SZULCZYNSKI, B., GEBICKI, J. (2017). Currently commercially available chemical sensors employed for detection of volatile organic compounds in outdoor and indoor air. *Environments*, 4(1), 21.
13. EMELIN, E.V., NIKOLAEV, I.N. (2006). Sensitivity of MOS sensors to hydrogen, hydrogen sulfide, and nitrogen dioxide in different gas atmospheres. *Measurement Techniques*, 49(5), 524-528.
14. GEBICKI, J., KLOSKOWSKI, A., CHRZANOWSKI, W., STEPNOWSKI, P., NAMIESNIK, J. (2016). Application of ionic liquids in amperometric gas sensors. *Critical reviews in analytical chemistry*, 46(2), 122-138.
15. STRANG, C.R., LEVINE, S.P., HERGET, W.F. (1989). A preliminary evaluation of the Fourier transform infrared (FTIR) spectrometer as a quantitative air monitor for semiconductor manufacturing process emissions. *American Industrial Hygiene Association Journal*, 50(2), 70-77.
16. MOSELEY, P.T. (1997). Solid state gas sensors. *Measurement Science and technology*, 8(3), 223.

17. MOLHAVE, L., NIELSEN, G.D. (1992). Interpretation and limitations of the concept “Total volatile organic compounds” (TVOC) as an indicator of human responses to exposures of volatile organic compounds (VOC) in indoor air. *Indoor Air*, 2(2), 65-77.
18. LUSZTYK, E., YANG, W., WON, D.Y. (2005). *Methodology for the Analysis of Volatile Organic Compounds in Emission Testing of Building Materials*.
19. WOLKOFF, P., CLAUSEN, P.A., JENSEN, B., NIELSEN, G.D., WILKINS, C.K. (1997). Are we measuring the relevant indoor pollutants?. *Indoor Air*, 7(2), 92-106.
20. SALTHAMMER, T., UHDE, E. (2009). *Organic indoor air pollutants: Occurrence, measurement, evaluation*. John Wiley & Sons.
21. WILLIAMS, R., KAUFMAN, A., GARVEY, S. (2015). Next generation air monitoring (NGAM) VOC sensor evaluation report. *EPA: Washington, DC, USA*.
22. LEIDINGER, M., SAUERWALD, T., REIMRINGER, W., VENTURA, G., SCHUTZE, A. (2014). Selective detection of hazardous VOCs for indoor air quality applications using a virtual gas sensor array. *Journal of Sensors and Sensor Systems*, 3(2), 253-263.
23. GOLDSTEIN, A.H., GALBALLY, I.E. (2009). Known and unexplored organic constituents in the Earth's atmosphere. *Geochimica et Cosmochimica Acta Supplement*, 73, A449.
24. HELSEL, D.R., HIRSCH, R.M. (2002). *Statistical methods in water resources* (Vol. 323). Reston, VA: US Geological Survey.
25. GILBERT, R.O. (1987). *Statistical methods for environmental pollution monitoring*. John Wiley & Sons.
26. MU, D.Y., SCOW, K.M. (1994). Effect of trichloroethylene (TCE) and toluene concentrations on TCE and toluene biodegradation and the population density of TCE and toluene degraders in soil. *Appl. Environ. Microbiol.*, 60(7), 2661-2665.

27. YANG, W.B., CHEN, W.H., YUAN, C.S., YANG, J.C., ZHAO, Q.L. (2012). Comparative assessments of VOC emission rates and associated health risks from wastewater treatment processes. *Journal of Environmental Monitoring*, 14(9), 2464-2474.
28. GEBICKI, J., DYMERSKI, T. (2016). *Application of Chemical Sensors and Sensor Matrixes to Air Quality Evaluation*.
29. WISLON, A.D. (2012). Review of electronic-nose technologies and algorithms to detect hazardous chemicals in the environment. *Procedia Technology*, 1, 453-463.
30. World Health Organization, Kobenhavn. (2010). *WHO guidelines for indoor air quality: selected pollutants*.
31. OFFICE OF RESOURCE CONSERVATION AND RECOVERY, U.S. EPA. (1998). *Volatile organic compounds in water, soil, soil gas, and air by direct sampling ion trap mass spectrometry (DSITMS): Method 8265*.
32. STATIONARY SOURCE COMPLIANCE DIVISION, U.S. EPA. (2001). *Determination of volatile matter content, water content, density, volume solids and weight solids of surface coatings. – Method 24*.
33. AMERICAN SOCIETY OF HEATING, REFRIGERATING AND AIR-CONDITIONING ENGINEERS. (2012). *Indoor air quality guide: Best practices*.
34. SMITH, J.E., HEATH, L.S. (2008). Carbon stocks and stock changes in US forests. In: *US Department of Agriculture. US Agriculture and Forestry Greenhouse Gas Inventory: 1990-2005. Technical Bulletin No. 1921. Washington, DC: Office of the Chief Economist: 65-80, C1-C7*.
35. TUCKER, W.G. (2001). Volatile organic compounds. *Indoor air quality handbook*, 31, 1-20.

36. MOYER, R.C., MCGUIGAN, J.R., RAO, R.P. (2012). *Contemporary financial management*. Nelson Education.
37. CHEN, G., CHENG, S.W. (1998). Max chart: combining X-bar chart and S chart. *Statistica Sinica*, 263-271.
38. DRAPELA, K., DRAPELOVA, I. (2011). Application of Mann-Kendall test and the Sen's slope estimates for trend detection in deposition data from Bílý Kříž (Beskydy Mts., the Czech Republic) 1997-2010. *Beskydy*, 4(2), 133-146.
39. GILBERT, R.O. (1987). *Statistical methods for environmental pollution monitoring*. John Wiley & Sons.
40. GOLYANDINA, N., NEKRUTKIN, V., ZHILJAVSKY, A.A. (2001). *Analysis of time series structure: SSA and related techniques*. Chapman and Hall/CRC.
41. MANE, T.F., TAKENS, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980* (pp. 366-381). Springer, Berlin, Heidelberg
42. PANKRATZ, A. (2009). *Forecasting with univariate Box-Jenkins models: Concepts and cases* (Vol. 224). John Wiley & Sons.
43. BOX, G.E.P., JENKINS, G.M., REINSEL, G.C., LJUNG, G.M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons
44. ZOU, K.H., TUNCALI, K., SILVERMAN, S.G. (2003). Correlation and simple linear regression. *Radiology*, 227(3), 617-628.
45. MILLER, R.G. (1997). *Beyond ANOVA: basics of applied statistics*. Chapman and Hall/CRC.
46. EMISSION MEASUREMENT CENTER, U.S. EPA. (2002). *Volatile Organic Compounds by Gas Chromatography – Method 18*.

47. EMISSION MEASUREMENT CENTER, U.S. EPA. (2002). *Determination of Volatile Organic Compound Leaks – Method 21*.
48. FOUST, A.S., WENZEL, L.A., CLUMP, C.W. MAUS, L. (2008). FOUST, A.S., WENZEL, L.A., CLUMP, C.W. MAUS, L. (2008). *Principles of Unit Operations*.